Universidad de Oviedo

Programa de Doctorado en Materiales

THE SUBMILLIMETER GALAXY MAGNIFICATION BIAS: AN UNDEREXPLOITED COSMOLOGICAL PROBE

EL SESGO DE MAGNIFICACIÓN DE LAS GALAXIAS SUBMILIMÉTRICAS: UNA SONDA COSMOLÓGICA INFRAUTILIZADA

TESIS DOCTORAL

Marcos Muñiz Cueli

Septiembre 2022

Universidad de Oviedo

Programa de Doctorado en Materiales

THE SUBMILLIMETER GALAXY MAGNIFICATION BIAS: AN UNDEREXPLOITED COSMOLOGICAL PROBE

EL SESGO DE MAGNIFICACIÓN DE LAS GALAXIAS SUBMILIMÉTRICAS: UNA SONDA COSMOLÓGICA INFRAUTILIZADA

TESIS DOCTORAL

Directores de tesis

Dr. D. Joaquín González-Nuevo González

Dra. Dña. Laura Bonavera

Universidad de Oviedo

RESUMEN DEL CONTENIDO DE TESIS DOCTORAL

1.- Título de la Tesis	
Español: El sesgo de magnificación de las galaxias submilimétricas: una sonda cosmológica infrautilizada	Inglés: The submillimeter galaxy magnification bias: an underexploited cosmological probe

2.- Autor	
Nombre: Marcos Muñiz Cueli	DNI:
Programa de Doctorado: Materiales	
Órgano responsable: Comisión Académica Programa de Doctorado en Materiales	

RESUMEN (en español)

El análisis del Universo a gran escala en las últimas décadas ha evidenciado la necesidad de utilizar diferentes observables para dar respuesta a los problemas sin resolver de la cosmología moderna y avanzar hacia el establecimiento de un modelo físico de concordancia. En esta línea, la presente tesis propone el sesgo de magnificación de las galaxias submilimétricas como una sonda innovadora e independiente que contribuya a arrojar luz sobre el denominado Universo oscuro.

La luz proveniente de fuentes lejanas es desviada debido a la distribución de masa (o lentes) entre el observador y las mismas, lo que puede dar lugar a distorsiones, imágenes múltiples o una amplificación de su flujo. Cuando la muestra de fuentes de fondo se escoge adecuadamente y se le impone un umbral mínimo de flujo, el resultado es un sesgo observacional conocido como sesgo de magnificación, el cual se manifiesta cuantitativamente a través de la observación de un exceso de fuentes lejanas alrededor de las lentes respecto a lo esperado en ausencia de las últimas. Debido a las propiedades físicas de las galaxias submilimétricas (alto *redshift*, baja emisión en el óptico y pendiente acuciada del número de cuentas), la señal se vuelve especialmente significativa cuando este tipo de fuentes es utilizado como muestra de fondo, dando nombre así al observable que da título a esta tesis.

La correlación observada entre el número de galaxias de fondo y lente puede ser interpretada dentro del marco teórico de las lentes gravitacionales, lo que pone de manifiesto su dependencia explícita con parámetros cosmológicos esenciales para la comprensión del Universo a gran escala, como la densidad promedio de materia, la varianza del campo de densidad lineal o la tasa de expansión del espacio. Además, supuesto el modelo de halos para explicar el proceso no lineal de formación de estructuras, este observable muestra una dependencia adicional con cómo las galaxias pueblan los halos de materia oscura (distribución de ocupación de halos o DOH), la densidad numérica de los mismos o su concentración. Mediante la explotación de las herramientas de la inferencia Bayesiana, se pueden cotejar las medidas con su predicción teórica y obtener restricciones observacionales sobre todas las cantidades mencionadas.

La presente tesis estudia la correlación cruzada entre galaxias submilimétricas de fondo extraídas del catálogo de la misión *Herschel*-ATLAS y galaxias lente ópticas. Mediante análisis de creciente complejidad, se examinarán los efectos de la DOH en la cosmología y la densidad numérica de halos, así como el potencial de nuestro observable para restringir estas cantidades. La influencia de la metodología de estimación será analizada, así como la utilización de muestras de galaxias lente de distinto carácter (fotométrico y espectroscópico). Además, se buscarán mejoras en los resultados a través de un estudio tomográfico que divida la muestra lente en distintos *bins* de *redshift* y permita la evolución temporal de cantidades como la densidad de energía oscura. Finalmente, se resumirán las líneas actuales de trabajo y el amplio abanico de posibilidades futuras de esta *sonda cosmológica infrautilizada*.

Universidad de Oviedo

RESUMEN (en Inglés)

The study of the large-scale Universe in the last few decades has made it clear that different observables are needed to give an answer to the unsolved problems in modern cosmology as well as to make progress toward the establishment of a physical concordance model. On this note, this thesis proposes the submillimeter galaxy magnification bias as a novel and independent cosmological probe that contributes to shed light on the so-called dark Universe.

The light from distant sources is deflected due to the mass distribution between them and the observer, which may give rise to distortions, multiple images or magnifications. When the background sample of sources is properly chosen and a minimum flux threshold is imposed, the result is an observational bias known as magnification bias, which manifests itself in a quantitative manner via the observation of an excess of background sources around those in the foreground with respect to the expected situation in absence of the latter. Due to the optimal physical properties of submillimeter galaxies (high redshift, low emission in the optical band and steep number counts), the signal becomes especially significant when these sources are used as the background sample, thus naming the observable in the title of this thesis.

The observed correlation between the number of foreground and background galaxies can be interpreted within the theoretical framework of gravitational lensing, which highlights its explicit dependence on cosmological parameters that are essential to the understanding of the large-scale Universe, like the average matter density, the variance of the linear overdensity field or the expansion rate of space. Furthermore, under the halo model of non-linear structure formation, this observable shows an additional dependence on how galaxies populate dark matter halos (the halo occupation distribution or HOD), their number density or their concentration. Through the tools of Bayesian inference, the measurements can be checked against their theoretical prediction so as to obtain observational constraints on all the aforementioned quantities.

The present PhD thesis studies the cross-correlation between submillimeter background galaxies extracted from the catalog of the *Herschel*-ATLAS survey and optical foreground galaxies. By means of analyses of increasing complexity, the effects of the HOD on cosmology and the number density of dark matter halos will be examined, along with the potential of the submillimeter galaxy magnification bias to constraint these quantities. The influence of the estimation methodology on the results will be studied, together with the use of foreground galaxy samples of a different nature (photometric and spectroscopic). Improvements on our findings will be pursued through a tomographic analysis that divides the foreground sample in different bins of redshift to allow for the time evolution of quantities like the dark energy density. Lastly, the ongoing lines of work will be summarized along with the wide range of future possibilities of this *underexploited cosmological probe*.

SR. PRESIDENTE DE LA COMISIÓN ACADÉMICA DEL PROGRAMA DE DOCTORADO EN MATERIALES

"Sometimes science is more art than science, Morty. A lot of people don't get that".

- Rick Sanchez, *Rick and Morty.*

Abstract

The study of the large-scale Universe in the last few decades has made it clear that different and complementary observables are needed to give an answer to the unsolved problems in modern cosmology as well as to make progress toward the establishment of a physical concordance model. On this note, the present thesis proposes the submillimeter galaxy magnification bias as a novel and independent cosmological probe that can contribute to shed light on the so-called dark Universe.

The light from distant sources is deflected due to the mass distribution between them and the observer, which may give rise to distortions, multiple images or magnifications. When the background sample of sources is properly chosen and a minimum flux threshold is imposed, the result is an observational effect known as magnification bias, which manifests itself in a quantitative manner via an excess of background sources around those in the foreground with respect to the expected situation in absence of the latter. Due to the optimal physical properties of submillimeter galaxies (high redshift, low emission in the optical band and steep number counts), the signal becomes especially significant when these sources are used as the background sample, thus naming the observable in the title of this thesis.

The observed correlation between the number of foreground and background galaxies can be interpreted within the theoretical framework of gravitational lensing, which highlights its explicit dependence on cosmological parameters that are essential to the understanding of the large-scale Universe, like the average matter density, the variance of the linear overdensity field or the expansion rate of space. Furthermore, under the halo model of non-linear structure formation, this observable shows an additional dependence on how galaxies populate dark matter halos (the halo occupation distribution or HOD), their number density or their concentration. Through the tools of Bayesian inference, the measurement of this magnification bias signal can be checked against its theoretical prediction so as to obtain observational constraints on all the aforementioned quantities.

The present PhD thesis studies the cross-correlation between submillimeter background galaxies extracted from a catalog of the *Herschel*-ATLAS survey and optical foreground galaxies. By means of analyses of increasing complexity, the effects of the HOD on cosmology and the number density of dark matter halos will be examined, along with the potential of the submillimeter galaxy magnification bias to constrain these quantities. The influence of the estimation methodology on the results will be studied, together with the use of foreground galaxy samples of a different nature, namely photometric and spectroscopic. Improvements on our findings will be pursued through a tomographic analysis that divides the foreground sample in different bins of redshift to allow for the time evolution of quantities like the dark energy density. Lastly, the ongoing lines of work will be summarized along with the wide range of future possibilities of this *underexploited cosmological probe*.

Resumen

El análisis del Universo a gran escala en las últimas décadas ha evidenciado la necesidad de combinar observables diferentes y complementarios para dar respuesta a los problemas sin resolver de la cosmología moderna y avanzar hacia el establecimiento de un modelo físico de concordancia. En esta línea, la presente tesis propone el sesgo de magnificación de las galaxias submilimétricas como una sonda innovadora e independiente que contribuya a arrojar luz sobre el denominado Universo oscuro.

La luz proveniente de fuentes lejanas es desviada debido a la distribución de masa (o lentes) entre el observador y las mismas, lo que puede dar lugar a distorsiones, imágenes múltiples o una amplificación de su flujo. Cuando la muestra de fuentes de fondo se escoge adecuadamente y se le impone un umbral mínimo de flujo, el resultado es un efecto observacional conocido como sesgo de magnificación, el cual se manifiesta cuantitativamente a través de un exceso de fuentes lejanas alrededor de las lentes respecto a lo esperado en ausencia de las últimas. Debido a las propiedades físicas de las galaxias submilimétricas (alto *redshift*, baja emisión en el óptico y pendiente acuciada del número de cuentas), la señal se vuelve especialmente significativa cuando este tipo de fuentes es utilizado como muestra de fondo, dando nombre así al observable que figura en el título de esta tesis.

La correlación observada entre el número de galaxias de fondo y lente puede ser interpretada dentro del marco teórico de las lentes gravitacionales, lo que pone de manifiesto su dependencia explícita con parámetros cosmológicos esenciales para la comprensión del Universo a gran escala, como la densidad promedio de materia, la varianza del campo de densidad lineal o la tasa de expansión del espacio. Además, supuesto el modelo de halos para explicar el proceso no lineal de formación de estructuras, este observable muestra una dependencia adicional con cómo las galaxias pueblan los halos de materia oscura (distribución de ocupación de halos o DOH), la densidad numérica de los mismos o su concentración. Mediante la explotación de las herramientas de la inferencia Bayesiana, se puede cotejar la medida del sesgo de magnificación con su predicción teórica y obtener restricciones observacionales sobre todas las cantidades mencionadas.

La presente tesis estudia la correlación cruzada entre galaxias submilimétricas de fondo extraídas de un catálogo de la misión *Herschel*-ATLAS y galaxias lente ópticas. Mediante análisis de creciente complejidad, se examinarán los efectos de la DOH en la cosmología y la densidad numérica de halos, así como el potencial de nuestro observable para restringir estas cantidades. La influencia de la metodología de estimación será estudiada, así como la utilización de muestras de galaxias lente de distinto carácter, a saber, fotométrico y espectroscópico. Además, se buscarán mejoras en los resultados a través de un análisis tomográfico que divida la muestra lente en distintos *bins* de *redshift* y permita la evolución temporal de cantidades como la densidad de energía oscura. Finalmente, se resumirán las líneas actuales de trabajo y el amplio abanico de posibilidades futuras de esta *sonda cosmológica infrautilizada*.

Contents

1. Introduction

"Ergo vivida vis animi pervicit, et extra processit longe flammantia moenia mundi atque omne immensum peragravit mente animoque".[1]

- Lucretius, *De rerum natura.*

The understanding of the Universe as a whole, however much it resembles the dream of a true megalomaniac, has largely benefited from the emergence of precision cosmology at the beginning of the century. The analysis of the Cosmic Microwave Background (CMB) anisotropies by the *Planck* mission, which built upon the prior work of the COsmic Background Explorer (COBE) and the Wilkinson Microwave Anisotropy Probe (WMAP), has provided very tight constraints on cosmological parameters (Planck Collaboration et al., 2020a) and, therefore, on the kind of Universe we inhabit. However, in order to establish a true concordance cosmological model, as well as to try and tackle the parameter degeneracies inherent to CMB analyses, additional data is required from complementary cosmological probes like type IA supernovae (Leibundgut and Sullivan, 2018), cluster abundances (Abbott et al., 2020), the Lyman α forest (Palanque-Delabrouille et al., 2015) or gravitational lensing (Abbott et al., 2022), to name a few with associated recent works.

Along this line, this thesis is centered on proposing the *submillimeter galaxy magnification bias* as an independent and complementary cosmological probe that can contribute to unravel the mysteries shrouding our Universe. This introductory chapter is aimed at providing a first cursory approach to the essentials of this work. It consists of a brief historical motivation, an exposition of the main goals and associated published contributions and a detailed description of the structure of the thesis.

1.1 Background and motivation

The phenomenon of magnification bias, a consequence of gravitational lensing, modifies the observed number density of sources in the sky in a flux-limited sample according to the so-called magnification field due to the mass of the lenses, that is, the objects between the sources and the observer. This effect, which depends directly on the logarithmic slope of the background source number counts, can induce a higher-than-average number of background sources close to the foreground lenses, which can be quantified via a non-negligible cross-correlation between the number of foreground-background pairs within a given angular distance.

This association between high- and low-z objects was first remarked by the American physicist Halton Arp in the 1960s via the detection of high-z quasars very close to low-z galaxies (see Arp, 1987, and references therein), which led him to believe that the redshift of the former could not possibly have a cosmological origin. Although such a claim does not hold in light of the overwhelming evidence against it, there was not a convincing explanation at the time for such apparent associations. However, further studies reported an ever-increasing number of positive correlations between high-z sources and low-z objects, like those found by Fugmann (1990) between bright, radio-loud, high-z active galactic nuclei from the 1-Jansky catalog (Kuehr et al., 1981) and low-z galaxies from the Lick catalog (Seldner et al., 1977). A long series of subsequent papers up to the beginning of the millenium,

[1] *"The living force of his soul gained the day: on he passed far beyond the flaming walls of the world and traversed throughout in mind and spirit the immeasurable universe".*

a non-exhaustive list of which can be found in Scranton et al. (2005) or Wang et al. (2011), reported a significant scatter in this kind of observational correlations, ranging from positive associations to null or negative ones. The results were deemed controversial or inconclusive given the discrepancy between the claimed detections (significant only at 2 or 3σ) and the early theoretical estimations. Indeed, it was clear that magnification bias due to individual low-z galaxies could certainly not account for this effect, since they could only increase the number density of background quasars in a region of a few arcsec around them (Bartelmann and Schneider, 2001; Schneider et al., 2006), but even lensing by the large-scale structure traced by the foreground galaxies did not seem sufficient to explain the observed signal.

As argued by Schneider et al. (2006), at the start of the century it was a matter of debate whether these discrepancies were due to the simplifying assumptions of the theoretical models or to purely observational effects. However, the 8σ detection by Scranton et al. (2005) of a positive correlation between ~ 200000 quasars and ~ 13 million galaxies from the Sloan Digital Sky Survey (SDSS; York et al., 2000) shed light on the matter and obtained, for the first time, the expected amplitude and angular dependence of magnification by large-scale structure. Their analysis also allowed them to conclude that previous deviations from theoretical models had most likely been due to systematic effects in the data that could not be controlled. Aware of these results, Wang et al. (2011) turned their attention to the most relevant factors determining the amplitude of these weak lensing-induced cross-correlations and considered the fact that submillimeter galaxies could constitute an ideal background sample for magnification bias studies given their high redshift distribution and steep number counts. Indeed, one should mention that Almaini et al. (2005) and Blake et al. (2006) had already studied the cross-correlation between background submillimeter galaxies and foreground optical galaxies on the grounds of the good properties of these special sources, but their findings were conflicting, with a significant signal for the former and null results for the latter.

Regardless of the exact conclusions, the necessity of deep wide-area submillimeter surveys was anticipated at the time so that a reliable detection could be carried out. This condition was fulfilled with the advent of the *Herschel* Space Observatory (Pilbratt et al., 2010), whose wide-area surveys provided a large sample of submillimeter sources. This allowed Wang et al. (2011) to measure the cross-correlation between submillimeter galaxies detected by *Herschel* and low-z galaxies, using preliminary data from the Spectral and Photometric Imaging REceiver (SPIRE; Griffin et al., 2010) instrument on *Herschel* and to find convincing evidence of a weak lensing-induced cross-correlation. It was at this point that the precursors of the ideas behind this PhD thesis originated: with a much larger sample size, and using both a photometric and a spectroscopic foreground sample of galaxies along with background submillimeter galaxies detected by the *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al., 2010), González-Nuevo et al. (2014) found a $> 10\sigma$ significance on a spatial cross-correlation, much higher than any other reported by that time. In addition, extensive realistic simulations demonstrated that this cross-correlation could be explained by the magnification bias effect on the background submillimeter galaxies. In a sense, the work of González-Nuevo et al. (2014) showed that the abandoned[2] quasar-galaxy cross-correlations could be resurrected with the right background sample, effectively rendering our observable, the submillimeter galaxy magnification bias, *an underexploited cosmological probe*.

[2]Weak lensing methods based on shear, as will be discussed in Chapter 2, have been favored to probe the galaxy-mass correlation, mainly because of the early controversial results we have discussed, which have to do with the largely varying logarithmic slope of the number counts of quasars, which in turn yielded signals with a low significance.

1.2 Objectives and published contributions

The main goal of this PhD thesis is *to assess the power of the submillimeter galaxy magnification bias as a cosmological probe.* The central idea is to select suitable foreground and background galaxy samples and to estimate the associated angular cross-correlation function, an unambiguous manifestation of the phenomenon of magnification bias when the samples in question do not overlap in redshift. With the choice of an appropriate framework, this estimate can be checked against its theoretical model and the probability distribution for the parameters it depends on can be obtained via a Bayesian estimation approach. To this end, two parallel (but related) secondary objectives can be defined:

- **Observational determination of cosmological and astrophysical quantities**

  Although the foreground-background angular cross-correlation function depends directly on the main parameters of the assumed cosmological model, there is an additional dependence on quantities like the number density of dark matter halos or the mean number of galaxies in a halo when a suitable physical model is adopted. It is thus natural to ask oneself whether these quantities can also be properly constrained alongside cosmology and what their effect is in the main cosmological results.

- **Analysis and refinement of the estimation methodology**

  The procedure of estimating the angular cross-correlation function is all but straightforward. We need to study whether different binning or tiling schemes can affect the results and, if so, how we can correct these biases to maintain consistency. Furthermore, a tomographic analysis carried out by dividing the foreground sample into several redshift bins is likely to improve the constraints from a single-bin study and opens the possibility of studying the redshift evolution of the quantities mentioned in the previous objective, including the dark energy density.

This PhD thesis gathers the main results from the following published contributions, in which I have taken part:

- L. Bonavera, J. González-Nuevo, **M. M. Cueli**, et al. (2020). Cosmology with the submillimetre galaxies magnification bias: Proof of concept. *Astronomy and Astrophysics*, 639:A128, cited throughout the thesis as Bonavera et al. (2020).

- **M. M. Cueli**, L. Bonavera, J. González-Nuevo, et al. (2021). A direct and robust method to observationally constrain the halo mass function via the submillimeter magnification bias: Proof of concept. *Astronomy and Astrophysics*, 645:A126, cited throughout the thesis as Cueli et al. (2021).

- J. González-Nuevo, **M. M. Cueli**, L. Bonavera, et al. (2021). Cosmological constraints on the magnification bias on sub-millimetre galaxies after large-scale bias corrections. *Astronomy and Astrophysics*, 646:A152, cited throughout the thesis as González-Nuevo et al. (2021).

- L. Bonavera, **M. M. Cueli**, J. González-Nuevo, et al. (2021). Cosmology with the submillimetre galaxies magnification bias. Tomographic analysis. *Astronomy and Astrophysics*, 656:A99, cited throughout the thesis as Bonavera et al. (2021).

- **M. M. Cueli**, L. Bonavera, J. González-Nuevo, et al. (2022). Tomography-based observational measurements of the halo mass function via the submillimeter magnification bias. *Astronomy and Astrophysics*, 62:A44, cited throughout the thesis as Cueli et al. (2022).

All five contributions have pursued either the main or secondary objectives, as will be discovered throughout the thesis.

1.3 Outline of the thesis

This PhD thesis has been structured as follows. Chapter 2 is dedicated to establishing the theoretical background of this work. It deals with our current understanding of cosmology, the phenomenon of weak gravitational lensing (with special focus on magnification bias and the weak lensing-induced angular cross-correlation function), and the halo model, the underlying framework that provides us with extremely valuable physical information. Chapter 3 lays out the methodology followed throughout this thesis. The H-ATLAS survey is motivated and described as the leap forward that was needed for submillimeter astrophysics and our source of background objects. Both the foreground and background galaxy samples we used are detailed along with the estimation procedure of the angular cross-correlation function. The last section is dedicated to the Bayesian framework of parameter estimation, with special attention to Markov chain Monte Carlo algorithms. Chapters 4 and 5 exhibit the main results of the published contributions that have been enumerated. The former deals with the non-tomographic analyses of Bonavera et al. (2020) and Cueli et al. (2021), namely preliminary constraints on cosmology, the halo occupation distribution and the halo mass function using a single foreground redshift bin. The results of González-Nuevo et al. (2021) are also included, since they aim to correct the large-scale biases that could contaminate the cross-correlation data and to update the aforementioned cosmological constraints. The second of these two chapters lays out our findings from the extension of our work to a tomographic setup, be it for cosmological (Bonavera et al., 2021) or halo-mass-function (Cueli et al., 2022) purposes, with special focus on the improvements with respect to the proof-of-concept papers. Chapter 6 is a brief segment dedicated to ongoing work and future prospects beyond this thesis, namely the possible improvements on both the methodological and theoretical side and the open lines of work. Chapter 7 concludes the thesis by summarizing the main results and ideas.

2. Theoretical background

This chapter establishes the theoretical framework of the thesis and translates the discussion laid out in the introduction into a quantitative language. The first section describes the basic principles of cosmology that will be necessary in the next segments. The physical phenomenon of weak gravitational lensing will be detailed in the second section, with special focus on magnification bias and its exploitation for the probing of the correlation between galaxies and matter. The last segment of this chapter is dedicated to the halo model as our underlying framework for the description of the process of non-linear structure formation.

2.1 The cosmological framework

"Le silence éternel de ces espaces infinis m'effraie".[1]

- Blaise Pascal, *Les Pensées.*

A somewhat theatrical scientist might state that cosmology is as broad a subject as the Universe. Throwing a veil over such a comment, a plethora of books have indeed been written on just the basics of this cruel science that presents us only the one Universe to observe and theorize. One cannot even help but wonder if such a peculiarity is to blame for it not being considered a proper science until the 20th century. Nevertheless, over the last few decades it has come of age as one of the most attractive fields of study in modern physics. In this section, we will make a brief summary of the very foundations that will be needed throughout the thesis.

2.1.1 The Robertson-Walker metric

The history of cosmology[2], from the ancient Greek civilization to the present days, has slowly (although not steadily) transitioned toward the belief that human beings are not privileged inhabitants of the Universe. The unlikeliness of a special place in space is summarized in the location principle, *"a revolutionary manifesto proclaiming that mankind is no longer king of the cosmic castle"* (Harrison, 2000). This claim is more usually referred to in the literature as the Copernican principle, but the latter might however convey the idea that a center of the Universe should exist elsewhere.

Although it might appear harmless by itself, the location principle is of paramount importance when supplemented with the assumption that the Universe is isotropic on sufficiently large scales. The detection of the nearly isotropic CMB radiation (Penzias and Wilson, 1965) elevates this assumption to the category of observational evidence. The scientific community is considerably convinced that, *"when averaged over sufficiently large scales, there exists a mean motion of radiation and matter in the Universe with respect to which all averaged observable properties are isotropic"* (Bartelmann and Schneider, 2001). Once local density perturbations are averaged over, all directions in the sky appear to be equivalent.

However, there is not a necessary relation between isotropy and homogeneity. A Universe can be isotropic around a given point but not homogeneous, as is the case of a cone in three-dimensional

[1] *"The eternal silence of these infinite spaces frightens me".*
[2] See, for instance, Harrison (2000) for an excellent and comprehensive summary.

Euclidean space. It is patently obvious that we, as static observers, could never prove the (large-scale) homogeneity of the Universe, since we would need to be able to move infinitely fast from one point to another. The only path we can walk down is the adoption of the location principle, which culminates in the following syllogistic argument: on large scales, the Universe is isotropic; since the existence of a privileged location is unlikely, the Universe is certain to be homogeneous. This is the ubiquitous *cosmological principle*, which was given its name in 1931 by Edward Milne and constitutes the foundation of modern cosmology. This claim implies the existence of fundamental or comoving observers at each location, defining a cosmological "rest frame" at each point in space.

Before the development of Einstein's theory of General Relativity in 1915 (Einstein, 1915), the geometry of spacetime and the role of gravity in the large-scale structure of the Universe seemed to have no relation. However, the two became inextricably linked after the physicist's most successful creation was brought to light[3], a gentle introduction to which can be found in Carroll (2003). The current description of the cosmological framework, which rests on the validity of the cosmological principle, entails a spatially homogeneous and isotropic background Universe, valid on large scales, and a theory for structure formation.

General Relativity regards spacetime as a four-dimensional differentiable Lorentzian manifold. The metric tensor, of covariant components $g_{\mu\nu}$, is the mathematical object of interest and is considered a dynamical tensor field whose behavior is governed by Einstein's field equations,

$$G_{\mu\nu} = 8\pi G\, T_{\mu\nu}, \tag{2.1}$$

which relate the Einstein tensor, of covariant components $G_{\mu\nu}$, to the energy-momentum tensor on the manifold, of covariant components $T_{\mu\nu}$. If $\mathcal{M}$ is the manifold describing the large-scale Universe, then $\mathcal{M} = \mathbb{R} \times \Sigma$, where $\mathbb{R}$ describes the temporal direction and Σ is the three-dimensional homogeneous and isotropic spatial manifold. In the jargon of General Relativity, one says that $\mathcal{M}$ can be foliated into spacelike three-dimensional slices that are maximally symmetric[4]. In the 1930s, H. P. Robertson and A. G. Walker independently showed (Robertson, 1935; Walker, 1937) that there could only be three such spacetimes, for which the metric tensor read

$$ds^2 \equiv g_{\mu\nu} dx^\mu dx^\nu = c^2 dt^2 - a^2(t)\,[d\chi^2 + f_K^2(\chi)\,(d\theta^2 + \sin^2\theta\, d\phi^2)\,], \tag{2.2}$$

where

$$f_K(\chi) = \begin{cases} \sin\left(K^{1/2}\chi\right)/\sqrt{K} & \text{if} \quad K > 0 \\ \chi & \text{if} \quad K = 0 \\ \sinh\left((-K)^{1/2}\chi\right)/\sqrt{-K} & \text{if} \quad K < 0 \end{cases} \tag{2.3}$$

and $a(t)$ is called the cosmic scale factor. The metric has been expressed in coordinates $x^\mu = (t, \chi, \theta, \phi)$ such that (χ, θ, ϕ) are spherical coordinates in the three-dimensional space of constant curvature K and t is the cosmic time, that is, the universally agreed time according to which locally-defined physical quantities are uniform and evolve on a three-dimensional hypersurface. This metric describes the geometry of spacetime for the only three possible spatially homogeneous and isotropic Universes, namely open ($K > 0$), flat ($K = 0$) and closed ($K < 0$), and is known as the *Roberston-Walker metric*. It is also common to find this metric written in terms of conformal time, τ, defined via $dt \equiv a(t)d\tau$, in

[3]The reader is encouraged to read Abraham Pais' scientific biography of Einstein to discover the history of how General Relativity came to be (Pais, 1982).

[4]In other words, Σ has the maximum possible number of Killing vectors.

terms of which it reads:

$$ds^2 = a^2(\tau)\left[c^2 d\tau^2 - \left(d\chi^2 + f_K^2(\chi)\,[d\theta^2 + \sin^2\theta d\,\phi^2\,]\right)\right]. \tag{2.4}$$

2.1.2 The Friedmann equations

The time evolution of the scale factor appearing in (2.2) is determined by introducing this metric tensor into Einstein's field equations (2.1) and considering the energy-momentum tensor of a perfect fluid, as constrained by homogeneity and isotropy (Weinberg, 1972), that is,

$$T_{\mu\nu} = \left(\bar{\rho} + \frac{\bar{P}}{c^2}\right)V_\mu V_\nu + \bar{P} g_{\mu\nu}, \tag{2.5}$$

where $\bar{\rho} = \bar{\rho}(t)$ and $\bar{P} = \bar{P}(t)$ are the energy density and pressure in the rest frame of the cosmic fluid and V^μ is its four-velocity relative to a given observer. It should be highlighted that both $\bar{\rho}$ and $\bar{P}$ should be understood as the sum of the contributions of all different components of the cosmic inventory (radiation, matter and dark energy). This results in the two independent differential equations for the scale factor,

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\bar{\rho}(t) - \frac{Kc^2}{a^2} \tag{2.6}$$

and

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\bar{\rho}(t) + \frac{3\bar{P}(t)}{c^2}\right), \tag{2.7}$$

where the overhead dot denotes a derivative with respect to cosmic time, which are known as the *Friedmann equations*[5]. Metrics of the form (2.2) satisfying the above equations are usually referred to as Friedmann-Robertson-Walker metrics.

To solve these equations, we need to know the equation of state for the fluid. Under the assumption of a perfect fluid, it is customary to choose a barotropic[6] equation of the form

$$\bar{P}(t) = w(t)\bar{\rho}(t)c^2, \tag{2.8}$$

where $w(t)$ is called the barotropic index of the fluid. The conservation of the energy-momentum tensor leads to the equation

$$\frac{\dot{\bar{\rho}}}{\bar{\rho}} = -3(1 + w(t))\frac{\dot{a}}{a}, \tag{2.9}$$

which can be integrated if $w(t)$ is assumed to be constant to yield

$$\bar{\rho}(t) = \bar{\rho}(t_0)\left[\frac{a(t)}{a(t_0)}\right]^{-3(1+w)}, \tag{2.10}$$

where t_0 denotes the present time. For the two most popular components of the Universe, namely matter (m) and radiation (r), the barotropic index equals 0 and 1/3, respectively (Mo et al., 2010), resulting in a mean energy density of

$$\bar{\rho}_{\rm m}(t) = \bar{\rho}_{\rm m}(t_0)\left[\frac{a_0}{a(t)}\right]^3 \qquad \bar{\rho}_{\rm r}(t) = \bar{\rho}_{\rm r}(t_0)\left[\frac{a_0}{a(t)}\right]^4, \tag{2.11}$$

[5]Since knowing the dependence of the mean energy density on the scale factor is enough to solve for $a(t)$, the first of these equations is sometimes referred to as *the* Friedmann equation.

[6]A barotropic fluid is one for which the density is only a function of pressure.

where a_0 denotes $a(t_0)$. The dark energy equation of state is unknown, but a popular parametrization of its running barotropic index, which will be used throughout this thesis is given by (Chevallier and Polarski, 2001; Linder, 2003)

$$w_{\rm DE}(t) = w_0 + w_a(1 - a(t)), \tag{2.12}$$

where w_0 and w_a are constants. The resulting dark energy density is

$$\bar{\rho}_{\rm DE}(t) = \bar{\rho}_{\rm DE}(t_0) f(t), \tag{2.13}$$

where

$$f(t) \equiv \left[\frac{a_0}{a(t)}\right]^{3(1+w_0+w_a)} e^{-3w_a(a_0-a(t))}, \tag{2.14}$$

which reduces to the well-known cosmological constant (Λ) term for $w_0 = -1$ and $w_a = 0$, for which $\bar{\rho}_{\rm DE}(t) \equiv \bar{\rho}_\Lambda(t) = \bar{\rho}_\Lambda(t_0)$.

Defining the Hubble parameter as $H(t) \equiv \dot{a}/a$ and introducing the above results into the first Friedmann equation, (2.6) can be cast into the form

$$H^2(t) = H_0^2\left(\Omega_r\left[\frac{a_0}{a(t)}\right]^4 + \Omega_m\left[\frac{a_0}{a(t)}\right]^3 + \Omega_k\left[\frac{a_0}{a(t)}\right]^2 + \Omega_{\rm DE} f(t)\right), \tag{2.15}$$

where we have defined the present-time matter, radiation and dark energy density parameters as

$$\Omega_{\rm i} \equiv \bar{\rho}_{\rm i}(t_0)/\rho_{\rm crit}(t_0), \tag{2.16}$$

where $i \in \{\rm r, m, DE\}$ and the present-time curvature (k) density parameter as

$$\Omega_{\rm k} \equiv -Kc^2/(a_0^2 H_0^2), \tag{2.17}$$

where $H_0 \equiv H(t_0)$ is the Hubble constant[7]. The critical density introduced to define the density parameters is given by $\rho_{\rm crit}(t) \equiv 3H^2(t)/8\pi G$ and is intrinsically linked to the curvature of spacetime, since $\Omega_0 \equiv \Omega_r + \Omega_m + \Omega_{\rm DE}$ is the current total energy density of the Universe in units of $\rho_{\rm crit}(t_0)$ and it must satisfy $\Omega_k = 1 - \Omega_0$. As of this point, and as customary, we consider the scale factor normalized to unity at present time, so that $a_0 \equiv 1$.

As a last comment, it should be mentioned that the matter component is assumed to be made up of dark matter (dm) and baryons (b), although it can include massive neutrinos (ν) at the time their masses become relevant. Likewise, radiation includes photons and those neutrino eigensates that can be considered relativistic at the time of consideration. The current concordance cosmological model, known as ΛCDM, regards dark energy as a cosmological constant, that is, having a non-evolving energy density, and dark matter as cold and pressureless. Simple extensions could imply, for instance, a dark energy density that evolves according to (2.13), namely the w_0CDM and w_0w_aCDM models, characterized by $w_0 \neq 0, w_a = 0$ and $w_0 \neq 0, w_a \neq 0$, respectively. Therefore, to specify the time evolution of the ubiquitous scale factor of a given Friedmann-Robertson-Walker Universe, we need to know the value of the density parameters at a given time (at t_0, for instance), which is the motivation behind the herculean efforts for their observational determination.

[7]Throughout the thesis, the equivalent adimensional Hubble constant, h, will be used, defined as $H_0 \equiv 100h$ km s^{-1}Mpc^{-1}.

2.1.3 Redshift and cosmological distances

The single concept that permeates every section of astrophysics and cosmology is that of redshift. Although it is more common to find a description of this quantity based strictly on purely physical arguments, in my view the following derivation (based on Carroll, 2003) provides a cleaner and more rigorous explanation.

Let $W^\mu = (1, 0, 0, 0)$ denote the four-velocity of a comoving observer. It is straightforward to check that, for Friedmann-Robertson-Walker metrics, the tensor

$$K_{\mu\nu} = a^2(g_{\mu\nu} + W_\mu W_\nu) \tag{2.18}$$

satisfies $\nabla_{(\sigma} K_{\mu\nu)} = 0$ and is thus a Killing tensor[8], which induces a conserved quantity along geodesics, namely

$$K^2 \equiv K_{\mu\nu} V^\mu V^\nu = a^2[V_\mu V^\mu + (W_\mu V^\mu)^2], \tag{2.19}$$

where $V^\mu = dx^\mu/d\lambda$ is the four-velocity of the particle. For a massless particle like a photon, we can always choose the affine parameter λ such that its momentum four-vector p^μ is

$$p^\mu = \frac{dx^\mu}{d\lambda}. \tag{2.20}$$

Since photons satisfy $p^\mu p_\mu = 0$, this implies

$$W_\mu\, p^\mu = \frac{K}{a}. \tag{2.21}$$

Therefore, since a comoving observer with four-velocity W^μ measures the energy of a photon with momentum four-vector p^μ to be

$$E_o \equiv \hbar\omega_o = p_\mu W^\mu = \frac{K}{a}, \tag{2.22}$$

then its observed frequency, ω_o, will deviate from its emission frequency, ω_e, as the Universe expands according to

$$\frac{\omega_o}{\omega_e} = \frac{a_e}{a_o}, \tag{2.23}$$

where a_e and a_o are the scale factors at the time when the photon was emitted and observed, respectively. The fractional change in the wavelength of the photon is quantified through the *cosmological redshift* between the two events,

$$z \equiv \frac{\lambda_o - \lambda_e}{\lambda_e} = \frac{a_o}{a_e} - 1. \tag{2.24}$$

If the observation is made at present time, then we arrive at the fundamental relation

$$a = \frac{1}{1+z}, \tag{2.25}$$

which relates the redshift of an object to the scale factor at the time the observed photon was emitted. It should be noticed that this redshift is not a Doppler shift (although the temptation to be interpreted as such is understandable): it is the change in the metric along the path of the photon that leads to this redshift, not the relative velocities between source and observer[9]. Its interpretation as a Doppler

[8] $\nabla_{(\sigma} K_{\mu\nu)} \equiv (\nabla_\sigma K_{\mu\nu} + \nabla_\sigma K_{\nu\mu} + \nabla_\mu K_{\sigma\nu} + \nabla_\sigma K_{\nu\mu} + \nabla_\nu K_{\sigma\mu} + \nabla_\nu K_{\mu\sigma})/6$, where ∇_σ denotes the covariant derivative with the Levi-Civita connection.

[9] We cannot even speak of relative velocities between two particles at different points in an arbitrary spacetime!

effect is justified as long as we observe galaxies over distances that are not very large, that is, small when compared to the so-called Hubble radius, c/H_0. In this sense, over sufficiently short distances, one can define a recession velocity $v \equiv cz$, which allows redshift to be used as a proxy for distance as we are about to see, although it should chiefly be thought of as a measure of time.

Having defined the concept of redshift, let us now consider the propagation of light through a Friedmann-Robertson-Walker Universe. Given that photons travel along null geodesics and that symmetry arguments demand that these be purely radial (that is, θ and ϕ must be constant), they must satisfy $c^2dt^2 = a^2(t)d\chi^2$. Choosing the location of the observer at $\chi = 0$, this yields $cdt = -a(t)d\chi$, whose physical interpretation is clear: the light we observe today was emitted at an earlier time which is obtained via the integration of the above differential equation.

Since we appear to have several choices to describe the location of a source (its comoving radial coordinate χ, the time t at which the light from it that we observe today was emitted, the scale factor $a(t)$ at that time or the corresponding redshift z), it is tempting to pose oneself the following question: what is *the* distance to the source? The perhaps unsatisfying answer is that there is no answer because the question is wrongly posed. We are biased by the Euclidean idea that there is a uniquely defined distance. When dealing with an arbitrary spacetime, we find ourselves in the face of curvature and, most importantly, the fact that observations measure distances along the backward light cone of the observer and not at a fixed time. Therefore, there exist several inequivalent notions of distance, which we proceed to discuss.

The proper distance at time t from a source at comoving radial coordinate χ to the observer (at $\chi = 0$) is defined as the distance along the corresponding spatial hypersurface ($dt = 0$), so that

$$d_P(t) = a(t)\chi. \tag{2.26}$$

In a sense, the proper distance can be interpreted as the distance that would be measured instantaneously by adding up the measurements from a large number of comoving observers along the path of photon, which clearly makes it completely impractical as a distance measure. Furthermore, since the comoving radial coordinate χ is not observable, we need to find a way to remove it from our equations. From the condition of radial null geodesics, it is clear that a source that emits a photon at time t and is detected at t_0 by the observer in a Friedmann-Robertson-Walker metric is assigned a comoving radial coordinate of

$$\chi(t) = c\int_t^{t_0} \frac{dt}{a(t)} = c\int_{a(t)}^{1} \frac{da}{a^2\,H(a)} = \int_0^{z(t)} \frac{dz}{H(z)}, \tag{2.27}$$

which is called the comoving distance to the source. It is the distance between the worldline of the source and ours on the spatial hypersurface $t = t_0$ and coincides with the proper distance at present, as expected.

Although they will not be explicitly used in this thesis, it is customary to mention the two other distance measures that are usually found in the literature. They share the common property of being defined as the distances we would infer if space was Euclidean and static and are called the luminosity distance and the angular diameter distance, whose expression is given by (Bartelmann and Schneider, 2001)

$$d_A(t) = a(t)f_K[\chi(t)] \qquad d_L(t) = \frac{d_A(t)}{a^2(t)}, \tag{2.28}$$

respectively. One should additionally mention that, for small enough redshifts, all four distance mea-

sures coincide and fulfill the famous Hubble-Lemaître law,

$$v = H_0\, d, \tag{2.29}$$

where d denotes the "distance" in quotes, given the previous discussion on the validity of speaking about a recession velocity.

2.1.4 The growth of structure

The homogeneous and isotropic Friedmann-Robertson-Walker metric only applies to the Universe on sufficiently large scales or at very early times, as indicated by the small temperature anisotropies of the CMB, corresponding to $z \sim 1100$. The structures we observe today, such as galaxies and galaxy clusters, are thought to have formed from small density fluctuations at a very early time via gravitational instability. The origin of these perturbations is still a matter of debate, although the theory of inflation has become increasingly popular over the last few decades as a process that would have amplified adiabatic fluctuations of a quantum nature, usually assumed to be uncorrelated and following a Gaussian distribution.

Let us consider the density contrast of component i of the Universe at a given point in space and time, defined as

$$\delta_i(\boldsymbol{x}, t) \equiv \frac{\rho_i(\boldsymbol{x}, t) - \bar{\rho}_i(t)}{\bar{\rho}_i(t)}, \tag{2.30}$$

where $\bar{\rho}_i(t)$ is the mean density of component i in the Universe at time t and $\boldsymbol{x}$ is the comoving spatial coordinate. In cosmology, this is to be regarded as a random field of which we wish to know its n-point correlation functions. Since the evolution of the density contrast will be strongly dependent on the physical scale, it is customary to work in Fourier space, so that we are interested in the evolution of $\delta_i(\boldsymbol{k}, t)$. A fully rigorous treatment of the perturbative growth of *small* fluctuations requires the tools of General Relativity, because the Newtonian framework cannot be applied to perturbations in relativistic particles or to modes that are not well inside the horizon size[10]. The main idea is to perform small perturbations around the Friedmann-Robertson-Walker metric, which will correspond to perturbations in the energy-momentum tensor via Einstein's field equations resulting in coupled differential equations for the perturbations in each component. Since this is a tedious task well beyond the scope of this thesis, we will only summarize the results that are important to us and refer the reader to Liddle and Lyth (2000), Dodelson (2003) or Mo et al. (2010), for instance, for a detailed description of the relativistic theory of small perturbations. The use (and abuse) of the word "small" here refers to the validity of the perturbative approach. Indeed, as long as $|\delta| \ll 1$, the Fourier modes of the overdensity field remain Gaussian and uncorrelated and we are in the so-called linear regime, for which one can compute the evolution numerically.

Both relativistic and Newtonian perturbation theory show that small cold dark matter perturbations that are well inside the Horizon grow according to

$$\delta_{\mathrm{dm}}(\boldsymbol{k}, t) = D(t)\delta_{\mathrm{dm}}(\boldsymbol{k}, t_0) \tag{2.31}$$

[10]The comoving horizon size of the Universe at time t is defined as the comoving distance a photon can have traveled from the Big Bang by time t, that is, $d_H(t) = c/(a(t)H(t))$. A Fourier mode is said to be inside the horizon at time t if its comoving wavelength satisfies $\lambda < d_H(t)$.

after matter-radiation equality[11], where

$$D(t) \propto H(t) \int_0^t \frac{t'}{a^2(t')H^2(t')} \tag{2.32}$$

is the linear growth factor, normalized to 1 at t_0. This result can be easily obtained within Newtonian perturbation theory via the combination of the Vlasov and Poisson equations. However, it is not possible to frame the evolution of super-horizon modes (for which causality cannot be considered instantaneous) within a non-relativistic theory. Furthermore, the non-negligible contribution of radiation manifests itself before matter-radiation equality via the Mészáros effect (Meszaros, 1974), effectively suppressing growth on small scales. The overall consequences of these effects, along with the back-reaction due to the inclusion of baryons can be accounted for with the introduction of the *transfer function*, defined as

$$T(k) \equiv \frac{D(t_i)}{D(t)} \frac{\delta_{\rm dm}(\boldsymbol{k}, t)}{\delta_{\rm dm}(\boldsymbol{k}, t_i)}, \tag{2.33}$$

where t_i is a very early time when all Fourier modes of interest were still outside the horizon. The computation of the transfer function is usually carried out numerically using Boltzmann codes like CAMB (Lewis et al., 2000) or CLASS (Lesgourgues, 2011) that solve the coupled differential equations for the density contrast in a relativistic framework. There are, however, accurate analytic approximations available (Eisenstein and Hu, 1998, 1999), which we have used throughout the thesis, as represented in Figure 2.1.

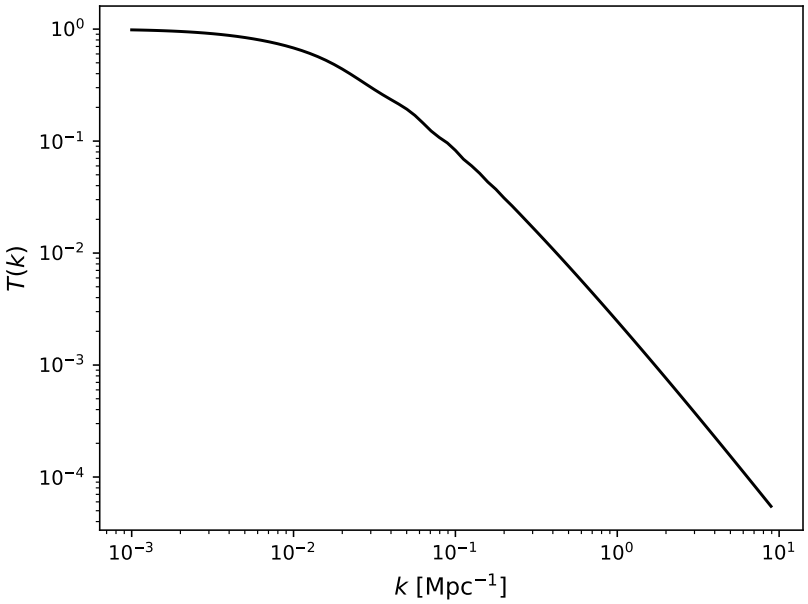

Figure 2.1: Dark matter transfer function for a flat ΛCDM cosmology, according to the fit by Eisenstein and Hu (1998). The cosmological parameters have been fixed to *Planck*'s most recent values (Planck Collaboration et al., 2020a).

With knowledge of the transfer function, the evolution of small dark matter perturbations is given

[11] Matter-radiation equality is defined as the time $t_{\rm eq}$ where $\Omega_m(t_{\rm eq}) = \Omega_r(t_{\rm eq})$, after which the energy density of radiation started its strong decrease.

by

$$\delta_{\rm dm}(\boldsymbol{k}, t) = T(k)\frac{D(t)}{D(t_i)}\delta_{\rm dm}(\boldsymbol{k}, t_i) \tag{2.34}$$

and the linear dark matter *power spectrum*, that is, the Fourier transform of the two-point correlation function of the dark matter density contrast in the linear regime, is thus given by

$$P^{\rm lin}_{\text{dm-dm}}(k, z) = T^2(k)\frac{D^2(z)}{D^2(z_i)}P_{\text{dm-dm}}(k, z_i), \tag{2.35}$$

where $z_i \equiv z(t_i)$.

The missing piece of the puzzle is the primordial (untransferred) power spectrum, $P_{\text{dm-dm}}(k, z_i)$. In the absence of a complete theory for the origin of density fluctuations and following the general prediction of inflationary models, this quantity is assumed to be a power law, that is,

$$P_{\text{dm-dm}}(k, z_i) = \mathcal{N}\, k^{n_s}, \tag{2.36}$$

where n_s is known as the spectral index. Current observational constraints on this parameter (Planck Collaboration et al., 2020a) hint at a value of $n_s = 1$, usually known as the Harrison-Zeldovich or scale-invariant power spectrum. Figure 2.2 shows the theoretical linear dark matter power spectrum at $z = 0$.

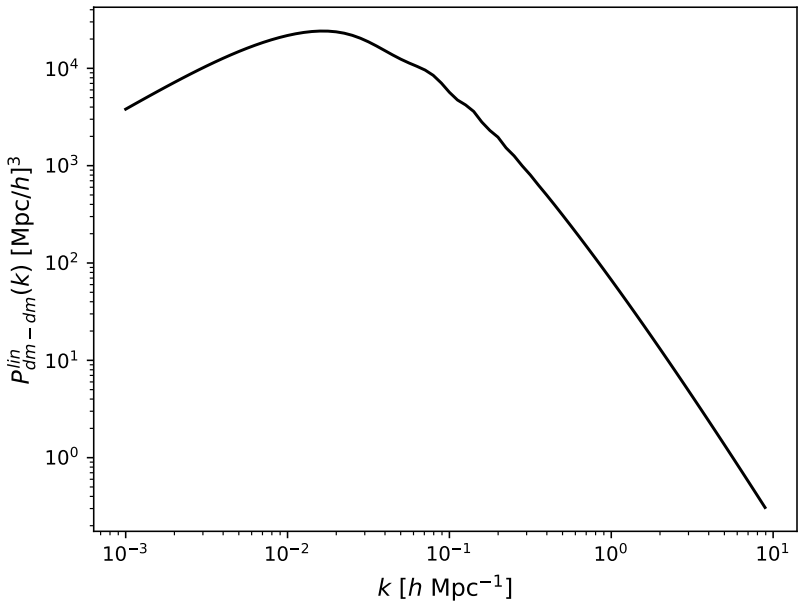

Figure 2.2: Linear dark matter power spectrum at $z = 0$ for a flat ΛCDM cosmology, assuming a Harrison-Zeldovich primordial spectrum and a transfer function modeled through the fit by Eisenstein and Hu (1998). The cosmological parameters have been fixed to *Planck*'s most recent values (Planck Collaboration et al., 2020a).

Before the end of this section, I would like to address the use of the word "small" in the last paragraphs, which might have made the reader feel uneasy. How do we compute the power spectrum when perturbations become large enough? Although there are several perturbative approaches that are valid up to the quasi-linear regime (see Sugiyama, 2014, and references therein) and some fitting formulae (Peacock and Dodds, 1996; Smith et al., 2003; McDonald et al., 2006), the most popular

procedure is resorting to N-body simulations, be it dark matter-only or with the complicated inclusion of baryons, where one can compute the correlation function directly. Apart from all this, there is a fairly successful analytical model to describe the process of structure formation in the non-linear regime that will allow us to extract physical information from our observable: *the halo model*. Before introducing it, let us however discuss the physical phenomenon underlying magnification bias and and how the need for a model of the non-linear power spectrum naturally arises in this context.

2.2 Weak gravitational lensing

Long before the birth of General Relativity, Isaac Newton posed himself a rhetorical question in his *Opticks* (Newton, 1718):

> *"Do not Bodies act upon Light at a distance, and by their action bend its Rays; and is not this action... strongest at the least distance?"*

This quote, albeit a tad obscure to the modern reader, makes it clear that Newton hypothesized that light itself should feel a gravitational pull. He was thus the first one to propose the existence of *gravitational lensing*. Although he never made any attempt at finding an observable proof of this conjecture, the german physicist Johann von Soldner was able to derive a quantitative result regarding the deflection angle of light under Newton's corpuscular theory, certainly well before photons were conceived. Indeed, assuming light was made up of particles with a non-zero (but almost negligible) mass, he showed that Newtonian gravity implied that a particle of light in the gravitational field of a very massive object would be subject to a scattering with a deflection angle of[12]

$$\alpha_{\mathrm{N}} = \frac{2GM}{\xi\, c^2}, \tag{2.37}$$

where G is Newton's gravitational constant, M is the mass of the object creating the gravitational field, c is the speed of light and ξ is the distance of closest approach or impact parameter.

