UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Fisica e Astronomia G. Galilei

CORSO DI DOTTORATO DI RICERCA IN: ASTRONOMIA

CICLO XXIX

TOWARD A UNIVERSAL MODEL FOR THE MASS ACCRETION HISTORY OF DARK MATTER HALOS IN COSMOLOGICAL SIMULATIONS

Coordinatore: Ch.mo Prof. Giampaolo Piotto

Supervisore: Ch.mo Prof. Giuseppe Tormen

Dottorando: Giacomo Baso

Contents

Abstract

It is now well established that a large percentage of the energy density of the universe is in the form of non-baryonic dark matter, a still unidentified type of matter that does not emit or interact with electromagnetic radiation. At the present time, most of the dark matter is virialized in large structures called 'halos' which formed via hierarchical clustering, a series of subsequent mergers of smaller halos originating from the growth of the perturbations of the density field of the early universe. The study of this tree of mergers, and of its main branch, is of primary importance in understanding the properties of halos at the present time.

The primary tools for the study of the evolution of structures in the non-linear regime are large numerical simulations, that evolve some suitable initial conditions by numerical integration of the gravity equations. We will present our set of simulations, partly developed in the context of this work. By exploiting the large statistic and dynamical range provided, we will present our refinement and expansion of a previous model for the mass accretion history of the halos, greatly expanding its applicability. In particular, our model will allow us to characterize both the median mass accretion history as well as the full halo-to-halo distribution, and we will discuss some applications. Studying the scatter of the distribution, we will present a preliminary analysis of the percentile distributions of mass accretion histories. Despite non-conclusive results, we will provide a characterization that can be useful in checking the validity of methods to generate synthetic merger trees.

We will argue for the universality of the model, which allows us to apply our results to massive neutrino cosmologies. Multiple experiments in recent years confirmed the existence of flavor oscillations in the propagation of neutrino fluxes, a phenomenon usually interpreted as the effect of a non-zero mass for the neutrinos together with a mixing of the flavor and mass eigenstates. The presence of non-zero neutrino masses has severe cosmological implications, causing in particular a slowdown in the growth and evolution of the structures on small scales. We will illustrate how to modify our model to account for these effects.

Sommario

È ormai ben accettato che una grande percentuale della densità di energia dell'universo è sotto forma di materia oscura non barionica, un tipo ancora identificato di materia che non emette o interagisce con la radiazione elettromagnetica. Attualmente, la maggior parte della materia oscura è virializzata in grandi strutture chiamate 'aloni' formatisi tramite clustering gerarchico, una serie di fusioni successive di aloni più piccoli originatisi dalla crescita delle perturbazioni del campo di densità dell'universo primordiale. Lo studio di questo albero di fusioni, e del suo ramo principale, è di primaria importanza per la comprensione delle proprietà degli aloni al tempo attuale.

Gli strumenti principali per lo studio dell'evoluzione delle strutture in regime non lineare sono grandi simulazioni numeriche, che evolvono opportune condizioni iniziali per integrazione numerica delle equazioni della gravità. Presenteremo il nostro set di simulazioni, in parte sviluppate nel contesto di questo lavoro. Sfruttando l'ampia statistica e gamma dinamica fornita, presenteremo il nostro raffinamento ed espansione di un modello precedente per la storia di accrescimento di massa degli aloni, ampliando notevolmente la sua applicabilità. In particolare, il nostro modello ci permette di caratterizzare sia la storia di formazione mediana sia la distribuzione completa degli aloni, e ne discuteremo qualche applicazione. Studiando lo scatter della distribuzione, presenteremo un'analisi preliminare delle distribuzioni percentili delle storie di accrescimento di massa. Nonostante i risultati non siano conclusivi, forniremo una caratterizzazione che può risultare utile per controllare la validità di metodi per generare alberi di fusioni sintetici.

Discuteremo dell'universalità del modello, che ci permette di applicare i nostri risultati a cosmologie di neutrini massivi. Diversi esperimenti negli ultimi anni hanno confermato l'esistenza di oscillazioni di sapore nella propagazione di flussi di neutrini, un fenomeno di solito interpretato come l'effetto di una massa non nulla per i neutrini insieme ad un mixing degli autostati di sapore e di massa. La presenza di masse non nulle dei neutrini ha grandi implicazioni cosmologiche, provocando in particolare un rallentamento della crescita e dell'evoluzione delle strutture su piccola scala. Illustreremo come modificare il nostro modello per tenere conto di questi effetti.

1 Introduction

In the standard cosmological model the Universe is assumed to be highly homogeneous at early times, with very small perturbations originating from quantum fluctuations in the inflaton quantum field. The structures we see today are the results of the evolution of these perturbations in the expanding universe, via the mechanism known as gravitational instability. In this chapter we will present the Newtonian perturbation theory in the linear regime for a static universe, which we will then expand to include relativistic effects and the metric expansions of the universe. We will discuss the non-linear evolution in following sections.

1.1 The Linear growth of perturbations

Let us consider a non-relativistic ideal fluid with density $\rho(\vec{r})$ and velocity $\vec{u}(\vec{r})$ under the influence of a gravitational field with potential $\phi(\vec{r})$. The time evolution of the fluid is given by the continuity equation, the Euler equation and the Poisson equation, together with the equation of state of the fluid:

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} + \rho\vec{\nabla}\cdot\vec{u} = 0 \tag{1.1}$$

$$\frac{\mathrm{D}\vec{u}}{\mathrm{D}t} = -\frac{1}{\rho}\vec{\nabla}p - \vec{\nabla}\phi \tag{1.2}$$

$$\nabla^2\phi = 4\pi G\rho \tag{1.3}$$

$$p = p\left(\rho, S\right) \tag{1.4}$$

where $p(\vec{r})$ is the isotropic pressure, S is the specific entropy and

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \vec{u}\cdot\nabla \tag{1.5}$$

is the convective (lagrangian) time derivative. This description of a single non-relativistic fluid can be expanded to include a smooth background of relativistic particles or vacuum energy by adding new

density terms in the Poisson equation. In the standard scenario of adiabatic and isoentropic systems

$$\frac{dS}{dt} = 0 \tag{1.6}$$

and we are only left with the equation of state to specify for having a complete description of the fluid. The unperturbed solution of the above equations is a static, homogeneous and isotropic universe:

$$\begin{cases} \rho = \rho_b = \text{const} \\ p = p_b = \text{const} \\ \vec{u} = 0 \\ \phi = \phi_b = \text{const} \\ S = \text{const} \end{cases} \tag{1.7}$$

This solution, despite being inconsistent (a constant potential ϕ_b implies $\nabla^2\phi = 0$, which is in contrast with $\rho \neq 0$) is a useful starting point for introducing small perturbations in the system:

$$\begin{cases} \rho = \rho_b + \delta\rho = \rho_b(1+\delta) \\ p = p_b + \delta p \\ \vec{u} = \delta\vec{u} \\ \phi = \phi_b + \delta\phi \end{cases} \tag{1.8}$$

Inserting these solutions in the continuity+Euler+Poisson system, linearizing and removing the unperturbed solutions we obtain:

$$\frac{\partial\delta\rho}{\partial t} + \rho_b\vec{\nabla}\cdot\delta\vec{v} = 0 \tag{1.9}$$

$$\frac{\partial\delta\vec{v}}{\partial t} = -\frac{c_s^2}{\rho_b}\vec{\nabla}\delta\rho - \vec{\nabla}\delta\phi \tag{1.10}$$

$$\nabla^2\delta\phi = 4\pi G\delta\rho \tag{1.11}$$

where c_s is the adiabatic sound speed:

$$c_s^2 = \left(\frac{\partial p}{\partial\rho}\right)_S \tag{1.12}$$

In the Fourier representation, if we look for plain wave solutions with the form

$$\delta f = \delta f \exp\left(i\vec{k}\cdot\vec{r} + i\omega t\right) \tag{1.13}$$

we can rewrite the system as

$$\omega\delta_k + \vec{k}\cdot\vec{u}_k = 0 \tag{1.14}$$

$$\omega\delta\vec{u}_k = -\vec{k}\left(c_s^2\delta_k + \delta\phi_k\right) \tag{1.15}$$

$$\delta\phi_k = -\frac{4\phi G\delta_k\rho_b}{k^2} \tag{1.16}$$

which can be combined in a single equation for ω:

$$\omega^2 = k^2 c_s^2 - 4\pi G \rho_b \tag{1.17}$$

This defines a characteristic length, called Jeans length:

$$\lambda_J = \frac{2\pi}{k_J} = c_s \sqrt{\frac{\pi}{G\rho_b}} \tag{1.18}$$

which expresses the distance a sound wave can travel in a gravitational free-fall time $t_{\rm ff} \simeq \left(G\rho\right)^{-1/2}$. Rewriting Equation 1.17 with the definition Equation 1.18 we obtain

$$\omega^2 = k^2 c_s^2 \left[1 - \left(\frac{\lambda}{\lambda_J}\right)^2\right] \tag{1.19}$$

and we can distinguish two regimes, $\lambda < \lambda_J$ for which $\omega^2 > 0$ and the solution is a propagating acoustic wave, and $\lambda > \lambda_J$ for which $\omega^2 < 0$ and the solution is a stationary wave with an amplitude that either increases (growing mode) or decreases (decaying mode) with time exponentially. The growing mode reflects the gravitational instability that allows the density perturbations in the early universe to become the structures we see today.

1.1.1 Solutions for an expanding Universe

To discuss the time evolution of the perturbations in an expanding universe, it is better to use the comoving coordinates $\vec{x}$ defined as

$$\vec{r} = a(t)\,\vec{x} \tag{1.20}$$

where $a(t)$ is the scale factor and the $\vec{r}$ are the proper coordinates. The proper velocity can then be written as

$$\vec{u} \equiv \dot{\vec{r}} = \dot{a}\vec{x} + a\dot{\vec{x}} = H\vec{r} + \vec{v} \tag{1.21}$$

where $\vec{v} = a\dot{\vec{x}}$ is the peculiar (intrinsic) velocity and $H\vec{r}$ is the Hubble drag, with

$$H = \frac{\dot{a}}{a} \tag{1.22}$$

being the Hubble parameter. In the comoving reference frame, the perturbation equations can be written as

$$\frac{\partial \delta}{\partial t} + \frac{1}{a}\vec{\nabla}\cdot\vec{v} + 3H\delta = 0 \tag{1.23}$$

$$\frac{\partial \vec{v}}{\partial t} + H\vec{v} = -\frac{c_s^2}{a}\vec{\nabla}\delta - \frac{1}{a}\vec{\nabla}\phi \tag{1.24}$$

$$\nabla^2\phi = 4\pi G \rho_b a^2 \delta \tag{1.25}$$

Switching to Fourier space and combining them in a single equation, we find

$$\ddot{\delta}_k + 2H\dot{\delta}_k + \delta_k\left(\frac{k_c^2 c_s^2}{a^2} - 4\pi G\rho_b\right) = 0 \tag{1.26}$$

where we used the comoving wavenumber $k_c = ak$. Again, this equation defines two regimes for the solutions separated by the Jeans scale λ_J. However, the growing mode (which we are interested in) is not exponential any more, because of the gravitational collapse being slowed down by the Hubble expansion. We find

$$\delta_- \propto H(t) \tag{1.27}$$

for the decaying mode and

$$\delta_+ \propto H(t)\int_0^t \frac{dt'}{a^2(t')\,H^2(t')} \propto H(z)\int_z^\infty \frac{1+z'}{E^3(z')}dz' \tag{1.28}$$

for the growing mode, where

$$E(z) = \left[\Omega_\Lambda + (1+\Omega_0)(1+z)^2 + \Omega_m(1+z)^3 + \Omega_r(1+z)^4\right]^{1/2} \tag{1.29}$$

and Ω_X is the fractional density of the specie X at the current time. For an Einstein-de Sitter universe, or at sufficiently high redshift such that EdS is a good approximation:

$$\delta_- \propto t^{-1}; \quad \delta_+ \propto t^{2/3} \tag{1.30}$$

In general, the growing mode can be obtained from Equation 1.28 numerically, but a good approximation exists (Lahav and Suto, 2004):

$$\delta_+(z) = \frac{g(z)}{1+z} \tag{1.31}$$

where

$$g(z) \approx \frac{5}{2}\frac{\Omega_m(z)}{\Omega_m^{4/7}(z) - \Omega_\Lambda(z) + \left[1+\frac{\Omega_m(z)}{2}\right]\left[1+\frac{\Omega_\Lambda(z)}{70}\right]} \tag{1.32}$$

and

$$\Omega_m(z) = \frac{\Omega_m(1+z)^3}{E^2(z)} \tag{1.33}$$

$$\Omega_\Lambda(z) = \frac{\Omega_\Lambda}{E^2(z)} \tag{1.34}$$

1.1.2 Relativistic effects

So far we assumed the cosmic fluid to be composed of non-relativistic matter. For a universe dominated by radiation, Equation 1.26 is slightly modified:

$$\ddot{\delta}_k + 2H\dot{\delta}_k + \delta_k\left(\frac{k_c^2 c_s^2}{a^2} - \frac{32}{3}\pi G\rho_b\right) = 0 \tag{1.35}$$

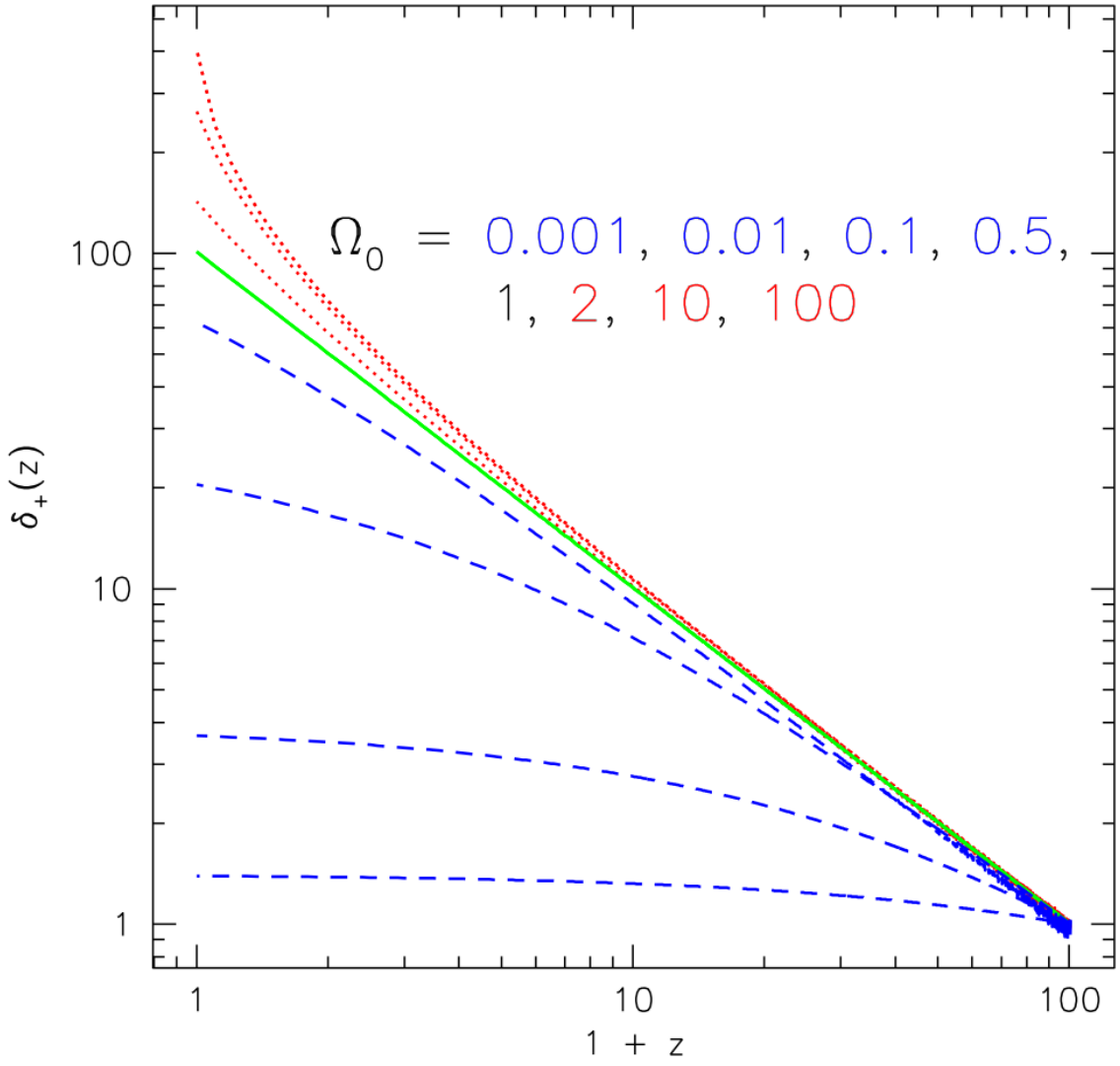

Figure 1.1: Growth of perturbations in matter-dominated universes for different values of Ω. The solid green line represents the $\Omega = 1$ solution, the dashed blue lines the $\Omega < 1$ solutions and the dotted red lines the $\Omega > 1$.

where the different coefficient in front of the gravitational term is due to the contribution of pressure to gravity, which is negligible in the non-relativistic case. Since the radiation pressure is $p = \frac{1}{3}\rho c^2$, the sound speed (Equation 1.12) is close to the speed of light:

$$c_s = \frac{c}{\sqrt{3}} \tag{1.36}$$

Considering the standard background solutions

$$\begin{cases} \rho(t) = \frac{3}{32\pi G t^2} \\ a \propto t^{1/2} \end{cases} \tag{1.37}$$

and looking for a power law solution $\delta_k(t) \sim t^\alpha$ we find:

$$\alpha^2 = 1 - \left(\frac{\lambda_J}{\lambda}\right)^2 \tag{1.38}$$

where in this case the Jeans length is

$$\lambda_J(t) = \frac{2\pi}{k_J} = c_s\sqrt{\frac{3\pi}{8G\rho_b(t)}} \tag{1.39}$$

If $\lambda \ll \lambda_J$ the solution is imaginary and the wave is an acoustic oscillation, while if $\lambda \gg \lambda_J$ the growing solution of Equation 1.35 is

$$\alpha = 1 \quad \Rightarrow \quad \delta_k(t) \propto t \propto a^2 \tag{1.40}$$

It is useful to write the temporal dependency of the Jeans length explicitly. From Equation 1.39, Equation 1.37 and Equation 1.36:

$$\lambda_J = 2\pi c_s t = 2ct\frac{\pi}{\sqrt{3}} \tag{1.41}$$

We remind the reader that in an expanding universe the cosmological horizon is defined as

$$R_H(t) = a(t)\int_0^t \frac{c\,dt'}{a(t')} = a(t)\int_0^t c\,d\tau \tag{1.42}$$

that corresponds to the scale above which there is no causal contact and only gravity acts. Before the epoch of the equivalence $a \propto t^{1/2}$ and we obtain:

$$R_H\left(t < t_{\rm eq}\right) = 2ct \tag{1.43}$$

which is smaller than the Jeans length of the radiation. This means that before the equivalence no perturbation can grow inside the horizon for the radiation field.

1.1.3 Stagnation effect

Let us now consider the evolution of perturbations in a system composed by two fluids. We are interested in particular in the perturbations of the dark matter fluid before the equivalence, when the universe is dominated by radiation. The equation describing the evolution of the perturbations is similar to Equation 1.26, but with both the density contributions in the source term:

$$\ddot{\delta}_{k;m} + 2H\dot{\delta}_{k;m} + k^2 c_{s,m}^2 - \left(4\pi G\rho_m \delta_{k;m} + \frac{32\pi}{3} G\rho_r \delta_{k;r}\right) = 0 \tag{1.44}$$

where we used the subscript m for the dark matter and r for the radiation. Inside the horizon, as we have seen in the previous paragraph, the radiation perturbations can be neglected. Similarly, if we only consider scales above the Jeans length for dark matter (which is very small), we can also neglect the dark matter pressure term. Equation 1.44 then becomes:

$$\ddot{\delta}_{k;m} + 2H\dot{\delta}_{k;m} - 4\pi G\rho_m \delta_{k;m} = 0 \tag{1.45}$$

where, however, we must remember that the evolution of $H = \dot{a}/a$ is driven by the radiation. It is convenient to rewrite this equation in terms of the new variable $x = a/a_{\mathrm{eq}}$. The time derivatives become:

$$\frac{d}{dt} = \frac{dx}{dt}\frac{d}{dx} = \frac{\dot{a}}{a_{\mathrm{eq}}}\frac{d}{dx} \tag{1.46}$$

$$\frac{d^2}{dt^2} = \frac{d}{dt}\left(\left(\frac{\dot{a}}{a_{\mathrm{eq}}}\frac{d}{dx}\right) = \frac{\ddot{a}}{a_{\mathrm{eq}}}\frac{d}{dx} + \frac{\dot{a}}{a_{\mathrm{eq}}}\frac{d}{dt}\left(\frac{d}{dx}\right) = \frac{\ddot{a}}{a_{\mathrm{eq}}}\frac{d}{dx} + \left(\frac{\dot{a}}{a_{\mathrm{eq}}}\right)^2 \frac{d^2}{dx^2}\right. \tag{1.47}$$

Indicating with a $'$ the derivatives with respect to x, Equation 1.45 can be written:

$$\frac{\ddot{a}}{a_{\mathrm{eq}}}\delta_k' + \left(\frac{\dot{a}}{a_{\mathrm{eq}}}\right)^2 \delta_k'' + 2\frac{\dot{a}}{a}\frac{\dot{a}}{a_{\mathrm{eq}}}\delta_k' - 4\pi G\rho_m \delta_k = 0 \tag{1.48}$$

Expressing the density with the new variable x, we first notice that by definition at the equivalence the radiation and dark matter density are equal, $\rho_m(a_{\mathrm{eq}}) = \rho_r(a_{\mathrm{eq}})$. Knowing that $\rho_m \sim a^{-3}$ and $\rho_r \sim a^{-4}$, it means that

$$\frac{\rho_m(a)}{\rho_r(a)} = \frac{a}{a_{\mathrm{eq}}} \equiv x \tag{1.49}$$

Using the first Friedmann equation for $\Omega = 1$, which is a good approximation before the equivalence:

$$\left(\frac{\dot{a}}{a}\right)^2 = H^2 = \frac{8\pi G}{3}(\rho_m + \rho_r) = \frac{8\pi G}{3}\rho_m\left(\frac{x+1}{x}\right) \tag{1.50}$$

which leads us to

$$\rho_m = \frac{3H^2 x}{8\pi G(x+1)} \tag{1.51}$$

Moreover:

$$\frac{\dot{a}}{a_{\mathrm{eq}}} = \frac{\dot{a}}{a}\frac{a}{a_{\mathrm{eq}}} = Hx \tag{1.52}$$

Using the second Friedmann equation:

$$\frac{\ddot{a}}{a_{\rm eq}} = -\frac{4piG}{3}\frac{a}{a_{\rm eq}}\left(\rho + \frac{3p}{c^2}\right) = -\frac{4\pi G}{3}\frac{a}{a_{\rm eq}}\left(\rho_m + 2\rho_r\right) = -\frac{4\pi G}{3}x\rho_m\left(\frac{x+2}{x}\right) = -\frac{H^2 x\,(x+2)}{2\,(x+1)} \tag{1.53}$$

where we substituted the radiation pressure $p = \frac{1}{3}\rho c^2$ and Equation 1.51 for the density. Inserting Equation 1.51, Equation 1.52 and Equation 1.53 into Equation 1.48, we obtain:

$$\delta_k'' + \frac{2+3x}{2x+(1+x)}\delta_k' - \frac{3}{2x+(1+x)}\delta_k' = 0 \tag{1.54}$$

Equation 1.54 is a hypergeometric equation which has a growing solution (Meszaros, 1974):

$$\delta_{k;m+} = 1 + \frac{3}{2}x \tag{1.55}$$

Therefore, from the moment at which the perturbations enter the horizon a_H up to the equivalence, they can only grow by a factor

$$\frac{\delta_k\left(a_{\rm eq}\right)}{\delta_k\left(a_H\right)} = \frac{1+3/2}{1+3a_H/2a_{\rm eq}} \leq \frac{5}{2} \tag{1.56}$$

This result shows that dark matter perturbations almost don't grow inside the horizon before the equivalence, and is called the stagnation effect. Outside the horizon, radiation perturbations above the Jeans length drive the evolution, and dark matter perturbations can grow. The overall effect is a small-scale depression in the power spectrum of the perturbations, from an initial scale-free power spectrum $P_{\rm in}\left(k\right) \sim k^{n_s}$, with $n_s \approx 0.96$ (Planck Collaboration, 2014), to a peaked power spectrum with negative logarithmic slope at scales smaller than the scale of the horizon at the equivalence.

1.2 Spherical collapse

The evolution of perturbations presented in the previous section only covers the linear regime, where $\delta \ll 1$. However, most of the large scale structures today are in the highly non-linear regime, like galaxies and dark matter halos. The spherical collapse model addresses the gravitational collapse and virialization of isolated overdensities in the non-linear regime for collisionless systems with spherical symmetry.

1.2.1 Solution for the perturbation

Let us consider an Einstein-de Sitter universe, where $\Omega_m = 1$ and $\Omega_\Lambda = 0$. An initially comoving top-hat spherical perturbation can be considered as an universe with $\Omega > 1$ embedded in the background flat universe, and we can use the Friedmann equations to derive the solutions for the perturbation and the background. Defining the variable η such that

$$dt = d\eta\frac{R\left(t\right)}{c} \tag{1.57}$$

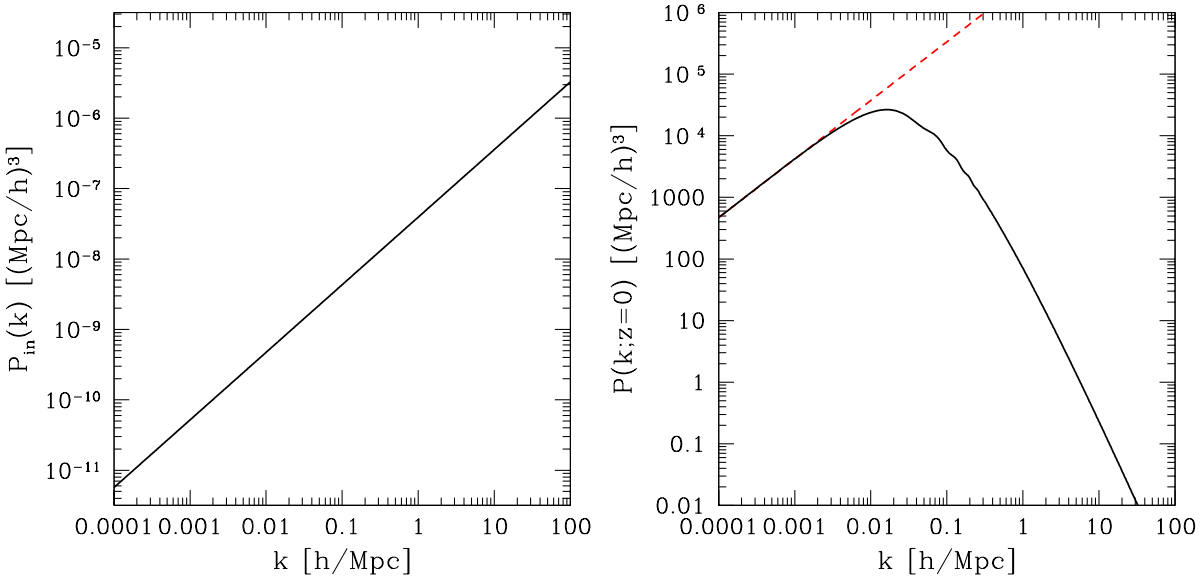

Figure 1.2: Initial and present-day power spectrum of density perturbations. The initial power spectrum is a power law with logarithmic slope $n_s \approx 0.96$. The present-day power spectrum is depressed at small scales by the stagnation effect. Other effects, like the Baryonic Acoustic Oscillations, are also visible. The red dashed line is the extension of the non-suppressed large scale part of the spectrum.

where $R(t)$ is the radius of the spherical perturbation, the first Friedmann equation can be written as:

$$\left[\frac{dR}{d\eta}\right]^2 = \left[\frac{R(t)}{c}\right]^2 \left[\frac{8\pi G}{3c^2}\rho R^2 - k\right] = \frac{8\pi G}{3c^2}\rho R^4 - kR^2 \tag{1.58}$$

If we assume that there is no *shell-crossing*, the mass conservation is:

$$\rho(t) = \rho_0 \left(\frac{R_0}{R}\right)^2 \tag{1.59}$$

Defining

$$R_* = \frac{G}{c^2}\frac{4\pi}{3}\rho_0 R_0^3 = \frac{GM_0}{c^2} \tag{1.60}$$

and plugging in the mass conservation equation, the Friedmann equation becomes:

$$\left[\frac{d}{d\eta}\left(\frac{R}{R_*}\right)\right]^2 = 2\frac{R}{R_*} - k\left(\frac{R}{R_*}\right)^2 \tag{1.61}$$

For the perturbation $k = 1$ and the solution is

$$\frac{R(\eta)}{R_*} = 1 - \cos\eta \tag{1.62}$$

which we can use to derive $t(\eta)$:

$$t(\eta) = \int_0^\eta d\eta' \frac{R(\eta')}{c} = \frac{R_*}{c}\int_0^\eta d\eta' (1 - \cos\eta') = \frac{R_*}{c}(\eta - \sin\eta) \tag{1.63}$$

The combined solutions for $R(\eta)$ and $t(\eta)$ give $R(t)$ in the form of the parametric equation for the cycloid. Considering only one period, the radius of the perturbation initially grows up until it reaches a maximum (called *turn-around* point) for $\eta_{\rm TA} = \pi$, $t_{\rm TA} = \pi \frac{R_*}{c}$ and $R_{\rm TA} = 2R_*$. The perturbation then collapses and the radius becomes zero at $\eta_{\rm coll} = 2\pi$, $t_{\rm coll} = 2\pi \frac{\pi R_*}{c} = 2t_{\rm TA}$.

1.2.2 Solution for the density contrast

For the background, the Friedmann equation is simply

$$\dot{a}^2(t) = \frac{8\pi G}{3}\rho_b a^2 \tag{1.64}$$

which has the classic solution

$$a \propto t^{2/3} \tag{1.65}$$

$$\rho_b = \frac{1}{6\pi G t^2} \tag{1.66}$$

Combining the solutions for the perturbation and for the background, we obtain the evolution of the density contrast:

$$1+\delta(\eta) = \frac{\rho(\eta)}{\rho_b(\eta)} = \frac{3M_0}{4\pi R^3(\eta)} 6\pi G t^2(\eta) = \frac{GM_0}{c^2 R_*}\frac{9}{2}\frac{(\eta-\sin\eta)^2}{(1-\cos\eta)^3} = \frac{9}{2}\frac{(\eta-\sin\eta)^2}{(1-\cos\eta)^3} \tag{1.67}$$

To recover the linear solution, we can expand the terms $(\eta-\sin\eta)^2$ and $(1-\cos\eta)^3$ for $\eta \ll 1$ with a third order Taylor expansion:

$$(\eta-\sin\eta)^2 \simeq \left[\eta - \left(\eta - \frac{\eta^3}{3!} + \frac{\eta^5}{5!} - \dots\right)\right]^2 \tag{1.68}$$

$$(1-\cos\eta)^3 \simeq \left[1 - \left(1 - \frac{\eta^2}{2!} + \frac{\eta^4}{4!} - \dots\right)\right]^3 \tag{1.69}$$

Then substituting in Equation 1.67:

$$1+\delta_{\rm lin}(\eta) \simeq \frac{9}{2}\frac{\left(\frac{\eta^3}{6} - \frac{\eta^5}{120}\right)^2}{\left(\frac{\eta^2}{2} - \frac{\eta^4}{24}\right)^3} \simeq \frac{9}{2}\frac{\frac{\eta^6}{36}\left(1-\frac{\eta^2}{10}\right)}{\frac{\eta^6}{8}\left(1-\frac{\eta^2}{4}\right)} \simeq \left(1-\frac{\eta^2}{10}\right)\left(1+\frac{\eta^2}{4}\right) = 1 + \frac{3}{20}\eta^2 \tag{1.70}$$

where we linearized to leading orders and used the fact that $(1-x)^{-1} \simeq (1+x)$ for $x \ll 1$. Similarly, expanding Equation 1.63:

$$t(\eta) = \frac{R_*}{c}(\eta - \sin\eta) \simeq \frac{R_*}{c}\frac{\eta^3}{6} \tag{1.71}$$

which combined with the previous equation gives

$$\delta_{\rm lin}(t) \simeq \frac{3}{20}\left(\frac{6ct}{R_*}\right)^{2/3} \tag{1.72}$$

This important result can be used to establish a correspondence between the prediction of the linear theory and the actual value of the density contrast in the non-linear regime. For example, at the *turn-around*, that is for $\eta = \pi$ and $t = \pi R_* / c$, we have:

$$1 + \delta_{\mathrm{TA}} = \frac{9}{2}\frac{\pi^2}{8} = 5.552 \tag{1.73}$$

$$\delta_{\mathrm{lin}} = \frac{3}{20}(6\pi)^{2/3} = 1.062 \tag{1.74}$$

that shows that the non-linearities cause the perturbation to grow faster than the linear solution. A more useful value is the prediction of the linear theory at collapse:

$$\delta_{\mathrm{lin}} = \frac{3}{20}(12\pi)^{2/3} = 1.68647 \tag{1.75}$$

which is also called the Press-Schechter threshold or the critical value for collapse δ_c, and is the basis of the model presented in section 1.3.

1.2.3 Virialization

So far we defined the collapse as the moment at which the density diverges. However, this does not happen in practice, as the no shell-crossing constrain fails before the collapse. Instead, the perturbation undergoes violent relaxation, the mass shells cross and oscillate around an equilibrium and the perturbation virializes. At the turn-around, since we are considering a self gravitating system, the total energy is

$$E_t(t_{\mathrm{TA}}) = U_{\mathrm{TA}} = -\alpha\frac{GM_0}{R_{\mathrm{TA}}} \tag{1.76}$$

where α is a parameter that depends on the mass distribution. For the conservation of energy, the total energy of the system at the turn-around is equal to the total energy at collapse:

$$E_t(t_{\mathrm{coll}}) = E_t(t_{\mathrm{vir}}) = T_{\mathrm{vir}} + U_{\mathrm{vir}} = -\alpha\frac{GM_0}{R_{\mathrm{TA}}} \tag{1.77}$$

Assuming α to remain constant, and using the virial theorem $2T + U = 0$:

$$E_{\mathrm{vir}} = T_{\mathrm{vir}} + U_{\mathrm{vir}} = \frac{U_{\mathrm{vir}}}{2} = -\alpha\frac{GM_0}{2R_{\mathrm{vir}}} = -\alpha\frac{GM_0}{R_{\mathrm{TA}}} \tag{1.78}$$

from which we obtain

$$R_{\mathrm{vir}} = \frac{1}{2}R_{\mathrm{TA}} \tag{1.79}$$

The density contrast at the time at which $R(\eta) = R_{\mathrm{TA}}/2$ is:

$$1 + \delta(t_{\mathrm{vir}}) = \frac{\rho(t_{\mathrm{vir}})}{\rho_b(t_{\mathrm{vir}})} = \frac{\rho(t_{\mathrm{vir}})}{\rho(t_{\mathrm{TA}})}\frac{\rho_b(t_{\mathrm{TA}})}{\rho_b(t_{\mathrm{vir}})}\frac{\rho(t_{\mathrm{TA}})}{\rho_b(t_{\mathrm{TA}})} = \left[\frac{R_{\mathrm{vir}}}{R_{\mathrm{TA}}}\right]^{-3}\left[\frac{t_{\mathrm{TA}}}{t_{\mathrm{vir}}}\right]^{-2}[1 + \delta(t_{\mathrm{TA}})] \tag{1.80}$$

Using the fact that $R_{\rm vir} = R_{\rm TA}/2$, $t_{\rm vir} = 2t_{\rm TA}$ and $1+\delta(t_{\rm TA}) = 9/16\pi^2$, we finally obtain:

$$\Delta_{\rm vir} \equiv 1+\delta(t_{\rm vir}) = 18\pi^2 \simeq 178 \tag{1.81}$$

which is the exact value of the non-linear overdensity for spherical virialized structures in an Einstein-de Sitter universe.

1.2.4 Solutions in a flat LCDM universe

The values $\delta_c = 1.686$ and $\Delta_{\rm vir} = 178$, albeit frequently used in literature, are only valid for a universe with $\Omega_m = 1$, $\Omega_\Lambda = 0$, or at very high redshift. If the cosmological constant is non-zero, there is an additional term $\frac{\Lambda}{3}R$ in Equation 1.58 that greatly complicates the computation. An useful approximation for the linear critical overdensity for collapse is given by (Lacey and Cole, 1993):

$$\delta_c = \frac{3}{20}(12\pi)^{2/3}[\Omega_m]^{0.0055} \approx 1.686\,[\Omega_m]^{0.0055} \tag{1.82}$$

which is accurate to better than 1%, and shows that the dependence on cosmology is very weak. For the non-linear overdensity, a commonly used approximation (Bryan and Norman, 1998) is:

$$\Delta_{\rm vir}(z) = \frac{18\pi^2 + 82x - 39x^2}{1+x} \tag{1.83}$$

where $x = \Omega_m(z) - 1$ and

$$\Omega_m(z) = \frac{\Omega_m(1+z)^3}{E^2(z)} \tag{1.84}$$

$$E(z) = \left[\Omega_\Lambda + (1+\Omega_0)(1+z)^2 + \Omega_m(1+z)^3 + \Omega_r(1+z)^4\right]^{1/2} \tag{1.85}$$

For $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, we have:

$$\delta_c = 1.675; \qquad \Delta_{\rm vir}(z=0) = 337 \tag{1.86}$$

We remind the reader that we defined $\Delta_{\rm vir}$ as the overdensity with respect to the background density $\rho_b = \rho_c\Omega_m$. However, and quite confusingly, the use of the overdensity with respect to the critical density is also common in the literature, in which case the value $\Delta^c_{\rm vir}(z=0) = 101$ should be used instead.

1.3 Excursion sets and mass function

We have seen in the previous sections that small perturbations on scales larger than the Jeans scale grow following a function δ_+ defined by Equation 1.31 and Equation 1.32. The linear theory holds until $\delta \sim 1$, but following the exact evolution of a perturbation in the non-linear regime is difficult.

We need a statistical way to characterize the collapsed objects population without following the non-linear dynamics in detail, in particular we need a method to partition the density field at some early time into a set of disjoint patches, each of which will form a single collapsed object at some later time. Then we can calculate the statistical properties of this partition. In order to do this, we can exploit the correspondence we found between the prediction of the linear theory in the non-linear regime and the non-linear solution obtained with the spherical collapse model, namely the value of $\delta_c \simeq 1.686$ for a virialized object. This correspondence allows us to use the linear theory only throughout the entire evolution of the perturbations, and consider δ_c as a threshold for the collapse:

$$\delta\left(\vec{x}, t_{\mathrm{coll}}\right) = \delta\left(\vec{x}\right) \frac{\delta_+\left(t_{\mathrm{coll}}\right)}{\delta_+\left(t_0\right)} = \delta_c \tag{1.87}$$

Sometimes it is useful to consider the growth of perturbations from a different point of view. Instead of thinking of an evolving field $\delta\left(\vec{x}, t\right)$ and a static threshold δ_c, we can transfer the temporal dependence to the threshold by multiplying Equation 1.87 by $\delta_+\left(t\right)/\delta_+\left(t_0\right)$. This approach is equivalent to considering the linear fluctuations field $\delta\left(\vec{x}\right)$ rescaled at the time t_0, and a "moving barrier"

$$\delta_c\left(t\right) = \delta_c \frac{\delta_+\left(t_0\right)}{\delta_+\left(t\right)} \tag{1.88}$$

that is decreasing with time. In this context, let us define the collapse with an idea first proposed by Press and Schechter (1974): the infinitesimal mass element in $\vec{x}$ is part of a halo of mass M or greater at the time t if the linear fluctuation $\delta_f\left(\vec{x}; R\right)$ centered in $\vec{x}$ and filtered on a sphere with radius $R \propto M^{1/3}$ has a value greater or equal than the required threshold:

$$\vec{x} \in M \quad \Longrightarrow \quad \delta_f\left(\vec{x}; R\right) \geq \delta_c\left(t\right) \tag{1.89}$$

1.3.1 Brownian paths

The filtered density field $\delta_f\left(\vec{x}; R\right)$ is the convolution of the pointwise field $\delta\left(\vec{x}\right)$ with a suitable window function $W\left(\vec{x}, R\right)$ of typical amplitude R:

$$\delta_f\left(\vec{x}; R\right) = \frac{1}{\left(2\pi\right)^3} \int d^3k \widehat{\delta}\left(k\right) \widehat{W}\left(kR\right) = \frac{1}{2\pi^2} \int_0^\infty dk k^2 \widehat{\delta}\left(k\right) \widehat{W}\left(kR\right) \approx \frac{1}{2\pi^2} \int_0^{k_f} dk k^2 \widehat{\delta}\left(k\right) \equiv \delta_f\left(\vec{x}; k_f\right) \tag{1.90}$$

where $k = \pi/R$ is the wavenumber corresponding to the filtering radius R. In the same way, we can also define the mass variance as the variance of the filtered density field:

$$\sigma_R^2 = \left\langle \delta^2\left(\vec{x}; R\right) \right\rangle = \frac{1}{2\pi^2} \int_0^\infty dk k^2 P\left(k\right) \widehat{W}^2\left(kR\right) \approx \frac{1}{2\pi^2} \int_0^{k_f} dk k^2 P\left(k\right) \equiv S\left(k_f\right) \tag{1.91}$$

$S\left(k_f\right)$ is a monotonically increasing function of k_f, such that $S\left(k_f = 0\right) = 0$ and $S\left(k_s \to +\infty\right) \to +\infty$. We can then construct for every point $\vec{x}$ the path in the bidimensional space $\left(S\left(k_f\right), \delta_f\left(\vec{x}; k_f\right)\right)$ of the

perturbation δ_f centered in $\vec{x}$ and filtered on the scale k_f. Every path will begin in $(S,\delta_f) = (0,0)$ for a null fluctuation and an infinite filtering radius, and it will depart from the origin point stochastically, depending on the mass distribution around $\vec{x}$. If we choose a top-hat window function in Fourier space, every step in the path will be uncorrelated. The trajectory of δ_f is then a brownian motion, described by the diffusion equation

$$\frac{\partial Q}{\partial S} = \frac{1}{2}\frac{\partial^2 Q}{\partial \delta_f^2} \tag{1.92}$$

where $Q(\delta_f, S)$ is the probability distribution in the stochastic variable δ_f for the trajectories with a given value of $\sigma^2(k_s) = S$. For a free brownian path the solution to the diffusion equation is a gaussian probability distribution:

$$Q(\delta_f, S) = \frac{1}{\sqrt{2\pi S}} \exp\left(-\frac{\delta_f^2}{2S}\right) \tag{1.93}$$

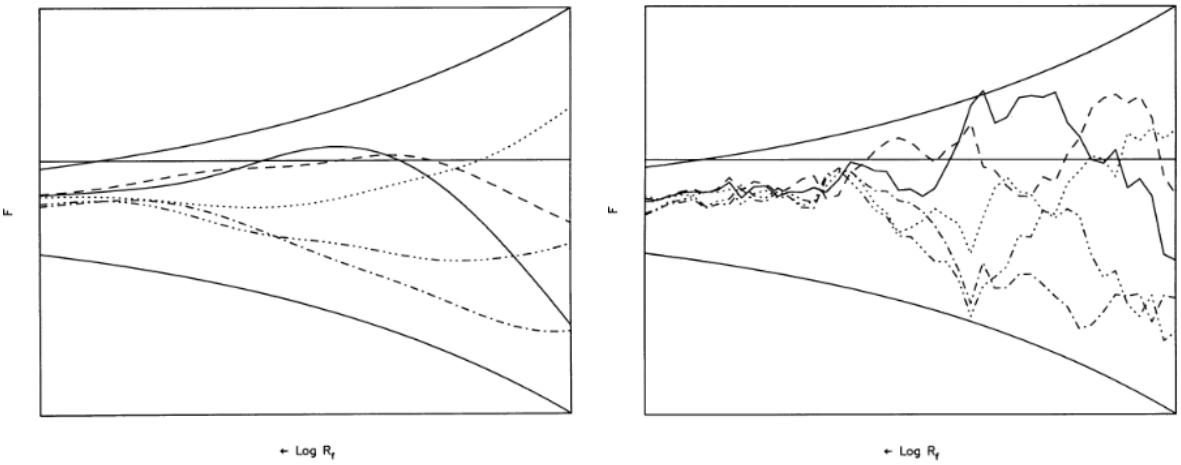

Figure 1.3: Trajectories of the filtered density field as a function of the filter scale R_f, for five different positions and a gaussian (left panel) and top-hat (right panel) window function. With a gaussian filtering the trajectories are correlated in k, while for a top-hat they are not and the motion is a true random walk. From Bond et al. (1991).

1.3.2 The excursion sets model

Using the brownian paths we can construct a model, called *excursion sets* model, first proposed by Bond et al. (1991). The fundamental idea is the following: trajectories that, starting from the origin and with increasing S, reach for the first time the moving barrier $\delta_f(S) = \delta_c(t)$ in correspondence to a given S are related to fluid elements which at the time t belong to collapsed objects with radius $R(S)$ and mass $M = \frac{4\pi}{3}\rho_b R(S)^3$. If a trajectory does not cross the barrier, the corresponding mass element in $\vec{x}$ does not belong yet to a collapsed object. With the passing of time the barrier lowers

and new trajectories will cross it. To compute the mass function of the collapsed objects, which is the mass distribution of the virialized structures at each time, we need to be able to count the types of trajectories with respect to the value of $\delta_c(t)$. Let us consider, at a time t, a critical scale k_0 and a corresponding mass variance S_0. There are three kinds of trajectories:

1. trajectories that reached the barrier for some value $k_f < k_0$ and are still over the barrier:

$$\delta_f(k_0) \geq \delta_c(t) \tag{1.94}$$

2. trajectories that are under the barrier at $k_f = k_0$, but that have reached the barrier for some lower value of k_f:

$$\delta_f(k_0) < \delta_c(t) \quad \text{but} \quad \exists\, k_f < k_0 \quad \text{such that} \quad \delta_f\left(k_f\right) \geq \delta_c(t) \tag{1.95}$$

3. trajectories that have always been under the barrier:

$$\delta_f\left(k_f\right) < \delta_c(t) \quad \forall k_f < k_0 \tag{1.96}$$

In particular, we want to count the third kind of trajectories, as the total number of trajectories that are under the barrier at $k_f = k_0$, given by Equation 1.93, minus the trajectories of the second kind. Since the path of the trajectories is determined by the addition of independent Fourier modes (thanks to the choice of the top-hat window function in Fourier space), at every step a trajectory has equal probability of moving up or down. This means that for every trajectory of the second kind there is another, virtual, with an equal probability that reaches the threshold in the same point $(S, \delta_c(t))$ coming from above instead of from below, obtained by reflecting about the $\delta_c(t)$ axis the part of the trajectory before the first crossing of the threshold. This virtual trajectory corresponds to a brownian path that begins in $\left(S, \delta_f\right) = (0, 2\delta_c(t))$, and satisfies the same diffusion equation with solution Equation 1.93 but translated by $2\delta_c(t)$. The probability associated to the second kind of trajectories is then

$$Q_2\left(\delta_f, S, \delta_c(t)\right) d\delta_f = \frac{1}{\sqrt{2\pi S}} \exp\left[-\frac{\left(\delta_f - 2\delta_c(t)\right)^2}{2S}\right] d\delta_f \tag{1.97}$$

The probability associated to the trajectories of the third kind is then given by the difference

$$\begin{aligned} Q_3\left(\delta_f, S, \delta_c(t)\right) d\delta_f &= \left[Q\left(\delta_f, S\right) - Q_2\left(\delta_f, S, \delta_c(t)\right)\right] d\delta_f \\ &= \frac{1}{\sqrt{2\pi S}} \left\{ \exp\left(-\frac{\delta_f^2}{2S}\right) - \exp\left[-\frac{\left(\delta_f - 2\delta_c(t)\right)^2}{2S}\right] \right\} \end{aligned} \tag{1.98}$$

The fraction of trajectories that did not reach the threshold $\delta_c(t)$ within a time t is the cumulative probability of the previous expression:

$$P_3(S, \delta_c(t)) = \int_{-\infty}^{\delta_c(t)} Q_3\left(\delta_f, S, \delta_c(t)\right) d\delta_f \tag{1.99}$$

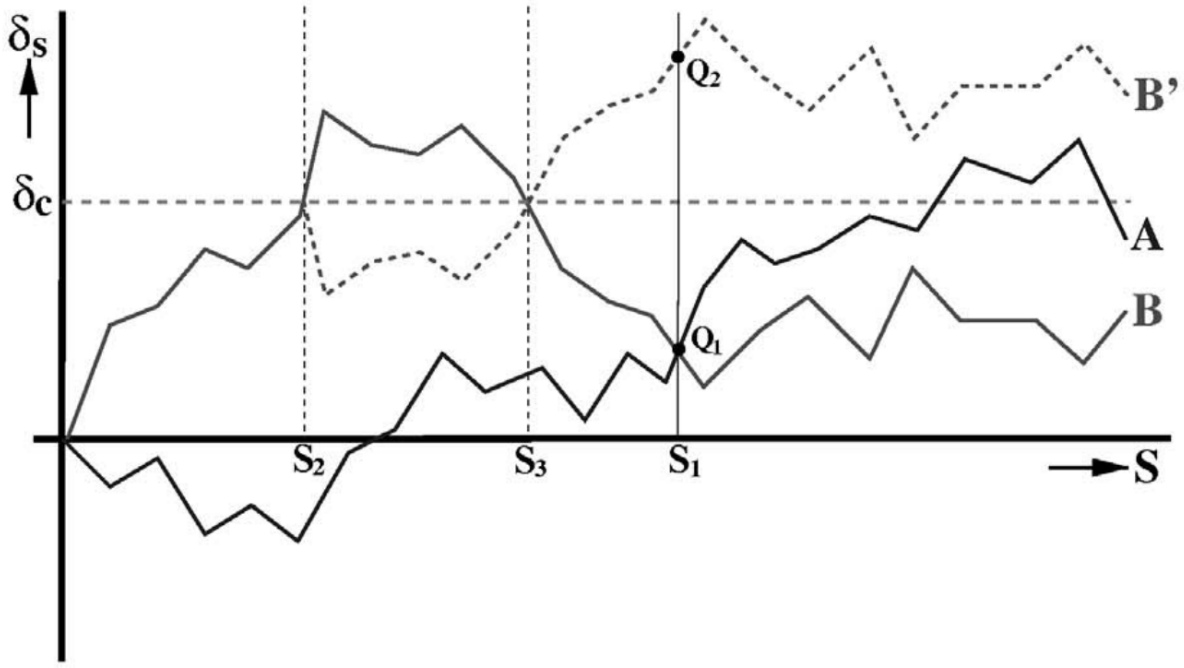

Figure 1.4: Schematic representation of random walk trajectories in (S,δ_f) space. Each trajectory corresponds to a mass element in the initial density field, and δ_f is the overdensity at the location of that mass element when the initial density field, linearly extrapolated to the present time, is smoothed with a top-hat filter in Fourier space of mass $M = 4/3\pi\rho_b R^3$ given by $S = \sigma^2(R)$. The horizontal dashed lines indicates the critical density for spherical collapse, $\delta_f = \delta_c$. Trajectory B' is obtained by mirroring trajectory B in $\delta_f = \delta_c$ for $S \geq S_2$, and, since the trajectories are Markovian, is equally likely as trajectory B.

the complement of which is the fraction of trajectories that have already reached the barrier at an earlier time t:

$$\overline{P_3}(S,\delta_c(t)) = 1 - P_3(S,\delta_c(t)) \tag{1.100}$$

The last expression represents the numerical fraction of fluid elements that at the time t belong to collapsed object of mass M associated to a variance $< S$, that is the mass fraction that at the time t is in objects with variance $< S$. This is exactly the definition of cumulative mass function at the time t, expressed in the variable S:

$$\overline{P_3}(S,\delta_c(t)) = P(<S,t) \tag{1.101}$$

1.3.3 Mass function

The differential mass function, that is the probability that at the time t a fluid element belongs to a collapsed object with mass between M and $M+dM$, is the probability distribution of the trajectories that reach the barrier for the first time at time t and variance S, and it is obtained differentiating the

cumulative distribution Equation 1.101 with respect to S:

$$p\left(S,\delta_c\left(t\right)\right) \equiv \frac{\partial \overline{P_3}\left(S,\delta_c\left(t\right)\right)}{\partial S} = -\frac{P_3\left(S,\delta_c\left(t\right)\right)}{\partial S} = -\frac{\partial}{\partial S}\int_{-\infty}^{\delta_c(t)} Q_3\left(\delta_f,S,\delta_c\left(t\right)\right)d\delta_f \tag{1.102}$$

Differentiating under the integral sign and using Equation 1.93 to substitute the derivative with respect to S with the derivative with respect to δ_f, we obtain

$$p\left(S,\delta_c\left(t\right)\right) = -\frac{1}{2}\int_{-\infty}^{\delta_c(t)} \frac{\partial^2 Q_3}{\partial \delta_f^2} d\delta_f = -\frac{1}{2}\left.\frac{\partial Q_3}{\partial \delta_f}\right|_{-\infty}^{\delta_c(t)} \tag{1.103}$$

and using Equation 1.98 for Q_3:

$$\frac{\partial Q_3}{\partial \delta_f} = \frac{1}{\sqrt{2\pi S}}\left\{-\frac{\delta_f}{S}\exp\left[-\frac{\delta_f^2}{2S}\right] + \frac{\left(\delta_f - 2\delta_c\left(t\right)\right)}{S}\exp\left[-\frac{\left(\delta_f - 2\delta_c\left(t\right)\right)^2}{2S}\right]\right\} \tag{1.104}$$

This function in $-\infty$ is vanishing because it is dominated by the negative exponentials, therefore we only need to evaluate it in δ_f:

$$\begin{aligned} p\left(S,\delta_c\left(t\right)\right) \equiv \frac{df\left(S\right)}{dS} = -\frac{1}{2}\left.\frac{\partial Q_3}{\partial \delta_f}\right|^{\delta_c(t)} &= \frac{-1}{2\sqrt{2\pi S}}\frac{\left(-2\delta_c\left(t\right)\right)}{S}\exp\left[-\frac{\delta_c^2\left(t\right)}{2S}\right] \\ &= \frac{\delta_c\left(t\right)}{\sqrt{2\pi S^3}}\exp\left[-\frac{\delta_c^2\left(t\right)}{2S}\right] \end{aligned} \tag{1.105}$$

This expression gives the mass fraction in structures with variance S. If we want to express it in terms of the variable M, we can use the conservation of probability law

$$\frac{df\left(M\right)}{dM} = \frac{df\left(S\right)}{dS}\left|\frac{dS}{dM}\right| \tag{1.106}$$

Given that

$$\left|\frac{dS}{dM}\right| = \frac{S}{M}\left|\frac{d\ln S}{d\ln M}\right| = \frac{2S}{M}\left|\frac{d\ln\sigma}{d\ln M}\right| \tag{1.107}$$

there follows

$$\frac{df\left(M\right)}{dM} = \sqrt{\frac{2}{\pi}}\frac{\delta_c\left(t\right)}{M\sigma\left(M\right)}\left|\frac{d\ln\sigma}{d\ln M}\right|\exp\left[-\frac{\delta_c^2\left(t\right)}{2\sigma^2\left(M\right)}\right] \tag{1.108}$$

The number density of objects with mass M in a volume V containing a total mass M_V is

$$\frac{dn\left(M\right)}{dM} = \frac{1}{V}\frac{df\left(M\right)}{dM}\frac{M_V}{M} = \sqrt{\frac{2}{\pi}}\frac{\rho_b}{M^2}\frac{\delta_c\left(t\right)}{\sigma\left(M\right)}\left|\frac{d\ln\sigma}{d\ln M}\right|\exp\left[-\frac{\delta_c^2\left(t\right)}{2\sigma^2\left(M\right)}\right| \tag{1.109}$$

This is the classic Press-Schechter mass function (Press and Schechter, 1974), but obtained with the excursion sets formalism. It is notable that it is possible to rewrite Equation 1.109 in a way that is independent of the perturbation spectrum, using the variable $\nu = \delta_c\left(t\right)/\sigma\left(M\right)$:

$$\frac{df}{d\ln\nu} = \sqrt{\frac{2}{\pi}}\nu\exp\left(-\frac{\nu^2}{2}\right) \tag{1.110}$$

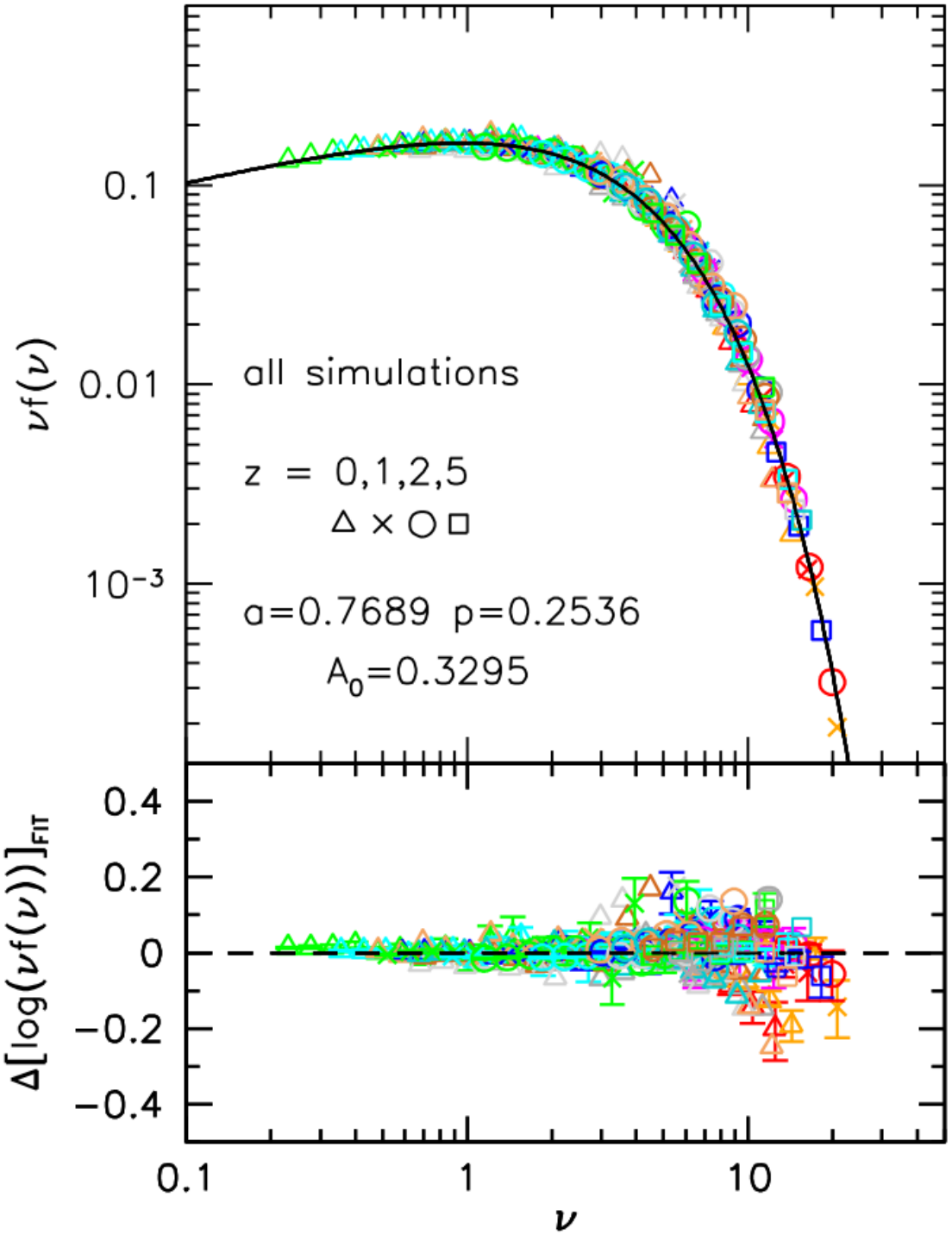

Figure 1.5: Universality of the mass function. Symbols show measurements from several simulations with different cosmologies at four redshifts: $z = 0, 1, 2, 5$. From Despali et al. (2016).

1.4 Dark matter halos

In this chapter we presented the evolution of the primordial perturbations in the density field, first in the linear regime when $\delta_k \ll 1$ and then in the non-linear regime $\delta \gtrsim 1$, up until the collapse and virialization. Such virialized objects are commonly called "halos", and represent the building blocks of the large scale structures in the universe. It can be shown that in a universe with a power-law primordial power spectrum $P(k) \propto k^n$ with index $n > -3$ the halo formation is *hierarchical*, i.e. smaller halos form first and larger halos form by subsequent mergers of the smaller ones. At each time, there is a characteristic non-linear mass M^* such that most of the halos that have formed have mass $M \lesssim M^*$. It is easy to see that

$$\sigma^2\left(M^*(t)\right) = \delta_c^2(t) = \delta_c^2\left[\frac{\delta_+(t_0)}{\delta_+(t)}\right]^2 \tag{1.111}$$

which means that for our universe $M^* \approx 4.944 \times 10^{12}\ \mathrm{M_\odot\ h^{-1}}$ at the present time.

As a first approximation, dark matter halos are mostly spherical, and they are found in numerical simulations to have an equilibrium density profile well described by the simple form (Navarro et al., 1997)

$$\frac{\rho(r)}{\rho_b} = \frac{\delta_0}{(r/r_s)(1+r/r_s)^2} \tag{1.112}$$

which is called *NFW profile*. In this formula ρ_b is the background density of the universe and δ_0 and r_s two parameters, called characteristic overdensity (related to the total mass of the halo) and scale radius. Using the spherical overdensity model, the virial radius of the halo is defined as the radius within which the mean density is Δ_{vir} times that of the background. The enclosed mass of the profile is then

$$M(r) = 4\pi\rho_b\delta_0 r^3\left[\ln(1+cx) - \frac{cx}{1+cx}\right] \tag{1.113}$$

where $x = r/R_{\mathrm{vir}}$, and

$$c = \frac{R_{\mathrm{vir}}}{r_s} \tag{1.114}$$

is called concentration. From Equation 1.113 evaluated at R_{vir} we obtain a relation between the characteristic overdensity and the concentration:

$$\delta_0 = \frac{\Delta_{\mathrm{vir}}}{3}\frac{c^3}{\ln(1+c) - c/(1+c)} \tag{1.115}$$

Thus, for a given cosmology, the NFW profile is completely characterized by its mass M and its concentration c, or equivalently by r_s and δ_0. Moreover, there exists a natural relation between the mass and the concentration, which is found in both simulations and galaxy clusters, for which more massive halos statistically have lower concentrations, albeit with significant scatter. Since smaller halos form earlier, when the universe was denser, this relation is thought to derive from the fact that the inner regions of dark matter halos retain an imprinting of the density of the universe at the time of their

formation, with denser cores leading to higher concentrations. Why the equilibrium density profile of dark matter halos is universal, what are the mechanisms that lead to its formation and how precisely the initial conditions and the assembly history influence its parameters are fundamental questions in the current cosmological research.

Cosmological simulations 2

Although the physics behind gravitational collapse is simple, in the sense that only gravitational interactions are involved, the evolution of the cosmic density field is in general complicated. This complexity arises because the initial density field contains perturbations over a wide range of scales, and non-linear evolution couples structures of different scale, as can be seen with a higher order perturbation theory. Thus, to follow the evolution of the cosmic density field in detail, we need to resort to numerical simulations. Despite the name, numerical (N-body) simulations are not simulations but rather integrations of the gravitational forces acting on a discretization of the density field inside a periodic box, represented by N point-like clumps of mass (particles). Currently, the larger simulations have hundreds of billions of particles, within boxes 100-1000 Mpc h^{-1} on a side.

2.1 GIF2

The GIF2 simulation (Gao et al., 2004) follows the evolution of 400^3 dark matter particles in a periodic cube of side 110 Mpc h^{-1} from an initial redshift $z = 49$ to the present time. The cosmological parameters adopted are $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\sigma_8 = 0.9$ and $h = 0.7$. The simulation was run in two step, from the initial redshift up to $z = 2.2$ with the parallel `SHMEM` version of the `HYDRA` code (MacFarland et al., 1998) and then completed with `GADGET` (Springel et al., 2001) which has better performances in the heavily clustered regime. The GIF2 simulation, processed with the codes described in section 2.3, is the main datased used by Giocoli et al. (2012), which is the basis of our work.

2.2 Le SBARBINE

Our main dataset consist in a set of 6 simulations ran between 2013 and 2014 by our group in Padova using the publicly available code `GADGET-2` (Springel, 2005), and are summarized in Table 2.1. The adopted cosmology was chosen from the 2013 Planck results (Planck Collaboration XVI): $\Omega_m = 0.30711$,

name	box [Mpc h^{-1}]	z_i	m_p [$M_\odot$ h^{-1}]	soft [kpc h^{-1}]	N_{h-tot}	$N_{h>1000}$
Ada	62.5	124	1.94×10^7	1.5	2264847	36532
Bice	125	99	1.55×10^8	3	2750411	44833
Cloe	250	99	1.24×10^9	6	3300880	54421
Dora	500	99	9.92×10^9	12	3997898	58193
Emma	1000	99	7.94×10^{10}	24	4739379	38574
Flora	2000	99	6.35×10^{11}	48	5046663	5285

Table 2.1: Main features of Le SBARBINE simulations: name, box size, initial redshift, particle mass, softening length and number of halos (total and of the subset of halos with more than 1000 particles) identified by the halo finder (subsection 2.3.1) at redshift $z = 0$.

$\Omega_\Lambda = 0.69289$, $\Omega_b = 0.04825$, $h_0 = 0.6777$, $s_8 = 0.8288$. They are collectively named Le SBARBINE simulations and were first presented in Despali et al. (2016).

Each simulation is Dark-Matter only and follows the evolution of 1024^3 particles, with increasing box sizes (from 62.5 Mpc h^{-1} to 2000 Mpc h^{-1}) in order to probe a large dynamical range, with both significant statistic and resolution at the low as well as the high end of the mass function, without the technical limitations of a single extremely large simulations (mainly the need for enormous processing power and storage space). Overall, the simulations contain close to 22 million halos, around 240000 of which have more than 1000 particles, with masses ranging from 2×10^{10} to 4×10^{15} M$_\odot$ h^{-1}.

For all simulations initial conditions were generated by the public code `N-GenIC` (Springel, 2003), starting from a standard *glass* distribution (Springel, 2005; White, 1994) and a suitable power spectrum generated with the public code `CAMB` (Lewis et al., 2000b). The random seed for each simulation was chosen carefully, to obtain independent realization and to follow the initial power spectrum even at large scales.

All runs (except for Ada, see subsection 2.2.2) and the post-processing were performed in Padova on the cluster 'Nemo': a SuperServer Twin 2U Dual Xeon Sandy Bridge composed by four independent node servers each equipped with two Xeon Sandy Bridge 8 Core E5-2670 and 128 GB of RAM, for a total of 64 cores or 128 CPU-threads and 512 GB of RAM.

2.2.1 COSMO

To complement the primary set of 6 simulations described in the previous paragraph, a secondary set was run by Despali et al. (2016) with different cosmological parameters, to test the universality of the halo mass function with respect to the cosmological model. Each of these simulations (collectively called COSMO hereafter) has 512^3 dark matter particles and was run twice, with a box size of 150 and 1000 Mpc h^{-1} (with a corresponding softening length of 7.2 and 48 kpc h^{-1} respectively) to ensure good resolution both for intermediate and high-mass halos. The cosmological parameters used are summarized in Table 2.2. In this work, we applied the same post-processing pipeline we

name	Ω_m	Ω_Λ	σ_8	box [Mpc h^{-1}]	m_p [$M_\odot$ h^{-1}]
Tea	0.2	0.8	0.7	150	1.396×10^9
Tea-big	0.2	0.8	0.7	1000	4.135×10^{11}
Tina	0.2	0.8	0.9	150	1.396×10^9
Tina-big	0.2	0.8	0.9	1000	4.135×10^{11}
Vera	0.4	0.6	0.7	150	2.791×10^9
Vera-big	0.4	0.6	0.7	1000	8.271×10^{11}
Viola	0.4	0.6	0.9	150	2.791×10^9
Viola-big	0.4	0.6	0.9	1000	8.271×10^{11}
Wanda (wmap7)	0.272	0.728	0.81	150	1.898×10^9
Wanda-big (wmap7)	0.272	0.728	0.81	1000	5.624×10^{11}

Table 2.2: Details of the secondary set of simulations COSMO. Each contains 512^3 dark matter particles with initial conditions generated at redshift $z = 99$. For all simulations the Hubble parameter is $h = 0.6777$, apart from the WMAP7 cosmology for which $h = 0.704$.

will describe for Le SBARBINE in section 2.3 to the COSMO simulations, and we will apply the same analysis to check the dependence on cosmology of our model for the MAH, as it will be presented in subsection 3.2.2.

2.2.2 Ada

In the context of this work, we followed the evolution of the simulation with the smallest box size, named Ada. Given the small scale, for which we expected stronger non linear effects, we choose to start the simulation at higher redshifts. Moreover, for the same reason, the time-step used in the integration of the equation of motion performed by `GADGET` is expected to be smaller. Due to the increased computational resources needed, we ran this simulation on a different system, an IBM Power7 computing cluster with 48 cores (each with a 4 way SMT) and 640 GB of RAM, with a peak performance of 1190.4 GFlops.

To assess the computational resources needed for the run, we conducted an analysis of the performance scaling of `GADGET` in terms of total CPU-time and memory consumption, with a set of 28 low-resolution simulations, named HyperCube. Firstly, we defined a baseline reference simulation with 256^3 particles, 200 Mpc h^{-1} box size, 32 processors and an appropriate softening length (of ~ 20 kpc h^{-1}) defined by the relation (commonly used in the literature):

$$\log\left(\varepsilon\ \left[\text{kpc h}^{-1}\right]\right) = -2.253 + \frac{1}{3}\log\left(m_p\ \left[\text{Mpc h}^{-1}\right]\right) \tag{2.1}$$

where ε is the softening and the mass resolution m_p is derived from the number of particles, the box size and the adopted cosmology (which for these simulations is the same as for Le SBARBINE):

$$m_p = \left(\frac{\text{box}}{N}\right)^3 \times \Omega_m \times \rho_c \tag{2.2}$$

where N^3 is the number of particles. Given this reference simulation, we ran several others changing one of the parameters and keeping the other fixed to asses the performance scaling for that parameter, plus some other simulations with multiple parameter changed to check our findings. We used 50, 100, 200 and 400 Mpc h^{-1} for the box size, 128^3, 256^3 and 512^3 particles, 2, 3, 5, 10, 20 and 40 kpc h^{-1} for the softening length and 8, 16 and 32 processor cores. We also tested different versions of `GADGET-3` (non public) and of the compiler, but we decided for a standard implementation of `GADGET-2` with gcc and openmpi compilers, which proved faster than the other versions, while however lacking the on-the-fly post-processing capabilities of `GADGET-3`. We found a scaling for the total CPU-time:

$$T \sim N^{3.46} \varepsilon^{-0.54} L^0 = N^4 L^{-0.54} \tag{2.3}$$

where L is the box size and we used the scaling for the softening

$$\varepsilon \sim N^{-1} L \tag{2.4}$$

The scaling with the box size, besides the contribution from Equation 2.4, is actually small but non-zero. However, it doesn't appear to be neither power-law nor even monotonic, so we choose to set it to zero and to use a constant box size for future comparisons. The resulting formula is included in a Google Spreadsheet (Baso, 2014), which we used to plan the run for Ada. The predicted time slightly underestimated the actual time it took to run the simulation, probably due to bottlenecks in the messaging protocol for the parallel processors, but resulted in a decent forecast nonetheless.

2.3 Post-processing pipeline

The output of `GADGET-2` consists in a series of 'snapshots' which contain the position, velocities and potential of each particle at the time corresponding to the snapshot, as well as an unique ID for each particle. We chose to save 56 snapshots, as a compromise between time resolution and use of computational resources, equally spaced in $\log(1+z)$ starting from $z = 99$ to $z = 0$, except for Ada which started at $z = 124$ and has an additional snapshot. See Table 2.3 for the full list, as well as the corresponding redshifts. From the snapshots we need to identify the halos and follow their evolution through the simulation, which is done with a pipeline of specialized codes both preexisting and specifically developed in this work.

2.3.1 Halo finder

The first code in the post-processing pipeline is the spherical overdensity halo finder `Denhf` (Tormen et al., 2004). Since halos are overdense regions, we start by computing the local density at each particles as the cube of the inverse of the distance to the tenth closest neighboring particle. We then grow

# snap	z	# snap	z	# snap	z	# snap	z
0	99	14	10.30	28	3.859	42	1.141
1	50	15	9.401	29	3.583	43	1.019
2	35	16	8.810	30	3.322	44	0.9041
3	20	17	8.252	31	3.076	45	0.7959
4	18.81	18	7.726	32	2.844	46	0.6937
5	17.68	19	7.230	33	2.626	47	0.5974
6	16.62	20	6.762	34	2.420	48	0.5065
7	15.62	21	6.320	35	2.225	49	0.4209
8	14.67	22	5.904	36	2.042	50	0.3401
9	13.78	23	5.511	37	1.869	51	0.2639
10	12.94	24	5.141	38	1.706	52	0.1920
11	12.15	25	4.792	39	1.552	53	0.1242
12	11.40	26	4.463	40	1.407	54	0.0603
13	10.69	27	4.152	41	1.270	55	0

Table 2.3: List of the simulation snapshots and corresponding redshifts. Except for the initial 4 snapshots, they are logarithmically equally spaced in redshift.

a sphere around the particle with the highest density, until it encloses a mean density of $\Delta(z)$ times the background density, where Δ is the virial overdensity for the spherical collapse in a Λ CDM cosmology (Eke et al., 1996) and has a value of $\Delta(z=0) \approx 319$ at redshift zero in our cosmology. Having identified the first halo, we remove all its particles from the list, find the next particle with the higher local density and proceed in the same fashion until all halos have been identified, with the additional caveat that an halo must have at least 10 particles to be considered. For every halo identified this way, the code also computes which of the particles are self-bound, save a full list of their IDs, and also save some information like the full number of particles, the number of bound particles, the virial radius, the average peculiar velocity, the velocity dispersion, the total energy, the position of the center of mass and the index of the most bound particle. This catalog is constructed for every snapshot.

2.3.2 Merger trees

In a hierarchical clustering scenario, halos grow by a mechanism where two or more smaller halos (called 'progenitors') merge to form a bigger one, with the addition of a contribution from smooth accretion from the field. The same process happens again many times, resulting in a 'tree' of progenitors for every final halo. The progenitor providing the largest mass contribution to the parent halo is termed the *main progenitor*. Starting from a host halo at $z = z_0$, we iterate this procedure back in time, thus obtaining a complete merger tree down to the mass resolution of 10 particles per halo. The *main branch* of a merger tree is defined as the branch tracing the main progenitor of the main progenitor, etc. We also define as 'satellites' those progenitors who merge directly into the main

branch. A schematic representation of the merger tree is illustrated in Figure 2.1 (Giocoli et al., 2008). For simplicity, we will call the set of halos in the main branch at different snapshots with the name 'main progenitor', identifying them with a single halo whose mass increases from snapshot to snapshot while accreting the satellites. We will use $m_{\rm MP}(z)$ to denote its mass as a function of redshift, M_0 to denote the final mass, and we will refer to the normalized value $m_{\rm MP}(z)/M_0$ as the 'mass accretion history' (hereafter MAH) of the halo.

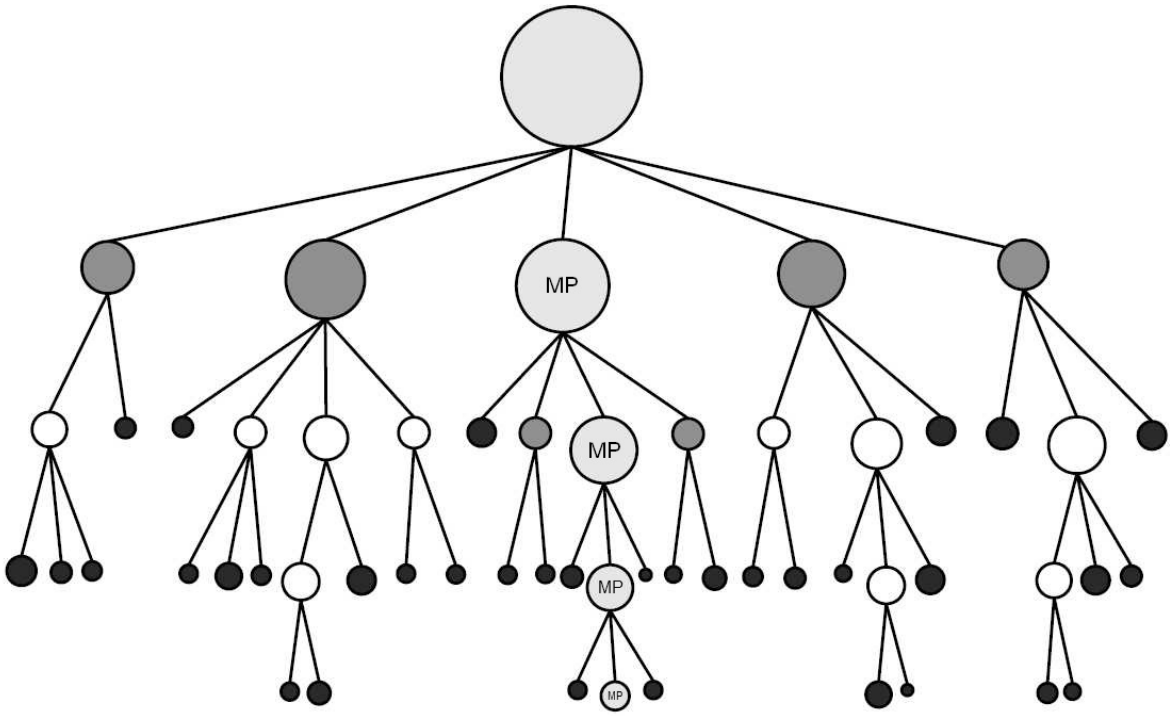

Figure 2.1: Schematic representation of the merger tree of a halo. Light gray circles represent the main branch, dark grey circles indicate satellites. Taken from Giocoli et al. (2008).

Operationally, we define as progenitors only halos that donate at least half of their mass during the merger, as it is done in Tormen et al. (2004) with the code `Mergertree`. However, in this work we rewrote the code from scratch to improve its performance and to include a new intermediary step in the construction of the merger tree, the particle history: from the halo catalogs, the code extracts the particles' membership at every snapshot, i.e. the ID of the halo each particle is contained in. From the particles' membership it is then trivial to construct the full merger tree and the main branch. While not used elsewhere in the remaining of this work, the particle history is an information that is 'orthogonal' to the particle list in the halo catalogs, and it can be especially useful in the study of the internal structure of halos. Caching this information represents the foundation of new efficient codes in future works.

The procedure to derive the MAH that we illustrated, however, suffers from a resolution issue that needs to be discussed. If we had an infinitely small time-step the MAH would be perfectly reconstructed, up to the mass resolution of the simulations. In practice, not only we use a finite time-step for the gravitational evolution of the particles in the simulation, but we also save the position and velocity of the particles for very few of such steps. In particular, in our simulations we only have 56 snapshots at our disposal, in which we can identify halos, thus we are limited to reconstruct the MAH in these time jumps, and we run the risk of missing important details of the mergers. This could potentially lead to a misidentification of the main branch and therefore to a bias in the MAH. Unfortunately, saving lots of snapshots is normally not feasible from a computational standpoint, so one is forced to accept a compromise.

2.3.3 Density profiles

The next step in our pipeline is to compute density profiles for every halo, which can be fitted with Equation 1.112 to derive the concentration. From the center of mass, the code computes the mean density inside a sphere with increasing radius at ~ 30 logarithmically equally spaced intervals in units of the virial radius, which is then converted to a radial profile. This procedure is however quite sensitive to the particle resolution of the innermost regions, especially for low-mass halos. Therefore we only computed the profile for halos with at least 1000 particles within the virial radius, a limit that also guarantees a reliably accurate estimation for the mass of the halos.

2.3.4 Formation redshift

The formation time of a halo is usually defined as the earliest time at which the main progenitor has a mass that exceeds a fraction f of the final mass. Widely used values are $f = 0.5$ (Lacey and Cole, 1994) and $f = 0.04$ (Zhao et al., 2009). In this work we computed the formation redshift of all halos for several values of f, namely 0.01, 0.02, 0.04, 0.1, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.7, 0.75, 0.8 and 0.9, for five different identification redshifts z_0, namely 0, 0.506, 1.019, 2.041 and 4.152. We will use these as the basis for the calibration of the model presented in section 3.1.

The formation redshifts thus defined, however, can only assume the discrete values listed in Table 2.3, which also underestimate the actual redshift which is usually somewhere in the middle of two snapshots, when the main progenitor crosses the mass threshold. To obtain a more accurate estimation, we linearly interpolated the MAH to the logarithm of $1 + z$, and defined the error on this estimate as the error of a uniform distribution in the interval of z between the two corresponding snapshots, $\Delta z = \frac{z_{i+1} - z_i}{\sqrt{12}}$. The choice of a uniform distribution is motivated by the fact that we have no prior information as to when a merging can happen in the considered interval.

2.4 Halo selection

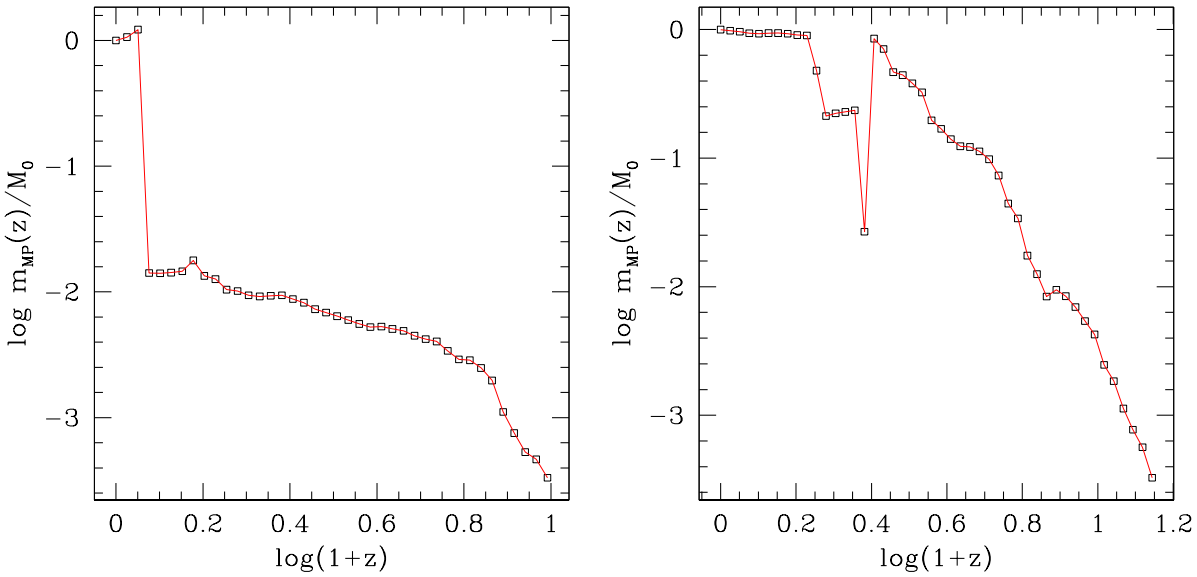

Figure 2.2: Mass accretion histories of two halos excluded by our relaxation criteria. The left panel shows a halo which accretes most of its mass in the last three snapshots; the right panel shows a halo with a prominent gap in its MAH.

The process of halo formation, described in subsection 2.3.2, has as a consequence that the internal structure of dark matter halos is often disturbed and far from equilibrium, which can lead to biased estimates of their mean structural properties (Neto et al., 2007; Macciò et al., 2008). Similarly, the mass accretion histories of such transient halos are far outliers in their distribution and escape modelization, a problem that we will address in subsection 3.2.1. We restricted our analysis to a subset of relaxed halos with the following criteria:

1. the total energy of the halo must be negative
2. the mass of the main progenitor $m_{\rm MP}(z)$ must not exceed by more than 10% the final mass of the halo
3. the distance of the center of mass from the potential minimum must be smaller than 5% of the virial radius
4. the main progenitor must not lose more than 10% of its mass between two consecutive snapshots
5. the halo must have accreted at least $10^{-1.5} \simeq 3.2\%$ of its final mass 4 snapshots before the snapshot of identification

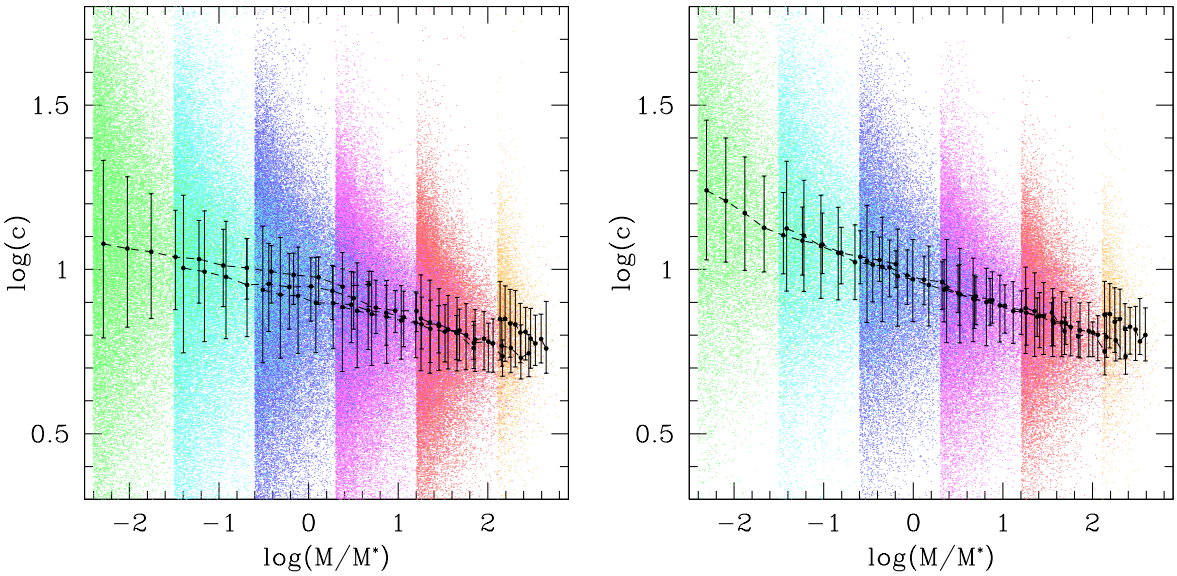

Figure 2.3: Mass-concentration relation for halos in Le SBARBINE simulations. Left panel: all halos with more than 1000 particles at $z = 0$. Right panel: halos that pass all relaxation criteria. Colors from green to orange represent the different simulations, from Ada to Flora, in ascending order of mass. Black points connected by dashed lines represent the median concentration at fixed mass for each simulation, with errorbars representing the 1st and 3rd quartiles.

in addition to the stated minimum of 1000 particles per halo. Two examples of halos excluded by the last two criteria are shown in Figure 2.2. Across all simulations, 128646 halos pass all five criteria at redshift $z = 0$, slightly more than 54%. At higher redshifts the percentage of relaxed halos is similar, with 46% at $z = 0.506$, 56% at $z = 1.019$, 59% at $z = 2.041$ and 64% at $z = 4.152$.

We can check the validity of our relaxation criteria with the median mass-concentration relation for each simulation in Le SBARBINE. As we presented in section 1.4 the concentration is the main parameter, along with the mass, that specifies the density profile of dark matter halos:

$$\frac{\rho(r)}{\rho_b} = \frac{\delta_0}{(r/r_s)\,(1 + r/r_s)^2} \tag{1.112}$$

where r_s is the scale radius and $c = R_{\rm vir}/r_s$. Ideally, the same relation should hold in the regions where the mass range of multiple simulations overlap. However, this is not the case if we include non-relaxed halos, as can be seen in Figure 2.3 (left panel), where there is considerable spread between the median concentrations at fixed mass for different simulations. A much tighter agreement can be obtained using only halos that passes all of our relaxation criteria (right panel).

Modeling the distribution of the mass accretion histories 3

We saw at the end of section 1.4 that the origin of the equilibrium density profile of dark matter halos is still an open question. The relations between mass and concentration seem to indicate that the assembly history of halos may play a role in the determination of the final profile (Navarro et al., 1996; Tormen, 1998; Giocoli et al., 2010), and indeed many authors have linked the concentration to the formation redshift (Bullock et al., 2001; Wechsler et al., 2002; Zhao et al., 2009; Correa et al., 2015) or to the full assembly history (Lu et al., 2006; Ludlow et al., 2013). For example, Zhao et al. (2009) proposed the relation

$$c = 4\left[1+\left(\frac{t}{3.75\,t_{0.04}}\right)^{8.4}\right]^{1/8} \tag{3.1}$$

where $t_{0.04}$ is the time at which the main progenitor first accretes the 0.04% of its final mass. Moreover, the rate at which the dark matter halos grow in mass sets, amongst others, the rate at which baryons can cool to form luminous objects. It is then clear that accurately modeling the mass accretion history of dark matter halos is paramount in our efforts of understanding the properties of large scale structures. Many models exist in the literature, either as empirical formulas (Wechsler et al., 2002; Zhao et al., 2009; Giocoli et al., 2012; van den Bosch et al., 2014) or based on the Press-Schechter formalism (van den Bosch, 2002; Giocoli et al., 2007). However, all of these models either only describe the mean or median MAH, or they are very poor fit of the tails of the distribution of the halo-to-halo MAHs (subsection 3.2.1).

The main goal of this work is to present a reliable, universal model that provides a good description of both the median as well as the full halo-to-halo distribution.

3.1 Previous model

The fitting formula for the MAH proposed by Giocoli et al. (2012, hereafter G12) is based on the definition of a scaled variable for the formation redshift:

$$\omega_{\rm f} = \frac{\delta_c(z_{\rm f}) - \delta_c(z_0)}{\sqrt{\sigma^2(fM_0) - \sigma^2(M_0)}} \tag{3.2}$$

where $z_{\rm f}$ is the formation redshift corresponding to a mass fraction of f, z_0 is the identification redshift, $\delta_c(z)$ is the initial overdensity required for spherical collapse in linear theory and $\sigma^2(M)$ is the variance of the linear fluctuation field when smoothed with a top-hat filter of scale $R = \left(3M/4\pi\rho_b\right)^{1/3}$, where ρ_b is the comoving density of the background. This variable is intimately related to the "peak height" variable ν presented in subsection 1.3.3:

$$\nu = \frac{\delta_c(z)}{\sigma(M_0)} \tag{3.3}$$

that allowed us to write a universal mass function. Using this scaled variable, G12 showed that the cumulative distribution $p\left(>\omega_{\rm f}\middle| f, M_0, z_0\right)$ is very similar for all values of M_0 and z_0, a result that we confirmed, albeit with small deviations for halos identified at $z_0 = 4.152$. Figure 3.1 shows two examples of the distributions of $\omega_{\rm f}$ for $f = 0.5$ and $f = 0.1$ for different values of z_0, for halos taken from all simulations and with all masses. Orange open diamonds represent the overall distribution obtained stacking halos from all of Le SBARBINE, with all masses and at all identification redshifts. Barring a small deviation for $z_0 = 4.152$, the distributions for fixed z_0 are in close agreement with each other and with the overall distribution, hence we will be using the overall distribution as our main dataset throughout this work.

For the cumulative distribution of $\omega_{\rm f}$, G12 proposed the the 1-parameter function

$$P(>\omega_{\rm f}) = \frac{\alpha_{\rm f}}{e^{\omega_{\rm f}^2/2} + \alpha_{\rm f} - 1} \tag{3.4}$$

with the parameter $\alpha_{\rm f}$ depending on f. Given several values of f, one can find the corresponding values for $\alpha_{\rm f}$, and derive an empirical relation $\alpha_{\rm f}\left(f\right)$ which we will call a 'scaling relation' for $\alpha_{\rm f}$, that G12 found to be:

$$\alpha_{\rm f} = 0.815 e^{-2f^3} / f^{0.707} \tag{3.5}$$

Combining Equation 3.5 with Equation 3.4 they obtained a model with no free parameters, which provided a good description, at least around the peak, for the distribution of $\omega_{\rm f}$ for every choice of f and every values of M_0 and z_0. We repeated the same procedure with the much bigger dataset provided by Le SBARBINE simulations, and found:

$$\alpha_{\rm f} = 1.259 e^{-2f^3} / f^{0.658} \tag{3.6}$$

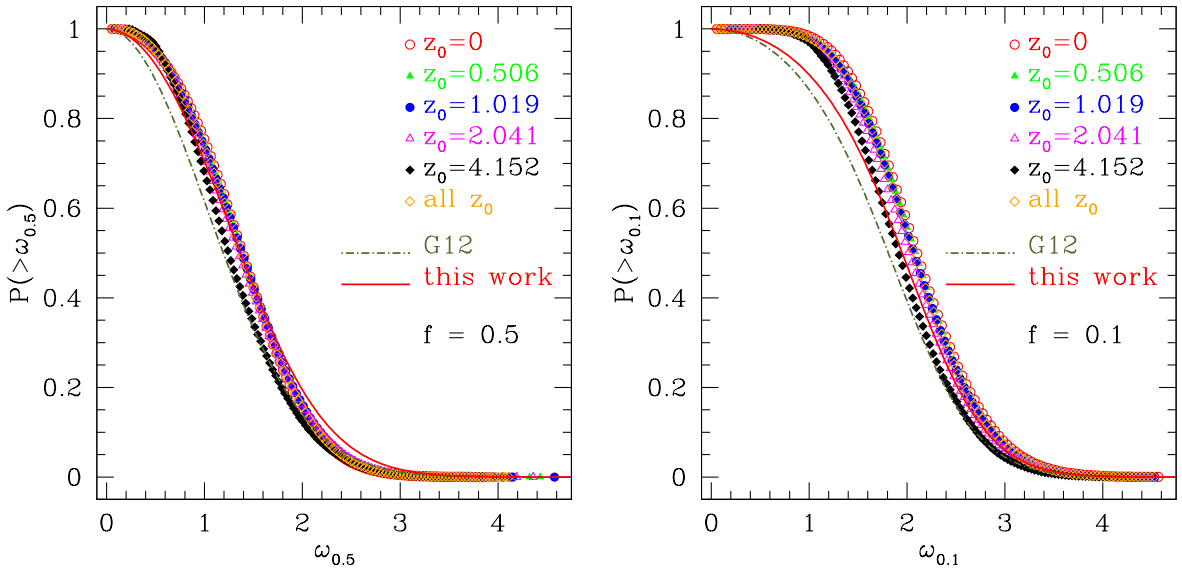

Figure 3.1: Cumulative distribution of the scaled variable ω_f for $f = 0.5$ and $f = 0.1$ for different values of the identification redshift z_0. Halos are taken from Le SBARBINE and stacked, and halos whit all masses are used. The green dot-dashed curve is the original fit of Equation 3.4 and Equation 3.5 from G12. The red solid curve is the same equation but with our calibration.

Figure 3.2 shows the two calibrations. The left panel is the result from G12, with shaded regions around the fit representing the 1 and 2σ contours of $\Delta\chi^2$, where $\chi^2(\alpha_f) = \Sigma_i \left(P_i - P\left(>\omega_{i,f}, \alpha_f\right)^2\right)$. The right panel shows our new calibration, with a shaded region representing the 99.73% confidence interval derived from a bootstrap with 10000 realizations.

From Equation 3.2 we can derive the median formation redshift at fixed f with

$$\delta_c(z_f) = \delta_c(z_0) + \tilde{\omega}_f \sqrt{\sigma^2\left(fM_0\right) - \sigma^2(M_0)} \tag{3.7}$$

where $\tilde{\omega}_f$ is the median of the distribution described by Equation 3.4:

$$\tilde{\omega}_f = \sqrt{2\ln(\alpha_f + 1)} \tag{3.8}$$

Similarly, the median main progenitor mass at fixed $z > z_0$ satisfies

$$\sigma^2\left(\tilde{f}M_0\right) = \sigma^2(M_0) + \frac{\left[\delta_c(z) - \delta_c(z_0)\right]^2}{\tilde{\omega}_f^2} \tag{3.9}$$

where, however, both $\sigma^2\left(fM_0\right)$ and $\tilde{\omega}_f$ depend on f, so the expression must be solved implicitly. However, they both trace the same relation, so we will focus only on the distribution (and the median) of z_f at fixed f. Figure 3.3 shows the median relation from the model compared with the data at $M_0 = M^*$[1] and $z_0 = 0$, the original result by G12 with data from GIF2 and our new calibration with the

[1] M^* is the typical non-linear mass at present time and it is defined as $\sigma^2\left(M^*\right) = \delta_c^2(z=0)$. In our cosmology, $M^* \approx 4.944 \times 10^{12}\,\mathrm{M_\odot\,h^{-1}}$

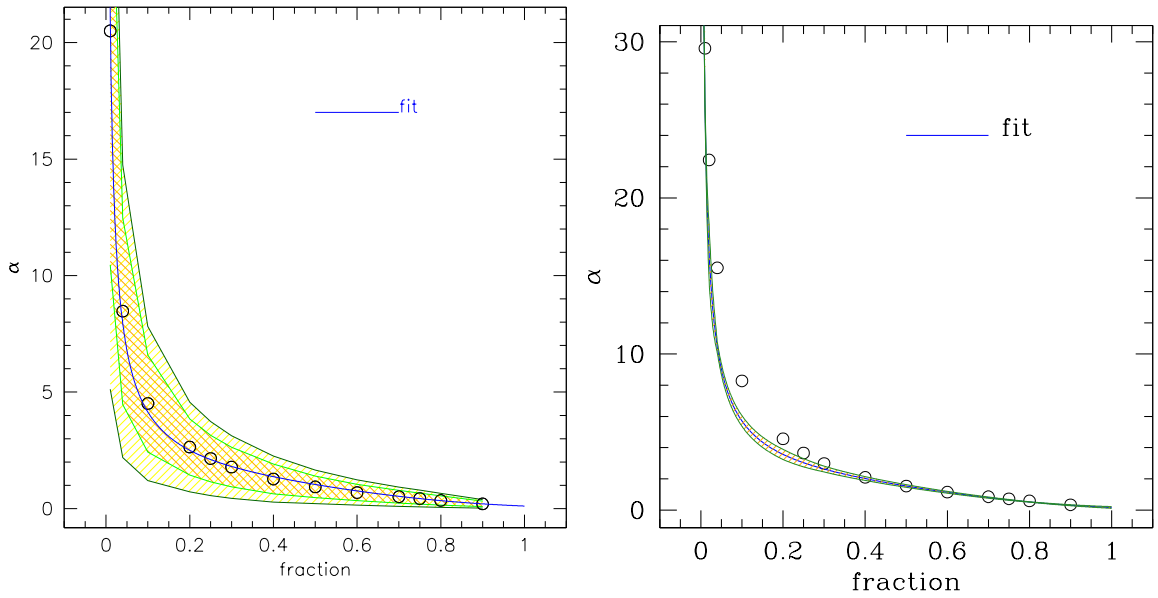

Figure 3.2: Scaling relation for α_f. The left panel is taken from G12, obtained with the GIF2 simulation. The right panel is our new calibration, obtained with Le SBARBINE. The blue curve represent Equation 3.6, with a green shaded region around representing the 99.73% confidence interval.