Seemingly unaware of this result, in 1915 Albert Einstein tackled the issue of the bending of light with the powerful machinery of his brand-new theory of gravity. Indeed, the propagation of light rays along null geodesics in a weak gravitational field yielded a result which was twice as large as the "Newtonian" prediction (Einstein, 1915):

$$\alpha_{\mathrm{E}} = \frac{4GM}{\xi\, c^2}. \tag{2.38}$$

At the time, this provided the opportunity to test the validity of the (poorly understood) theory of General Relativity, since a light ray that grazed the surface of the Sun should be deflected by about $\alpha_{\mathrm{E}} \approx 1.75$ arcsec, compared to the $\alpha_{\mathrm{N}} \approx 0.87$ arcsec of the Newtonian counterpart. It was soon realized that a total solar eclipse would provide very suitable conditions, since others stars would be observable projected near the surface of the Sun and a 1919 expedition to Brazil confirmed the larger value in what is sometimes considered as the most important step towards the acceptance of General Relativity (Eddington, 1920).

The phenomenon of gravitational lensing can indeed be interpreted within General Relativity, since the geodesics followed by light rays are no longer straight lines in a spacetime that has been deformed from a flat geometry due to the presence of mass. Sufficiently strong gravitational fields may cause extreme distortions resulting in an observer seeing a multiple image of the same object in the sky. The tidal nature of a gravitational field can additionally modify the shape and size of the cross section of a light bundle, inducing apparent changes to the observer in the lensed image of an object regarding its flux or position. Given the observational character of this physical phenomenon and the dependence of its related observables on the total mass density and cosmological parameters, it comes as no surprise

[12] Soldner's result is usually referred to as *the Newtonian prediction* of the deflection of light for lack of a better name, but it is obviously wrong on a fundamental level. *"What photons do in a Newtonian world is really anyone's guess"* (Coles and Lucchin, 2003).

that it has been extensively exploited for cosmological purposes since the first discovery of a multiple-image system in 1979, known as the twin quasars (Walsh et al., 1979), depicted in Figure 2.3. In fact, the field of gravitational lensing has flourished into one of the most active areas of astrophysics and cosmology as the never-ending literature on the subject can bear witness; see, for instance, Schneider et al. (1992), Bartelmann and Schneider (2001), Petters et al. (2001), Schneider et al. (2006), Ellis (2010), Dodelson (2017) or Cervantes-Cota et al. (2019) for general reviews on the history, theory and applications of gravitational lensing.

Figure 2.3: Image of "QSO B0957+0561", also known as the twin quasars, taken by the Hubble Space Telescope. Credits: ESA/Hubble & NASA.

Nevertheless, it should have dawned on the attentive reader that the title of this section contains an additional adjective. Indeed, in contrast to extreme light distorsions that cause multiple images of single objects or elongated arcs in galaxy clusters (phenomena popularly referred to as related to *strong* lensing), weaker magnifications and distorsions only manifest themselves when one works with large samples of astrophysical objects. *Weak* gravitational lensing deals with the study of these effects in a statistical sense, deriving average properties of the cosmic density field provided that the number density of sources is high enough and/or that one observes a sufficiently large portion of the sky. Let us discover how it works.

2.2.1 The principles of lensing by large-scale structure

A careful and rigorous treatment of the propagation of light beams in the Universe requires the formalism of General Relativity, where the world line of particles in an arbitrary spacetime is described as a four-vector. An appropriate choice of a coordinate system can nonetheless simplify the matter and allow us to effectively describe the separation (transverse to the line of sight) between two infinitesimally close light rays by a two-dimensional vector. The treatment laid out in this subsection will not be fully thorough, since the mathematics involved go beyond the scope of this thesis, but it does aim to shed light on the key ideas. The reader is referred to Schneider et al. (1992), Seitz et al. (1994) and Bartelmann and Schneider (2001) for a more detailed and technical discussion.

Let us consider a fiducial light ray traveling through an arbitrary spacetime. Denoting by $x_0^\mu(\lambda)$ its position four-vector, where λ is a carefully chosen affine parameter, its evolution is dictated by the

geodesic equation (Weinberg, 1972; Misner et al., 1973). If $x^\mu(\boldsymbol{\theta}, \lambda)$ is the position four-vector of a ray of the beam that forms an angle $\boldsymbol{\theta}$ with the fiducial one at the position of a given observer and provided that $|\boldsymbol{\theta}|$ is small, the separation four-vector

$$\xi^\mu(\boldsymbol{\theta}, \lambda) \equiv x^\mu(\boldsymbol{\theta}, \lambda) - x^\mu(\mathbf{0}, \lambda), \tag{2.39}$$

where $x^\mu(\mathbf{0}, \lambda) = x_0^\mu(\lambda)$, can be linearized in $\boldsymbol{\theta}$ and further decomposed (Bartelmann and Schneider, 2001) into

$$\xi^\mu(\boldsymbol{\theta}, \lambda) = -\xi_1(\boldsymbol{\theta}, \lambda) E_1^\mu(\lambda) - \xi_2(\boldsymbol{\theta}, \lambda) E_2^\mu(\lambda) - \xi_0(\boldsymbol{\theta}, \lambda) \tilde{k}^\mu(\lambda), \tag{2.40}$$

where $E_1^\mu(\lambda)$ and $E_2^\mu(\lambda)$ are orthonormal unit four-vectors that span a two-dimensional plane perpendicular to both the four-velocity W_O^μ of the observer and the dimensionless wave vector $\tilde{k}^\mu$ along every event of the ray[13]. Therefore, the two-dimensional vector $\boldsymbol{\xi}(\boldsymbol{\theta}, \lambda)$ (with components $\xi_1(\boldsymbol{\theta}, \lambda)$ and $\xi_2(\boldsymbol{\theta}, \lambda)$) represents the transverse separation of two light rays, while the $\xi_0(\boldsymbol{\theta}, \lambda)$ component accounts for deviations along the direction of the beam. The linearity in $\boldsymbol{\theta}$ and the choice of the affine parameter imply that the transverse separation vector must satisfy the geodesic deviation equation (Bartelmann and Schneider, 2001):

$$\frac{d^2 \boldsymbol{\xi}(\boldsymbol{\theta}, \lambda)}{d\lambda^2} = \mathcal{T}(\lambda)\, \boldsymbol{\xi}(\boldsymbol{\theta}, \lambda), \tag{2.41}$$

where $\boldsymbol{\xi}(\boldsymbol{\theta}, \lambda)$ is regarded as a function of λ and

$$\mathcal{T}(\lambda) = \begin{pmatrix} \mathcal{R}(\lambda) + \mathrm{Re}[\mathcal{F}(\lambda)] & \mathrm{Im}[\mathcal{F}(\lambda)] \\ \mathrm{Im}[\mathcal{F}(\lambda)] & \mathcal{R}(\lambda) - \mathrm{Re}[\mathcal{F}(\lambda)] \end{pmatrix} \tag{2.42}$$

is the optical tidal matrix, which describes how spacetime influences the propagation of light via the quantities $(\mathcal{R})$ and $(\mathcal{F})$, computed from the Ricci and Weyl tensor, respectively.

Since our interest is gravitational lensing by the large-scale structure of the Universe, let us consider the metric tensor of a weakly inhomogeneous expanding Universe,

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = a^2(\tau) \left[\left(1 + \frac{2\Phi}{c^2} \right) c^2 d\tau^2 - \left(1 - \frac{2\Phi}{c^2} (d\chi^2 + f_K^2(\chi) d\Omega) \right) \right], \tag{2.43}$$

which is characterized by the peculiar Newtonian potential, Φ, and is written in terms of conformal time, τ, and comoving radial distance, χ. If the affine parameter is chosen to be $\lambda \equiv \chi$, it proves convenient to define the comoving separation vector

$$\boldsymbol{x}(\boldsymbol{\theta}, \chi) \equiv \boldsymbol{\xi}(\boldsymbol{\theta}, \chi)) / a(\chi) \tag{2.44}$$

between a fiducial ray and a ray that deviates by an angle $\boldsymbol{\theta}$ at the observer. The tidal matrix (2.42) can be explicitly computed for the above metric and the geodesic deviation equation (2.41) describing the evolution of $\boldsymbol{x}(\boldsymbol{\theta}, \chi)$ yields (Bartelmann and Schneider, 2001; Schneider et al., 2006)

$$\frac{d^2 \boldsymbol{x}(\boldsymbol{\theta}, \chi)}{d\chi^2} + K \boldsymbol{x}(\boldsymbol{\theta}, \chi) = -\frac{2}{c^2} \left[\nabla_\perp \Phi(\boldsymbol{x}(\boldsymbol{\theta}, \chi), \chi) - \nabla_\perp \Phi^{(0)}(\chi) \right], \tag{2.45}$$

where K is the curvature of the Universe, $\nabla_\perp$ denotes the transverse comoving gradient operator and $\Phi^{(0)}(\chi)$ is the peculiar Newtonian potential along the fiducial ray. The above equation is further

[13] This can be done via the parallel transport of $\tilde{k}^\mu$, W_O^μ, E_1^μ and E_2^μ from the observer along the curve defined by the fiducial ray (Bartelmann and Schneider, 2001).

equipped with the initial conditions

$$\boldsymbol{x}(\boldsymbol{\theta}, 0) = \mathbf{0} \qquad \frac{d\boldsymbol{x}}{d\chi}(\boldsymbol{\theta}, 0) = \boldsymbol{\theta} \tag{2.46}$$

and the formal solution can be written via a Green function, yielding

$$\boldsymbol{x}(\boldsymbol{\theta}, \chi) = f_K(\chi)\boldsymbol{\theta} - \frac{2}{c^2}\int_0^\chi d\chi'\, f_K(\chi - \chi')\Big[\nabla_\perp \Phi(\boldsymbol{x}(\boldsymbol{\theta}, \chi'), \chi') - \nabla_\perp \Phi^{(0)}(\chi')\Big]. \tag{2.47}$$

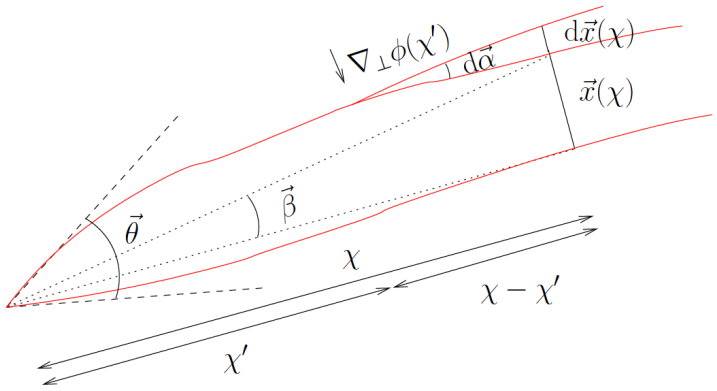

Figure 2.4: Kilbinger (2015). Pictorial representation of the propagation of two light rays (red lines) that converge on the observer under the angle $\boldsymbol{\theta}$. The unperturbed geodesics (dotted lines) are modified by a deflector at comoving distance χ'.

In the absence of lensing, the comoving separation $\boldsymbol{x}$ between two rays at comoving distance χ would be observed under an angle $\boldsymbol{\beta} = \boldsymbol{x}/f_K(\chi)$. This defines the angular separation under which the unperturbed or unlensed transverse comoving separation $\boldsymbol{x}$ is seen and relates it to the apparent angle $\boldsymbol{\theta}$ via the lens equation:

$$\boldsymbol{\beta}(\boldsymbol{\theta}, \chi) = \boldsymbol{\theta} - \frac{2}{c^2}\int_0^\chi d\chi'\, \frac{f_K(\chi - \chi')}{f_K(\chi)}\Big[\nabla_\perp \Phi(\boldsymbol{x}(\boldsymbol{\theta}, \chi'), \chi') - \nabla_\perp \Phi^{(0)}(\chi')\Big]. \tag{2.48}$$

The above equation allows us to compute the deflection of a light ray as it travels to the observer in the presence of density perturbations. Indeed, $\boldsymbol{\beta}(\boldsymbol{\theta}, \chi)$ is the unlensed angular position of a light ray at comoving distance χ that converges on the observer under an angle $\boldsymbol{\theta}$ with respect to an arbitrarily chosen fiducial ray.

Equation (2.48) can be locally linearized in $\boldsymbol{\theta}$ and the Jacobian matrix,

$$\mathcal{A}(\boldsymbol{\theta}, \chi) \equiv \frac{\partial \boldsymbol{\beta}}{\partial \boldsymbol{\theta}}(\boldsymbol{\theta}, \chi), \tag{2.49}$$

is given by

$$\mathcal{A}_{ij}(\boldsymbol{\theta}, \chi) = \delta_{ij} - \frac{2}{c^2}\int_0^\chi d\chi' \frac{f_K(\chi - \chi')f_K(\chi')}{f_K(\chi)} \frac{\partial^2 \Phi}{\partial x_i \partial x_k}(\boldsymbol{x}(\boldsymbol{\theta}, \chi'), \chi')\, \mathcal{A}_{kj}(\boldsymbol{\theta}, \chi'). \tag{2.50}$$

This expression can be tackled via a perturbative approach, expanding the Jacobian matrix in terms of the gravitational potential, $\mathcal{A}(\boldsymbol{\theta}, \chi) = \mathcal{A}^{(0)}(\boldsymbol{\theta}, \chi) + \mathcal{A}^{(1)}(\boldsymbol{\theta}, \chi) + \ldots$ To linear order, we have

$$\mathcal{A}_{ij}(\boldsymbol{\theta}, \chi) = \delta_{ij} - \frac{2}{c^2}\int_0^\chi d\chi' \frac{f_K(\chi - \chi')f_K(\chi')}{f_K(\chi)} \frac{\partial^2 \Phi}{\partial x_i \partial x_j}(f_K(\chi')\boldsymbol{\theta}, \chi')), \tag{2.51}$$

which corresponds to the so-called Born approximation, in which the integration is carried out along the unperturbed ray. As a consequence, defining the deflection or lensing potential as

$$\psi(\boldsymbol{\theta},\chi) \equiv \frac{2}{c^2}\int_0^{\chi} d\chi' \frac{f_K(\chi-\chi')}{f_K(\chi)f_K(\chi')}\Phi(f_K(\chi')\boldsymbol{\theta},\chi'), \tag{2.52}$$

the Jacobi matrix is expressed as $\mathcal{A}_{ij} = \delta_{ij} - \partial^2\psi/\partial x_i \partial x_j$. More precisely, it can be parametrized as

$$\mathcal{A}(\boldsymbol{\theta},\chi) = \begin{pmatrix} 1-\kappa(\boldsymbol{\theta},\chi)-\gamma_1(\boldsymbol{\theta},\chi) & -\gamma_2(\boldsymbol{\theta},\chi) \\ -\gamma_2(\boldsymbol{\theta},\chi) & 1-\kappa(\boldsymbol{\theta},\chi)+\gamma_1(\boldsymbol{\theta},\chi) \end{pmatrix} \tag{2.53}$$

in terms of the *convergence* and *shear*, respectively:

$$\kappa(\chi,\boldsymbol{\theta}) = \frac{1}{2}\left[\frac{\partial^2}{\partial\theta_1^2}+\frac{\partial^2}{\partial\theta_2^2}\right]\psi(\chi,\boldsymbol{\theta}) \qquad \gamma(\chi,\boldsymbol{\theta}) = \frac{1}{2}\left[\frac{\partial^2}{\partial\theta_1^2}-\frac{\partial^2}{\partial\theta_2^2}\right]\psi(\chi,\boldsymbol{\theta}) + i\frac{\partial^2\psi(\chi,\boldsymbol{\theta})}{\partial\theta_1\partial\theta_2}, \tag{2.54}$$

the latter usually expressed as a complex scalar field. The Jacobian matrix can be rewritten as

$$\mathcal{A}(\boldsymbol{\theta},\chi) = (1-\kappa(\boldsymbol{\theta},\chi))\begin{pmatrix}1&0\\0&1\end{pmatrix} - \begin{pmatrix}\gamma_1(\boldsymbol{\theta},\chi) & \gamma_2(\boldsymbol{\theta},\chi)\\ \gamma_2(\boldsymbol{\theta},\chi) & -\gamma_1(\boldsymbol{\theta},\chi)\end{pmatrix}, \tag{2.55}$$

which sheds light on the physical meaning of convergence and shear. The convergence field, κ, induces an isotropic focusing, causing lensed images to only be rescaled by a constant factor in all directions. The shear field, γ, on the other hand, induces a shape distortion and a anisotropic focusing, streching the shape of the source along a certain direction. Figure 2.5 shows a pictorial representation of these two effects on a circular source.

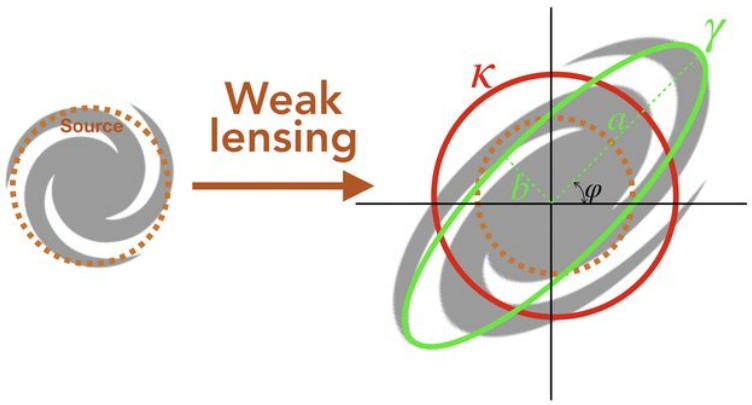

Figure 2.5: (Shuntov, 2019). Schematic representation of the effect of convergence (κ) and shear (γ) on a circular source.

Comparing these results from the standard ones from the conceptually simpler gravitational lens theory (Bartelmann and Schneider, 2001; Schneider et al., 2006), lensing by three-dimensional large-scale structure can be effectively regarded as if we had a lens plane at a comoving distance χ from the observer with surface mass density $\kappa = (\psi_{,11}+\psi_{,22})/2$ and shear $\gamma = (\psi_{,11}-\psi_{,22})/2 + i\psi_{,12}$, where $\psi_{,ij} \equiv \partial^2\psi/\partial\theta_i\partial\theta_j$.

It should therefore be noticed that the convergence and the lensing potential are related through a two-dimensional Poisson equation. To introduce the three-dimensional density contrast, one can artificially add the term $\partial^2/\partial\chi^2$ to this two-dimensional Laplacian, since it vanishes when the integration

along the line of sight is carried out (Bartelmann and Schneider, 2001). Consequently,

$$\kappa(\boldsymbol{\theta},\chi) = \frac{1}{c^2}\int_0^{\chi} d\chi' \frac{f_K(\chi-\chi')}{f_K(\chi)f_K(\chi')}\nabla^2\Phi(f_K(\chi')\boldsymbol{\theta},\chi'), \tag{2.56}$$

where $\nabla^2 \equiv \partial^2/\partial\theta_1^2 + \partial^2/\partial\theta_2^2 + \partial^2/\partial\chi^2$. Using the three-dimensional Poisson equation in comoving coordinates,

$$\nabla_\chi^2\Phi = 4\pi G a^2 \bar{\rho}\,\delta_{\rm m}, \tag{2.57}$$

we have

$$\kappa(\boldsymbol{\theta},\chi) = \frac{3}{2}\frac{H_0^2}{c^2}\int_0^{\chi} d\chi' \frac{f_K(\chi')f_K(\chi-\chi')}{f_K(\chi)}\delta_{\rm m}(f_K(\chi')\boldsymbol{\theta},\chi')E^2(\chi')\,a^2(\chi'), \tag{2.58}$$

where we have defined

$$E(z) \equiv \frac{H(z)}{H_0}. \tag{2.59}$$

Equation (2.58) gives the convergence or *effective* surface mass density for a source at comoving distance χ as an integral over the density contrast along the unperturbed light path to the observer weighted by a series of factors related to the comoving distance and scale factor. Since it corresponds to a source at a fixed redshift z or comoving distance χ, it needs to be modified for a redshift distribution of sources with $p_z(z)dz = p_\chi d\chi$, so that

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_h} d\chi\, p_\chi(\chi)\kappa(\boldsymbol{\theta},\chi) = \frac{3}{2}\frac{H_0^2}{c^2}\int_0^{\chi_h} d\chi W_\chi(\chi)\, f_K(\chi)\,\delta_{\rm m}(f_K(\chi)\boldsymbol{\theta},\chi)\,E^2(\chi)\,a^2(\chi), \tag{2.60}$$

where

$$W_\chi(\chi) \equiv \int_\chi^{\chi_h} d\chi' p_\chi(\chi')\frac{f_K(\chi'-\chi)}{f_K(\chi)} \tag{2.61}$$

is called the lens efficiency factor and $\chi_h \equiv \lim_{z\to\infty}\chi(z)$. This expression represents the mean effective surface mass density for a population of source galaxies with comoving distance distribution $p_\chi(\chi)$.

The convergence field is everything we need for our purposes. As commented at the beginning of this section, weak gravitational lensing is characterized by small distortions in shape and size, which mathematically translates into the Jacobi matrix being close to the identity. As we will see next, this will allow us to only take the convergence field into account as far as our observable is concerned.

2.2.2 Magnification bias

In the absence of emission, absorption or scattering, the phase-space density of photons is conserved during gravitational lensing as a consequence of Liouville's theorem (Liouville, 1838; Gibbs, 1902) of Hamiltonian optics. This implies directly that surface brightness, that is, the flux density per unit solid angle, must be conserved along the path of a photon. In other words, the *observed* surface brightness at an angular position $\boldsymbol{\theta}$ in the sky must equal the *true* surface brightness at the unlensed angular position $\boldsymbol{\beta}(\boldsymbol{\theta},\chi)$ of the photon.

Taking this into account, let us consider the weak lensing by large-scale structure of the light coming from a background distribution of sources. We define the *magnification* associated with a lensed image at $\boldsymbol{\theta}$ of a sufficiently small source at comoving distance χ observed at an angle $\boldsymbol{\beta}$ as the ratio of the image flux to that of the unlensed source, which is thus given by

$$\mu(\boldsymbol{\theta},\chi) \equiv \frac{S_{\boldsymbol{\theta}}}{S_{\boldsymbol{\beta}}} = \frac{d\Omega_{\boldsymbol{\theta}}}{d\Omega_{\boldsymbol{\beta}}} = \frac{1}{|\mathcal{A}(\boldsymbol{\theta},\chi)|} = \frac{1}{(1-\kappa(\boldsymbol{\theta},\chi))^2 - |\gamma(\boldsymbol{\theta},\chi)|^2}. \tag{2.62}$$

The above equation once again highlights the fact that lensed images are distorted in shape and size and that magnification increases the flux of sources by the same amount it enlarges the solid angle they subtend. However, if a flux threshold is introduced in a source sample, it can effectively create a mismatch between the two effects and introduce an observational *magnification bias*. Indeed, either there are enough faint sources that go over the flux threshold to compensate the decrease in surface density or there are too few, which causes the enlargement of the solid angle to dominate. The former case, represented in Figure 2.6, constitutes the cornerstone of this thesis and will be described below.

Let $n_{\mathrm{b}_0}(S, z)$ be the intrinsic differential number counts of a background galaxy population, that is, the number of sources per unit solid angle, redshift and flux in the absence of lensing, that is,

$$n_{\mathrm{b}_0}(S, z) \equiv \frac{dN_{\mathrm{b}_0}}{d\Omega\, dz\, dS}. \tag{2.63}$$

Equation (2.62) implies that the background differential number counts at a point $\boldsymbol{\theta}$ with magnification $\mu(\boldsymbol{\theta}, z)$ are modified to

$$n_{\mathrm{b}}(S, z; \boldsymbol{\theta}) = \frac{1}{\mu^2(\boldsymbol{\theta}, z)} n_{\mathrm{b}_0}\left(\frac{S}{\mu(\boldsymbol{\theta}, z)}, z\right) \tag{2.64}$$

in the presence of lensing.

If we introduce a flux threshold in the background sample, the integrated or cumulative number counts are also changed due to magnification bias, yielding

$$n_{\mathrm{b}}(> S, z; \boldsymbol{\theta}) \equiv \int_S^\infty n_{\mathrm{b}}(\tilde{S}, z; \boldsymbol{\theta})\, d\tilde{S} = \frac{1}{\mu(\boldsymbol{\theta}, z)} n_{\mathrm{b}_0}\left(> \frac{S}{\mu(\boldsymbol{\theta}, z)}, z\right). \tag{2.65}$$

Equation (2.65) highlights the two competing effects of magnification bias, namely the dilution that decreases the number density and the lowering of the effective flux threshold that allows the detection of sources that are fainter than S.

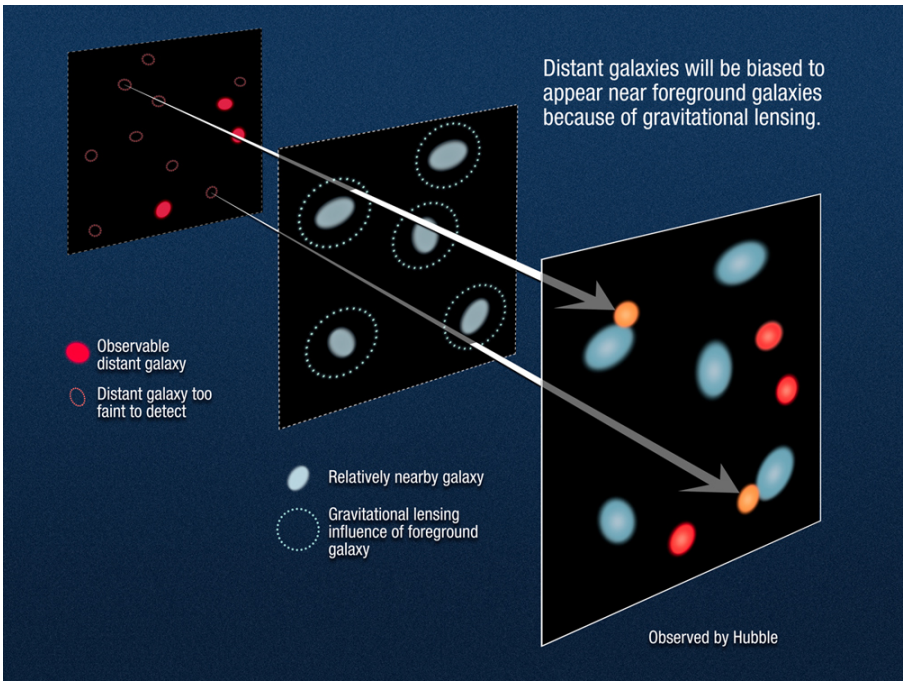

Figure 2.6: Schematic representation of the phenomenon of magnification bias. An excess of background galaxies, which would otherwise be too faint to detect, can be seen around foreground galaxies. Credits: NASA, ESA and A. Feild.

Let us see what this implies in a simple but useful case; assume that the intrinsic background integral number counts follow a redshift-independent power law of the form

$$n_{\mathrm{b}_0}(>S,z) = AS^{-\beta}, \tag{2.66}$$

Therefore, the ratio of lensed to unlensed integral number counts is given by

$$\frac{n_{\mathrm{b}}(>S,z;\boldsymbol{\theta})}{n_{\mathrm{b}_0}(>S,z)} = \mu^{\beta-1}(\boldsymbol{\theta},z), \tag{2.67}$$

Since we will be working with redshift-integrated quantities, the fluctuation in the number density of background sources at an angular position $\boldsymbol{\theta}$ due to magnification bias in a flux-limited sample is given by

$$\delta n^{\mu}_{\mathrm{b}}(\boldsymbol{\theta}) \equiv \frac{n_{\mathrm{b}}(>S;\boldsymbol{\theta}) - n_{\mathrm{b}_0}(>S)}{n_{\mathrm{b}_0}(>S)} = \int_0^{\infty} dz\, p_{z_{\mathrm{b}}}(z)\mu^{\beta-1}(\boldsymbol{\theta},z) - 1 \equiv \mu^{\beta-1}(\boldsymbol{\theta}) - 1, \tag{2.68}$$

where $p_{z_{\mathrm{b}}}$ is the unit-normalized redshift distribution of the background population. The above equation emphasizes the importance of the steepness of the intrinsic number counts of the background population. Indeed, for $\beta = 1$, there is no observable effect, but values of $\beta > 1$ (< 1) imply an increase (decrease) in the number density of objects with respect to the case without lenses. As we will see in the next subsection, this effect can be quantified through the foreground-background cross-correlation function.

Two comments should nonetheless be made before going further. Firstly, since we are dealing with weak lensing, the fact that the Jacobi matrix is close to the identity implies that $|\mu - 1| \ll 1$, which justifies the use of a power law (2.66) to describe the unlensed source counts, since we are only probing them over a small range in flux. Secondly, since both κ and $|\gamma|$ are much smaller than unity in the weak lensing limit, we can approximate the fluctuations in the number density of background sources due to magnification bias by

$$\delta n^{\mu}_{\mathrm{b}}(\boldsymbol{\theta}) \approx 2(\beta-1)\,\kappa(\boldsymbol{\theta}) \tag{2.69}$$

to first order in the perturbations via a Taylor expansion.

2.2.3 Probing the galaxy-matter correlation

The phenomenon of weak lensing has been extensively exploited for probing the large-scale structure by means of the technique of *cosmic shear*. This effect consists in the statistical allignment of the orientation of galaxy images, which is measured as a correlation between their shapes. Since the first detections of this effect at the beginning of the millenium (Bacon et al., 2000; Van Waerbeke et al., 2000; Wittman et al., 2000; Rhodes et al., 2001), cosmic shear has benefited from a considerable refinement of its methodology to analyze the results from recent weak lensing surveys such as the Dark Energy Survey (DES; Dark Energy Survey Collaboration et al., 2016), the Kilo-Degree Survey (KiDS; de Jong et al., 2013) and the Hyper Suprime-Cam survey (HSC; Aihara et al., 2018) and to prepare for the corresponding data resulting from future, more precise surveys such as Euclid (Laureijs et al., 2011), LSST (LSST Science Collaboration et al., 2009) or WFIRST (Spergel et al., 2015). An excellent review on cosmic shear is provided by Kilbinger (2015) and recent results on the matter can be found in Hildebrandt et al. (2017), Troxel et al. (2018), Hikage et al. (2019), Hildebrandt et al. (2020) and Amon et al. (2022).

Contrary to cosmic shear, which only correlates projected ellipticities of foreground and background

galaxies, shear can be exploited through galaxy-galaxy lensing, which correlates the distortion in the shape of background galaxies with the positions of foreground galaxies. This effect, usually quantified through the mean tangential shear as first found by Brainerd et al. (1996), probes the cross-correlation between galaxies and mass, singling out the mass distribution along the line of sight to background sources, assumed to be traced by the foreground galaxies. Recent work on galaxy-galaxy lensing includes Mandelbaum et al. (2013), Baxter et al. (2016a), Baxter et al. (2016b), Joudaki et al. (2018), van Uitert et al. (2018), Prat et al. (2022) and references therein.

However, the cross-power spectrum between galaxies and mass can also be studied through correlating the positions of foreground and background galaxies. The fluctuations in the number of background sources surrounding a sample of foreground sources (with respect to the case without lensing) probes the mass density field traced by the foreground sources. Indeed, as discussed in the introduction, the observed overdensity of bright high-redshift quasars around low-redshift galaxies cannot be explained by the lensing of individual galaxies. However, if galaxies trace the underlying (dark) matter distribution, then the observed correlation can be clearly interpreted as due to the magnification of the large-scale matter distribution. Let us see how magnification bias, central to this thesis, induces this cross-correlation.

The angular cross-correlation between the (redshift-integrated) number density of foreground and background sources is defined as

$$w(\boldsymbol{\theta}) \equiv \langle \delta n_{\rm f}(\boldsymbol{\phi})\, \delta n_{\rm b}(\boldsymbol{\phi}+\boldsymbol{\theta})\rangle_{\boldsymbol{\phi}}, \tag{2.70}$$

where each of the terms describing the fluctuations in the number counts can be divided into the contributions of pure clustering (c) and magnification bias (μ), so that

$$w(\boldsymbol{\theta}) = \langle \delta n^c_{\rm f}(\boldsymbol{\phi})\, \delta n^c_{\rm b}(\boldsymbol{\phi}+\boldsymbol{\theta})\rangle_{\boldsymbol{\phi}} + \langle \delta n^c_{\rm f}(\boldsymbol{\phi})\, \delta n^\mu_{\rm b}(\boldsymbol{\phi}+\boldsymbol{\theta})\rangle_{\boldsymbol{\phi}} + \tag{2.71}$$

$$+ \langle \delta n^\mu_{\rm f}(\boldsymbol{\phi})\, \delta n^c_{\rm b}(\boldsymbol{\phi}+\boldsymbol{\theta})\rangle_{\boldsymbol{\phi}} + \langle \delta n^\mu_{\rm f}(\boldsymbol{\phi})\, \delta n^\mu_{\rm b}(\boldsymbol{\phi}+\boldsymbol{\theta})\rangle_{\boldsymbol{\phi}}. \tag{2.72}$$

The first term describes the cross-correlation due to the intrinsic clustering of both populations, which vanishes if the distributions do not overlap, as is the case throughout this thesis. The second term is due to the magnification of background sources by the foreground sample and is precisely the phenomenon we are interested in. The third term describes the lensing of foreground sources due to the large-scale structure traced by the background, which again is not present since there is no overlap in the redshift distributions. The fourth term is due to magnification by the large-scale structure in front of both the foreground and background sources, which can be considered negligible (Lapi et al., 2012), especially with respect to the main contribution. Therefore, the angular foreground-background cross-correlation is given by

$$w(\boldsymbol{\theta}) = \langle \delta n^c_{\rm f}(\boldsymbol{\phi})\, \delta n^\mu_{\rm b}(\boldsymbol{\phi}+\boldsymbol{\theta})\rangle_{\boldsymbol{\phi}} \tag{2.73}$$

for the entirety of this thesis.

The fluctuations in the number density of foreground sources are thus purely due to clustering and are computed as follows. If the number density of galaxies above a given flux threshold and at a certain comoving distance χ is given by $n(\mathbf{x},\chi)$, then the projected number density of foreground sources at angular position $\boldsymbol{\phi}$ is given by

$$n_{\rm f}(\boldsymbol{\phi}) = \int d\chi\, S_{\rm f}(\chi)\, n(f_K(\chi)\boldsymbol{\phi},\chi), \tag{2.74}$$

where $S_{\rm f}(\chi)$ is the foreground sample selection function. Since the mean number density of these galaxies in the sky is

$$\bar{n}_{\rm f} = \int d\chi\, S_{\rm f}(\chi) \int d\boldsymbol{\phi}\, n(f_K \boldsymbol{\phi}, \chi), \tag{2.75}$$

then

$$\delta n^c_{\rm f}(\boldsymbol{\phi}) \equiv \frac{n_{\rm f}(\boldsymbol{\phi}) - \bar{n}_{\rm f}}{\bar{n}_{\rm f}} = \int d\chi\, p_{\chi_{\rm f}}(\chi)\, \delta_{\rm g}(f_K(\chi)\boldsymbol{\phi}, \chi), \tag{2.76}$$

where $p_{\chi_{\rm f}}(\chi) \equiv S_{\rm f}(\chi)\bar{n}_{\rm f}(\chi)/\bar{n}_{\rm f}$ is the (unit-normalized) comoving distance distribution of foreground galaxies, $\bar{n}_{\rm f}(\chi)$ is the mean number density of foreground galaxies at comoving distance χ and $\delta_{\rm g}$ is the galaxy density contrast.

As we saw in the previous subsection, the fluctuations in the number density of background sources due to magnification bias is given by

$$\delta n^{\mu}_{\rm b}(\boldsymbol{\phi}) = \mu^{\beta-1}(\boldsymbol{\phi}) - 1 \approx 2(\beta - 1)\kappa(\boldsymbol{\theta}), \tag{2.77}$$

an approximation which holds in the weak lensing regime.

The foreground-background angular cross-correlation function is thus expressed in terms of two two-dimensional fields, since

$$w(\boldsymbol{\theta}) = 2(\beta - 1)\, \langle \delta n^c_{\rm f}(\boldsymbol{\phi} + \boldsymbol{\theta})\, \kappa(\boldsymbol{\phi}) \rangle_{\boldsymbol{\phi}}. \tag{2.78}$$

It should be noticed that $\delta n^c_{\rm f}$ is a two-dimensional projection of the galaxy overdensity field, δ_g (2.76), and κ is a two dimensional projection of the (dark) matter[14] overdensity field, $\delta_{\rm dm}$ (2.60). In this scenario, we can make use of Limber's approximation (Limber, 1953) to write the above expression in terms of the three-dimensional dark matter-galaxy cross-power spectrum. Indeed, if $g(f_K(\chi)\boldsymbol{\theta}, \chi)$ and $h(f_K(\chi)\boldsymbol{\theta}, \chi)$ are homogeneous and isotropic three-dimensional random fields, the projections

$$\bar{g}(\boldsymbol{\theta}) \equiv \int d\chi\, q_g(\chi)\, g(f_K(\chi)\boldsymbol{\theta}, \chi) \qquad \bar{h}(\boldsymbol{\theta})) \equiv \int d\chi\, q_h(\chi)\, h(f_K(\chi\boldsymbol{\theta}, \chi), \tag{2.79}$$

where $q_g(\chi)$ and $q_h(\chi)$ are certain weight functions, are two-dimensional homogeneous and isotropic random fields and Limber's equation states that the two-dimensional cross-correlation function is given by (Bartelmann and Schneider, 2001; Schneider et al., 2006)

$$\langle \bar{g}(\boldsymbol{\phi} + \boldsymbol{\theta})\, \bar{h}(\boldsymbol{\phi}) \rangle = \int_0^{\chi_h} d\chi\, q_g(\chi)\, q_h(\chi) \int_0^{\infty} \frac{dk}{2\pi}\, k\, P_{gh}(k, \chi)\, J_0(f_K(\chi)\theta\, k), \tag{2.80}$$

where P_{gh} denotes the cross-power spectrum of the two three-dimensional fields and J_0 is the zeroth-order Bessel function of the first kind. It should be noted that this cross-correlation can only depend on $\theta \equiv |\boldsymbol{\theta}|$.

By identifying $\bar{g} \equiv \delta n^c_{\rm f}$ and $\bar{h} \equiv \kappa$, the corresponding weight functions are

$$q_g(\chi) = p_{\chi_{\rm f}}(\chi) \qquad q_h(\chi) = \frac{3H_0^2}{2c^2} W_\chi(\chi) f_K(\chi) E^2(\chi) a^2(\chi) \tag{2.81}$$

and, therefore, the foreground-background cross-correlation function under the Limber approximation

[14] Given the redshifts of interest in this thesis, the power spectrum of dark matter and that of dark matter plus baryons differ by less than 2% (Angulo et al., 2013).

reads (Cooray and Sheth, 2002)

$$w(\theta) = 2(\beta - 1) \int_0^{\chi_H} d\chi \, p_{\chi_\mathrm{f}}(\chi) W_\chi^{\mathrm{lens}}(\chi) \int_0^\infty \frac{dk}{2\pi} \, k \, P_{\text{g-dm}}(k, \chi) \, J_0(f_K(\chi)\theta \, k), \tag{2.82}$$

where

$$W_\chi^{\mathrm{lens}}(\chi) = \frac{3}{2} \frac{H_0^2}{c^2} E^2(\chi) a^2(\chi) \int_\chi^{\chi_H} d\chi' p_{\chi_\mathrm{b}}(\chi') \frac{f_K(\chi) f_K(\chi' - \chi)}{f_K(\chi')}. \tag{2.83}$$

Noting that the redshift and comoving distance galaxy distributions are related via $p_z(z)dz = p_\chi(\chi)d\chi$, we can express the integrals in terms of redshift and the final result is

$$w(\theta) = 2(\beta - 1) \int_0^\infty \frac{dz}{f_K^2(\chi(z))} \, p_{z_\mathrm{f}}(z) W_z^{\mathrm{lens}}(z) \int_0^\infty \frac{dl}{2\pi} \, l \, P_{\text{g-dm}}\left(\frac{l}{f_K(\chi(z))}, \chi(z) \right) J_0(l\theta), \tag{2.84}$$

where

$$W_z^{\mathrm{lens}}(z) = \frac{3}{2} \frac{H_0^2}{c^2} \left[\frac{E(z)}{1+z} \right]^2 \int_z^\infty dz' p_{z_\mathrm{b}}(z') \frac{f_K(\chi(z')) f_K(\chi(z') - \chi(z'))}{f_K(\chi(z'))} \tag{2.85}$$

and, for each z, we have defined $l \equiv k f_K(\chi(z))$.

For a given cosmological model (and assuming the knowledge of the redshift distribution of foreground and background sources), all quantities in the above expressions are computable (semi-)analytically except for the non-linear galaxy-dark matter cross-power spectrum. The following subsection addresses this issue by introducing the halo model of large-scale structure.

2.3 The halo model

As we have emphasized, the description of the non-linear clustering of matter cannot be fully addressed through perturbation theory. Increasingly complex N-body simulations have been carried out in the last decades (see Vogelsberger et al., 2020, for an overview of the most recent cosmological simulations, with special focus on the methodology) in order to scrutinize the corresponding n-point correlation functions of the cosmic density field. The theoretical approach adopted in this thesis to describe the non-linear power spectrum is nonetheless based on a model that is purely analytical and that provides a way to easily study the variation of our observable with respect to cosmological and astrophysical quantities of interest.

This framework, known as the *halo model*, has its origins in a fairly old paper by Jerzy Newman and Elizabeth Scott (Neyman and Scott, 1952), in which a theory for the spatial distribution of galaxies was built based on realizations of a homogeneous random point process. The Neyman-Scott process, as it is known, describes clustering in two steps: firstly, a homogeneous Poisson process distributes a set of parent points (the cluster centers in their original paper); secondly, a certain density function generates a cloud of points scattered around each of these parent points. The halo model results from extending the idea behind the Neyman-Scott process to the commonly accepted theory that galaxies form within dark matter halos (White and Rees, 1978). Indeed, the theory by Neyman and Scott was entirely written in real coordinate space, dealt only with discrete statistics and did not account for realistic cluster-cluster correlations. The true (continuous, more realistic and based on Fourier space) halo model was first described by Scherrer and Bertschinger (Scherrer and Bertschinger, 1991), followed by the works of Seljak (2000), Ma and Fry (2000), Peacock and Smith (2000) and Scoccimarro et al. (2000). The most important aspects treated in these papers were gathered in the extensive and well-known review by A. Cooray and R. K. Sheth Cooray and Sheth (2002).

The idea, much as that of Neyman and Scott, rests on the assumption that all mass in the Universe is bound up into distinct units (see Figure 2.7), which we call halos, and one can think of spatial statistics again in two steps provided that the size of halos is small compared to the distance between them. The correlations in the mass density field on small scales are thus dominated by the spatial distribution of matter within halos, whereas on large scales it is only the spatial distribution of the halos themselves that is important. Since the beginning of the century, this analytical approach to the (dark) matter density distribution has been exploited in a vast number of works, a non-exhaustive list

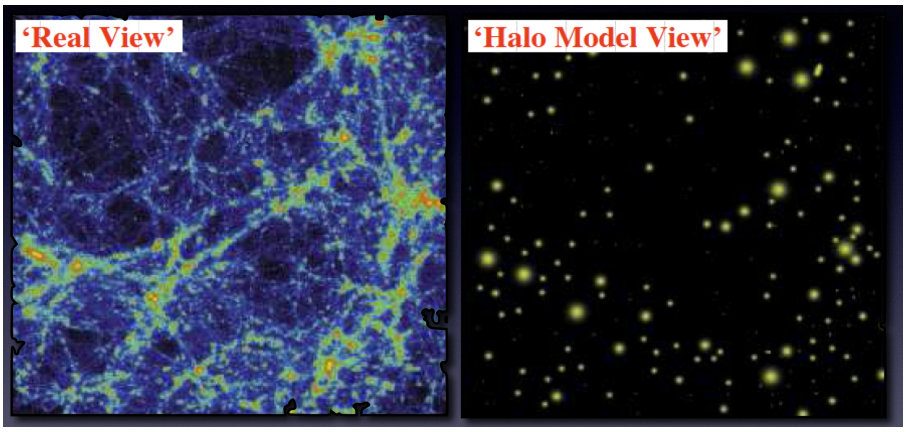

Figure 2.7: van den Bosch, F. (2020). Halo model picture of the dark matter distribution. Retrieved from `http://www.astro.yale.edu/vdbosch/astro610_lecture13.pdf`

of which can be found in Murray et al. (2021) and references therein.

The halo model allows us to analyze the n-point correlation functions of a variety of physical quantities, like the dark matter density field, the galaxy number density or the velocity. It is thus crucial to agree upon a suitable definition and to gather everything we know about the spatial distribution, abundance and density profile of halos, given that they are the central piece of this approach. Within the standard paradigm of galaxy formation, structure grows in a hierarchical process from small density inhomogeneities that initially behave as a Gaussian random field. The very early decoupling of dark matter from the primeval plasma led it to start clustering and form halos through gravitational instability long before big bang nucleosynthesis. Indeed, the high temperature of baryonic matter at the time prevented it from forming gravitationally self-bound objects due to radiation pressure until the period of recombination. Only when the thermal bath cools down sufficiently can baryons decouple from photons and fall into dark matter potential wells. Stars and galaxies thus form within dark matter halos, which merge to form ever more massive ones in a hierarchical process (Lacey and Cole, 1993, 1994; Tormen et al., 2003). Within this picture, there is thus no question about the relevance of dark matter halos for the probing of large-scale structure and cosmology and this realization unavoidably leads to the search for a quantitative understanding of their statistical properties.

2.3.1 The halo mass function

The first to turn their attention to a theoretical model for the abundance of collapsed massive objects were William H. Press and Paul Schechter in a seminal paper (Press and Schechter, 1974) that set the conceptual foundations of a crucial framework. Indeed, despite its simplicity and shortcomings, the underlying ideas put forward by the Press-Schechter model would remain unaltered in all other analytic theories that would follow up years later; this is the reason why it will be discussed in detail in this subsection.

According to linear perturbation theory, as described in 2.1.4, the dark matter overdensity field evolves as

$$\delta_{\rm dm}(\boldsymbol{x}, z) = \delta_{\rm dm}(\boldsymbol{x}, 0) D(z) \tag{2.86}$$

for sub-horizon scales after matter-radiation equality. Let us denote by $\delta_0(\boldsymbol{x}) \equiv \delta_{\rm dm}(\boldsymbol{x}, 0)$ the overdensity field linearly extrapolated to the present. Obviously, when $\delta_{\rm dm}(\boldsymbol{x}, z)$ is Fourier-transformed, modes with a large enough wavenumber will substantially differ from the linearly extrapolated value. However, the so-called spherical collapse model (Gunn and Gott, 1972) allows us to identify, at a given time, regions where the linearly-extrapolated field crosses a threshold value with regions that have formed a virialized object by that time. Indeed, regions where $\delta_0(\boldsymbol{x}) > \hat{\delta}_c(z)/D(z) \equiv \delta_c(z)$ will have collapsed by redshift z to form virialized objects, where $\hat{\delta}_c(z)$ is the collapse threshold, for which we adopt the fit by Kitayama and Suto (1996):

$$\hat{\delta}_c(z) = \frac{3}{20}(12\pi)^{2/3}(1 + 0.0123 \log_{10} \Omega_m(z)). \tag{2.87}$$

The problem now is how to assign a mass to such regions. Since an object of a certain mass M forms from a sufficiently overdense region in the initial density field, one is naturally drawn to assume that such a region is linked to a peak in the linear density field after smoothing it with a window function of a characteristic scale $R \propto M^{1/3}$. However, this simple association poses the so-called cloud-in-cloud problem, since the same mass element could be identified with peaks in the overdensity field smoothed

using two different length scales, and some of them should thus be excluded.

The Press-Schechter theory is a simple approach to perform such an identification. The initial idea is again to consider the smoothed overdensity field linearly extrapoled to the present,

$$\delta_0(\boldsymbol{x}; R) = \int \delta_0(\boldsymbol{r})\, W(\boldsymbol{x}+\boldsymbol{r}; R)\, d^3\boldsymbol{r}, \tag{2.88}$$

where $W(\boldsymbol{x}; R)$ is the window function of characteristic scale R corresponding to a mass

$$M = 4\pi/3\, \bar{\rho}_{\rm m}(t_0) R^3 \tag{2.89}$$

for a spherical top-hat filter. It should be emphasized once more that cosmology regards this overdensity field as a random field, which we shall denote by δ_0^R. The crucial step is now to *postulate* that the probability that $\delta_0^R > \delta_c(z)$ at an arbitrary position $\boldsymbol{x}$ equals the mass fraction that at redshift z is contained within collapsed objects (halos) of mass greater than M. Assuming that δ_0 is a Gaussian random field, δ_0^R is too and, therefore,

$$\mathcal{P}(\delta_0^R > \delta_c(z)) = \frac{1}{\sqrt{2\pi\sigma^2(M)}} \int_{\delta_c(z)}^{\infty} \exp\left[-\left(\frac{(\delta_0^R)}{\sqrt{2}\sigma(M)} \right) \right] d\delta_0^R = \frac{1}{2}\,\mathrm{erfc}\left[\frac{\delta_c(z)}{\sqrt{2}\,\sigma(M)} \right], \tag{2.90}$$

where $\mathrm{erfc}(x)$ is the complementary error function and

$$\sigma^2(M) = \langle \delta_0^2(\boldsymbol{x}; R) \rangle = \frac{1}{2\pi^2} \int_0^{\infty} P(k)\, \tilde{W}(\boldsymbol{k}R)\, k^2\, dk \tag{2.91}$$

is the variance of the smoothed overdensity field linearly extrapolated to the present, $P(k)$ and $\tilde{W}(\boldsymbol{k}; R)$ being the linear power spectrum and the Fourier transform of the window function, respectively[15].

Press and Schechter soon realized that their postulate was flawed; indeed, for a power-law power spectrum, $P(k) \propto k^n$, it is straightforward to show that $\sigma^2(M) \propto M^{-(n+3)/3}$, and thus $\sigma^2(M) \to \infty$ when $M \to 0$ provided that $n > -3$ (as overwhelmingly suggested by observational evidence). As a consequence, $\mathcal{P}(\delta_0^R > \delta_c(z)) \to 1/2$ in the same limit, which would mean that only half of the total mass of the Universe would be contained within collapsed objects. In order to solve this inconvenience, they introduced a factor of 2 into their formula via the (hand-waving) argument that the matter in regions that were initially underdense would eventually fall into collapsed objects forming from initially overdense regions, doubling their masses without influencing the shape of the mass function. This reasoning led to the following expression for the (comoving) number density of halos with masses in the range $(M, M+dM)$ at redshift z:

$$\begin{aligned} n(M,z)dM &= \frac{\bar{\rho}_{\rm m}(t_0)}{M} \frac{\partial F(>M)}{\partial M} dM = 2\frac{\bar{\rho}_{\rm m}(t_0)}{M} \frac{\partial \mathcal{P}[\delta_0^R > \delta_c(z)]}{\partial \sigma} \left| \frac{d\sigma}{dM} \right| dM = \\ &= \sqrt{\frac{2}{\pi}} \frac{\bar{\rho}_{\rm m}(t_0)}{M^2} \frac{\delta_c(z)}{\sigma(M)} \exp\left[-\frac{1}{2} \frac{\delta_c^2(z)}{\sigma^2(M)} \right] \left| \frac{d\ln\sigma}{d\ln M} \right| dM. \end{aligned} \tag{2.92}$$

This expression is known as the Press-Schechter *halo mass function* (HMF) and it is usually expressed via the variable[16]

$$\nu \equiv \frac{\delta_c^2(z)}{\sigma^2(M)}, \tag{2.93}$$

[15] The cosmological parameter $\sigma_8 \equiv \sqrt{\langle\, \delta_0^2(\boldsymbol{x}; 8\,\mathrm{Mpc}/h) \rangle}$ sets the normalization of the linear matter power spectrum.