Le SBARBINE. Circles indicate the median z at fixed f, while the shaded region encloses 95% and 50% (blue solid and dashed lines, respectively) of the MAHs of the halos. G12 also showed a comparison with previous models (Zhao et al., 2009; van den Bosch, 2002).

3.2 Revised model

The combination of Equation 3.7, Equation 3.8 and Equation 3.6 provides a good description of the median relation between the formation time z_f and the mass fraction f required for the formation. However, if one is interested in the full distribution of z_f at fixed f, one would find that Equation 3.4 is actually a rather poor description of the tails of the distribution of ω_f, as it is shown in Figure 3.4, especially for low values of f. To describe the full differential distribution of ω_f, G12 proposed a second function of the form

$$p(\omega_f) = A_0 \omega_f^{0.63 f^{-2/3}} e^{-\gamma_f \omega_f^{\beta_f}} \tag{3.10}$$

where

$$A_0 = \frac{\beta_f \gamma_f^{B_0}}{\Gamma(B_0)} \tag{3.11}$$

is a normalization factor and

$$B_0 = \frac{1 + 0.63 f^{-2/3}}{\beta_f} \tag{3.12}$$

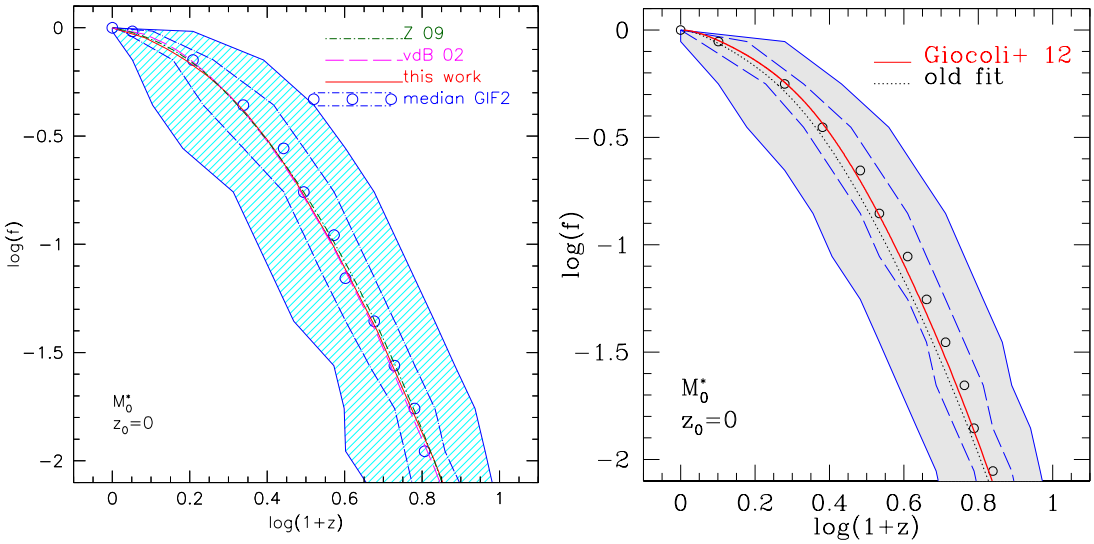

Figure 3.3: Left panel: mass accretion history of halos in the GIF2 simulations, taken from G12. Open circles show the median z at fixed $f = m_{\rm MP}(z)/M_0$, blue solid curves enclose 95% of the data and blue dot-dashed curves enclose 50% of the data. The green dot-dashed and magenta dashed curves show the model of Zhao et al. (2009) and van den Bosch (2002), respectively. The red curve is the result of Equation 3.7 combined with Equation 3.8 and Equation 3.5. Right panel: the same but with data from Le SBARBINE simulations and Equation 3.6 for $\alpha_{\rm f}(f)$. Black dotted curve shows the result from Equation 3.5 for comparison.

For $\gamma_{\rm f}$ and $\beta_{\rm f}$, similarly to the procedure used for $\alpha_{\rm f}(f)$, they found the scaling relations

$$\gamma_{\rm f} = 0.12 + \frac{1}{290 f^{1.4}} + 4.3\left(f - 0.24\right)^2 \tag{3.13}$$

$$\beta_{\rm f} = 3.05 e^{-0.6\gamma_{\rm f}} + \frac{e^{3.2\gamma_{\rm f}}}{3800} \tag{3.14}$$

which are shown in Figure 3.5. As can be seen in the left and central panels, however, both $\gamma_{\rm f}(f)$ and $\beta_{\rm f}(f)$ are non-monotonic functions, therefore there is not an one-to-one correspondence between $\gamma_{\rm f}$ and $\beta_{\rm f}$ and Equation 3.14 can only be a rather crude approximation.

3.2.1 Calibration with Le SBARBINE

Given the relatively small size of the GIF2 simulation, the study of the tails of the distribution of $\omega_{\rm f}$ made by G12 was restricted to about $p(\omega_{\rm f}) \gtrsim 10^{-2}$ by the limited statistic available. The much bigger dataset from Le SBARBINE provided a great opportunity to reach a much deeper level and improve on these previous results. Figure 3.6 shows the distribution of the relaxed halos in Le SBARBINE, compared to Equation 3.30 and our reparametrization of Equation 3.10 described in the following (Equation 3.15). The adherence of the data to the latter is much better, as well as the difference with

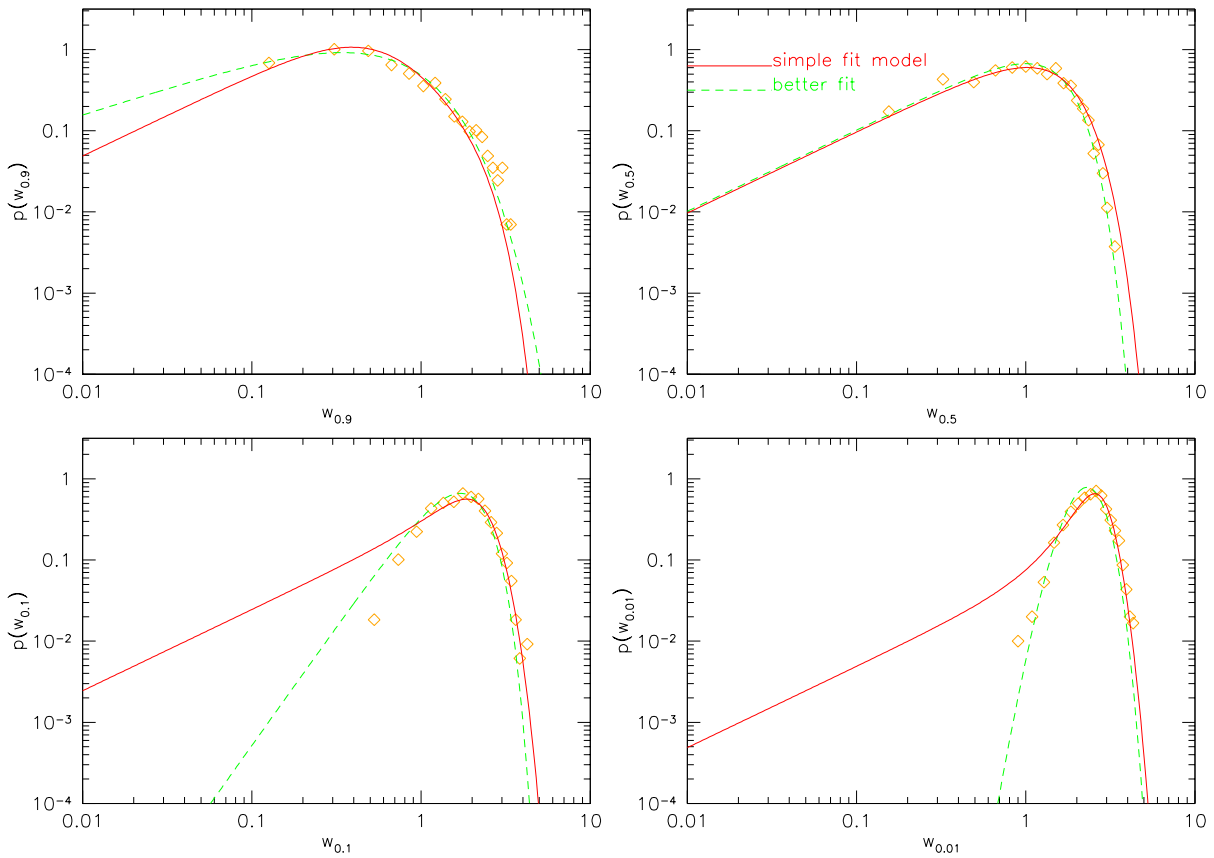

Figure 3.4: Differential distributions of ω_f for $f = 0.01$, $f = 0.1$, $f = 0.5$ and $f = 0.9$, taken from G12. The red solid curve is the fit of the differential of Equation 3.4, while the green dashed curve is the fit of Equation 3.10.

respect to Equation 3.4 for low values of ω_f. In this region, there are still some discrepancies in the form of an excess for low values of f and a defect for high values, with $f \simeq 0.25$ being the threshold. Such discrepancies are highly influenced by the relaxation criteria used, as can be seen in Figure 3.7, where we show the distributions of all halos with more than 1000 particles and of the halos that pass the first three relaxation criteria, related to the internal structure, but possibly not the last two MAH-related criteria (section 2.4). It is clear that adding non-relaxed halos to the sample leads to strong deviations from the distribution of relaxed halos, deviations that escape a simple modelization. In particular, there appears to be a fairly large number of halos with large values of ω_f that don't pass the internal structure relaxation criteria, and they would pollute the distribution if included. The same can be said for very low values of ω_f (possibly related to just-formed halos), with the addition that some of them remains after the internal structure criteria are applied, and the full set of criteria is needed to properly remove these outliers. A more detailed analysis of the distribution of non-relaxed halos will be presented in a future work (Gambarotto et al., in preparation).

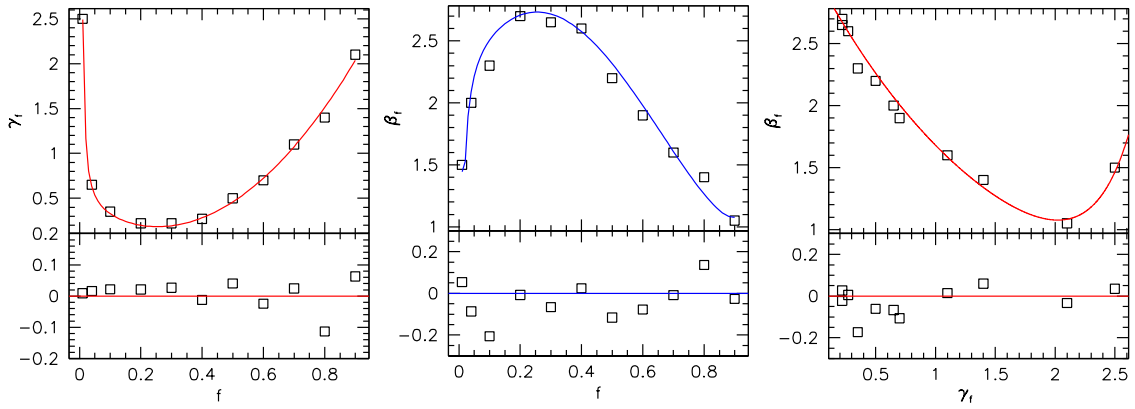

Figure 3.5: Scaling relations for γ_f and β_f. Bottom panels show the residuals of the fitted values with respect to Equation 3.13 and Equation 3.14. Taken from G12.

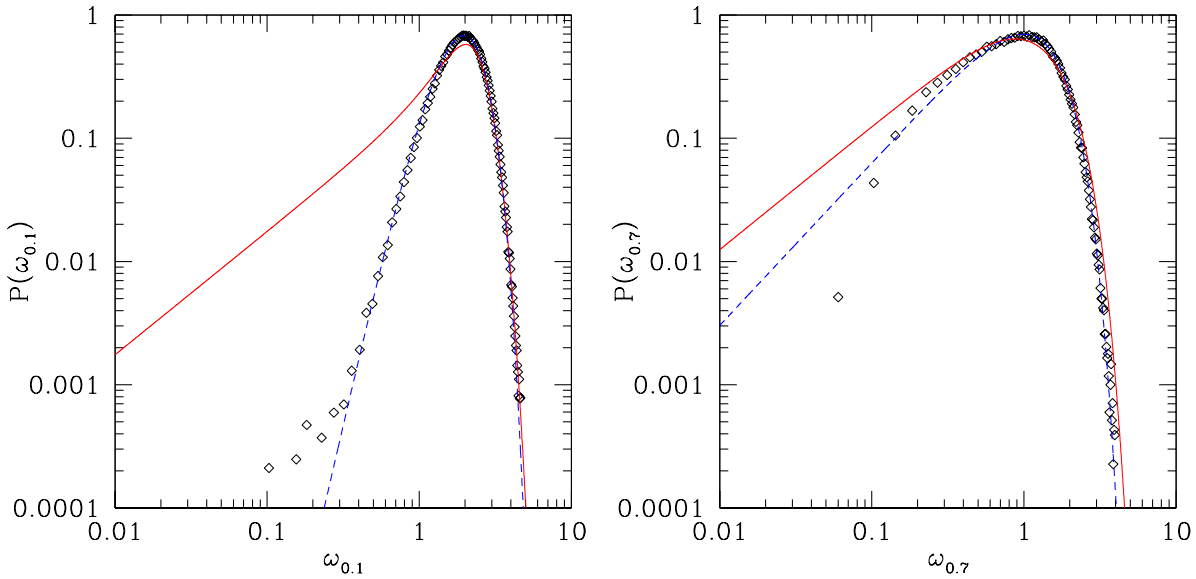

Figure 3.6: Differential distributions of ω_f for $f = 0.1$, and $f = 0.7$ from Le SBARBINE simulations (open symbols). The red solid curve is the fit of the differential of Equation 3.4, while the blue dashed curve is the fit of Equation 3.15.

As mentioned, we decided to change the parametrization of Equation 3.10 as to better fit the results, in particular to avoid an excessive steepness in the scaling relations that can lead to an enhancement of the errors when computing derived parameters, such as the median, and to minimize the covariance of the parameters. After some trial and error, we found

$$p(\omega_f) = \frac{\beta}{\Gamma(\alpha)\gamma}\left(\frac{\omega_f}{\gamma}\right)^{\alpha\beta-1} e^{-\left(\frac{\omega_f}{\gamma}\right)^{\beta}} \tag{3.15}$$

where $\Gamma(x)$ is the Euler Gamma function and where we explicitly freed the parameter α_f in the power-

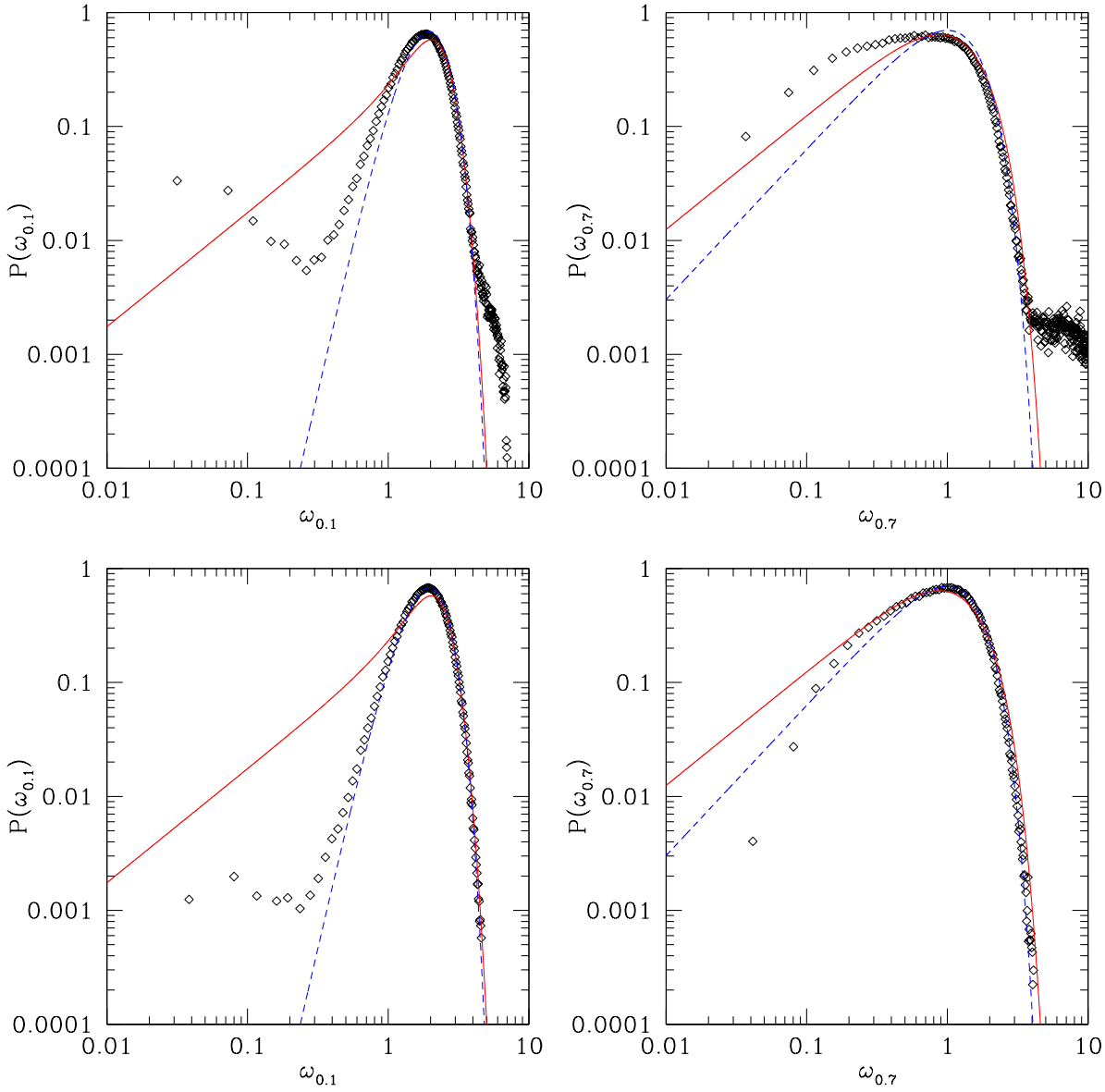

Figure 3.7: Same as Figure 3.6, including non relaxed halos. Top panels: all halos; bottom panels: subsample obtained applying only the first three relaxation criteria, related to the internal structure of the halos.

law component, to satisfy such requests. To simplify the notation, we dropped the explicit dependence of the parameters on f, and we will consider it implicit hereafter. For the fitting we used the orthogonal distance regression (ODR) procedure, as implemented by the python `scipy.odr` package (Jones et al., 2001–), which is an interface for the ODRPACK library (Brown et al., 1990). It uses a modified trust-region Levenberg-Marquardt-type algorithm to estimate the function parameters and, unlike a ordinary least square fit, it also accounts for errors in the independent variable. We remind the reader that we introduced this kind of errors in subsection 2.3.4 on the estimate of z_f (and therefore in ω_f) of each halo as a consequence of the interpolation between discrete snapshots. After binning ω_f, we used the weighted mean as the center of the bin and its standard error as the error of the center. We also assigned a poissonian error to the counts in each bin, and normalized both counts and errors to $\int_0^\infty p(\omega)\,d\omega = 1$. Besides the best-fitting values of the parameters α, β and γ the `scipy.odr` package also provides the standard errors, the 3-parameters covariance matrix, the residual variance, the sum of square errors and several intermediate computations for the fit.

The results of the ODR fitting are shown in Figure 3.8 and Figure 3.9, with residuals in Figure 3.10. Most of the distributions are very well fitted by Equation 3.15, with some deviations especially for very low and high values of f. The estimated parameters are reported in Table 3.1, along with standard errors, and the covariance matrices are reported in Table 3.2. Note that the standard errors are a factor from 1.5 to 5 (depending on f) higher than the errors derived from the covariance matrices. We interpret this to be due to Equation 3.15 being a very good approximation, but not the exact underlying distribution of the data. To construct a confidence interval, represented with orange regions in the figure, we performed a bootstrap on the data. Starting with our sample of 400000 halos, we took 10000 resamplings and applied to each the fitting procedure, obtaining 10000 sets of parameters. The envelope of 95% of the curves derived from these parameters is the confidence region, and it is represented with a light blue shade in the figures, in particular in the insert in the $f = 0.5$ panel in Figure 3.8 and in Figure 3.10.

A different estimation (not shown) can be made with a Monte Carlo simulation. In this case, one would take the estimated parameters as if they were the 'true' values and then draw from the corresponding distribution to construct a synthetic dataset. We took 100000 of such realizations, each with the same number of data points as our original sample, and found similar results for the confidence region. However, in this case the standard errors were very close to the errors derived from the covariance matrix, which reinforces our supposition that Equation 3.15 is not the exact underlying distribution of our data. For this reason we also disregarded the (substantially different, not shown) confidence interval derived from the covariance matrix:

$$\Delta p(\omega) = \left[\sum_i \sum_j \frac{\partial p(\omega)}{\partial x_i} \frac{\partial p(\omega)}{\partial x_j} C_{ij} \right]^{\frac{1}{2}} \tag{3.16}$$

f	α_f	β_f	γ_f	σ_α	σ_β	σ_γ
0.01	6.814859	0.869383	1.735133	0.829242	0.124273	0.110805
0.02	6.098033	0.902954	1.757322	0.156480	0.026239	0.023823
0.04	4.880313	1.012731	1.844040	0.121146	0.025790	0.024481
0.10	3.204840	1.216150	2.020827	0.058611	0.018817	0.020319
0.20	2.047594	1.415909	2.239350	0.029420	0.013841	0.018482
0.25	1.691647	1.497920	2.356822	0.025640	0.013951	0.021019
0.30	1.428261	1.561030	2.468992	0.021844	0.013351	0.022787
0.40	1.069415	1.643741	2.680843	0.016041	0.011693	0.025480
0.50	0.934892	1.608178	2.696833	0.020741	0.016146	0.038889
0.60	0.932266	1.485350	2.492989	0.025771	0.020060	0.044504
0.70	1.081704	1.250469	2.055180	0.043104	0.030737	0.051317
0.75	1.271615	1.055325	1.741658	0.058498	0.036932	0.048821
0.80	1.753582	0.720183	1.328632	0.110152	0.049325	0.048568
0.90	9.769217	0.003123	0.412727	1.911203	0.003246	0.041638

Table 3.1: Fitted parameters from the ODR fitting, with standard errors.

where C is the covariance matrix, $x_{i=1,2,3} = \alpha, \beta, \gamma$ are the parameters of the distribution and the partial derivatives are:

$$\frac{1}{p}\frac{\partial p}{\partial \alpha} = \beta \log\left(\frac{\omega}{\gamma}\right) \tag{3.17}$$

$$\frac{1}{p}\frac{\partial p}{\partial \beta} = \left(\alpha - \left(\frac{\omega}{\gamma}\right)^{\beta}\right)\log\left(\frac{\omega}{\gamma}\right) \tag{3.18}$$

$$\frac{1}{p}\frac{\partial p}{\partial \gamma} = \frac{1}{\gamma}\left(1 - \alpha\beta + \beta\left(\frac{\omega}{\gamma}\right)^{\beta}\right) \tag{3.19}$$

From the covariance matrices, it is apparent that there is a strong covariance between the parameters α_f, β_f and γ_f, quantified by the global correlation coefficient (James and Roos, 1975):

$$\rho_i = \sqrt{1 - \left(C_{ii}C^{-1}{}_{ii}\right)^{-1}} \simeq 0.977 - 0.999 \tag{3.20}$$

where C is the covariance matrix and C^{-1} its inverse. To characterize the correlation, following Sheth et al. (2003), we look at combination of these parameters, such as the mean $\overline{\omega}$ and the most probable value $\widehat{\omega}$:

$$\overline{\omega} = \int_0^\infty \omega\, p(\omega)\, \mathrm{d}\omega = \frac{\Gamma\left(\alpha + \frac{1}{\beta}\right)}{\Gamma(\alpha)}\gamma \tag{3.21}$$

$$\left.\frac{\mathrm{d}p(\omega)}{\mathrm{d}\omega}\right|_{\widehat{\omega}} = 0 \quad \Rightarrow \quad \widehat{\omega} = \left(\alpha - \frac{1}{\beta}\right)^{\frac{1}{\beta}}\gamma \tag{3.22}$$

Equating Equation 3.21 and Equation 3.22 to the measured values from the (unbinned) data, we obtain a set of two equations from which we can derive two of the parameters given the third. For

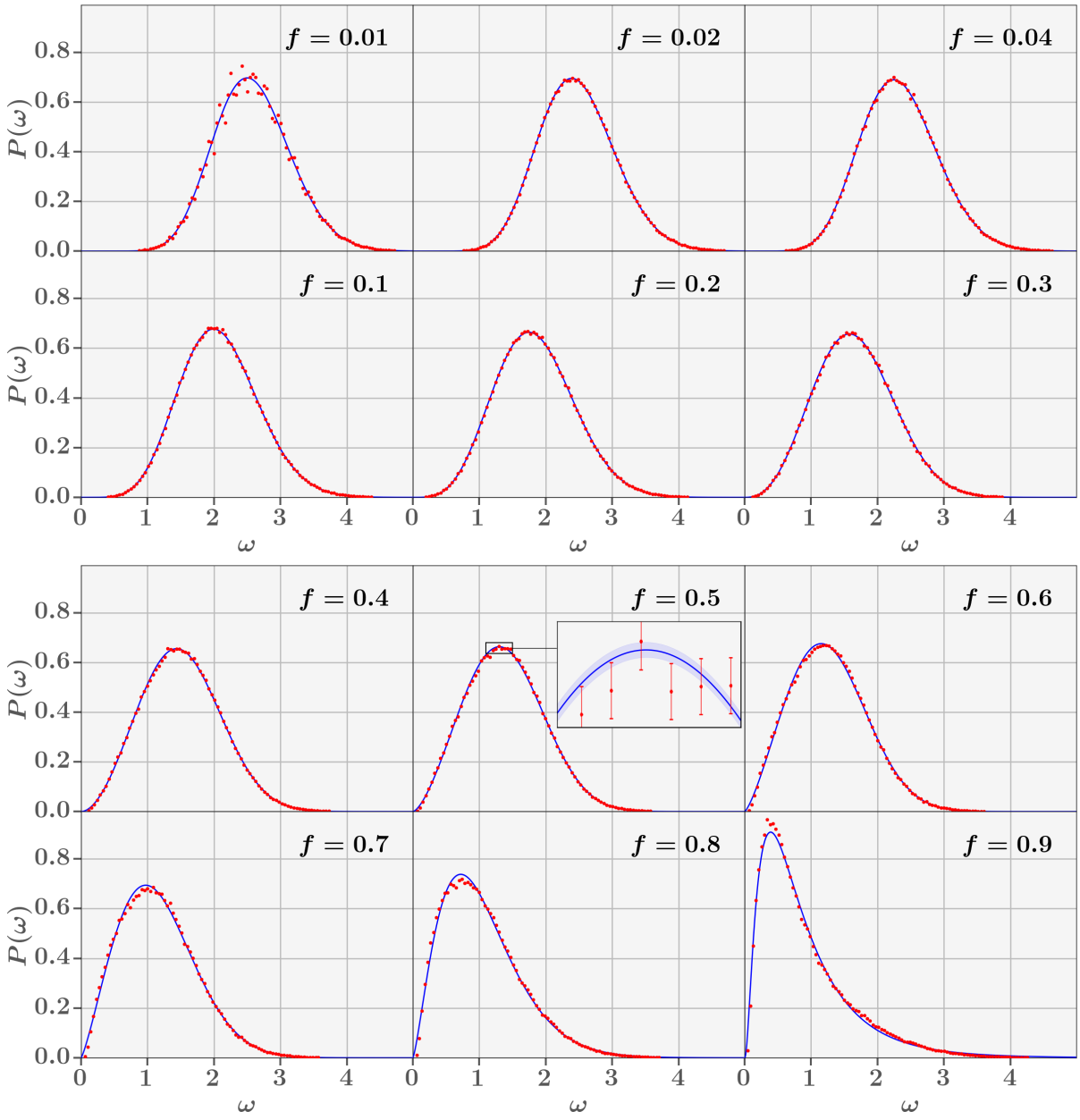

Figure 3.8: Distributions of ω_f with fitting results for different values of f. Red points represent the data, and the solid blue curve represents the best fit of Equation 3.15 obtained with the ODR procedure. The insert in the $f = 0.5$ panel is a zoom-in of the region around the peak of the distribution enclosed in the black rectangle. The insert also contains the 95% confidence region obtained with a bootstrap, represented with a light blue shade, and 1σ error bars, as described in the main text. Despite the magnification, errors in the independent variable are too small to be visible.

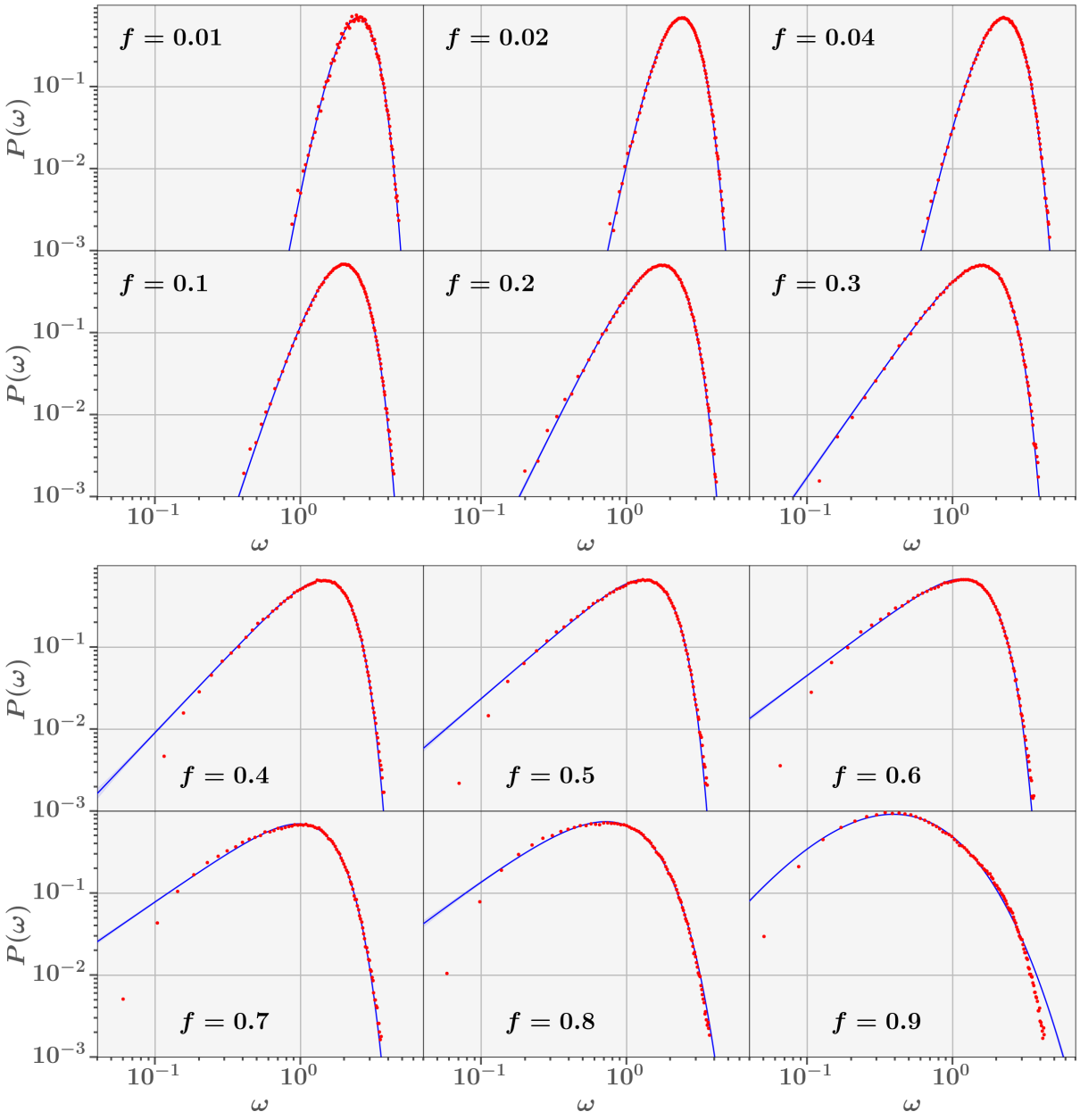

Figure 3.9: Same as Figure 3.8, in the log-log plane.

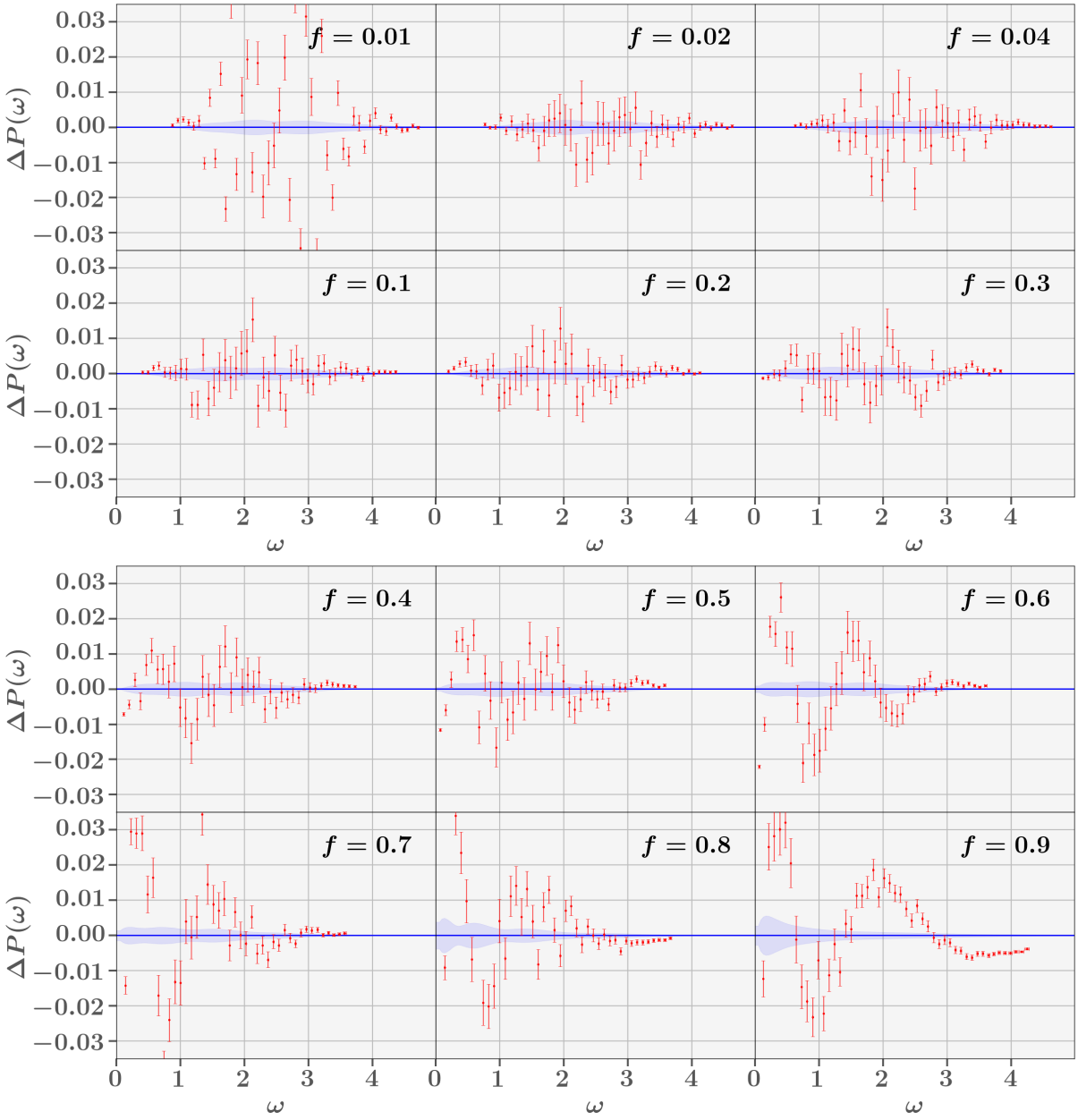

Figure 3.10: Residuals of the fit of Equation 3.15 to the distribution of ω_f for different values of f. Light blue regions are the 95% confidence region obtained with a bootstrap. Error bars represent 1σ errors. For clarity, only half of the points are shown.

f	C_{11}	C_{12}	C_{13}	C_{21}	C_{22}	C_{23}	C_{31}	C_{32}	C_{33}
0.01	2.974e-02	-4.451e-03	-3.956e-03	-4.451e-03	6.678e-04	5.948e-04	-3.956e-03	5.948e-04	5.309e-04
0.02	2.134e-02	-3.574e-03	-3.233e-03	-3.574e-03	6.001e-04	5.441e-04	-3.233e-03	5.441e-04	4.947e-04
0.04	1.088e-02	-2.313e-03	-2.186e-03	-2.313e-03	4.931e-04	4.672e-04	-2.186e-03	4.672e-04	4.444e-04
0.10	2.891e-03	-9.264e-04	-9.930e-04	-9.264e-04	2.980e-04	3.206e-04	-9.930e-04	3.206e-04	3.475e-04
0.20	7.598e-04	-3.564e-04	-4.707e-04	-3.564e-04	1.682e-04	2.228e-04	-4.707e-04	2.228e-04	2.999e-04
0.25	4.175e-04	-2.263e-04	-3.366e-04	-2.263e-04	1.236e-04	1.842e-04	-3.366e-04	1.842e-04	2.806e-04
0.30	2.633e-04	-1.601e-04	-2.696e-04	-1.601e-04	9.836e-05	1.655e-04	-2.696e-04	1.655e-04	2.865e-04
0.40	1.126e-04	-8.136e-05	-1.748e-04	-8.136e-05	5.983e-05	1.273e-04	-1.748e-04	1.273e-04	2.841e-04
0.50	7.789e-05	-5.993e-05	-1.425e-04	-5.993e-05	4.720e-05	1.103e-04	-1.425e-04	1.103e-04	2.738e-04
0.60	7.960e-05	-6.125e-05	-1.341e-04	-6.125e-05	4.823e-05	1.039e-04	-1.341e-04	1.039e-04	2.374e-04
0.70	1.303e-04	-9.217e-05	-1.517e-04	-9.217e-05	6.625e-05	1.082e-04	-1.517e-04	1.082e-04	1.847e-04
0.75	2.249e-04	-1.412e-04	-1.841e-04	-1.412e-04	8.963e-05	1.166e-04	-1.841e-04	1.166e-04	1.566e-04
0.80	6.321e-04	-2.822e-04	-2.747e-04	-2.822e-04	1.268e-04	1.237e-04	-2.747e-04	1.237e-04	1.229e-04
0.90	1.845e-01	-3.132e-04	-4.011e-03	-3.132e-04	5.322e-07	6.824e-06	-4.011e-03	6.824e-06	8.758e-05

Table 3.2: Covariance matrices of the ODR fitting.

example, given α, we can derive β from

$$\frac{\Gamma\left(\alpha+\frac{1}{\beta}\right)}{\Gamma(\alpha)\left(\alpha-\frac{1}{\beta}\right)^{\frac{1}{\beta}}}=\frac{\overline{\omega}}{\widehat{\omega}} \tag{3.23}$$

which must be solved implicitly, and γ from

$$\gamma=\frac{\widehat{\omega}}{\left(\alpha-\frac{1}{\beta}\right)^{\frac{1}{\beta}}} \tag{3.24}$$

The same equations can be used to derive α and γ given β, while β given γ can be derived from

$$\frac{\Gamma\left[\left(\frac{\widehat{\omega}}{\gamma}\right)^{\beta}+\frac{2}{\beta}\right]}{\Gamma\left[\left(\frac{\widehat{\omega}}{\gamma}\right)^{\beta}+\frac{1}{\beta}\right]}=\frac{\overline{\omega}}{\gamma} \tag{3.25}$$

and α from

$$\alpha=\left(\frac{\widehat{\omega}}{\gamma}\right)^{\beta}+\frac{1}{\beta} \tag{3.26}$$

Such relations are very good descriptions of the correlations of the parameters derived from the Monte Carlo simulation, as can be seen in Figure 3.11.

Other usable combinations of the parameters (not addressed in this work) are the higher moments:

$$\mu_n=\left\langle\omega^n\right\rangle=\int_0^{\infty}\omega^n p(\omega)\,\mathrm{d}\omega=\frac{\Gamma\left(\alpha+\frac{n}{\beta}\right)}{\Gamma(\alpha)}\gamma^n \tag{3.27}$$

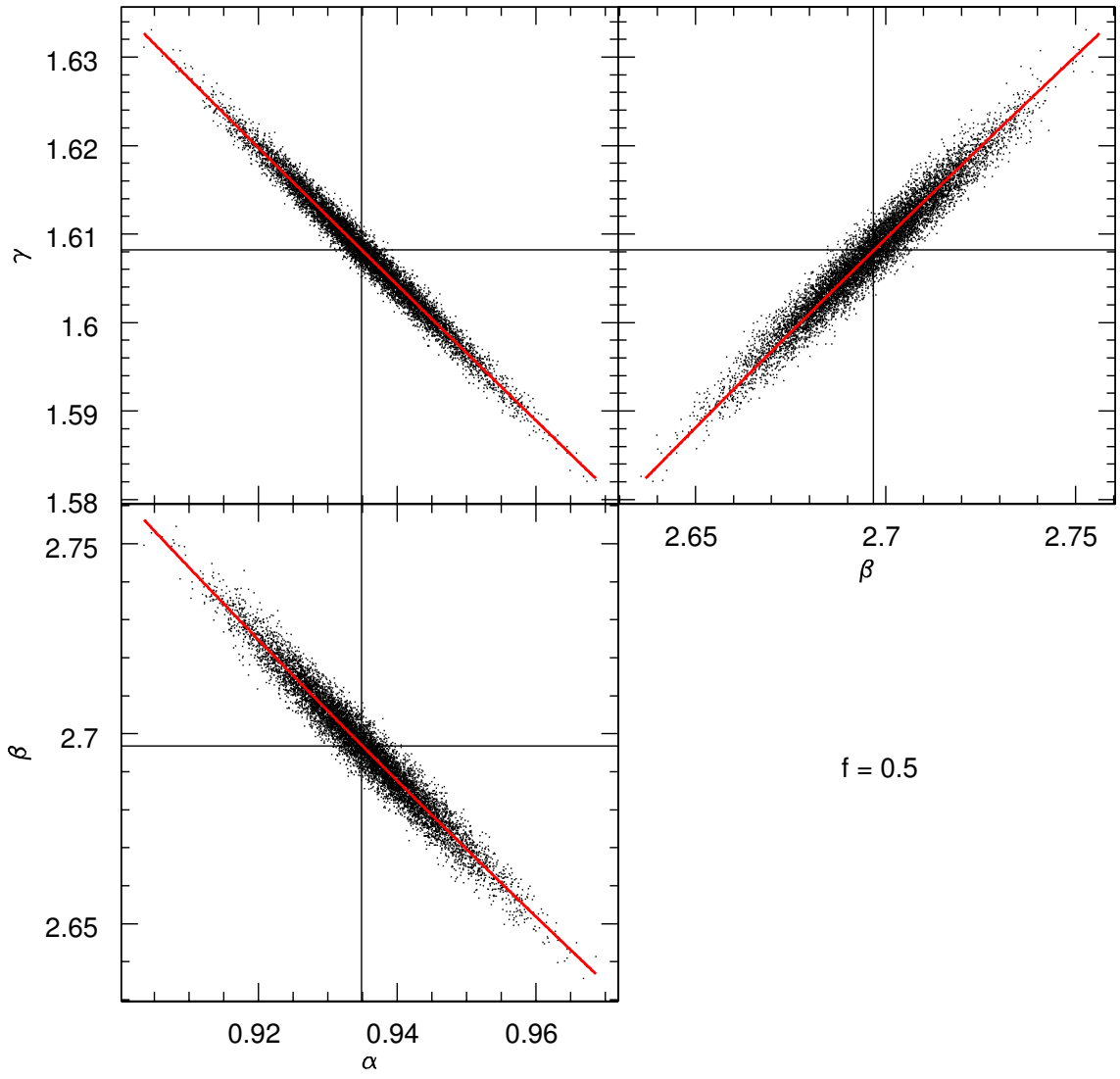

Figure 3.11: Correlations between the fitted parameters in the Monte Carlo Simulation. Vertical and horizontal lines represent the original value of the fit. Red lines represent the relations derived from Equation 3.21 and Equation 3.22.

or their central versions:

$$c_n = \left\langle (\omega - \langle\omega\rangle)^n \right\rangle = \sum_{k=0}^{n} (-1)^k \binom{n}{k} \left\langle \omega^{n-k} \right\rangle \langle\omega\rangle^k \tag{3.28}$$

from which one can also compute the standardized moments. Note that theoretically one can also use such combinations of parameters for fitting the distribution with the method-of-moments, where one would derive the distribution's parameters directly from the unbinned sample, exploiting the constrains mentioned above in at least the same number as the parameters to derive, in our case three of one's choice. This kind of procedure could provide useful for example if one were to repeat the same analysis but with much poorer data.

3.2.2 Scaling relations

Given the parameters α_f, β_f and γ_f and their standard errors, derived from the ODR fitting procedure, we derived scaling relations as in G12 (Equation 3.13 and Equation 3.14). The choice of the functions is somewhat arbitrary, and of course depends on the parametrization of Equation 3.15. We found the followings to be good descriptions of the relations between the fitted parameters and f:

$$\alpha_\mathrm{f} = -13.03 + 12.65 f^{-0.1083} + 3.556 f^{3.517} \tag{3.29}$$

$$\beta_\mathrm{f} = 0.8998 + 3.386 f - 4.210 f^2 + 1.837 f^3 - 2.486 f^4 \tag{3.30}$$

$$\gamma_\mathrm{f} = 1.720 + 3.679 f - 11.57 f^2 + 45.32 f^3 - 76.88 f^4 + 37.73 f^5 \tag{3.31}$$

which we obtained with an ordinary least square procedure. The fitting results are shown in Figure 3.12. From the figure, it appears clear that the points corresponding to $f = 0.01$, $f = 0.02$ and especially $f = 0.9$ are outliers and have much bigger errors than the other points, and any attempt to include them (e.g. changing the fitting functions) proved unsuccessful or introduced significant distortions, worsening the results. Therefore, we decided to exclude $f = 0.01$, $f = 0.02$ and $f = 0.9$ from the fit, and to restrict all our findings on the mass accretion histories to the interval $0.04 \lesssim f \lesssim 0.9$. From the covariance matrix we derived a confidence region (same as in Equation 3.16) corresponding to 1σ, shown with an orange shade in the figure. The lower panels show the residuals in units of the confidence interval at that point.

We used the scaling relations also to check for dependencies of our model from the adopted cosmology. Using the COSMO simulations (subsection 2.2.1), we constructed the stacked distribution (in mass and identification redshift) in the same way as for Le SBARBINE, applied the same fitting procedures and found the fitted parameters. As can be seen in Figure 3.12, these parameters (represented with empty red circles) are in remarkable agreement with the ones derived from Le SBARBINE (full black squares), with only a small discrepancy for β and γ that amounts to a $\sim 5\%$ difference in

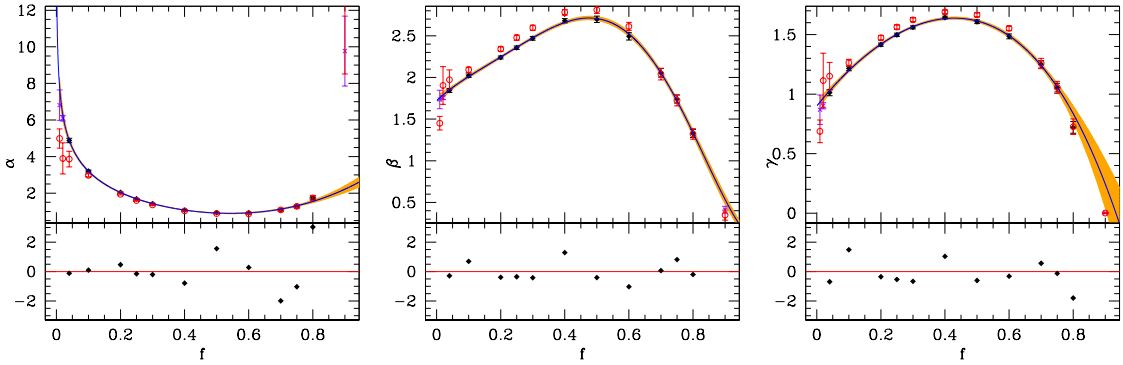

Figure 3.12: Scaling relations for α_f, β_f and γ_f. Black points represent the fitted parameters derived from Le SBARBINE simulations used for the calibration. Purple crosses represent the points excluded from the calibration, namely those corresponding to $f = 0.01$, $f = 0.02$ and $f = 0.9$. Red circles represent the fitted parameters derived from the COSMO simulations. Blue solid lines represent the best fit of Equation 3.29, Equation 3.30 and Equation 3.31 for respectively the left, central and right panel. The 1σ confidence intervals derived from the covariance matrices are represented with an orange shade. Lower panels: residual difference of the black points from the best-fitting curve, in units of the confidence interval at that point.