[16] Special attention should be paid to the variables found in the literature, since others authors define ν as the square root of our value or directly work with the variance, σ^2.

in terms of which it reads

$$n(M,z) = \frac{\bar{\rho}_{\rm m}(t_0)}{M^2} f_{\rm PS}(\nu) \left| \frac{d\ln\nu}{d\ln M} \right|, \tag{2.94}$$

where

$$f_{\rm PS}(\nu) = \sqrt{\frac{\nu}{2\pi}} e^{-\nu/2}. \tag{2.95}$$

The above fit performed reasonably well when compared to numerical simulations up until the 1990s, where important deviations began to manifest themselves for halos of masses above and below the characteristic mass $M^*(z)$, defined as

$$\sigma(M^*) = \delta_c(z). \tag{2.96}$$

However, the Press-Schechter approach provides some insight into the hierarchical model of nonlinear structure formation, since halos of mass $M > M^*(z)$ will not have had the possibility to form in a substantial number by redshift z. As a consequence, the value $\nu = 1$ defines a characteristic scale, which is one reason why the ν variable is introduced. Furthermore, the so-called excursion set formalism put forward by Bond et al. (1991) introduced a framework based on Markovian random walks from which the same expression for the Press-Schechter HMF could be derived without the *ad hoc* necessity of a factor of 2.

The aforementioned discrepancies of the Press-Schechter model with respect to simulations made it clear that new models were needed for the HMF, the first of which was that of Ravi K. Sheth and Giuseppe Tormen (Sheth and Tormen, 1999), obtained by applying a spherical overdensity (SO) halo finder algorithm to a set of cosmological simulations. The two-parameter fit they obtained reads

$$n(M,z) = \frac{\bar{\rho}_{\rm m}(t_0)}{M^2} f_{\rm ST}(\nu) \left| \frac{d\ln\nu}{d\ln M} \right|, \tag{2.97}$$

where

$$f_{\rm ST}(\nu; a, p) = A(p) \sqrt{\frac{a\nu}{2\pi}} \left[1 + \left(\frac{1}{a\nu} \right)^p \right] e^{-a\nu/2} \tag{2.98}$$

and the best-fit values were found to be $a = 0.707$ and $p = 0.3$. The function $A(p)$ is determined by requiring that the all the mass in the Universe be bound up in dark matter halos, which mathematically translates to the condition that

$$\int_0^\infty \frac{f_{\rm ST}(\nu)}{\nu} d\nu = 1 \tag{2.99}$$

or, equivalently, that

$$A(p) = \left[1 + \frac{2^{-p}}{\sqrt{\pi}} \Gamma(1/2 - p) \right]^{-1}, \tag{2.100}$$

where Γ is the gamma function. The Sheth-Tormen (ST) fit was shown by Jenkins et al. (2001) to perform very well when tested against numerical simulations for different cosmologies from galaxy to cluster scales and up to redshift 5, although they provided an alternative fit to improve the high-mass end of the function. Furthermore, they showed that the HMF could be expressed in a universal form when appropriately rescaled, meaning that the function

$$f(\nu, z) \equiv \frac{M^2}{\bar{\rho}_{\rm m}(t_0)} n(M,z) \left| \frac{d\ln M}{d\ln\nu} \right| \tag{2.101}$$

remained unchanged for varying cosmologies and redshifts. Therefore, the introduction of the ν variable is not only due to mathematical aesthetics, but to the argument that, when parametrizing the HMF

in terms of $f(\nu)$ (as motivated by the Press-Schechter formalism), it is this function that seems to encapsulate all the information.

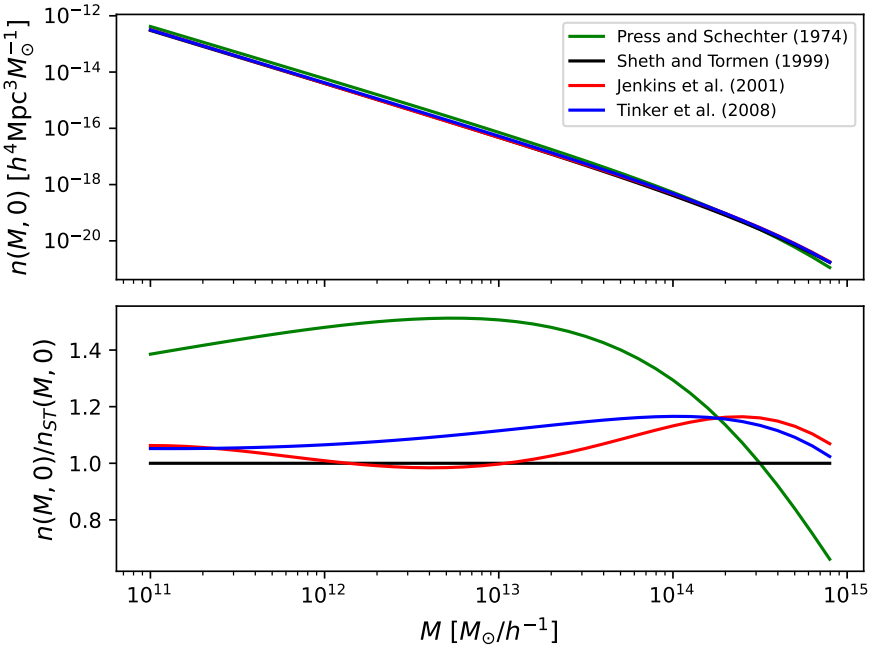

Figure 2.8: Top panel: $z = 0$ HMF for different models in a flat ΛCDM cosmology with *Planck*'s best-fit values (Planck Collaboration et al., 2020a). Bottom panel: ratio of each fitting function to the ST model. The mass range of validity for both the Jenkins et al. (2001) and Tinker et al. (2008) fits have been extrapolated for a direct comparison.

Since then, the literature has been inundated with a large number of functional fits to the HMF based on numerical simulations of ever-increasing precision and for different mass and redshift ranges, some of which are depicted in Figure 2.8. A number of them have confirmed universality within a few percent (Reed et al., 2003, 2007; Warren et al., 2006) while others have quantified small departures from it (Tinker et al., 2008; Crocce et al., 2010; Courtin et al., 2011; Watson et al., 2013). However, as shown by Despali et al. (2016), the non-universal behavior could be related to the way halos are defined or identified. Indeed, there is no standard definition of a halo. When a halo catalog is built from a numerical simulation, two different methods are commonly used: friends-of-friends (FOF; Davis et al., 1985) and spherical overdensity (SO; Lacey and Cole, 1994) algorithms. FOF algorithms are based on a parameter known as the linking length, l, and demand that all particles within a distance l from at least one other particle belong to the same halo. FOF halos are thus arbitrarily shaped. In contrast, SO algorithms identify halos by drawing a sphere around particles and stopping when the mean density within it equals an previously chosen overdensity threshold, Δ, effectively defining halos as spherical. Both methods have benefits and drawbacks and departures from universality have been found for the two kinds of algorithm. However, as shown by Despali et al. (2016), the universality of $f(\nu)$ could hold within a few percent if SO halos are identified using the virial overdensity value, Δ_{vir}, as the cutoff overdensity. For other common threshold values, like $\tilde{K}\bar{\rho}_{\mathrm{m}}$ or $\tilde{K}\rho_{\mathrm{crit}}$, (where $\tilde{K}$ is a constant), non-universal trends are induced. Additionaly, as pointed out by Courtin et al. (2011), not introducing the redshift dependence of the linear collapse threshold, $\hat{\delta}_c(z)$, could play a role in the deviations from universality.

Moreover, baryonic processes such as radiative cooling or feedback from supernovae and active

galactic nuclei are known to induce non-negligible modifications to the HMF based on dark matter-only simulations. Indeed, Cui et al. (2014) showed that the addition of feedback from active galactic nuclei causes a general decrease in the number density of halos, which is more noticeable at low masses and redshift, a trend which has been confirmed through simulations with a higher resolution (Sawala et al., 2013; Bocquet et al., 2015; Castro et al., 2020). Additionally, phenomena beyond the Standard Model of particle physics have been shown to be able to induce deviations in the HMF, namely the presence of massive neutrinos (Castorina et al., 2014) or a coupling between dark matter and dark energy (Cui et al., 2012; Giocoli et al., 2013).

Although the ST model will be the main fit we will use, we will compare it with that of Tinker et al. (2008), which reads

$$f_{\rm T}(\nu) = \tilde{A}\left[1 + \left(B\sqrt{\nu}\right)^{\tilde{p}}\right]e^{-C\nu}, \tag{2.102}$$

where the best-fit values for $z = 0$ were found to be

$$\tilde{A} = 0.093 \quad B = \frac{2.57}{\hat{\delta}(0)} \quad C = \frac{1.19}{\hat{\delta}^2(0)} \quad \tilde{p} = 1.47. \tag{2.103}$$

2.3.2 Deterministic halo bias

The HMF described in the above subsection allows us to compute the mean number of dark matter halos in *any* given comoving volume of the Universe. We can however ask ourselves how this expectation value changes when the selected region has a given background overdensity. In other words, since halos are biased tracers of the underlying density field (Cole and Kaiser, 1989), their overdensity, $\delta_{\rm h}$, must depend on that of the background density field, δ_b. How can we compare the halo two-point correlation function to that of dark matter?

In principle, this bias is non-local, non-linear and stochastic, but it is very common to assume that the halo density contrast is related to the dark matter overdensity field via a relation of the form

$$\delta_{\rm h}(\boldsymbol{x}, z|M) = b_1(M, z)\delta_b(\boldsymbol{x}, z). \tag{2.104}$$

The simplest method to obtain this relation is the so-called peak-background split (Kaiser, 1984; Bardeen et al., 1986). Based on the fact that a high density perturbation is bound to collapse earlier if it is found in a region of large-scale overdensity, we can consider separating the full overdensity field into a short-wavelength part corresponding to "peaks" ($\delta_{\rm h}$) and a long-wavelength background ($\delta_{\rm b}$) that modulates their number density:

$$\delta(\boldsymbol{x}, z) = \delta_b(\boldsymbol{x}, z) + \delta_{\rm h}(\boldsymbol{x}, z), \tag{2.105}$$

where δ_b is small enough to be treated by linear theory (see Figure 2.9). The small-scale component, the one that eventually forms halos, is affected by the large-scale part, since the effective threshold for halo formation will be modified according to

$$\hat{\delta}_c'(\boldsymbol{x}, z) = \hat{\delta}_c(z) - \delta_b(\boldsymbol{x}, z), \tag{2.106}$$

and the background density "felt" by this short-wavelength component will be

$$\bar{\rho}'(\boldsymbol{x}, z) = \bar{\rho}(z)(1 + \delta_b(\boldsymbol{x}, z)). \tag{2.107}$$

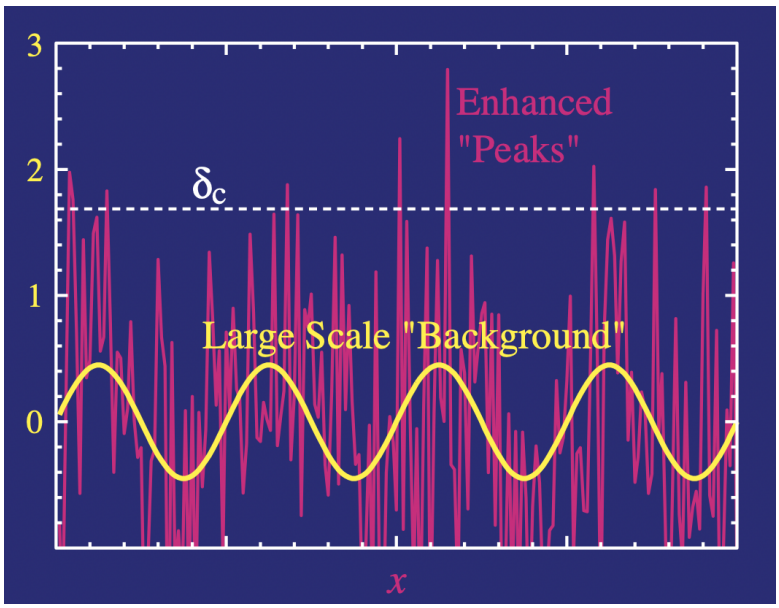

Figure 2.9: Hu, W. (2002). One dimensional representation of the peak-background split. Retrieved from `http://background.uchicago.edu/~whu/Presentations/trieste_lecture3.pdf`.

Let us now consider the HMF at a given point $\boldsymbol{x}$ where the background component is $\delta_b(\boldsymbol{x}, z)$ and expand it around the value $\delta_b = 0$ up to linear order:

$$\begin{aligned} n(M, z, \boldsymbol{x}) &= n(M, z) + \left.\frac{\partial n(M, z, \boldsymbol{x})}{\partial \delta_b}\right|_{\delta_b=0} \delta_b = \\ &= n(M, z)\left[1 + \left(1 - \left.\frac{d \ln f(\nu)}{d\nu}\right|_{\delta_b=0}\right)\delta_b\right] \end{aligned} \tag{2.108}$$

Since the halo overdensity is defined as

$$\delta_h(\boldsymbol{x}, z|M) \equiv \frac{n(M, z, \boldsymbol{x}) - n(M, z)}{n(M, z)}, \tag{2.109}$$

we arrive at equation (2.104), where

$$b_1(M, z) = 1 - \left.\frac{d \ln f(\nu)}{d\nu}\right|_{\delta_b=0} \tag{2.110}$$

is called the *linear (deterministic) halo bias*. Halos are therefore biased tracers of the underlying large-scale matter field: they tend to occur in regions that are overdense with matter. A more rigorous derivation of this relation was carried out by Mo and White (1996) using the conditional mass function and the excursion set formalism of Bond et al. (1991).

For the two HMF models that we will consider throughout this thesis, the linear bias factor reads:

$$b_1(M, z) = 1 + \frac{a\nu - 1}{\hat{\delta}_c(z)} + \frac{2p/\hat{\delta}_c(z)}{1 + (a\nu)^p} \qquad \text{for the ST fit} \tag{2.111}$$

$$b_1(M, z) = 1 + \frac{2C\nu}{\hat{\delta}_c(z)} - \frac{\tilde{p}\hat{\delta}_c(z)}{1 + [1/(B\sqrt{\nu})^{\tilde{p}}]} \qquad \text{for the Tinker fit} \tag{2.112}$$

This linear bias approximation only holds for large scales, when the long-wavelength mode can be

treated within the linear regime. Non-linear corrections are needed on small scales, as discussed by Desjacques et al. (2018).

2.3.3 Halo density profiles

Having discussed the abundance and clustering of dark matter halos, we now turn to analyzing their internal structure, namely their density profile. It is a reasonable first approximation to consider dark matter halos as spherically symmetric, in which case their mass distribution is described by a density profile of the form $\rho(r)$. For a power-law power spectrum, $P(k) \propto k^n$, with $n > -3$, as is the case for cold dark matter models, the typical density profile around a particle in the linear regime of a density field will be $\delta(r) \sim r^{-(n+3)}$, according to the scaling in linear theory (Mo et al., 2010). Assuming an initial density perturbation of this form, a careful treatment under the spherical collapse model (with shell-crossing) yields a dark matter halo profile of

$$\rho(r) \propto \begin{cases} r^{-2} & \text{for} \quad n \leq -1 \\ r^{-(3n+9)/(n+4)} & \text{for} \quad n > -1 \end{cases} . \tag{2.113}$$

Therefore, as a first approximation, the model predicts a density profile for virialized dark matter halos that resembles that of Singular Isothermal Spheres (SIS):

$$\rho_{\rm SIS}(r) = \frac{\sigma_v^2}{2\pi G r^2}, \tag{2.114}$$

which physically corresponds to a set of massive particles (stars in a galaxy) following a Maxwell-Boltzmann velocity distribution for all radii with a one-dimensional dispersion

$$\sigma_v = \sqrt{\frac{GM}{2r_h}}, \tag{2.115}$$

where $r_{\rm h}$ is the so-called virial radius of the halo, introduced to effectively truncate it so that the mass enclosed within it does not diverge. It is defined as the radius within which the mean matter density equals a given theshold for virialization, $\rho_{\rm h}$, for which there is no consensus. Indeed, as commented in 2.3.1, some authors use the so-called virial overdensity value $\rho_{\rm h}(z) = \Delta_{\rm vir}(z)\bar{\rho}_{\rm m}(z)$, while others opt for a constant times the mean or the critical density of the Universe ($\rho_{\rm h} = \tilde{K}\bar{\rho}$ and $\rho_{\rm h} = \tilde{K}\rho_{\rm crit}$, respectively). As a consequence, the halo mass definition (that is, the overdensity criterion) is a crucial aspect to always be considered, since physical conclusions could very well depend on it. For the entirety of the thesis, we use the virial overdensity value, for which we adopt the fit by Weinberg and Kamionkowski (2003):

$$\Delta_{\rm vir}(z) = 18\pi^2(1 + a\,\Theta^b(z)), \tag{2.116}$$

where $\Theta(z) \equiv 1/\Omega_m(z) - 1$,

$$a \equiv 0.399 - 1.309(\,|w_{\rm DE}|^{0.426} - 1) \tag{2.117}$$

$$b \equiv 0.941 - 0.205(\,|w_{\rm DE}|^{0.938} - 1), \tag{2.118}$$

and $w_{\rm DE}$ is the barotropic index of dark energy. Although it is sometimes used to analytically model galaxies due to its simplicity, the SIS profile is, at best, an approximation and numerical simulations are needed to derive the density profile of dark matter halos. A seminal paper by J. Navarro, C. Frenk

and S. White (Navarro et al., 1996) used high-resolution numerical simulations in a CDM model to show that spherically-identified dark matter halos were described by a universal profile of the form

$$\rho_{\rm NFW}(r) = \frac{\rho_s}{(r/r_s)(1+r/r_s)^2}, \tag{2.119}$$

where r_s and ρ_s are the scale radius and density, respectively. The Navarro-Frenk-White (NFW) profile is shallower than the SIS model near the halo center and steeper for larger radii ($r \gtrsim r_s$) and provides an excellent fit for dark matter halos of all scales in N-body simulations, albeit with some systematic deviations. Once the profile is truncated at a virial radius r_h, it is straightforward to find the mass of an NFW halo enclosed within it, yielding the following relation between ρ_s and the mass of the virialized halo (or, equivalently, its mean density ρ_h):

$$\rho_s = \frac{\rho_h}{3}\frac{c^3}{\ln(1+c) - c/(1+c)}, \tag{2.120}$$

where $c \equiv r_h/r_s$ is called the halo concentration parameter. As a consequence, for a given cosmological model, an NFW halo is uniquely determined by its mass and its concentration parameter or, equivalently, by the scale radius and density. This is the halo profile model which will be used throughout the thesis. It should be stressed once more that the mass M of the halo depends on the definition of the virial radius, that is, on the chosen overdensity threshold.

As explored in detail by numerical simulations, the concentration of dark matter halos is found to be strongly related to formation time and history (Wechsler et al., 2002; Zhao et al., 2003b), in that halos forming at an earlier time are typically more concentrated and that more massive halos are less concentrated (Navarro et al., 1997; Bullock et al., 2001; Zhao et al., 2003a; Neto et al., 2007; Macciò et al., 2007). Several authors have therefore been led to develop analytical models to compute the mean concentration parameter of a halo of mass M at a given time t. One of the simplest ones, which we use throughout the thesis, is that of Bullock et al. (2001), which computes the typical concentration of a halo of mass M at redshift z via the expression

$$c(M,z) = K\frac{a(z)}{a_c}, \tag{2.121}$$

where K is a constant and a_c is the scale factor at the time of collapse, an epoch defined such that the characteristic mass M^* equals a fixed fraction F of the mass of the halo, that is,

$$M^*(a_c) = FM. \tag{2.122}$$

Using N-body simulations, they find best-fit values of $F = 0.01$ and $K = 4.0$.

Since we will be working in Fourier space, it will be necessary to have an expression for the Fourier transform of the NFW profile, which is given by (Cooray and Sheth, 2002)

$$\begin{aligned}\tilde{u}(k,z|M) = \frac{4\pi\rho_s r_s^3(M,z)}{M}\Big[&\sin kr_s\big[\mathrm{Si}([1+c]kr_s) - \mathrm{Si}(kr_s)\big] - \\ &- \frac{\sin ckr_s}{[1+c]kr_s} + \cos kr_s\big[\mathrm{Ci}([1+c]kr_s) - \mathrm{Ci}(kr_s)\big]\Big],\end{aligned} \tag{2.123}$$

where

$$\mathrm{Si}(x) \equiv \int_0^x \frac{\sin t}{t}dt \qquad \text{and} \qquad \mathrm{Ci}(x) \equiv -\int_x^\infty \frac{\cos t}{t}dt. \tag{2.124}$$

Although we will not use other halo profiles throughout the thesis, there is a long list of interesting and popular fits derived via the introduction of different baryonic effects in numerical simulations or motivated by the well-known cusp-core problem. A brief summary of these models can be found in (Freundlich et al., 2020).

2.3.4 The Halo Occupation Distribution

If we aim to describe the non-linear clustering of (dark) matter within the halo model, we need no further information. However, within our current understanding of galaxy formation, the gas cooling preceding star formation can only occur in potential wells such as those of dark matter halos. Galaxies are thus thought to form and evolve within these halos and, therefore, they are biased tracers of the underlying dark matter distribution. However, it is not simple to establish a link between galaxies and dark matter halos, since the natural approach of hydrodynamic simulations is strongly dependent on the baryonic processes of galaxy formation, which are still poorly understood. Therefore, other methods are explored to quantify the connection between galaxies and dark matter halos, the most popular of which is the Halo Occupation Distribution (HOD; Ma and Fry, 2000; Peacock and Smith, 2000; Seljak, 2000; Scoccimarro et al., 2000; Berlind et al., 2003), which characterizes the relation between the dark matter and galaxy distributions by means of the the probability $P(N_g|M)$ that a halo of mass M hosts N_g galaxies. Indeed, the computation of galaxy correlation functions requires the knowledge of moments of the form

$$\langle N_g^k \rangle_M = \sum_{N_g=0}^{\infty} N_g^k P(N_g|M) \tag{2.125}$$

and the HOD is therefore essential to describe galaxy clustering. Although the probability $P(N_g|M)$ should in principle be derived from a theory of galaxy formation, the absence of a complete framework makes it customary to adopt mathematical models that are (partially) motivated by the occupation number statistics in hydrodynamic simulations. Furthermore, and inspired by galaxy formation models, it is useful to separate the contribution of central and satellite galaxies within halos, so that the mean number of galaxies in a halo of mass M is

$$\langle N_g \rangle_M = \langle N_{c_g} \rangle_M + \langle N_{s_g} \rangle_M. \tag{2.126}$$

Since two-point galaxy correlation functions involve the quantity $\langle N_g(N_g-1)\rangle$, we have

$$\begin{aligned} \langle N_g(N_g-1)\rangle_M &= \langle N_{c_g}^2 \rangle_M + 2\langle N_{c_g} N_{s_g}\rangle_M + \langle N_{s_g}^2\rangle_M - \langle N_{c_g}\rangle_M - \langle N_{s_g}\rangle_M = \\ &= \langle N_{s_g}(N_{s_g}-1)\rangle_M + 2\langle N_{c_g}\rangle_M \langle N_{s_g}\rangle_M, \end{aligned} \tag{2.127}$$

where we have assumed that the occupation statistics of central and satellite galaxies are independent. If we now assume that the probability distribution of the number of satellites follows a Poisson distribution, as motivated, for instance, by the occupation statistics of subhalos found in N-body simulations (Kravtsov et al., 2004),

$$P(N_{s_g}|M) = \frac{\lambda^{N_s} e^{-\lambda}}{N_{s_g}!}, \tag{2.128}$$

where $\lambda \equiv \langle N_{s_g} \rangle_M$, then the HOD is completely determined by the mean occupation number of both central and satellite galaxies, since all higher-order moments of its distribution satisfy

$$\langle N_{s_g}(N_{s_g} - 1) \dots (N_s - r) \rangle_M = \langle N_{s_g} \rangle_M^{r+1}. \tag{2.129}$$

The simplest model for the mean occupation number of galaxies is motivated by results from smoothed particle hydrodynamic simulations and semianalytic calculations (Kauffmann et al., 1999; Seljak, 2000; Scoccimarro et al., 2000; White et al., 2001; Yoshikawa et al., 2001; Berlind et al., 2003) and parametrizes the halo occupation statistics for galaxies brighter than a certain luminosity threshold as the sum of a step function for central galaxies and a power law for satellite galaxies. The three-parameter $(M_{\min}, M_1, \alpha)$ model of Zehavi et al. (2005), depicted in Figure 2.10, is:

$$\langle N_g \rangle_M = \langle N_{c_g} \rangle_M + \langle N_{s_g} \rangle_M = \left[1 + \left(\frac{M}{M_1} \right)^{\alpha} \right] \Theta(M - M_{\min}), \tag{2.130}$$

where $M_{\min}$ is the mean minimum mass a halo needs to have in order to host galaxies, M_1 is the mean halo mass at which there is exactly one satellite and α is the logarithmic slope of the satellite occupation number. This is the fit that will be used throughout the thesis.

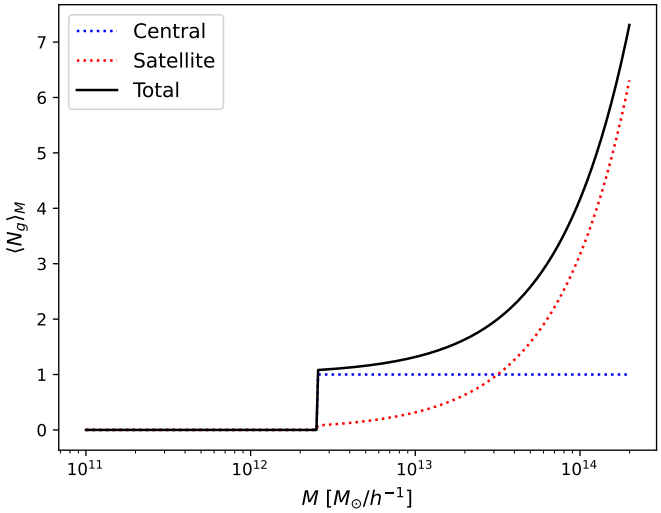

Figure 2.10: Mean number of galaxies in a halo as a function of its mass, according to the model of Zehavi et al. (2005). The black solid line represents the total number of galaxies, while the blue and red dotted lines show the contribution from central and satellite galaxies.

A more complicated model, which is rather popular in the literature, represents a more realistic smooth transition of the mean number of central galaxies from 0 to 1 and adds a cut-off mass for satellites. This five-parameter $(\hat{M}_{\min}, \sigma_{\log M}, M_{\text{cut}}, \hat{M}_1, \alpha)$ fit is described by Zheng et al. (2005):

$$\begin{aligned} \langle N_g \rangle_M &= \langle N_{c_g} \rangle_M + \langle N_{s_g} \rangle_M = \\ &= \frac{1}{2} \left[1 + \operatorname{erf}\left(\frac{\log M - \log \hat{M}_{\min}}{\sigma_{\log M}} \right) \right] + \left[\frac{M - M_{\text{cut}}}{\hat{M}_1} \right]^{\alpha} \Theta(M - M_{\text{cut}}), \end{aligned} \tag{2.131}$$

where $\hat{M}_{\min}$ is the characteristic mean minimum halo mass for central galaxies, $\sigma_{\log M}$ describes the transition width due to the scatter in the luminosity-mass relation of central galaxies, M_{cut} is the mean halo mass below which there are no satellites, $\hat{M}_1$ sets the normalization of the satellite occupation

number and α is, as in the previous case, its logarithmic slope.

Other alternative parametrizations for halo occupation statistics involve the conditional luminosity function, $\Phi_g(L|M)$, that is, the mean number of galaxies of luminosity L within halos of mass M. This approach (Yang et al., 2003; van den Bosch et al., 2003) has the advantage of being valid for any luminosity sample, not only those with a threshold value, so that

$$\langle N_g \rangle_M = \int_{L_1}^{L_2} \Phi_g(L|M) dL \tag{2.132}$$

and, furthermore, connects the galaxy luminosity function, $\Phi_g(L)$, with the HMF:

$$\Phi_g(L) = \int_0^\infty \Phi_g(L|M)\, n(M,z)\, dM. \tag{2.133}$$

By means of an analysis of a galaxy group catalog, Yang et al. (2008) showed that the conditional luminosity function for central galaxies is well parametrized by a log-normal distribution,

$$\Phi_{c_g}(L|M) = \frac{1}{\sqrt{2\pi\sigma_c^2}} \exp\left[-\left(\frac{\ln L/L_c}{\sqrt{2\sigma_c^2}} \right)^2 \right] \frac{1}{L}, \tag{2.134}$$

whether that of satellites fits a modified Schechter function:

$$\Phi_{s_g}(L|M) = \frac{\phi_s}{L_s} \left(\frac{L}{L_s} \right)^{\alpha_s} \exp[-(L/L_s)^2], \tag{2.135}$$

where the parameters $L_c, L_s, \sigma_c, \phi_s$ and α_s all depend on halo mass. This approach has been exploited in several studies of galaxy clustering and galaxy-galaxy lensing (Cacciato et al., 2009, 2013; van den Bosch et al., 2013) and satellite galaxy kinematics (More et al., 2008; Lange et al., 2018; van den Bosch et al., 2019).

2.3.5 The clustering of dark matter and galaxies

Having already introduced all the ingredients of the halo model, we now proceed to construct the non-linear power spectra of dark matter and galaxies and, most importantly for our purposes, the dark matter-galaxy cross-power spectrum.

Let us begin by imagining that, at a given redshift z, space is divided up into a large number of comoving volumes, ΔV_i, that are small enough to contain one halo at most. Therefore, the dark matter field at comoving position $\boldsymbol{x}$ can be written as a sum over all such volumes,

$$\rho_{\rm dm}(\boldsymbol{x}, z) = \sum_i \mathcal{N}_i(z) M_i u(\boldsymbol{x} - \boldsymbol{x}_i, z|M_i), \tag{2.136}$$

where $\mathcal{N}_i(z)$ is the occupation number in volume ΔV_i at redshift z, which can be either 0 or 1, and we have written the density profile of each halo in terms of its normalized profile $u(\boldsymbol{x} - \boldsymbol{x}_i, z|M_i)$. Defining the dark matter overdensity field as $\delta_{\rm dm}(\boldsymbol{x}, z) = \rho_{\rm dm}(\boldsymbol{x}, z)/\bar{\rho}_{\rm m}(z) - 1$, then its two-point correlation function is given by

$$\begin{aligned} \xi_{\text{dm-dm}}(\boldsymbol{x}, \boldsymbol{x}+\boldsymbol{r}, z) &\equiv \langle \delta_{\rm dm}(\boldsymbol{x}, z)\, \delta_{\rm dm}(\boldsymbol{x}+\boldsymbol{r}, z) \rangle = \\ &= \sum_{i,j} \langle \mathcal{N}_i(z) \mathcal{N}_j(z) \frac{M_i}{\bar{\rho}_{\rm m}(z)} \frac{M_j}{\bar{\rho}_{\rm m}(z)} u(\boldsymbol{x} - \boldsymbol{x}_i, z|M_i) u(\boldsymbol{x} + \boldsymbol{r} - \boldsymbol{x}_j, z|M_j) \rangle - 1, \end{aligned} \tag{2.137}$$

where $\langle \ldots \rangle$ denotes an ensemble average, which can be suitably computed as a spatial average following the ergodic principle (Mo et al., 2010)[17]. The above expression can be split into the so-called *1-halo* and *2-halo* terms by taking the terms with $i = j$ and $i \neq j$ of the summation, respectively. Indeed, when $i = j$, we are taking into account correlations from the same halo, while the $i \neq j$ terms arise from contributions between different halos. Before proceeding, let us realize that

$$\langle \mathcal{N}_i(z) M_i u(\boldsymbol{x} - \boldsymbol{x}_i, z|M_i) \rangle = \int_0^\infty dM n(M,z) M \Delta V_i u(\boldsymbol{x} - \boldsymbol{x}_i, z|M), \tag{2.138}$$

since both quantities describe the average dark matter density at $\boldsymbol{x}$ at redshift z due to the halo in cell ΔV_i. Using the above property, the 1-halo term reads

$$\begin{aligned}
\xi^{\text{1h}}_{\text{dm-dm}}(\boldsymbol{x}, \boldsymbol{x}+\boldsymbol{r}, z) &= \sum_i \left\langle \mathcal{N}_i^2(z) \frac{M_i^2}{\bar{\rho}_{\text{m}}^2(z)} u(\boldsymbol{x} - \boldsymbol{x}_i, z|M_i)\, u(\boldsymbol{x}+\boldsymbol{r} - \boldsymbol{x}_i, z|M_i) \right\rangle = \\
&= \sum_i \int_0^\infty dM \frac{M^2}{\bar{\rho}_{\text{m}}^2(z)} n(M,z)\, \Delta V_i\, u(\boldsymbol{x} - \boldsymbol{x}_i, z|M)\, u(\boldsymbol{x}+\boldsymbol{r} - \boldsymbol{x}_i, z|M) = \\
&= \int_0^\infty dM \frac{M^2}{\bar{\rho}_{\text{m}}^2(z)} n(M,z) \int d^3\boldsymbol{y}\, u(\boldsymbol{x} - \boldsymbol{y}, z|M)\, u(\boldsymbol{x}+\boldsymbol{r} - \boldsymbol{y}, z|M)
\end{aligned} \tag{2.139}$$

where we have interchanged the summation with an integral:

$$\sum_i \Delta V_i\, u(\boldsymbol{x} - \boldsymbol{x}_i, z|M)\, u(\boldsymbol{x}+\boldsymbol{r} - \boldsymbol{x}_i, z|M) \to \int d^3\boldsymbol{y}\, u(\boldsymbol{x} - \boldsymbol{y}, z|M)\, u(\boldsymbol{x}+\boldsymbol{r} - \boldsymbol{y}, z|M). \tag{2.140}$$

The 2-halo term involves the cross terms of the summation and is a little more convolved, since halos are clustered themselves. Indeed, incluiding the -1 term, it reads

$$\begin{aligned}
\xi^{\text{2h}}_{\text{dm-dm}}(\boldsymbol{x}, \boldsymbol{x}+\boldsymbol{r}, z) &= \sum_i \sum_{j \neq i} \left\langle \mathcal{N}_i(z)\, \mathcal{N}_j(z) \frac{M_i}{\bar{\rho}_{\text{m}}(z)} \frac{M_j}{\bar{\rho}_{\text{m}}(z)} u(\boldsymbol{x} - \boldsymbol{x_i}, z|M_i) u(\boldsymbol{x}+\boldsymbol{r} - \boldsymbol{x_j}, z|M_j) \right\rangle - 1 = \\
&= \sum_i \sum_{j \neq i} \int_0^\infty dM_1 \frac{M_1}{\bar{\rho}_{\text{m}}(z)} n(M_1,z) \int_0^\infty dM_2 \frac{M_2}{\bar{\rho}_{\text{m}}(z)} n(M_2,z)\, \Delta V_i\, \Delta V_j\, \times \\
&\times \left[1 + \xi_{\text{h-h}}(\boldsymbol{x_i} - \boldsymbol{x_j}, z|M_1, M_2)\right] u(\boldsymbol{x_1} - \boldsymbol{x_i}, z|M_1)\, u(\boldsymbol{x_2} - \boldsymbol{x_j}, z|M_2) - 1 = \\
&= \int_0^\infty dM_1 \frac{M_1}{\bar{\rho}_{\text{m}}(z)} n(M_1,z) \int_0^\infty dM_2 \frac{M_2}{\bar{\rho}_{\text{m}}(z)} n(M_2,z)\, \times \\
&\times \int d\boldsymbol{y_1} \int d\boldsymbol{y_2}\, u(\boldsymbol{x_1} - \boldsymbol{y_1}, z|M_1)\, u(\boldsymbol{x_2} - \boldsymbol{y_2}, z|M_2)\, \xi_{\text{h-h}}(\boldsymbol{y_1} - \boldsymbol{y_2}, z|M_1, M_2),
\end{aligned} \tag{2.141}$$

where $\xi_{\text{h-h}}$ denotes the halo-halo correlation function. According to the discussion in 2.3.2 with respect to halo bias, we can approximate

$$\xi_{\text{h-h}}(\boldsymbol{y}_1 - \boldsymbol{y}_2, z|M_1, M_2) \approx b_1(M_1,z) b_1(M_2,z) \xi^{\text{lin}}_{\text{dm-dm}}(\boldsymbol{y_1} - \boldsymbol{y_2}) \tag{2.142}$$

for large scales. Given the convolution integrals appearing in the model, it is more convenient to work in Fourier space. Let

$$\rho_{\text{dm}}(\boldsymbol{k}, t) = \sum_i \mathcal{N}_i(z) M_i \tilde{u}(\boldsymbol{k}|M_i, z) e^{-\boldsymbol{k}\cdot x_i} \tag{2.143}$$

be the Fourier transform of the density field defined in (2.136), where $\tilde{u}(\boldsymbol{k}, z|M)$ is the Fourier transform

[17] It should be noted that the dark matter density contrast is defined here with respect to the mean *total matter* density.

of the density profile, as discussed in 2.3.3. The non-linear prediction of the halo model for the dark matter power spectrum is thus

$$P_{\text{dm-dm}}(k,z) = P^{\text{1h}}_{\text{dm-dm}}(k,z) + P^{\text{2h}}_{\text{dm-dm}}(k,z), \tag{2.144}$$

where

$$P^{\text{1h}}_{\text{dm-dm}}(k,z) = \int_0^\infty dM\, M^2 \frac{n(M,z)}{\bar{\rho}^2_{\text{m}}(z)} |\tilde{u}_{\text{dm}}(k,z|M)|^2 \tag{2.145}$$

and

$$P^{\text{2h}}_{\text{dm-dm}}(k,z) = P^{\text{lin}}_{\text{dm-dm}}(k,z) \int_0^\infty dM_1\, M_1 \frac{n(M_1,z)}{\bar{\rho}_{\text{m}}(z)} b_1(M_1,z)\tilde{u}_{\text{dm}}(k,z|M_1) \times \\ \times \int_0^\infty dM_2\, M_2 \frac{n(M_2,z)}{\bar{\rho}_{\text{m}}(z)} b_1(M_2,z)\tilde{u}_{\text{dm}}(k,z|M_2) \tag{2.146}$$

This quantity, however, was not quite what we were looking for, but the halo model prescription for its construction is equally valid for the galaxy-dark matter cross-power spectrum via the introduction of the HOD. Although the exact details are not complicated and are certainly parallel to the ones we have just described for the dark matter power spectrum, we refer the reader to Cooray and Sheth (2002) or Mo et al. (2010) for a detailed derivation and we ourselves present the result without proof.

The non-linear prediction of the halo model for the galaxy-dark matter cross-power spectrum is given by

$$P_{\text{g-dm}}(k,z) = P^{\text{1h}}_{\text{g-dm}}(k,z) + P^{\text{2h}}_{\text{g-dm}}(k,z), \tag{2.147}$$

where

$$P^{\text{1h}}_{\text{g-dm}}(k,z) = \int_0^\infty dM\, M \frac{n(M,z)}{\bar{\rho}_{\text{m}}(z)} \frac{\langle N_g \rangle_M}{\bar{n}_g(z)} |\tilde{u}_{\text{dm}}(k,z|M)||\tilde{u}_{\text{g}}(k,z|M)|^{q-1} \tag{2.148}$$

and

$$P^{\text{2h}}_{\text{g-dm}}(k,z) = P^{\text{lin}}_{\text{dm-dm}}(k,z) \Big[\int_0^\infty dM\, M \frac{n(M,z)}{\bar{\rho}_{\text{m}}(z)} b_1(M,z)\tilde{u}_{\text{dm}}(k,z|M) \Big] \cdot \\ \cdot \Big[\int_0^\infty dM\, n(M,z) b_1(M,z) \frac{\langle N_g \rangle_M}{\bar{n}_g(z)} \tilde{u}_g(k,z|M) \Big], \tag{2.149}$$

$\tilde{u}_g(k,z|M)$ being the Fourier transform of the galaxy density profile, which can be taken to be equal to that of dark matter as a good approximation (Sheth and Diaferio, 2001). The exponent q needs to be set to 1 for central galaxies and to 2 for satellites (Cooray and Sheth, 2002) and $\bar{n}_g(z)$ is the mean number density of galaxies at redshift z.

The above equations provide the final necessary information for the foreground-background angular cross-correlation function, for which we derived an expression in 2.2.3. The theoretical model for our observable is now established and we can move on to the description of the data and to the computation of the corresponding estimator.

3. Galaxy samples and measurements

Having already established the theoretical background, the methodology followed throughout the thesis will be thoroughly described in this chapter. A first section covers, nonetheless, a review on the history and physical properties of the protagonists of this dissertation, submillimeter galaxies, with the aim of providing a motivation behind their use for our purposes. Additionally, this segment allows us to introduce and specify the main characteristics of the *Herschel*-ATLAS submillimeter survey, from which our background galaxy sample is extracted. The second section describes the selection and characterization procedures for both the foreground and background galaxies we have used. The methodology employed for the measurement of the cross-correlation function is detailed in the third section, including a discussion about the different procedures available to this end. Lastly, the fourth section covers our approach for the Bayesian estimation of parameters, based on a Markov chain Monte Carlo algorithm.

3.1 Submillimeter galaxies and the *Herschel*-ATLAS survey

The term *submillimeter galaxy* historically referred to a galaxy detected at 850 μm with a flux density of $S_{850} > 2 - 5$ mJy. The definition, although not written in stone, has recently been applied more broadly to galaxies with flux densities of $S > 1$ mJy from 250 μm to 2 mm (Casey et al., 2014). Let us go on a walk through the history of the brightest and best understood kind of dusty star-forming galaxies at high redshift and discover the physical properties that make them the optimal background sample for magnification bias studies.

One should start by mentioning that, compared to its optical or ultraviolet counterparts, astrophysical observations in the submillimeter band, loosely defined as wavelengths from ~ 200 μm to 1 mm, had a late flowering due to the technical challenge of building detectors that were sensitive enough at these energy scales. Furthermore, emission and absorption in our atmosphere allows reasonable detections only from high and dry locations on Earth, airborne observatories or from space. The first space telescope closing in on the shorter end of this wavelength range was the InfraRed Astronomy Satellite (IRAS; Neugebauer et al., 1984), which performed an all-sky map at infrared wavelengths of 60 and 100 μm. Launched in January 1983, the mission lasted ten months and detected around 250000 infrared sources, the vast majority of which were not bright enough to have been seen in previous optical catalogs. One of its key discoveries was that the total luminosity of a significant number of local galaxies was dominated by infrared emission. Among them, Houck et al. (1985) found a sample of 9 sources with exceptionally high (~ 50) infrared-to-optical luminosity ratios that were invisible or extremely faint on the plates of the Palomar Sky Survey. This new population of local galaxies, with infrared luminosities of $10^{12} L_\odot < L_{\rm IR} < 10^{13} L_\odot$, was termed UltraLuminous InfraRed Galaxies (ULIRGS) and are currently thought to be mergers of two or more equal-mass galaxies (de Jong et al., 1984; Lonsdale et al., 1984; Joseph and Wright, 1985; Veilleux et al., 2002). The Infrared Space Observatory (ISO; Lemke et al., 1996) satellite, launched in 1995, found a more numerous population of local galaxies with large infrared luminosities ($L_{\rm IR} > 10^{11} L_\odot$), a small ($\sim 12$) subset of which were ULIRGs. It should be stressed that local ULIRGs are very rare when compared to Milky Way-like galaxies and do not substantially contribute to the cosmic star formation rate at $z = 0$. As we will shortly discover, distant infrared-bright galaxies paint a completely different picture.

The advent of the Far-InfraRed Absolute Spectrophotmeter (FIRAS) on the COsmic Background Explorer (COBE; Boggess et al., 1992) satellite in the early 1990s provided further indisputable evidence that far-infrared and submillimeter physics was not to be overlooked. The contribution of infrared and submillimeter wavelengths above 150 μm to the energy density of the Universe was shown to be comparable to the optical and ultraviolet bands. The amount of light absorbed by the dusty interstellar medium, heated and then re-radiated in the (rest-frame) far infrared and submillimeter was therefore as abundant as the one directly observed in the ultraviolet and optical (Puget et al., 1996; Fixsen et al., 1998; Dwek et al., 1998; Dole et al., 2006). Relying only on observations in the latter bands would thus result in missing approximately half the star formation activity. The analysis of this (unresolved) cosmic infrared background (CIB) along with optical galaxy surveys led to the key conclusion that there had to exist either a population of dust-enshrouded galaxies or a number of regions enshrouded in dust within optically-detected galaxies where newly formed stars were born.

Amidst all these hypotheses, the implementation of the Submillimeter Common-User Bolometer Array (SCUBA) on the James Clerk Maxwell Telescope in 1997 (Holland et al., 1999) brought about a revolution in submillimeter astronomy and shed light on the matter. Operating simultaneously at 450 μm and 850 μm atop the Mauna Kea volcano in Hawaii, it had unmatched sensitivity at the time, with a field of view of around 5 arcmin^2. The first deep-field SCUBA maps detected a large population of star-forming galaxies not known at the time and that were extremely bright at far-infrared and submillimeter wavelenghts, while nearly invisible in the optical band. In particular, the work of Smail et al. (1997), Barger et al. (1998) and Hughes et al. (1998) showed that the surface density of dusty submillimeter sources was largely above the expectation for non-evolving galaxy populations and that optical surveys might have substantially underestimated the star formation density in the distant Universe. As pointed out by Blain et al. (1999), these submillimeter analyses showed that optical studies were missing a large population of luminous, strongly dust-obscured and high-redshift sources, a claim which was supported by the works of Cimatti et al. (1998), Ivison et al. (1998a) and Ivison et al. (1998b).

A comprehensive review on the knowledge about the physics of submillimeter galaxies in the early days of SCUBA was done by Blain et al. (2002), where it was also pointed out that the analysis of the properties of individual galaxies was, however, an extremely complicated task given the large beamsize of SCUBA and the difficulty in identifying multi-wavelength counterparts. However, the realization that submillimeter sources should be likewise identified as faint radio galaxies (Ivison et al., 1998b, 2000; Smail et al., 2000; Ivison et al., 2002; Ivison et al., 2007) made it possible to refine some of the SCUBA positions and detect a population of galaxies with a median redshift of $z \sim 2.2$ (Chapman et al., 2004, 2005), star formation rates larger than $500\,M_\odot/\mathrm{yr}$[1] and far-infrared luminosities $L_{\mathrm{FIR}} \gtrsim 10^{12.5} L_\odot$, thus confirming the suspicion that these submillimeter galaxies indeed sat at high redshifts. Figure 3.1 (Ivison et al., 2000) compares the 850 μm map of the central region of the Abell 1835 cluster with a radio image of the same region taken by the Very Large Array observatory (Thompson et al., 1980), where one can observe that four out of the five submillimeter sources have an associated radio emission, underlining the usefulness of deep radio maps for the identification of submillimeter galaxies. An optical image of this region overlaid with the SCUBA 850 μm contours is also shown to highlight the faint optical emission of these sources.

At this point, and given their large (rest-frame) infrared luminosities, it is only natural to think of

[1]For the purpose of comparison, note that the Milky Way has a star formation rate of $\sim 2\,M_\odot/\mathrm{yr}$ (Kennicutt, 1998; Madau and Dickinson, 2014).

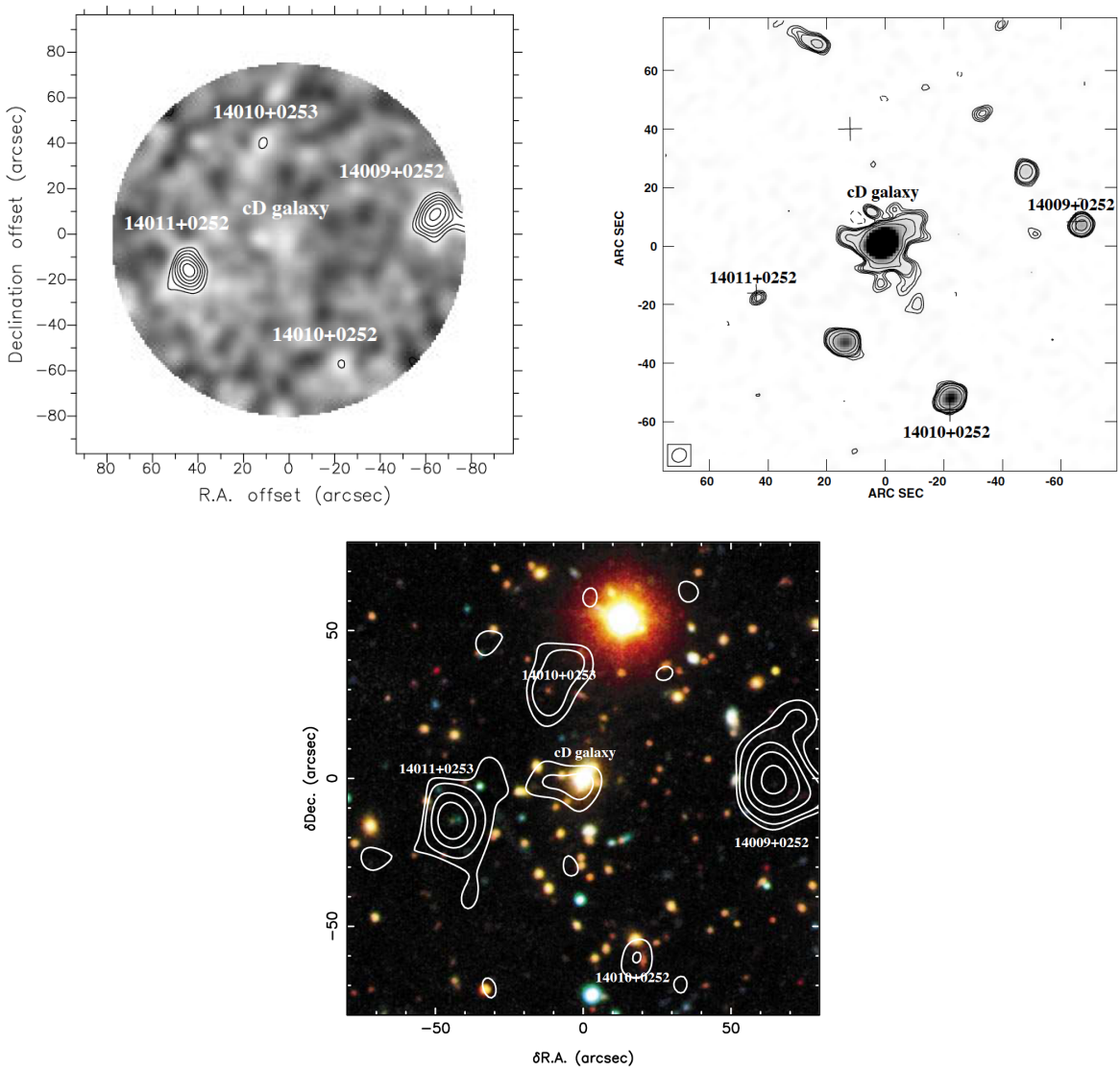

Figure 3.1: (Ivison et al., 2000) Top left: SCUBA 850 μm image of the central region of the massive cluster Abell 1835. Top right: Radio image of the same region taken by the Very Large Array at 1.4 GHz, with crosses at the positions of the submillimeter galaxies identified by SCUBA. Bottom: Background 3-color optical image of Abell 1835 overlaid with the SCUBA 850 μm contours, smoothed to a full-width half maximum of 20 arcsec.

submillimeter galaxies as the high-redshift analogs of local ULIRGs. Although they are often regarded as so, the picture is, however, more complicated (Casey et al., 2014; Hodge and da Cunha, 2020) and, moreover, their number density is about two orders of magnitude larger than local ULIRGs (Cowie et al., 2004; Lonsdale et al., 2006) and their contribution to the volume-averaged cosmic star formation rate density at $z = 2 - 4$ is $\sim 20\%$ (Michałowski et al., 2010), to be compared to the negligible contribution of their alleged local counterparts.

Subsequent instruments continued scrutinizing the far-infrared and submillimeter bands with the aid of data from other wavelengths to further examine the properties of these dusty, star-forming galaxies (see Casey et al., 2014, for a review of far-IR, millimeter and submillimeter instruments and our physical understanding of these galaxies). For our purposes, it suffices to start by mentioning the Submillimeter High Angular Resolution Camera-II (SHARC-II), which operated at the Caltech Submillimeter Observatory on the Mauna Kea volcano at wavelengths of 350 μm and 450 μm. From

2002 to 2013, it mainly performed far-IR follow-ups of submillimeter-selected SCUBA galaxies with a slightly smaller ($9''$) beamsize. In particular, Kovacs et al. (2006) and Coppin et al. (2008) constrained the Spectral Energy Distribution (SED) of submillimeter galaxies detected by SCUBA at 850 μm assuming a modified blackbody model, the former incorporating additional information from the MAx-planck Millimeter BOlometer (MAMBO; Kreysa et al., 1998) and the BOLOmeter CAMera (BOLOCAM; Laurent et al., 2005) instruments. Figure 3.2 (in black) shows a schematic representation of the infrared SED of a submillimeter galaxy at $z = 2$ with $L_{\rm IR} = 3 \cdot 10^{12} L_{\odot}$, as presented by Pope et al. (2008) after a fit using multiwavelength information in the Great Observatories Origis Deep Survey-North field (GOODS-N; Dickinson et al., 2004), one of the most studied regions of the sky. The template SEDs of two local ULIRGs (Polletta et al., 2007) are shown in red for comparison, as well as vertical shaded bands representing the sensitivities of different far-infrared observatories.

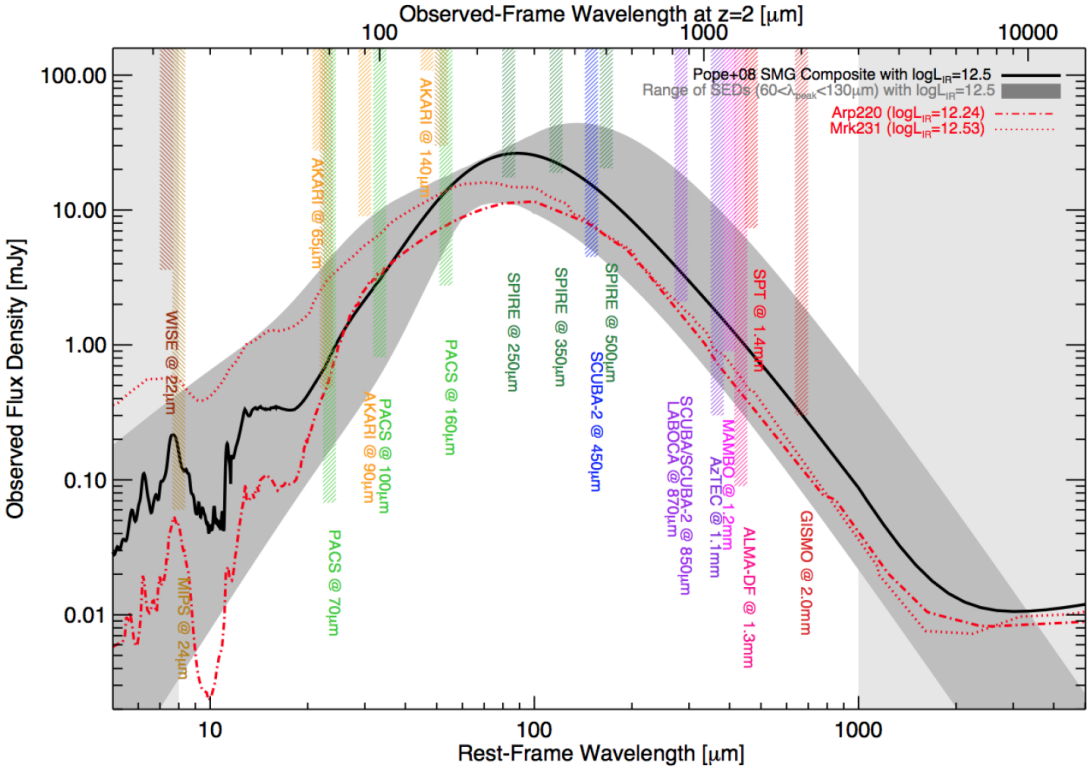

Figure 3.2: Schematic picture of the infrared SED of a submillimeter galaxy at $z = 2$ with $L_{\rm IR} = 10^{12.5} L_{\odot}$ (in black; Pope et al., 2008). In gray, a band showing the plausible range of SEDs for galaxies of the same luminisoty with peak SED wavelengths from 70 to 130 μm. The red lines depict model SEDs for the local ULIRGs Arp 220 and Mrk 231. Vertical shaded bands represent the 3σ detection limit of certain instruments.

As for airborne telescopes, The Balloon-borne Large Aperture Submillimeter Telescope (BLAST; Devlin et al., 2009) was the first submillimeter telescope launched on a balloon, operating above most of the water vapor of our atmosphere, at an altitude of 35 km. The most important of its scientific flights took place in 2006 in the Antarctic circle and provided the first wide-area (~ 10 deg^2) maps of the sky at 250, 350 and 500 μm, within the Extended Chandra Deep Field South (ECDF-S). Although its features were best suited for CIB measurements, it led to the discovery of an important number of individual luminous galaxies in the far-IR. Indeed, Dunlop et al. (2010) determined the redshift of 20 galaxies selected by BLAST at 250 μm within the central part of the GOODS-South field, also mapped by the Large Apex BOlometer CAmera (LABOCA; Siringo et al., 2009; Weiß et al., 2009) at

870 μm and the AzTEC bolometer array camera (Wilson et al., 2008; Scott et al., 2010) at 1.1 mm. Furthermore, Chapin et al. (2011) characterized the temperatures and far-IR luminosities of 69 sources observed in the ECDF-S and with a significant emission in the BLAST bands by combining the data with the aforementioned LABOCA map of the same region at 870 μm.

In the category of space observatories, even though the entire sky was surveyed by IRAS at 60 and 100 μm, very little was found out about the dust in early-type galaxies (Bregman et al., 1998). Furthermore, IRAS measurements were shown by Devereux and Young (1990) to miss about 90% of the dust in late-type galaxies, since it effectively detected only the small fraction of dust that was warm enough to radiate significantly in the far-infrared. Although in a much less critical manner, the ISO and *Spitzer* (Werner et al., 2004) telescopes also suffered from this problem (Bendo et al., 2003). Their efforts and breakthroughs are nonetheless unquestionable, since they showed that $\sim 70\%$ of the comoving star formation rate density at $0.5 < z < 3.0$ is obscured by dust (Chary and Elbaz, 2001; Le Floc'h et al., 2005; Pérez-González et al., 2005). However, submillimeter astronomers were still lacking a large-area survey with the ability to provide information about the dust content and the dust-obscured star formation activity. In essence, scientists were awaiting the kind of leap made in the optical with the Two Degree Field Galaxy Redshift Sky Survey (2dF; Colless, 1999) and SDSS (York et al., 2000), but this time in the far-infrared and submillimeter bands. This was the motivation behind the *Herschel* space observatory.

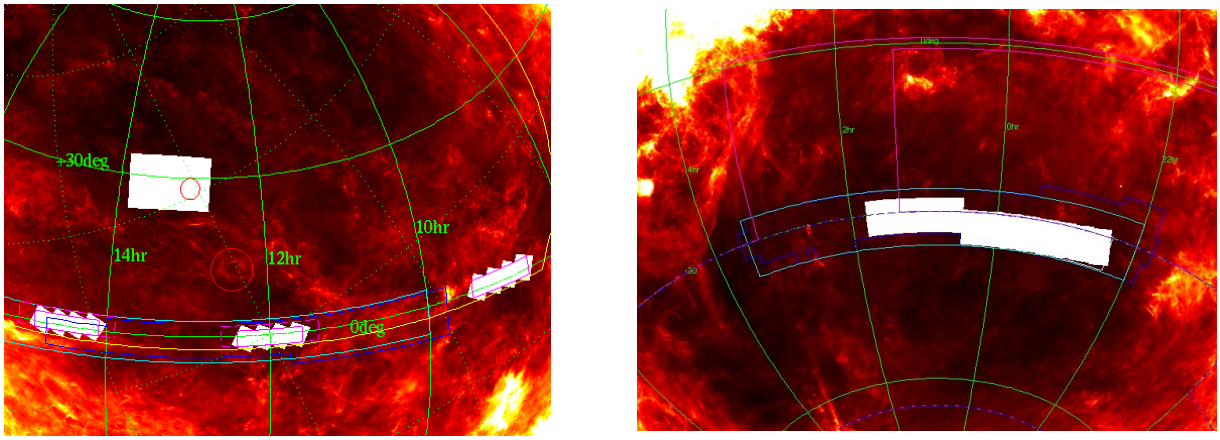

Figure 3.3: Fields surveyed by H-ATLAS, superimposed on the IRAS 100μm sky map tracing the galactic dust distribution. Left panel: NGP and equatorial fields. Right panel: SGP field (Eales et al., 2010).

Herschel (Pilbratt et al., 2010) was a space observatory born with the clear objective of providing a large-area survey of the sky in the submillimeter band, well above the ~ 10 deg^2 covered by BLAST or the very modest 0.25 deg^2 and 125 arcmin^2 by the Hubble Deep Field North (HDF-N; Borys et al., 2003) SCUBA map and the SCUBA HAlf-DEgree Extragalactic Survery (SHADES; Mortier et al., 2005), respectively. Launched in 2009, the mission was designed to extend the work of earlier infrared telescopes like IRAS, ISO, *Spitzer* or AKARI (Murakami et al., 2007) and provide a wider spectral coverage further into the infrared and submillimeter bands to analyze the wave band from 60 to 500 μm, which, at the time, had only been modestly explored in a large-area survey by BLAST. *Herschel* was equipped with three instruments which complemented each other, namely the Photodector Array Camera and Spectrometer (PACS; Poglitsch et al., 2010), the Spectral and Photometric Imaging REceiver (SPIRE; Griffin et al., 2010) and the Heterodyne Instrument for the Far Infrared (HIFI; de Graauw et al., 2010), providing photometric data in six bands with central wavelengths of 70,

100, 160, 250, 350 and 500 μm. It had a much better resolution and sensitivity than its predecessors ($\sim 18''$ at 250 μm) and the spectral coverage of SPIRE was well suited to detect the cold dust that previous observatories were insensitive to. *Herschel* conducted two large-area submillimeter surveys using SPIRE and PACS, namely the *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al., 2012) and the *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al., 2010), the latter covering the largest area, $\sim 660\,\mathrm{deg}^2$.

The H-ATLAS survey was assigned 600 hours of observing time in the 100, 160, 250, 350 and 500 μm photometric bands. Given the incredible success in understanding the distribution of nearby galaxies of large optical spectroscopic surveys of the local Universe at the time, like 2dF or SDSS, the necessity of a complementary submillimeter survey was imperative to address the dust phase of the interstellar medium and dust-obscured star formation. It was also clear at the time that the SPIRE spectral coverage would outperform upcoming local-Universe ground-based submillimeter surveys like SCUBA-2 (Holland et al., 2013) and LABOCA, since these would operate mostly at 850 μm, where nearby galaxies are intrinsically faint. The *original* goal of H-ATLAS was therefore to provide a survey over a very large area of the sky in order to measure the dust content and the dust-obscured star formation for tens of thousands of *local* galaxies. Interestingly, a signification fraction of the sources detected by H-ATLAS was found to lie at high redshifts, as shown, for instance, by the early works of Amblard et al. (2010), Lapi et al. (2011), González-Nuevo et al. (2012) and Pearson et al. (2013). As pointed out by Valiante et al. (2016), this is mainly due to the high sensitivity of *Herschel* as well as to the large and negative k-correction of star-forming galaxies in the submillimeter waveband (Blain and Longair, 1993).

The fields surveyed by H-ATLAS were chosen to minimize the dust emission in the Milky Way (as determined from the IRAS 100 μm maps) but, most importantly, with a view to maximizing the complementary data obtained in other wavelengths. Five fields were selected, one close to the North Galactic Pole (NGP), three on the celestial equator and two near the South Galactic Pole (SGP). In particular, and as shown by Figure 3.3, the NGP field is a rectangular block of ~ 180 deg^2 centered at right ascension and declination coordinates of 13^{h} 18^{m} and $+29°$ 13' (J200). This region was entirely surveyed by SDSS. The three fields along the equator are located at right ascension coordinates of 9^{h}, 12^{h} and 14.5^{h}, amounting to a total area of ~ 160 deg^2. These equatorial fields were observed by SDSS, 2dF and, most importantly, by the Galaxy And Mass Assembly (GAMA; Driver et al., 2009) spectroscopic survey, much deeper than SDSS, which is why they will be referred to as the G09, G12 and G15 fields throughout this thesis. Lastly, the SGP field is a region of ~ 318 deg^2 centered at right ascension and declination coordinates of 0^{h} 6^{m} and -32° 44' (J2000). It was entirely covered by the 2dF survey. Figure 3.4 shows the SPIRE maps of the SGP field; more precisely, the full 250 μm map, the number of scans and individual 250, 350 and 500 μm maps of a blown-up region of the entire field.

It should be mentioned that, on top of the aforementioned photometric data, there is a sizeable amount of imaging data in other wavelengths, since these fields were also observed by the GALaxy evolution EXplorer (GALEX; Martin et al., 2005) in the FUV and NUV filters, the Wide-field Infrared Survey Explorer (WISE; Wright et al., 2010) in the W1, W2, W3 and W4 filters, the VISTA Kilo-Degree Infrared Galaxy Survey (VIKING; Edge et al., 2013) the UKIRT Infrared Deep Sky Survey Large Area Survey (UKIDSS-LAS; Lawrence et al., 2007) in the near-infrared Z, Y, J, H and K_S filters, and the VST Kilo-Degree Survey (KiDS; de Jong et al., 2013) in the u, g, r and i optical bands.

The advent of the H-ATLAS survey provided the data needed to strongly support the claim that submillimeter galaxies constituted an optimal sample for gravitational lensing studies. For instance,

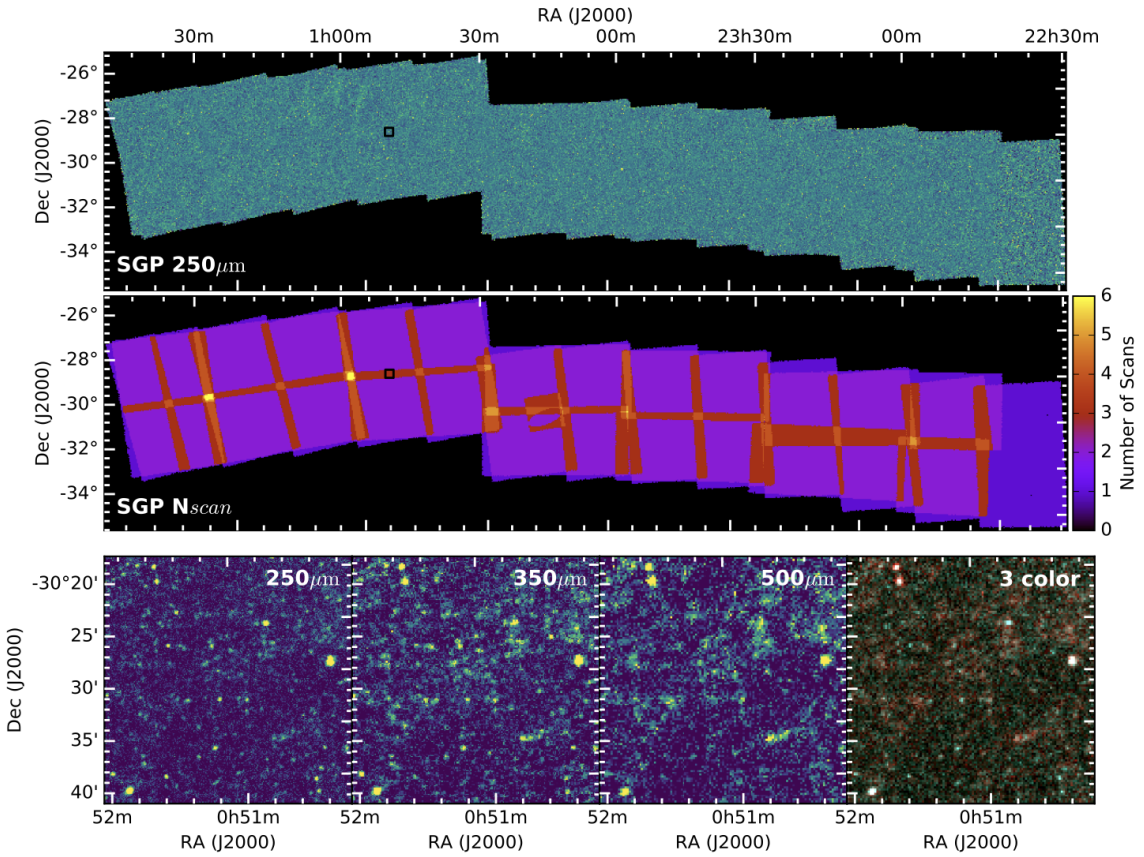

Figure 3.4: SPIRE map of the SGP field (Smith et al., 2017). The two top panels depict the 250 μm and the number of observations. The bottom panels show SPIRE images at 250, 350 and 500 μm of the blown-up region of the top panels, as well as a 3-color image combining all three wavelengths.

their very steep (unlensed) number counts had already been observed in galaxies detected by SCUBA at 850 μm (Coppin et al., 2006) or by BLAST at 250, 350 and 500 μm (Patanchon et al., 2009), but this was confirmed with H-ATLAS galaxies and given a clearer physical meaning. Lapi et al. (2011) showed that the Granato et al. (2004) model for the evolution of spheroidal galaxies (further developed by Lapi et al., 2006) provided a good fit to the number counts at the wavelengths observed by H-ATLAS. Submillimeter galaxies are thus mainly interpreted as proto-spheroidal galaxies, that is, elliptical galaxies and bulges of disk galaxies in the process of forming most of their stellar mass. This is shown by Lapi et al. (2011), since the submillimeter counts are seen to dominate the number counts over a limited flux range, as shown by Figure 3.5. This key property allowed Negrello et al. (2010) to suggest a simple but effective method to identify strongly-lensed submillimeter galaxies above 100 mJy at 500 μm, which would later be refined by González-Nuevo et al. (2012) for fainter flux densities. Moreover, and as already mentioned, submillimeter galaxies detected by H-ATLAS were shown to predominantly lie at high redshifts, with median values of $z_{\rm med} \sim 2.0 - 2.5$ (Lapi et al., 2011; González-Nuevo et al., 2012). Lastly, since most H-ATLAS lenses are typically elliptical galaxies with negligible emission in the submillimeter (Auger et al., 2009; Negrello et al., 2010; Treu et al., 2011), there is not a significant cross-contamination between lenses and background sources.

All these properties have been exploited for (strong) gravitational lensing studies in the submillimeter band in subsequent works making use H-ATLAS, among which we can highlight Bussmann et al. (2012), Bussmann et al. (2013), Fu et al. (2012), Wardlow et al. (2013), Calanog et al. (2014), Nayyeri et al. (2016), González-Nuevo et al. (2019) and Bakx et al. (2020), along with the papers

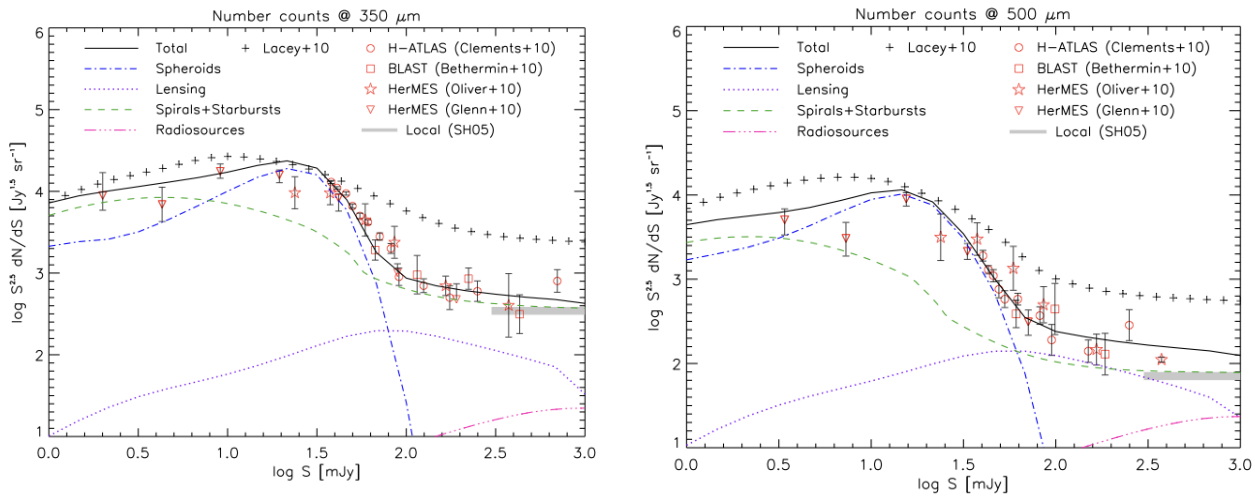

Figure 3.5: (Lapi et al., 2011) Observed Euclidean-normalized counts at 350 (left) and 500 (right) μm (Clements et al., 2010; Oliver et al., 2010; Glenn et al., 2010; Béthermin et al., 2010) with the predictions of the full model (solid black line) of Lapi et al. (2011), which updates that of Lapi et al. (2006) and contains contributions of several galaxy populations. Note the dominant contribution of spheroids over a limited flux range.

that this thesis is based upon. During the strong lensing analysis by González-Nuevo et al. (2012), a serendipitous discovery was made: high-redshift H-ATLAS galaxies tended to lie along the line of sight of foreground sources. While this had already been observed to a certain point with the first submillimeter sources detected by SCUBA (Almaini et al., 2005) and magnification bias had been suggested as a plausible explanation for this effect, the necessity for deeper and wide-area submillimeter surveys increasing the background source density was deemed imperative (Blake et al., 2006) for further studies at the time, as emphasized in the introduction. This condition was clearly fulfilled by the *Herschel* Space Observatory.

The reader may consult the details of the H-ATLAS map-making, source extraction and catalog generation process in Ibar et al. (2010), Pascale et al. (2011), Rigby et al. (2011), Valiante et al. (2016), Bourne et al. (2016) and Maddox and Dunne (2020). For our purposes, it is sufficient to mention the 1σ SPIRE noise for source detection, which includes both confusion and instrumental noise. For the equatorial regions, it is 7.4, 9.4 and 10.2 mJy per beam at 250, 350 and 500 μm, respectively. For the NGP and SGP fields, the corresponding values are 11.0, 11.1 and 12.3 mJy per beam at the same wavelengths, although they reach lower values in given regions.

3.2 The foreground and background galaxy samples

The background and foreground galaxy samples used throughout this thesis have been extracted from the H-ATLAS (Valiante et al., 2016; Smith et al., 2017) and GAMA II (Baldry et al., 2010; Driver et al., 2011; Baldry et al., 2014; Liske et al., 2015) catalogs, respectively, except for an additional foreground sample used in 4.3, obtained from the 16th data release of SDSS (Blanton et al., 2017; Ahumada et al., 2020)). They will be described in detail in this section.

The background sample for 4.1, 4.2, 5.1 and 5.2 , that gather the results from Bonavera et al. (2020), Cueli et al. (2021), Bonavera et al. (2021) and Cueli et al. (2022) is made up of all H-ATLAS sources that overlap with the fields observed by GAMA, which will be described later. The main selection criterion is a 4σ detection limit at 250 μm, so that $S_{250} > 29$ mJy, but an additional 3σ detection limit at 350 μm was introduced to improve the robustness in the estimation of photometric redshifts. Indeed, this last criterion effectively introduces a minimum redshift in the sample, since sufficiently local submillimeter galaxies must satisfy $S_{250} > S_{350}$. The photometry from the three SPIRE bands (250, 350 and 500 μm) and the (mostly upper-limit) data at 100 and 160 μm from PACS are fit to the SED template of SMM J2135-0102 ("the Cosmic Eyelash"), a gravitationally-lensed submillimeter galaxy at $z = 2.3$ (Ivison et al., 2010; Swinbank et al., 2010, see Figure 3.6). As shown by Lapi et al. (2011), González-Nuevo et al. (2012) and Ivison et al. (2016) using the few spectroscopic redshifts available and a set of different SED templates, this SED provides the best overall fit to the data, with a median value for $\Delta z/(1+z) \equiv (z_{\rm ph}-z_{\rm sp})/(1+z_{\rm sp})$ of -0.07, a dispersion of 0.153 and no outliers. This allows us to perform a photometric redshift selection of $1.2 < z < 4.0$ to ensure that the background sample of galaxies does not overlap with the lenses. In the end, we are left with ~ 58000 background galaxies, $\sim 24\%$ of the initial sample.

However, this "fiducial" redshift distribution is affected by the selection of background galaxies with photometric redshifts between 1.2 and 4.0, since, given the errors in the estimates, there is a nonzero probability that sources are left out of or mistakengly introduced in the final redshift distribution. Following Bianchini et al. (2016) and González-Nuevo et al. (2017) (both motivated by Budavári et al., 2003), the redshift distribution of background galaxies selected by our photometric window function is estimated as

$$\mathcal{P}(z|W) = \mathcal{P}(z) \int dz_{\rm ph}\, W(z_{\rm ph}) \mathcal{P}(z_{\rm ph}|z), \tag{3.1}$$

where

$$W(z_{\rm ph}) = \begin{cases} 1 & \text{if} \quad z_{\rm ph} \in [1.2, 4.0] \\ 0 & \text{if} \quad z_{\rm ph} \notin [1.2, 4.0] \end{cases} \tag{3.2}$$

is the window function, $\mathcal{P}(z)$ is the fiducial redshift distribution and $\mathcal{P}(z_{\rm ph}|z)$ is the error function, parametrized as a Gaussian distribution with zero mean and dispersion $(1+z)\sigma_{\Delta z/(1+z)}$, for which we adopt the value of $\sigma_{\Delta z/(1+z)} = 0.153$ found by Ivison et al. (2016). Figure 3.7 (in green) shows the final background redshift distribution after the window function is applied. As expected, the main effect of taking into account the errors in the estimation of photometric redshifts is to broaden the distribution beyond the selection limits. A small overlap around $z \sim 0.75$ appears, although it only comprises about 0.34 % of the foreground sources and is deemed statistically negligible (see Lapi et al., 2011; González-Nuevo et al., 2012, 2014, 2017, for a more detailed analysis).

It should be noted that the background sample for 4.3, which gathers the results from González-Nuevo et al. (2021) is, however, different. Although the flux density criteria and the process of pho-

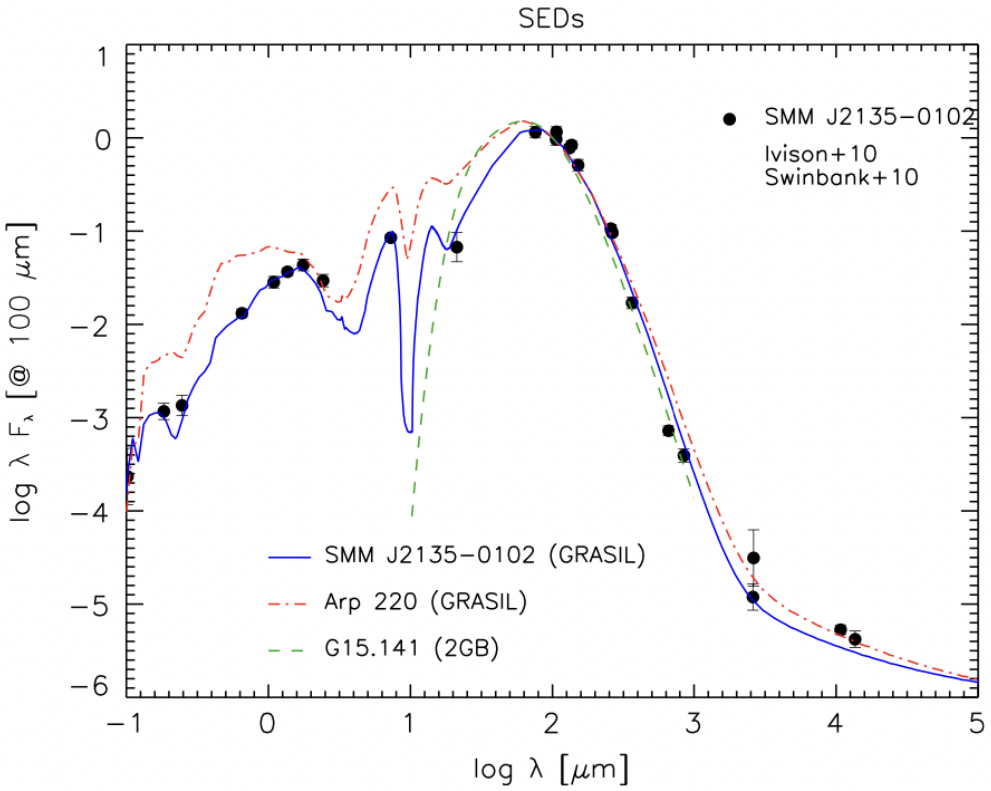

Figure 3.6: Rest-frame SED of SMM J2135-0102 ("The Cosmic Eyelash"), as modeled by the spectrophotometric code GRASIL (blue solid line; Lapi et al., 2011). The corresponding data from Ivison et al. (2010), Swinbank et al. (2010) and the SEDs of the local Arp220 and $z = 4.24$ G15.141 (Cox et al., 2011) galaxies are also shown for comparison.

tometric redshift estimation remain the same, the initial sample of background galaxies is made up of all H-ATLAS galaxies that overlap with either the GAMA or the SDSS surveys. In principle, we have two different background samples, since two different cross-correlation functions are measured (H-ATLAS/GAMA and H-ATLAS/SDSS), but the redshift distributions remain practically the same, which is the reason why they are not depicted separately in Figure 3.7.

The foreground galaxy sample for 4.1, 4.2, 5.1 and 5.2 was extracted from the GAMA II spectroscopic survey, as is the case of one of the two foreground samples in 4.3. The motivation behind this choice is twofold; firstly, and as stressed by González-Nuevo et al. (2014), gravitational lensing analyses require accurate redshift measurements for the foreground sources, especially if the goal is to perform a tomographic analysis. Secondly, the GAMA survey was coordinated with H-ATLAS to maximize the common area. Indeed, both surveys observed the three equatorial regions (G09, G12 and G15), although the SGP field was only partially surveyed by GAMA. Due to the different scanning strategies, the overlap is not perfect, resulting in a total common area of ~ 207 deg^2 (~ 147 deg^2 for the equatorial fields and ~ 60 deg^2 for SGP), surveyed down to a limit of $r \approx 19.8$ mag. For visual purposes, 3.8 shows an image of H-ATLAS and GAMA sources in a region of the G09 field, where one can subtly observe how empty regions of the foreground map tend to correspond to voids in the background map.

The GAMA foreground sample for the non-tomographic analyses (4.1, 4.2 and 4.3) is made up of $\sim$ 150.000 galaxies with reliable redshift measurements between 0.2 and 0.8. Although we do not expect an important lensing effect from the matter distribution at $z < 0.2$ (Lapi et al., 2012), this number

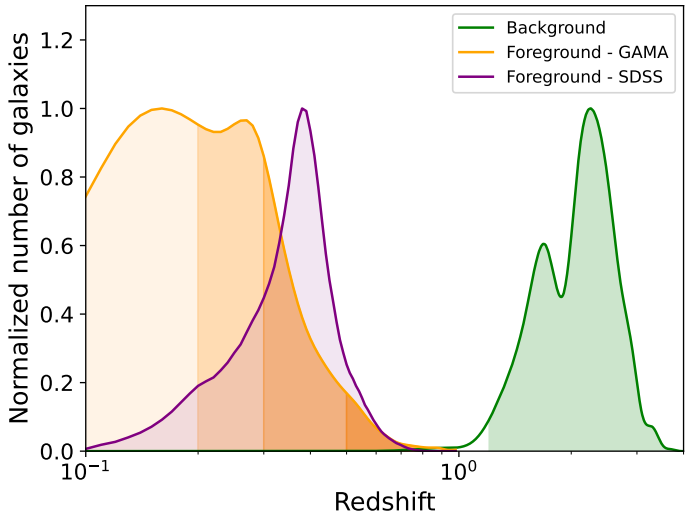

Figure 3.7: Normalized redshift distribution of the background (in green) and foreground (in orange and purple for the GAMA and SDSS data set, respectively) samples of galaxies. The shaded regions represent the selected redshift ranges: $1.2 < z < 4.0$ for the background (in green) and $0.1 < z < 0.2$, $0.2 < z < 0.3$, $0.3 < z < 0.5$ and $0.5 < z < 0.8$ for the four different foreground bins (in orange).

is increased to $\sim$ 225.000 in 5.1 and 5.2 to include potentially useful information at $0.1 < z < 0.2$. For these tomographic analyses, the foreground sample was divided into four redshift bins: $0.1 - 0.2$, $0.2 - 0.3$, $0.3 - 0.5$ and $0.5 - 0.8$, with a similar number of sources in each of them. The redshift distribution for these GAMA sources is shown in Figure 3.7 (in orange). It should be stressed once again that these galaxies have spectroscopic redshifts, which eliminates any potential misplacement of a source into the wrong redshift bin.

As already mentioned at the beginning of this section, in 4.3 we used a second foreground sample

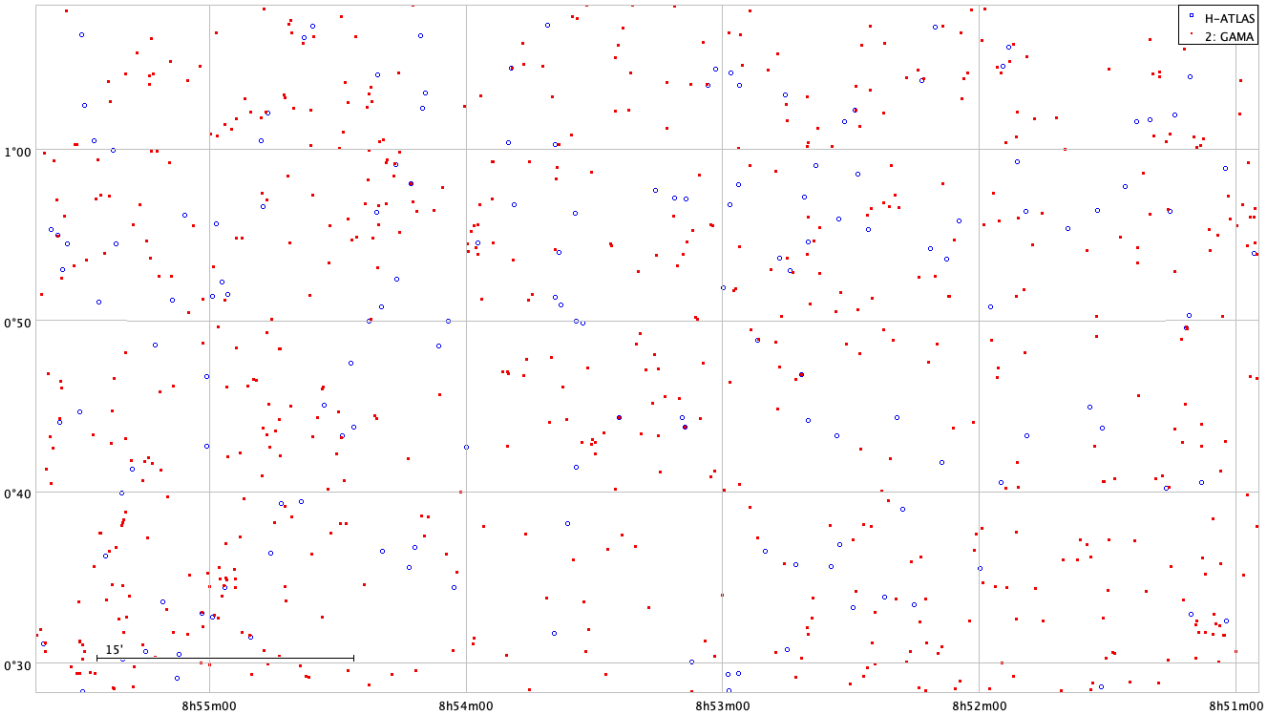

Figure 3.8: Image of H-ATLAS and GAMA sources in a region of the G09 field.

from SDSS in order to assess the effect of a larger number density of potential lenses. The downside is, of course, that these sources have photometric redshifts, although this was not deemed important since we are working with one wide redshift bin in a non-tomographic analysis. SDSS surveyed the H-ATLAS equatorial fields and the NGP region, resulting in a common area of ~ 317 deg^2 (note that the SGP field was not observed by SDSS). This second foreground sample comprises ~ 962.000 galaxies, namely those with a photometric redshift between 0.2 and 0.8 and an error of $\Delta z/(1+z) < 1$. The final redshift distribution of SDSS lenses, taking into account the errors in the estimation of photometric redshifts as in the background sample, is shown in purple in Figure 3.7.

3.3 Measurement of the cross-correlation function

This section describes the methodology followed for the measurement of the angular cross-correlation function. Given the H-ATLAS scanning strategy, the fields have been surveyed with either the characteristic diamond shape shown in Figure 3.9 or with an approximately rectangular shape. Following González-Nuevo et al. (2014), the underlying idea of cross-correlation measurements is to count the number of foreground-background pairs within a given angular separation and comparing it with the case of randomly distributed background sources, averaging over a given number of independent minimal areas. Therefore, in principle, we have the freedom to choose the full fields, tile-like regions of 4×4 deg^2 (shown in red in Figure 3.9) or "minitiles" of 2×2 deg^2 (shown in blue) as these minimal areas. Each choice has its advantages and drawbacks, which will be described in this section and were the main goal of study in the discussion laid out in 4.3. It should be noted that this procedure is not the same as the one adopted by González-Nuevo et al. (2017), which divided the full common fields of the foreground and background samples in circular regions with a given maximum overlap, which had the disadvantage of leaving portions of these regions without sources to deal with border effects.

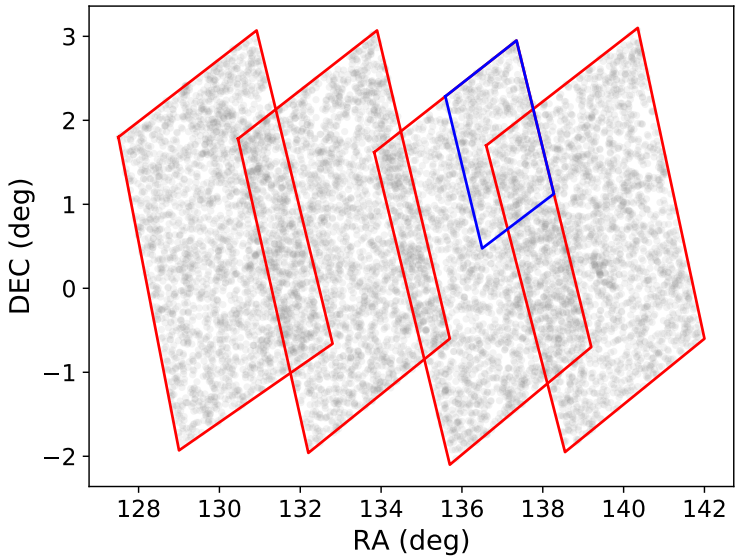

Figure 3.9: Representation of the tiles (in red) and minitiles (in blue) schemes for the G09 field. The pattern is analogous for the rest of equatorial fields.

For each minimal area, the angular cross-correlation between the full background and foreground redshift bin i was computed through a modified version of the Landy and Szalay (1993) estimator (Herranz, 2001):

$$\tilde{w}^i(\theta) = \frac{\mathrm{D}_\mathrm{f}^i\mathrm{D}_\mathrm{b}(\theta) - \mathrm{D}_\mathrm{f}^i\mathrm{R}_\mathrm{b}(\theta) - \mathrm{D}_\mathrm{b}\mathrm{R}_\mathrm{f}^i(\theta) + \mathrm{R}_\mathrm{f}^i\mathrm{R}_\mathrm{b}(\theta)}{\mathrm{R}_\mathrm{f}^i\mathrm{R}_\mathrm{b}(\theta)}, \tag{3.3}$$

where each of the terms are normalized foreground-background pair counts of galaxies (see Blake et al., 2006, for a discussion on different estimators for the cross-correlation function). For instance, $\mathrm{D}_\mathrm{f}^i\mathrm{D}_\mathrm{b}(\theta)$ is the normalized number of foreground-background galaxy pairs in the data within an angular separation of θ (effectively computed as belonging to the angular bin $\theta - \Delta\theta/2, \theta + \Delta\theta/2$) taking into account only bin i. Likewise, $\mathrm{D}_\mathrm{f}^i\mathrm{R}_\mathrm{b}(\theta)$ is the normalized number of galaxy pairs within angular distance θ, where the background sources are taken from a random mock catalog and the foreground galaxies are taken from bin i in the data.

For each minimal area, we computed the angular cross-correlation function and the statistical error coming from ten different random catalog realizations. The final measurement of the cross-correlation at a given angular separation is the average over all minimal areas, with an uncertainty corresponding to the standard error of the mean, $\sigma_\mu = \sigma/\sqrt{n}$, where σ is the standard deviation of the complete sample and n is the number of independent minimal areas. Alternative approaches to estimate uncertainties, like jackknife methods, were tested and shown to provide equivalent results (González-Nuevo et al., 2021). As expected, cosmic variance is observed to be the dominant source of uncertainty.

Regarding the different choices of minimal areas, it is clear that they must affect the results. Indeed, given a field of finite area, one must take into account the variation in the number of galaxies with respect to a completely homogeneous distribution when constructing random catalogs for a given field. Otherwise, the DR, RD and RR terms in (3.3) could introduce a bias in the cross-correlation function, an effect which is seen to be most important at large scales. There are known to be two main sources of a large-scale bias to this respect, namely the integral constraint and surface density variations on the galaxy samples, which we proceed to describe.

Given a certain field, the number of detected galaxies in it will certainly be higher or lower than what we would have in a fair sample of the Universe, thus affecting our estimates in the random catalog of the field. Although this would induce a weaker or stronger correlation in that particular field, averaging over a large number of fields tends to introduce an artificial weakening of the observed clustering, since sources sufficiently close to edges of the corresponding field are less likely to have pairs at large distances. This effectively causes the estimated cross-correlation to be biased low by a constant (Adelberger et al., 2005), so that

$$w_{\rm ideal}(\theta) = \tilde{w}(\theta) + \mathrm{IC}. \tag{3.4}$$

The integral constraint is commonly estimated numerically summing over angular bins (Infante, 1994; Roche and Eales, 1999):

$$\mathrm{IC} \approx \frac{\sum_i \mathrm{RR}(\theta_i)\, w_{\rm ideal}(\theta_i)}{\sum_i \mathrm{RR}(\theta_i)} \tag{3.5}$$

However, there are additional sources of large-scale biases that create an opposite effect to that of the integral constraint. As argued by Blake and Wall (2002), a varying source density will spuriously enhance the cross-correlation measurement since the number of close galaxy pairs depend on the local surface density (that is, $\mathrm{DD} \propto \overline{\sigma^2}$, where σ is the surface density), while galaxy pairs in the random distribution are linked to the global average surface density ($\mathrm{RR} \propto \bar{\sigma}^2$). A non-zero variance will thus introduce a higher correlation with respect to the real value.

The treatment of surface density variations in galaxy surveys is a complicated matter, since it can affected by several factors. Regarding the varying density our background sample, and given the characteristics of H-ATLAS, a well-known bias is the instrumental noise due to the scanning strategy. Indeed, when there is an overlap between scanned regions, fainter sources are detected with respect to the rest of the field, producing a spurious overdensity. To correct this, we followed the procedure of Amvrosiadis et al. (2019) to build random catalogs. The idea is to discard any simulated background galaxy whose flux density is smaller than four times the local noise, taken as the instrumental plus the confusion noise at that position, an information that is available for this survey. This generates a new random catalog for the background sources only, which is used to estimate the $\mathrm{D_f^i R_b}(\theta)$ and $\mathrm{R_f^i R_b}(\theta)$ terms in (3.3).

As for the foreground sample, we followed a simple approach that entailed smoothing the surface

density map with a Gaussian kernel and applying the H-ATLAS survey maps to neglect border effects. This filtered density map can then be used to construct unbiased random catalogs for the foreground, meaning that simulated sources will take into account large-scale overdensities in the data. Nevertheless, a caveat arises naturally in this process: the precise value of the standard deviation for the Gaussian kernel is not known; this issue will be described in detail when our results are presented in Chapter 4.

3.4 Bayesian estimation of parameters

After the release of the extremely successful CosmoMC code (Lewis and Bridle, 2002), statistics based on Markov chain Monte Carlo (MCMC) algorithms has become standard for Bayesian parameter inference in cosmology. This section will describe these methods and explain the methodology we have followed to apply them to our data and extract statistical results[2].

Contrary to the frequentist school of thought, probability is regarded within the Bayesian framework as a measure of the degree of belief about a proposition. It is thus natural for probability in this context to be updated in light of new information. Bayes' theorem, which is an uncontroversial mathematical result easily derived from Kolmogorov's axioms of probability, is given additional meaning and is promoted as the basis for inference within Bayesian statistics. Indeed, the probability that a given hypothesis H is true given the observed data d and the *a priori* information I is given by

$$\mathcal{P}(H|d,I) = \frac{\mathcal{P}(d|H,I)\mathcal{P}(H|I)}{\mathcal{P}(d|I)}. \tag{3.6}$$

The interpretation of the above expression is in keeping with the idea of updating the prior probability distribution for the hypothesis, $\mathcal{P}(H|I)$, which describes the *a priori* probability that H is true, conditional on whatever external information I, and which represents our knowledge about H before the data are analyzed. The information provided by the data is taken into account through their sampling distribution, $\mathcal{P}(d|H,I)$, which measures how the plausibility of the hypothesis changes when new data are acquired. When considered as a function of the hypothesis only (that is, for fixed data), it is called the likelihood function and is usually written as $\mathcal{L}(H|d)$. The denominator, $\mathcal{P}(d|I)$, is the marginal likelihood (usually referred to as Bayesian evidence in cosmology), a central quantity for model comparison, but which plays the role of a normalization constant for the purposes of this thesis. The combination of these three quantities leads to the posterior probability distribution, $\mathcal{P}(H|d,I)$, which represents the new degree of belief that the hypothesis is true after the data are taken into account. Lastly, and due to mathematical simplicity, these quantities are usually converted to logarithm, so that (3.6) becomes

$$\log \mathcal{P}(H|d,I) = \log \mathcal{P}(d|H,I) + \log \mathcal{P}(H|I) - \log \mathcal{P}(d|I) \tag{3.7}$$

This formalism can be applied to the ubiquitous problem of estimating the parameters of a model given a set of data, as it will be done throughout this thesis. The theoretical model for the foreground-background cross-correlation function, $w(\theta)$, described in Chapter 2, depends on a series of physical parameters. For the purposes of the papers on which this thesis is based, the relevant parameters are those describing cosmology $(\Omega_m, \Omega_b, h, n_s, \sigma_8, \tau)$, the dark matter halo occupation distribution $(M_{\rm min}, M_1, \alpha)$ and the halo mass function (A, a, p and $\tilde{A}, B, C, \tilde{p}$ for the ST and Tinker models, respectively). Let us refer to the set of parameters to be estimated as $\{p_j\}_j$. The prior probability density functions (since we are dealing with continuous random variables) throughout the thesis will be assumed to be either uniform or Gaussian, so that

$$\mathcal{P}(p_j) = \begin{cases} \frac{1}{b_j - a_j} & \text{if} \quad p_j \in (a_j, b_j) \\ 0 & \text{if} \quad p_j \notin (a_j, b_j) \end{cases} \tag{3.8}$$

[2] A comprehensive and useful review on MCMC techniques can be found in Hogg and Foreman-Mackey (2018).

for the former and

$$\mathcal{P}(p_j) = \frac{1}{\sqrt{2\pi\sigma_j^2}} e^{-(\mu_j - p_j)^2/2\sigma_j^2} \tag{3.9}$$

for the latter. The ranges (a_j, b_j) of each parameter for the uniform case and the mean (μ_j) and standard deviation (σ_j) for the Gaussian case were chosen according to physical arguments, as will be discussed in Chapter 4, along with the delicate issue of prior selection, which will be given special attention in 4.2. Since the parameters of the model are assumed to be independent, the full prior density function is thus given by

$$\mathcal{P}(\{p_j\}_j) = \prod_j \mathcal{P}(p_j). \tag{3.10}$$

The theoretical prediction for the cross-correlation function is checked against the measurements described in 3.3, $\tilde{w}(\theta)$. The likelihood function is assumed to be Gaussian, meaning that the cross-correlation data at every angular scale θ follow a Gaussian distribution, which is justified for a large enough amount of measurements. Therefore, for a set of m measurements of the cross-correlation function at different angular values, the log-likelihood function is given by

$$\log \mathcal{L}(\{p_j\}_j | \, \theta_1, \ldots, \theta_m) = -\frac{1}{2} \sum_{i=1}^{m} \left[\log\left(2\pi\sigma_i^2\right) + \frac{[w(\theta_i; \{p_j\}_j) - \tilde{w}(\theta_i)]^2}{\sigma_i^2} \right], \tag{3.11}$$

where $\tilde{w}(\theta_i)$ is the mean cross-correlation function at angular scale θ_i and σ_i is its corresponding uncertainity, computed as described in 3.3.

It should be noted that (3.11) is valid for a single-bin nontomographic case, where the the entire foreground sample is taken into account. For a tomographic analysis that considers the cross-correlation between each of the redshift bins and the background, the log-likelihood is given by

$$\log \mathcal{L}(\{p_j\}_j | \, \theta_1, \ldots, \theta_m) = -\frac{1}{2} \sum_{i=1}^{m} \sum_{k=1}^{4} \left[\log\left(2\pi(\sigma_i^k)^2\right) + \frac{[w^k(\theta_i; \{p_j\}_j) - \tilde{w}^k(\theta_i)]^2}{(\sigma_i^k)^2} \right], \tag{3.12}$$

where $\tilde{w}^k(\theta_i)$ is the estimated cross-correlation function between bin k and the background at angular scale θ_i and σ_i^k is its corresponding uncertainty. It should be commented that the cross-correlation function is not measured at exactly the same angular scales in all redshift bins (as could appear from the above equation), but we have chosen to write it in this way so as not to overcomplicate the expression. Furthermore, in writing (3.12), we have assumed statistical independence among the different bins, which is justified for two reasons: firstly, there is no possible bin overlap, as the foreground redshifts are all spectroscopic; secondly, the probability of double lensing events or of an alignment of two lenses from different bins and a background galaxy is very low (Lapi et al., 2012; Bonavera et al., 2021).

Given the fact that the theoretical model is kept unchanged throughout the thesis, the Bayesian evidence factor in (3.7) is taken to be constant and the log-posterior density function is thus expressed as

$$\log \mathcal{P}(\{p_j\}_j | \theta_1, \ldots, \theta_m) = \log \mathcal{P}(\{p_j\}_j) + \log \mathcal{L}(\{p_j\}_j | \, \theta_1, \ldots, \theta_m) + \tilde{k}, \tag{3.13}$$

where $\tilde{k}$ is a constant whose value is meaningless for our purposes.