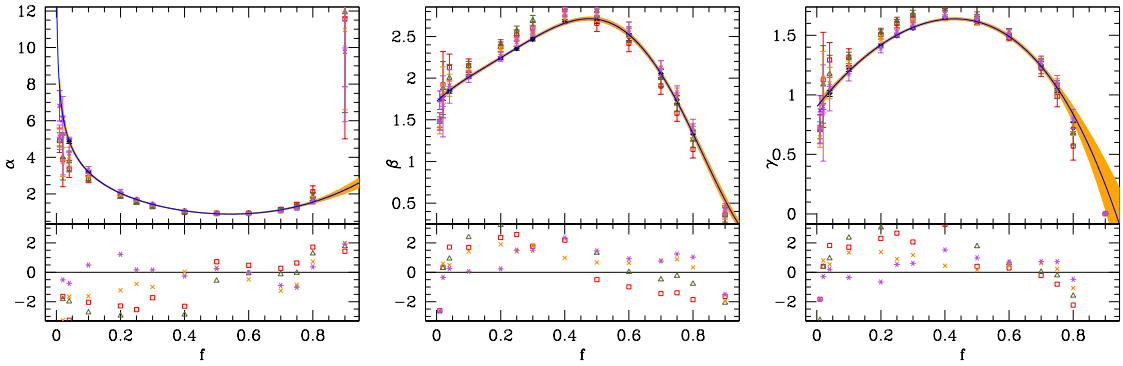

Figure 3.13: Same as Figure 3.12, with colored symbols representing the different COSMO simulations fitted separately. Residuals are in units of the standard errors.

the distribution $p(\omega)$. We also plotted the parameters for the COSMO simulations fitted separately in Figure 3.13, where we can see that in the validity range $0.04 < f < 0.9$ the distributions from the different simulations in the COSMO set are in good agreement with each other, despite having vastly different cosmological parameters.

3.3 The median mass accretion history

From Equation 3.4 it is easy to derive the median value of the distribution $\tilde{\omega} = \sqrt{2\ln(\alpha_f + 1)}$ (Equation 3.8). Unfortunately, Equation 3.15 doesn't have a closed form for the median, and one has to

either use an estimator or compute it numerically. Since Equation 3.15 is a generalized gamma distribution (Stacy and Mihram, 1965), we used the estimator of the median of the gamma distribution proposed by Banneheka and Ekanayake (2010), which we then converted to the median for the generalized gamma. In fact, using the change of variables

$$x = \left(\frac{\omega}{\gamma}\right)^{\beta} \tag{3.32}$$

Equation 3.15 reduces simply to

$$p(x) = \frac{1}{\Gamma(\alpha)} x^{\alpha-1} \mathrm{e}^{-x} \tag{3.33}$$

If $\mu = \alpha$ is the mean of Equation 3.33 and ν is its median (computed numerically) we find that for $\alpha \gtrsim 1$ the linear relation

$$\frac{\mu}{\mu - \nu} \simeq 0.21 + 3\alpha \tag{3.34}$$

holds almost perfectly. Making ν explicit, the proposed estimator is then:

$$\nu_{BE} = \alpha \frac{3\alpha - 0.79}{3\alpha + 0.21}, \qquad \alpha \geq 1 \tag{3.35}$$

and the corresponding median for Equation 3.15 is:

$$\widetilde{\omega} = \gamma \left(\alpha \frac{3\alpha - 0.79}{3\alpha + 0.21}\right)^{\frac{1}{\beta}} \tag{3.36}$$

Unfortunately, in the range $0.436 < f < 0.648$ the value of α derived from Equation 3.29 falls indeed under 1, with a relative error for ν_{BE} up to 2.68% for $f = 0.543$ (Figure 3.14, left panel). However, when converted to $\widetilde{\omega}$, the errors reduces to $\lesssim 0.35\%$ in the full region $0.01 < f < 0.9$ (Figure 3.14, right panel), therefore we will utilize Equation 3.36 for the median in the rest of this work.

Figure 3.15 shows the median $\log(f) - \log(1+z)$ relation derived from the model compared to the distribution of MAHs for halos with $M_0 = M^*$ and identified at $z_0 = 0$. More plots, with masses ranging from $M^*/256$ to $64M^*$ and $z_0 = 0, 0.506, 1.019, 2.041$ and 4.152 are available in Appendix A. The formal 1σ confidence region was computed as

$$\sigma^2_{\widetilde{\omega}} = \left|\frac{\partial \widetilde{\omega}}{\partial \alpha}\right|^2 \sigma^2_{\alpha} + \left|\frac{\partial \widetilde{\omega}}{\partial \beta}\right|^2 \sigma^2_{\beta} + \left|\frac{\partial \widetilde{\omega}}{\partial \gamma}\right|^2 \sigma^2_{\gamma} \tag{3.37}$$

where σ_α, σ_β and σ_γ are the confidence intervals derived in subsection 3.2.2. Following van den Bosch et al. (2014), we also plotted in Figure 3.16 the MAHs linearly as a function of the lookback time to accentuate the late time behavior of the MAHs, which is generally of more interest for studying galaxy formation. It is clear that the model is inadequate outside of the stated region $0.04 \lesssim f \lesssim 0.9$. At lower values, the model departs significantly from the data and start to overestimate the formation redshift. At higher values, the confidence interval starts to diverge, and so does the median (not shown). In the central part, however, the model is in very good agreement with the data, of the same order or better than the previous model (Equation 3.4).

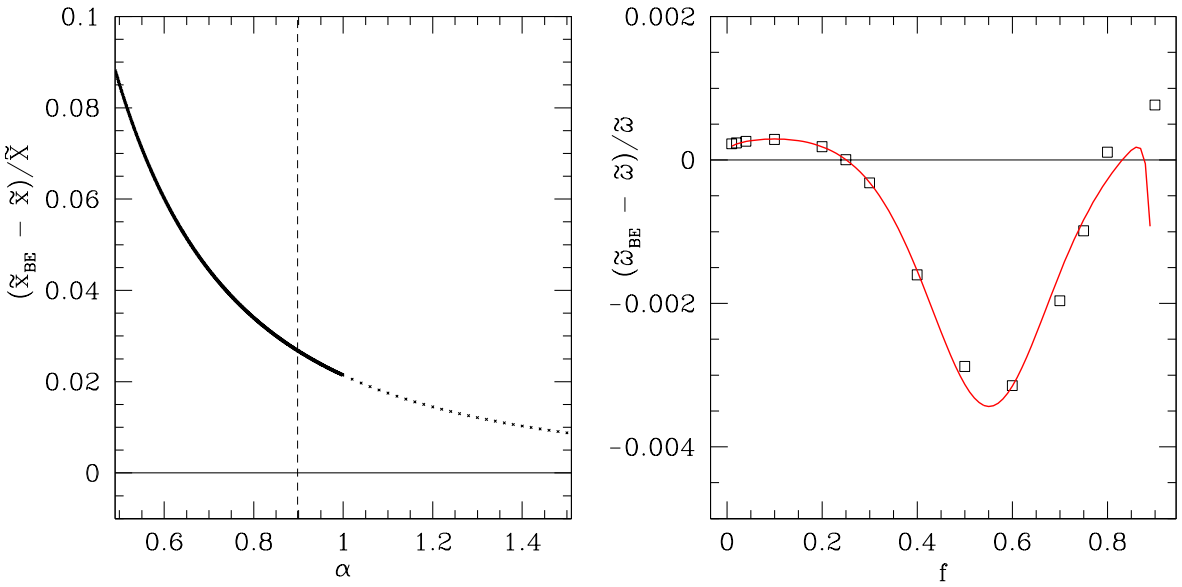

Figure 3.14: Left panel: relative error of the estimator for the median of a gamma distribution, as a function of the parameter α. The vertical dashed line indicate the lowest value for α in our work. Right panel: relative error of the estimator for the median of our generalized gamma distribution, as a function of the parameter f. Open squares represent values obtained using the fitted parameters. The red curve represent values obtained using the scaling relations.

3.4 Scatter in the mass accretion histories

In order to characterize the halo-to-halo scatter, we also performed a preliminary analysis of the percentile distribution of the MAH tracks, where we define as 'track' the trajectory of a halo in the $f - z$ plane. Our motivation is to understand how a track move in the distribution along its evolution. Does a halo who is more massive than the median at the formation or at the first identification stay that way at all redshifts? Or does it move randomly in the distribution?

3.4.1 Global scatter

To quantify the movement of a halo in the distribution, we defined the quantity $\Delta(z)$ as the typical scatter of tracks, as follows. First, we fixed the identification redshift z_0 and the mass bin for the final mass M_0. For every halo in that bin, we computed the percentile of its track with respect to the distribution at fixed z $p(f|z)$ for every z[2], obtaining the percentile track $p_i(z)$. From the percentile track we subtracted its median value, obtaining the residual track $\delta p_i(z) = p_i(z) - \overline{p}_i$ which oscillates around 0. After superimposing all δp_i we computed the interquartile range of the distribution $p(\delta p_i|z)$ at fixed z, which we called $\Delta(z)$. Figure 3.17 illustrates a sample of such distributions, high-

[2]We remind the reader that the MAH tracks are discrete in z, due to the discreteness of the snapshots of the simulations.

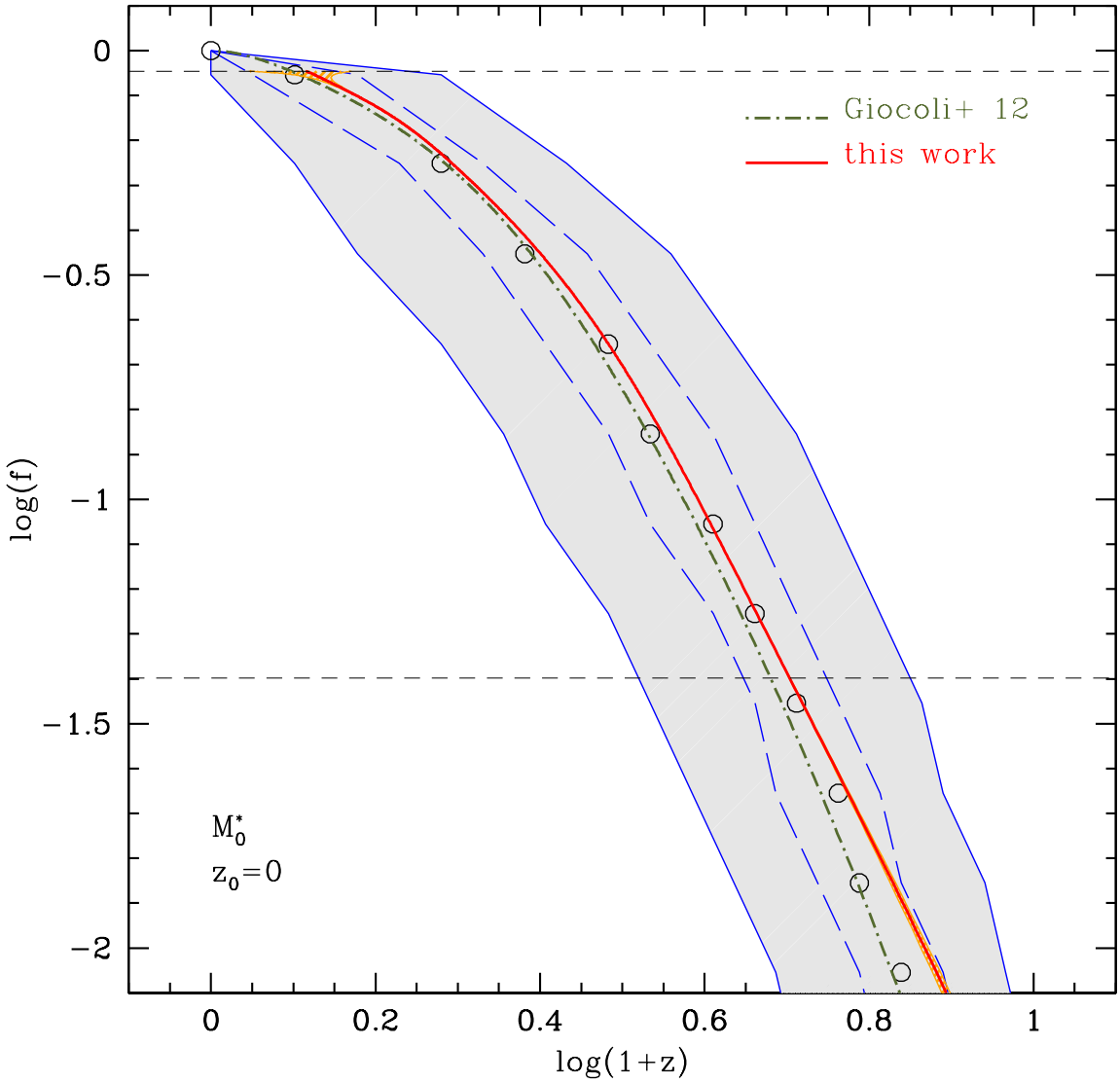

Figure 3.15: Distribution of MAHs for halos with $M_0 = M^*$ identified at $z_0 = 0$. Open circles represent the median of the distribution. The dot-dashed green line represent the median z_f obtained with Equation 3.7 and Equation 3.8, while the solid red line represent the same, but obtained with Equation 3.7 and Equation 3.36. The orange shade around the red line represents the 1σ confidence region. Horizontal dashed lines represent the validity region $0.04 < f < 0.9$.

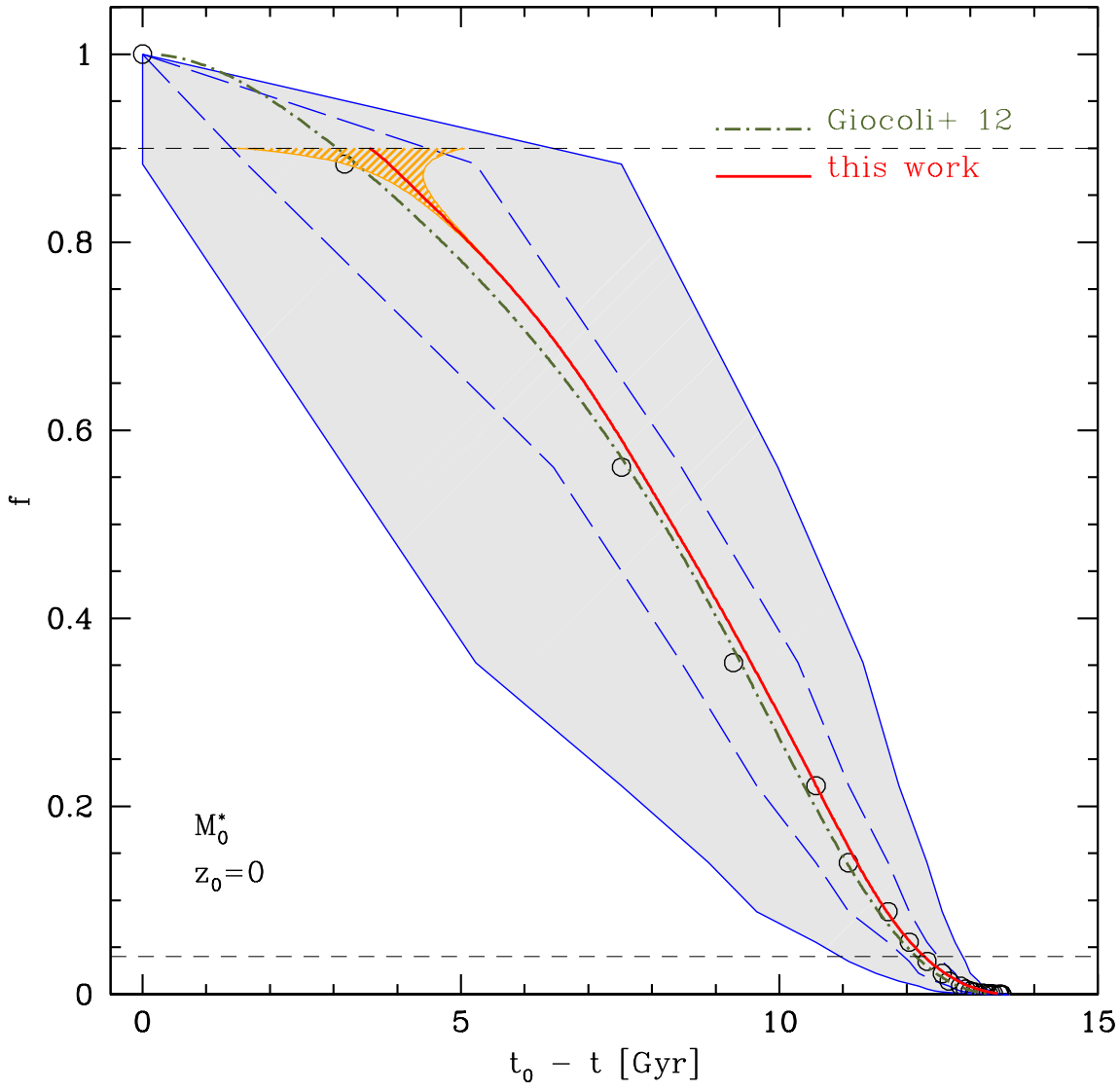

Figure 3.16: Same as Figure 3.15, but with f plotted linearly against the lookback time.

lighting some of the δp_i. Figure 3.18 shows $\Delta(z)$ for different mass bins and identification redshifts. Each curve has a pronounced minimum whose position and value depends on the mass of the final halo, and it is shifted for different identification redshifts.

If we assume that tracks typically remain roughly in the same region of the percentile distribution for the entirety of their evolutions, our interpretation of the shape of $\Delta(z)$ is that there is a characteristic redshift, possibly related to the formation redshift, at which a track is statistically closer in percentile to its typical value. In other words, the mass of the halo at the characteristic redshift is a better indication of its typical position in the distribution than the initial mass or the mass at low redshift. If the previous assumption does not hold, however, no such conclusion can be reached.

3.4.2 Individual properties

To characterize the behavior of the MAH tracks in a more general way, we looked at properties of individual tracks. For each track, we defined the followings:

- $\overline{p}$: the already defined median percentile of the track
- Δ: interquartile range of the distribution of the percentiles of the track
- ϵ: maximum excursion of the distribution of the percentiles of the track
- χ: number of "crossings" of the track, i.e. the number of times the track crosses the value p=0.5
- ρ: Pearson's correlation of the track
- $\overline{\mathrm{d}p}$: median of the differential of the percentile track
- δ: interquartile range of the distribution of the differential of the percentile track
- p_1: percentile at the first identification
- p_f: percentile at the formation redshift

For each of the above we computed the distribution, whose median and interquartile range are quoted in Table 3.3, and the correlations between each properties and all others, shown in Figure 3.19 and Figure 3.20. Note that almost every distribution is somewhat peaked, except for the distribution of p_1. This is mostly due to the mass resolution of the simulations, since the first identification correspond by definition to the time at which the halo first has ~10 particles. At fixed final mass, and therefore number of particles, this correspond effectively to a threshold for f (corresponding, for the simulation Ada and $M = M^*$, to $f_{th} = 4 \times 10^{-5}$), which translate to a cut in the distribution at fixed z.

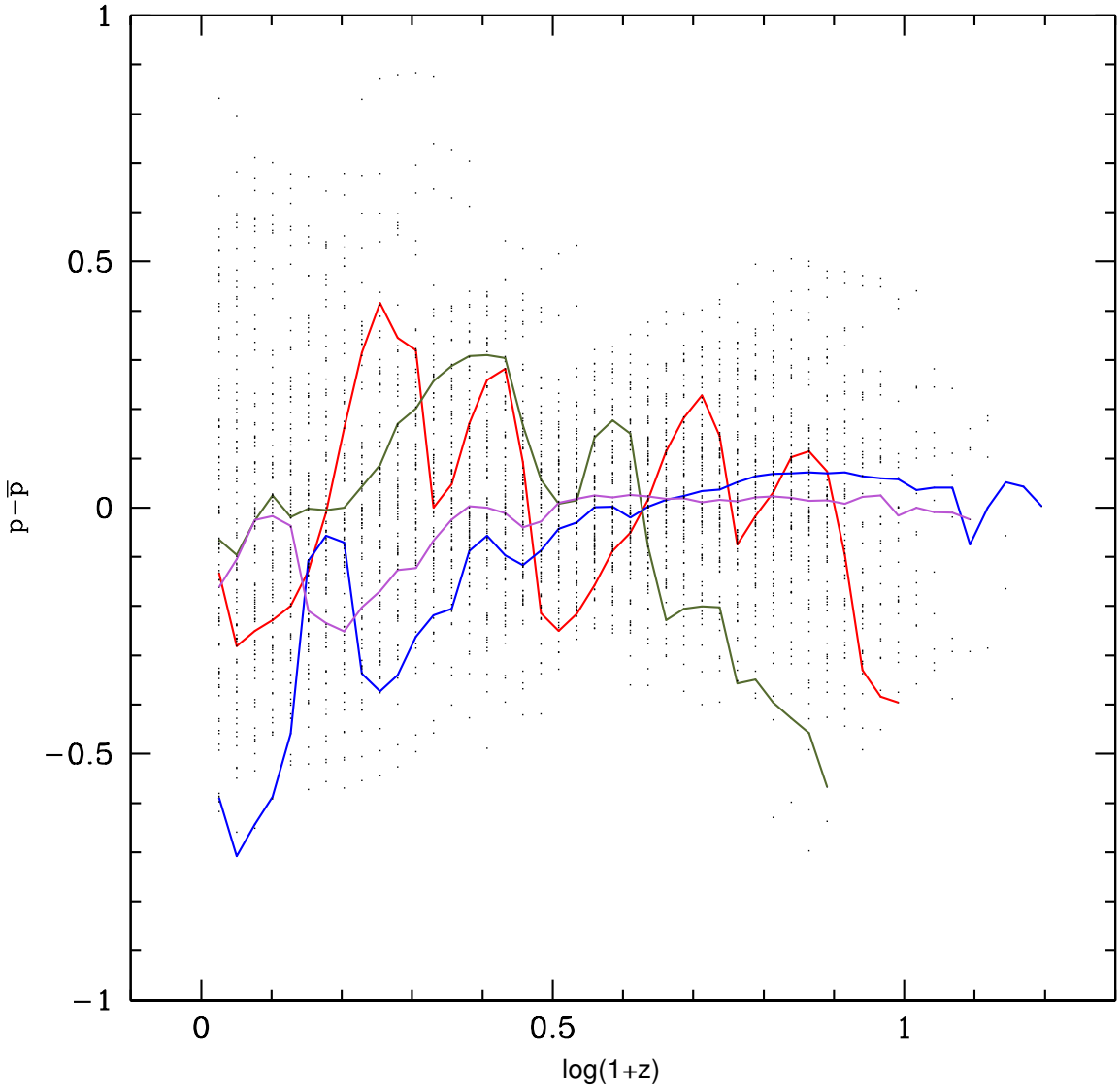

Figure 3.17: Residual tracks $\delta p_i = p_i - \overline{p}_i$. Black dots represent ~100 tracks superimposed, 4 of which, randomly chosen, are highlighted with colored curves.

From Figure 3.20, it is clear that there is no discernible correlation between any of the above properties, with the strongest linear correlation factor being $r \lesssim 0.6$. However, even if our results are not conclusive, they may serve as a characterization of the percentile properties of MAH tracks of halos in simulations. A similar analysis can be performed on synthetic MAHs, for example generated with the algorithm proposed by Giocoli et al. (2009), and the results compared with our findings to check the validity of the method.

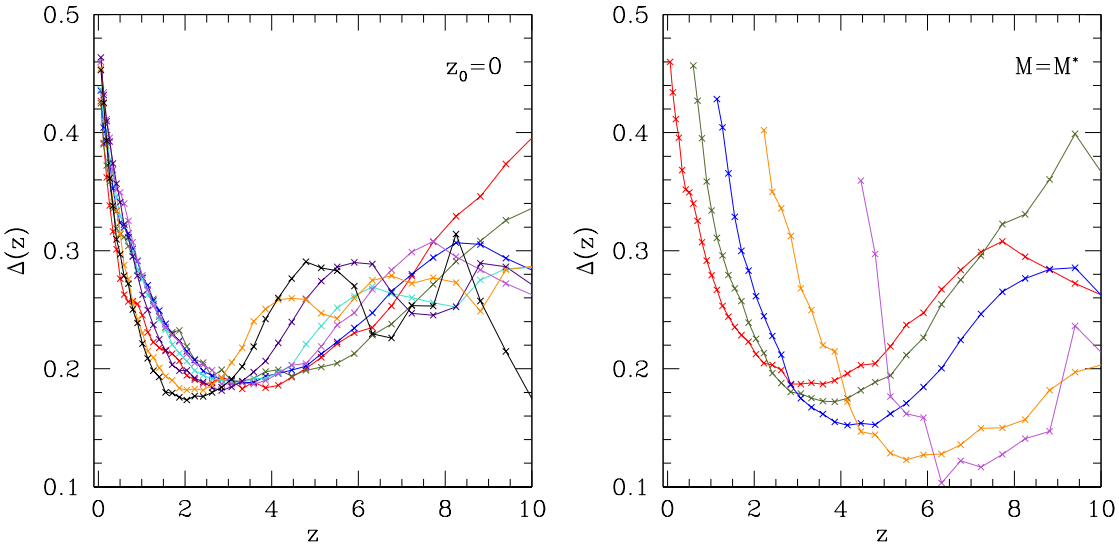

Figure 3.18: Percentile scatter $\Delta(z)$ for different mass (left panel) and identification redshifts (right panel). Colors are in ascending order of red, green, blue, cyan, magenta, yellow, orange and black, corresponding to $M = M*/256$, $M^*/64$, $M^*/16$, $M^*/4$, M^*, $4M^*$, $16M^*$ and $64M^*$, and $z_0 = 0$, 0.506, 1.019, 2.041 and 4.152.

	$\overline{p}$	Δ	ϵ	χ	ρ	$\overline{\mathrm{d}p}$	δ	p_1	p_f
median	0.488	0.298	0.767	3	-0.121	-0.0103	0.0876	0.220	0.509
IQR	0.362	0.168	0.216	3	0.820	0.0148	0.0413	0.331	0.477

Table 3.3: Median value and interquartile range of the distribution of the properties of MAH tracks.

3.5 The full distributions of mass accretion histories

Distributions of z_f at fixed f are shown in Figure 3.21, Figure 3.22, Figure 3.23 and in Appendix B for the identification redshift $z_0 = 0$. Data are taken from logarithmic slices of f, indicated in each panel, around a central value f_i, and are compared to the distribution extracted from the model

$$p(z_f) = p\left[\omega(z_f, f_i, M_0, z_0)\right] \frac{\mathrm{d}\omega}{\mathrm{d}z_f} \tag{3.38}$$

where the derivative $\frac{\mathrm{d}\omega}{\mathrm{d}z_f}$ is computed numerically for simplicity. The model is in general in very good agreement with the histogram of the data, with only a very slight shift to higher values of z_f, although this could be an artifact of the binning. Significant deviations begin to appear in the outskirts of the distribution, both in the far tails of $p(z_f)$ and for high/low values of f_i, which, given the previous discussions, is to be expected. Residuals, in term of the poissonian error of the data histograms, are provided both for each distribution at various f_i and as a color map for the full distribution. The color scheme used for the map is the colorblind friendly, black-and-white friendly, perceptually uniform colormap `viridis` (Smith and van der Walt, 2015).

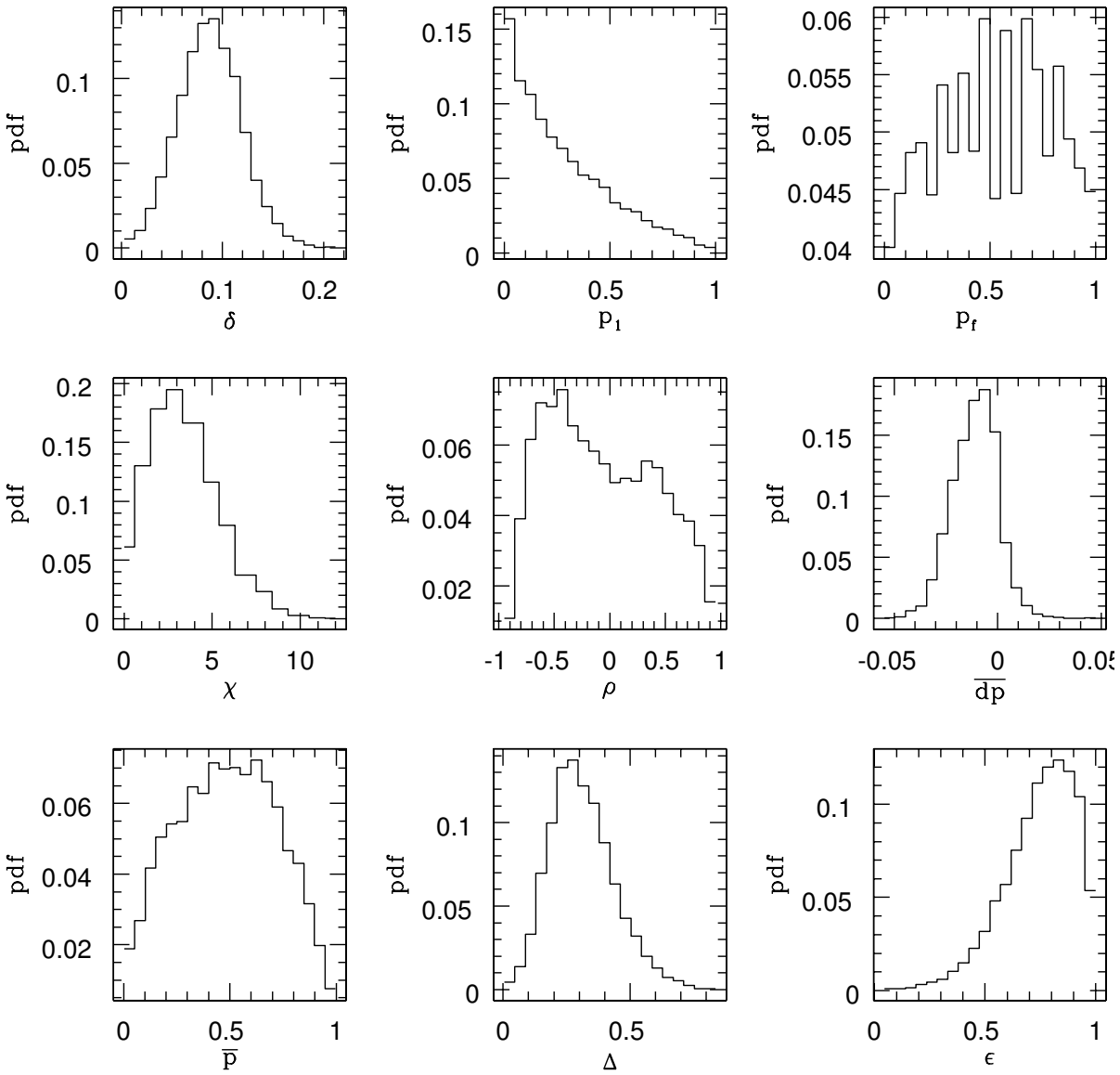

Figure 3.19: Probability distributions of the individual properties of MAH tracks described in the main text for halos with $M = M^*$ and $z_0 = 0$.

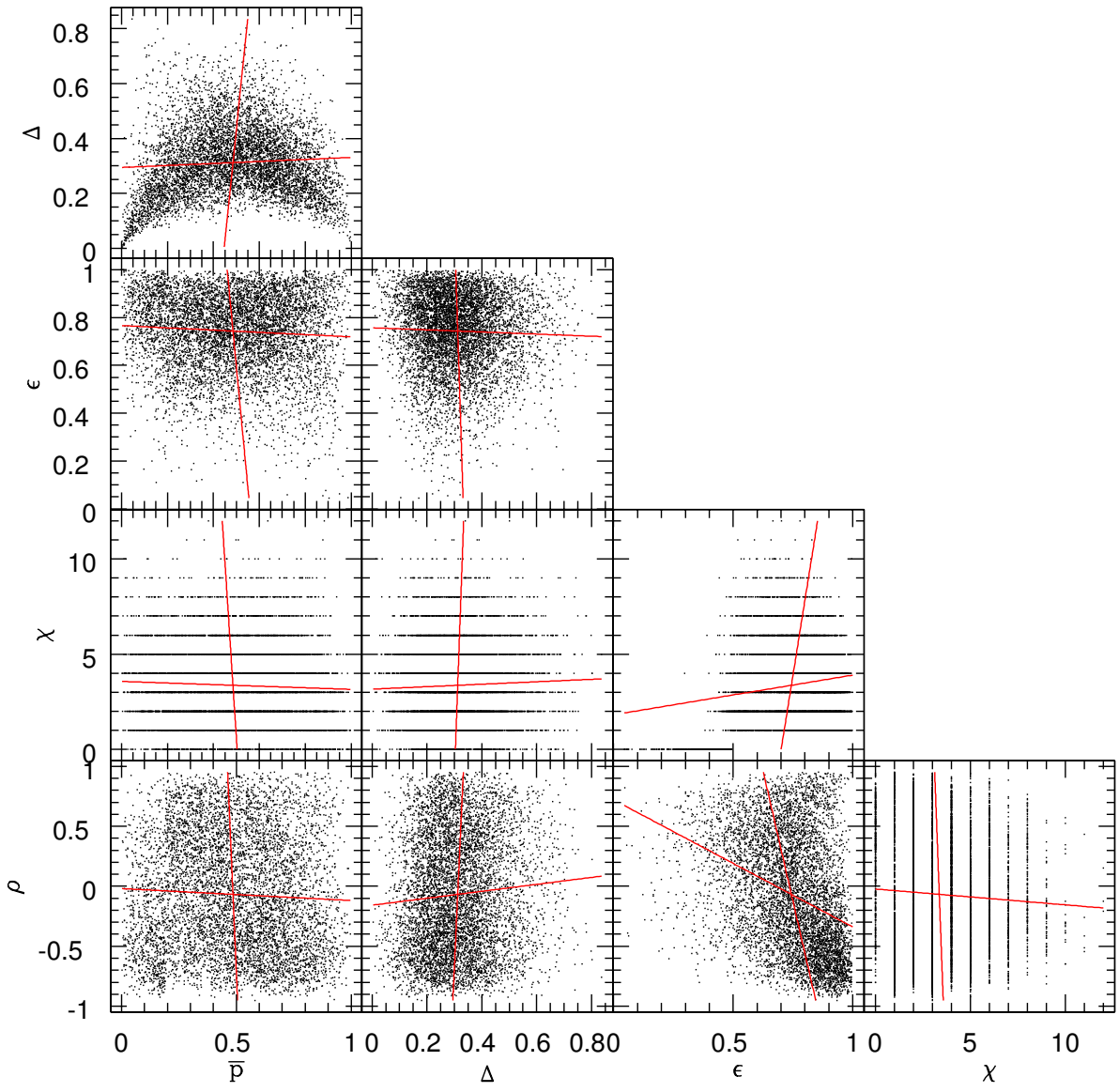

Figure 3.20: Correlations of the individual properties of MAH tracks described in the main text for halos with $M = M^*$ and $z_0 = 0$. Red lines represent the main directions of the linear regression.

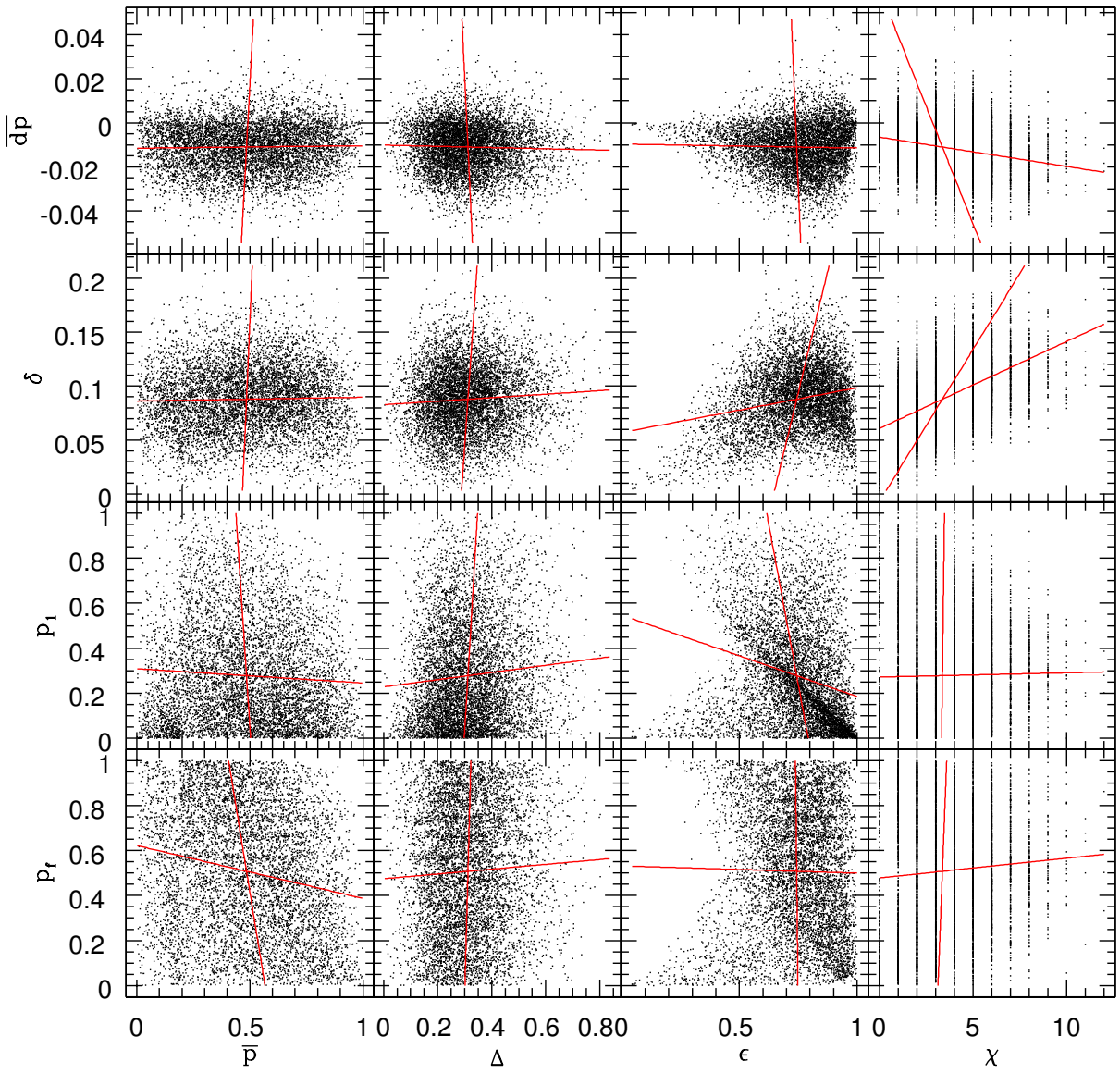

Figure 3.20: Correlations of the individual properties of MAH tracks described in the main text for halos with $M = M^*$ and $z_0 = 0$. Red lines represent the main directions of the linear regression. *(cont.)*

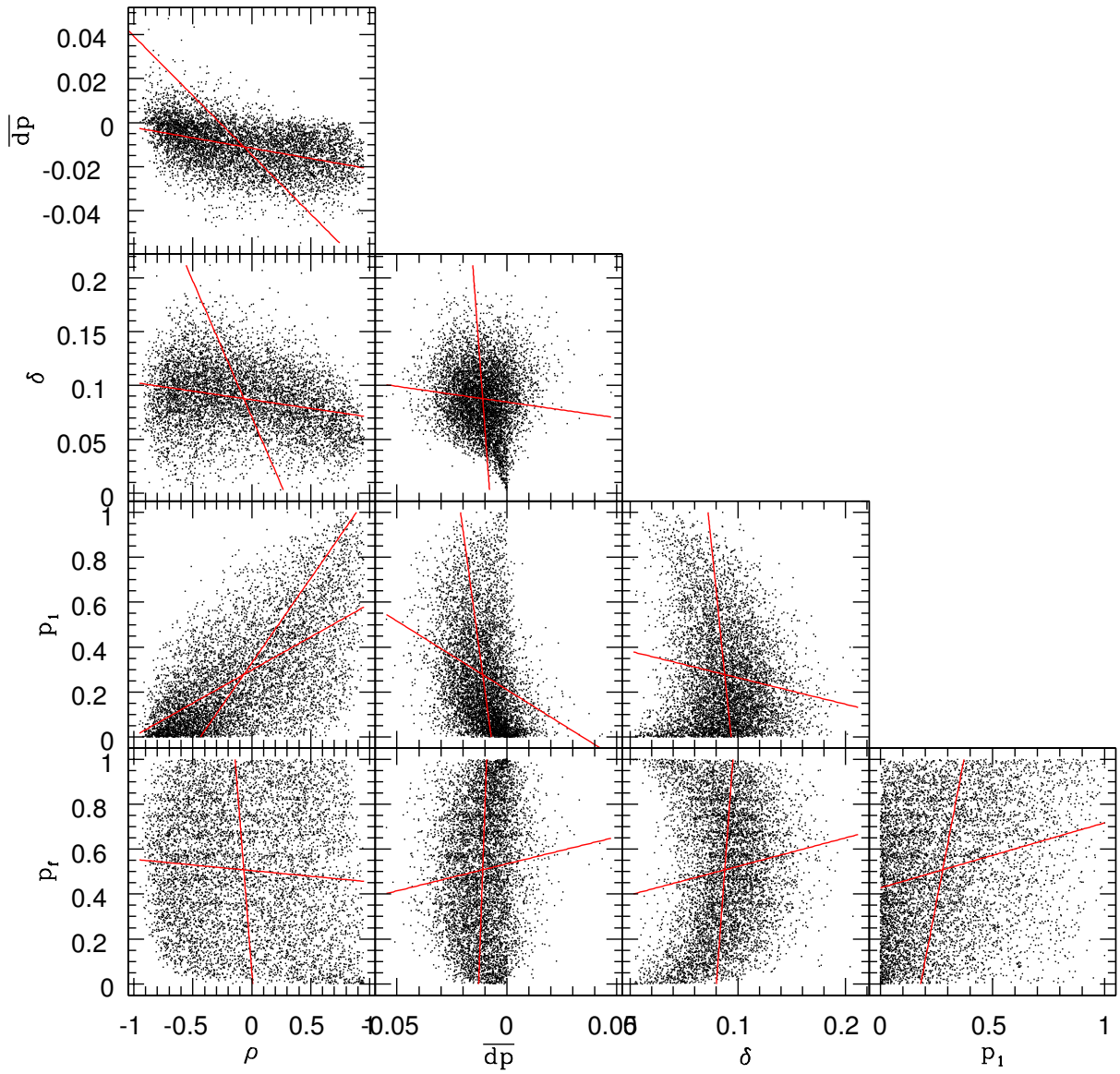

Figure 3.20: Correlations of the individual properties of MAH tracks described in the main text for halos with $M = M^*$ and $z_0 = 0$. Red lines represent the main directions of the linear regression. *(cont.)*

We also show in Figure 3.24 a direct comparison of the distribution of the formation redshift z_{50} derived from the model to the one extracted from the data. The correspondence is remarkable, with typical errors around 2%. The errors are more or less constant for all mass bins, from $M^*/256$ to $16M^*$, with slightly larger errors for $64M^*$, as shown in the figures in Appendix C. The bin width barely influences these results, with similar errors for $\Delta = 2$ and somewhat larger errors for $\Delta = 1.1$: the effect of the increased precision in the determination of the mass for the data is countered by the much lowered statistics.

Finally, we show in Figure 3.25 the mean formation redshift $z_{0.50}$ as a function of the final mass at $z_0 = 0$, derived from the distributions constructed with the model. At low masses our model overestimates the formation redshift by $\Delta \log(1 + z_{0.50}) \approx 0.01$, or by about 2.4%, a similar error to previous model like the one presented by Giocoli et al. (2007). However, while their errors stay more or less constant over five decades of mass, the errors of our model decrease steadily towards higher masses. Similar conclusions can be drawn when comparing to other models, like the ones presented by McBride et al. (2009) and Giocoli et al. (2012).

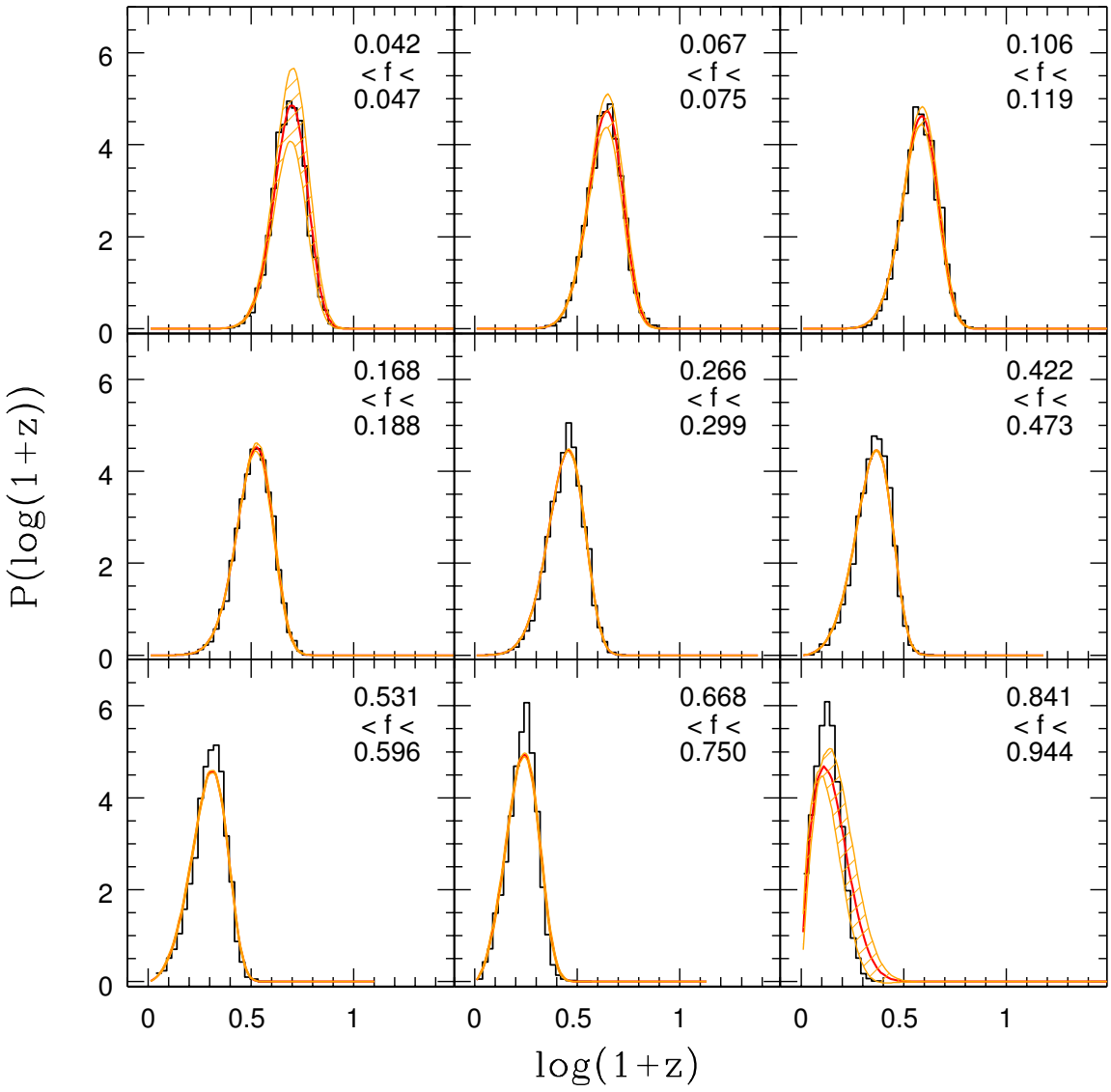

Figure 3.21: Distribution of z_f at fixed f for halos with $M_0 = M^*$ and $z_0 = 0$. The data are represented with histograms, while the distribution derived from the model is represented as a red line, with an orange shade representing the formal 1σ confidence region.

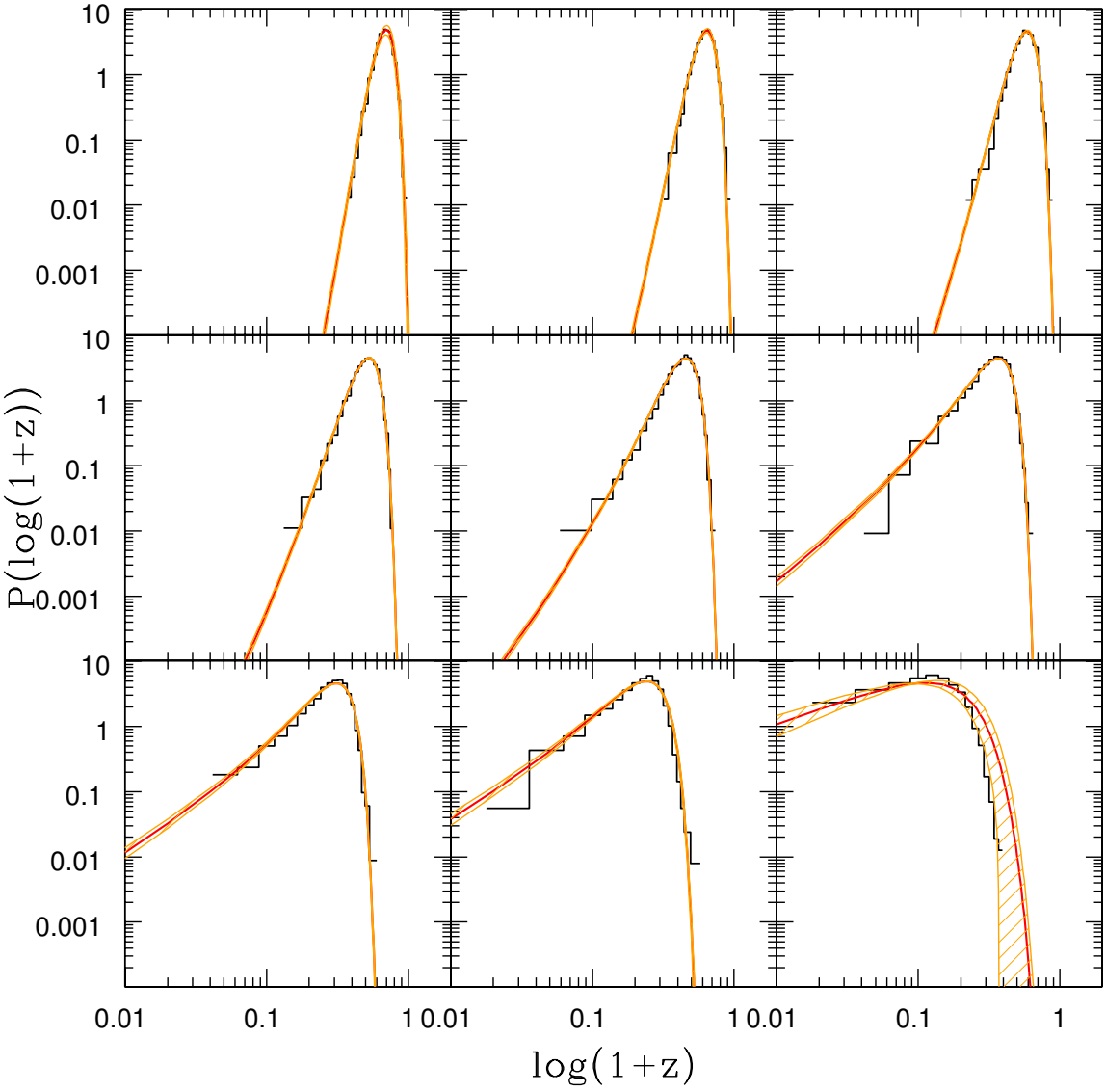

Figure 3.22: Same as Figure 3.21, but in the log-log plane.

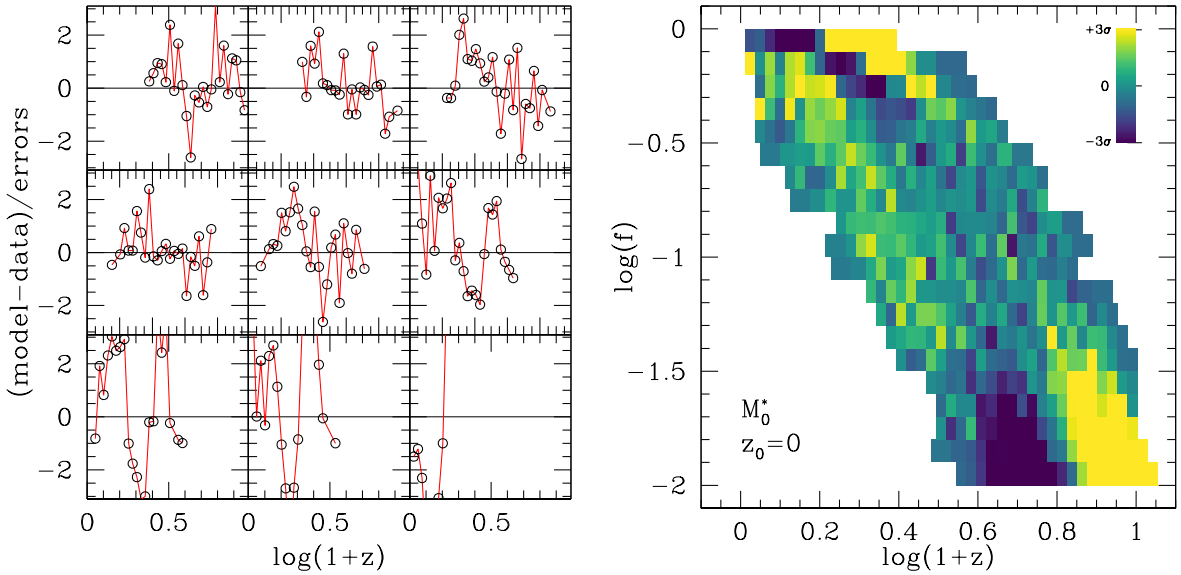

Figure 3.23: Left panel: residuals between the model and the data, as in Figure 3.21, in units of 1σ poissonian errors. Only the points inside the 3σ interval are shown. Right panel: color map of the residuals for the full distribution. The color scale varies between -3σ (dark blue) and 3σ (light yellow).

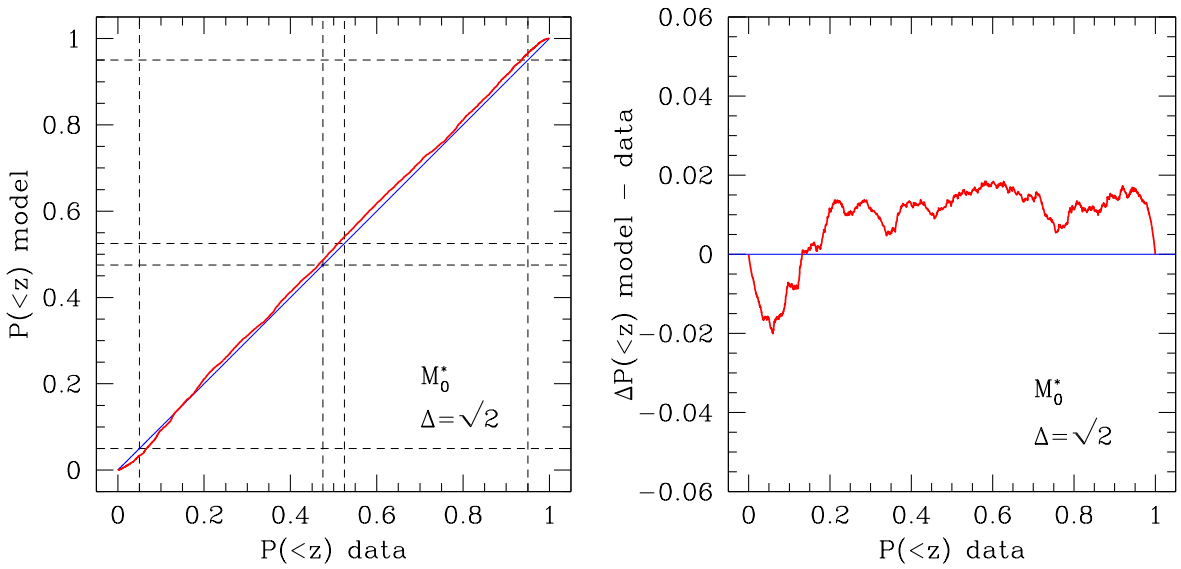

Figure 3.24: Left panel: Comparison of the cumulative distributions of the formation redshift z_{50} derived from the data and the model, for $M_0 = M^*$ and $z_0 = 0$. Data are taken in a mass bin with a multiplicative width $\Delta = \sqrt{2}$, i.e. the limits of the bin are $M^*/\sqrt{2}$ and $\sqrt{2}M^*$. The red line represents the cumulative derived from the model as a function of the cumulative derived from the data, with the blue line as a reference, representing perfect correspondence. Vertical and horizontal dashed lines mark the lower, central and higher 5% percentiles. Right panel: difference of the two cumulatives.

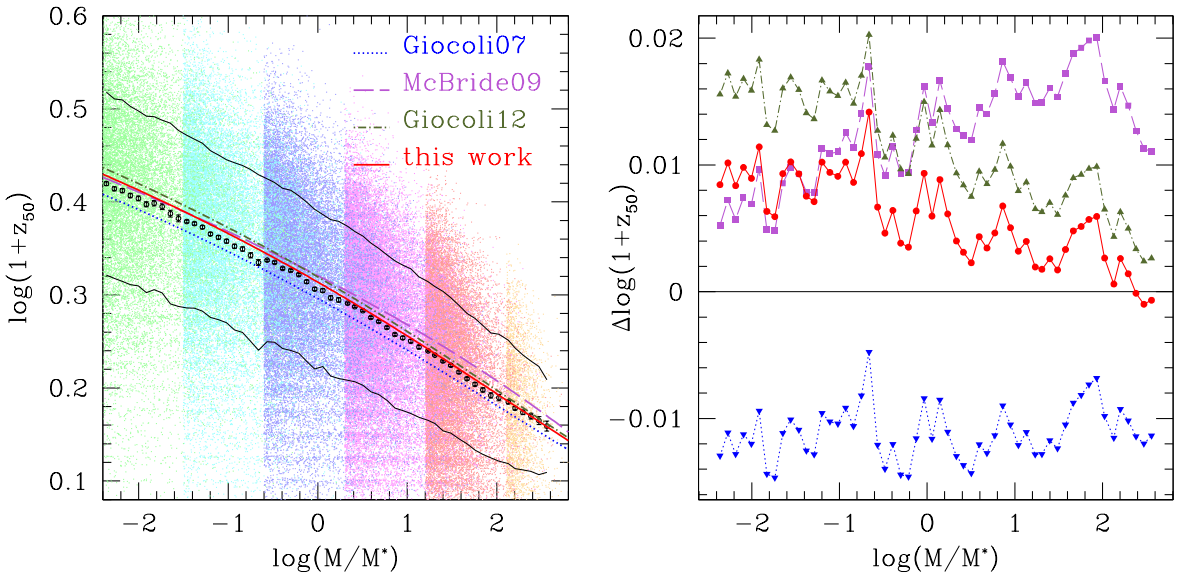

Figure 3.25: Left panel: mean formation redshift $z_{0.50}$ as a function of the final mass at $z_0 = 0$. Colored points represent halos from Le SBARBINE, color-coded from green (Ada) to orange (Flora). Due to their different box size, each simulation probes a different scale, while maintaining a good statistic in the range 10^{10}–10^{15} $M_\odot$ h^{-1}. Open circles represent the mean at fixed mass from the data, with errorbars representing the standard error of the mean. Black solid curves represent the 1σ halo-to-halo variance. The red solid line represents the mean formation redshift derived from our model, compared to the results of Giocoli et al. (2007) (blue dotted line), McBride et al. (2009) (purple dashed line) and Giocoli et al. (2012) (dash-dotted green line). Right panel: corresponding logarithmic difference between the data and the models.

Applications 4

4.1 General applications

Many models exist for the median or mean halo properties that can be derived from the mass accretion histories, but the study of the full distribution, or just the scatter, is often still little explored. Many astrophysical observables as well have median relations that are by now well characterized, but show significant scatter. The star formation rate, color and fraction of star-forming galaxies in groups are good examples of such observables (Poggianti et al., 2009). A question still without a clear answer is if and how the scatter of these properties depends on the mass of the hosting halo, on the mass of its progenitor at a previous significant redshift, like z_f or $z = 2$, corresponding to the peak of the star formation history (Madau and Dickinson, 2014), and on the scatter of the distribution at that redshift. The latter is easily derived from our model. Having defined a cosmology and computed the appropriate $\delta_c(z)$ and $\sigma^2(M)$, the distribution of progenitor masses at fixed z is

$$p_f(f) = p(\omega_z)\frac{\mathrm{d}\omega}{\mathrm{d}f} \tag{4.1}$$

with the derivative computed numerically or analytically from Equation 3.2. Given the distribution, the scatter easily follows as

$$\sigma_z^2 = \int_0^\infty f^2 p_z(f)\,\mathrm{d}f \tag{4.2}$$

and can be computed for a desired range of masses. The resulting $\sigma_z - M_0$ relation can then be compared to similar relations for various galaxy properties, looking for correlations.

Another possible application of our results consists in the study of the properties of "extremal halos", defined as the halos whose main progenitor is extremal in the distribution at fixed redshift, for example being more massive than 95% of all other halos. Given our model, one can compute thresholds that can be used to select halos, of which it would be interesting to study the properties, like the concentration, fraction of substructures, shape, and the various proxies for the relaxation defined in section 2.4. We present here, as a proof-of-concept, the mass-concentration relation for halos in the lower, central and higher 5% of the distribution of redshift at fixed mass fraction $f = 0.5$.

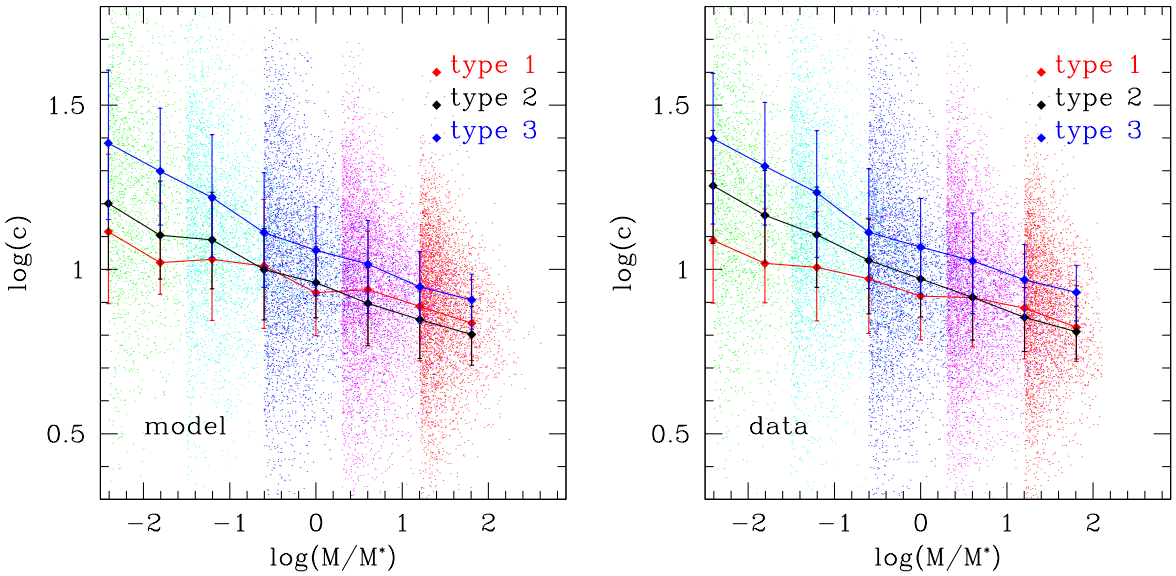

Figure 4.1: Median concentration at fixed mass for halos of type 1 (red), 2 (black) and 3 (blue). Colored points represent halos from Le SBARBINE. Halo are selected using the procedure described in the main text for the model (left panel) and the data (right panel).

The idea is similar to the mentioned extremality with respect to the mass distribution at fixed redshift, which however can be done not only fixing the redshift for all halos (for example $z = 2$) but also using a different value for the redshift for each individual halo, for example their formation redshift z_{50}. Regardless, the prescription for the distribution at fixed mass fraction is is much easier to implement.

We proceed as follows. For every halo, we construct the cumulative distribution $P\left(< z; f = 0.5\right)$ relative to its final mass M_0 and identification redshift z_0, and determine the values of the 0.05, 0.475, 0.525 and 0.95 percentiles. We then define the halo as type 1 if its formation redshift z_{50}, measured from its MAH, is lower then the 0.05 percentile, type 2 if it is between the 0.475 and 0.525 percentile, type 3 if it is higher than the 0.95 percentile, and type 0 otherwise. A similar selection can be done by comparing the formation redshift of each halo to their actual distribution, computed from the data inside a mass bin with a certain (preferably small) width. Halos are then categorized in type 1, 2, 3 and 0 in the same way.

It is worth noticing that defining the extremality with the data relies on having a large statistics for the construction of the distribution of formation redshifts, to compute reasonably accurate percentiles. This is not always the case if one wants to keep the width of the mass bin as small as possible, especially at large masses, and one is then forced to accept a compromise between noise and statistical significance. By contrast, using the procedure we described for our model each halo is compared to its own theoretical distribution, which can be done regardless of statistics, with the only assumption being the reliability of the model. This is possible because in the scaled variable ω_f there is no

dependence on mass in the distribution, so the model is built on the distribution of all halos. Using our model is equivalent to comparing each halo to the distribution of all others, rescaled to eliminate the mass dependence.

In Figure 4.1 we show the median concentration at fixed mass for halos of type 1, type 2 and type 3 separately, using the two selection procedures described. The results agree quite well, showing that our model provide a good description of the distribution of the formation redshift (already seen in Figure 3.24). On the other end, it also means that the increase in precision obtained by not binning is either countered by the not perfect accuracy of the model or washed out by the large halo-to-halo variance.

Lastly, taking advantage of the universality of the model discussed in subsection 3.2.2, in the rest of this chapter we will focus on applications to mass accretion histories in non-standard cosmologies. In particular we will focus on cosmologies with a component of massive neutrinos, an area of research that is rapidly growing in interest in the present day cosmological investigation.

4.2 Cosmological interest in massive neutrinos

It became clear in recent years that neutrinos experience flavor oscillations via a mechanism proposed by Pontecorvo (1957), an intuition first confirmed by Cleveland et al. (1998) observing relative abundances in the solar flux. Many experiments in the following years confirmed such results, e.g. Fukuda et al. (1998), Ahmad et al. (2001), Abe et al. (2008), Smy (2013). Barring exotic effects due to violations of equivalence principle (Gasperini, 1988), violations of the Lorentz group symmetry (Coleman and Glashow, 1997), torsion (de Sabbata and Gasperini, 1981) and others, the presence of flavor oscillations is usually interpreted as an effect of the mixing of the mass and flavor eigenstates, which in turn implies that neutrinos have non-vanishing masses. However, the Standard Model of particle physics does not give any prediction for their masses, on the contrary neutrinos explicitly don't couple to the Higgs field (Peskin and Schroeder, 1995). Oscillation experiments themselves don't provide measurements on the neutrino masses, but rather on the mass splits (Forero et al., 2012):

$$\Delta m^2_{21}\left(10^{-5}\mathrm{eV}^2\right) = 7.62^{+0.58}_{-0.50} \qquad \Delta m^2_{31}\left(10^{-3}\mathrm{eV}^2\right) = 2.55^{+0.19}_{-0.24} \tag{4.3}$$

Other kind of experiments, like observations of beta decays Drexlin et al. (2013) put an upper bound to the mass of the lightest neutrino of $m_0 < 2.05 - 2.3$ eV which, combined to the mass splits, restricts the sum of the neutrino masses $M_\nu = \sum_i m_i$ to the approximate range

$$0.06\ \mathrm{eV} \lesssim M_\nu \lesssim 6\ \mathrm{eV} \tag{4.4}$$

The value of M_ν is also of cosmological interest in an universe where the neutrinos are massive. From early universe physics we know that there is a cosmic background of relic neutrinos with a number density per flavor $n_\nu = \frac{3}{11} n_\gamma$, where n_γ is the photon number density, and an energy density in the ultra-relativistic regime of

$$\rho_\nu = \frac{7}{8}\left(\frac{4}{11}\right)^{4/3} N_{\rm eff} \tag{4.5}$$

where $N_{\rm eff} = 3.046$ is the effective number of neutrinos (Mangano et al., 2005). The cosmic neutrinos decouple roughly when the universe has temperature $T_{\rm dec} \approx 1$ MeV, around 1 second after the Big Bang (Lesgourgues and Pastor, 2006). Even if the neutrinos are massive, their mass is so small that the decoupling happens in the ultra-relativistic regime, and they initially behave like radiation. If the neutrinos are indeed massive, from the mass splits Equation 4.3 we see that at least two of the three mass eigenstates are non-relativistic today, and contribute to the present matter energy density as

$$\Omega_\nu = \frac{\sum_i m_i}{93.14 h^2 \text{ eV}} \tag{4.6}$$

with a corresponding matter fraction

$$f_\nu = \frac{\rho_\nu}{\rho_{\rm CDM} + \rho_{\rm b} + \rho_\nu} = \frac{\Omega_\nu}{\Omega_m} \tag{4.7}$$

The corresponding bounds to the energy density from Equation 4.4 are

$$0.0013 \lesssim \Omega_\nu \lesssim 0.13 \qquad 0.0043 \lesssim f_\nu \lesssim 0.43 \tag{4.8}$$

where we assumed $h \approx 0.7$ and $\Omega_m \approx 0.3$. More stringent upper bounds are imposed by cosmological observations, in particular by the CMB power spectrum (Planck Collaboration et al., 2016):

Parameter	TT	TT+len	TT+len+ext	TT,TE,EE	TT,TE,EE+len	TT,TE,EE+len+ext
$\sum m_\nu$ [eV]	<0.715	<0.675	<0.234	<0.492	<0.589	<0.194

corresponding to:

$$\Omega_\nu \lesssim 0.0043 - 0.0157 \qquad f_\nu \lesssim 0.014 - 0.052 \tag{4.9}$$

The non-relativistic transition happens at

$$1 + z_{nr} \approx 1890 \left(\frac{m}{1 \text{ eV}}\right) \tag{4.10}$$

after which neutrinos start to behave like hot dark matter with a free-streaming wavenumber

$$k_{FS}(t) = \left(\frac{4\pi G \bar{\rho}(t)\, a^2(t)}{v_{\rm th}^2(t)}\right)^{1/2} = 0.8\, \frac{\sqrt{\Omega_\Lambda + \Omega_m (1+z)^3}}{(1+z)^2} \left(\frac{m}{1 \text{ eV}}\right) h \text{ Mpc}^{-1} \tag{4.11}$$

where $v_{\rm th} = \frac{\langle p \rangle}{m}$ is the thermal velocity and the last equality is valid during matter or Λ domination. Since the comoving free-streaming length λ_{FS}/a decreases like $\left(a^2 H\right)^{-1} \propto t^{-1/3}$, the comoving free-streaming wavenumber has a minimum at the non-relativistic transition:

$$k_{nr} \simeq 0.018\, \Omega_m^{1/2} \left(\frac{m}{1\ \text{eV}}\right)\ h\ \text{Mpc}^{-1} \tag{4.12}$$

4.3 Effects of neutrinos on the evolution of perturbations

The net effect of the neutrino free-streaming is to damp the matter perturbation at small scales, slowing the growth of structures and reducing the matter power spectrum at $k > k_{nr}$. Therefore, this could have consequences for the formation of halos and their history, leading in particular to delayed formation times. A comparison of the matter power spectra for ΛCDM and ΛMDM (Λ Mixed Dark Matter) models is shown in Figure 4.2. We computed the power spectra with the public code `CAMB` (Lewis et al., 2000a), version 0.1.2 – May 2016. The ΛMDM power spectrum shows a clear suppression at scales smaller than ~ 100 Mpc$/h$. The suppression at small scales can be quantified as (Lesgourgues and Pastor, 2006):

$$\frac{\Delta P(k)}{P(k)} = \frac{P\left(k; f_\nu\right) - P\left(k; f_\nu = 0\right)}{P\left(k; f_\nu = 0\right)} \approx -8 f_\nu \tag{4.13}$$

although this approximation can be rather crude for large values of f_ν (Kiakotou et al., 2008).

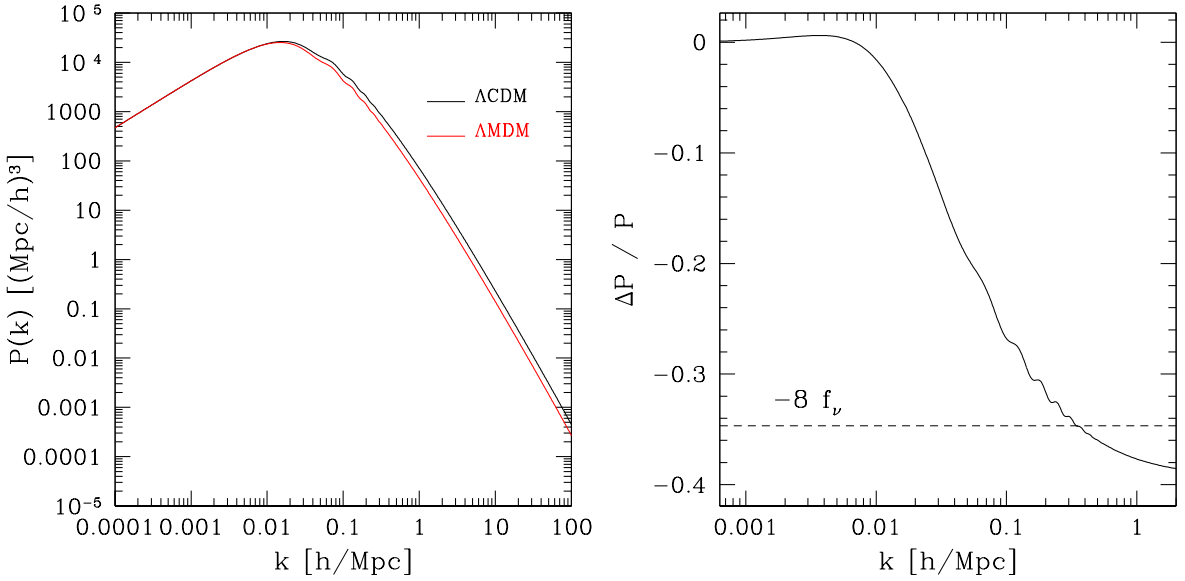

Figure 4.2: Comparison between the matter power spectra for a ΛCDM and a ΛMDM model, at redshift $z = 0$. $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, $h = 0.7$, $\Omega_\nu = 0$ and 0.013 ($f_\nu = 0.0433$).

In the $k >> k_{nr}$ regime the linear growth of perturbation is described by

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} - 4\pi G\rho_0\left(1 - f_\nu\right)\delta = 0 \tag{4.14}$$

whose growing solution can be approximated as

$$\delta_+(a) \simeq \left[ag(a)\right]^{1-\frac{3}{5}f_\nu} \tag{4.15}$$

which is a suppression of the ΛCDM solution (Equation 1.31), reducing the power-law index by $-\frac{3}{5}f_\nu$. In the limit $k << k_{nr}$ the growth factor is unsuppressed, but in the intermediate regime one needs to rely on semi-empirical approximations (e.g. Kiakotou et al., 2008) or compute the solution numerically. The main difficulty here is that in a ΛMDM model the description of the growth of structures can't be separated in a scale and in a time component, because the time evolution depends on k as well.

4.4 Numerical computations

In order to apply our model to a ΛMDM universe, we first need to compute the fundamental ingredients of the rescaled variable

$$\omega_\mathrm{f} = \frac{\delta_c(z_\mathrm{f}) - \delta_c(z_0)}{\sqrt{\sigma^2\left(fM_0\right) - \sigma^2(M_0)}} \tag{3.2}$$

Namely, the mass variance $\sigma^2(M)$, and thus the power spectrum $P\left(k; f_\nu\right)$:

$$\sigma_R^2 = \left\langle \delta^2(\vec{x}; R)\right\rangle = \frac{1}{2\pi^2}\int_0^\infty dk k^2 P(k)\,\widehat{W}^2(kR) \approx \frac{1}{2\pi^2}\int_0^{k_f} dk k^2 P(k) \equiv S\left(k_f\right) \tag{1.91}$$

and the growth factor $\delta_+(a,k)$, which is used to compute the time-dependent critical overdensity:

$$\delta_c(t) = \delta_c\frac{\delta_+(t_0)}{\delta_+(t)} \tag{1.88}$$

We used `CAMB` to compute both the power spectrum up to $k_\mathrm{max} = 100\ \mathrm{Mpc}/h$ and the growth factor for 5 sets of cosmological parameters with different values of f_ν, detailed in Table 4.1. If the option "get_transfer" is enabled, CAMB produces the matter $\delta(k)/k^2$ in the synchronous gauge at the desired redshift, as well as the individual contributions from CDM, baryons, photon and massive neutrinos. Running CAMB for multiple redshifts in the range 0 – 20 (with step 0.01) we obtained a table of $\delta(k,z)$ which we normalized dividing by $\delta(k, z=0)$, obtaining the growth factor in the redshifts and scales of interest. The time-dependent critical overdensity is then

$$\delta_c(k,z) = \delta_c\frac{\delta(k, z=0)}{\delta(k,z)} \tag{4.16}$$

where $\delta_c = 1.675$ from Equation 1.82 for the chosen cosmological parameters. As a matter of fact, the value of δ_c also depends on the neutrino masses because the spherical collapse is influenced by

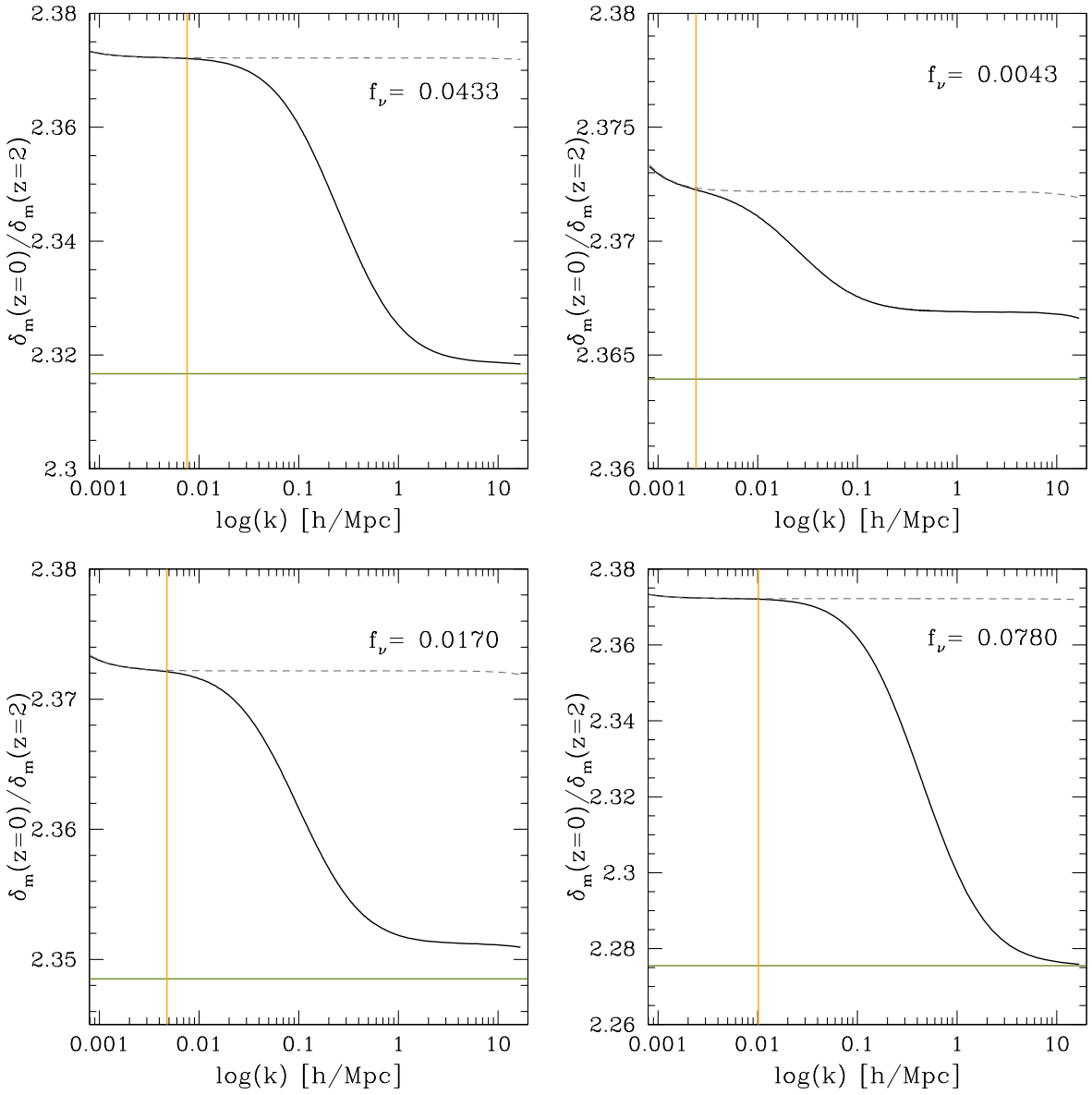

Figure 4.3: Ratios of the matter overdensity at $z = 0$ and at $z = 2$ in ΛMDM models, in the scale range $10^{-3}\ h/\mathrm{Mpc} < k < 10\ h/\mathrm{Mpc}$. The gray dashed curve represents the model with $f_\nu = 0$. The dark green horizontal line represents the small scale approximation Equation 4.15. The orange vertical line represents the scale of the non-relativistic transition k_{nr}.

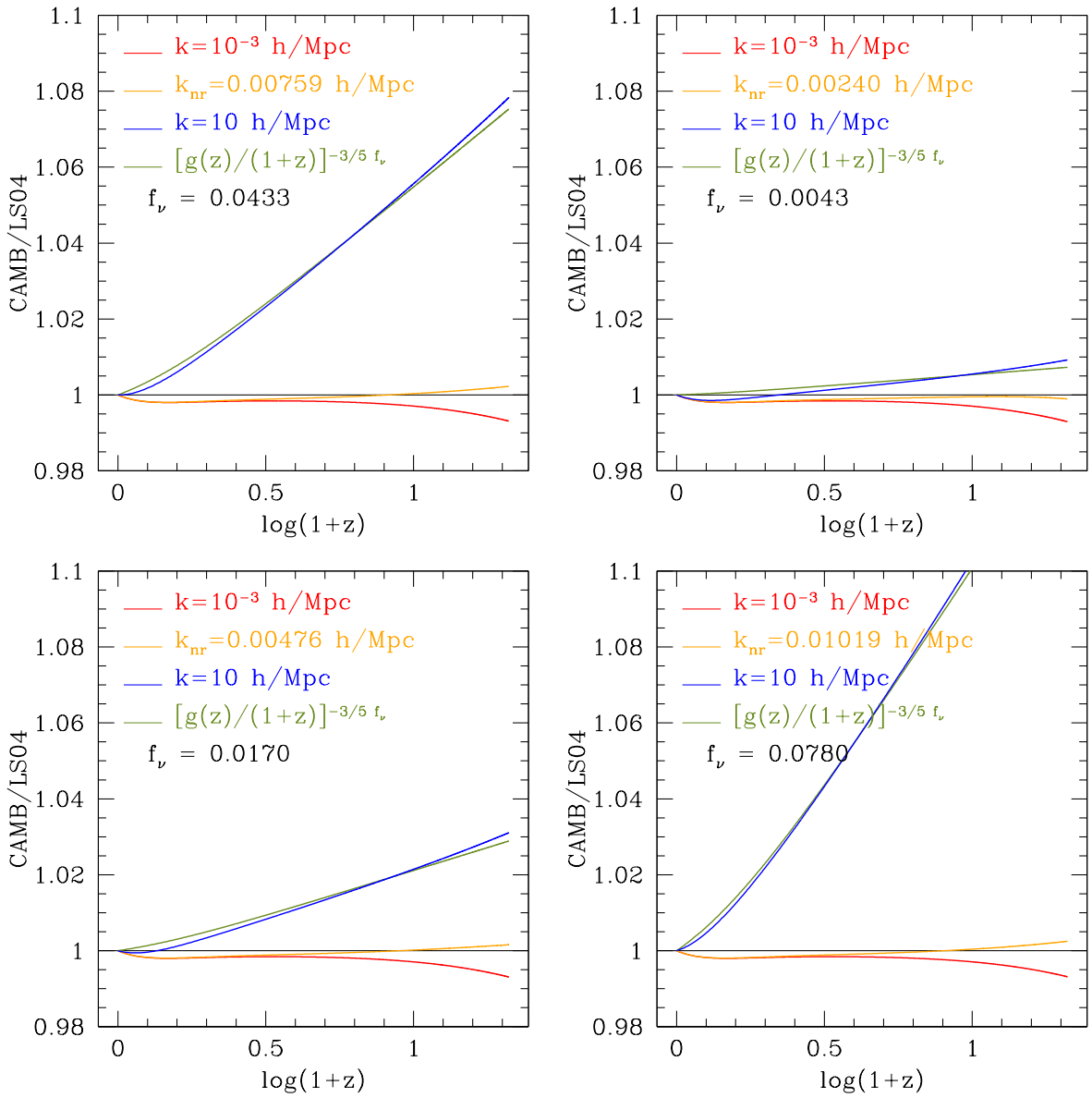

Figure 4.4: Growth factors in ΛMDM models. Red, orange and blue curves represent the ratio of the growth factors at the stated scales and the ΛCDM approximation. The dark green curve represents the small scales approximation Equation 4.15.

name	$\sum m_\nu$ [eV]	Ω_ν	f_ν	M^* [$M_\odot$ h^{-1}]
Manuela	0.5933	0.0130	0.0433	3.054×10^{12}
Micaela	0.0593	0.0013	0.0043	3.868×10^{12}
Martina	0.2328	0.0051	0.0170	3.675×10^{12}
Marcela	1.0680	0.0234	0.0780	2.171×10^{12}
Manola	0	0	0	3.902×10^{12}

Table 4.1: Sets of parameters used in CAMB. For every set $\Omega_\Lambda = 0.7$, $\Omega_m = 0.3$, $\Omega_b = 0.048$, $h = 0.7$ and $\sigma_8 = 0.8$.

the presence of massive neutrinos. However, the difference is only $\lesssim$ 0.5-1% (LoVerde, 2014), so we decided to use the ΛCDM value (Equation 1.82) for simplicity.

Figure 4.3 shows the ratio of the matter overdensities $\delta(k,0)/\delta(k,z)$ for $z = 2$. More plots for various z are available in Appendix D. The scale of the non-relativistic transition k_{nr} is shown with a vertical orange line, and marks the beginning of the intermediate scales. The growth factor at k_{nr}, as well as at large ($k = 10^{-3}$ $h/$Mpc) and small ($k = 10$ $h/$Mpc) scales, divided by the ΛCDM approximation Equation 1.31 (Lahav and Suto, 2004) is shown in Figure 4.4. The agreement of the numerical results with both the small and large scale approximations is remarkable, but the numerical approach also works in the intermediate regime. This same regime is also relevant for the mass variance, corresponding to the peak of the power spectrum (Figure 4.2) and to the mass range

$$M_k = \frac{4\pi}{3}\rho_b R_k^3 = \frac{4\pi}{3}\rho_b \left(\frac{\pi}{k}\right)^3 \approx 10^{13} - 10^{19}\,M_\odot\,h^{-1} \tag{4.17}$$

Including the scale (mass) dependence of the critical overdensity in the definition of ω:

$$\omega_\nu = \frac{\delta_c(z, k(M_0)) - \delta_c(z_0, k(M_0))}{\sqrt{\sigma^2(fM_0) - \sigma^2(M_0)}} \tag{4.18}$$

we can apply our model as usual. Note that there is no guarantee that the universality of the model holds for ΛMDM models, since we only verified it with ΛCDM simulations. However, given the smallness of the f_ν parameter, deviations from universality (if any) are expected to be small, with negligible effects on the results.

Figure 4.5 shows the mean formation redshift – mass relations for ΛMDM cosmologies derived from our model, multiplied by by an arbitrary factor of $M_0^{-0.05}$ to enhance readability. As expected, the effect of the neutrinos is more pronounced at smaller halo masses, with predicted differences in the formation redshift up to 9.4% for $M_0 = 10^{-2}M^*$ in the case of neutrino masses of order unity, with respect to the $\sum m_\nu = 0$ model. For smaller neutrino masses, the difference decreases to 3.2% and 0.3% for $\sum m_\nu \simeq 0.6$ and 0.23 eV, respectively. For neutrino masses smaller than ~ 0.1 eV the difference in the formation redshift is negligible ($\lesssim$ 0.1%) even at small halo masses. At $M_0 = M^*$ the differences decrease to 7.8%, 2.6% and 0.2%, and become negligible around $10^2 M^*$.

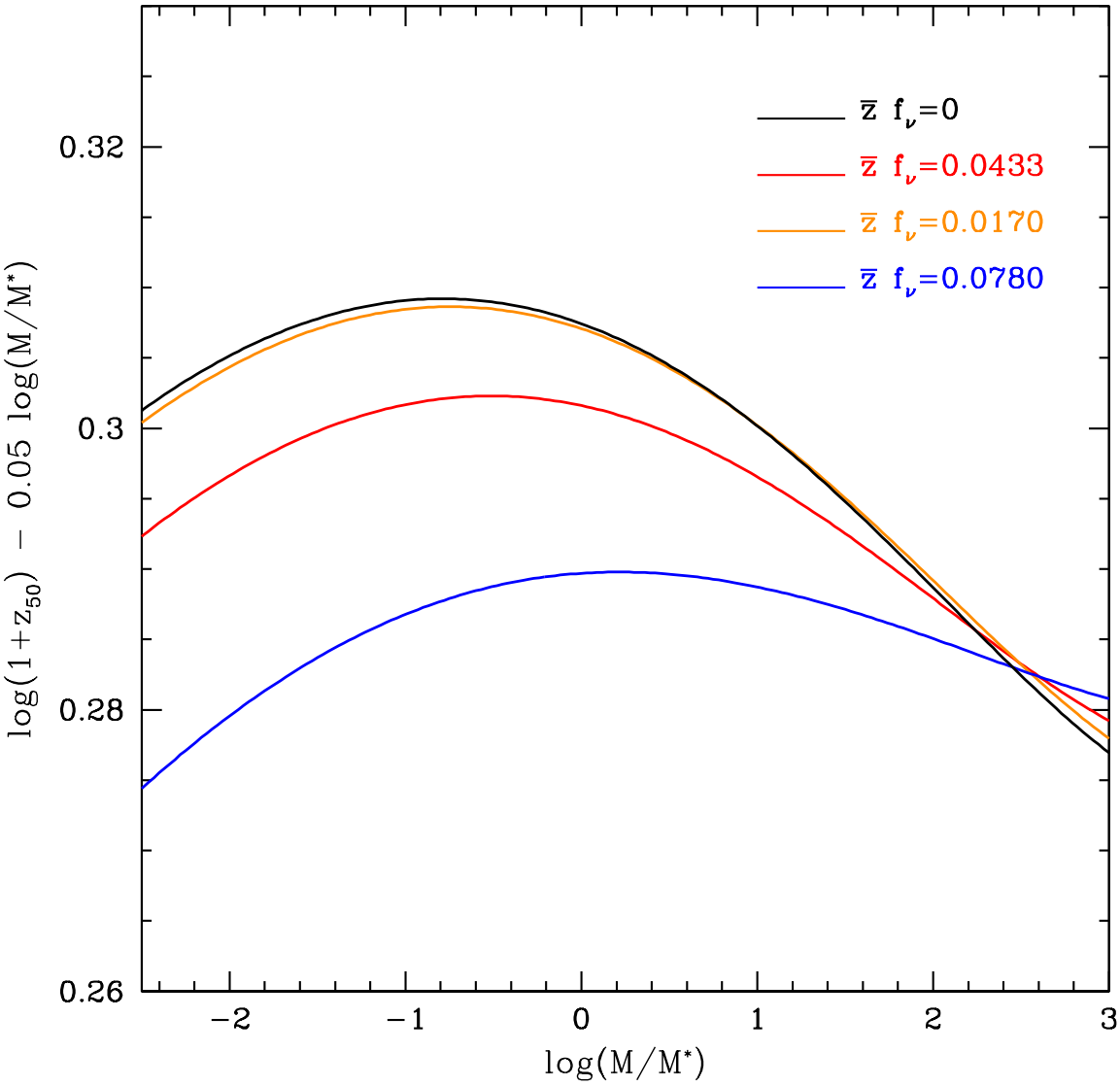

Figure 4.5: Mean $z_{0.50} - M_0$ relation for ΛMDM models. $f_\nu = 0.0043$ is too close to $f_\nu = 0$ to be distinguishable and it is not shown.

The natural next step in this analysis would be a direct comparison with N-body simulations with ΛMDM cosmologies. However, despite a few cosmological simulations exist (although not public), these are either low resolution (Viel et al., 2010; Brandbyge et al., 2010) or are not suitable for a comparison with our model. This is the case, for example, for the DEMNUni simulations (Castorina et al., 2015), a set of 4 large scale simulations with $\sum m_\nu$ = 0.00, 0.17, 0.30, 0.53 characterized by a comoving volume of $2000^3\ h^{-3}\,\mathrm{Mpc}^3$ filled with 2048^3 dark matter and, where present, 2048^3 neutrino particles. The corresponding mass resolution of 8×10^{10} $\mathrm{M}_\odot$ h^{-1} means that resolved halos, i.e. those with more than 1000 particles, would be in the mass range $\gtrsim 16M^*$, where the effects of the presence of massive neutrinos on the growth of structure are starting to become negligible. Given the growing interest of the scientific community in massive neutrinos, however, we expect that more large scale simulations with a massive neutrino component will arise in the near future. Until then, we hope to provide to the interested reader a useful tool to forecast the effects of neutrinos on the growth of structures.

Conclusion 5

The mass accretion history of dark matter halos is considered one of the main processes that influence the shape of their present day internal structure (Navarro et al., 1996; Tormen, 1998; Zhao et al., 2009; Giocoli et al., 2010), as well as one of the main driver of the formation and evolution of galaxies. It is then clear that accurately modeling the mass accretion history of dark matter halos is of primary importance in our efforts of understanding the properties of large scale structures. While many models for the mass accretion history already exist in the literature (Wechsler et al., 2002; van den Bosch, 2002; Giocoli et al., 2007; Zhao et al., 2009; Giocoli et al., 2012; van den Bosch et al., 2014), they either describe only the mean or median MAH, or they are very poor fit of the tails of the halo-to-halo distribution of the MAHs. Our main goal is to compensate for this shortcoming.

In this work, we presented the cosmological framework of the formation and linear and non-linear evolution of large scale structures, and the primary tools used to study the non-linear evolution of the density perturbations, namely cosmological simulations. We detailed our own set of simulations, partly developed in the context of this work, and the post-processing algorithm used to construct halo catalogs and merger trees. We exploited the large statistic and dynamical range of this sample and, starting from the model for the mass accretion history proposed by Giocoli et al. (2012), we presented a revised model that accurately describe both the median as well as the tails of the distribution. For the median, we implemented the estimator proposed by Banneheka and Ekanayake (2010) and tested the accuracy over full range of our parameters, finding a maximum error of about 0.35%. For the tails, we illustrated the deviations introduced by the presence non-relaxed halos and we accurately defined relaxation criteria to eliminate such deviations, as well as to obtain consistent properties across all simulations. A more detailed analysis of the distribution of non-relaxed halos will be presented in a future work (Gambarotto et al., in preparation).

For the calibration of the model we applied a robust fitting procedure with a careful consideration of the errors, on both the data and the fitted parameters. The results suggest that our choice of the distribution and its parametrization provides a good approximation, but not the exact underlying

distribution of the data. This finding is supported by the outcome of bootstrap and Monte Carlo simulations. We used the calibration to check the dependence of the model on the adopted cosmology, applying the same analysis on a secondary set of simulations with different cosmological parameters. We found good agreement between the fitted parameters of the main and secondary simulations, confirming the universality of the model.

Studying the scatter of the distribution, we performed a preliminary analysis of the percentile distributions of the MAH tracks. We found that, in the assumption that a halo remains mostly in the same region of the percentile distribution during all of its evolution, there exists a characteristic redshift at which the mass of the halo is statistically closer to its typical value in the percentile distribution. We make no claim on the validity of such assumption. From the percentile distributions we also defined and extracted several properties of the individual halos, in an attempt to characterize the behavior of their mass evolution. Despite not founding significant correlations between such properties, they may serve as a characterization that can be useful when checking the validity of algorithms to construct synthetic merger trees.

From the model we extracted conditional distributions for the formation redshift at fixed mass fraction f, and we found an overall good agreement with the data in the restricted range $0.04 \lesssim f \lesssim 0.9$, with only a slight overestimation of the formation redshift. Outside of this range, the model significantly departs from the data and it is not viable. We believe that the wide applicability of our model, for every mass, identification redshift, formation redshift, fraction and cosmology, together with the ability to describe the full distribution of the MAHs, more than compensate these minor drawbacks.

We discussed possible applications of our model, presenting as an example a proof-of-concept procedure to select halos that are outliers in their distribution, of which one can study properties like the mass-concentration relation and how they differ from those for halos in the central region of their distribution. Motivated by the universality of the model, we then focused to the applicability to Λ Mixed Dark Matter cosmologies. We outlined the main cosmological effects of massive neutrinos on the evolution of matter density perturbations and shown how to include appropriate modifications in our model. We found that at small halo masses, the formation redshift can be lowered by up to 10% for neutrino masses of around 1 eV, the current weaker upper limit from cosmological observations (Planck Collaboration et al., 2016). This effect is less pronounced at larger masses, and becomes negligible at around 10^{15} $M_\odot$ h^{-1}. Given the still early stage of the present day research on massive neutrinos, the complications in the implementation of the related scale-dependent physics in simulations and the large computational resources needed, we could not compare our results against existing simulations, although we expect that such comparison will become possible in the near future. We therefore propose our findings as an useful tool to forecast the effects of neutrinos on the growth of structures.