The goal is therefore to sample the posterior probability density function given by (3.13) to obtain marginalized credible intervals for each of the parameters as well as probability contours. For instance,

the expectation value of the set of parameters $\{p_j\}_j$ (expressed as the vector $\vec{p}$), which is defined as

$$E[\vec{p}] \equiv \int \vec{p}\, \mathcal{P}(\{p_j\}_j|\theta_1, \ldots, \theta_m)\, d\vec{p} \tag{3.14}$$

can be computed as a sum over a sufficiently large number (N) of samples:

$$E[\vec{p}] \approx \frac{1}{N} \sum_i^N \vec{p}^{\,i}, \tag{3.15}$$

where $\vec{p}^{\,i}$ is the vector of parameters corresponding to sample i. MCMC methods sample the posterior distribution by generating random Markovian walks or Markov chains in parameter space so that, with a sufficiently large number of steps, a representative set of samples can be drawn from the chain. In essence, each state of the chain, $\vec{p}^{\,i}$, depends only on the previous state, $\vec{p}^{\,i-1}$, which makes it Markovian. Different MCMC algorithms are characterized by the (stochastic) process involved in deciding the next step of the chains. In this thesis we have made use of the *emcee* software package (Foreman-Mackey et al., 2013), a Python-based implementation of the Goodman and Weare (2010) so-called affine invariant MCMC ensemble sampler, which we will describe next.

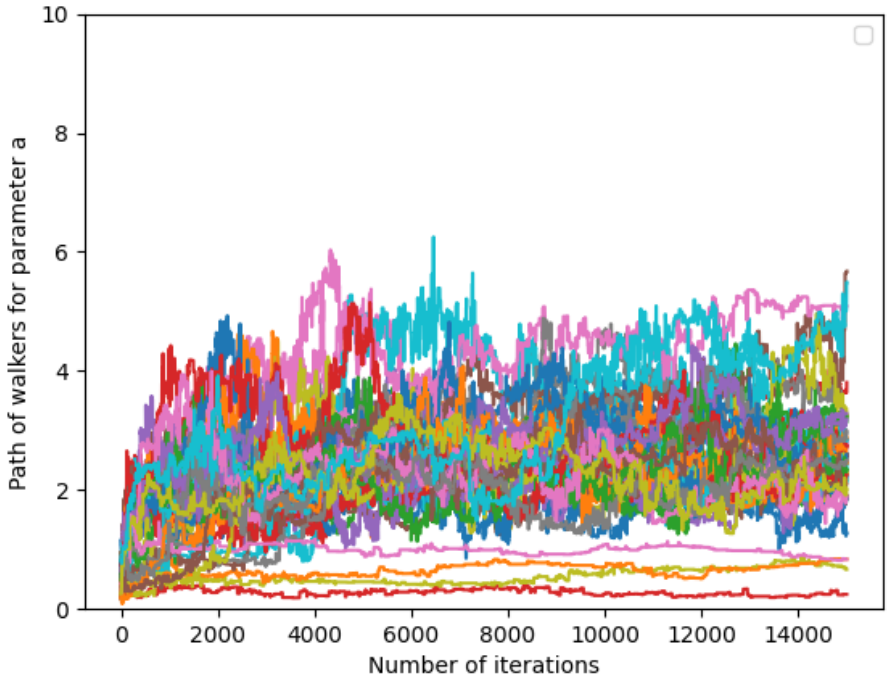

Figure 3.10: Example of how a set of walkers travels through parameter space as the number of MCMC iterations increases.

The *emcee* algorithm does not differ substantially from the simplest and best-known MCMC algorithm, Metropolis-Hastings (Metropolis et al., 1953). The latter generates a single chain, $\{\vec{p}^{\,i}\}_i$, for which, given the state at iteration step i ($\vec{p}^{\,i}$), a new proposal position ($\vec{p}^{\,\prime}$) is drawn from a so-called transition probability distribution, chosen to be easy to sample. The new position is accepted or rejected according to whether the ratio of the posterior density function at the new and old positions is larger than a random number between 0 and 1 sampled from a uniform distribution. In other words, $\vec{p}_{i+1} = \vec{p}^{\,\prime}$ if $\mathcal{P}(\vec{p}^{\,\prime}|\theta_1, \ldots, \theta_m)/\mathcal{P}(\vec{p}^{\,i}|\theta_1, \ldots, \theta_m) > r$ and $\vec{p}^{\,i+1} = \vec{p}^{\,i}$ otherwise. The *emcee* algorithm is

based on the same idea, but simultaneously evolves the state of an ensemble of K chains, $\{\vec{p}_1^{\,i}, \ldots, \vec{p}_K^{\,i}\}_i$, known as walkers, where the transition probability distribution for each walker is based on the state of the remaining $K-1$. A specific example of how a set of walkers updates their one-dimensional state as the algorithm advances is shown in Figure 3.10. Since *emcee* was shown to outperform standard Metropolis-Hastings methods, it was adopted for the entirety of the thesis.

However, a number of subtleties unavoidably arise when MCMC algorithms are used. Chains are biased random walks, in the sense that because of the local Markovian property, nearby samples in a chain are not independent. For them to be, they must be sufficiently separated. Furthermore, how many iteration steps are necessary for the algorithm to produce reliable results? While this is an extremely difficult problem for which there is no reassuring answer, Goodman and Weare (2010) recommend using the so-called integrated autocorrelation time, τ_i, to quantify the effects of sampling errors and to give an answer to the preceding questions. The estimation of this quantity is notoriously complicated, but we have followed the method by Foreman-Mackey et al. (2013) to compute it numerically. Although the details are not important for our purposes, the key concept is that τ_i is the number of iteration steps required for the chain to produce an independent sample, effectively forgetting where it began. This quantity can then be used to apply a thinning of the chain, effectively splitting it into independent samples and a burn-in, which removes the first iteration steps of the sample, which are clearly not representative of the distribution.

A last important comment should be made: *emcee* is not an optimizer, but a sampler. As such, the maximum of the full posterior distribution is not a reliable estimate of the parameters, since there could be considerable N-dimensional degeneracies and it cannot be assigned an uncertainty (more precisely, a credible region) in parameter space. The complete information is encoded into the full N-dimensional posterior distribution.

4. Non-tomographic analyses

This chapter summarizes the results from the first three papers that constitute this thesis, namely Bonavera et al. (2020), Cueli et al. (2021) and González-Nuevo et al. (2021). The common denominator is the use of the foreground-background angular cross-correlation function in a non-tomographic setup, that is, using a single broad foreground redshift bin. Nevertheless, the objectives are related but clearly distinct. Indeed, Bonavera et al. (2020) was a proof-of-concept work whose main goal was to assess the potential of the submillimeter galaxy magnification bias to constrain cosmology. While Cueli et al. (2021) had the same preliminary nature, its main goal was to provide observational constraints on the number density of dark matter halos and to study the potential influence of the HOD parameters. González-Nuevo et al. (2021), on the contrary, had a more methodological character and focused on correcting the large-scale biases that could be contaminating the cross-correlation signal and studying the potential improvements with respect to Bonavera et al. (2020).

4.1 Preliminary results on cosmology and the HOD

This first section lays out the main results from Bonavera et al. (2020), to be regarded as the starting point of the thesis. The chief goal of this work, of preliminary nature, was to probe the extent to which the phenomenon of magnification bias on submillimeter galaxies could potentially constrain some of the cosmological parameters, as well as those from the HOD, in a non-tomographic study. Furthermore, two different angular binning schemes for the measurement of the cross-correlation function were adopted, one that is considered "default" and another one that is "shifted" by half a bin in order to study the potential influence on the conclusions; the data are represented in Figure 4.1 and compared to the cross-correlation measured by González-Nuevo et al. (2017). The background and foreground galaxy samples were extracted from the H-ATLAS survey and the GAMA-II catalog, as detailed in 3.2 and the cross-correlation function was estimated using a single foreground redshift bin within the "minitiles" scheme, as described in 3.3.

To gather as much information as possible, we proceeded in two steps. Firstly, we carried out two MCMC runs (one for each binning strategy) to derive the posterior probability distributions of just the HOD parameters within a fixed flat ΛCDM[1] cosmology according to the best-fit values from *Planck* (Planck Collaboration et al., 2020a), that is,

$$\Omega_m = 0.315 \qquad \Omega_b = 0.049 \qquad h = 0.674 \qquad n_s = 0.965 \qquad \sigma_8 = 0.811, \tag{4.1}$$

with uniform prior distributions for the HOD parameters:

$$\alpha \sim \mathcal{U}\,[0.5, 1.37] \qquad \log M_{\rm min} \sim \mathcal{U}\,[11.6, 13.6] \qquad \log M_1 \sim \mathcal{U}\,[13.0, 14.5], \tag{4.2}$$

where the masses are expressed in $M_\odot/h$, as will be done throughout this thesis. The ranges of the uniform distributions were chosen according to the available literature, namely Sifón et al. (2015) and (Viola et al., 2015), since the galaxy sample they worked with was similar to ours: it covered the overlapping region between KiDS and GAMA, which has an area of ~ 100 deg^2, and their galaxies

[1]Note that the reionization optical depth, which completes the list of parameters of the minimal ΛCDM model, does not affect the matter power spectrum and is thus not considered here.

spanned a redshift range of $0 - 0.5$, with an average value of 0.25. Associating the stellar mass distribution of satellite galaxies found by Sifón et al. (2015) with the masses of the corresponding halos via the recipe of Pantoni et al. (2019) and considering the halo masses found by Viola et al. (2015), the above uniform prior ranges were fixed.

These first two MCMC runs were complemented with an additional run on the auto-correlation function of the foreground sample in order to have an independent estimation of the HOD parameters characterizing these galaxies, although this analysis has an important caveat that will be discussed further ahead.

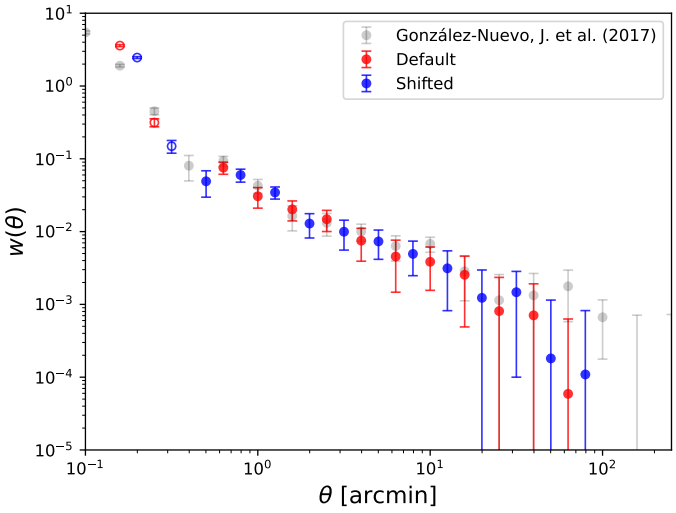

Figure 4.1: Cross-correlation data for 4.1 with the default (red) and shifted (blue) binning schemes. The data from González-Nuevo et al. (2017) are also depicted in faint gray for comparison. The empty data points were not taken into consideration since they are not associated to the weak-lensing regime.

Secondly, we aimed to include Ω_m, h and σ_8 in the MCMC analysis along with the HOD parameters with the following prior distributions:

$$\Omega_m \sim \mathcal{U}\,[0.1, 0.8] \qquad \sigma_8 \sim \mathcal{U}\,[0.6, 1.2] \qquad h \sim \mathcal{U}\,[0.5, 1.0], \tag{4.3}$$

to study the potential constraints on cosmology. The rest of cosmological parameters (n_s and Ω_b), whose influence on the cross-correlation function was not likely to be observed, were kept fixed to the aforementioned *Planck* values. Since the main goal of this paper was to study the capability of the submillimeter galaxy magnification bias to constrain cosmology with as little information as possible, we decided to choose uniform priors for Ω_m, σ_8 and h. However, to assess potential improvements, this joint analysis was carried out for two different choices of the HOD prior distributions, namely the uniform distributions described in (4.2) and the following (more restricting) Gaussian distributions,

$$\alpha \sim \mathcal{N}\,[0.92, 0.15] \qquad \log M_{\min} \sim \mathcal{N}\,[12.40, 0.10] \qquad \log M_1 \sim \mathcal{N}\,[13.95, 0.30], \tag{4.4}$$

also based on the analysis of Sifón et al. (2015) and Viola et al. (2015). These cases constitute the third, fourth, fifth and sixth[2] MCMC runs of the paper, respectively (since the two binning schemes were analyzed for each choice of HOD prior distribution).

Before diving into the results, it will prove useful (for future reference in the understanding of our

[2]Note that the MCMC run on the auto-correlation function is not being counted.

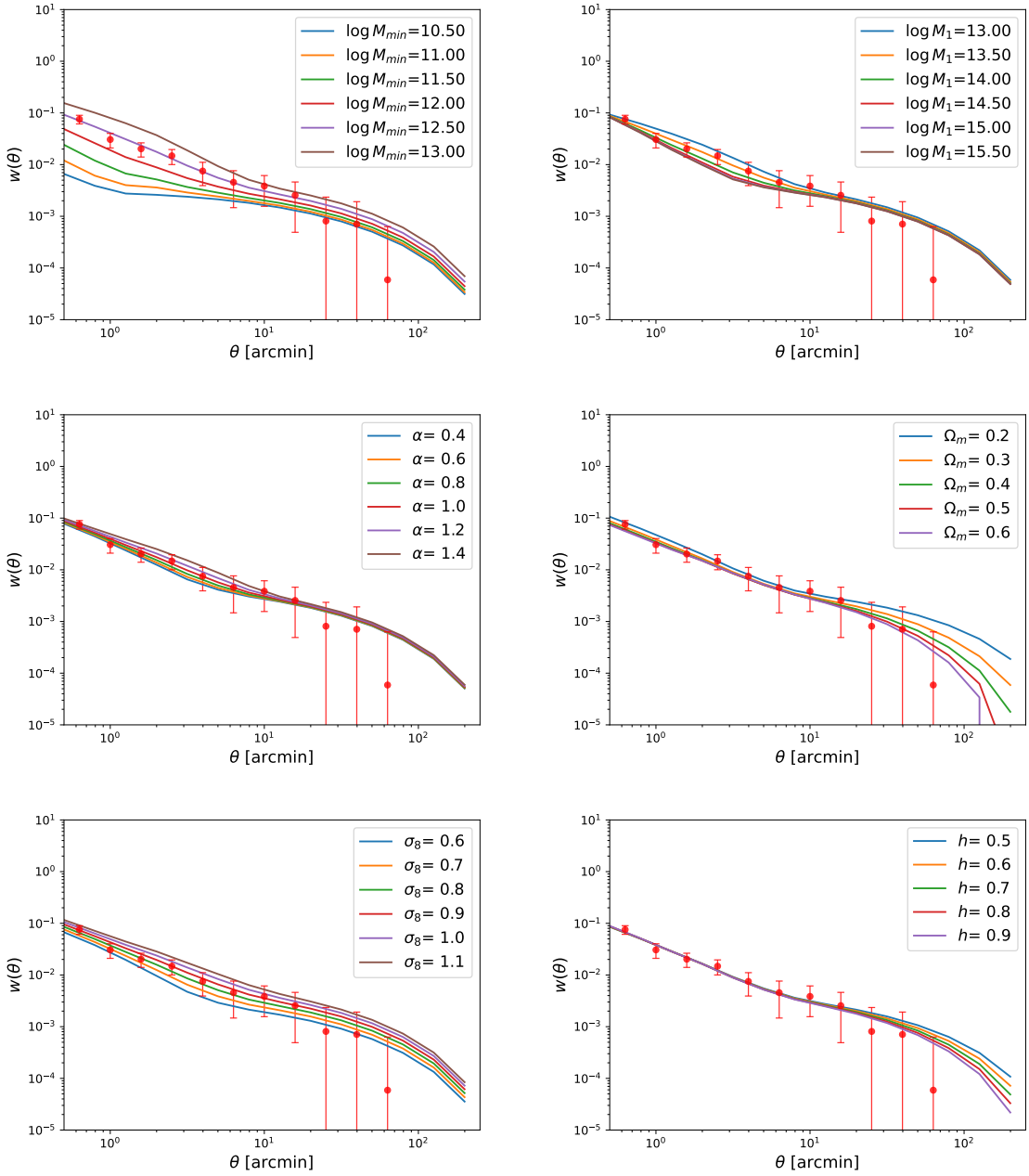

Figure 4.2: Variation of the theoretical cross-correlation function with respect to the HOD and cosmological parameters. The data are also plotted for visual purposes.

findings) to plot the the angular cross-correlation function as a function of the HOD and cosmology. Figure 4.2 shows the variation with respect to each of the six parameters, while keeping the rest fixed to the fiducial values $\alpha = 0.9$, $\log M_{\rm min} = 12.4$, $\log M_1 = 13.6$, $\Omega_m = 0.32$, $\sigma_8 = 0.81$ and $h = 0.67$. Comments will be made on this figure as the results are interpreted.

4.1.1 Fixed flat ΛCDM cosmology

For the first two MCMC runs, namely the fixed flat ΛCDM cosmology model for both data sets (default and shifted), the mean[3], mode and 68% marginalized credible intervals are summarized in Table 4.1,

[3]Throughout the thesis, a hyphen will be used in tables to highlight that the quantity in question is not statistically meaningful.

Table 4.1: Parameter prior distributions and summarized marginalized statistical results for the first and second MCMC runs of 4.1: a fixed flat ΛCDM model with uniform priors for the default and shifted data sets.

		Default			Shifted		
Parameter	Prior	Mean	Mode	68% CI	Mean	Mode	68% CI
$\log M_{\rm min}$	$\mathcal{U}[11.60, 13.60]$	12.42	12.46	$[12.31, 12.62]$	12.41	12.45	$[12.34, 12.58]$
$\log M_1$	$\mathcal{U}[13.00, 14.50]$	-	-	$[13.00, 14.50]$	13.65	13.54	$[13.11, 13.94]$
α	$\mathcal{U}[0.50, 1.37]$	-	-	$[0.50, 1.37]$	-	-	$[0.50, 1.37]$

while the marginalized posterior distributions and contour plots[4] for the HOD parameters are depicted in Figure 4.3.

The mean minimum halo mass, $M_{\rm min}$, is well constrained with the two data sets, with mean values of $\log M_{\rm min} = 12.42^{+0.20}_{-0.11}$ and $12.41^{+0.17}_{-0.07}$ at 68% credibility[5] for the default and shifted cases, respectively. These values are in agreement with the results from González-Nuevo et al. (2017) and Bonavera et al. (2019), as well as with typical values from the literature (Sifón et al., 2015; Viola et al., 2015). The α parameter, however, is equally unconstrained for both data sets. As for M_1, the behavior is somewhat different depending on the data set; for the default case, it is not possible to set any constraints on M_1, whereas a mean value of $\log M_1 = 13.65^{+0.29}_{-0.54}$ can be meaningfully found for the shifted data set.

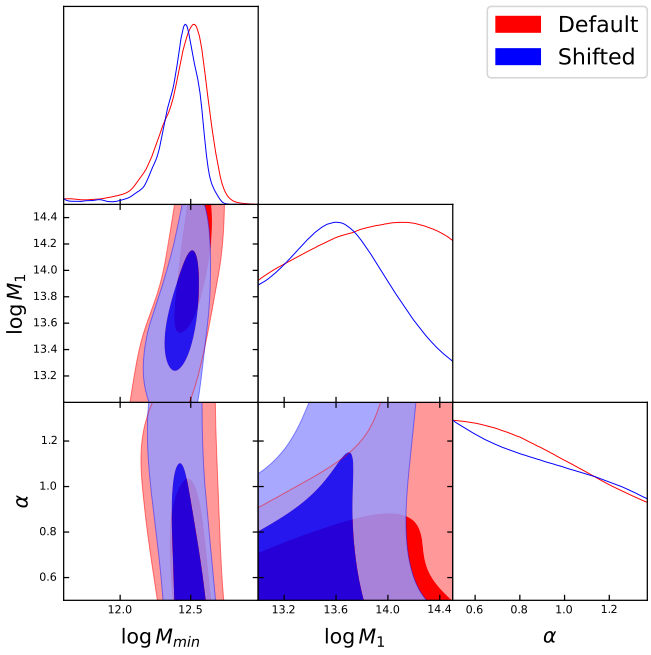

Figure 4.3: Marginalized posterior distributions and contour plots for the first and second MCMC runs of 4.1: a fixed flat ΛCDM model with uniform priors for both the default and shifted data sets.

[4]Unless otherwise stated, the probability contours are set to 0.393 and 0.865 for a direct comparison with the one-dimensional 1σ and 2σ credible intervals.

[5]Credible intervals will be 68% throughout the rest of the thesis and will be computed as highest posterior density regions, unless otherwise specified.

Table 4.2: Parameter prior distributions and summarized marginalized statistical results for the additional MCMC run of 4.1 on the auto-correlation function of the foreground sample.

Parameter	Prior	Mean	Mode	68% CI
$\log M_{\min}$	$\mathcal{U}$[9.00,15.00]	12.362	12.360	$[12.355, 12.364]$
$\log M_1$	$\mathcal{U}$[9.00,15.00]	13.64	13.56	$[13.55, 13.67]$
α	$\mathcal{U}$[0.10,1.50]	1.17	1.18	$[1.12, 1.23]$

All these aspects can be understood via the aforementioned sensitivity plots of Figure 4.2. Indeed, among the HOD parameters, $M_{\min}$ is the one to which our observable is most sensitive, which explains the tight constraints. However, for a fixed value of this parameter, the proposed prior ranges for both M_1 and α do not leave much room to adjust the model to the data, given the substantially lower variations they produce in the cross-correlation function. As for the difference in M_1 between both data sets, we believe it might be related to the first cross-correlation point at $\theta \sim 0.5$ arcmin, notably lower for the shifted case than its default counterpart (as seen in Figure 4.1).

Regarding the additional MCMC run on the angular auto-correlation function of the foreground sample, the mean, mode and 68% marginalized credible intervals are summarized in Table 4.2. The constraints on the HOD are extremely tight with respect to the cross-correlation analysis, with mean values of $\log M_{\min} = 12.362^{+0.002}_{-0.007}$, $\log M_1 = 13.64^{+0.03}_{-0.09}$ and $\alpha = 1.17^{+0.06}_{-0.05}$. This is to be expected, given the fact the the angular auto-correlation function has much smaller error bars and, most importantly, that the corresponding MCMC analysis was carried out with the additional constraint that the number density of galaxies should equal that of the sample.

One might now be tempted to use these very well-constrained HOD values for the cross-correlation analysis, but this is not correct and is related to the caveat that was mentioned at the beginning of this section. The angular cross-correlation function, as explained in 2.2.3, only probes the matter distribution that acts as a lens, while the auto-correlation function characterizes the clustering of the (foreground) sample and, therefore, probes the entire galaxy sample. As a consequence, the HOD need not coincide and should not be mixed so as not to introduce an uncontrollable bias in the analysis.

4.1.2 Semi-free flat ΛCDM cosmology

Uniform priors on the HOD parameters

The marginalized posterior distributions and contour plots for the HOD and cosmological parameters corresponding to the third and fourth MCMC runs are depicted in Figure 4.4, while the corresponding mean, mode and 68% marginalized credible intervals are summarized in Table 4.3. These results correspond to the MCMC runs on the HOD and cosmology, with uniform priors on both sets of parameters and for the two binning schemes: default and shifted, respectively. Concerning the HOD parameters, they are in perfect agreement with those obtained in the previous case, which also means that only the average minimum halo mass, $M_{\min}$, is well constrained, with a mean value of $\log M_{\min} = 12.53^{+0.29}_{-0.16}$ and $12.50^{+0.23}_{-0.17}$ for the default and shifted case, respectively. The behavior of both M_1 and α is maintained, including the better constraining of the M_1 parameter in the shifted binning case.

As regards cosmology, the results are clearly very robust against changes in the binning of the cross-

Table 4.3: Parameter prior distributions and summarized marginalized statistical results for the third and fourth MCMC runs of 4.1: a semi-free flat ΛCDM model with uniform priors for the default and shifted data sets.

		Default			Shifted		
Parameter	Prior	Mean	Mode	68% CI	Mean	Mode	68% CI
$\log M_{\min}$	$\mathcal{U}[11.60, 13.60]$	12.53	12.61	$[12.37, 12.82]$	12.50	12.53	$[12.33, 12.73]$
$\log M_1$	$\mathcal{U}[13.00, 14.50]$	-	-	$[13.00, 14.50]$	-	-	$[13.24, 14.16]$
α	$\mathcal{U}[0.50, 1.37]$	-	-	-	-	-	-
Ω_m	$\mathcal{U}[0.10, 0.80]$	0.54	-	$[0.46, 0.80]$	0.52	-	$[0.43, 0.80]$
σ_8	$\mathcal{U}[0.60, 1.20]$	0.78	0.74	$[0.63, 0.85]$	0.74	0.77	$[0.63, 0.89]$
h	$\mathcal{U}[0.50, 1.00]$	-	-	$[0.50, 1.00]$	-	-	$[0.50, 1.00]$

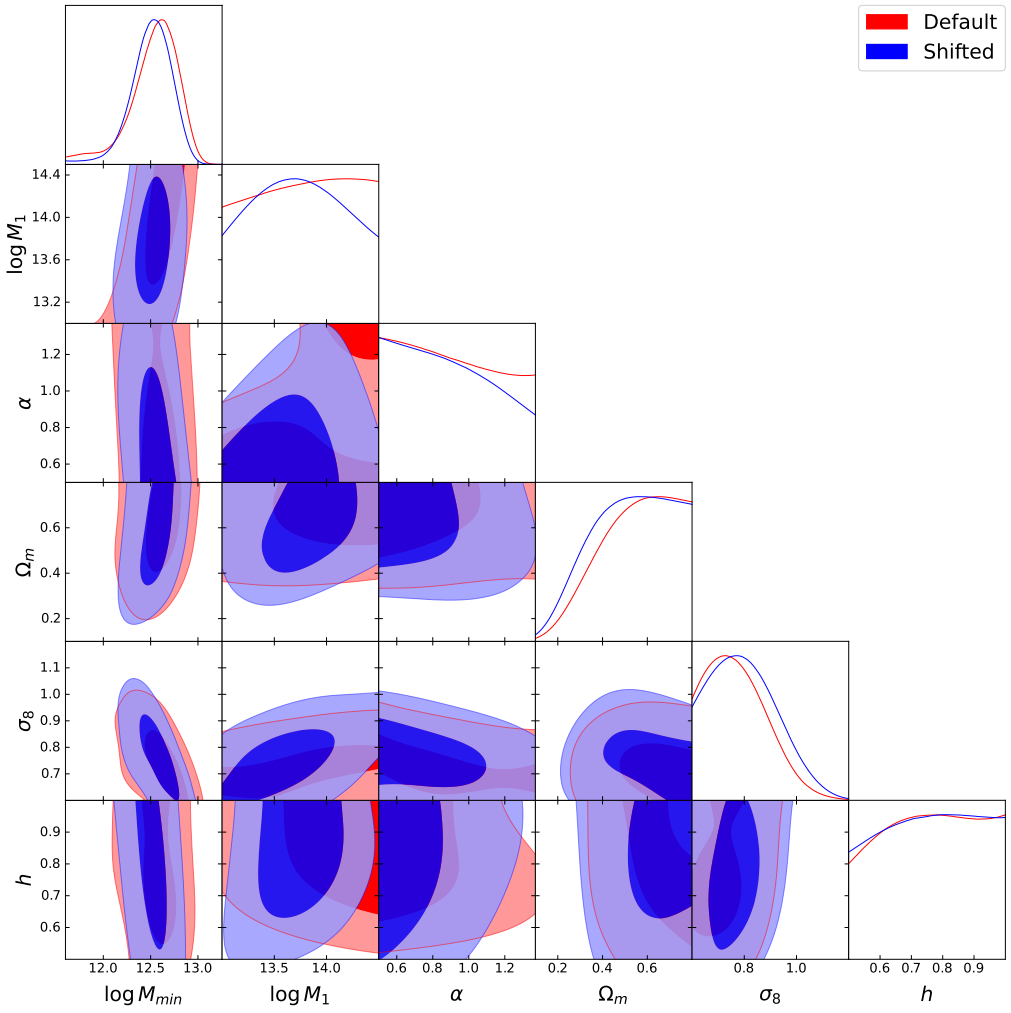

Figure 4.4: Marginalized posterior distributions and contour plots for the third and fourth MCMC runs of 4.1: a semi-free flat ΛCDM model with uniform priors for the default and shifted data sets.

correlation data. The σ_8 parameter is reasonably well-constrained, with a mean value of $\sigma_8 = 0.78^{+0.07}_{-0.15}$ and $0.80^{+0.09}_{-0.17}$ for the default and shifted case, respectively, as well as a 95% upper limit of $\sigma_8 < 0.98$ and < 1.02. This is reasonable, given the fact that σ_8 is the cosmological parameter to which the cross-correlation function is most sensitive and, most importantly, on all angular scales, as shown by Figure 4.2. On the contrary, the Ω_m parameter can only (although interestingly) be assigned a lower limit of $\Omega_m > 0.24$ at 95% credibility in the shifted-binning case, which can also be understood via the aforementioned sensitivity plot; indeed, large values for Ω_m are needed to fit the low cross-correlation signal at the largest scales, specially the last data point. However, the adimensional Hubble constant, h, is unconstrained for both data sets, since our observable seems to only be sensitive to this parameter on the largest scales, where, as of now, the uncertainties cannot be reduced further.

Gaussian priors on the HOD parameters

The marginalized posterior distributions and contour plots for the HOD and cosmological parameters are depicted in Figure 4.5, while the corresponding mean, mode and 68% marginalized credible intervals are summarized in Table 4.4.

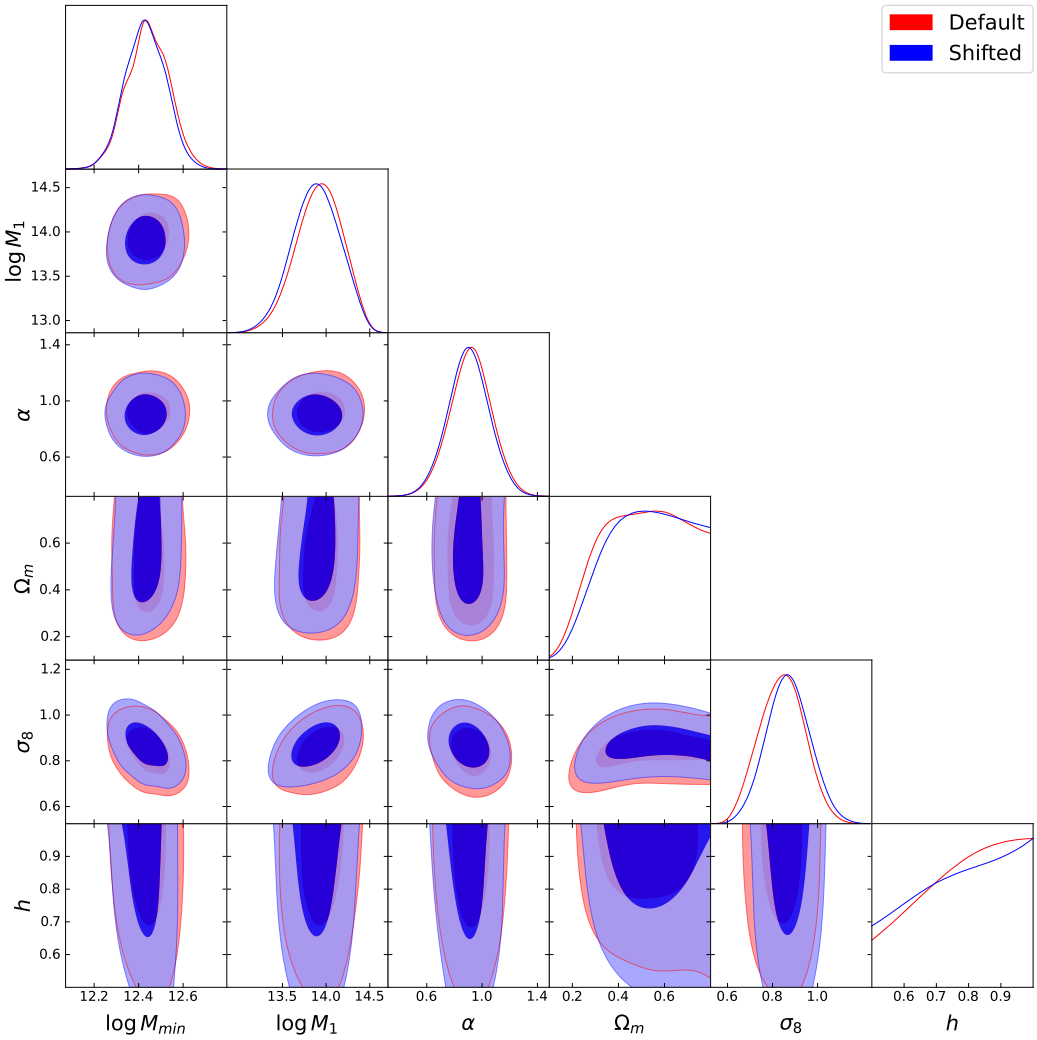

Figure 4.5: Marginalized posterior distributions and contour plots for the fifth and sixth runs of 4.1: a semi-free flat ΛCDM model with Gaussian/uniform priors on the HOD/cosmology for the default and shifted data sets.

Table 4.4: Parameter prior distributions and summarized marginalized statistical results for the fifth and sixth MCMC runs of 4.1: a semi-free flat ΛCDM model with Gaussian/uniform priors on the HOD/cosmology for the default and shifted data sets.

		Default			Shifted		
Parameter	Prior	Mean	Mode	68% CI	Mean	Mode	68% CI
$\log M_{\rm min}$	$\mathcal{N}[12.40, 0.10]$	12.44	12.43	$[12.35, 12.54]$	12.43	12.43	$[12.34, 12.52]$
$\log M_1$	$\mathcal{N}[13.95, 0.30]$	13.92	13.95	$[13.67, 14.21]$	13.89	13.89	$[13.62, 14.17]$
α	$\mathcal{N}[0.92, 0.15]$	0.92	0.92	$[0.77, 1.07$	0.90	0.90	$[0.75, 1.05]$
Ω_m	$\mathcal{U}[0.10, 0.80]$	0.51	0.57	$[0.35, 0.73]$	0.52	0.51	$[0.39, 0.75]$
σ_8	$\mathcal{U}[0.60, 1.20]$	0.84	0.86	$[0.74, 0.94]$	0.87	0.87	$[0.77, 0.97]$
h	$\mathcal{U}[0.50, 1.00]$	-	-	$[0.50, 1.00]$	-	-	$[0.50, 1.00]$

Compared to the previous case, the more constraining priors on the HOD, which practically do not deviate from their fiducial distributions, contribute slightly to the reduction of the uncertainty for the σ_8 parameter, yielding mean values of $\sigma_8 = 0.84^{+0.10}_{-0.10}$ and $0.87^{+0.10}_{-0.10}$ for the default and shifted cases, respectively. The tendency of the Ω_m parameter remains unchanged with the 95% lower limit of $\Omega_m > 0.23$ and > 0.25, while the Hubble constant is still unconstrained. We can conclude that the issue with Ω_m that was present in the previous case does not vanish with a better constraining of the HOD, which is indicative of the need of better data and/or methodology.

4.1.3 Comparison with other results and next steps

Although this preliminary analysis constituted a proof of concept, it is nonetheless informative to compare our findings with constraints from the most popular cosmological probes to identify the key aspects. In particular, we concentrated on the CMB lensing analysis from *Planck* (Planck Collaboration et al., 2020b) and the (tomographic) cosmic shear results from CFHTLens (Joudaki et al., 2017), HSC (Hamana et al., 2020), the combination of KiDS and VIKING (Hildebrandt et al., 2020) and the first-year data of DES (Troxel et al., 2018).

Let us focus on the $\Omega_m - \sigma_8$ plane. Figure 4.6 shows the contour plots from the aforementioned cosmological probes along with our results for the cases of uniform priors (left panel) and Gaussian priors (right panel) on the HOD for the default (gray) and shifted (black dotted) binning. The first aspect to be noted is the improvement in terms of uncertainty when the prior distribution on the HOD is restricted, which is indicative of the fact that better constraints on the HOD are certain to be decisive. Secondly, and although they are still preliminary, our results do not show the typical degeneracy between Ω_m and σ_8 characterizing cosmic shear results; future, more constraining results on the submillimeter magnification bias could potentially help break this degeneracy.

The cosmological constraints at this point are clearly not competitive, but improvements are certainly expected by following a series of steps. Firstly, this preliminary analysis was based on measurements of the cross-correlation function using only one (relatively large) foreground redshift bin. As

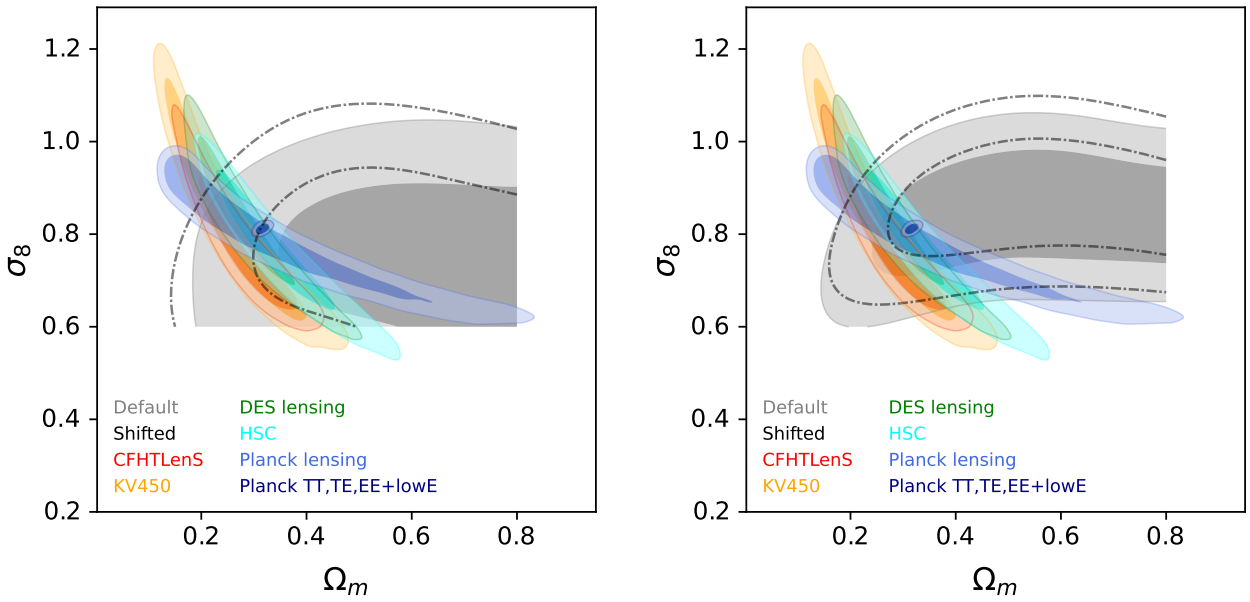

Figure 4.6: Contour plots on the $\Omega_m - \sigma_8$ plane comparing our results for the cases of uniform (left) and Gaussian (right) prior distributions for the HOD parameters with those from *Planck* (blue for lensing only and dark blue for temperature and polarization), CFHTLens (red), KV450 (orange), DES lensing (green) and HSC (cyan). The results from the default and shifted cases are depicted in grey and dotted black, respectively. The probability contours are set to 0.68 and 0.95.

demonstrated in González-Nuevo et al. (2017), a tomographic study that divides the foreground sample into several bins can be performed and is very likely to improve constraints via the joint analysis of effectively more than one observable. The obvious downside to this approach comes from the reduction of the sample size for each redshift bin, which leads to the second step.

Indeed, the statistical sample can be increased by enlarging both the lens and background source populations. Regarding the lenses, surveys like SDSS or DES are characterized by a much larger source density than GAMA, but with the disadvantage of mainly providing photometric redshift measurements, which are less accurate by nature. This is a crucial aspect for our cross-correlation study where the foreground and background samples need to be clearly separated in redshift, even more so for a tomographic setup. Future more precise catalogs of lenses might be available from experiments like the Javalambre Physics of the Accelerating Universe Astrophysical Survey (J-PAS; Benitez et al., 2014) or *Euclid*. As for background sources, the possibility of considering the catalog of the whole area covered by *Herschel*, that is, the *Herschel* Extragalactic Legacy Project (HELP; Shirley et al., 2019) or the *Herschel*/SPIRE Point Source Catalog (HSPSC; Schulz et al., 2017) is likely to provide a larger number of submillimeter galaxies.

4.2 Observational constraints on the HMF

This section discusses the results from Cueli et al. (2021), whose main goal was to study the potential of the submillimeter galaxy magnification bias to derive observational measurements of the number density of dark matter halos and to assess the influence of the HOD parameters on these constraints. To do so, we performed a non-tomographic analysis using the same background and foreground galaxy samples as in 4.1, although the cross-correlation function was estimated with the "tiles" scheme[6].

This analysis was carried out by adopting two well-known HMF models, namely the ST and Tinker fits, already described in 2.3.1, although it could certainly be applied to any other. For each of the two cases, we decided to study the influence of the HOD distribution by first estimating the HMF parameters with fixed HOD values corresponding to

$$\alpha = 0.9 \qquad \log M_{\rm min} = 12.4 \qquad \log M_1 = 13.6, \tag{4.5}$$

as motivated by 4.1, and then redoing the analysis assuming Gaussian priors based on the literature, that is,

$$\alpha \sim \mathcal{N}\,[0.92, 0.15] \quad \log M_{\rm min} \sim \mathcal{N}\,[12.40, 0.10] \quad \log M_1 \sim \mathcal{N}\,[13.95, 0.30], \tag{4.6}$$

as in 4.1.

In all four MCMC runs, we assumed uniform priors for the HMF parameters. Contrary to the HOD (and to the situation in the previous section), however, the choice of the prior distributions was not at all straightforward. Although it should always be regarded as a delicate matter, the limits of (uniform) prior distributions are relatively simple to set when the parameters to be estimated have a clear physical meaning (for instance, the baryon density parameter must satisfy $0 < \Omega_b < \Omega_m$). This is not the case for the HMF parameters. If we recall the expressions for the ST and Tinker fits,

$$f_{\rm ST}(\nu; a, p) = A(p)\sqrt{\frac{a\nu}{2\pi}}\left[1 + \left(\frac{1}{a\nu}\right)^p\right]e^{-a\nu/2} \qquad f_{\rm T}(\nu; \tilde{A}, B, C, \tilde{p}) = \tilde{A}\big[1 + \big(B\sqrt{\nu}\big)^{\tilde{p}}\big]e^{-C\nu}, \tag{4.7}$$

mathematical (and physical) consistency demands that

$$A > 0 \quad a > 0 \quad \tilde{A} > 0 \quad B > 0 \quad C > 0, \tag{4.8}$$

but we have no information as to how large these parameters can be. Furthermore, p and $\tilde{p}$ can, in principle, have any sign. On the one hand, too strict a prior could bias the results but, on the other hand, too wide a range could waste computational time as well as explore meaningless values for the parameters. In the literature, the HMF parameters are usually found by means of an optimizer that finds the best fit of a model to numerical simulations through a χ^2 analysis, but no information is usually provided regarding the ranges that are explored. In particular, we have found no mention to the potential non-positivity of p or $\tilde{p}$ in HMF-related works. Consequently, even though simulation-based fits normally yield values of $p \approx 0.3$, we do no think there is a physically motivated reason to exclude negative values from its priors. Nevertheless, we decided to assume all parameters were positive for the aforementioned cases, which are to be considered the main runs of this work. The corresponding

[6]The reason behind this choice is the fact that this was thought to be the optimal method at the time; we will learn more about this further ahead.

prior distributions were

$$a \sim \mathcal{U}[0, 10] \qquad p \sim \mathcal{U}[0, 0.5]$$
$$\tilde{A} \sim \mathcal{U}[0, 5] \qquad B \sim \mathcal{U}[0, 5] \qquad C \sim \mathcal{U}[0, 5] \qquad \tilde{p} \sim \mathcal{U}[0, 5]. \tag{4.9}$$

However, motivated by the results of these runs as well as by the above discussion, three additional cases were studied for the ST fit, all of them with a fixed HOD for simplicity. Firstly, we considered the possibility of a free normalization parameter A, instead of the function $A(p)$, effectively going from a two-parameter to a three-parameter fit. This corresponds physically to relaxing the assumption that all matter should be contained within dark matter halos and it is often done when fitting HMF models to numerical simulations. The prior distributions for the HMF in this fifth MCMC run were chosen to be

$$A \sim \mathcal{U}[0, 1] \qquad a \sim \mathcal{U}[0, 10] \qquad p \sim \mathcal{U}[0, 0.5] \tag{4.10}$$

for the sake of comparison with the main case. Note that A must still satisfy the condition

$$\int_0^\infty \frac{f(\nu)}{\nu} d\nu < 1 \tag{4.11}$$

in this three-parameter fit for the HMF to have a physical meaning.

Secondly, the possibility of allowing p to be negative was investigated, both for a two-parameter and a three-parameter ST fit, with prior distributions

$$a \sim \mathcal{U}[0, 10] \qquad p \sim \mathcal{U}[-10, 0.5] \tag{4.12}$$

for the former and

$$A \sim \mathcal{U}[0, 1] \qquad a \sim \mathcal{U}[0, 10] \qquad p \sim \mathcal{U}[-10, 10] \tag{4.13}$$

for the latter.

Lastly, with a view to constraining the HMF itself at any redshift, the advantages of the Bayesian approach were exploited and the full posterior distribution was sampled to derive median values and credible intervals for the HMF. These observational measurements were compared to the traditional values obtained with N-body simulations.

4.2.1 Main results

The Sheth-Tormen model

Table 4.5 shows the summarized statistical results from the first MCMC run, namely the two-parameter ST fit with positive p and a fixed HOD, whereas the full corner plot is depicted in Figure 4.7 (in red). While the marginalized posterior distribution of the a parameter is well-constrained with a clear mode at $a = 0.88$ and a mean value of $a = 1.29^{+0.24}_{-0.87}$, the p parameter can only be assigned upper bounds, namely $p < 0.17$ and $p < 0.31$ at 68% and 95%, respectively. Therefore, the traditional values of the ST parameters found in numerical simulations are compatible given the wide uncertainties in the posterior distributions. However, the distribution of the p parameter seems to hint at smaller (even negative) values.

The upper-left panel of Figure 4.8 shows the posterior-sampled cross-correlation (solid red lines)

Table 4.5: Parameter prior distributions and summarized marginalized statistical results for the first MCMC run of 4.2: a two-parameter ST fit with positive p and a fixed HOD.

Parameter	Prior	Mean	Mode	68% CI
a	$\mathcal{U}[0.00,10.00]$	1.29	0.88	$[0.42, 1.53]$
p	$\mathcal{U}[0.00,0.50]$	-	-	$[0.00, 0.17]$

along with the data (black dots) and the lines corresponding to the traditional values of the ST fit (dotted black) and the marginal peak-like[7] values (dashed light red), that is, $a = 0.88$ and $p = 0.20$. Although it seems to perform well for the smallest scales, the low probability density for $\theta > 3$ arcmin hints at the fact that the model cannot fully explain the large-scale data. The upper-right panel of Figure 4.8 depicts the posterior-sampled HMF at $z = 0$ (solid red lines) and compares it with the fit from traditional (dotted black line) and marginal peak-like (dashed light red) values. Our results are compatible with the usual estimations within the uncertainties, but there appears to be a tendency

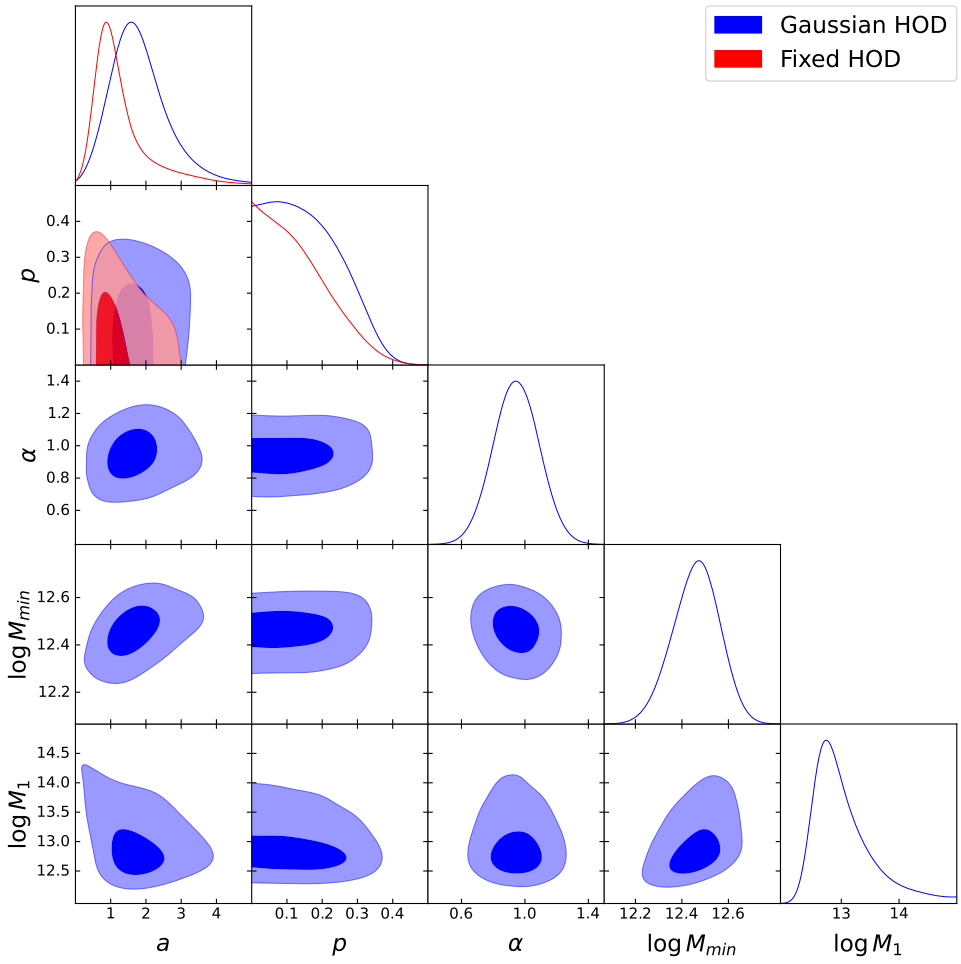

Figure 4.7: Marginalized posterior distributions and contour plots for the first (in red) and second (in blue) MCMC runs of 4.2: a two-parameter ST fit with positive p, with fixed HOD values for the former and Gaussian priors on the HOD for the latter.

[7]Only for visual purposes, we plotted the model with the so-called peak-like values, that is, using those parameter values that seemed to provide a good fit even though their marginalized distributions did not display clear peaks.

toward a smaller number of dark matter halos at large masses ($M > 10^{13.5} M_\odot/h$), for which there are interesting upper bounds. For smaller halo masses, the HMF is much better constrained.