Median MAHs at various masses and redshifts A

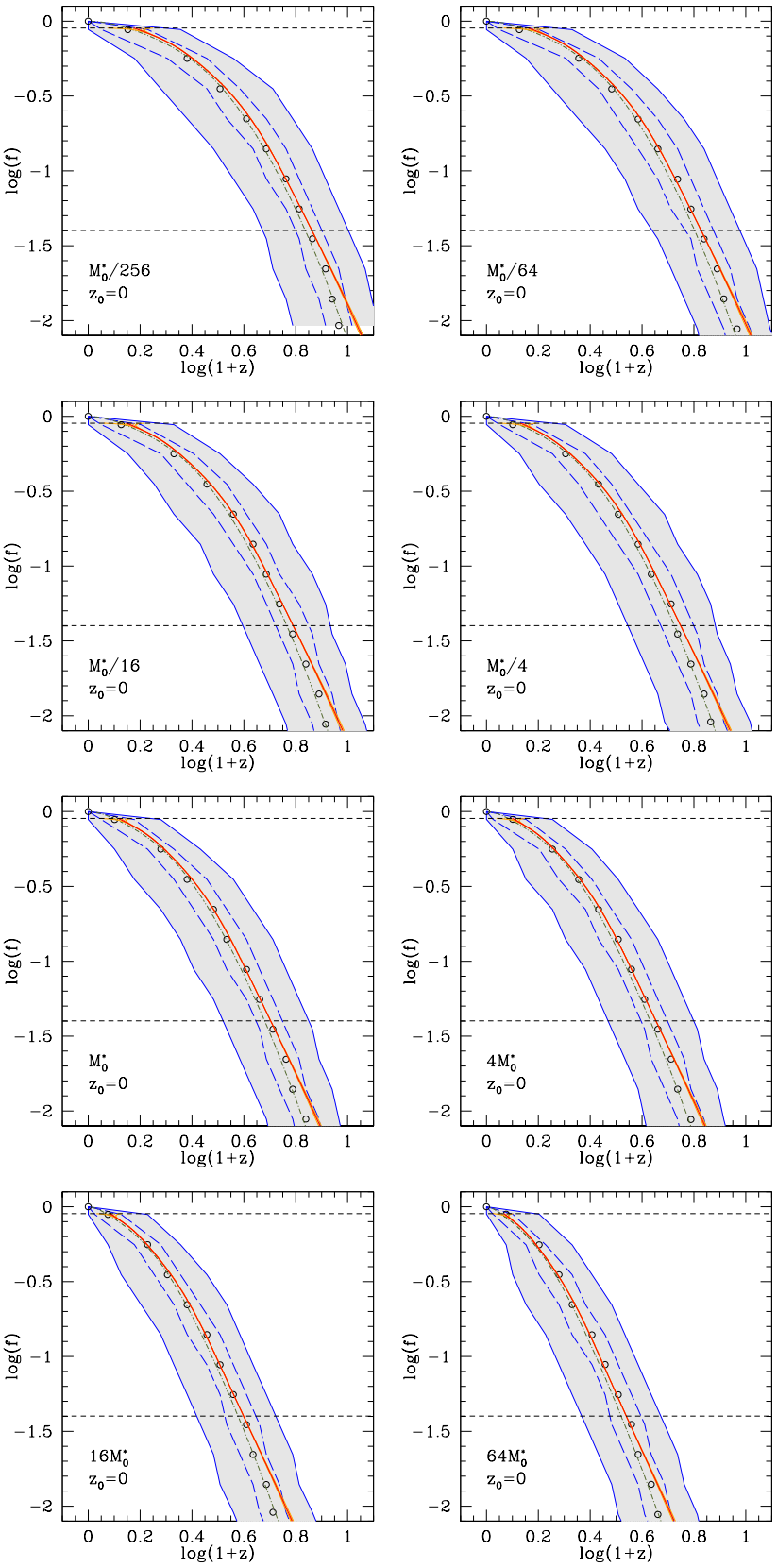

Figure A.1: Same as Figure 3.15, but with different masses.

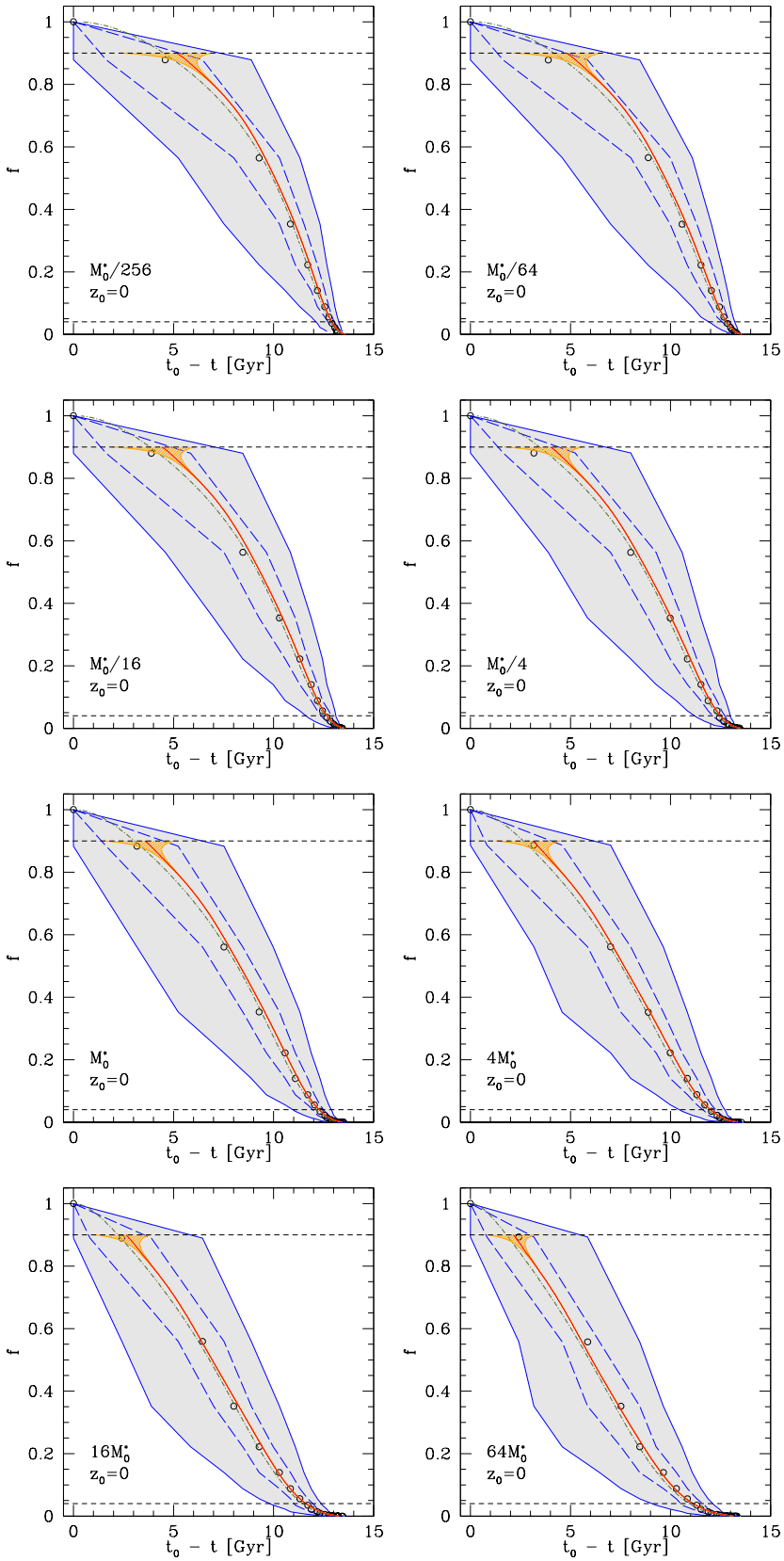

Figure A.2: Same as Figure 3.16, but with different masses.

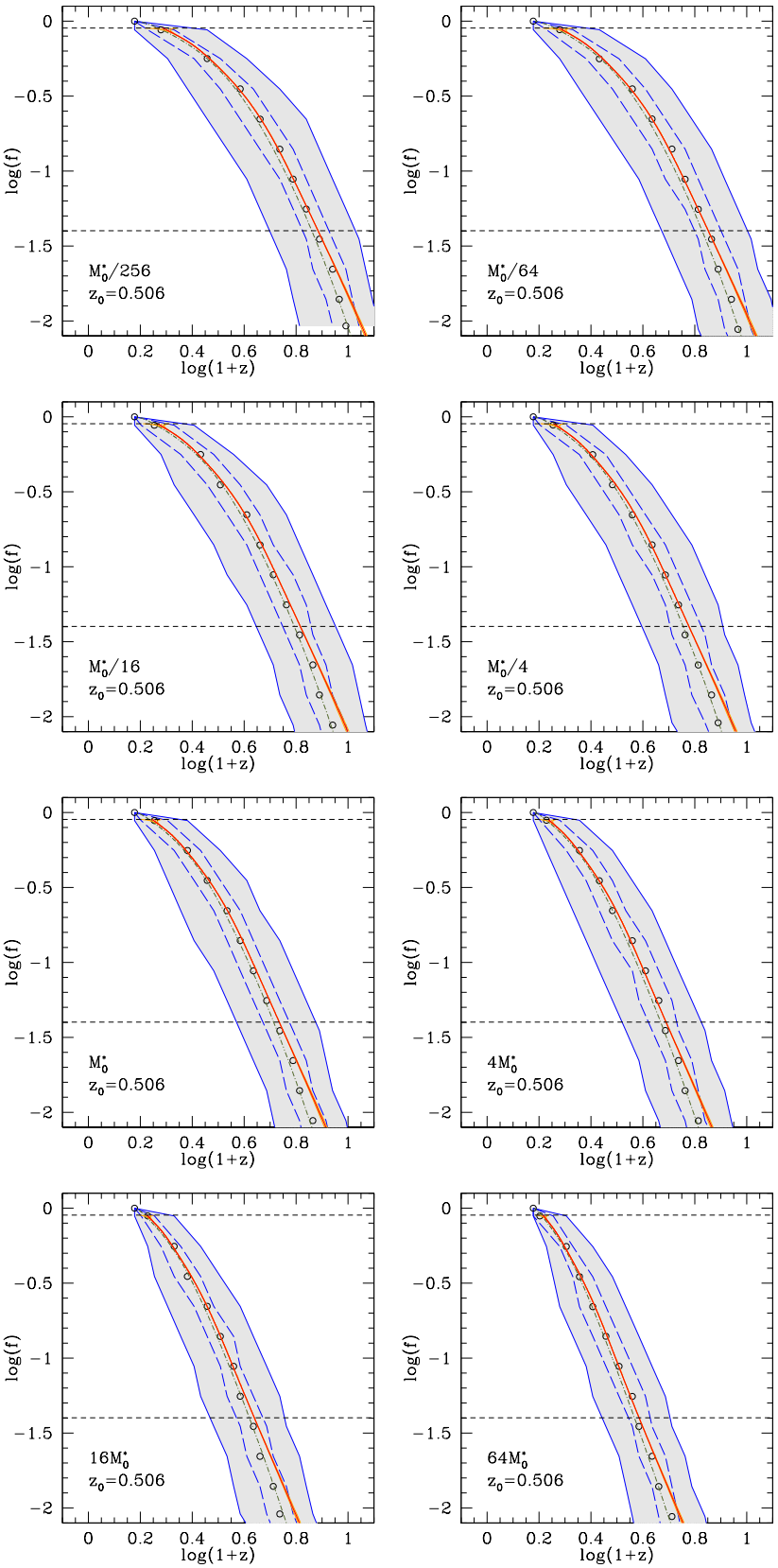

Figure A.3: Same as Figure 3.15, but with different masses and $z_0 = 0.506$.

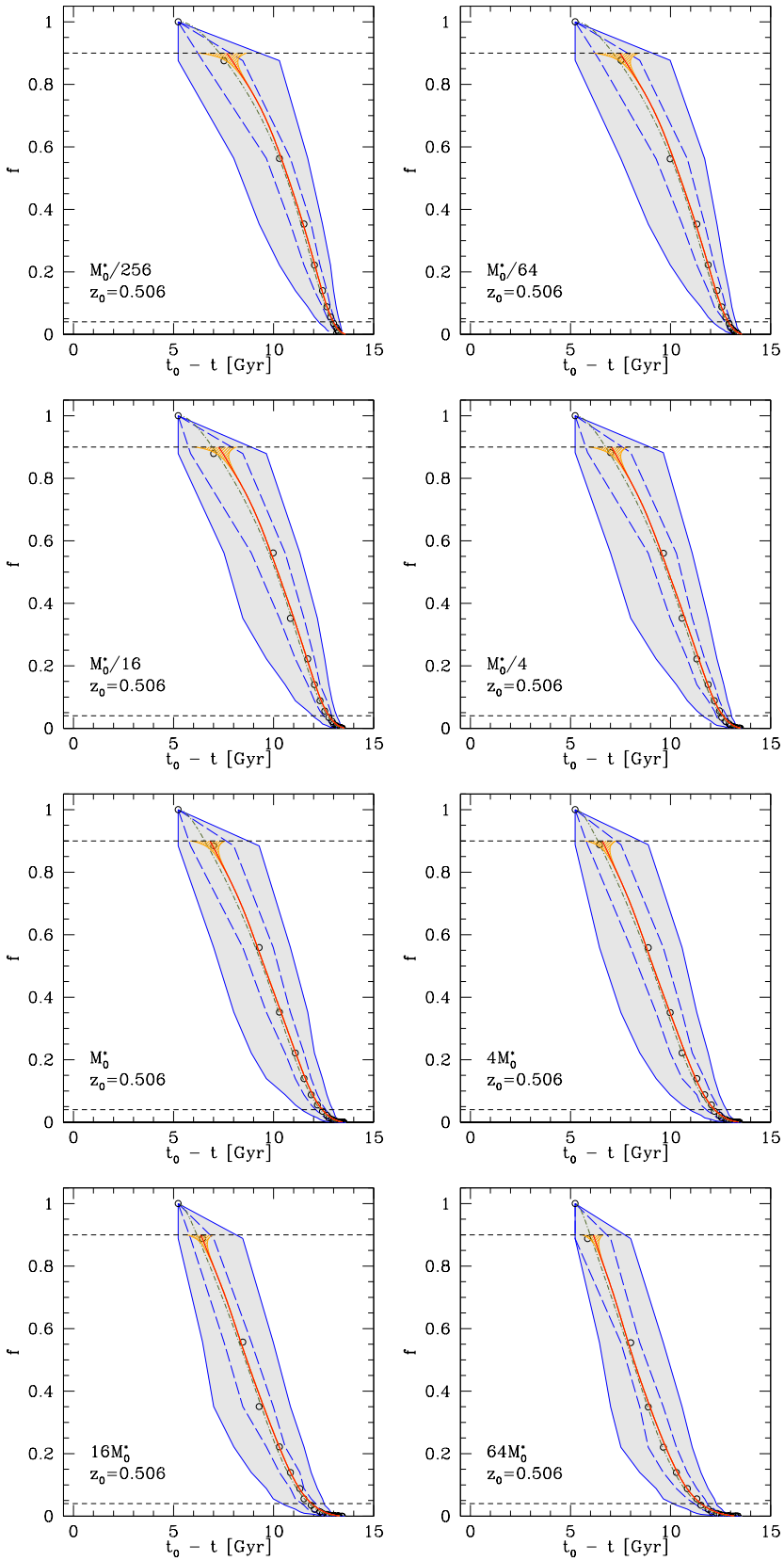

Figure A.4: Same as Figure 3.16, but with different masses and $z_0 = 0.506$.

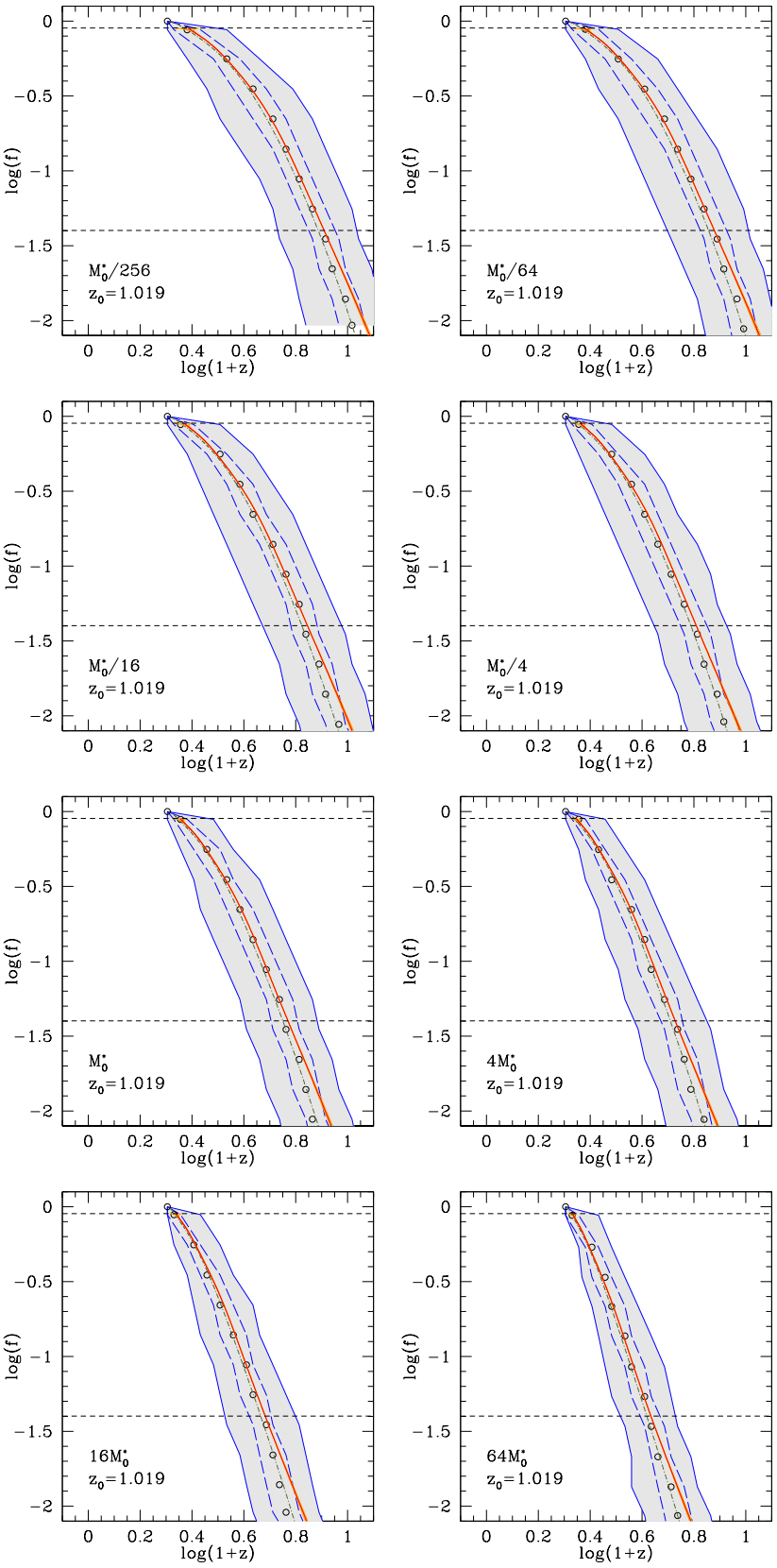

Figure A.5: Same as Figure 3.15, but with different masses and $z_0 = 1.019$.

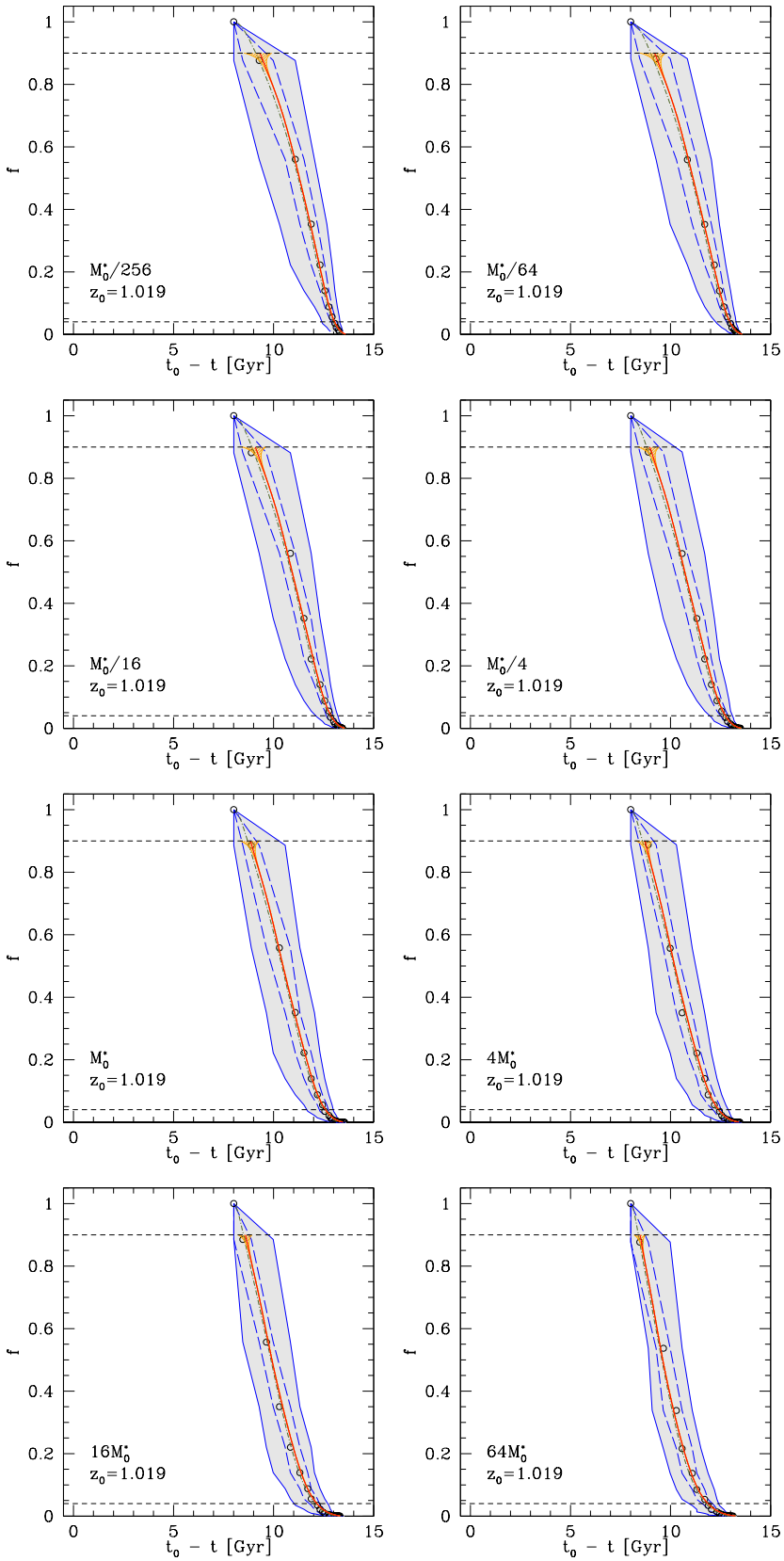

Figure A.6: Same as Figure 3.16, but with different masses and $z_0 = 1.019$.

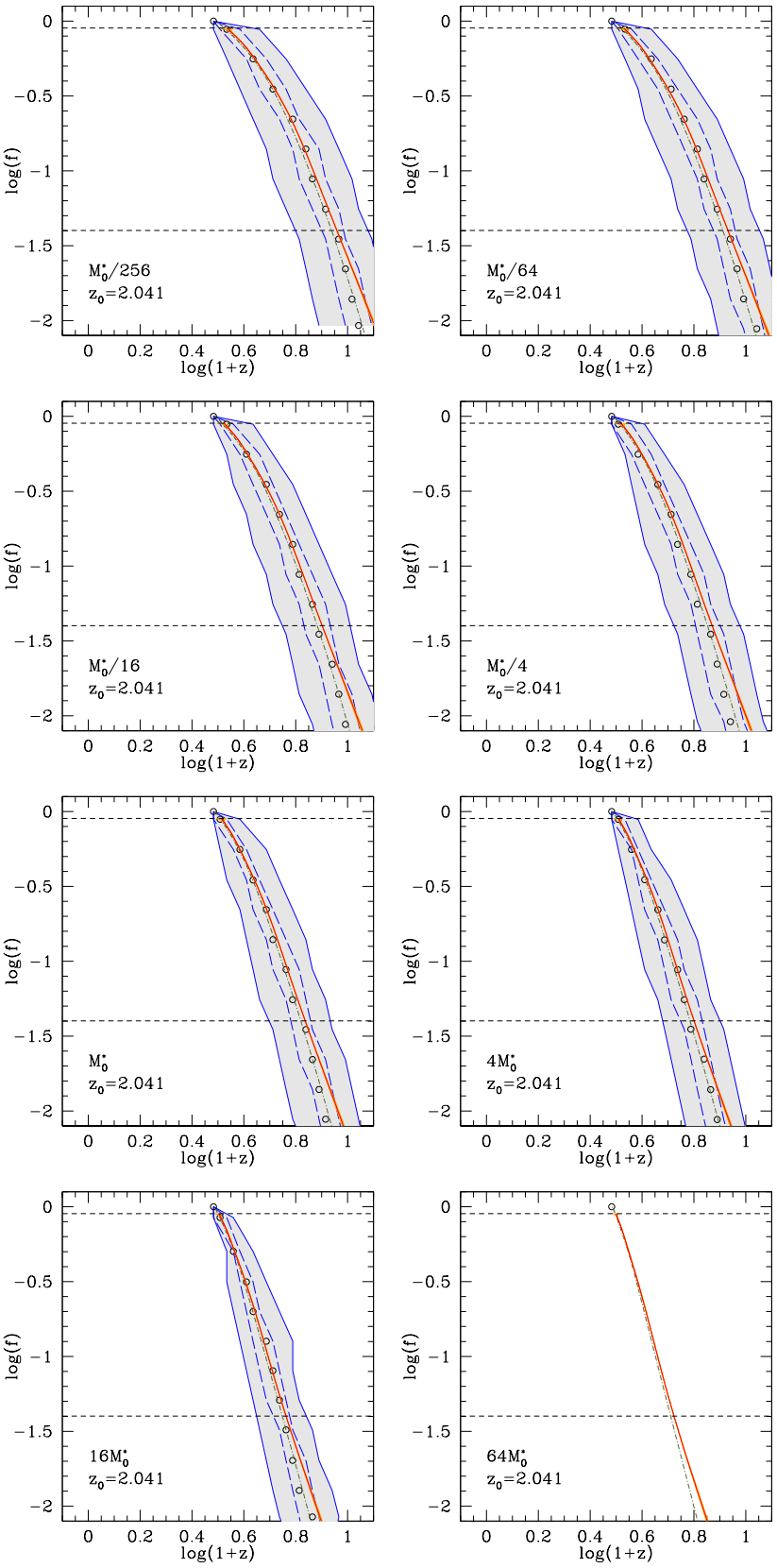

Figure A.7: Same as Figure 3.15, but with different masses and $z_0 = 2.041$.

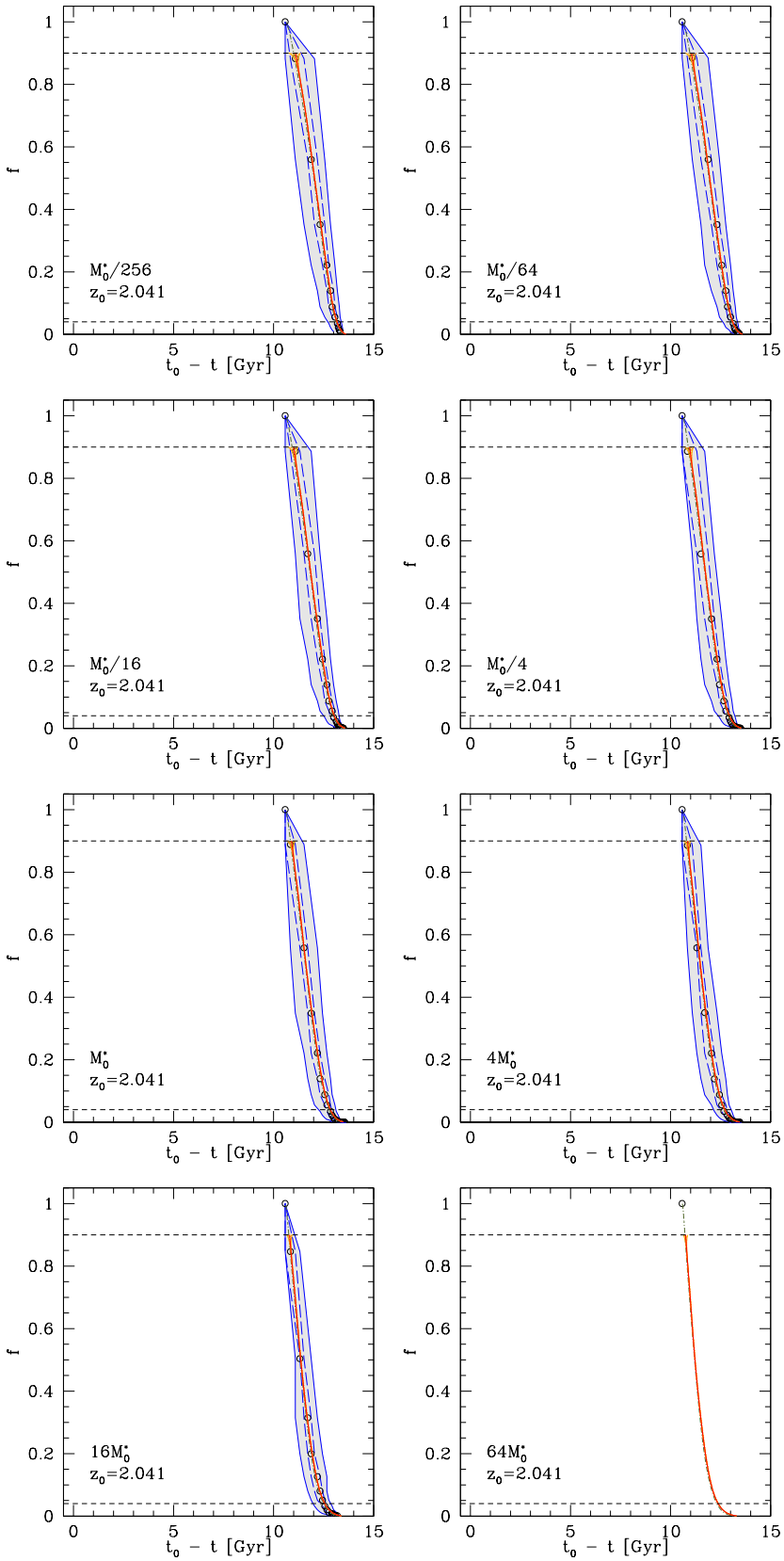

Figure A.8: Same as Figure 3.16, but with different masses and $z_0 = 2.041$.

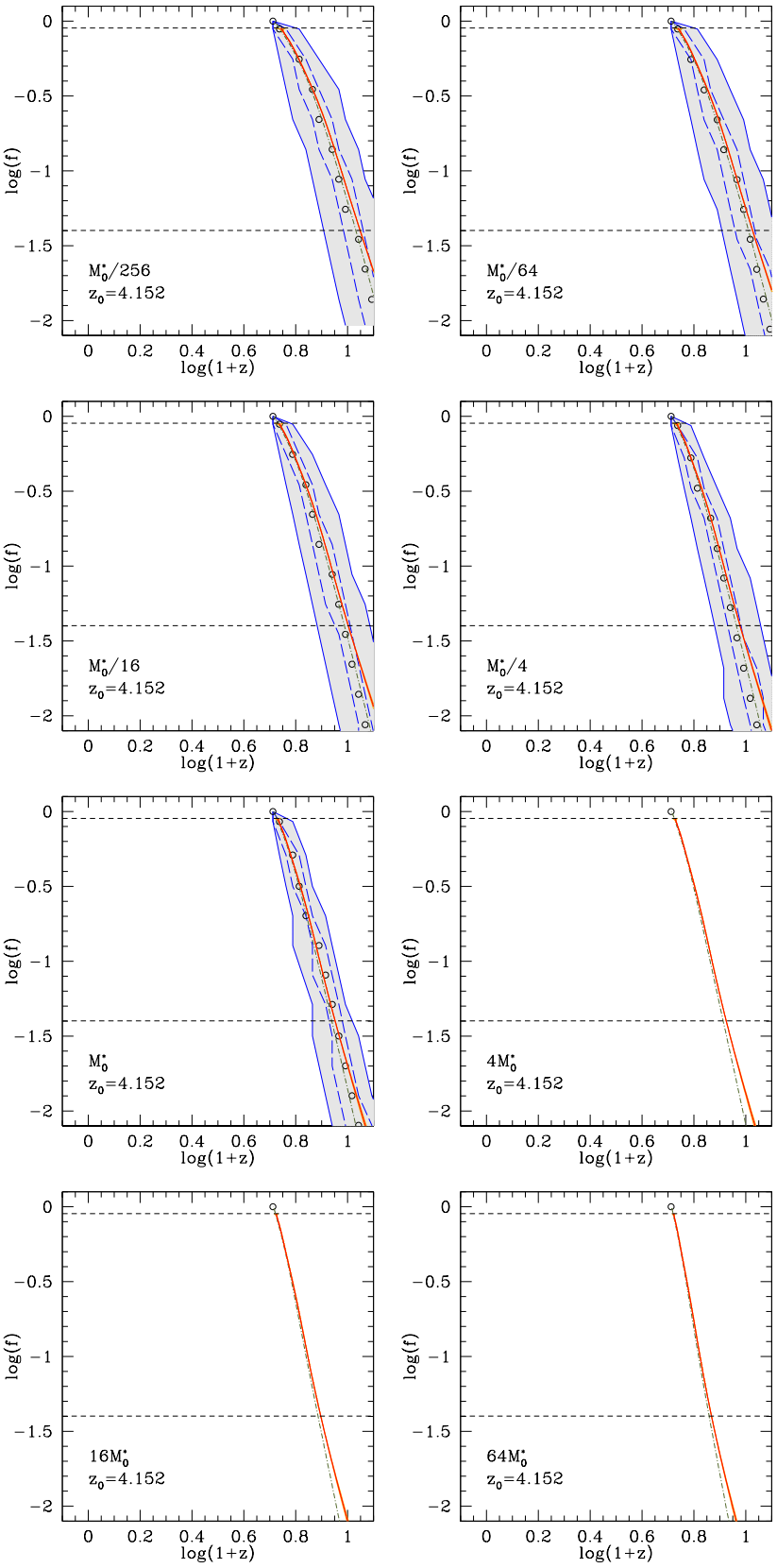

Figure A.9: Same as Figure 3.15, but with different masses and $z_0 = 4.152$.

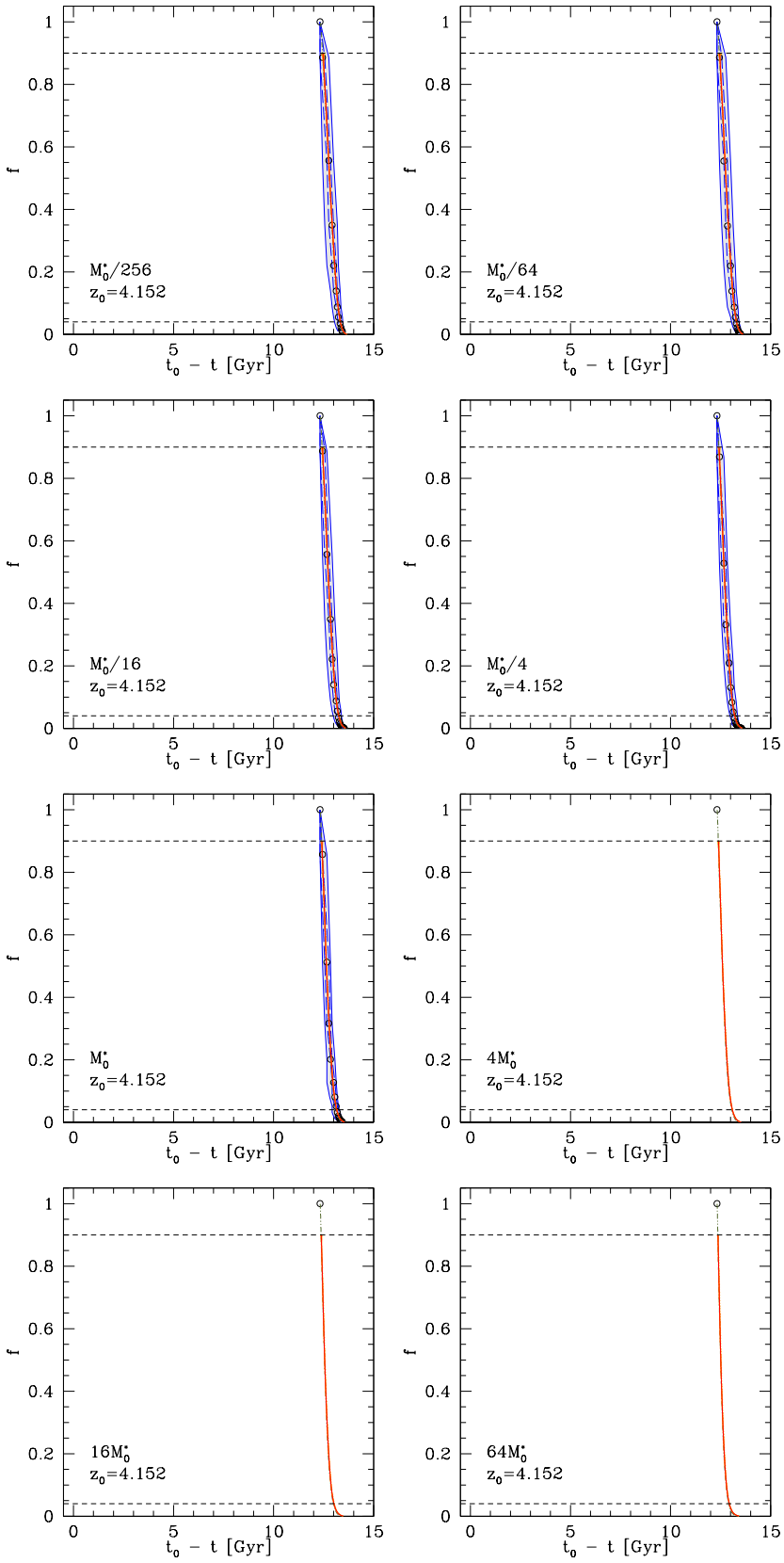

Figure A.10: Same as Figure 3.16, but with different masses and $z_0 = 4.152$.

Redshift distributions B

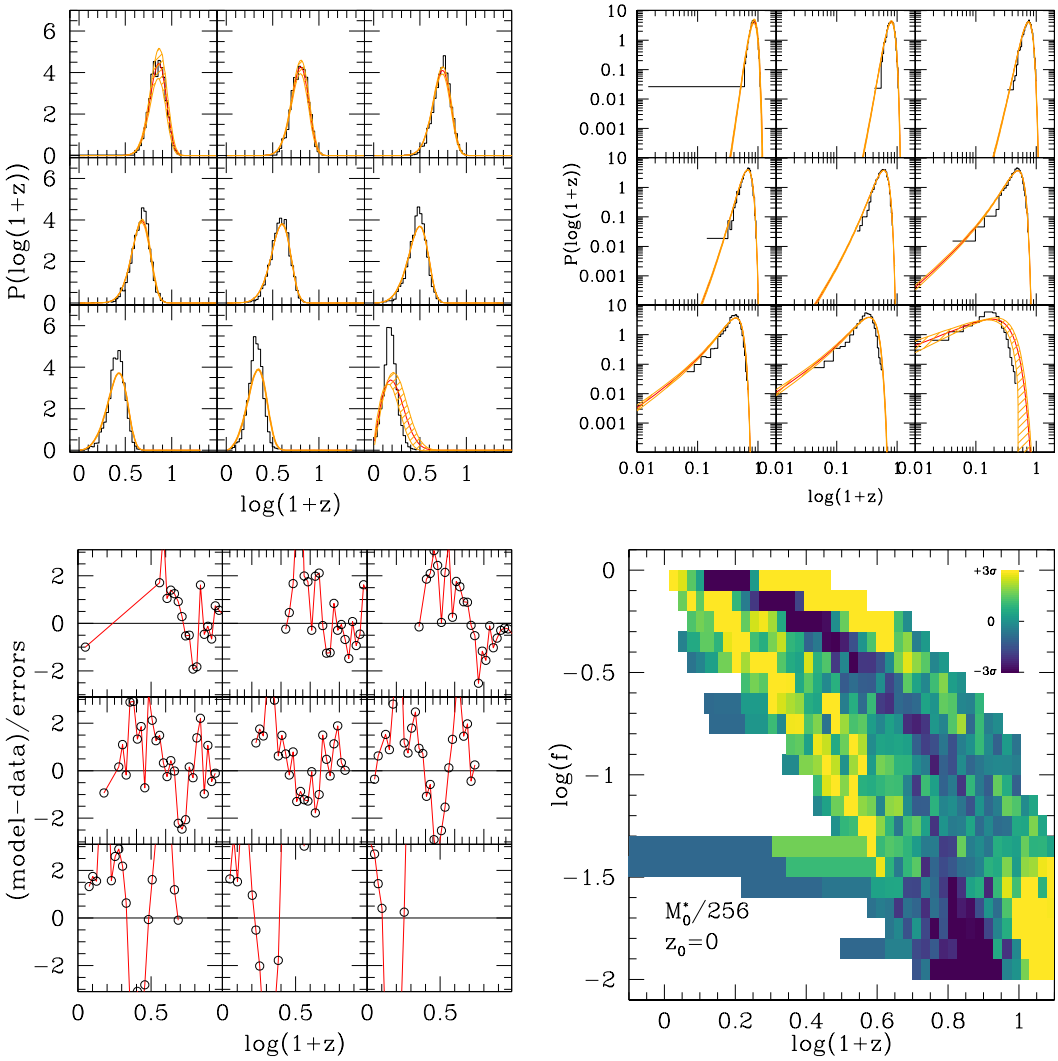

Figure B.1: Same as Figure 3.21, with $M_0 = M^*/256$.

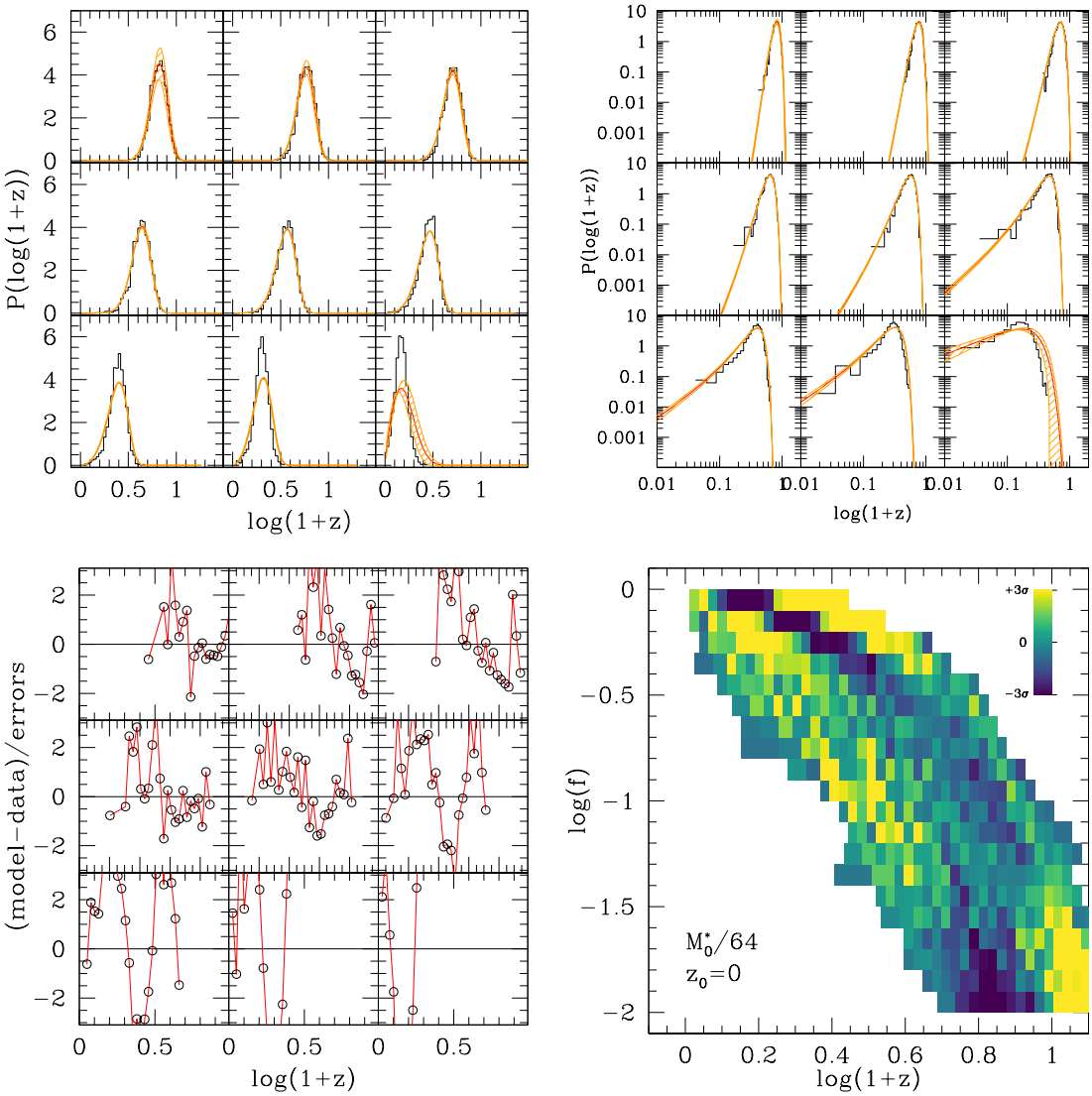

Figure B.2: Same as Figure 3.21, with $M_0 = M^*/64$.

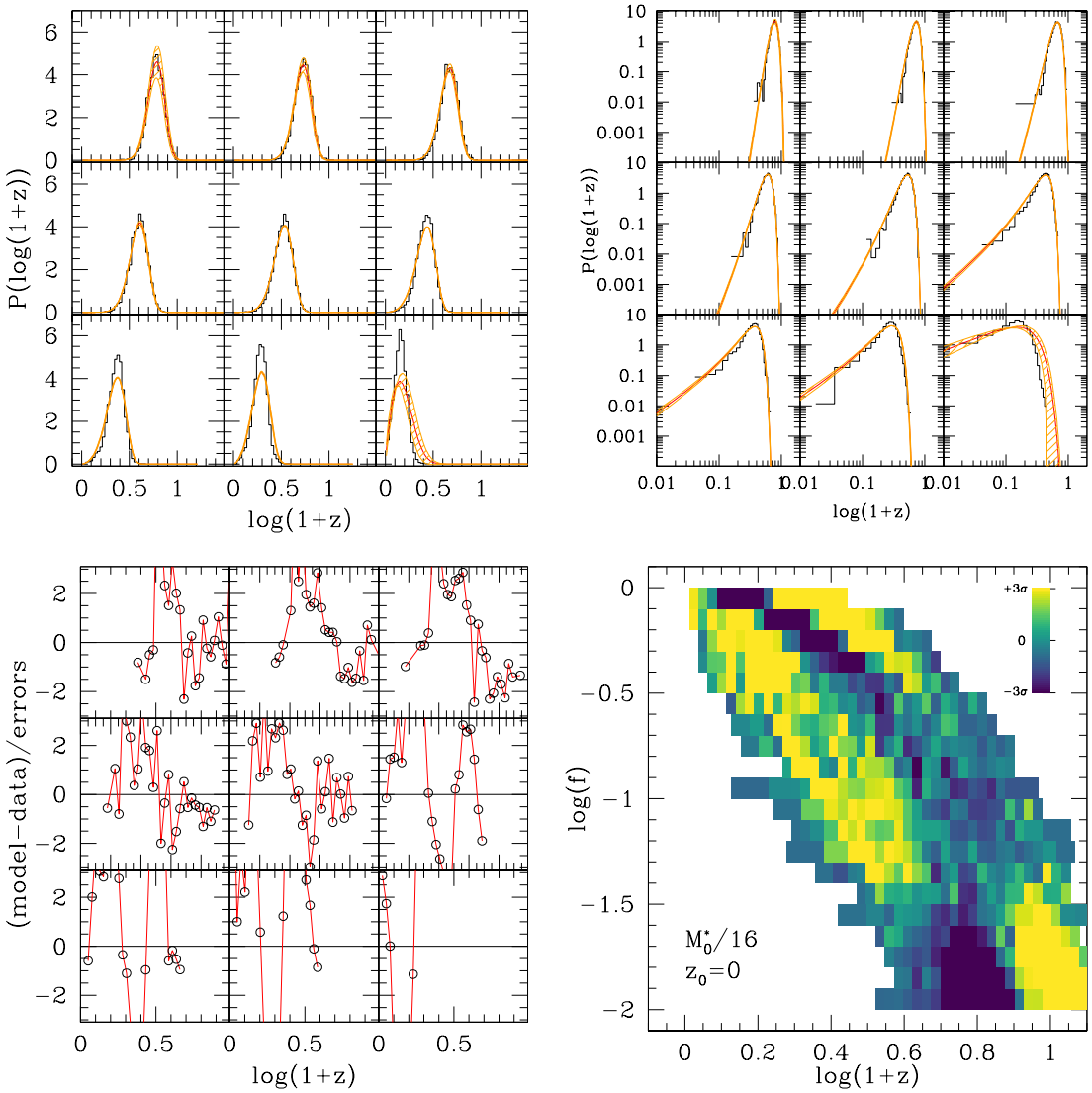

Figure B.3: Same as Figure 3.21, with $M_0 = M^*/16$.

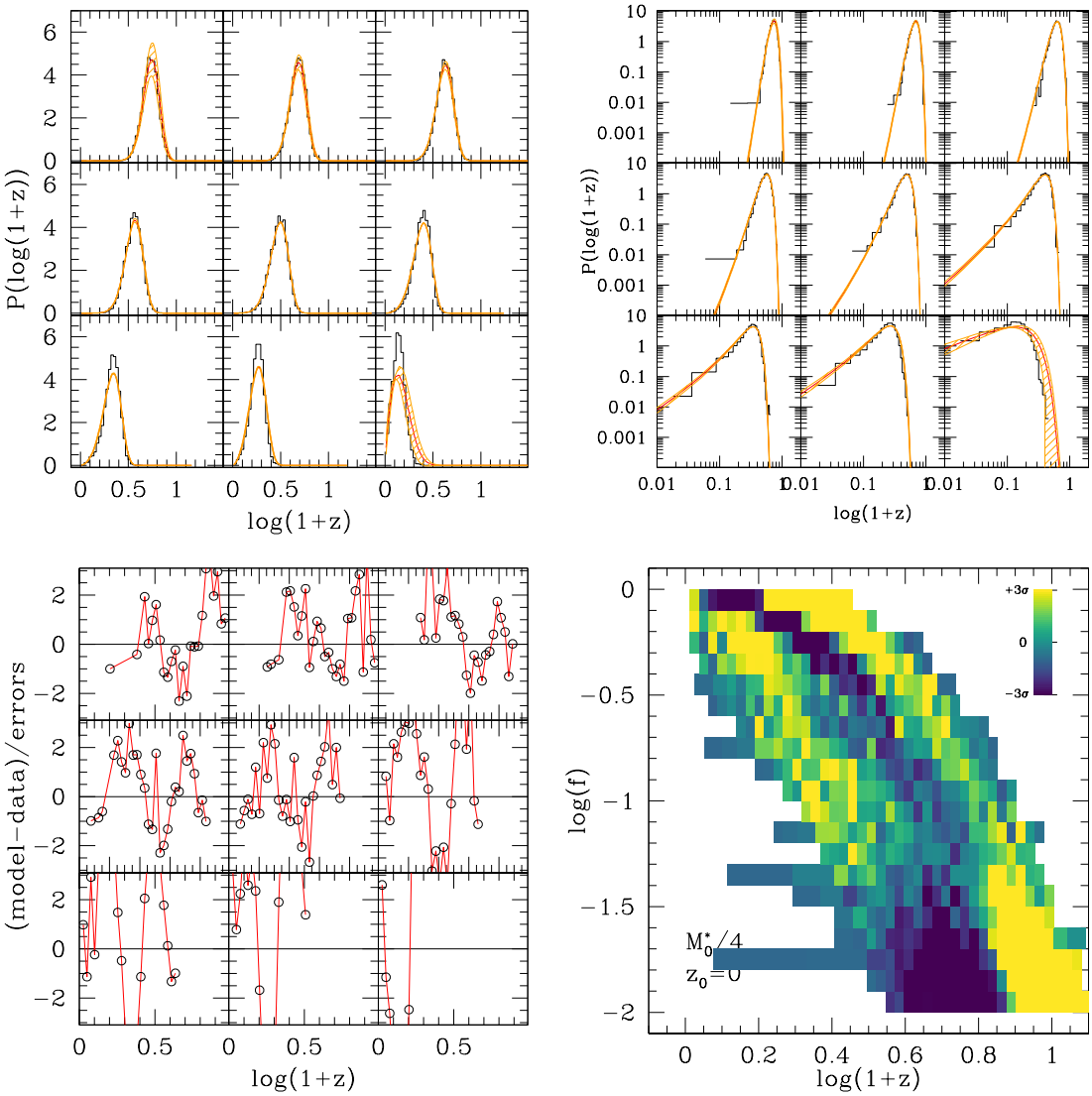

Figure B.4: Same as Figure 3.21, with $M_0 = M^*/4$.

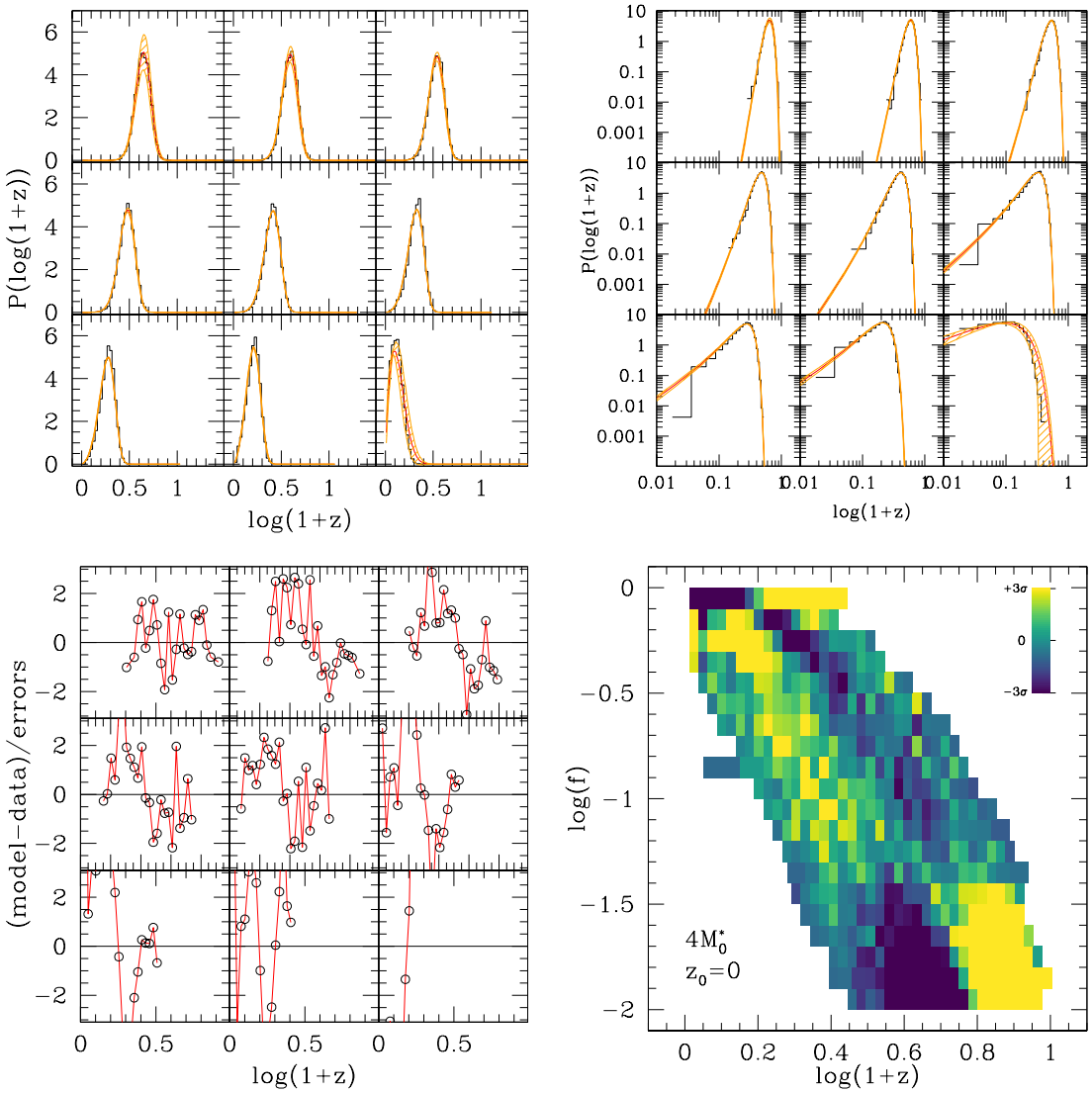

Figure B.5: Same as Figure 3.21, with $M_0 = 4M^*$.

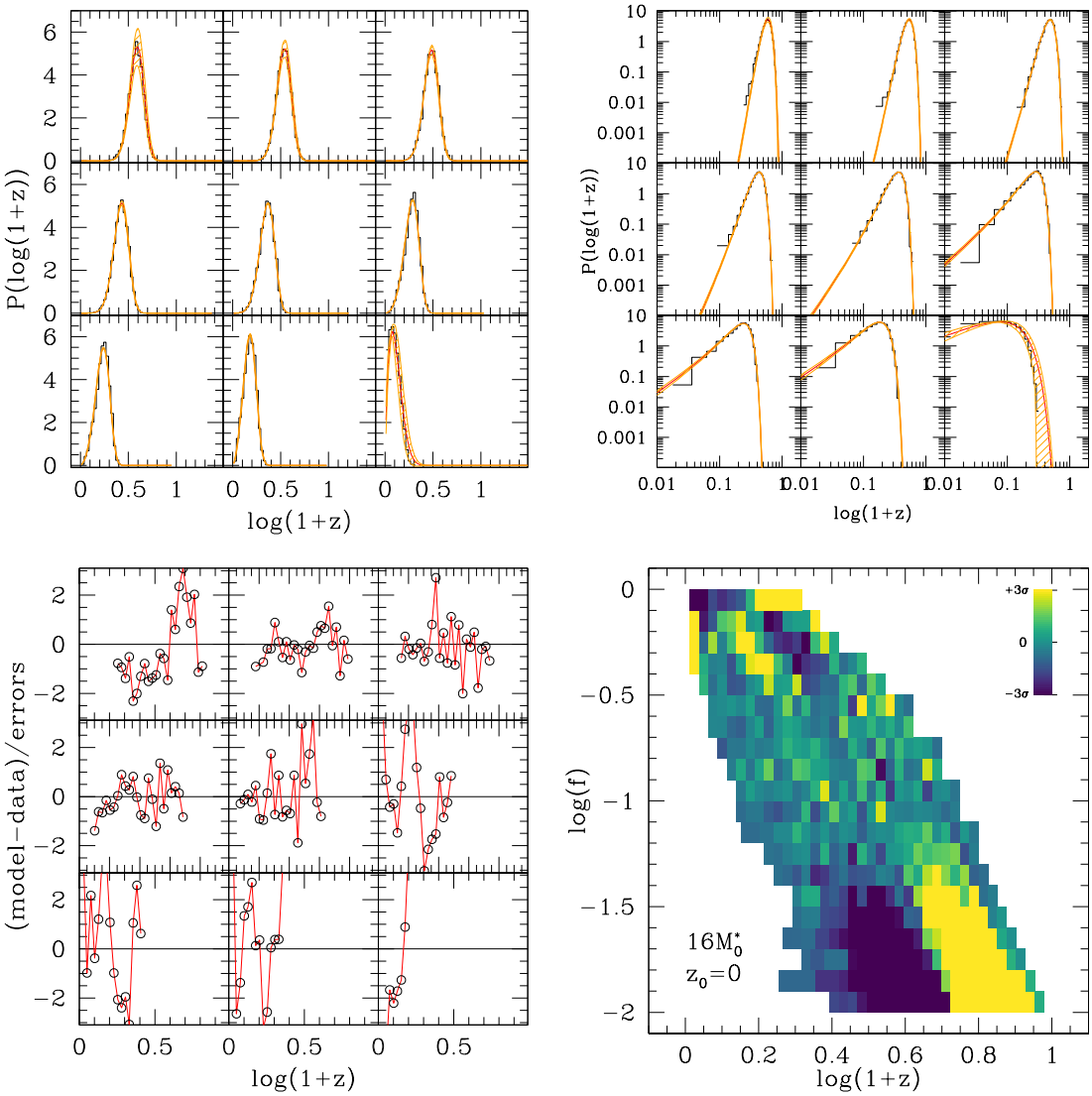

Figure B.6: Same as Figure 3.21, with $M_0 = 16M^*$.

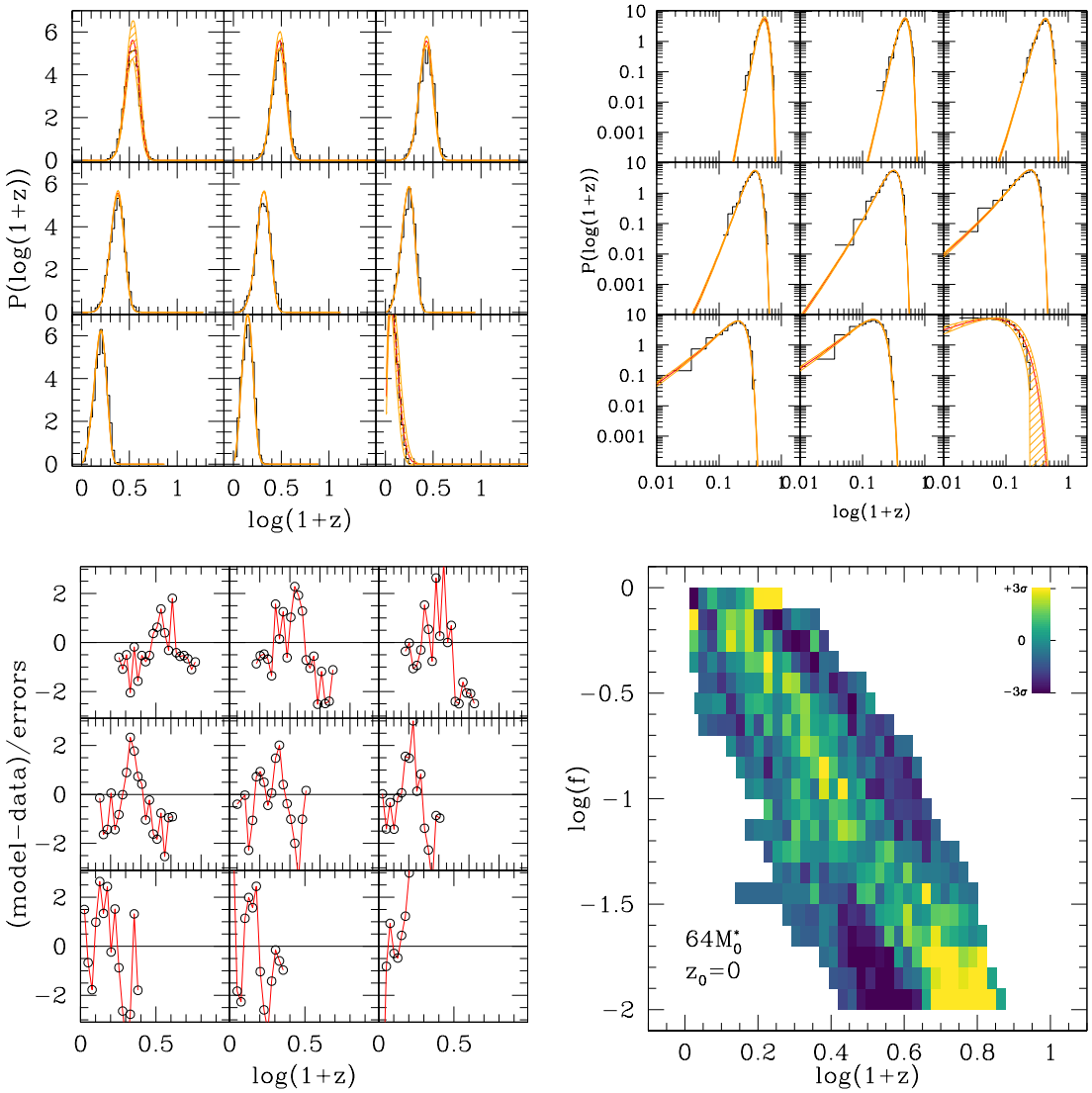

Figure B.7: Same as Figure 3.21, with $M_0 = 64M^*$.

Cumulative distributions C

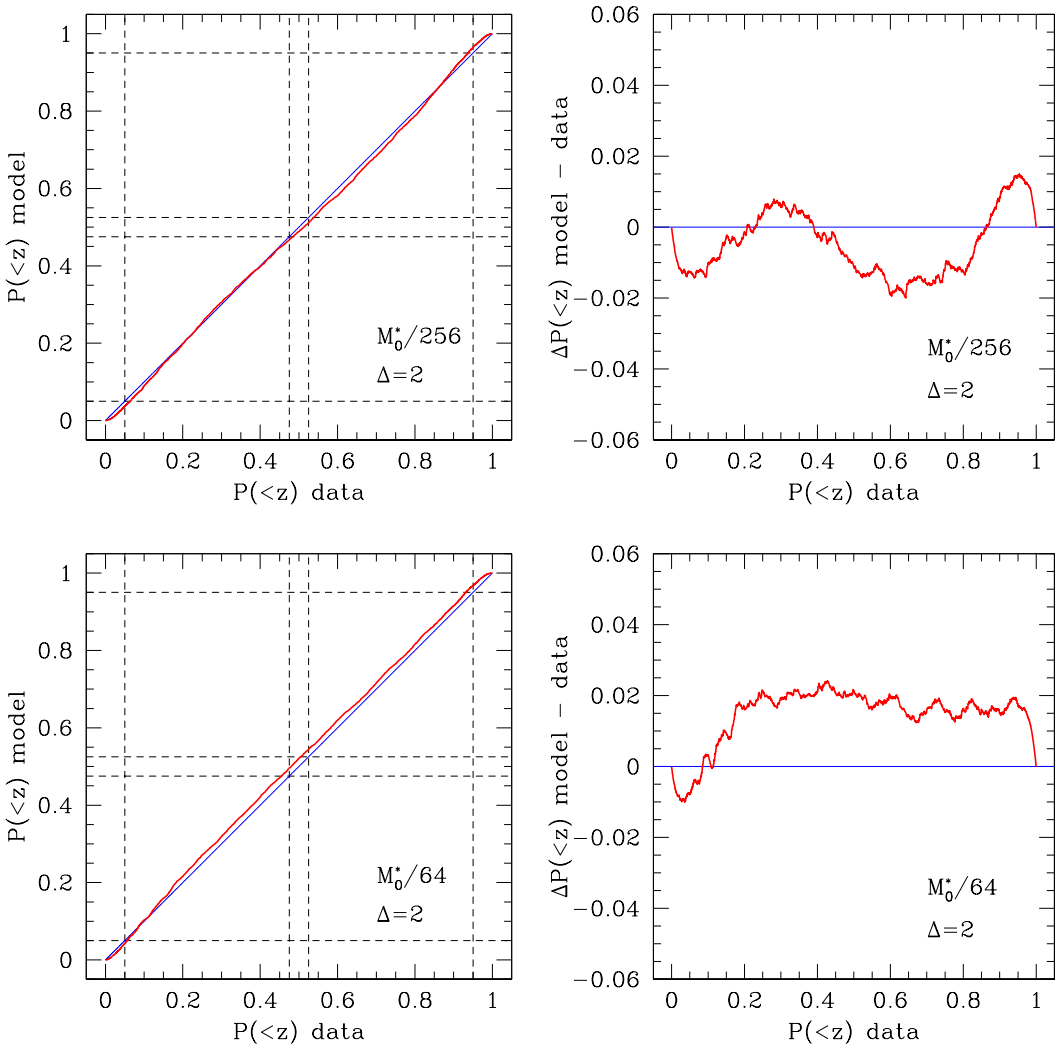

Figure C.1: Same as Figure 3.24, for $\Delta = 2$ and different masses.

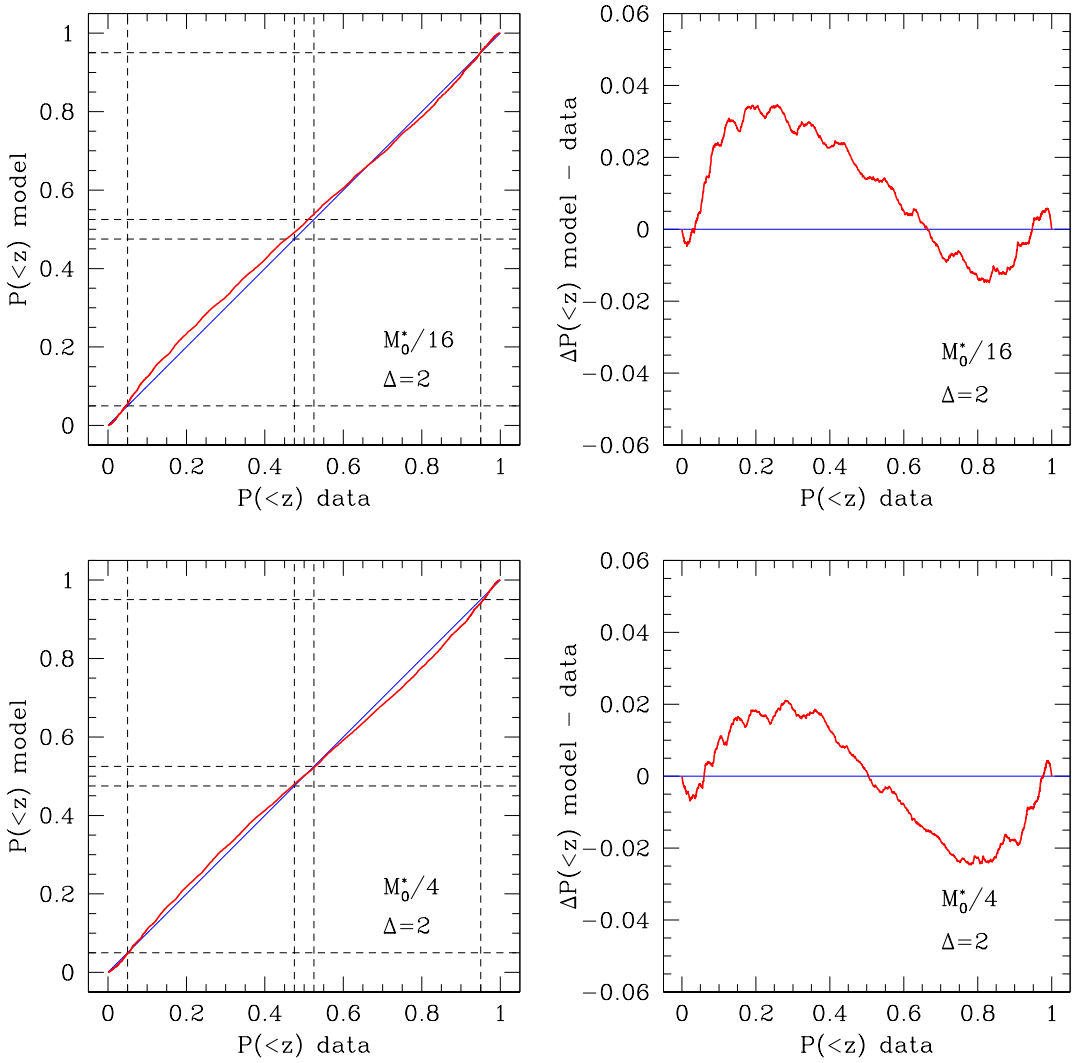

Figure C.1: Same as Figure 3.24, for $\Delta = 2$ and different masses. *(cont.)*

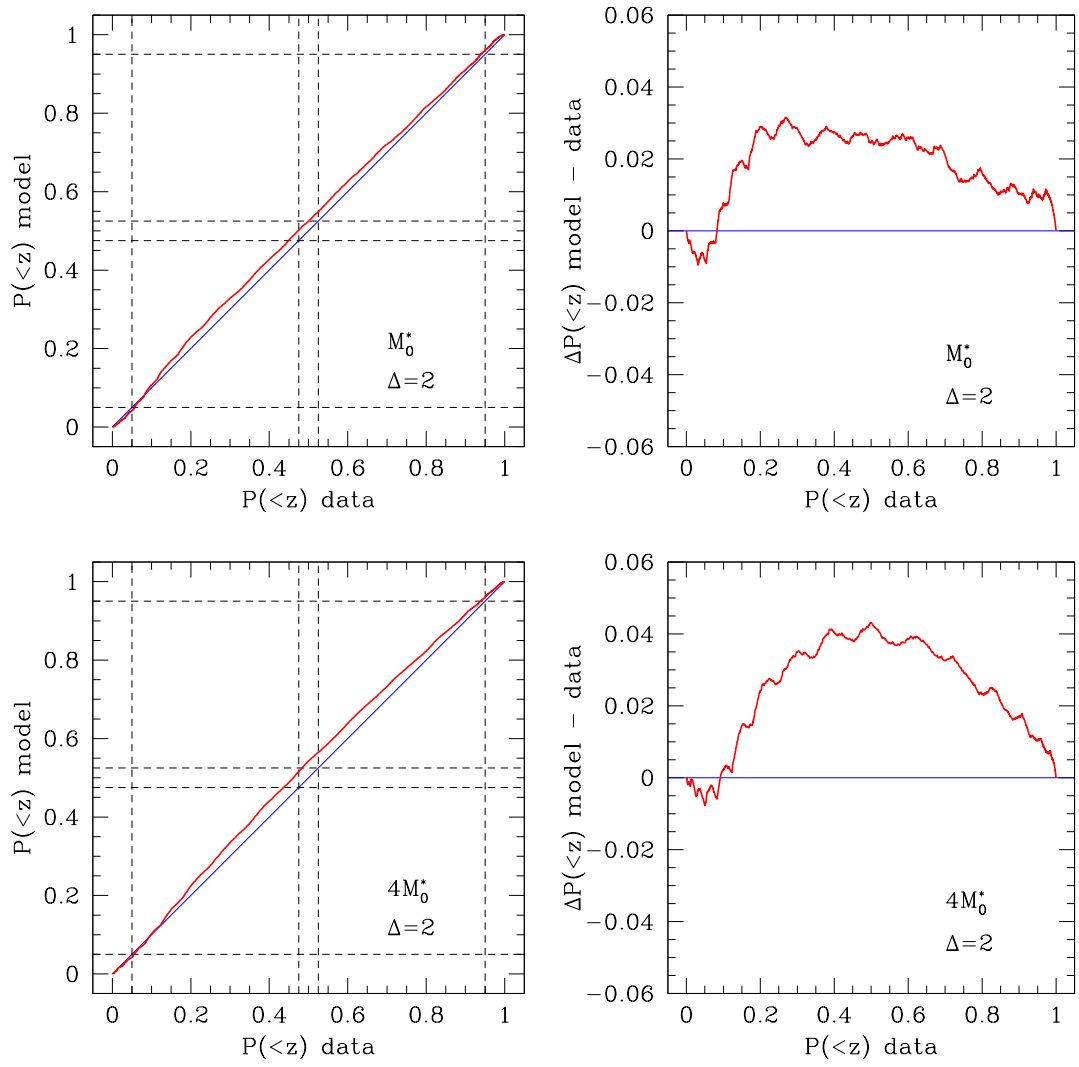

Figure C.1: Same as Figure 3.24, for $\Delta = 2$ and different masses. *(cont.)*

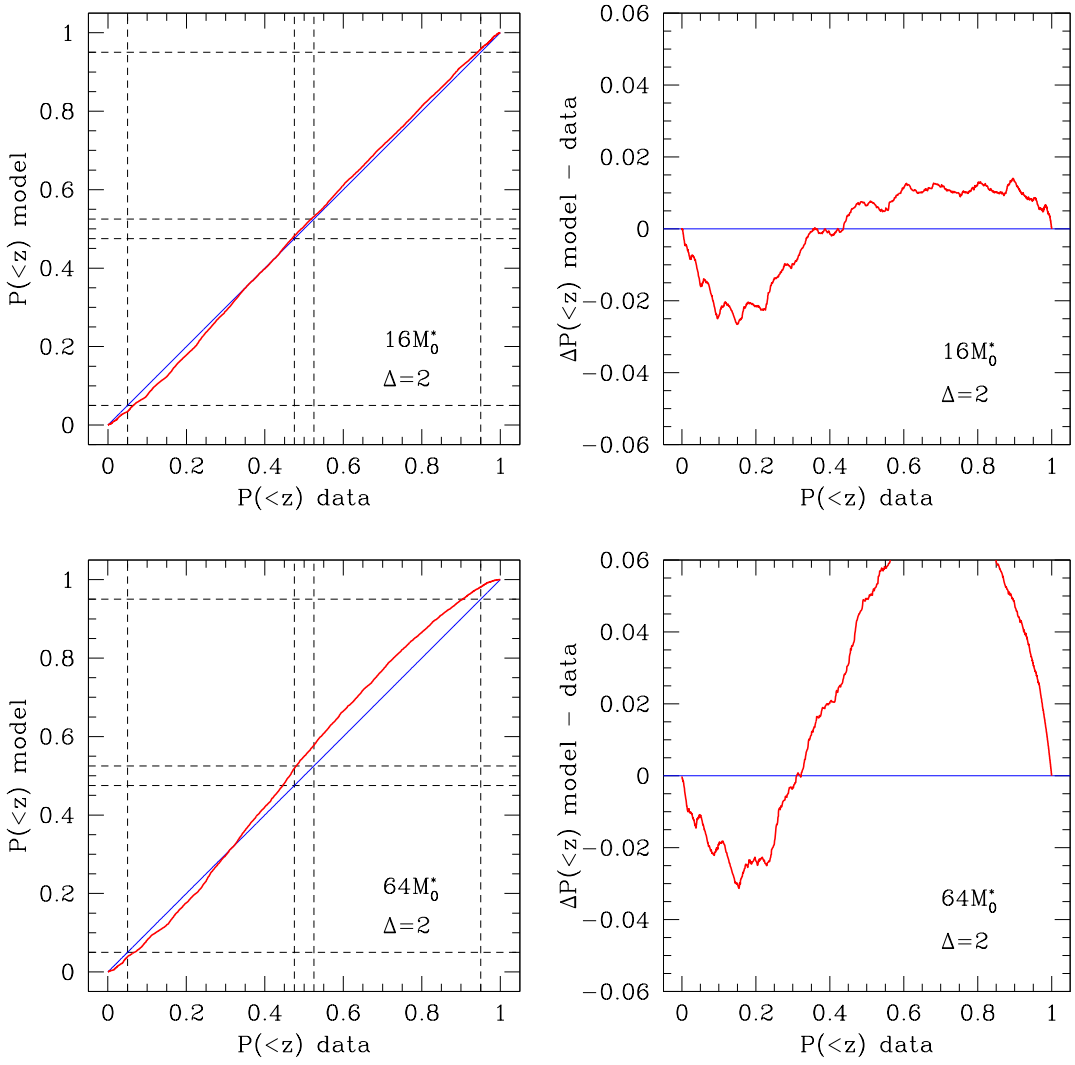

Figure C.1: Same as Figure 3.24, for $\Delta = 2$ and different masses. *(cont.)*

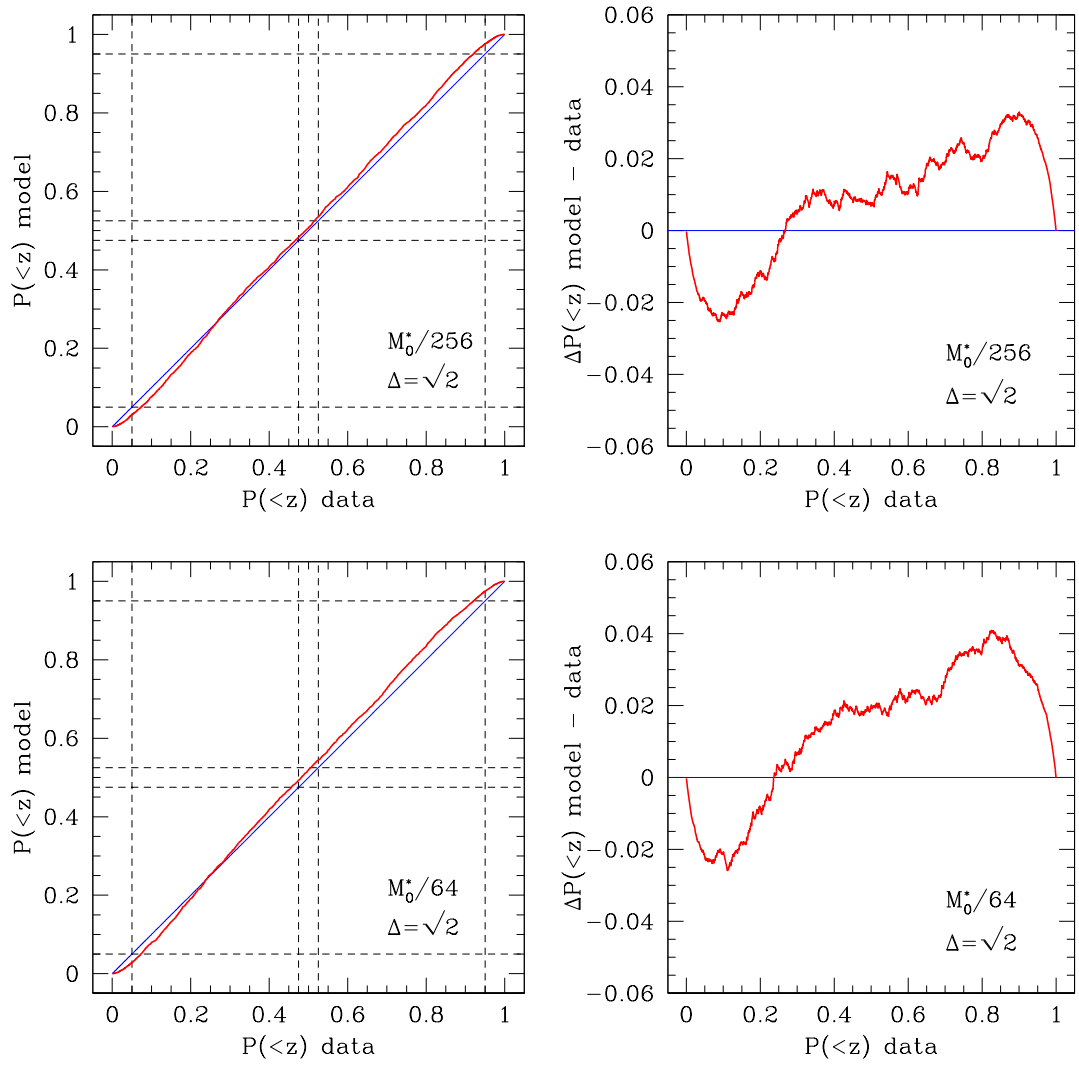

Figure C.2: Same as Figure 3.24, for $\Delta = \sqrt{2}$ and different masses.

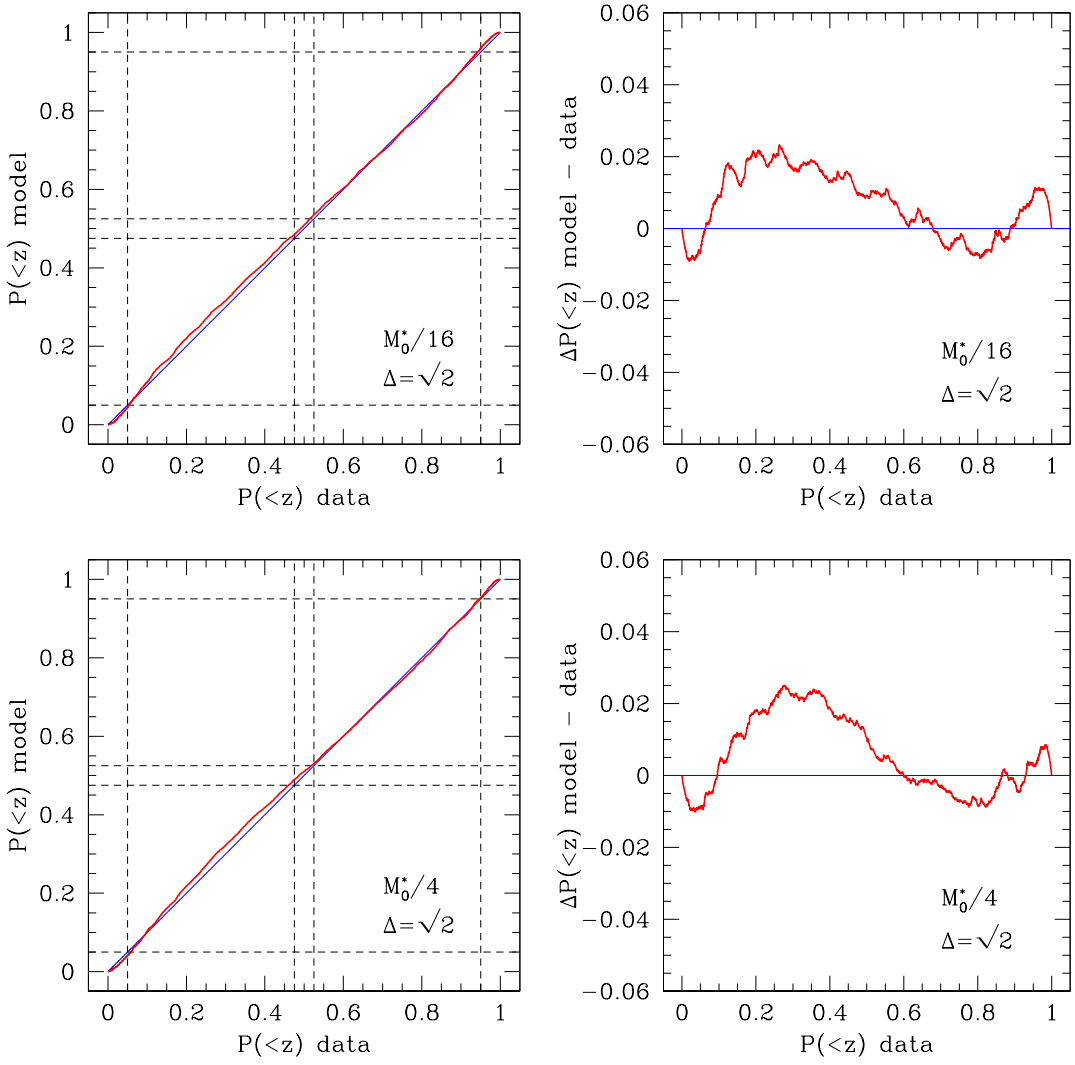

Figure C.2: Same as Figure 3.24, for $\Delta = \sqrt{2}$ and different masses. *(cont.)*

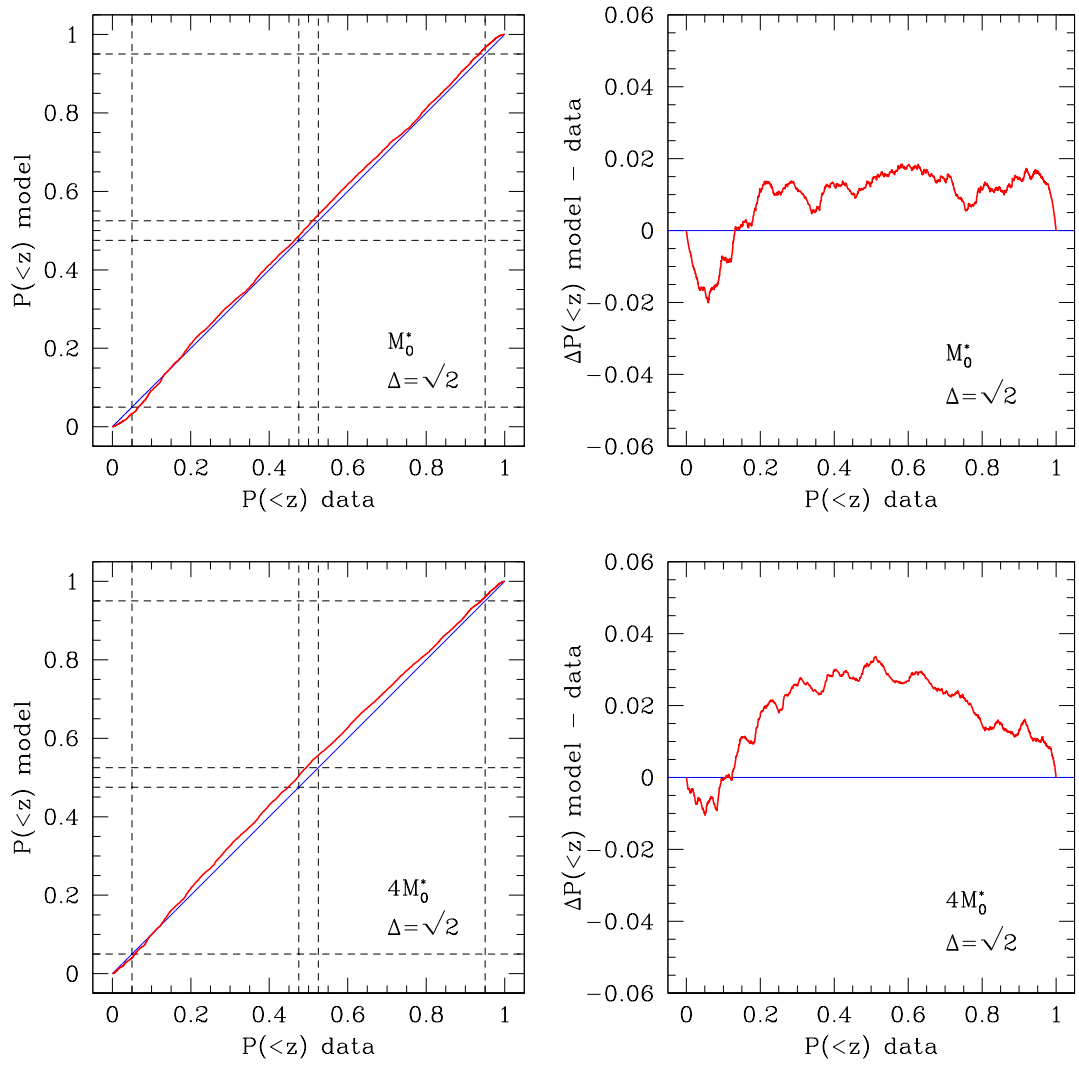

Figure C.2: Same as Figure 3.24, for $\Delta = \sqrt{2}$ and different masses. *(cont.)*

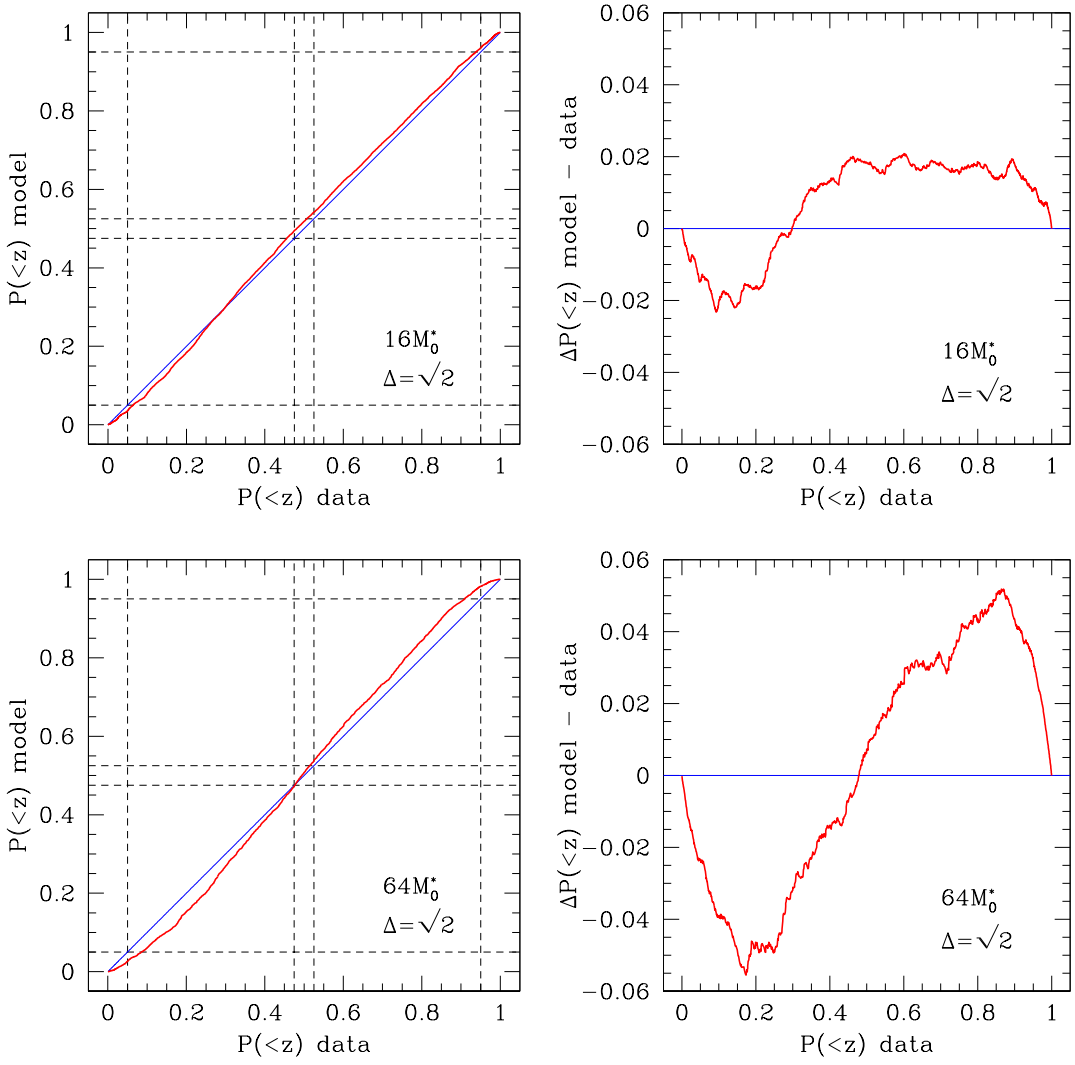

Figure C.2: Same as Figure 3.24, for $\Delta = \sqrt{2}$ and different masses. *(cont.)*

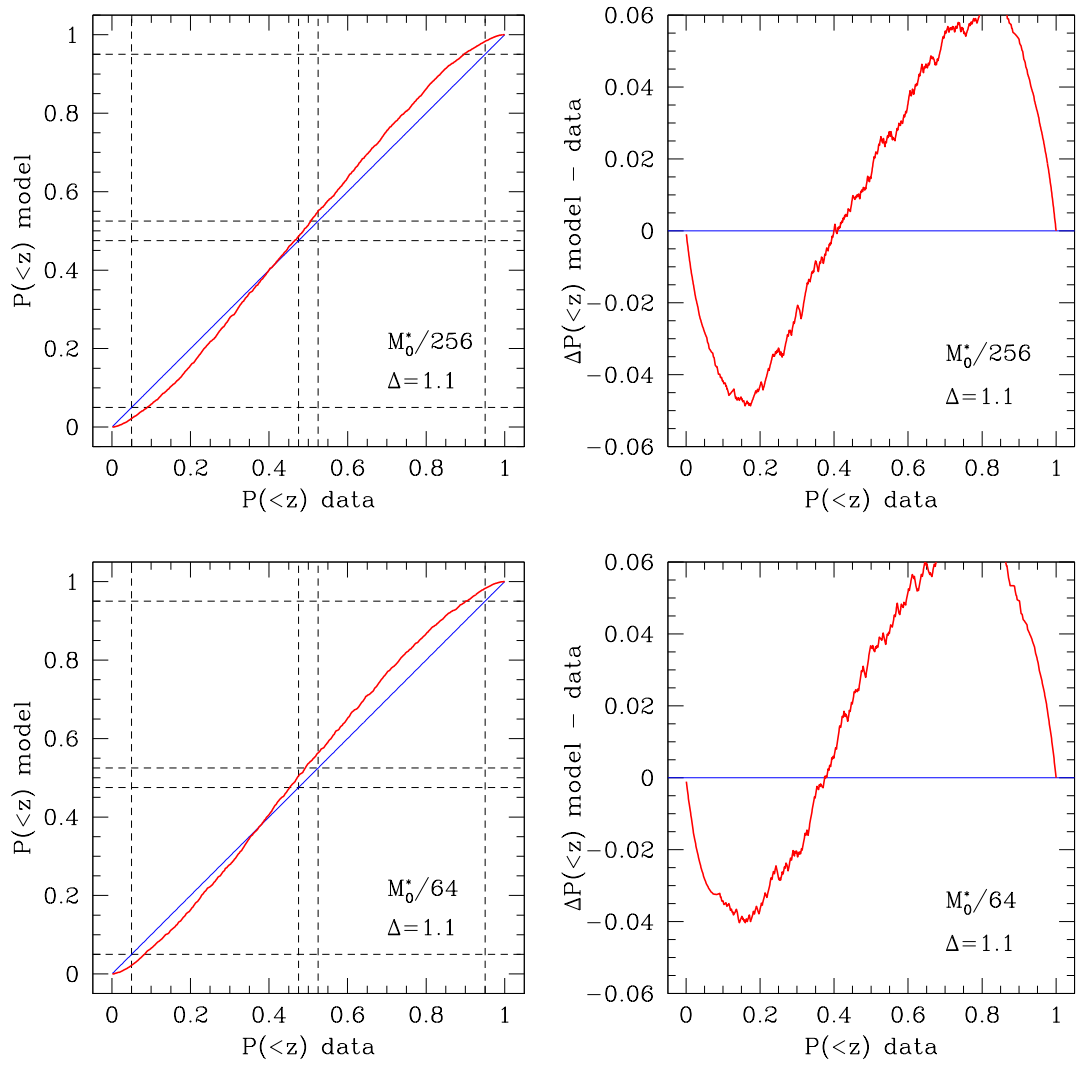

Figure C.3: Same as Figure 3.24, for $\Delta = 1.1$ and different masses.