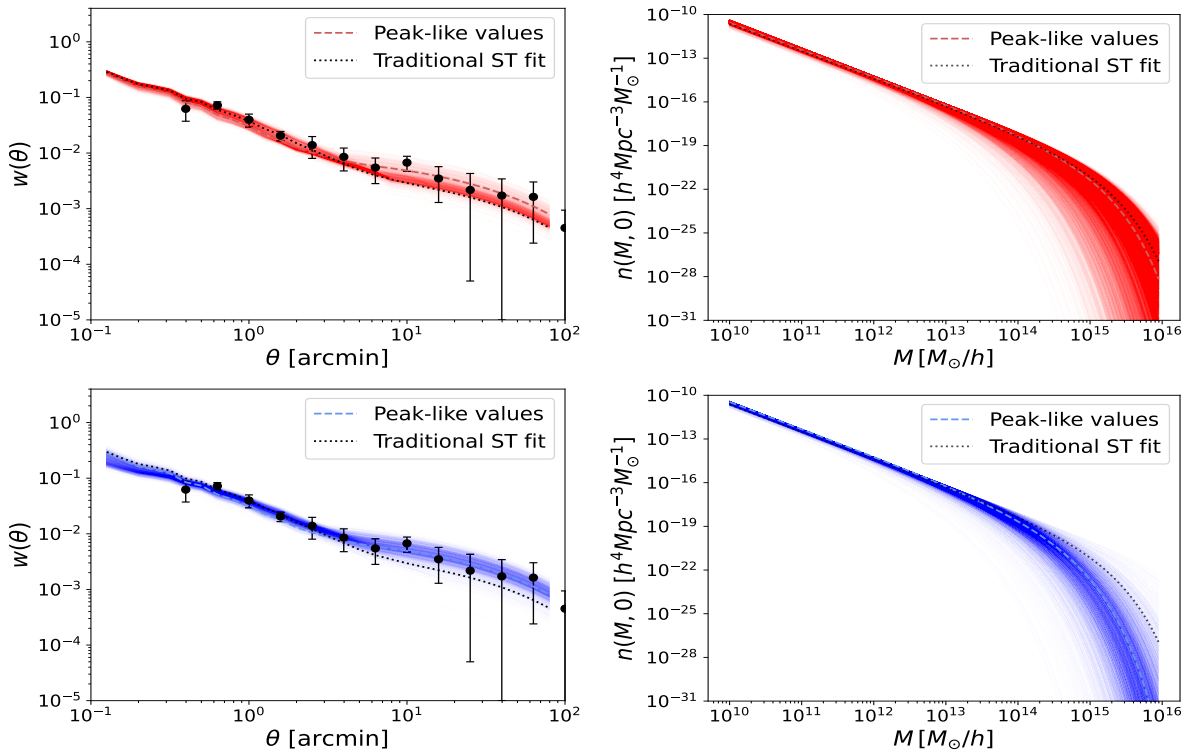

Figure 4.8: Posterior-sampled cross-correlation function (left) and $z = 0$ HMF (right) for the first (in red) and second (in blue) MCMC runs of 4.2: a two-parameter ST fit with positive p, with fixed HOD values for the former and Gaussian priors on the HOD for the latter. The black filled circles are the cross-correlation data. In all four panels, the traditional ST fit and the model using marginal peak-like values are plotted with a dotted black and a dashed line, respectively.

The results vary quantitatively, as expected, when the HOD parameters are introduced in the MCMC algorithm and thus allowed to change. Table 4.6 shows the corresponding summarized statistical results from the second MCMC run, namely the two-parameter ST fit with positive p and Gaussian priors on the HOD, whereas the full corner plot is depicted in Figure 4.7 (in blue). The marginalized posterior distribution of the a parameter is widened and displaced to the right, with a mode of $a = 1.58$

Table 4.6: Parameter prior distributions and summarized marginalized statistical results for the second MCMC run of 4.2: a two-parameter ST fit with positive p and Gaussian priors on the HOD.

Parameter	Prior	Mean	Mode	68% CI
a	$\mathcal{U}[0.00,10.00]$	1.88	1.58	$[0.87, 2.42]$
p	$\mathcal{U}[0.00, 0.50]$	0.15	0.07	$[0.00, 0.20]$
α	$\mathcal{N}[0.92, 0.15]$	0.95	0.94	$[0.80, 1.09]$
$\log M_{\rm min}$	$\mathcal{N}[12.40, 0.10]$	12.46	12.48	$[12.37, 12.57]$
$\log M_1$	$\mathcal{N}[13.95,0.3]$	13.03	12.74	$[12.44, 13.26]$

and a 68% credibility interval of $[0.87, 2.42]$. The p parameter can still only be assigned upper bounds of $p < 0.20$ and $p < 0.33$ at 68% and 95% credibility. As regards the HOD, the marginalized posterior distributions of both α and $M_{\rm min}$ hardly deviate from their priors, but that of M_1 does substantially, with a clear mode of $\log M_1 = 12.74$, more than 3σ away from the mean of its prior.

The lower-left and lower-right panels of Figure 4.8 show the posterior-sampled cross-correlation and $z = 0$ HMF (solid blue lines) along with the data (black dots) and the lines corresponding to the traditional values of the ST fit (dotted black) and the marginal peak values (dashed light blue). In contrast to the previous case, the model can now accommodate the data at all angular scales given the additional freedom provided by the HOD. The derived HMF, although still compatible with the traditional ST values hints at a smaller number density of halos with respect to the preceding case, especially for masses $M > 10^{13.4} M_\odot/h$.

In summary, when the HOD is fixed, the model does not seem to be able to fit the data at angular scales $\theta > 3$ arcmin. Although a sensitivity analysis shows that larger values of a would help in this direction, it would provide a poorer general fit, since it would cause the cross-correlation on small scales (which is better constrained by observations) to decrease. Only when the HOD is introduced in the MCMC analysis can the model accommodate the data by requiring that M_1 decrease, that is, that the mean number of satellite galaxies be larger.

The Tinker fit

Table 4.7 shows the summarized statistical results from the third MCMC run, namely the four-parameter Tinker fit with positive $\tilde{p}$ and a fixed HOD, whereas the full corner plot is depicted in Figure 4.9 (in red). The posterior distributions of the $\tilde{A}$ and C parameters are well constrained, with modes of $\tilde{A} = 0.16$ and $C = 0.63$ and mean values of $\tilde{A} = 0.20^{+0.09}_{-0.12}$ and $C = 0.78^{+0.22}_{-0.45}$. However, the $\tilde{p}$ parameter is unconstrained and B can only be assigned an upper bound of $B < 1.96$ at 68%. Furthermore, this issue is not resolved by widening the prior distributions (or even considering negative values for $\tilde{p}$, in which case B seems to become unconstrained). As a consequence, the statistical conclusions to be drawn are compromised, since they are dependent on the range of allowed values for B and $\tilde{p}$, for which we have no information. Although we suspect that the derived HMF is not too sensitive to this issue (specially since $\tilde{A}$ and C are robust against the aforementioned changes), we considered it delicate and decided not to treat it further.

Notwithstanding this discussion, the Tinker fit with a fixed HOD seems to explain the cross-correlation data at all angular scales in contrast to the ST fit, as shown by the upper-left panel of

Table 4.7: Parameter prior distributions and summarized marginalized statistical results for third MCMC run of 4.2: a four-parameter Tinker fit with positive $\tilde{p}$ and a fixed HOD.

Parameter	Prior	Mean	Mode	68% CI
$\tilde{A}$	$\mathcal{U}[0.00,5.00]$	0.20	0.15	$[0.08, 0.29]$
B	$\mathcal{U}[0.00,5.00]$	1.66	0.82	$[0.00, 1.96]$
C	$\mathcal{U}[0.00,5.00]$	0.78	0.56	$[0.33, 1.00]$
$\tilde{p}$	$\mathcal{U}[0.00,5.00]$	–	–	$[0.00, 5.00]$

Figure 4.10, where the posterior-sampled cross-correlation function (solid red lines) is shown along with the data (black dots) and the lines corresponding to the traditional values of the Tinker fit (black dotted) and the marginal peak-like values (light red dashed), that is, $\tilde{A} = 0.15$, $B = 0.82$, $C = 0.56$ and $\tilde{p} = 1.50$. As seen in the upper-right panel of the same Figure, the posterior-sampled HMF at $z = 0$ is in agreement with the traditional Tinker fit within the uncertainties, although it seems to prefer a steeper cutoff at large masses with respect to the ST fit and shows a wider spread for low and intermediate masses ($M < 10^{14} M_{\odot}/h$).

The results from the fourth run of the MCMC algorithm, that is, the four-parameter Tinker fit with positive $\tilde{p}$ and Gaussian priors on the HOD, are summarized in Table 4.8 and Figure 4.9 (in blue). The marginalized posteriors of the HMF parameters remain practically unchanged with respect to the previous case, with the exception of C, whose distribution is slightly widened with a mode displaced to $C = 0.63$. Regarding the HOD, the situation resembles that of the previous case; while the posterior distributions of α and $M_{\rm min}$ hardly deviate from the priors, that of M_1 is displaced toward smaller

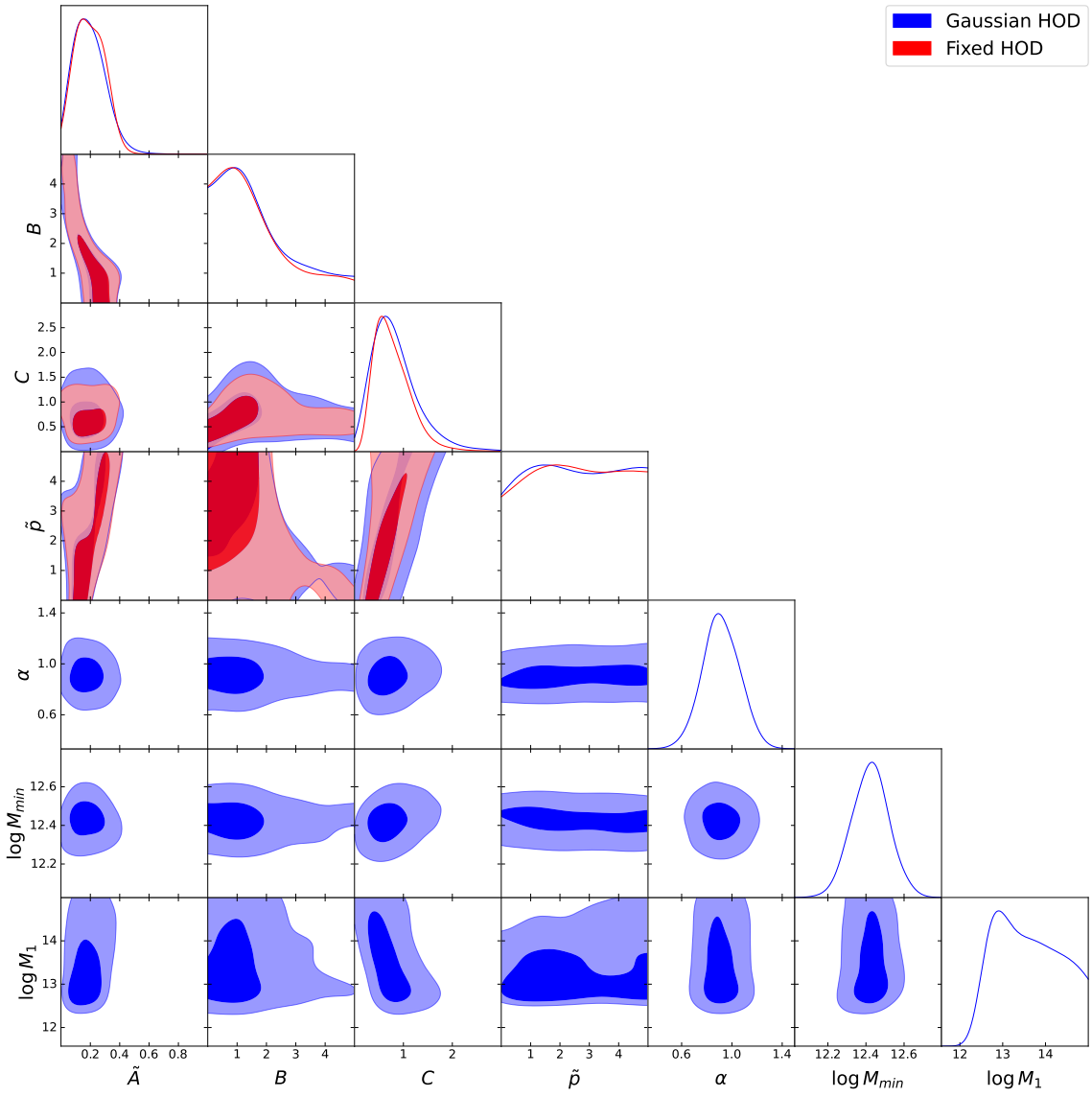

Figure 4.9: Marginalized posterior distributions and contour plots for the third (in red) and fourth (in blue) MCMC runs of 4.2: a four-parameter Tinker fit with positive $\tilde{p}$, with fixed HOD values for the former and Gaussian priors on the HOD for the latter.

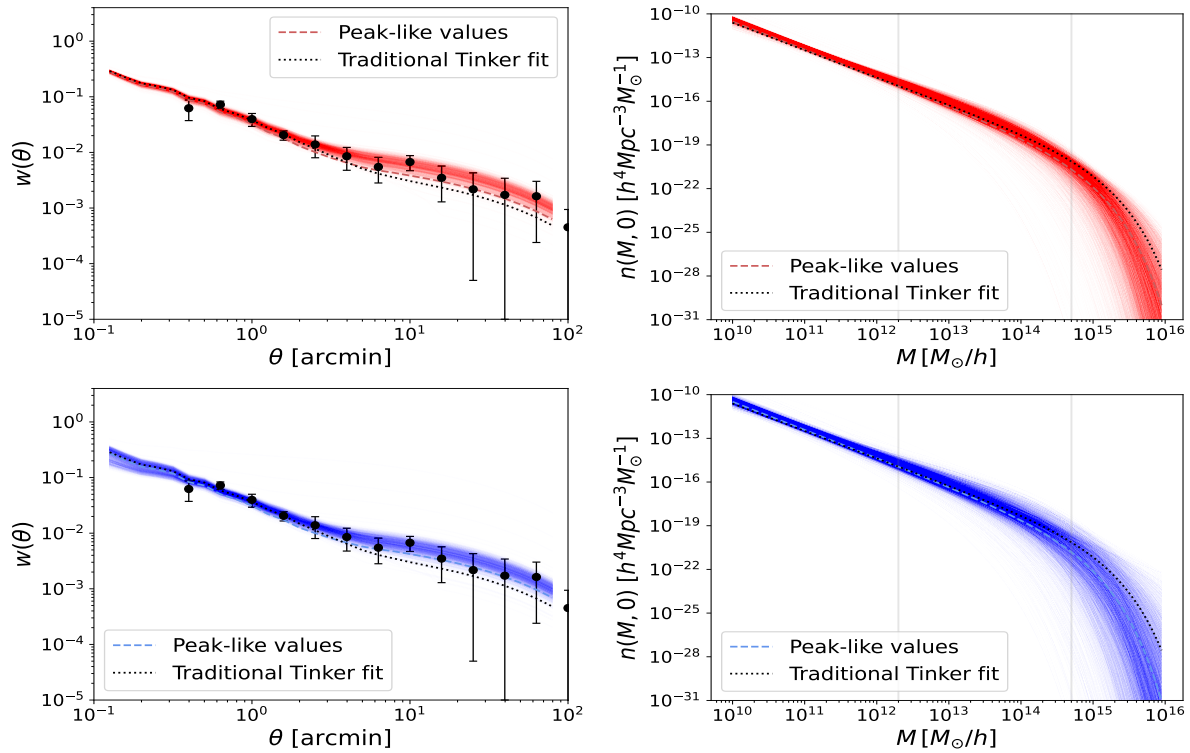

Figure 4.10: Posterior-sampled cross-correlation function (left) and $z = 0$ HMF (right) for the third (in red) and fourth (in blue) MCMC runs of 4.2: a four-parameter Tinker fit with positive $\tilde{p}$, with fixed HOD values for the former and Gaussian priors on the HOD for the latter. The black filled circles are the cross-correlation data. In all four panels, the traditional Tinker fit and the model using marginal peak-like values are plotted with a dotted black and a dashed line, respectively.

values, retaining however a high-probability tail around larger masses.

The lower-left and lower-right panels of Figure 4.10 show the corresponding posterior-sampled cross-correlation function and $z = 0$ HMF (solid red lines), respectively, along with the data (black

Table 4.8: Parameter prior distributions and summarized marginalized statistical results for the fourth MCMC run of 4.2: a four-parameter Tinker fit with positive $\tilde{p}$ and Gaussian priors on the HOD.

Parameter	Prior	Mean	Mode	68% CI
A	$\mathcal{U}$[0.00,5.00]	0.20	0.16	$[0.07, 0.28]$
B	$\mathcal{U}$[0.00,5.00]	1.71	0.91	$[0.00, 2.04]$
C	$\mathcal{U}$[0.00,5.00]	0.85	0.63	[0.27,1.11]
p	$\mathcal{U}$[0.00,5.00]	-	-	$[0.00, 5.00]$
α	$\mathcal{N}$[0.92,0.15]	0.91	0.89	[0.77,1.06]
$\log M_{\rm min}$	$\mathcal{N}$[12.40,0.10]	12.42	12.43	$[12.32, 12.52]$
$\log M_1$	$\mathcal{N}$[13.95,0.30]	13.56	12.91	[12.58,14.20]

dots) and the lines of traditional Tinker (dotted black) and marginal peak-like values (dashed faint blue), that is, $\tilde{A} = 0.16$, $B = 0.91$, $C = 0.63$ and $\tilde{p} = 1.50$. Given the fact that the data were already properly explained with a fixed HOD, we observe an expected increase in the spread of the derived HMF.

To summarize, the Tinker fit gives a robust description of the data against changes in the HOD, but at the expense of compromising the statistical conclusions given their dependence of the prior range of two of its parameters, namely B and $\tilde{p}$.

4.2.2 Further discussion

Non-normalization of the HMF

Table 4.9 shows the summarized results from the fifth MCMC run, namely the three-parameter ST fit with a fixed HOD and positive HMF parameters, whereas the full corner plot is depicted in Figure 4.11 (in red) and compared with that from the two-parameter fit (in blue). The marginalized distribution of A is well constrained, with a mean value of $A = 0.60^{+0.13}_{-0.13}$. With respect to the main two-parameter case, the posterior distribution of a is narrower and better determined, with a mean value of $a = 1.12^{+0.20}_{-0.47}$, while the p parameter is now unconstrained on both sides, hinting at the fact that negative values should perhaps be explored.

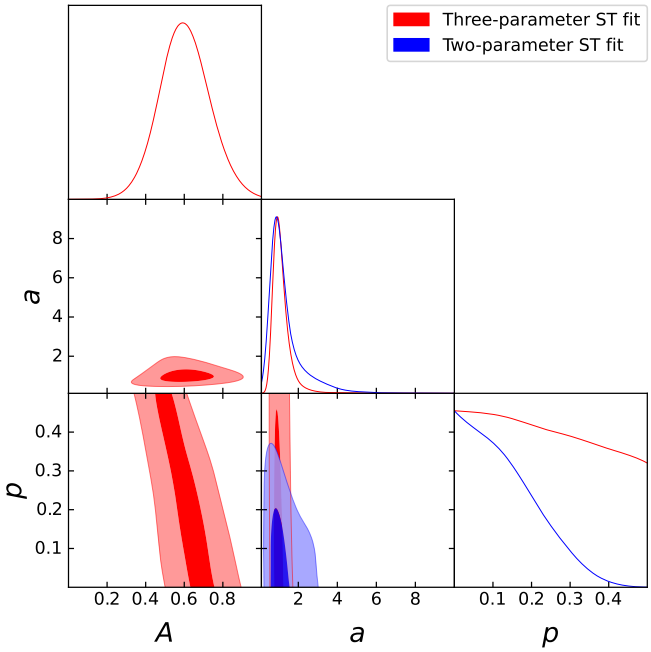

Figure 4.11: Marginalized posterior distributions and contour plots for the first (in blue) and fifth (in red) MCMC runs of 4.2: a two- and a three-parameter ST fit with positive p and a fixed HOD, respectively.

The left panel of Figure 4.12 (in blue) compares the posterior-sampled cross-correlation of this run with that of the two-parameter fit (in red). It becomes clear that the introduction of A as a free parameter in the MCMC algorithm while keeping p positive solves the issue of the two-parameter ST fit and is able to explain the cross-correlation data at large angular scales. In turn, this translates into a more constraining HMF, as shown by the gray band in the right panel of Figure 4.12, which is the area of overlap of both samplings.

Table 4.9: Parameter prior distributions and summarized marginalized statistical results for the fifth MCMC run of 4.2: a three-parameter ST fit with positive p and a fixed HOD.

Parameter	Prior	Mean	Mode	68% CI
A	$\mathcal{U}$[0.00,1.00]	0.60	0.59	$[0.47, 0.73]$
a	$\mathcal{U}$[0.00,10.00]	1.12	0.93	$[0.65, 1.32]$
p	$\mathcal{U}$[0.00,0.50]	-	-	$[0.00, 0.50]$

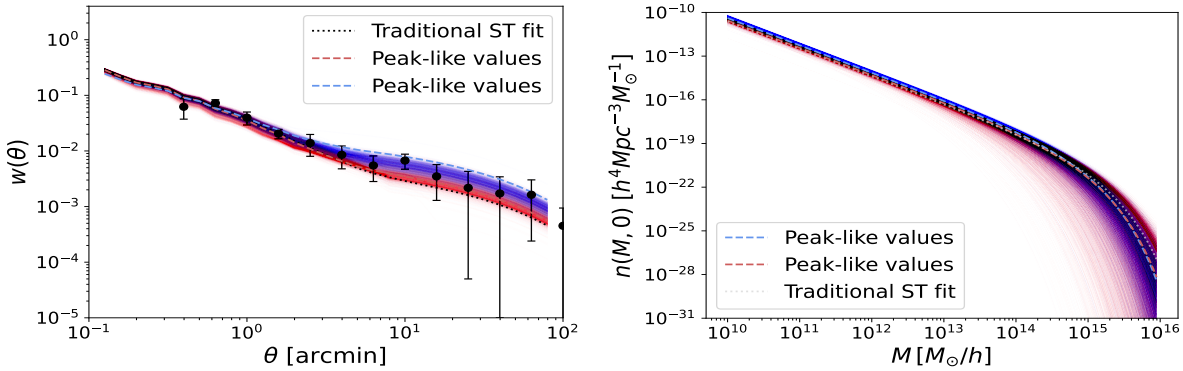

Figure 4.12: Posterior sampling (solid lines) and marginal peak-like values (dashed lines) for the first (in red) and fifth (in blue) MCMC runs of 4.2, that is, a two- and three-parameter ST fit with positive p and a fixed HOD, respectively. The left panels show the cross-correlation function (the black filled circles being our measurements), while the right panels display the $z = 0$ HMF. The dotted line corresponds to the traditional ST fit.

Non-positivity of HMF parameters

Tables 4.10 and 4.11 show the summarized results from the sixth and seventh MCMC runs, namely the two- and three-parameter ST fits with a fixed HOD and the possibility that p can be negative; the full corner plots are depicted in Figure 4.13 in blue and red, respectively.

Table 4.10: Parameter prior distributions and summarized marginalized statistical results for the sixth MCMC run of 4.2: a two-parameter ST fit with the possibility of negative p values and a fixed HOD.

Parameter	Prior	Mean	Mode	68% CI
a	$\mathcal{U}$[0.00,10.00]	2.86	1.46	$[0.32, 3.42]$
p	$\mathcal{U}$[-10.00,0.50]	-1.24	-0.43	$[-1.52, 0.31]$

In the two-parameter case, we observe clear peaks in the marginalized posteriors at values of $a = 1.46$ and $p = -0.43$ together with a strong degeneracy direction that induces the appearance of long tails in the distributions. As a consequence, and as can be seen in the right panel of Figure 4.14 (in blue), the posterior-sampled HMF at $z = 0$ has a much wider spread when compared to the main case of the first run (in red), especially at the large-mass end. Additionally, inspecting the left panel of the same figure allows us to conclude that, although including negative values of p appears

Table 4.11: Parameter prior distributions and summarized marginalized statistical results for the seventh MCMC run of 4.2: a three-parameter ST fit with the possibility of negative p values and a fixed HOD.

Parameter	Prior	Mean	Mode	68% CI
A	$\mathcal{U}$[0.00,5.00]	0.55	0.66	$[0.36, 0.87]$
a	$\mathcal{U}$[0.00,10.00]	1.30	1.29	$[0.74, 2.55]$
p	$\mathcal{U}$[-10.00,0.50]	-1.15	-1.25	$[-2.36, 0.11]$

to account for the cross-correlation signal at large angular scales, it is not as sufficient as varying the HOD parameters or including the normalization parameter A in the MCMC analysis.

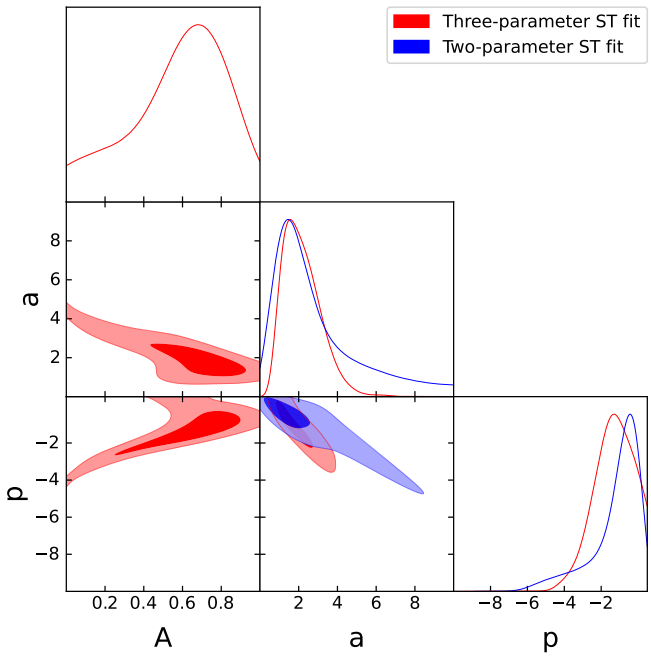

Figure 4.13: Marginalized posterior distributions and contour plots for the sixth (in blue) and seventh (in red) MCMC runs of 4.2: a two- and a three-parameter ST fit with the possibility of negative p values and a fixed HOD, respectively.

On the other hand, the three-parameter fit shows well-constrained marginalized posterior distributions for all the parameters. In particular, it shows a very symmetric distribution for p, as can be seen in Figure 4.13 (in red). This parameter shows once again a degeneracy with a, but this time there are not any long one-sided tails in the distribution. As summarized in Table 4.11, the mean values of the HMF parameters are $A = 0.55^{+0.32}_{-0.14}$, $a = 1.30^{+1.25}_{-0.56}$ and $p = -1.15^{+1.26}_{-1.21}$. Furthermore, as opposed to the two-parameter fit, the cross-correlation data are clearly explained at all angular scales, as can be seen in the left panel of Figure 4.15 (in blue). The derived HMF at $z = 0$ (depicted in the right panel of the same figure) predicts a larger number density of halos for intermediate masses, approximately in the range $10^{12} < M < 10^{15} M_{\odot}/h$.

If we compare both three-parameter cases (according to whether or not we allow p to be negative), the discrepancies are easily explained. Indeed, the p parameter appears to be driven by the data to

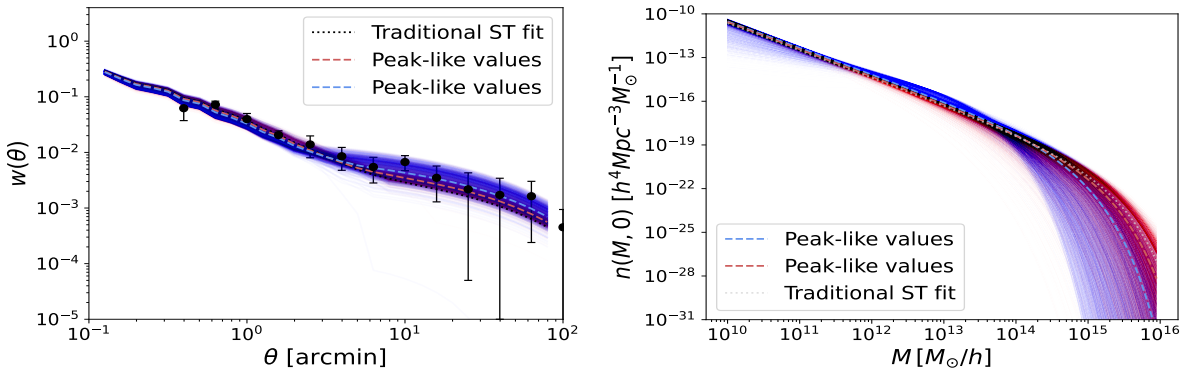

Figure 4.14: Posterior sampling (solid lines) and marginal peak-like values (dashed lines) for the first (in red) and sixth (in blue) MCMC runs of 4.2, that is, a two-parameter ST fit with a fixed HOD, with positive p for the former and the possibility of negative p values for the latter. The left panels show the cross-correlation function (the black filled circles being our measurements), while the right panels display the $z = 0$ HMF. The dotted line corresponds to the traditional ST fit.

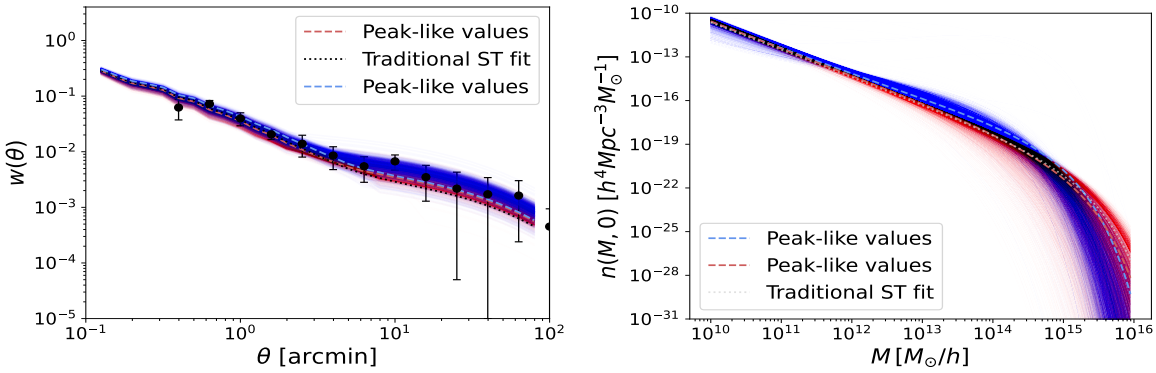

Figure 4.15: Posterior sampling (solid lines) and marginal peak-like values (dashed lines) for the first (in red) and seventh (in blue) MCMC runs of 4.2, that is, a two-parameter ST fit with a fixed HOD and with positive p and and three-parameter ST fit with a fixed HOD and the possibility of negative p values, respectively. The left panels show the cross-correlation function (the black filled circles being our measurements), while the right panels display the $z = 0$ HMF. The dotted line corresponds to the traditional ST fit.

take negative values and, in turn, a has to increase in order to counteract this effect, which is the reason why the three-parameter fit with positive parameters shows smaller values for a: the p parameter did not have a wide enough range to move. In summary, the introduction of A as a free parameter along with the possibility of negative values for p allows us to bypass the two issues that we have encountered in this section: the long tails in the posterior distribution of a and p and the lack of generality on the choice of prior range. This is the reason why this "free" case, more general than the rest, is the most reliable one.

Tabulation of the HMF

The information contained in the shaded bands shown in the plots of the previous subsections can be summarized at certain mass values (chosen from 10^{10} to $10^{15.5} M_\odot/h$) for a direct comparison among the different cases. Figure 4.16 shows the median, 68% and 95% credible intervals for the $z = 0$ HMF at the above masses for the $p > 0$ two-parameter ST fit with fixed HOD values (first MCMC run, in red), the $p > 0$ two-parameter ST fit with Gaussian priors on the HOD (second run, in green) and the

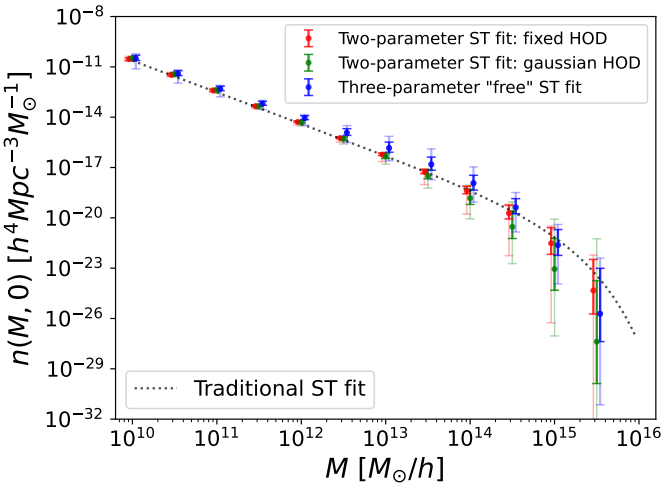

Figure 4.16: Credible intervals (68% in bold and 95% in faint colors) for the $z = 0$ HMF at different mass values when the full posterior distribution is sampled for the first (in red), second (in green) and seventh (in blue) MCMC runs of 4.2. The dotted line corresponds to the traditional ST fit.

three-parameter "free" ST fit with fixed HOD values (seventh run, in blue). Whereas all three cases are compatible with the traditional values of the ST fit at 2σ, there is a slight tendency (1σ discrepancy) toward a larger number density of dark matter halos for masses of $10^{11.5} < M < 10^{14.5} M_\odot/h$ for the "free" ST fit. The other two-parameter fits, although compatible, seem to hint at a smaller number density for large masses ($M > 10^{14} M_\odot/h$).

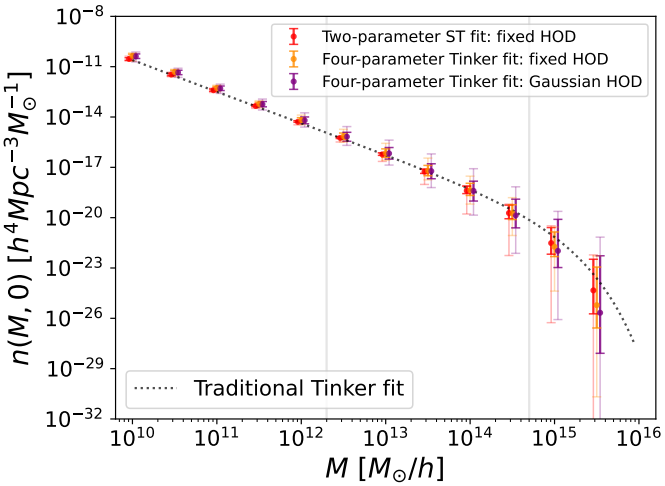

Figure 4.17: Credible intervals (68% in bold and 95% in faint colors) for the $z = 0$ HMF at different mass values when the full posterior distribution is sampled for the first (in red), third (in dark orange) and fourth (in purple) MCMC runs of 4.2. The dotted line corresponds to the traditional Tinker fit and the vertical lines show its range of validity.

Figure 4.17 shows the corresponding results for the two four-parameter Tinker fits, namely that with fixed HOD values and that with Gaussian priors, in dark orange and purple, respectively. The two-parameter main ST fit is also depicted in red for the sake of comparison. The Tinker fit provides a remarkable agreement with the associated traditional values (dotted black line) but there is a wider spread for low and intermediate masses ($M < 10^{14} M_\odot/h$) when compared with the ST fits.

4.3 Correction of large-scale measurement biases

This third section discusses the results from González-Nuevo et al. (2021). The main goal was to study the large-scale biases that arise when estimating the cross-correlation function and their quantification for different choices of tiling scheme. Once the corresponding contributions were computed, we corrected the data and analyzed the deviations from the cosmological constraints coming from the "raw" signal of 4.1. Furthermore, two different foreground galaxy samples were used, one with spectroscopic redshifts (from GAMA II) and another with photometric ones (from SDSS), as described in 3.2, with the aim of exploring the potential differences in the results using a sample of galaxies with a higher number density (in contrast to the less accurate redshift measurements), including a semi-tomographic analysis combining the two samples. Throughout the paper, a flat ΛCDM cosmology is assumed, with

$$\Omega_b = 0.049 \qquad n_s = 0.965 \tag{4.14}$$

following Planck Collaboration et al. (2020a).

Firstly, we analyzed the large-scale estimation biases according to the possible sources contaminating the cross-correlation signal that were described in 3.3. After correcting the data, five main MCMC runs were carried out, all with uniform prior distributions on both the HOD and cosmological parameters using slightly wider ranges than 4.1, namely

$$\begin{aligned} \alpha &\sim \mathcal{U}[0.5, 1.5] & \log M_{\min} &\sim \mathcal{U}[12.0, 14.0] & \log M_1 &\sim \mathcal{U}[12.5, 15.5] \\ \Omega_m &\sim \mathcal{U}[0.1, 0.8] & \sigma_8 &\sim \mathcal{U}[0.6, 1.2] & h &\sim \mathcal{U}[0.5, 1.0]. \end{aligned} \tag{4.15}$$

The first two MCMC runs dealt with the spectroscopic foreground sample and were carried out for the corrected data within the "minitiles" and "tiles" schemes. The photometric foreground sample was used for the third and four MCMC runs. Lastly, and only for the "minitiles" scheme for reasons that will become clear in this section, a joint analysis was carried out using both foreground samples at the same time to explore any potential improvements in the determination of cosmology.

4.3.1 Correction of the cross-correlation signal

As described in 3.3 and as depicted in Figure 3.9, there are different minimal areas over which the cross-correlation estimation can be averaged, each with its advantages and drawbacks. The "all" scheme allows us to count a considerable number of galaxy pairs (providing a large sample size within the minimal area) as well as to probe the highest angular scales, but at the expense of only four to five fields, thus being very affected by cosmic variance. The red diamond (the "tiles" scheme) is arguably the natural shape to choose given the observational strategy. The area of each tile (~ 16 deg^2) should in principle be large enough to avoid integral constraint biases, but a certain (small) overlap between tiles is needed to maintain a regular shape. Averaging over around 24 different tiles should, however, help diminish the effects of cosmic variance. Lastly, the "minitiles" scheme is made up of minimal areas (the blue diamonds) of ~ 4 deg^2. The obvious advantage is the availability of about 96 minitiles, but the maximum distances allowed in this case are very close to the angular scales we want to probe for cosmological constraints.

Figure 4.18 shows the "raw" cross-correlation signal for both the GAMA (spectroscopic) and SDSS (photometric) galaxy samples and for each tiling scheme[8] (note that the GAMA foreground sample

[8]The measurements with the "all" tiling scheme are practically the same for both foreground samples, which is why we

with the "minitiles" scheme was the one used in Bonavera et al., 2020). Whereas the measurements agree within the uncertainties for the smallest angular scales, there is a non-negligible discrepancy in the data above $\theta \sim 10$ arcmin. Since these scales are of interest to us for cosmological constraints, corrections need to be applied so that the analysis is consistent. However, it should be highlighted that the measurements using the "minitiles" scheme appear to be the most consistent when compared between the two foreground samples.

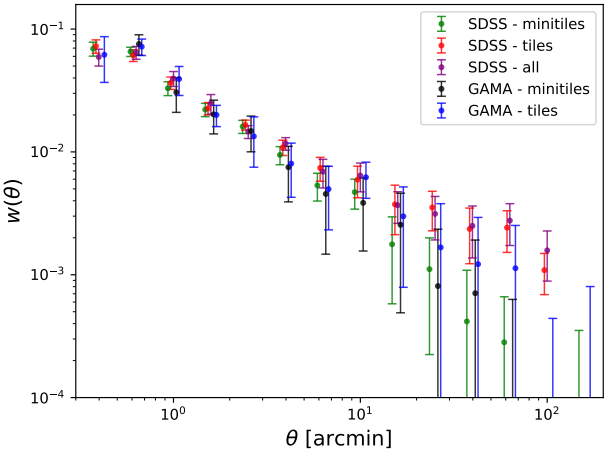

Figure 4.18: Cross-correlation measurements before large-scale bias corrections for both foreground galaxy samples (GAMA and SDSS) using the "tiles", "minitiles" and "all" schemes.

Integral constraint

As discussed in 3.3, the integral constraint is usually estimated via (3.5), for which we need to assume an "ideal" model for the cross-correlation function. As a first approximation, we used a power law, that is, $w_{\rm ideal}(\theta) = A\theta^{\gamma}$ and estimated the best-fit value of the parameters using only the data below 20 arcmin to be as independent as possible of cosmology, yielding $\gamma = -0.89$ and $A = 10^{-1.54}$ arcmin$^{0.89}$. With this model, the derived IC value for the "minitiles" scheme was IC $= 9 \cdot 10^{-4}$, a number that was shown to be robust against the choice of a smaller upper limit for the angular separation or different data sets. A second method for the "ideal" model assumed the best fit of 4.1, which can be considered biased low, yielding an equal result.

Repeating the analysis for the "tiles" scheme, we found a value of IC $= 5 \cdot 10^{-4}$ for both the power-law model and the best fit from 4.1. As expected, the correction for the "tiles" scheme is smaller than for the "minitiles" and only marginally affects the data above ~ 40 arcmin. For the "all" tiling scheme, the integral constraint is completely negligible.

Variations in the surface number density

A clear source of a varying background surface density four our sample is the well-known instrumental noise due to the scanning strategy of H-ATLAS. As described in 3.3 and as highlighted in the middle panel of Figure 3.4, spurious overdensities can arise due to the repeated scanning or observation of a given region. The correcting procedure entails the adoption of a specific method to generate the random catalogs for the estimation of the cross-correlation signal that takes this bias into account.

have decided to focus only on SDSS for simplicity.

Following Amvrosiadis et al. (2019), we started by choosing a random flux density among those in the background sample. For each randomly located galaxy on the field, we estimated the local noise as the sum of the instrumental and confusion noises (as detailed in Valiante et al., 2016, for the GAMA fields) and introduced a Gaussian perturbation to the flux density. If the simulated galaxy had a flux density at least four times larger than the estimated local noise (the same detection limit to reproduce the official H-ATLAS catalog), it was kept in the sample; otherwise, it was discarded. The process was repeated until the random catalog was completed. When this correction was applied to the cross-correlation estimation procedure, the signal was slightly displaced toward lower values for the largest angular scales. However, although this proved the instrumental noise was not a negligible bias, it also highlighted the fact that it was not enough to explain the stronger cross-correlation signal observed for the "tiles" scheme in the GAMA foreground sample and for the "all" scheme in both data sets.

The additional surface density variations we addressed had to do with the foreground samples, for which we adopted the following simple approach. A surface density map was created by adding 1 to the pixel value at the position of each galaxy of the sample. This map was then smoothed via a Gaussian kernel with a certain standard deviation and the H-ATLAS masks were applied to neglect the border effects that could arise due to the smoothing step. These surface density maps were then used to generate the random catalogs for the foreground samples. This process is useful to unveil any hidden large-scale density variations, but suffers from a serious issue: the exact value for the dispersion of the Gaussian kernel is not known. Although one can set some lower and upper limits on this quantity by considering the effects that too low or too high a value could induce, we decided to proceed using a standard deviation of 180 arcmin, but keeping in mind that it was arbitrarily chosen. Nevertheless, we verified that small variations around this value are only second order when taking into account the uncertainties in the measurements.

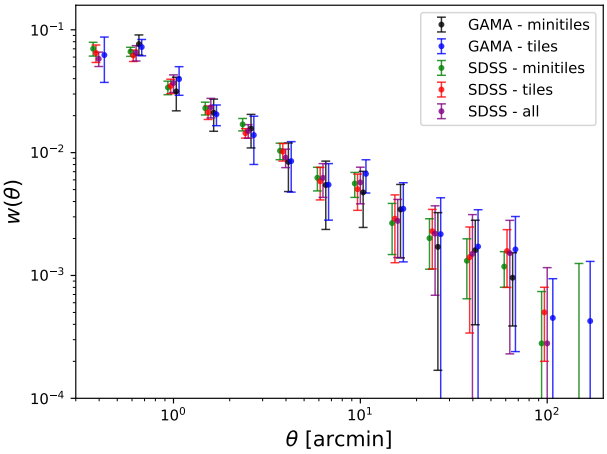

Figure 4.19: Cross-correlation measurements after large-scale bias corrections for both galaxy samples (GAMA and SDSS) using the "tiles", "minitiles" and "all" schemes.

Once all corrections are taken into account, the difference among the mean values of the cross-correlation function at each angular scale is much smaller than the corresponding uncertainties, as can be seen in Figure 4.19. The result is thus deemed robust against the tiling scheme at all angular scales. The use of minitiles as the minimal area for both samples provides the most satisfactory tiling scheme, given the larger number of them (that minimizes cosmic variance and provides a statistically more meaningful sample size) and the fact that only the integral constraint correction is actually needed,

effectively rendering innocuous the issue of the unknown dispersion of the Gaussian kernel.

Although in principle we would have 6 different cases to study in this paper (three tiling schemes for each foreground sample), the "all" tiling scheme was discarded given the fact that cosmic variance is the most relevant source of uncertainty and that the arbitrary choice of the Gaussian kernel is most important for this case. Furthermore, the agreement between the "all" and "tiles" schemes for both data sets is almost perfect, which justifies our choice and leaves us with four different MCMC runs to analyze.

4.3.2 Cosmological constraints

The GAMA foreground sample

The summarized statistical results from the first and second MCMC runs, that is, using the "minitiles" and "tiles" schemes for the GAMA foreground sample, are shown in Table 4.12, whereas Figure 4.20 depicts the corresponding corner plots (in black and blue, respectively).

The results are qualitatively similar to 4.1 and decently consistent between the two tiling schemes, with slighly larger uncertainties for the "minitiles" case. The average minimum halo mass is well constrained for both the "minitiles" and "tiles" schemes, with a mean value of $\log M_{\rm min} = 12.57^{+0.23}_{-0.17}$ and $12.61^{+0.19}_{-0.15}$, respectively. The α parameter is once again unconstrained for both cases, whereas M_1 can only be assigned lower bounds, namely $\log M_1 > 13.88$ and > 14.00, respectively, at 68%.

As regards cosmology, and in contrast to 4.1, the Ω_m parameter can now be constrained on both sides, with a mean value of $\Omega_m = 0.45^{+0.13}_{-0.21}$ and $0.42^{+0.14}_{-0.24}$ for the "minitiles" and "tiles" scheme, respectively, the corresponding modes being 0.38 and 0.31. This indicates that the large-scale corrections carried out in this paper have increased the constraining power of our observable and confirms the suspicion (motivated by the sensitivity plots of Figure 4.2) that a higher cross-correlation signal at the largest scales (as is obtained after the corrections) induces a lower value of Ω_m. The results for the σ_8 parameter are in agreement with 4.1, yielding a mean value of $\sigma_8 = 0.84^{+0.11}_{-0.18}$ and $0.82^{+0.08}_{-0.20}$ for the "minitiles" and "tiles" scheme, respectively, the modes being 0.83 and 0.75. The adimensional Hubble

Table 4.12: Parameter prior distributions and summarized marginalized statistical results for the first and second runs of 4.3: the GAMA spectroscopic foreground sample with the "minitiles" and "tiles" schemes, respectively.

		Minitiles			Tiles		
Parameter	Prior	Mean	Mode	68% CI	Mean	Mode	68% CI
$\log M_{\rm min}$	$\mathcal{U}[12.00, 14.00]$	12.57	12.61	$[12.40, 12.80]$	12.61	12.56	$[12.46, 12.80]$
$\log M_1$	$\mathcal{U}[12.50, 15.50]$	14.26	15.03	$[13.88, 15.50]$	14.37	14.71	$[14.00, 15.50]$
α	$\mathcal{U}[0.50, 1.50]$	-	-	$[0.50, 1.50]$	-	-	$[0.00, 0.50]$
Ω_m	$\mathcal{U}[0.10, 0.80]$	0.45	0.38	$[0.34, 0.58]$	0.42	0.31	$[0.18, 0.56]$
σ_8	$\mathcal{U}[0.60, 1.20]$	0.84	0.83	$[0.66, 0.95]$	0.82	0.75	$[0.62, 0.90]$
h	$\mathcal{U}[0.50, 1.00]$	-	-	$[0.50, 1.00]$	-	-	$[0.00, 1.50]$

constant, h is once again unconstrained.

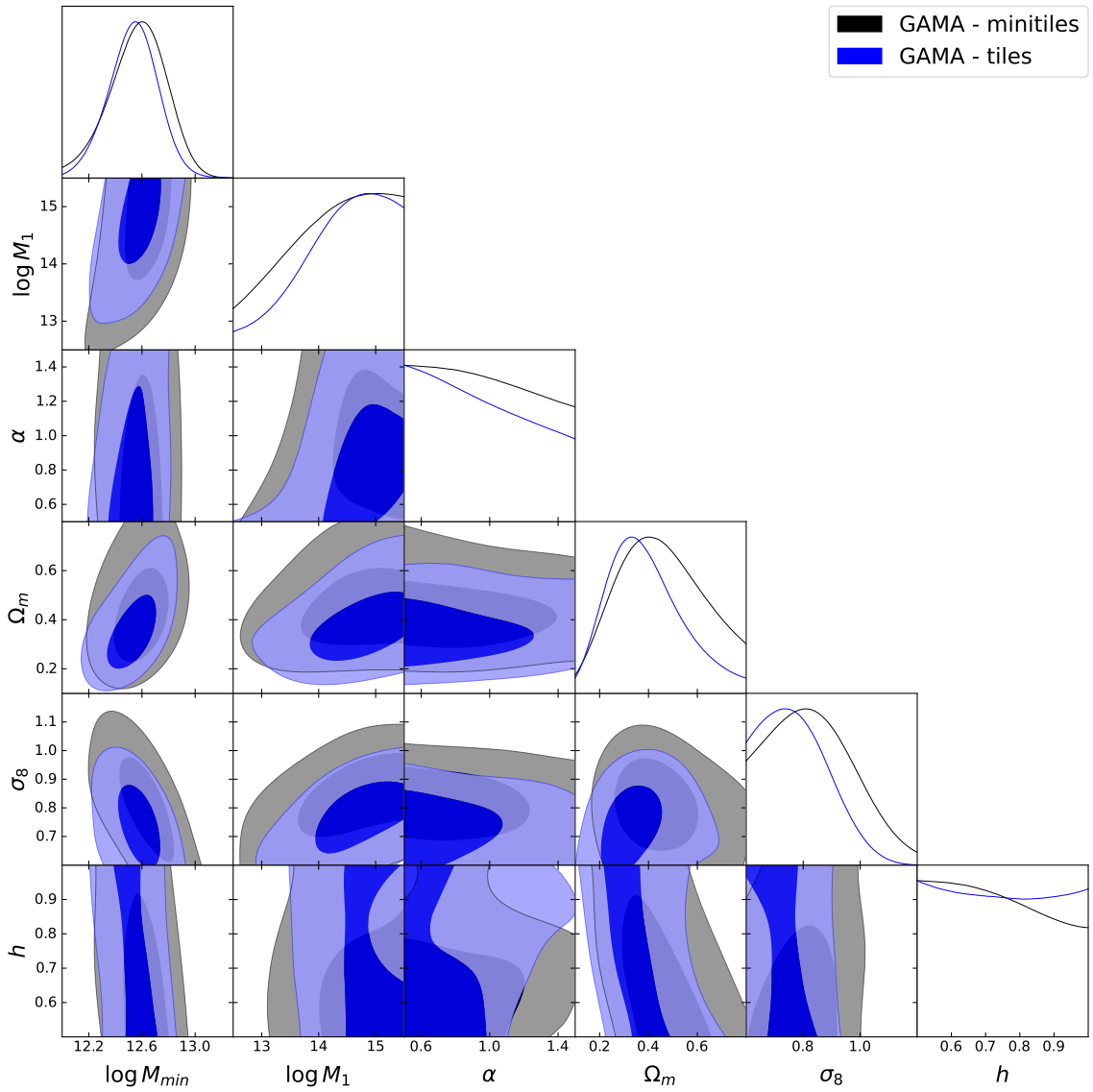

Figure 4.20: Marginalized posterior distributions and contour plots for the first (in black) and second (in blue) MCMC runs of 4.3: the GAMA spectroscopic foreground sample with the "minitiles" and "tiles" scheme, respectively.

The SDSS foreground sample

The corresponding statistical results from the third and fourth MCMC runs, that is, using the "minitiles" and "tiles" schemes for the SDSS foreground sample are shown in Table 4.13, whereas Figure 4.21 depicts the corresponding corner plots (in red and blue, respectively).

The results are once again qualitatively similar to 4.1 and, this time, remarkably consistent between the two tiling schemes. The average minimum halo mass is equally well-constrained, with mean values of $\log M_{\rm min} = 12.60^{+0.20}_{-0.13}$ and $12.61^{+0.20}_{-0.13}$ for the "minitiles" and "tiles" schemes, respectively. The posterior distribution of the M_1 parameter, however, differs substantially from that of the previous case. Although it is not very constraining, it has a clear peak at $\log M_1 = 13.60$ and 13.74 for the "minitiles" and "tiles" schemes, with a mean value of $\log M_1 = 13.81^{+0.53}_{-1.09}$ and $13.95^{+0.74}_{-0.95}$, respectively. The α parameter is unsurprisingly not constrained.

As far as cosmology is concerned, the matter density parameter, Ω_m, remains practically unchanged with respect to the GAMA foreground sample. Indeed, the posterior distribution peaks at $\Omega_m = 0.38$ and 0.39 for the "minitiles" and "tiles" schemes and the associated mean values are $\Omega_m = 0.46^{+0.11}_{-0.18}$ and $0.46^{+0.12}_{-0.19}$, respectively. The σ_8 parameter, however, is substantially drawn toward higher values than the previous case, with a mean of $\sigma_8 = 0.99^{+0.12}_{-0.11}$ and $0.98^{+0.16}_{-0.10}$ for the "minitiles" and "tiles" schemes. The Hubble constant, just like α, is unconstrained.

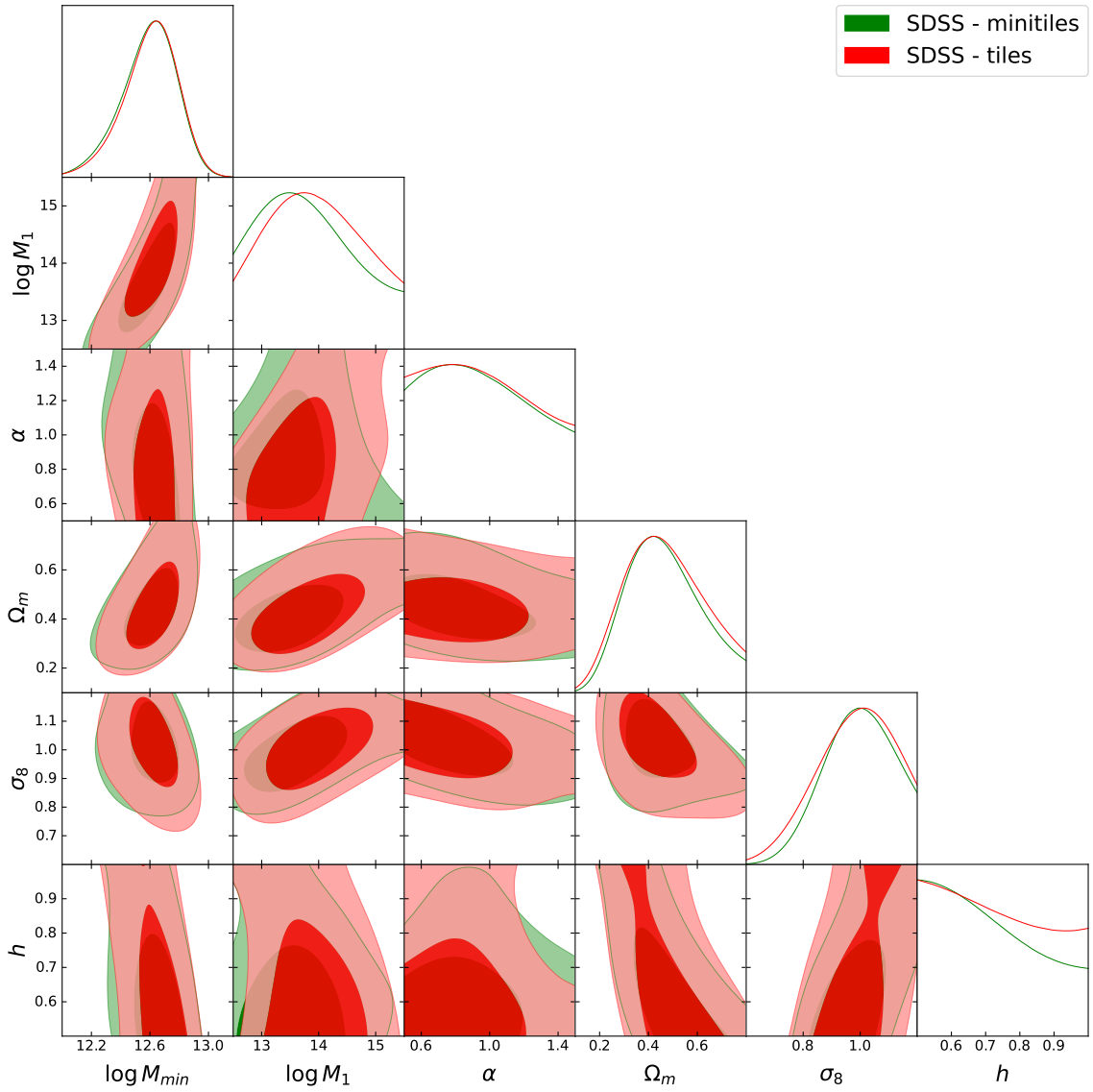

Figure 4.21: Marginalized posterior distributions and contour plots for the third (in green) and fourth (in red) MCMC runs of 4.3: the SDSS photometric foreground sample with the "minitiles" and "tiles" scheme, respectively

To summarize, the constraints on the HOD and cosmological parameters are relatively robust with respect to the tiling scheme within the same foreground sample, with remarkable consistency in the case of the photometric SDSS data set. The large-scale bias corrections provide a two-sided constraint on Ω_m, to be compared with only lower limit found in 4.1. However, there are non-negligible discrepancies in the M_1 and σ_8 parameters between the two foreground samples and, furthermore, the higher number density of the SDSS sample does not appear to increase the constraining power of our observable. Given the fact that the cross-correlation measurements for the two samples are almost equal, the differences in

Table 4.13: Parameter prior distributions and summarized marginalized statistical results for the third and fourth runs of 4.3: the SDSS photometric foreground sample with the "minitiles" and "tiles" schemes, respectively.

		Minitiles			Tiles		
Parameter	Prior	Mean	Mode	68% CI	Mean	Mode	68% CI
$\log M_{\rm min}$	$\mathcal{U}[12.00, 14.00]$	12.60	12.67	$[12.47, 12.80]$	12.61	12.66	$[12.48, 12.81]$
$\log M_1$	$\mathcal{U}[12.50, 15.50]$	13.81	13.60	$[12.72, 14.34]$	13.95	13.74	$[13.00, 14.69]$
α	$\mathcal{U}[0.50, 1.50]$	0.96	0.77	$[0.50, 1.11]$	0.96	0.73	$[0.50, 1.11]$
Ω_m	$\mathcal{U}[0.10, 0.80]$	0.46	0.38	$[0.28, 0.57]$	0.46	0.39	$[0.27, 0.58]$
σ_8	$\mathcal{U}[0.60, 1.20]$	0.99	0.98	$[0.88, 1.11]$	0.98	1.00	$[0.88, 1.14]$
h	$\mathcal{U}[0.50, 1.00]$	-	-	$[0.50, 0.77]$	-	-	$[0.50, 1.00]$

the parameter constraints appear to be related only to the redshift distributions. Indeed, as concluded by González-Nuevo et al. (2017), a preliminary tomographic analysis dividing the foreground sample in four redshift bins showed a strong time evolution of $M_{\rm min}$ and slight differences in M_1 among bins. When using a single (wide) foreground bin, the derived parameters are actually an average weighted by the corresponding redshift distribution. As a consequence, we expect that a tomographic analysis will yield more consistent results, although the GAMA spectroscopic sample will be more reliable for such a study given its precise redshift measurements.

The combined sample

Since a proper tomographic analysis was beyond the scope of this section, we performed a joint analysis of both foreground samples with a shared cosmology, meaning that only the HOD was allowed to vary between data sets. The photometric SDSS sample could then safely be used, since we were still considering a single redshift bin and photometric errors were shown not to be too important. This fifth MCMC run only made use of the "minitiles" scheme data, given the minimal large-scale bias corrections needed. Moreover, since both samples partially overlap in the sky, the measured cross-correlation data cannot be considered independent. In order to avoid taking into account the correlation between the samples, we restricted the SDSS foreground sample to just the NGP zone of the H-ATLAS survey (which was not covered by GAMA), as described in 3.1. As regards the cross-correlation measurements using this restricted SDSS sample, they agree with the original data set, but the uncertainties increase by a factor of ~ 2 at the largest angular scales. The redshift distribution, however, does not display any significant differences.

The summarized statistical results from this joint MCMC run are shown in Table 4.14, whereas Figure 4.22 depicts the corresponding corner plots in dark orange. Regarding the HOD, the results for the GAMA sample remain practically the same; indeed, the average minimum halo mass went from a mean value of $\log M_{\rm min} = 12.57^{+0.23}_{-0.17}$ in the preceding analysis to $\log M_{\rm min} = 12.55^{+0.22}_{-0.19}$ in the joint run. The M_1 parameter can still only be assigned a lower limit, with $\log M_1 > 13.82$ at 68% credibility, to be compared with the previous value of $\log M_1 > 13.88$. Concerning the SDSS sample, the results are

somewhat different: the $M_{\rm min}$ parameter barely changes, with a mean value of $\log M_{\rm min} = 12.49^{+0.25}_{-0.24}$ in the joint analysis and $\log M_{\rm min} = 12.57^{+0.23}_{-0.17}$ in the individual one. However, the distribution of M_1 is displaced toward smaller values, with a 68% upper limit of $\log M_1 < 13.73$ in this fifth run, to be compared with the 68% credible interval of $[12.72, 14.34]$ of the preceding analysis. This shift is most certainly due to the larger error bars in the cross-correlation measurements of the "new" SDSS lens sample used in the joint analysis. The distributions of the α parameters are hardly modified for both data sets.

With respect to cosmology, we obtained a remarkable improvement in the determination of the σ_8 parameter, with a mean of $\sigma_8 = 0.75^{+0.07}_{-0.10}$, to be compared with the values of $0.84^{+0.11}_{-0.18}$ and $0.99^{+0.12}_{-0.11}$ found for the individual GAMA and SDSS runs, respectively. The Ω_m parameter does not seem to benefit from the joint analysis, which could be due to the dependence of this parameter on the largest scales of the cross-correlation data (as shown by the sensitivity plots of Figure 4.2) and is thus expected to be influenced by the smaller-area SDSS sample used for this last MCMC run. The adimensional

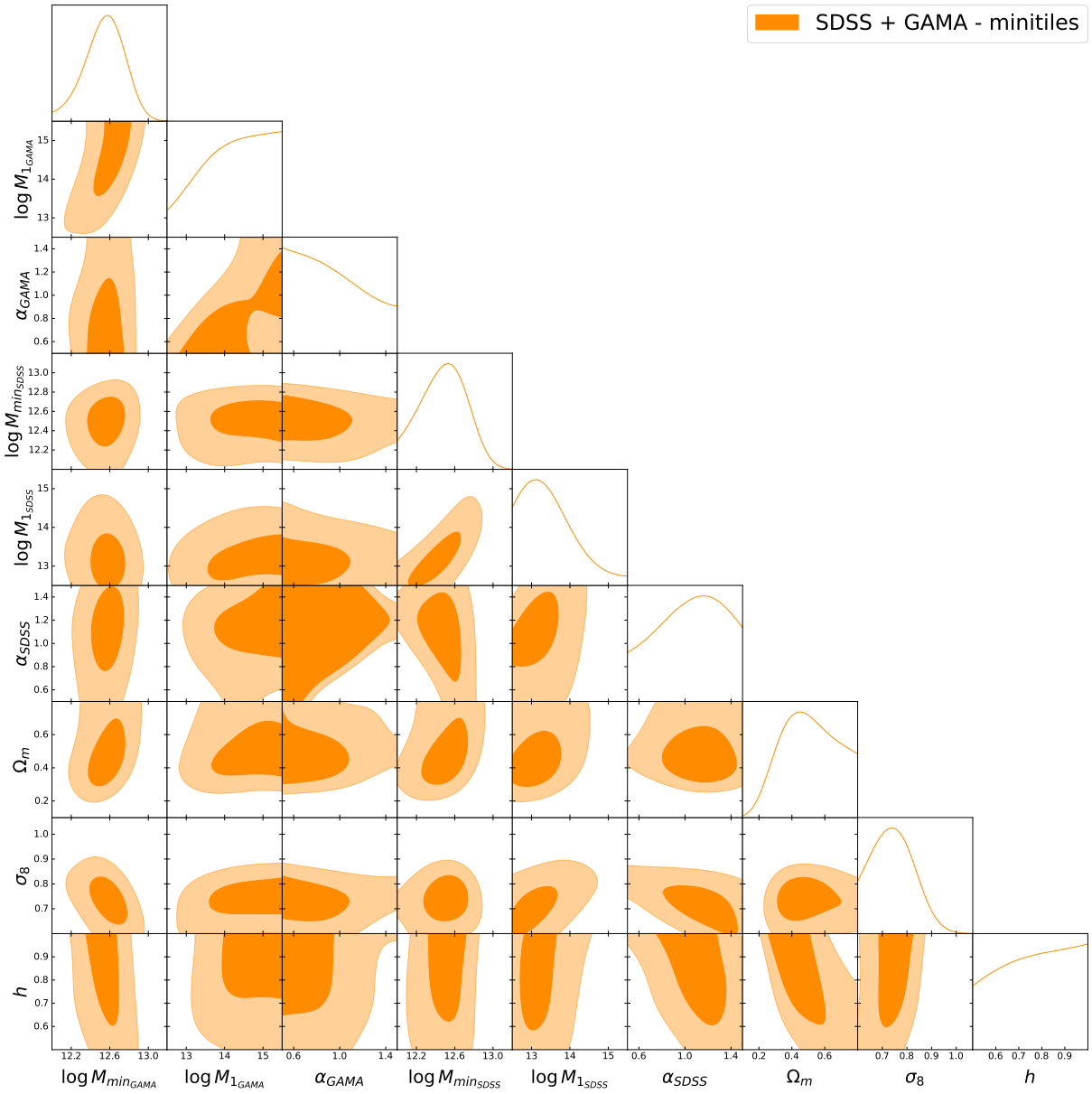

Figure 4.22: Marginalized posterior distributions and contour plots for the fifth MCMC run of 4.3: the SDSS and GAMA foreground samples in a joint analysis with the "minitiles" scheme.

Table 4.14: Parameter prior distributions and summarized marginalized statistical results for the fifth run of 4.3: the SDSS and GAMA foreground samples in a joint analysis with the "minitiles" scheme.

Parameter	Prior	Mean	Mode	68% CI
$\log M_{\mathrm{min}_{\mathrm{GAMA}}}$	$\mathcal{U}$[12.00,14.00]	12.55	12.59	[12.36, 12.77]
$\log M_{1_{\mathrm{GAMA}}}$	$\mathcal{U}$[12.50,15.50]	-	-	[13.82, 15.50]
α_{GAMA}	$\mathcal{U}$[0.50,1.50]	-	-	[0.50,1.06]
$\log M_{\mathrm{min}_{\mathrm{SDSS}}}$	$\mathcal{U}$[12.00,14.00]	12.49	12.54	[12.25, 12.74]
$\log M_{1_{\mathrm{SDSS}}}$	$\mathcal{U}$[12.50,15.50]	13.49	13.11	[12.50, 13.73]
α_{SDSS}	$\mathcal{U}$[0.50,1.50]	1.04	1.16	[0.85, 1.44]
Ω_m	$\mathcal{U}$[0.10,0.80]	0.50	0.45	[0.30,0.64]
σ_8	$\mathcal{U}$[0.60,1.20]	0.75	0.74	[0.65,0.82]
h	$\mathcal{U}$[0.50,1.00]	-	-	[0.69, 1.00]

Hubble constant, h, appears to be still prior-dominated.

4.3.3 Comparison with other results

As we did in 4.1.3 for the results of 4.1, we can compare our findings with those of the most popular cosmological probes; in particular, the CMB lensing analysis by *Planck* (Planck Collaboration et al., 2020b) and the cosmic shear studies from CFHTlens (Joudaki et al., 2017), HSC (Hamana et al., 2020), the combination of KiDS and VIKING (Hildebrandt et al., 2020) and the first-year data of DES (Troxel et al., 2018).

Figure 4.23 shows the contour plots from the aforementioned cosmological probes along with our results from the individual-sample (left panel) and joint (right panel) analyses. The results for the individual MCMC run with the GAMA foreground sample are shown with dashed lines, while those from the SDSS sample are shown in gray. With respect to 4.1, we can conclude that the marginalized constraints on Ω_m and σ_8 have benefited from the large-scale bias corrections and are more in agreement with the results from other probes. As for the $\Omega_m - \sigma_8$ plane, the spectroscopic GAMA sample yields results that are more in accordance with external findings, which might be indicative of the importance of precise redshift measurements for the lens sample. Moreover, the most relevant aspect is the restriction of the probability contour when both foreground samples are analyzed in a joint study. This appears to suggest that a proper tomographic analysis, as was anticipated in 4.1, could lead to tighter constraints both on this plane and on the corresponding marginalized distributions.

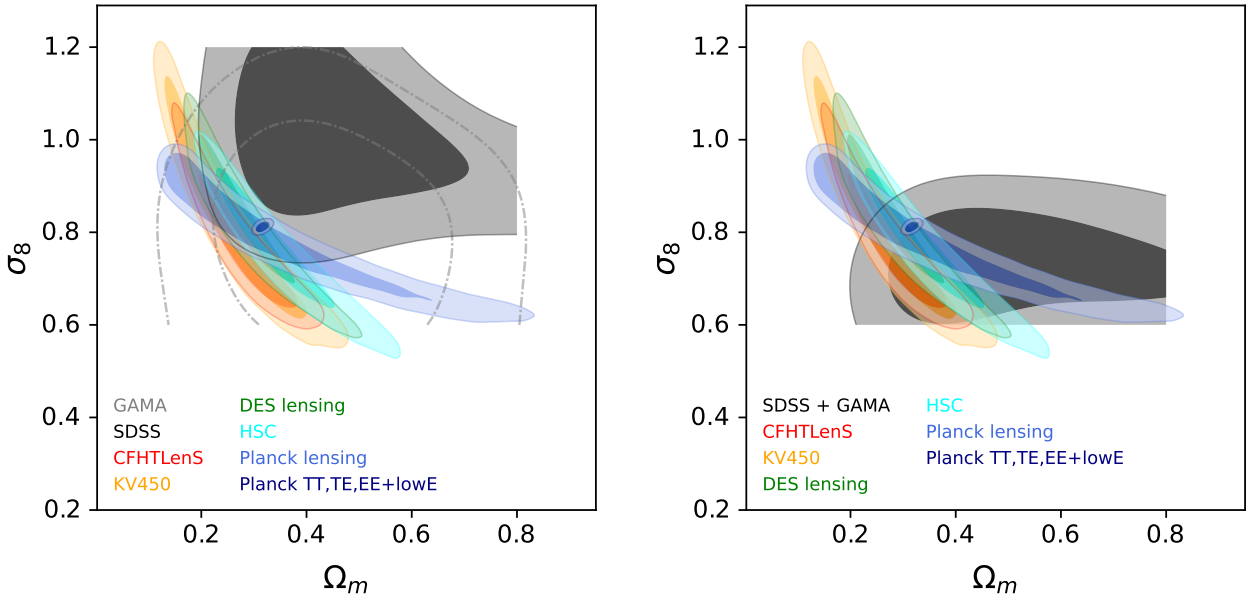

Figure 4.23: Contour plots on the $\Omega_m - \sigma_8$ plane comparing the results for the individual-sample (left) and joint (right) analyses of 4.3 with those from *Planck* (blue for lensing-only and dark blue for temperature and polarization), CFHTLens (red), KV450 (orange), DES lensing (green) and HSC (cyan). The results from the SDSS and GAMA foreground sample are depicted in solid and dashed gray, respectively. The probability contours are set to 0.68 and 0.95.

5. Tomographic analyses

The main conclusion drawn from the discussion of the previous chapter is that a non-tomographic approach can only yield so much accuracy with the current data. As is often done in shear studies, a tomographic analysis is likely to improve the constraints on the parameters. On that note, the main results from the last two published contributions to this thesis, namely Bonavera et al. (2021) and Cueli et al. (2022) will be laid out in this chapter. These works should be regarded as the natural extension of 4.1 and 4.2 to a tomographic setup with the additional large-scale corrections provided in 4.3. Indeed, the main goal of the discussion that will be laid out in 5.1 was to improve the cosmological constraints by allowing the HOD to evolve with redshift. Furthermore, the tomographic character made it possible to explore the possibility of a time-varying dark energy density, effectively probing the flat ΛCDM, ω_0CDM and $\omega_0\omega_a$CDM cosmological models, described in 2.1.2. Likewise, the main objective of the work that will be presented in 5.2 was to make improvements to the observational constraints on the number density of dark matter halos by allowing time-evolving HOD parameters. The possibility of a violation of universality in the HMF was also explored.

5.1 Cosmology and the HOD

This section discusses the tomographic extension of 4.1. Throughout this segment, the GAMA foreground sample (the same as in 4.1 but with additional sources above $z > 0.1$) was divided up into four redshift bins, as described in 3.2, with ranges

$$\text{Bin 1: } z \in [0.1, 0.2] \qquad \text{Bin 2: } z \in [0.2, 0.3] \qquad \text{Bin 3: } z \in [0.3, 0.5] \qquad \text{Bin 4: } z \in [0.5, 0.8]$$

to perform a tomographic analysis of the cross-correlation function, effectively performing a joint MCMC analyisis on four different observables while keeping the cosmology unchanged among them. Following the conclusions of 4.3, we adopted the "minitiles" scheme for the estimation of our observable, given the minimal large-scale corrections needed (and their non-arbitrary nature) and the optimal sample size for each angular scale. The corresponding cross-correlation data are shown in Figure 5.1.

Four different MCMC runs were carried out, with a flat cosmology assumed throughout the paper, where common values of

$$n_s = 0.965 \qquad \Omega_b = 0.049 \tag{5.1}$$

were set, following Planck Collaboration et al. (2020a). As in 4.1 and 4.3, a fixed ST model for the HMF with the traditional simulation-based values was used. The first case we studied was a fixed flat ΛCDM cosmology according to *Planck*'s best fit,

$$\Omega_m = 0.3089 \qquad \sigma_8 = 0.8159 \qquad h = 0.6774, \tag{5.2}$$

aiming to derive the posterior probability distribution of only the HOD parameters, with the peculiarity that this time they were allowed to vary among redshift bins in this paper. For all four runs, the prior distributions for the HOD were chosen to be uniform with slightly different values among redshift bins, except for the α parameter, which was assigned the same prior as in 4.1. Indeed, the preliminary tomographic analysis of González-Nuevo et al. (2017) allowed us to conclude that the average minimum

halo mass appeared to increase with redshift, thus motivating the following choice of prior distributions:

$$\begin{aligned}
&\alpha_1 \sim \mathcal{U}[0.5, 1.5] && \log M_{\min_1} \sim \mathcal{U}[10.0, 13.0] && \log M_{1_1} \sim \mathcal{U}[11.0, 15.5] \\
&\alpha_2 \sim \mathcal{U}[0.5, 1.5] && \log M_{\min_2} \sim \mathcal{U}[11.0, 13.0] && \log M_{1_2} \sim \mathcal{U}[12.0, 15.5] \\
&\alpha_3 \sim \mathcal{U}[0.5, 1.5] && \log M_{\min_3} \sim \mathcal{U}[11.5, 13.5] && \log M_{1_3} \sim \mathcal{U}[12.5, 15.5] \\
&\alpha_4 \sim \mathcal{U}[0.5, 1.5] && \log M_{\min_4} \sim \mathcal{U}[13.0, 15.5] && \log M_{1_4} \sim \mathcal{U}[13.0, 15.5],
\end{aligned}$$

where the subindex i denotes the redshift bin. As in 4.1, an autocorrelation analysis was also carried out, which will be just touched upon (and not counted as part of the main four).

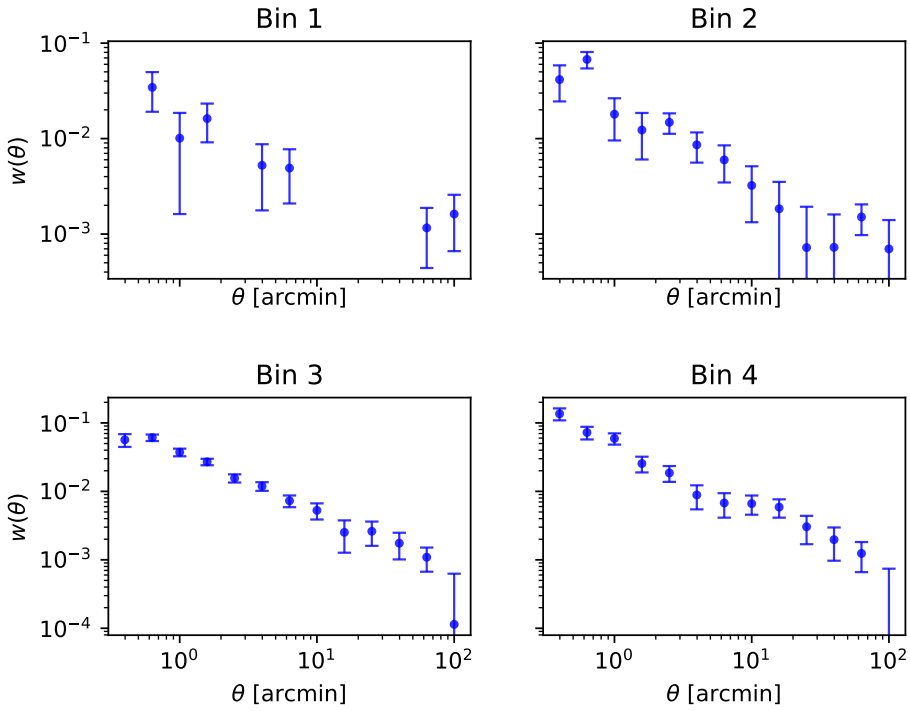

Figure 5.1: Angular cross-correlation measurements used in 5.1. The data for foreground bins 1 to 4 are depicted from left to right and from top to bottom.

The second case we studied was a semi-free flat ΛCDM model, where, in addition to the HOD, the cosmological parameters Ω_m, σ_8 and h were included in the MCMC analysis. The prior distributions were not modified with respect to 4.1, that is, we assumed

$$\Omega_m \sim \mathcal{U}[0.1, 0.8] \qquad \sigma_8 \sim \mathcal{U}[0.6, 1.2] \qquad h \sim \mathcal{U}[0.5, 1.0]. \tag{5.3}$$

The third and four cases dealt with analyzing the simple extensions of a flat CDM model with a cosmological constant, which can be carried out given the tomographic nature of this paper. Indeed, the flat w_0CDM and w_0w_aCDM models, described in 2.1.2, are characterized by one and two additional parameters, respectively, shich we introduced in the MCMC algorithm with prior distributions

$$w_0 \sim \mathcal{U}[-2.0, 0.0] \qquad w_a \sim \mathcal{U}[-3.0, 3.0]. \tag{5.4}$$

The results from all MCMC runs will lastly be compared to the ones from the non-tomographic analyses of the previous chapter, as well as with the findings of other independent cosmological probes.

5.1.1 A constant dark energy density

Fixed cosmology

The summarized statistical results from the first MCMC run are shown in Table 5.1, whereas Figures 5.2 (solid lines) and 5.3 show the marginalized distributions of the HOD parameters and the corner plots for each redshift bin, respectively. The first thing to notice is that bin 1 shows the poorest constraints on the HOD, with only an upper bound for $M_{\rm min}$ and a lower one for M_1, which is to be expected given the large error bars and few data points, as shown in Figure 5.1. The α parameter is, unsurprisingly, prior-dominated. These results were expected to a certain point, given the large error bars in the cross-correlation data for this redshift range.

Table 5.1: Parameter prior distributions and summarized marginalized statistical results for the first MCMC run of 5.1: a tomographic analysis with a fixed flat ΛCDM cosmology.

Parameter	Prior	Median	Mean	Mode	68% CI
α_1	$\mathcal{U}[0.5, 1.5]$	0.93	0.96	0.5	$[0.00, 1.12]$
$\log M_{min_1}$	$\mathcal{U}[10.0, 13.0]$	10.76	10.82	10.00	$[10.00, 11.09]$
$\log M_{1_1}$	$\mathcal{U}$[11.0,15.5]	13.93	13.79	14.63	$[13.28, 15.50]$
α_2	$\mathcal{U}[0.5, 1.5]$	0.97	0.98	0.5	$[0.50, 1.50]$
$\log M_{min_2}$	$\mathcal{U}[11.0, 13.0]$	11.76	11.76	11.80	$[11.37, 12.21]$
$\log M_{1_2}$	$\mathcal{U}$[12.0,15.5]	14.13	14.04	15.41	$[13.56, 15.41]$
α_3	$\mathcal{U}[0.5, 1.5]$	1.11	1.09	1.16	$[0.92, 1.43]$
$\log M_{min_3}$	$\mathcal{U}[11.5, 13.5]$	12.48	12.42	12.54	$[12.25, 12.80]$
$\log M_{1_3}$	$\mathcal{U}$[12.5,15.5]	13.10	13.18	12.95	$[12.54, 13.36]$
α_4	$\mathcal{U}[0.5, 1.5]$	1.07	1.05	1.50	$[0.88, 1.50]$
$\log M_{min_4}$	$\mathcal{U}[13.0, 15.5]$	13.66	13.64	13.69	$[13.51, 13.85]$
$\log M_{1_4}$	$\mathcal{U}$[13.0,15.5]	14.46	14.47	14.09	$[13.79, 15.03]$

Notwithstanding this issue, the average minimum halo mass shows a clear evolution with redshift, as can be seen more explicitly in the left panel of Figure 5.4. Indeed, the median of the marginalized posterior distribution of $\log M_{\rm min_i}$ is 10.76, 11.76, 12.48 and 13.66 from redshift bin 1 to 4, respectively. As for the M_1 parameter, the results are generally not very constraining, with just lower limits of $\log M_{1_i} > 13.28$ and > 13.56 at 68% for bins 1 and 2. Bins 3 and 4, however, have associated mean values of $\log M_{1_i} = 13.10^{+0.26}_{-0.56}$ and $14.46^{+0.57}_{-0.67}$ at 68% credibility, respectively. The α parameter is only relatively well-constrained in redshift bin 3, with a mean value of $\alpha_3 = 1.09^{+0.34}_{-0.17}$. These results, irrespective of their constraining nature, are all in agreement with the preliminary tomographic analysis of González-Nuevo et al. (2017) and the non-tomographic study of 4.1.

The redshift evolution of the average minimum mass is in keeping with the interpretation given by González-Nuevo et al. (2017). Indeed, working with a flux-limited sample induces an observational bias toward intrinsically brighter (and thus more massive) objects. Furthermore, there is a disagreement with respect to the auto-correlation analysis of the foreground sample (Bonavera et al., 2021). However, and as commented in 4.1, this is not unexpected, since the cross-correlation function takes into account the characteristics of the lenses, not the full sample of galaxies. In other words, the derived masses correspond only to the objects acting as lenses, not to the mean halo mass of all the galaxy sample. This makes it clear again that using the results from the auto-correlation analysis as the prior distributions for a cross-correlation study would not be correct, hence the choice of uniform priors.

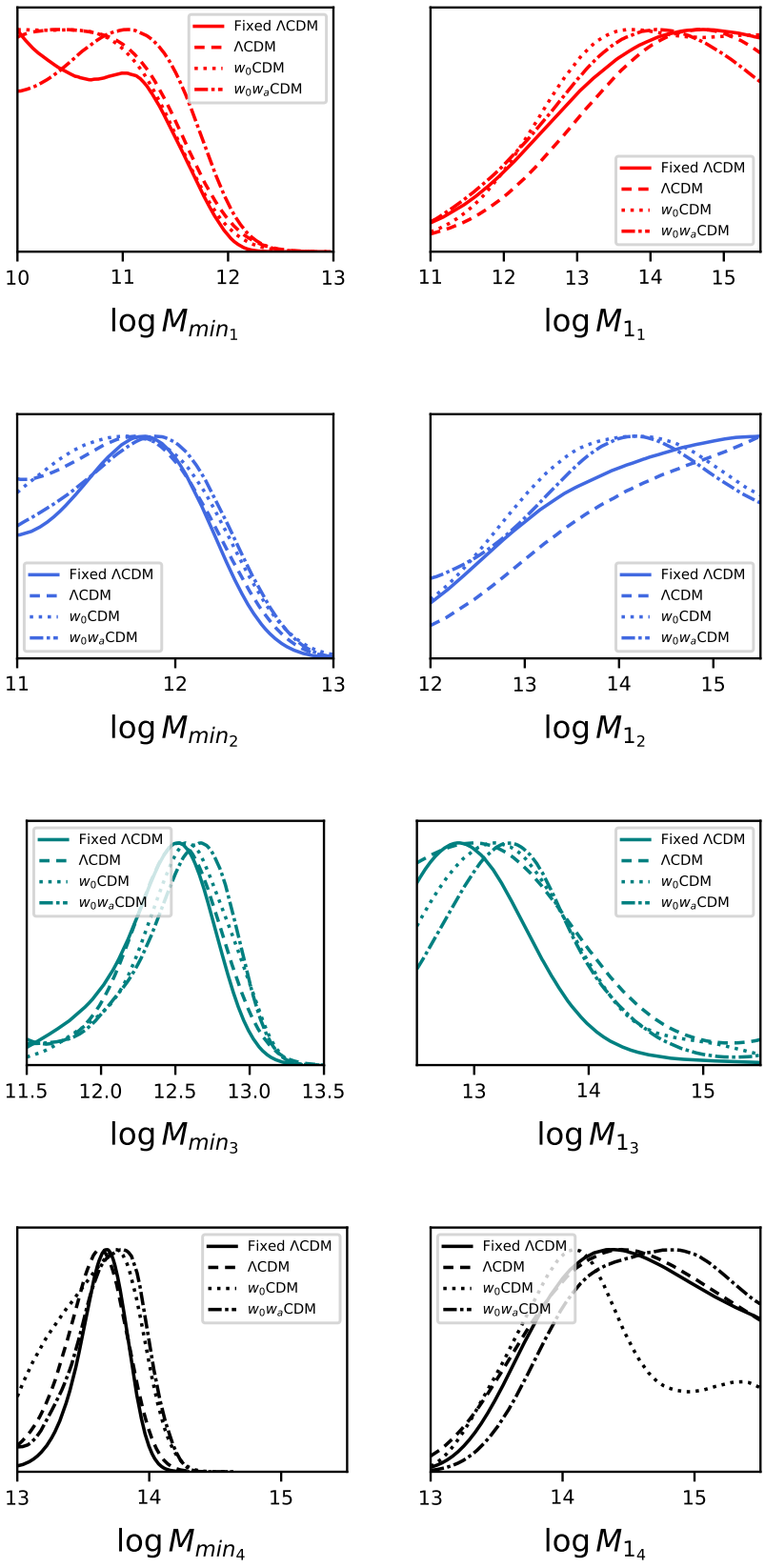

Figure 5.2: Marginalized posterior distributions for the M_{min_i} and M_{1_i} parameters in all four MCMC runs of 5.1: a tomographic analysis with a fixed flat ΛCDM cosmology (solid lines), a semi-free flat ΛCDM cosmology (dashed lines), a semi-free flat w_0CDM cosmology (dotted lines) and a semi-free flat w_0w_aCDM cosmology (dash-dotted lines).

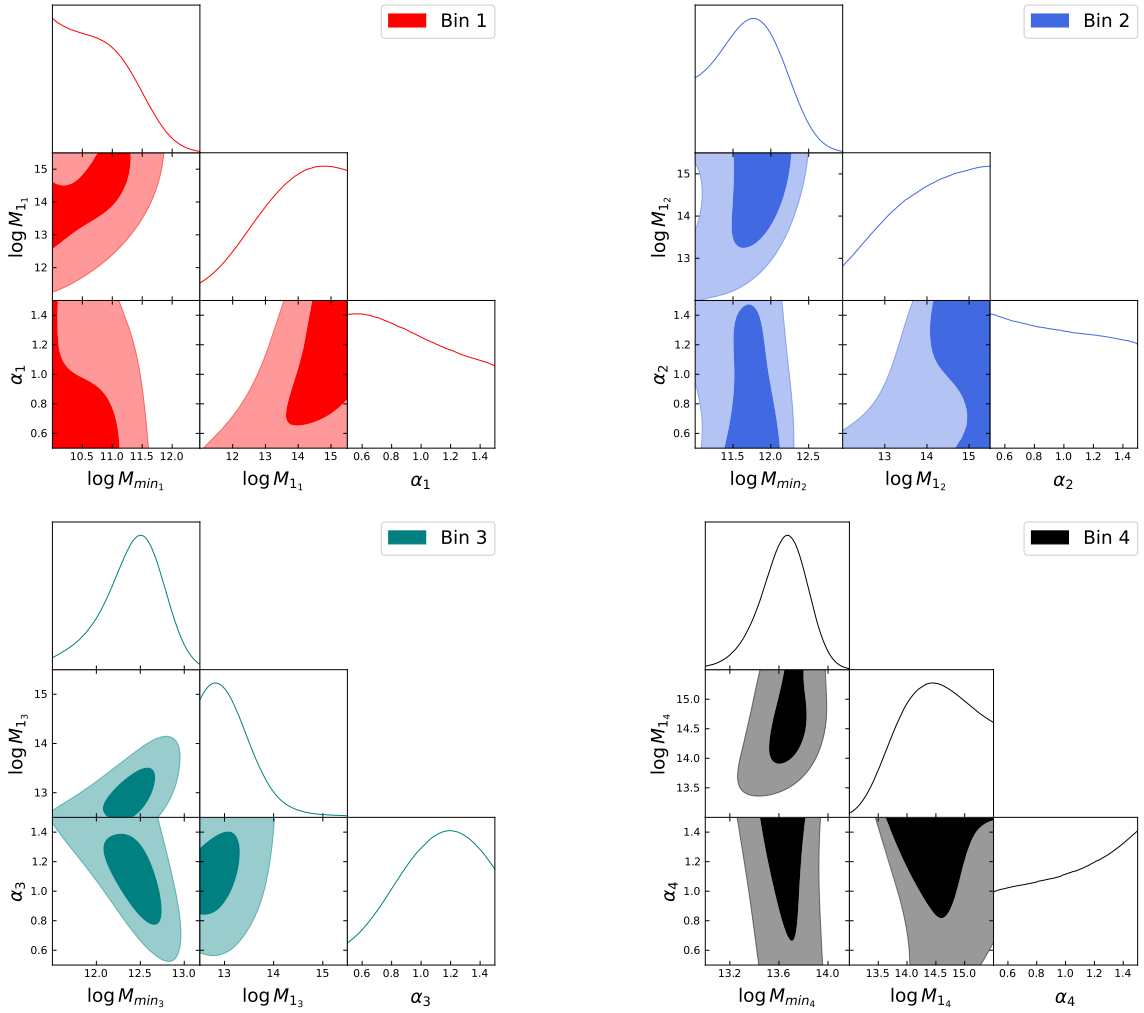

Figure 5.3: Marginalized posterior distributions and contour plots for the first MCMC run of 5.1: a tomographic analysis with a fixed flat ΛCDM cosmology.

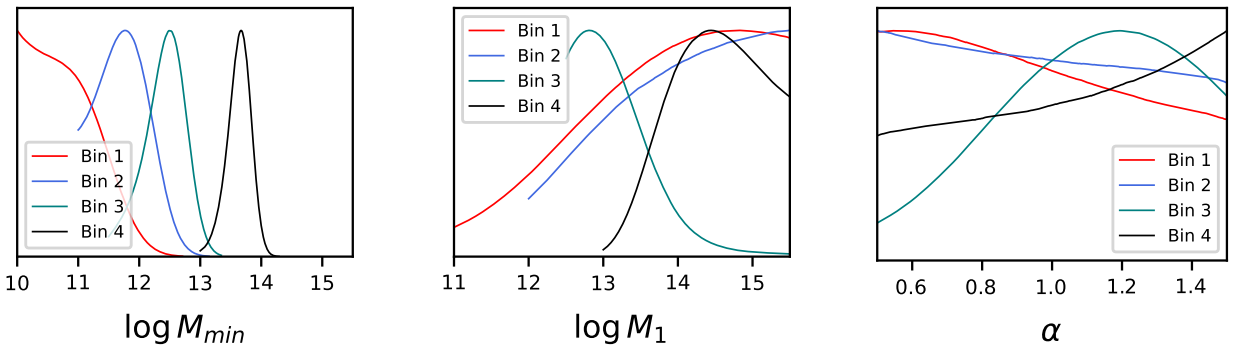

Figure 5.4: Marginalized posterior distributions for the first MCMC run of 5.1: a tomographic analysis with a fixed flat ΛCDM cosmology.

Semi-free cosmology

Table 5.2 summarizes the main statistical results from the second MCMC run, that is, a semi-free flat ΛCDM model, whereas Figures 5.5 and 5.6 show the corner plots for each redshift bin and the marginalized distribution of the cosmological parameters (with dashed lines), respectively. The HOD parameters show a general trend that is similar to that of the previous case, where the cosmology was

kept fixed. The α parameter is unconstrained in all bins, except for bin 2 in this case, where a mean value of $\alpha_2 = 0.93^{+0.14}_{-0.41}$ can be meaningfully assigned at 68% credibility. The marginalized distributions of the $M_{\min_i}$ and M_{1_i} parameters for this run are explicitly shown in Figure 5.2 (with dashed lines) and compared to those of the fixed-cosmology case (solid lines). Although with slight modifications, bin 1 presents once again the least constraining results, with practically just an upper bound on $M_{\min_1}$ and a lower bound on M_{1_1}. The results for the average minimum halo mass among bins yield once more a redshift evolution, with mean values of $\log M_{\min_i} = 10.87$, 11.74, 12.47 and 13.59 from redshift bin 1 to 4 and tighter constraints for the last two. As for M_{1_i}, the behavior is quantitatively similar to the previous case, with just lower bounds of $\log M_{1_1} > 13.49$ and $\log M_{1_2} > 13.72$ and mean values of $\log M_{1_3} = 13.46^{+0.18}_{-0.96}$ and $\log M_{1_4} = 14.44^{+0.67}_{-0.61}$.

The most interesting constraints, however, come from cosmology. Indeed, there is a substantial improvement with respect to the non-tomographic analysis of 4.1 regarding Ω_m and σ_8, as can be seen in Figure 5.6. The tomographic setup yields mean values of $\Omega_m = 0.33^{+0.08}_{-0.16}$ and $\sigma_8 = 0.87^{+0.13}_{-0.12}$, to be compared with the simple lower limit of $\Omega_m > 0.24$ at 95% and the mean value of $\sigma_8 = 0.78^{+0.07}_{-0.15}$

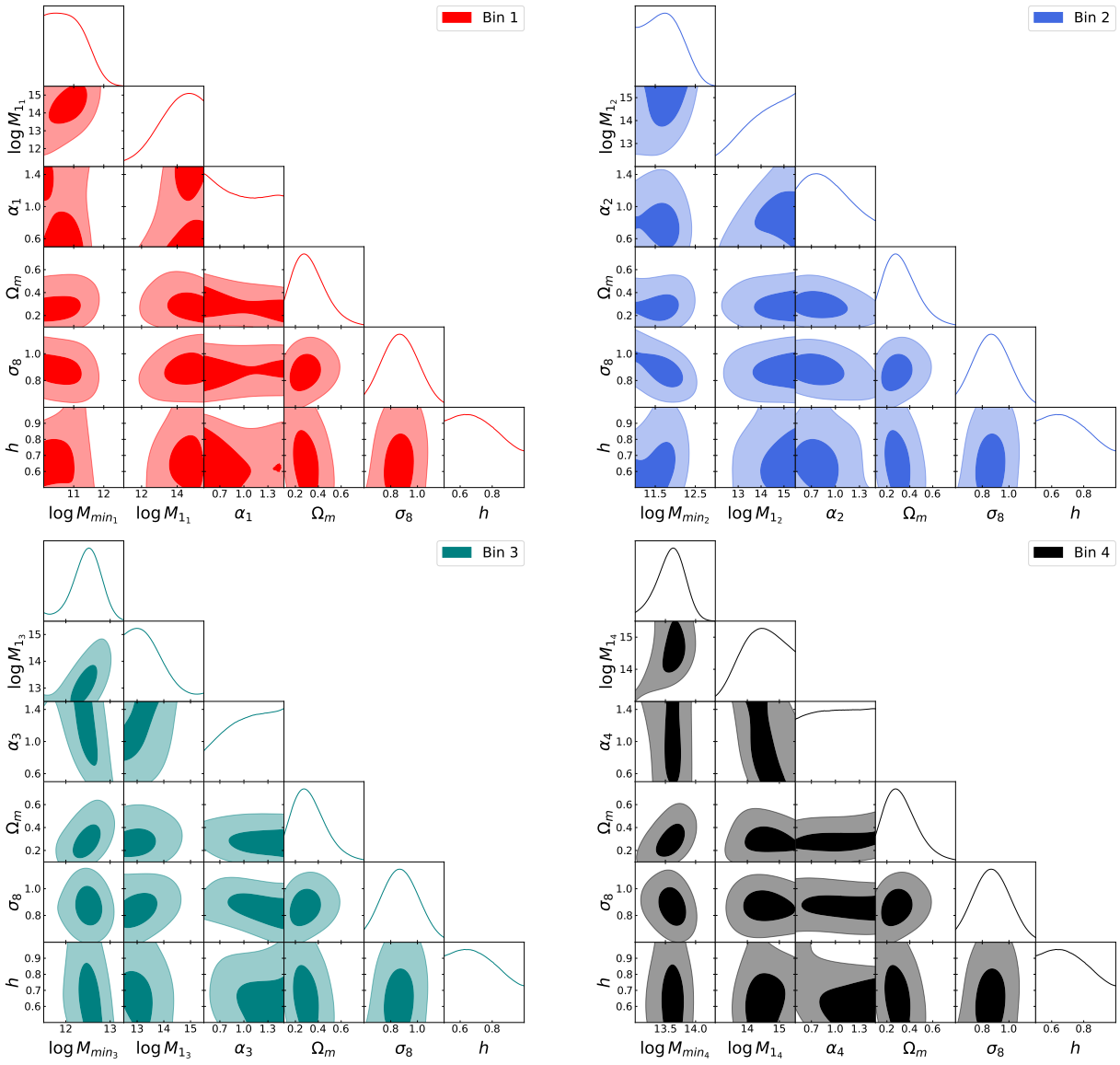

Figure 5.5: Marginalized posterior distributions and contour plots for the second MCMC run of 5.1: a tomographic analysis with a semi-free flat ΛCDM cosmology.

Table 5.2: Parameter prior distributions and summarized marginalized statistical results for the second MCMC run of 5.1: a tomographic analysis with a semi-free flat ΛCDM cosmology.

Parameter	Prior	Median	Mean	Mode	68% CI
α_1	$\mathcal{U}[0.5, 1.5]$	0.95	0.97	0.50	$[0.00, 1.50]$
$\log M_{min_1}$	$\mathcal{U}[10.0, 13.0]$	10.84	10.87	10.34	$[10.07, 11.38]$
$\log M_{1_1}$	$\mathcal{U}[11.0,15.5]$	14.04	13.91	14.43	$[13.49, 15.50]$
α_2	$\mathcal{U}[0.5, 1.5]$	0.90	0.93	0.80	$[0.52, 1.07]$
$\log M_{min_2}$	$\mathcal{U}[11.0, 13.0]$	11.75	11.74	11.92	$[11.00, 11.96]$
$\log M_{1_2}$	$\mathcal{U}[12.0,15.5]$	14.27	14.16	15.50	$[13.72, 15.50]$
α_3	$\mathcal{U}[0.5, 1.5]$	1.07	1.05	1.50	$[0.90, 1.50]$
$\log M_{min_3}$	$\mathcal{U}[11.5, 13.5]$	12.50	12.47	12.52	$[12.25, 12.80]$
$\log M_{1_3}$	$\mathcal{U}[12.5,15.5]$	13.37	13.46	13.41	$[12.50, 13.64]$
α_4	$\mathcal{U}[0.5, 1.5]$	1.01	1.01	1.50	$[0.00, 1.50]$
$\log M_{min_4}$	$\mathcal{U}[13.0, 15.5]$	13.61	13.59	13.65	$[13.43, 13.85]$
$\log M_{1_4}$	$\mathcal{U}[13.0,15.5]$	14.44	14.44	14.23	$[13.83, 15.12]$
Ω_m	$\mathcal{U}[0.1, 0.8]$	0.31	0.33	0.26	$[0.17, 0.41]$
σ_8	$\mathcal{U}[0.6, 1.2]$	0.87	0.87	0.87	$[0.75, 1.00]$
h	$\mathcal{U}[0.5,1.0]$	0.71	0.72	0.67	$[0.50, 0.79]$

that were found in 4.1. The preference for lower values of Ω_m is bound to be a consequence of the large-scale bias corrections carried out in 4.3 and the significantly tighter constraints with respect to 4.1 and 4.3 are likely due to the tomographic setup. As regards h, there is a hint of a peak at $h = 0.67$ in its marginalized distribution and a 68% upper bound on $h < 0.79$, but it is still poorly constrained.

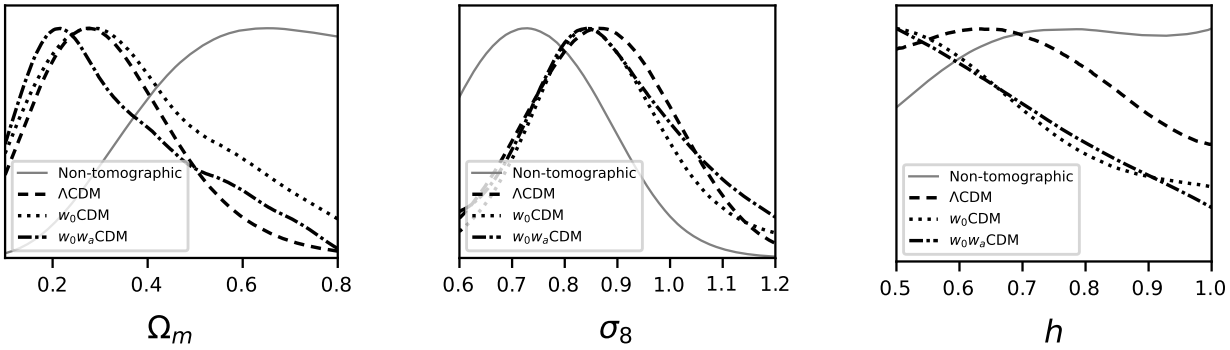

Figure 5.6: Comparison between the marginalized posterior distributions of the cosmological parameters from the MCMC runs of 5.1: a tomographic analysis with semi-free flat ΛCDM (dashed), w_0CDM (dotted) and w_0w_aCDM (dash-dotted) cosmologies. The non-tomographic results from 4.1 are also shown (solid lines).