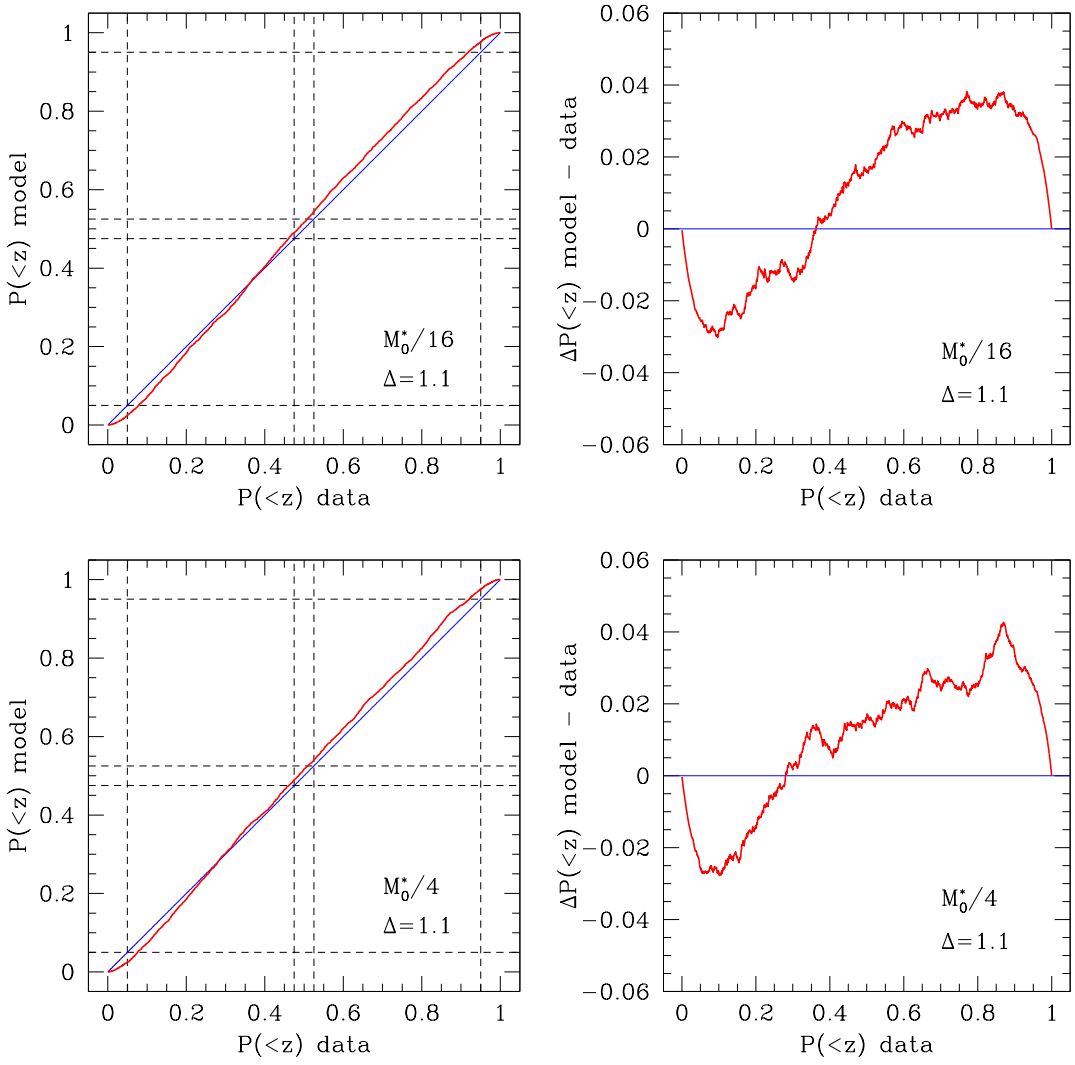

Figure C.3: Same as Figure 3.24, for Δ = 1.1 and different masses. *(cont.)*

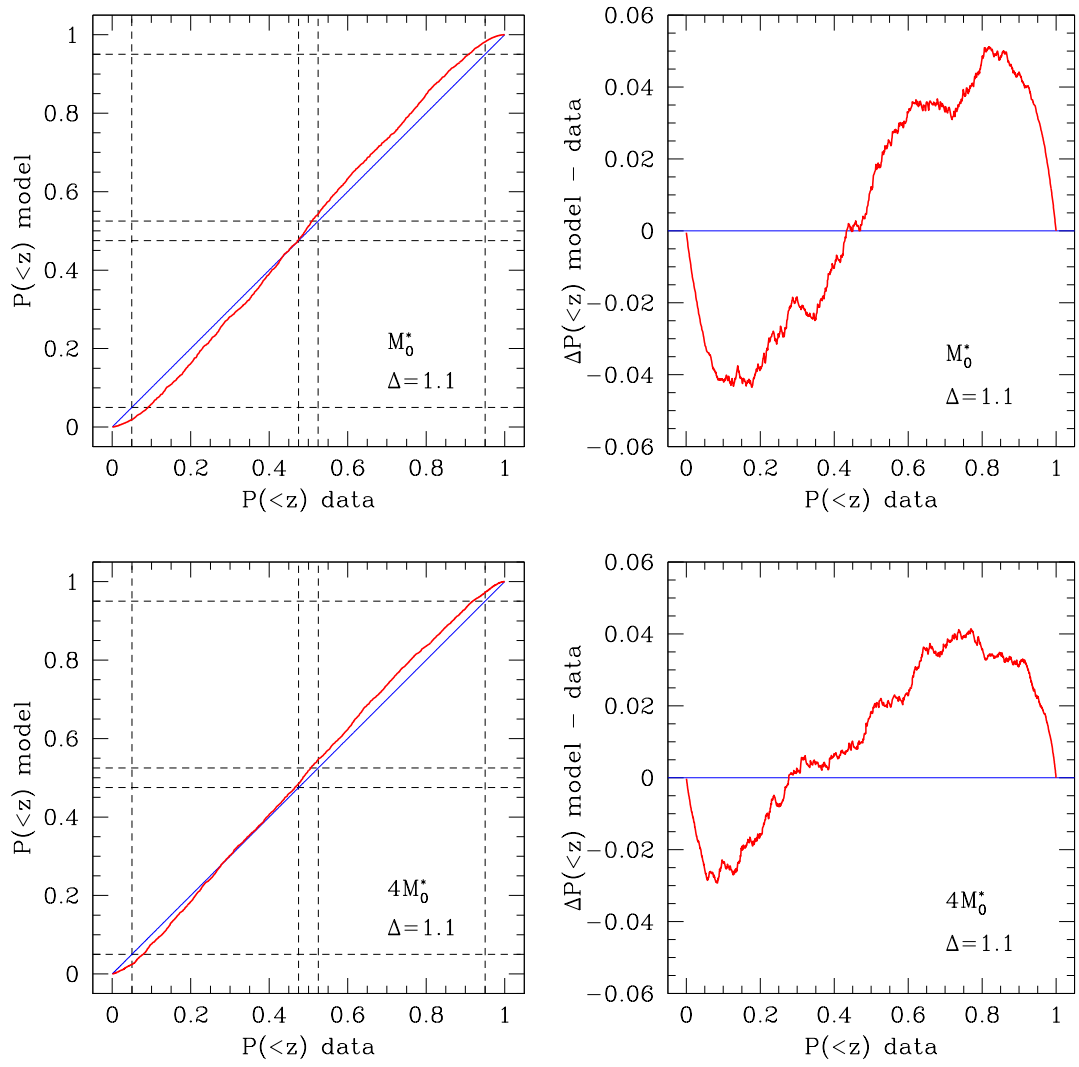

Figure C.3: Same as Figure 3.24, for $\Delta = 1.1$ and different masses. *(cont.)*

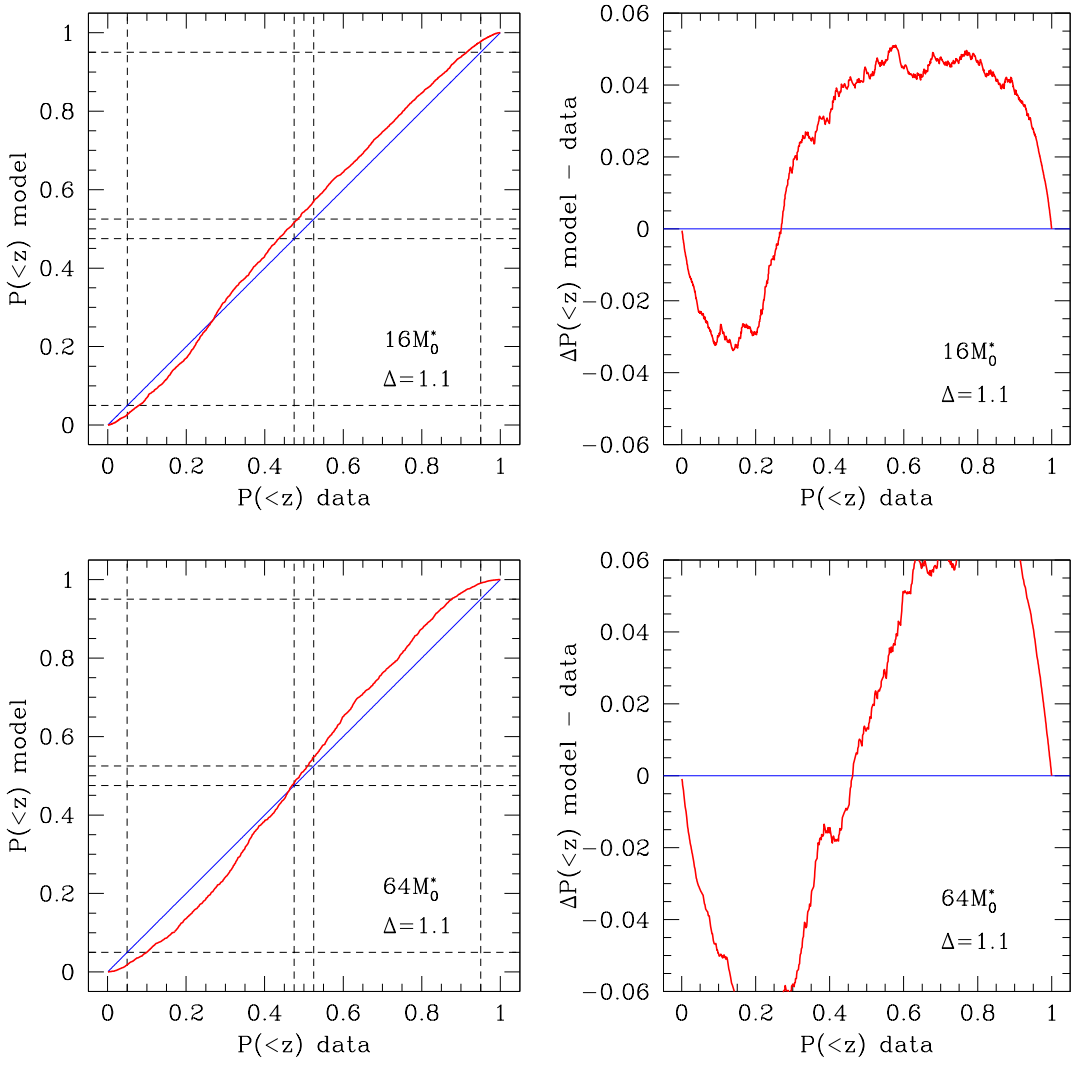

Figure C.3: Same as Figure 3.24, for $\Delta = 1.1$ and different masses. *(cont.)*

Growth factor D

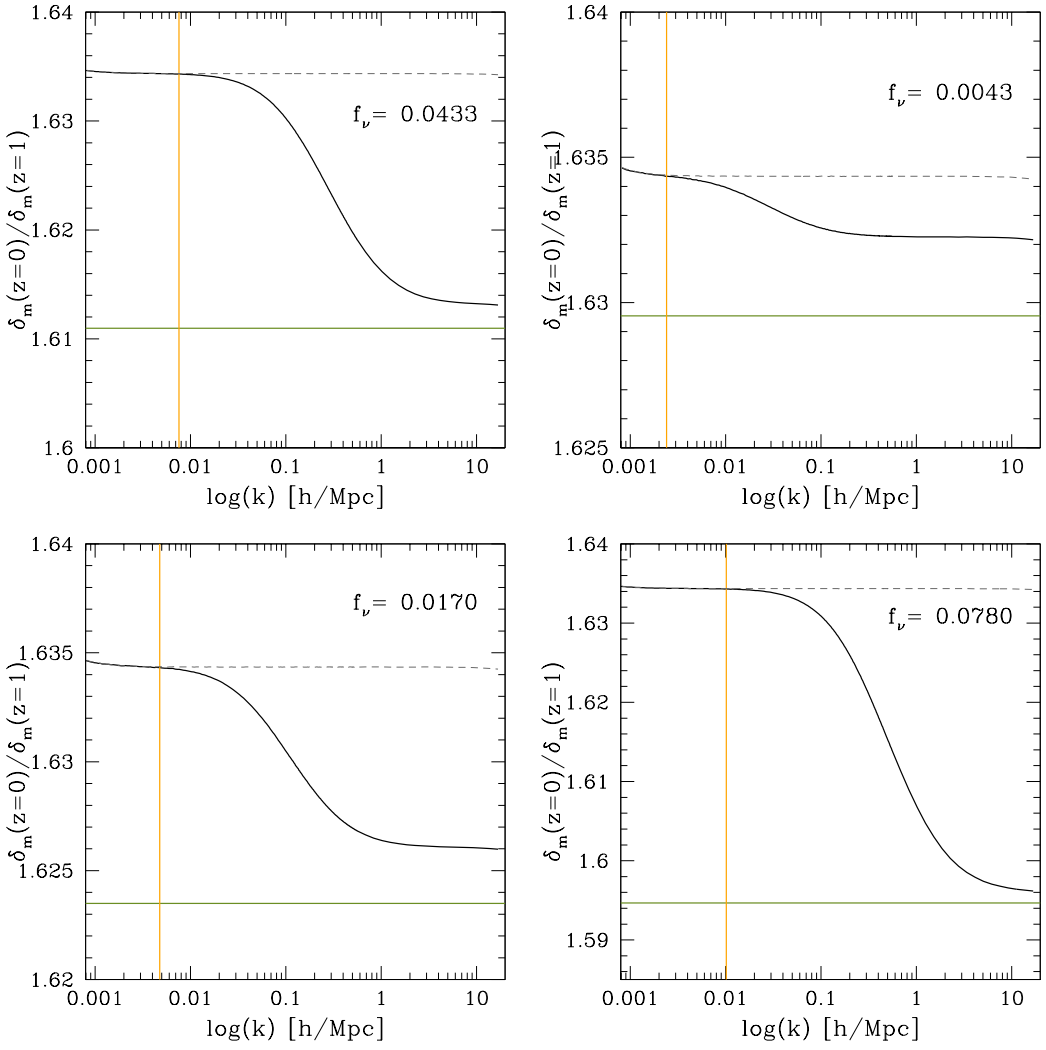

Figure D.1: Same as Figure 4.3, for $z = 1$.

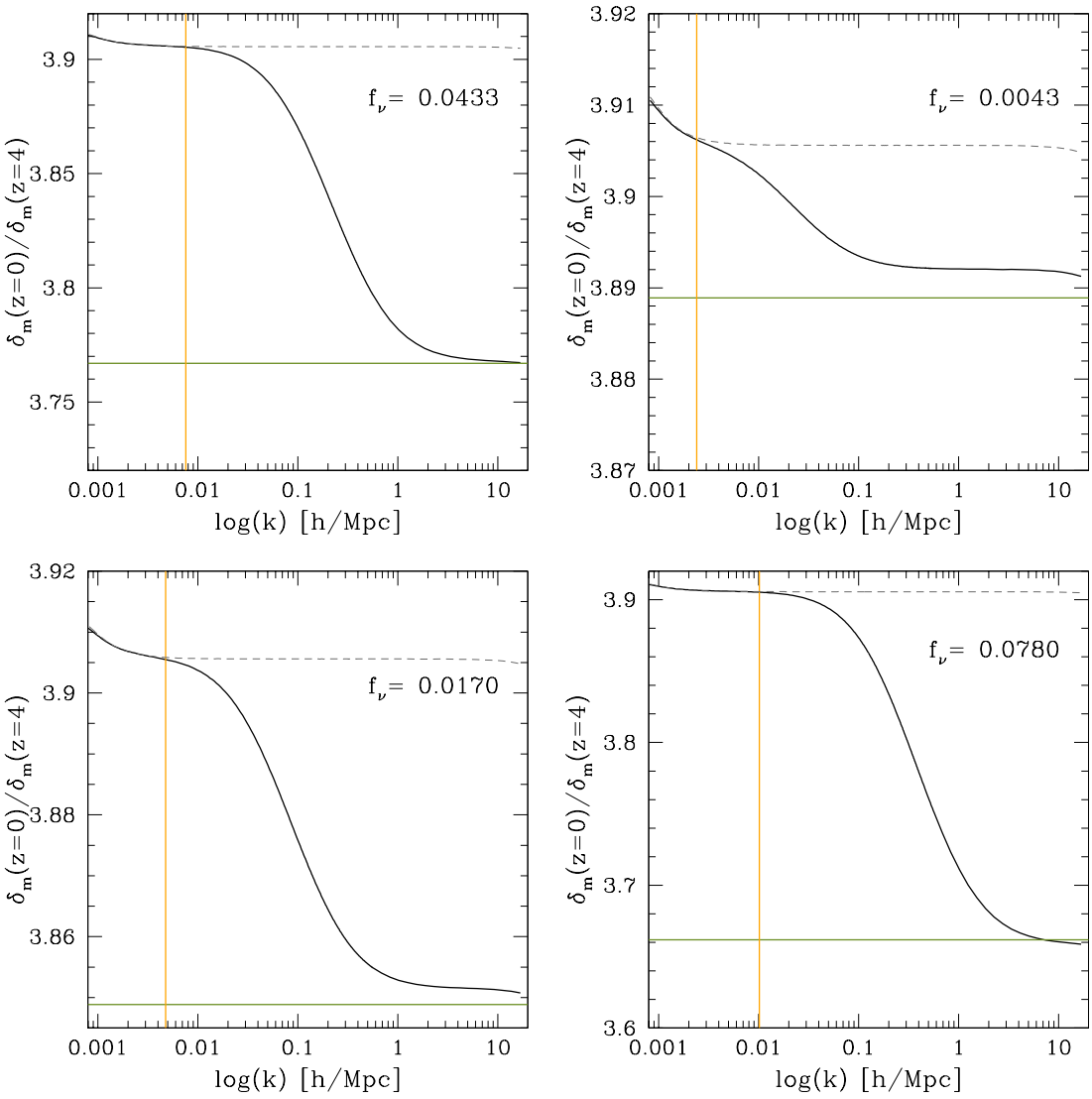

Figure D.2: Same as Figure 4.3, for $z = 4$.

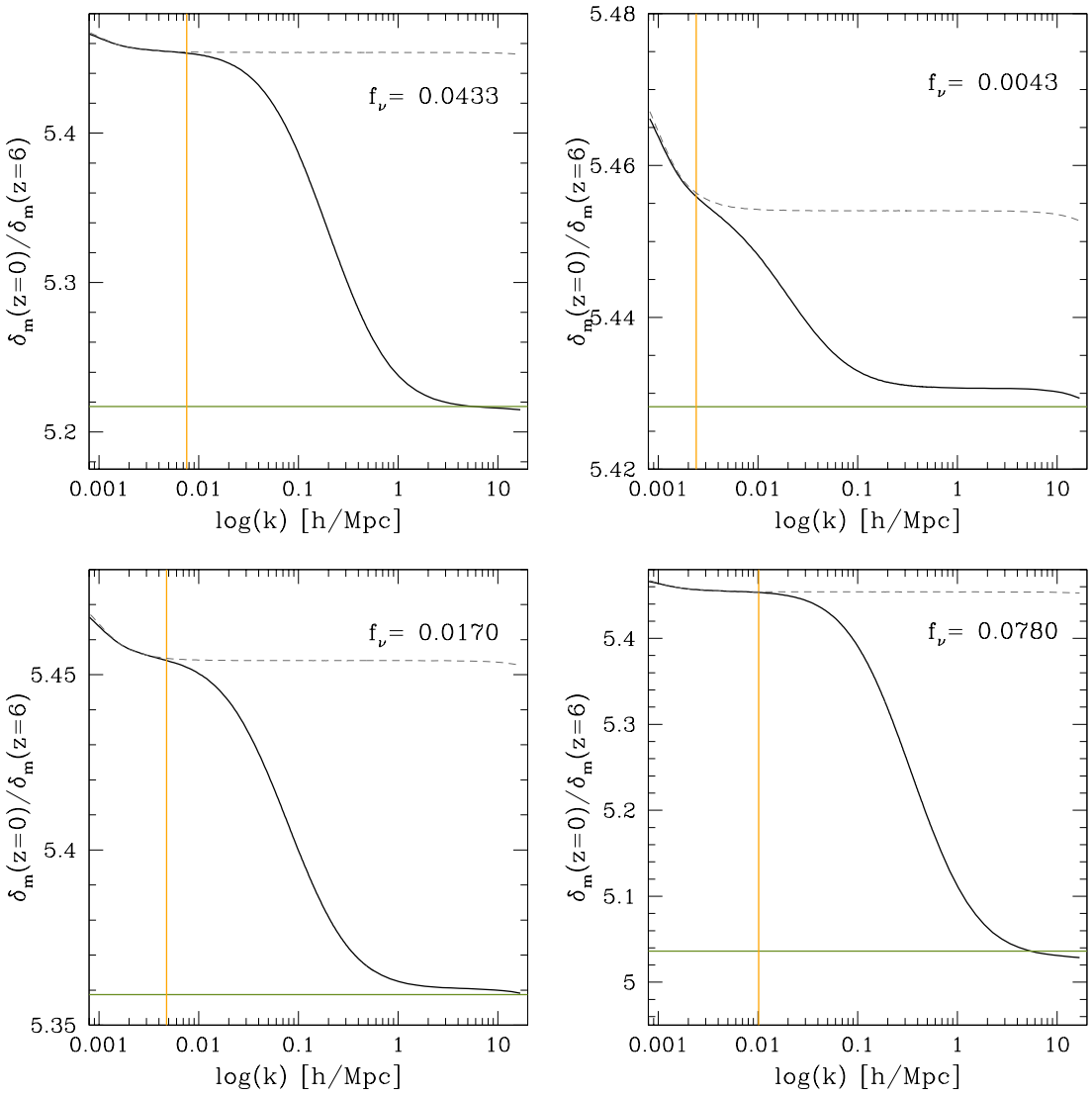

Figure D.3: Same as Figure 4.3, for $z = 6$.

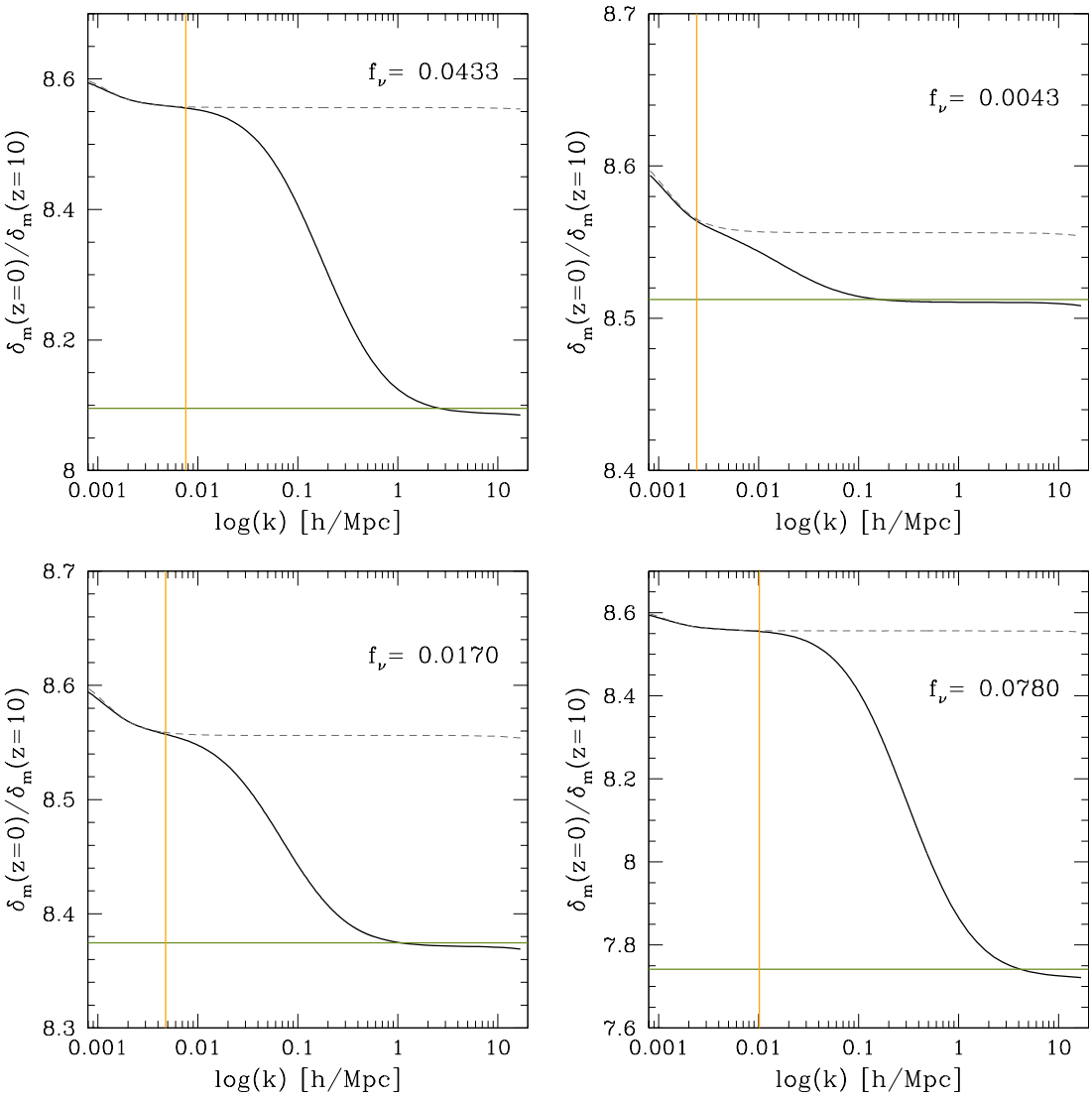

Figure D.4: Same as Figure 4.3, for $z = 10$.

Acknowledgment

I acknowledge the Department of Information Engineering of the University, that gave us access to their IBM computing cluster Power7/P770[1], which we used for running Ada.
I thank the nice people at http://stackoverflow.com and the StackExchange communities for their selfless assistance during many late night bug-hunting and compilation issues, and for providing an invaluable repository of information and tricks.
I thank the Python and the Open Source communities for the many, many libraries and tidbits of code that I used and adapted for this work.

I also thank my friends and colleagues for the many discussions and exchange of opinions on aesthetics and all graphical matters.
I want to thank especially Rosamaria, for proofreading and providing valuable feedback and positive criticism.

[1] http://acg.dei.unipd.it/index.php?option=com_content&view=article&id=13&Itemid=41

List of Figures

List of Tables

Bibliography

S. Abe, T. Ebihara, S. Enomoto, K. Furuno, Y. Gando, K. Ichimura, H. Ikeda, K. Inoue, Y. Kibe, Y. Kishimoto, M. Koga, A. Kozlov, Y. Minekawa, T. Mitsui, K. Nakajima, K. Nakajima, K. Nakamura, M. Nakamura, K. Owada, I. Shimizu, Y. Shimizu, J. Shirai, F. Suekane, A. Suzuki, Y. Takemoto, K. Tamae, A. Terashima, H. Watanabe, E. Yonezawa, S. Yoshida, J. Busenitz, T. Classen, C. Grant, G. Keefer, D. S. Leonard, D. McKee, A. Piepke, M. P. Decowski, J. A. Detwiler, S. J. Freedman, B. K. Fujikawa, F. Gray, E. Guardincerri, L. Hsu, R. Kadel, C. Lendvai, K.-B. Luk, H. Murayama, T. O'Donnell, H. M. Steiner, L. A. Winslow, D. A. Dwyer, C. Jillings, C. Mauger, R. D. McKeown, P. Vogel, C. Zhang, B. E. Berger, C. E. Lane, J. Maricic, T. Miletic, M. Batygov, J. G. Learned, S. Matsuno, S. Pakvasa, J. Foster, G. A. Horton-Smith, A. Tang, S. Dazeley, K. E. Downum, G. Gratta, K. Tolich, W. Bugg, Y. Efremenko, Y. Kamyshkov, O. Perevozchikov, H. J. Karwowski, D. M. Markoff, W. Tornow, K. M. Heeger, F. Piquemal, and J.-S. Ricol. Precision Measurement of Neutrino Oscillation Parameters with KamLAND. *Physical Review Letters*, 100(22):221803, June 2008. doi: 10.1103/PhysRevLett.100.221803.

Q. R. Ahmad, R. C. Allen, T. C. Andersen, J. D. Anglin, G. Bühler, J. C. Barton, E. W. Beier, M. Bercovitch, J. Bigu, S. Biller, R. A. Black, I. Blevis, R. J. Boardman, J. Boger, E. Bonvin, M. G. Boulay, M. G. Bowler, T. J. Bowles, S. J. Brice, M. C. Browne, T. V. Bullard, T. H. Burritt, K. Cameron, J. Cameron, Y. D. Chan, M. Chen, H. H. Chen, X. Chen, M. C. Chon, B. T. Cleveland, E. T. Clifford, J. H. Cowan, D. F. Cowen, G. A. Cox, Y. Dai, X. Dai, F. Dalnoki-Veress, W. F. Davidson, P. J. Doe, G. Doucas, M. R. Dragowsky, C. A. Duba, F. A. Duncan, J. Dunmore, E. D. Earle, S. R. Elliott, H. C. Evans, G. T. Ewan, J. Farine, H. Fergani, A. P. Ferraris, R. J. Ford, M. M. Fowler, K. Frame, E. D. Frank, W. Frati, J. V. Germani, S. Gil, A. Goldschmidt, D. R. Grant, R. L. Hahn, A. L. Hallin, E. D. Hallman, A. Hamer, A. A. Hamian, R. U. Haq, C. K. Hargrove, P. J. Harvey, R. Hazama, R. Heaton, K. M. Heeger, W. J. Heintzelman, J. Heise, R. L. Helmer, J. D. Hepburn, H. Heron, J. Hewett, A. Hime, M. Howe, J. G. Hykawy, M. C. Isaac, P. Jagam, N. A. Jelley, C. Jillings, G. Jonkmans, J. Karn, P. T. Keener, K. Kirch, J. R. Klein, A. B. Knox, R. J. Komar, R. Kouzes, T. Kutter, C. C. Kyba, J. Law, I. T. Lawson, M. Lay, H. W. Lee, K. T. Lesko, J. R. Leslie, I. Levine, W. Locke, M. M. Lowry, S. Luoma, J. Lyon, S. Majerus, H. B. Mak, A. D. Marino, N. McCauley, A. B. McDonald, D. S. McDonald, K. McFarlane, G. McGregor, W. McLatchie, R. M. Drees, H. Mes, C. Mifflin, G. G. Miller, G. Milton, B. A. Moffat, M. Moorhead, C. W. Nally, M. S. Neubauer, F. M. Newcomer, H. S. Ng, A. J. Noble, E. B. Norman, V. M. Novikov, M. O'Neill, C. E.

Okada, R. W. Ollerhead, M. Omori, J. L. Orrell, S. M. Oser, A. W. Poon, T. J. Radcliffe, A. Roberge, B. C. Robertson, R. G. Robertson, J. K. Rowley, V. L. Rusu, E. Saettler, K. K. Schaffer, A. Schuelke, M. H. Schwendener, H. Seifert, M. Shatkay, J. J. Simpson, D. Sinclair, P. Skensved, A. R. Smith, M. W. Smith, N. Starinsky, T. D. Steiger, R. G. Stokstad, R. S. Storey, B. Sur, R. Tafirout, N. Tagg, N. W. Tanner, R. K. Taplin, M. Thorman, P. Thornewell, P. T. Trent, Y. I. Tserkovnyak, R. van Berg, R. G. van de Water, C. J. Virtue, C. E. Waltham, J.-X. Wang, D. L. Wark, N. West, J. B. Wilhelmy, J. F. Wilkerson, J. Wilson, P. Wittich, J. M. Wouters, and M. Yeh. Measurement of the Rate of $\nu_e + d \rightarrow p + p + e^-$ Interactions Produced by ^{8}B Solar Neutrinos at the Sudbury Neutrino Observatory. *Physical Review Letters*, 87(7):071301, August 2001. doi: 10.1103/PhysRevLett.87.071301.

B. Banneheka and G. Ekanayake. A new point estimator for the median of gamma distribution. *Vidyodaya Journal of Science*, 14(1), jan 2010. URL http://vjs.sljol.info/articles/abstract/1534/.

G. Baso. `simulation planner`. http://goo.gl/N9NmTA, February 2014.

J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion set mass functions for hierarchical Gaussian fluctuations. *ApJ*, 379:440–460, October 1991. doi: 10.1086/170520.

J. Brandbyge, S. Hannestad, T. Haugbølle, and Y. Y. Y. Wong. Neutrinos in non-linear structure formation — the effect on halo properties. *JCAP*, 9:014, September 2010. doi: 10.1088/1475-7516/2010/09/014.

P.J. Brown, W.A. Fuller, American Mathematical Society, Institute of Mathematical Statistics, Society for Industrial, and Applied Mathematics. *Statistical Analysis of Measurement Error Models and Applications: Proceedings of the AMS-IMS-SIAM Joint Summer Research Conference Held June 10-16, 1989, with Support from the National Science Foundation and the U.S. Army Research Office*. Conm Ser : Vol.112. American Mathematical Society, 1990. ISBN 9780821851173. URL https://books.google.it/books?id=TRkcCAAAQBAJ.

G. L. Bryan and M. L. Norman. Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. *ApJ*, 495:80–99, March 1998. doi: 10.1086/305262.

J. S. Bullock, T. S. Kolatt, Y. Sigad, R. S. Somerville, A. V. Kravtsov, A. A. Klypin, J. R. Primack, and A. Dekel. Profiles of dark haloes: evolution, scatter and environment. *MNRAS*, 321:559–575, March 2001. doi: 10.1046/j.1365-8711.2001.04068.x.

E. Castorina, C. Carbone, J. Bel, E. Sefusatti, and K. Dolag. DEMNUni: the clustering of large-scale structures in the presence of massive neutrinos. *JCAP*, 7:043, July 2015. doi: 10.1088/1475-7516/2015/07/043.

B. T. Cleveland, T. Daily, R. Davis, Jr., J. R. Distel, K. Lande, C. K. Lee, P. S. Wildenhain, and J. Ullman. Measurement of the Solar Electron Neutrino Flux with the Homestake Chlorine Detector. *ApJ*, 496: 505–526, March 1998. doi: 10.1086/305343.

S. Coleman and S. L. Glashow. Cosmic ray and neutrino tests of special relativity. *Physics Letters B*, 405:249–252, February 1997. doi: 10.1016/S0370-2693(97)00638-2.

C. A. Correa, J. S. B. Wyithe, J. Schaye, and A. R. Duffy. The accretion history of dark matter haloes - III. A physical model for the concentration-mass relation. *MNRAS*, 452:1217–1232, September 2015. doi: 10.1093/mnras/stv1363.

V. de Sabbata and M. Gasperini. Macroscopical consequences of a propagating torsion potential. *Nuovo Cimento Lettere*, 30:503–506, April 1981.

G. Despali, C. Giocoli, R. E. Angulo, G. Tormen, R. K. Sheth, G. Baso, and L. Moscardini. The universality of the virial halo mass function and models for non-universality of other halo definitions. *MNRAS*, 456:2486–2504, March 2016. doi: 10.1093/mnras/stv2842.

G. Drexlin, V. Hannen, S. Mertens, and C. Weinheimer. Current Direct Neutrino Mass Experiments. *ArXiv e-prints*, June 2013.

V. R. Eke, S. Cole, and C. S. Frenk. Cluster evolution as a diagnostic for Omega. *MNRAS*, 282, September 1996. doi: 10.1093/mnras/282.1.263.

D. V. Forero, M. Tórtola, and J. W. F. Valle. Global status of neutrino oscillation parameters after Neutrino-2012. *Phys.Rev.D*, 86(7):073012, October 2012. doi: 10.1103/PhysRevD.86.073012.

Y. Fukuda, T. Hayakawa, E. Ichihara, K. Inoue, K. Ishihara, H. Ishino, Y. Itow, T. Kajita, J. Kameda, S. Kasuga, K. Kobayashi, Y. Kobayashi, Y. Koshio, M. Miura, M. Nakahata, S. Nakayama, A. Okada, K. Okumura, N. Sakurai, M. Shiozawa, Y. Suzuki, Y. Takeuchi, Y. Totsuka, S. Yamada, M. Earl, A. Habig, E. Kearns, M. D. Messier, K. Scholberg, J. L. Stone, L. R. Sulak, C. W. Walter, M. Goldhaber, T. Barszczxak, D. Casper, W. Gajewski, P. G. Halverson, J. Hsu, W. R. Kropp, L. R. Price, F. Reines, M. Smy, H. W. Sobel, M. R. Vagins, K. S. Ganezer, W. E. Keig, R. W. Ellsworth, S. Tasaka, J. W. Flanagan, A. Kibayashi, J. G. Learned, S. Matsuno, V. J. Stenger, D. Takemori, T. Ishii, J. Kanzaki, T. Kobayashi, S. Mine, K. Nakamura, K. Nishikawa, Y. Oyama, A. Sakai, M. Sakuda, O. Sasaki, S. Echigo, M. Kohama, A. T. Suzuki, T. J. Haines, E. Blaufuss, B. K. Kim, R. Sanford, R. Svoboda, M. L. Chen, Z. Conner, J. A. Goodman, G. W. Sullivan, J. Hill, C. K. Jung, K. Martens, C. Mauger, C. McGrew, E. Sharkey, B. Viren, C. Yanagisawa, W. Doki, K. Miyano, H. Okazawa, C. Saji, M. Takahata, Y. Nagashima, M. Takita, T. Yamaguchi, M. Yoshida, S. B. Kim, M. Etoh, K. Fujita, A. Hasegawa, T. Hasegawa, S. Hatakeyama, T. Iwamoto, M. Koga, T. Maruyama, H. Ogawa, J. Shirai, A. Suzuki,

F. Tsushima, M. Koshiba, M. Nemoto, K. Nishijima, T. Futagami, Y. Hayato, Y. Kanaya, K. Kaneyuki, Y. Watanabe, D. Kielczewska, R. A. Doyle, J. S. George, A. L. Stachyra, L. L. Wai, R. J. Wilkes, and K. K. Young. Evidence for Oscillation of Atmospheric Neutrinos. *Physical Review Letters*, 81:1562–1567, August 1998. doi: 10.1103/PhysRevLett.81.1562.

L. Gao, S. D. M. White, A. Jenkins, F. Stoehr, and V. Springel. The subhalo populations of ΛCDM dark haloes. *MNRAS*, 355:819–834, December 2004. doi: 10.1111/j.1365-2966.2004.08360.x.

M. Gasperini. Testing the principle of equivalence with neutrino oscillations. *Phys.Rev.D*, 38:2635–2637, October 1988. doi: 10.1103/PhysRevD.38.2635.

C. Giocoli, J. Moreno, R. K. Sheth, and G. Tormen. An improved model for the formation times of dark matter haloes. *MNRAS*, 376:977–983, April 2007. doi: 10.1111/j.1365-2966.2007.11520.x.

C. Giocoli, G. Tormen, and F. C. van den Bosch. The population of dark matter subhaloes: mass functions and average mass-loss rates. *MNRAS*, 386:2135–2144, June 2008. doi: 10.1111/j.1365-2966.2008.13182.x.

C. Giocoli, L. Pieri, G. Tormen, and J. Moreno. A merger tree with microsolar mass resolution: application to γ-ray emission from subhalo population. *MNRAS*, 395:1620–1630, May 2009. doi: 10.1111/j.1365-2966.2009.14649.x.

C. Giocoli, G. Tormen, R. K. Sheth, and F. C. van den Bosch. The substructure hierarchy in dark matter haloes. *MNRAS*, 404:502–517, May 2010. doi: 10.1111/j.1365-2966.2010.16311.x.

C. Giocoli, G. Tormen, and R. K. Sheth. Formation times, mass growth histories and concentrations of dark matter haloes. *MNRAS*, 422:185–198, May 2012. doi: 10.1111/j.1365-2966.2012.20594.x.

F. James and M. Roos. Minuit - a system for function minimization and analysis of the parameter errors and correlations. *Computer Physics Communications*, 10(6):343 – 367, 1975. ISSN 0010-4655. doi: http://dx.doi.org/10.1016/0010-4655(75)90039-9. URL http://www.sciencedirect.com/science/article/pii/0010465575900399.

E. Jones, T. Oliphant, P. Peterson, et al. SciPy: Open source scientific tools for Python, 2001–. URL http://www.scipy.org/. [Online; accessed 2016-11-09].

A. Kiakotou, Ø. Elgarøy, and O. Lahav. Neutrino mass, dark energy, and the linear growth factor. *Phys.Rev.D*, 77(6):063005, March 2008. doi: 10.1103/PhysRevD.77.063005.

C. Lacey and S. Cole. Merger rates in hierarchical models of galaxy formation. *MNRAS*, 262:627–649, June 1993.

C. Lacey and S. Cole. Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-Body Simulations. *MNRAS*, 271:676, December 1994.

O. Lahav and Y. Suto. Measuring our Universe from Galaxy Redshift Surveys. *Living Reviews in Relativity*, 7:8, December 2004. doi: 10.12942/lrr-2004-8.

J. Lesgourgues and S. Pastor. Massive neutrinos and cosmology. *Phys. Rep.*, 429:307–379, July 2006. doi: 10.1016/j.physrep.2006.04.001.

A. Lewis, A. Challinor, and A. Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *ApJ*, 538:473–476, August 2000a. doi: 10.1086/309179.

A. Lewis, A. Challinor, and A. Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *ApJ*, 538:473–476, August 2000b. doi: 10.1086/309179.

M. LoVerde. Spherical collapse in $\nu\Lambda$CDM. *Phys.Rev.D*, 90(8):083518, October 2014. doi: 10.1103/PhysRevD.90.083518.

Y. Lu, H. J. Mo, N. Katz, and M. D. Weinberg. On the origin of cold dark matter halo density profiles. *MNRAS*, 368:1931–1940, June 2006. doi: 10.1111/j.1365-2966.2006.10270.x.

A. D. Ludlow, J. F. Navarro, M. Boylan-Kolchin, P. E. Bett, R. E. Angulo, M. Li, S. D. M. White, C. Frenk, and V. Springel. The mass profile and accretion history of cold dark matter haloes. *MNRAS*, 432:1103–1113, June 2013. doi: 10.1093/mnras/stt526.

A. V. Macciò, A. A. Dutton, and F. C. van den Bosch. Concentration, spin and shape of dark matter haloes as a function of the cosmological model: WMAP1, WMAP3 and WMAP5 results. *MNRAS*, 391:1940–1954, December 2008. doi: 10.1111/j.1365-2966.2008.14029.x.

T. MacFarland, H. M. P. Couchman, F. R. Pearce, and J. Pichlmeier. A new parallel P^3M code for very large-scale cosmological simulations. *New Ast.*, 3:687–705, December 1998. doi: 10.1016/S1384-1076(98)00033-5.

P. Madau and M. Dickinson. Cosmic Star-Formation History. *ARA&A*, 52:415–486, August 2014. doi: 10.1146/annurev-astro-081811-125615.

G. Mangano, G. Miele, S. Pastor, T. Pinto, O. Pisanti, and P. D. Serpico. Relic neutrino decoupling including flavour oscillations. *Nuclear Physics B*, 729:221–234, November 2005. doi: 10.1016/j.nuclphysb.2005.09.041.

J. McBride, O. Fakhouri, and C.-P. Ma. Mass accretion rates and histories of dark matter haloes. *MNRAS*, 398:1858–1868, October 2009. doi: 10.1111/j.1365-2966.2009.15329.x.

P. Meszaros. The behaviour of point masses in an expanding cosmological substratum. *A&A*, 37: 225–228, December 1974.

J. F. Navarro, C. S. Frenk, and S. D. M. White. The Structure of Cold Dark Matter Halos. *ApJ*, 462:563, May 1996. doi: 10.1086/177173.

J. F. Navarro, C. S. Frenk, and S. D. M. White. A Universal Density Profile from Hierarchical Clustering. *ApJ*, 490:493, December 1997. doi: 10.1086/304888.

A. F. Neto, L. Gao, P. Bett, S. Cole, J. F. Navarro, C. S. Frenk, S. D. M. White, V. Springel, and A. Jenkins. The statistics of Λ CDM halo concentrations. *MNRAS*, 381:1450–1462, November 2007. doi: 10.1111/j.1365-2966.2007.12381.x.

M.E. Peskin and D.V. Schroeder. *An Introduction to Quantum Field Theory*. Advanced book classics. Addison-Wesley Publishing Company, 1995. ISBN 9780201503975. URL https://books.google.it/books?id=_H-oPv1raioC.

Planck Collaboration. Planck 2013 results. XVI. Cosmological parameters. *A&A*, 571:A16, November 2014. doi: 10.1051/0004-6361/201321591.

Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *A&A*, 594:A13, September 2016. doi: 10.1051/0004-6361/201525830.

B. M. Poggianti, A. Aragón-Salamanca, D. Zaritsky, G. De Lucia, B. Milvang-Jensen, V. Desai, P. Jablonka, C. Halliday, G. Rudnick, J. Varela, S. Bamford, P. Best, D. Clowe, S. Noll, R. Saglia, R. Pelló, L. Simard, A. von der Linden, and S. White. The Environments of Starburst and Post-Starburst Galaxies at z = 0.4-0.8. *ApJ*, 693:112–131, March 2009. doi: 10.1088/0004-637X/693/1/112.

B. Pontecorvo. Mesonium and anti-mesonium. *Sov.Phys.- JETP*, 6:429–431, 1957. URL www.scopus.com.

W. H. Press and P. Schechter. Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation. *ApJ*, 187:425–438, February 1974. doi: 10.1086/152650.

R. K. Sheth, M. Bernardi, P. L. Schechter, S. Burles, D. J. Eisenstein, D. P. Finkbeiner, J. Frieman, R. H. Lupton, D. J. Schlegel, M. Subbarao, K. Shimasaku, N. A. Bahcall, J. Brinkmann, and Ž. Ivezić. The Velocity Dispersion Function of Early-Type Galaxies. *ApJ*, 594:225–231, September 2003. doi: 10.1086/376794.

N. Smith and S. van der Walt. A better default colormap for matplotlib. SciPy2015 - Scientific Computing with Python, jul 2015. URL http://bids.github.io/colormap/.

M. Smy. Super-Kamiokande#700s Solar ν Results. *Nuclear Physics B Proceedings Supplements*, 235: 49–54, February 2013. doi: 10.1016/j.nuclphysbps.2013.03.010.

V. Springel. The cosmological simulation code GADGET-2. *MNRAS*, 364:1105–1134, December 2005. doi: 10.1111/j.1365-2966.2005.09655.x.

V. Springel, N. Yoshida, and S. D. M. White. GADGET: a code for collisionless and gasdynamical cosmological simulations. *New Ast.*, 6:79–117, April 2001. doi: 10.1016/S1384-1076(01)00042-2.

Volker Springel. `n-genic`. http://www.mpa-garching.mpg.de/gadget, 2003.

E. W. Stacy and G. A. Mihram. Parameter Estimation for a Generalized Gamma Distribution. *Technometrics*, 7(3):349–358, 1965. doi: 10.1080/00401706.1965.10490268. URL http://amstat.tandfonline.com/doi/abs/10.1080/00401706.1965.10490268.

G. Tormen. The assembly of matter in galaxy clusters. *MNRAS*, 297:648–656, June 1998. doi: 10.1046/j.1365-8711.1998.01545.x.

G. Tormen, L. Moscardini, and N. Yoshida. Properties of cluster satellites in hydrodynamical simulations. *MNRAS*, 350:1397–1408, June 2004. doi: 10.1111/j.1365-2966.2004.07736.x.

F. C. van den Bosch. The universal mass accretion history of cold dark matter haloes. *MNRAS*, 331: 98–110, March 2002. doi: 10.1046/j.1365-8711.2002.05171.x.

F. C. van den Bosch, F. Jiang, A. Hearin, D. Campbell, D. Watson, and N. Padmanabhan. Coming of age in the dark sector: how dark matter haloes grow their gravitational potential wells. *MNRAS*, 445: 1713–1730, December 2014. doi: 10.1093/mnras/stu1872.

M. Viel, M. G. Haehnelt, and V. Springel. The effect of neutrinos on the matter distribution as probed by the intergalactic medium. *JCAP*, 6:015, June 2010. doi: 10.1088/1475-7516/2010/06/015.

R. H. Wechsler, J. S. Bullock, J. R. Primack, A. V. Kravtsov, and A. Dekel. Concentrations of Dark Halos from Their Assembly Histories. *ApJ*, 568:52–70, March 2002. doi: 10.1086/338765.

S. D. M. White. Formation and Evolution of Galaxies: Les Houches Lectures. *ArXiv Astrophysics e-prints*, October 1994.

D. H. Zhao, Y. P. Jing, H. J. Mo, and G. Börner. Accurate Universal Models for the Mass Accretion Histories and Concentrations of Dark Matter Halos. *ApJ*, 707:354–369, December 2009. doi: 10.1088/0004-637X/707/1/354.