5.1.2 An evolving dark energy density

The w_0CDM model

The summarized statistical results from the third MCMC run are shown in Table 5.3, whereas Figure 5.7 shows the corresponding corner plots for each redshift bin. As far as the HOD is concerned, there are not any significant changes with respect to the previous cases, as is emphasized in Figure 5.2, which compares this run (dotted lines) with respect to the the cases we have already discussed, namely the fixed ΛCDM cosmology (solid lines) and the semi-free ΛCDM cosmology (dashed lines). The

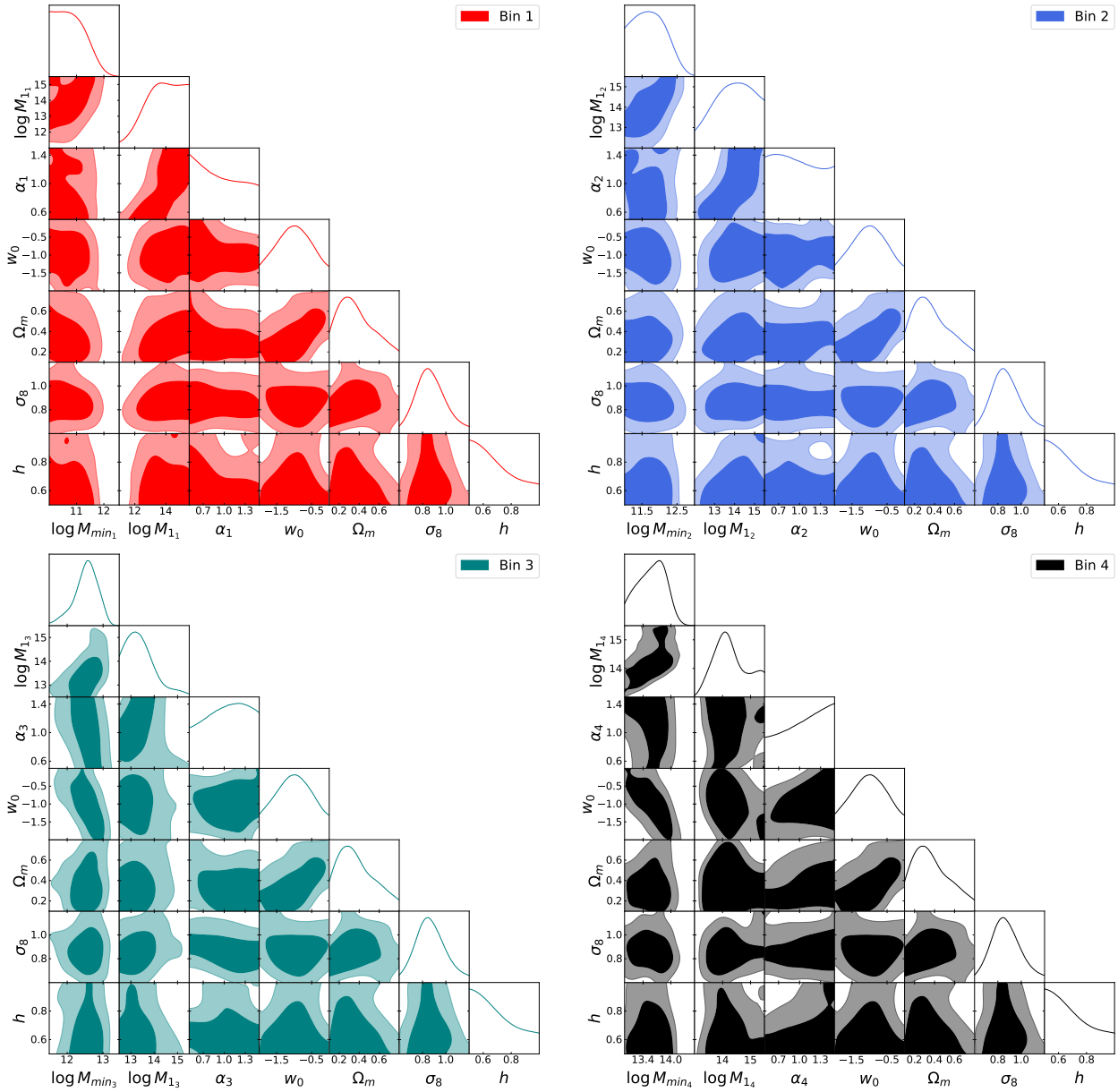

Figure 5.7: Marginalized posterior distributions and contour plots for third MCMC run of 5.1: a tomographic analysis with a semi-free flat w_0CDM cosmology.

average minimum halo mass shows can only be assigned an upper limit in bin 1, $\log M_{\min_1} < 11.11$ at 68% credibility, while the mean values show the expected redshift evolution, namely $\log M_{\min_2} = 11.75^{+0.35}_{-0.59}$, $\log M_{\min_3} = 12.53^{+0.37}_{-0.26}$ and $\log M_{\min_4} = 13.59^{+0.35}_{-0.28}$, at 68% credibility. Almost all the M_{1_i} parameters show a qualitatively similar behavior with respect to the previous cases, with a lower limit of $\log M_{1_1} > 13.23$ and mean values of $\log M_{1_2} = 13.95^{+1.23}_{-0.76}$ and $\log M_{1_3} = 13.43^{+0.30}_{-0.83}$. The M_{1_4} parameter, however, shows a suppression of the largest values with respect to the other models, with a mean value of $\log M_{1_4} = 14.27^{+0.40}_{-0.78}$.

As regards cosmology, the h parameter is once more not well constrained, with a certain tendency toward low values that was not present in the previous case. The marginalized distribution of σ_8 remains almost exactly unchanged, with a mean value of $\sigma_8 = 0.87^{+0.11}_{-0.14}$, thus allowing us to conclude that the constraining power with respect to this parameter is not diminished when considering a more general model for the dark energy equation of state. Concerning Ω_m, the distribution has a wider uncertainty with respect to the ΛCDM cosmology, but still remarkably better than the non-tomographic setup,

Table 5.3: Parameter prior distributions and summarized marginalized statistical results for the third MCMC run of 5.1: a tomographic analysis with a semi-free flat w_0CDM cosmology.

Parameter	Prior	Median	Mean	Mode	68% CI
α_1	$\mathcal{U}[0.5, 1.5]$	0.92	0.95	0.50	$[0.00, 1.12]$
$\log M_{min_1}$	$\mathcal{U}[10.0, 13.0]$	10.77	10.83	10.05	$[10.00, 11.11]$
$\log M_{1_1}$	$\mathcal{U}[11.0,15.5]$	13.84	13.77	13.74	$[13.23, 15.50]$
α_2	$\mathcal{U}[0.5, 1.5]$	0.96	0.98	0.59	$[0.50, 1.50]$
$\log M_{min_2}$	$\mathcal{U}[11.0, 13.0]$	11.72	11.75	11.67	$[11.16, 12.10]$
$\log M_{1_2}$	$\mathcal{U}[12.0,15.5]$	13.99	13.95	14.20	$[13.19, 15.18]$
α_3	$\mathcal{U}[0.5, 1.5]$	1.05	1.03	1.23	$[0.50, 1.50]$
$\log M_{min_3}$	$\mathcal{U}[11.5, 13.5]$	12.55	12.53	12.57	$[12.27, 12.90]$
$\log M_{1_3}$	$\mathcal{U}[12.5,15.5]$	13.34	13.43	13.17	$[12.60, 13.73]$
α_4	$\mathcal{U}[0.5, 1.5]$	1.10	1.07	1.50	$[0.92, 1.50]$
$\log M_{min_4}$	$\mathcal{U}[13.0, 15.5]$	13.60	13.59	13.75	$[13.31, 13.94]$
$\log M_{1_4}$	$\mathcal{U}[13.0,15.5]$	14.19	14.27	14.09	$[13.49, 14.67]$
Ω_m	$\mathcal{U}[0.1, 0.8]$	0.35	0.38	0.26	$[0.13, 0.47]$
σ_8	$\mathcal{U}[0.6, 1.2]$	0.86	0.87	0.85	$[0.73, 0.98]$
h	$\mathcal{U}[0.5,1.0]$	0.67	0.70	0.50	$[0.50, 0.75]$
w_0	$\mathcal{U}[-2.0, 0.0]$	-1.00	-1.00	-0.97	$[-1.56, -0.47]$

with a mean value of $\Omega_m = 0.38^{+0.09}_{-0.13}$. In turn, the marginalized distribution for the dark energy parameter, w_0, is remarkably symmetric, with a mean value of $w_0 = -1.00^{+0.56}_{-0.56}$. Therefore, the results are perfectly compatible with a cosmological constant.

The w_0w_aCDM model

Table 5.4 summarizes the main statistical results from the fourth MCMC run, whereas Figure 5.8 shows the corresponding corner plots for each redshift bin. Regarding the HOD parameters, they are once again rather consistent with respect to the previous cases, as shown by the dash-dotted line of Figure 5.2. The average minimum halo mass shows the already well-known redshift evolution, with mean values of $\log M_{\mathrm{min}_1} = 10.98^{+1.54}_{-0.66}$, $\log M_{\mathrm{min}_2} = 11.80^{+0.45}_{-0.50}$, $\log M_{\mathrm{min}_3} = 12.53^{+0.41}_{-0.22}$ and $\log M_{\mathrm{min}_4} = 13.69^{+0.30}_{-0.21}$. As for the M_{1_i} parameters, they are relatively better constrained, with mean values of $\log M_{1_1} = 13.71^{+1.64}_{-0.64}$, $\log M_{1_2} = 13.94^{+1.32}_{-0.66}$, $\log M_{1_3} = 13.44^{+0.36}_{-0.71}$ and $\log M_{1_4} = 14.58^{+0.73}_{-0.45}$.

Concerning cosmology, the σ_8 parameter remains unchanged with the additional freedom provided by the w_0w_aCDM model, confirming the results from the previous case: tomography-based constraints on σ_8 are not loosened when extending the base ΛCDM model. The mean value for this run is $\sigma_8 = 0.88^{+0.13}_{-0.16}$. Regarding Ω_m, the results are still less constraining than the semi-free ΛCDM case, but still far better than those from the non-tomographic analysis. The mean value is $\Omega_m = 0.34^{+0.07}_{-0.23}$. The seeming preference of low values for h is still present, although not much can be said about this. As for the dark energy equation-of-state parameters, the distribution of w_0 is as constraining as in the previous case, with a mean value of $w_0 = -1.09^{+0.43}_{-0.63}$, thus being a robust constraint against the additional freedom provided by w_a. The model is, as in the previous case, compatible with a cosmological constant. The w_a parameter displays a symmetric distribution with a mean value of $w_a = -0.19^{+1.67}_{-1.69}$, in agreement once more with ΛCDM.

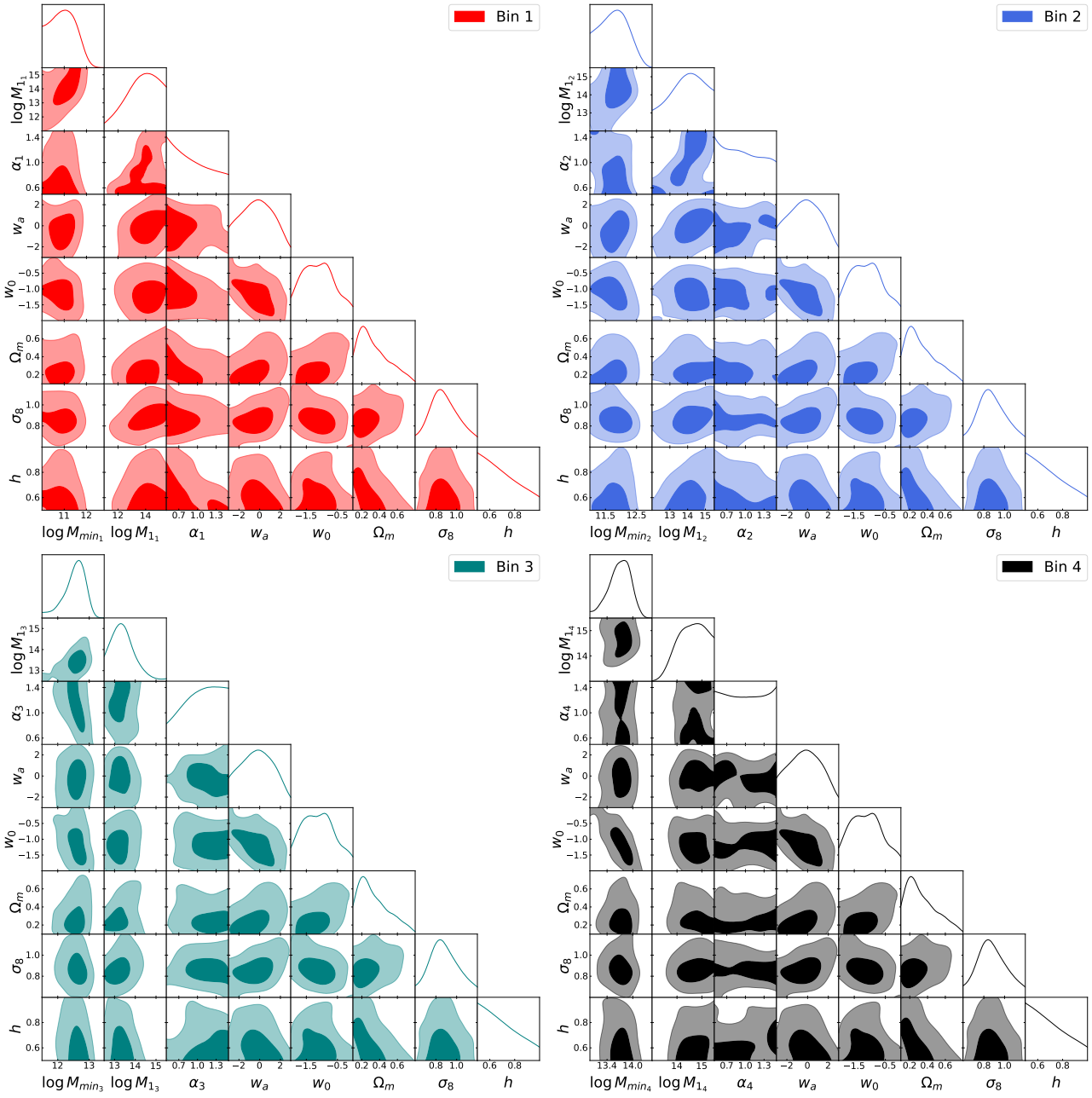

Figure 5.8: Marginalized posterior distributions and contour plots for the fourth MCMC run of 5.1: a tomographic analysis with a semi-free flat w_0w_aCDM cosmology.

5.1.3 Comparison with other results

As we did with the non-tomographic analysis of 4.1, it is informative to compare our results with the constraints obtained from the most popular alternative cosmological probes, some of which are known to be in tension.

Focusing first on the dark energy equation-of-state parameters, our constraints for w_0 are in perfect agreement with recent findings in the literature and, most importantly, they have a comparable precision when referring to *non-combined*[1] results from other experiments. For example, under a w_0CDM model, Allen et al. (2008) obtained a value of $w_0 = -1.14 \pm 0.31$ by measuring the X-ray gass mass

[1]Tighter constraints are obviously found via the combination of different observational probes, as commented in Bonavera et al. (2021) and references therein, but we would like to emphasize that we do not aim to pursue such a joint analysis ourselves.

Table 5.4: Parameter prior distributions and summarized marginalized statistical results for the fourth MCMC run of 5.1: a tomographic analysis with a semi-free flat w_0w_aCDM cosmology.

Parameter	Prior	Median	Mean	Mode	68% CI
α_1	$\mathcal{U}[0.5, 1.5]$	0.87	0.91	0.50	$[0.00, 1.06]$
$\log M_{min_1}$	$\mathcal{U}[10.0, 13.0]$	10.95	10.98	11.06	$[10.32, 11.52]$
$\log M_{1_1}$	$\mathcal{U}[11.0,15.5]$	13.82	13.71	14.08	$[13.07, 15.35]$
α_2	$\mathcal{U}[0.5, 1.5]$	0.94	0.96	0.50	$[0.50, 1.13]$
$\log M_{min_2}$	$\mathcal{U}[11.0, 13.0]$	11.79	11.80	11.88	$[11.30, 12.25]$
$\log M_{1_2}$	$\mathcal{U}[12.0,15.5]$	14.00	13.94	14.17	$[13.28, 15.26]$
α_3	$\mathcal{U}[0.5, 1.5]$	1.09	1.06	1.27	$[0.92, 1.50]$
$\log M_{min_3}$	$\mathcal{U}[11.5, 13.5]$	12.58	12.53	12.67	$[12.31, 12.94]$
$\log M_{1_3}$	$\mathcal{U}[12.5,15.5]$	13.37	13.44	13.30	$[12.73, 13.80]$
α_4	$\mathcal{U}[0.5, 1.5]$	1.00	1.00	1.50	$[0.50, 1.50]$
$\log M_{min_4}$	$\mathcal{U}[13.0, 15.5]$	13.711	13.69	13.80	$[13.48, 13.99]$
$\log M_{1_4}$	$\mathcal{U}[13.0,15.5]$	14.61	14.58	14.84	$[14.13, 15.31]$
Ω_m	$\mathcal{U}[0.1, 0.8]$	0.31	0.34	0.21	$[0.11, 0.41]$
σ_8	$\mathcal{U}[0.6, 1.2]$	0.87	0.88	0.84	$[0.72, 1.01]$
h	$\mathcal{U}[0.5,1.0]$	0.67	0.70	0.50	$[0.50, 0.76]$
w_0	$\mathcal{U}[-2.0, 0.0]$	-1.11	-1.09	-0.92	$[-1.72, -0.66]$
w_a	$\mathcal{U}[-3.0, 3.0]$	-0.21	-0.19	-0.20	$[-1.88, 1.48]$

fraction in a sample of 42 galay clusters, whereas Suzuki et al. (2012) found a value of $w_0 = -1.001^{+0.348}_{-0.398}$ within same model via the analysis of type IA supernovae. The use of *Planck*'s CMB TT, TE, EE and "low E" modes plus the lensing power spectra yielded $w_0 = -1.56^{+0.19}_{-0.39}$ for an w_0CDM model, while a value of $w_0 = -0.59^{+0.29}_{-0.26}$ is found for an w_0w_aCDM model using, for instance, the TT, TE, EE and

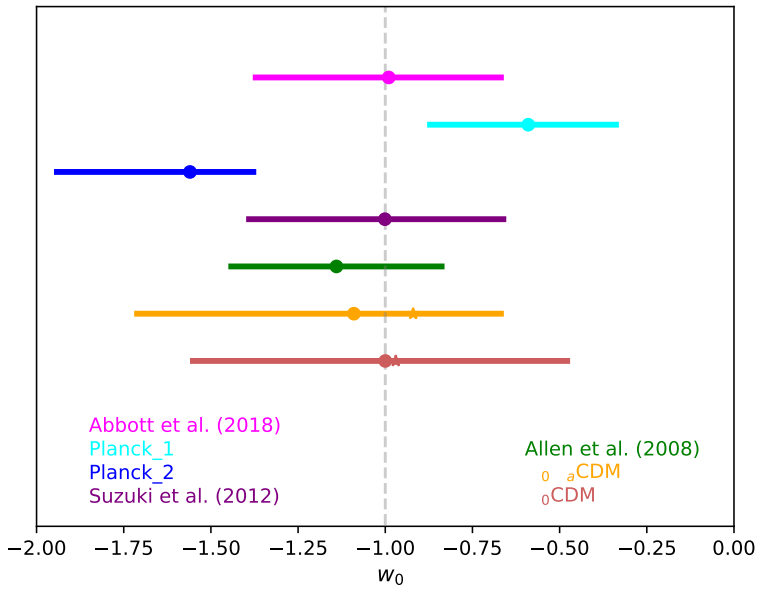

Figure 5.9: Mean (dot), median (triangle) and 68% credible intervals for the w_0 parameter found in the third and fourth MCMC runs of 5.1, namely the results from the w_0CDM (red) and w_0w_aCDM (yellow) models. For comparison, the corresponding uncombined results from Allen et al. (2008), Suzuki et al. (2012), Abbott et al. (2018) and Planck Collaboration et al. (2020a) (w_0CDM and w_0w_aCDM) as described in the text are shown in green, purple, magenta, cyan and blue, respectively.

"low E" modes plus BAO/RSD and lensing (Planck Collaboration et al., 2020a).

The first-year data release of DES (Abbott et al., 2018) provided a mean value of $w_0 = -0.99^{+0.33}_{-0.39}$ using only the correlation function of the different components of the ellipticities of galaxies, a value that has been updated to $w_0 = -0.98^{+0.32}_{-0.20}$ via the analysis of cosmic shear, galaxy-galaxy lensing and galaxy clustering (Abbott et al., 2022). A summary of these results, along with ours for comparison, is shown in Figure 5.9. Concerning the w_a parameter, *Planck*'s CMB TT, TE, EE and "low E" modes plus BAO/RSD and lensing yielded a mean value of $w_a = -0.72^{+0.62}_{-0.54}$. Furthermore, Alam et al. (2017) found a value of $w_a = -1.16 \pm 0.55$ with a cosmological analysis of BOSS galaxies together with *Planck*'s TT, TE, EE and "low E" modes plus BAO and a value of $w_a = -0.39 \pm 0.34$ with the addition of FS and supernovae. Additionally, they found a strong anti-correlation between w_0 and w_a, a hint of which (of around 30%) seems to be present in our results, as shown by Figure 5.8.

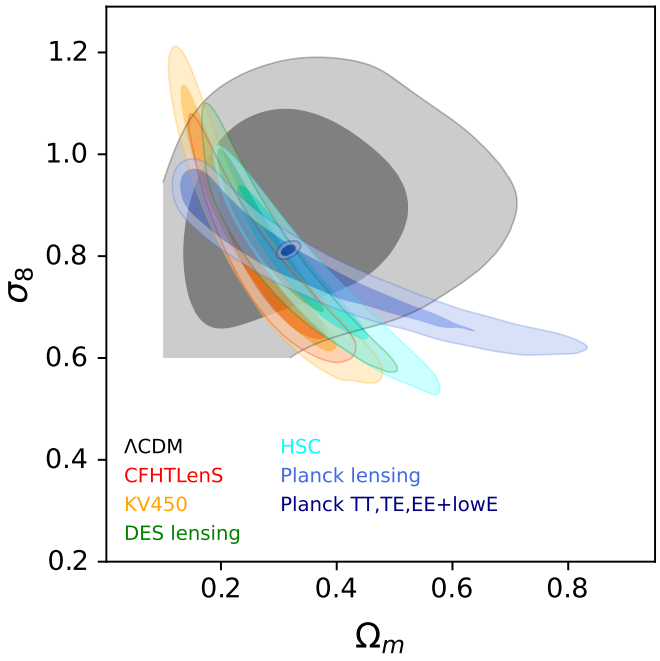

Figure 5.10: 68% and 95% posterior probability contour on the $\Omega_m - \sigma_8$ plane for the second run of 5.1: a tomographic analysis with a semi-free flat ΛCDM model (in gray). The corresponding results from other cosmological probes are shown for comparison, as described in the text.

As for Ω_m and σ_8, we focused on the CMB lensing analysis from *Planck* Planck Collaboration et al. (2020b) and the (tomographic) cosmic shear results from CFHTLens (Joudaki et al., 2017), HSC (Hamana et al., 2020), the combination of KiDS and VIKING (Hildebrandt et al., 2020) and the first-year data of DES (Troxel et al., 2018). Figure 5.10 shows, in green, the 68% and 95% contour plots on the $\Omega_m - \sigma_8$ plane obtained by Bonavera et al. (2021) within the semi-free ΛCDM model. For comparison, the results from CFHTLens, KiDS and VIKING, DES, HSC and *Planck* (both the TT, TE, EE plus low E modes and lensing) are depicted in cyan, gray, magenta, yellow and blue. Although it was already somewhat present in 4.1, the tomographic analysis confirms that the degeneracy direction usually found in cosmic shear measurements is not present in the exploitation of the submillimeter galaxy magnification bias. The identification of a "perpendicular" degeneracy direction is, however, still not possible due to the large uncertainties.

5.2 Observational constraints on the HMF

This section deals with the main results from Cueli et al. (2022), the tomographic extension of 4.2. Since this paper was developed in parallel to Bonavera et al. (2021), it made use of the same foreground sample from the GAMA II catalog, which was divided up into the four redshift bins described in 5.1, namely

$$\text{Bin 1: } z \in [0.1, 0.2] \qquad \text{Bin 2: } z \in [0.2, 0.3] \qquad \text{Bin 3: } z \in [0.3, 0.5] \qquad \text{Bin 4: } z \in [0.5, 0.8]$$

to adapt to the tomographic setup. Regarding the tiling strategy for the estimation of the cross-correlation function, given that we used the "tiles" scheme in the non-tomographic analysis of 4.2 before the large-scale bias corrections of 4.3 were applied, the non-tomographic analysis will be repeated here for both the "minitiles" and "tiles" strategies to assess any potential deviations. Following both the conclusions we will lay out and the results from 4.3, the "minitiles" scheme was later adopted for the subsequent tomographic MCMC runs.

Throughout this section, and given that no significant variations were to be expected, cosmology was kept fixed to *Planck*'s best fit values (Planck Collaboration et al., 2020a),

$$n_s = 0.965 \qquad \Omega_b = 0.049 \qquad \Omega_m = 0.3089 \qquad \sigma_8 = 0.8159 \qquad h = 0.6774, \tag{5.5}$$

as in 4.2. Following the conclusions of the non-tomographic analysis, the four-parameter Tinker fit for the HMF was abandoned in favor of the three-parameter ST model because of the the issues discussed in 4.2 regarding the apparent dependence of the statistical conclusions on the prior ranges.

As described above, the first two MCMC runs were performed in a non-tomographic setup with the corrected cross-correlation data using the "tiles" and "minitiles" schemes, respectively. The prior distributions for the HOD and HMF parameters were chosen to be uniform with the following ranges:

$$\begin{array}{lll} A \sim \mathcal{U}[0,1] & a \sim \mathcal{U}[0,10] & p \sim \mathcal{U}[-10,0.5] \\ \alpha \sim \mathcal{U}[0,1.5] & \log M_{\min} \sim \mathcal{U}[9,16] & \log M_1 \sim \mathcal{U}[9,16]. \end{array} \tag{5.6}$$

After the results from the above analysis and due to statistical rigor (to be commented next), the "minitiles" tiling scheme was chosen for the rest of the section. A tomographic MCMC run was then carried out under the assumption of universality of the HMF, that is, assuming that its parameters were not redshift-dependent. The prior distributions for the parameters included in the MCMC analysis in the third run were

$$\begin{array}{lll} A \sim \mathcal{U}[0,1] & a \sim \mathcal{U}[0,10] & p \sim \mathcal{U}[-10,0.5] \\ \alpha_1 \sim \mathcal{U}[0,1.5] & \log M_{\min_1} \sim \mathcal{U}[9,16] & \log M_{1_1} \sim \mathcal{U}[9,16] \\ \alpha_2 \sim \mathcal{U}[0,1.5] & \log M_{\min_2} \sim \mathcal{U}[9,16] & \log M_{1_2} \sim \mathcal{U}[9,16] \\ \alpha_3 \sim \mathcal{U}[0,1.5] & \log M_{\min_3} \sim \mathcal{U}[9,16] & \log M_{1_3} \sim \mathcal{U}[9,16] \\ \alpha_4 \sim \mathcal{U}[0,1.5] & \log M_{\min_4} \sim \mathcal{U}[9,16] & \log M_{1_4} \sim \mathcal{U}[9,16], \end{array} \tag{5.7}$$

where the allowed ranges for the α_i, $M_{\min_i}$ and M_{1_i} parameters were widened with respect to 5.1. Furthermore, we decided to carry out an additional tomographic analysis with the two-parameter ST fit. Indeed, although this model did not perform well in the non-tomographic setup of 4.2, this did not

have to be the case in a tomographic analysis, let alone with the additional freedom of the HOD. This two-parameter fit assumes that all the mass in the Universe is bound up into dark matter halos, which is more consistent with the underlying halo model, hence the motivation for this fourth run, for which the prior distributions were exactly the same as the previous one, with the normalization parameter A fixed by (2.99).

Lastly, with a view to assessing the plausibility of a non-universal HMF, we performed four additional non-tomographic runs in each of the foreground redshift bins independently. Assuming a three-parameter ST fit for the HMF (on the aforementioned grounds), the prior distributions for redshift bin i were thus

$$\begin{aligned} A_i &\sim \mathcal{U}[0,1] & a_i &\sim \mathcal{U}[0,10] & p_i &\sim \mathcal{U}[-10,0.5] \\ \alpha_i &\sim \mathcal{U}[0,1.5] & \log M_{\mathrm{min}_i} &\sim \mathcal{U}[9,16] & \log M_{1_i} &\sim \mathcal{U}[9,16]. \end{aligned} \tag{5.8}$$

The results from all MCMC runs will be discussed, with special focus on the potential improvements with respect to the non-tomographic analysis of 4.2.

5.2.1 The single-bin analysis revisited

The cross-correlation data in the non-tomographic setup for both the "tiles" and "minitiles" schemes are shown in Figure 5.11 in red and blue, respectively, along with the 68% full posterior sampling. The summarized statistical results from the first and second MCMC runs are shown in Table 5.5, whereas the associated corner plots are depicted in Figure 5.12, in red and blue, respectively.

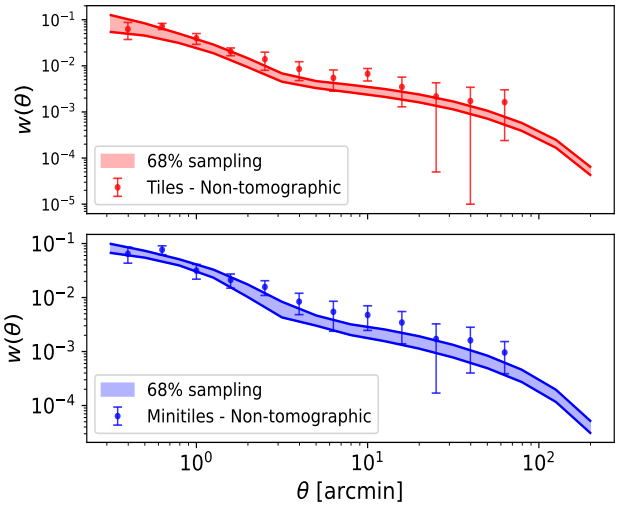

Figure 5.11: Posterior-sampled cross-correlation function (68% band) and measured data for the non-tomographic case with the "tiles" (top panel) and the "minitiles" scheme (bottom panel).

The results are very similar overall, with slightly larger uncertainties in the "tiles" case, especially for the M_1 parameter. This justifies the use of this tiling scheme in 4.2, which only aimed to provide a proof-of-concept method to observationally measure the HMF. The three HMF parameters display marginalized distributions with a clear mode, confirming the findings of 4.2 that the submillimeter galaxy magnification bias can constrain the ST parameters in a non-tomographic setup when p is allowed to take negative values and A is set free. As expected, the uncertainties increase with respect to a fixed HOD, but with a negligible variation in the distribution of p and only a widening of the

Table 5.5: Parameter prior distributions and summarized marginalized statistical results for the first and second MCMC runs of 5.2: a non-tomographic analysis with a three-parameter ST fit using the "tiles" and "minitiles" schemes, respectively.

		Minitiles			Tiles		
Parameter	Prior	Mean	Median	68% CI	Mean	Median	68% CI
$\log M_{\rm min}$	$\mathcal{U}[9, 16]$	12.02	12.35	$[11.72, 13.03]$	12.18	12.51	$[11.96, 13.09]$
$\log M_1$	$\mathcal{U}[9, 16]$	11.17	10.90	$[9.00, 11.72]$	12.07	11.77	$[9.00, 16.00]$
α	$\mathcal{U}[0, 1.5]$	0.88	0.93	$[0.68, 1.50]$	0.83	0.88	$[0.00, 1.50]$
A	$\mathcal{U}[0, 1]$	0.18	0.16	$[0.02, 0.24]$	0.21	0.18	$[0.03, 0.28]$
a	$\mathcal{U}[0, 10]$	3.34	2.85	$[0.68, 4.28]$	3.93	3.37	$[0.85, 5.16]$
p	$\mathcal{U}[-10, 0.5]$	-1.09	-0.99	$[-1.74, 0.14]$	-0.94	-0.83	$[-2.54, 0.50]$

curve for a, the mode remaining practically unchanged. As for the normalization parameter, A, the distribution is displaced towards lower values given the freedom provided by the HOD. More concretely, the medians of the distributions are $A = 0.16$, $a = 2.85$ and $p = -0.99$ for the "minitiles" scheme and $A = 0.18$, $a = 3.37$ and $p = -0.83$ for the "tiles" scheme. As shown in Table 5.5, the uncertainties are still large, which we expected given the non-tomographic nature of these first results.

Regarding the HOD, the mean minimum halo mass, $M_{\rm min}$, is well-constrained, with a mean value of $\log M_{\rm min} = 12.02^{+1.01}_{-0.30}$ and $12.18^{+0.91}_{-0.22}$ for the "minitiles" and "tiles" scheme, respectively, and median values of 12.35 and 12.51. The M_1 parameter shows a tendency toward low values, with an upper bound

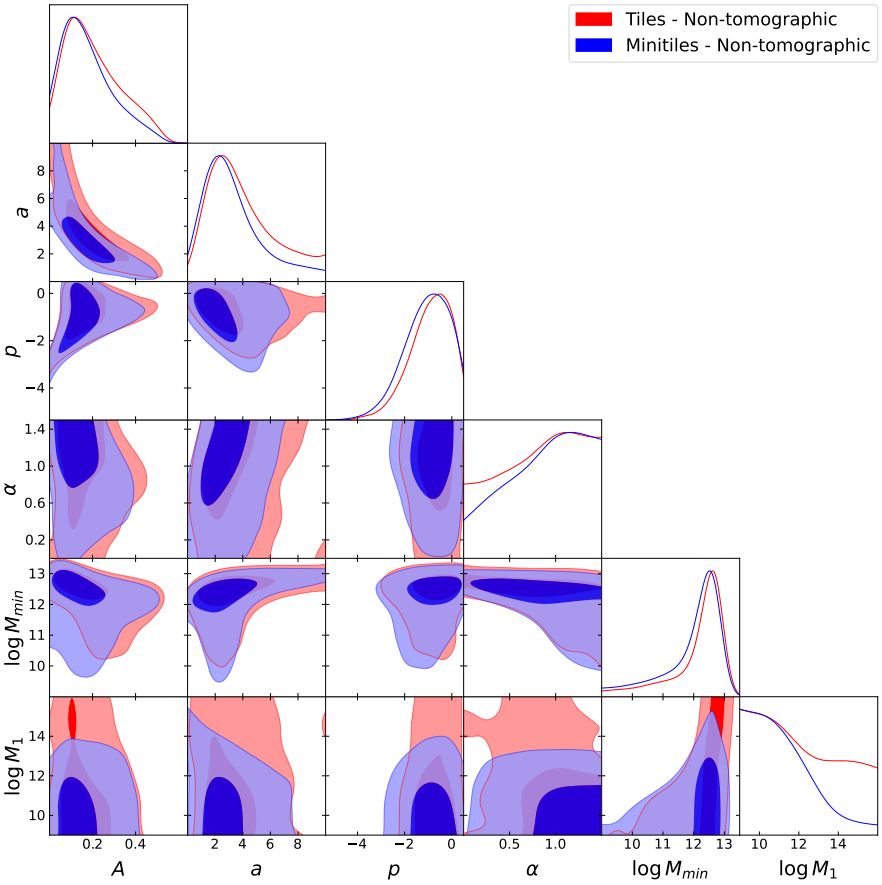

Figure 5.12: Marginalized posterior distributions and contour plots for the first and second MCMC runs of 5.2: a non-tomographic analysis with a three-parameter ST fit using the "tiles" (red) and "minitiles" (blue) schemes, respectively.

of $\log M_1 < 11.72$ at 68% credibility for the "minitiles" strategy, although it is basically unconstrained for the "tiles" scheme. The α parameter is, unexpectedly, not constrained.

For our purposes, which deal with the observational determination of the number density of halos, credible intervals for the HMF at any lens redshift can be computed via the sampling of the full posterior distribution. Figure 5.13 shows the corresponding median values and the 68 and 95 % credible intervals at several halo mass values for $z = 0.4$ for both the "tiles" (in red) and "minitiles" (in blue) schemes, along with the ST best fit found by Despali et al. (2016) under a *Planck* cosmology (green line). The first thing to notice is that the error bars are larger than those obtained in 4.2, but this is to be expected, since the HOD was fixed in that section for the final conclusions. As seen in Figure 5.13, there are not any significant differences between both tiling schemes aside from a slight tendency toward larger values of the HMF and the smaller overall uncertainties for the "minitiles". Given the large error bars, the results are compatible with the traditional values of the ST fit within the uncertainties, but the tendency found in 4.2 is hinted at again, with a lower number density of halos seemingly predicted toward the high-mass end.

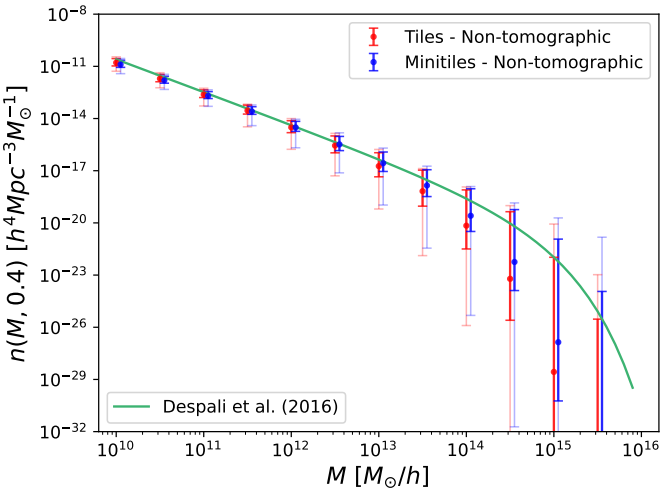

Figure 5.13: Credible intervals (68% in bold and 95% in faint colors) for the $z = 0.4$ HMF at different mass values when the full posterior distribution is sampled for the first (in red) and second (in blue) MCMC runs of 5.2: a non-tomographic analysis with a three-parameter ST fit using the "tiles" and "minitiles" schemes, respectively. The ST best fit by Despali et al. (2016) is drawn in green.

5.2.2 The tomographic setup

In light of the results from the previous subsection, which show little to no difference between both tiling schemes, and given the fact that "minitiles" imply a larger sample size for each angular scale (thus justifying the assumption of Gaussianity in the likelihood function), the "minitiles" scheme was adopted for the rest of the section. Nonetheless, the cross-correlation data in each redshift bin for both tiling schemes are depicted in Figure 5.14. It should be noted that, unlike in 4.3, there are some isolated significant differences in the data between both tiling schemes, mainly due to the reduction in sample size as a consequence of dividing up the sample into several redshift bins.

Three-parameter ST fit

The marginalized statistical results from the third MCMC run, that is, the three-parameter ST fit in a tomographic analysis, are shown in Table 5.6. For visual purposes, the part of the full corner plot

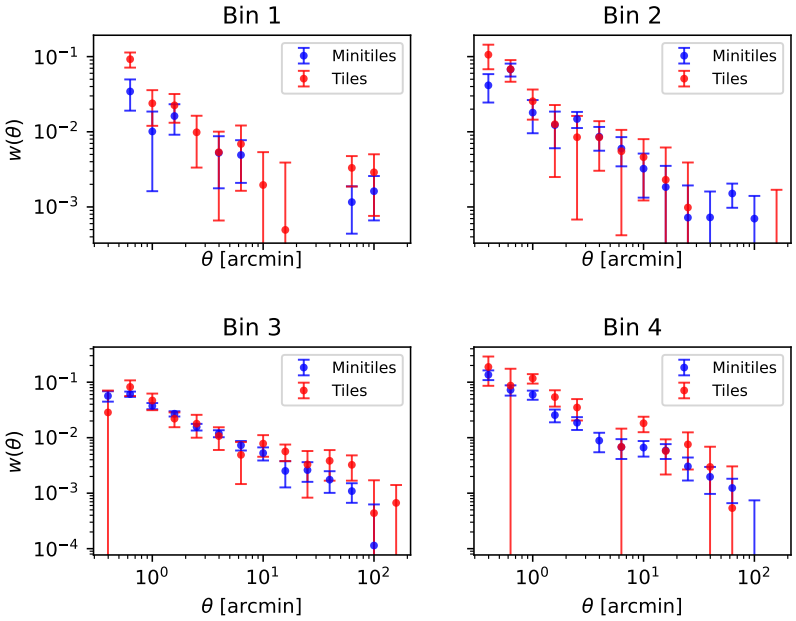

Figure 5.14: Angular cross-correlation measurements for the "tiles" (in red) and "minitiles" (in blue) schemes. The data for foreground bins 1 to 4 are depicted from left to right and from top to bottom.

concerning just the HMF parameters is shown separately in Figure 5.15 (in black, to be compared with the non-tomographic results, in blue), as are the marginalized posterior distributions of the HOD in Figure 5.16, with a different black line for each redshift bin. The full corner plots (one for each redshift bin) are depicted in Figure 5.17 (in black), where they are compared against the non-tomographic results (in blue).

The improvement with respect to the non-tomographic case is patently clear. As shown by Figure

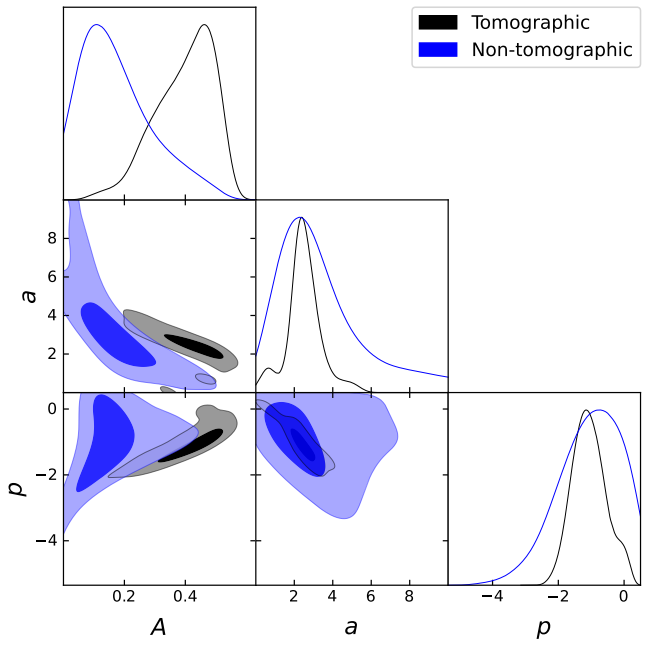

Figure 5.15: Marginalized posterior distributions and contour plots for the HMF parameters for the first and third runs of 5.2: a non-tomographic (in blue) and a tomographic (in black) analysis with a three-parameter ST fit.

Table 5.6: Parameter prior distributions and summarized marginalized statistical results for the third MCMC run of 5.2: a tomographic analysis with a three-parameter ST fit.

Parameter	Prior	Median	Mean	68% CI	95% CI
A	$\mathcal{U}[0,1]$	0.41	0.40	$[0.33, 0.52]$	$[0.21, 0.56]$
a	$\mathcal{U}[0,10]$	2.45	2.51	$[1.76, 3.24]$	$[1.93, 4.33]$
p	$\mathcal{U}[-10, 0.5]$	-1.09	-1.05	$[-1.64, -0.61]$	$[-2.00, 0.12]$
α_1	$\mathcal{U}[0, 1.5]$	0.62	0.65	$[0.00, 0.85]$	$[0.00, 1.50]$
$\log M_{min_1}$	$\mathcal{U}[9, 16]$	11.03	10.83	$[10.44, 11.69]$	$[9.29, 11.82]$
$\log M_{1_1}$	$\mathcal{U}[9,16]$	13.59	13.52	$[12.77, 16.00]$	$[10.88, 16.00]$
α_2	$\mathcal{U}[0, 1.5]$	0.89	0.79	$[0.00, 1.05]$	$[0.00, 1.50]$
$\log M_{min_2}$	$\mathcal{U}[9, 16]$	12.15	11.71	$[11.54, 12.71]$	$[9.00, 12.44]$
$\log M_{1_2}$	$\mathcal{U}[9,16]$	13.07	13.04	$[12.37, 16.00]$	$[10.12, 16.00]$
α_3	$\mathcal{U}[0, 1.5]$	1.34	1.23	$[1.25, 1.50]$	$[0.00, 1.50]$
$\log M_{min_3}$	$\mathcal{U}[9, 16]$	12.50	12.42	$[12.17, 12.94]$	$[11.46, 13.22]$
$\log M_{1_3}$	$\mathcal{U}[9,16]$	12.83	12.91	$[12.19, 13.40]$	$[11.84, 16.00]$
α_4	$\mathcal{U}[0, 1.5]$	0.71	0.75	$[0.00, 1.50]$	$[0.00, 1.50]$
$\log M_{min_4}$	$\mathcal{U}[9, 16]$	13.66	13.49	$[13.45, 13.89]$	$[11.64, 14.19]$
$\log M_{1_4}$	$\mathcal{U}[9,16]$	14.73	14.60	$[14.20, 16.00]$	$[12.50, 16.00]$

5.15, the posterior distributions of both a and p become significantly narrower, with mean values of $a = 2.51^{+0.73}_{-0.75}$ and $p = -1.05^{+0.44}_{-0.59}$ and median values of 2.45 and -1.09, respectively. Notice that the non-negligible probability region at $a < 1$, also present on the $A - a$ plane, is not due to the lack of convergence of the MCMC algorithm, but to a small number of walkers that remain in the vicinity of those values from the beginning of their path. Regarding the marginalized distribution of the normalization parameter, A, it is displaced toward larger values with respect to the non-tomographic case, with a mean value of $A = 0.40^{+0.12}_{-0.07}$, although it is neither narrowed nor widened. The conclusion drawn from this run is that tomography provides a substantial improvement in terms of uncertainties.

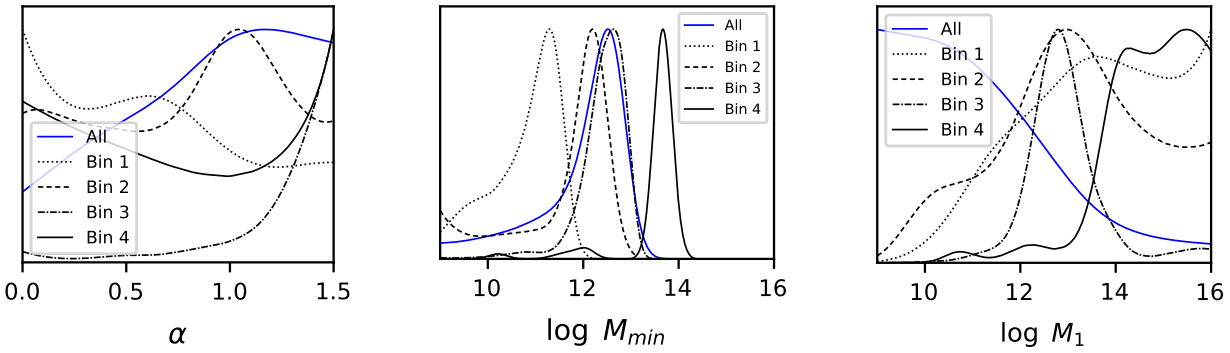

Figure 5.16: Marginalized posterior distributions for the α_i, $M_{\min_i}$ and M_{1_i} parameters in the third MCMC run of 5.2: the results from bin 1 to 4 are shown with a dotted, dashed, dash-dotted and solid lines, respectively. The non-tomographic results are depicted in blue, for comparison.

Concerning the behavior of the HOD, Figure 5.16 (in black) shows a clear time evolution of the mean minimum halo mass, $M_{\min_i}$, with mean values of $\log M_{\min_1} = 10.83^{+0.86}_{-0.39}$, $\log M_{\min_2} = 11.71^{+1.00}_{-0.17}$, $\log M_{\min_3} = 12.42^{+0.52}_{-0.25}$ and $\log M_{\min_4} = 13.49^{+0.40}_{-0.04}$. This result is in agreement with the findings of González-Nuevo et al. (2017) and 5.1, but the distributions are certainly narrower.

The remarkable improvement in terms of uncertainties with respect to the non-tomographic case is

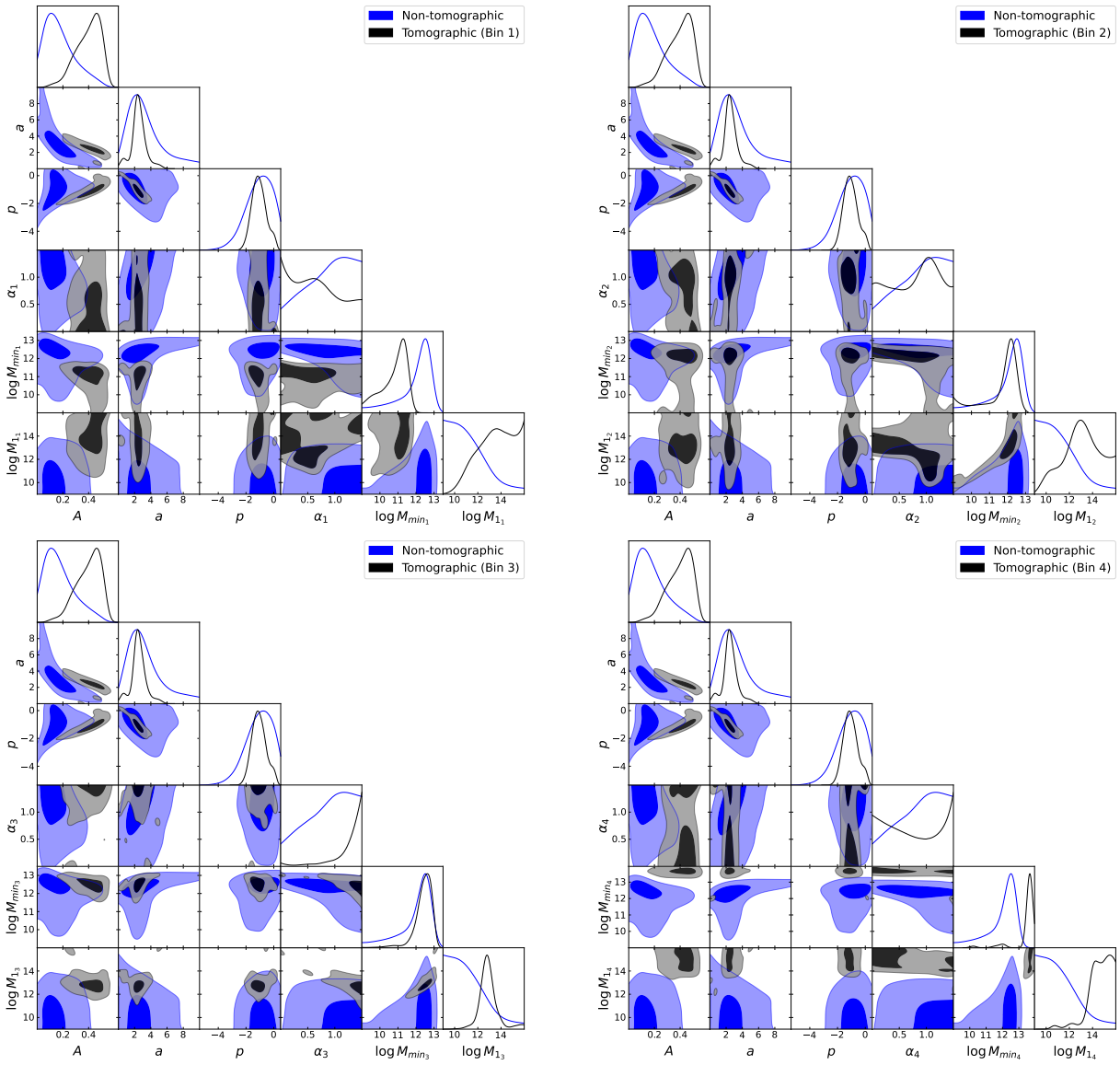

Figure 5.17: Marginalized posterior distributions and contour plots for the first and third runs of 5.2: a non-tomographic (in blue) and a tomographic (in black) analysis with a three-parameter ST fit.

highlighted in Figure 5.18, where the corresponding median values and 68% and 95% credible intervals for the $z = 0.4$ HMF are depicted in black and compared to those from the single-bin analysis, shown in blue. As expected, the uncertainties are substantially smaller at all halo masses. Most importantly, the error bars for masses below $\sim 10^{12} M_{\odot}/h$ show that our results from the three-parameter fit are not compatible with traditional simulation-based ST values at 95% credibility. In fact, under the assumption that the HMF is well fit by the three-parameter ST model, we predict a larger number density of low-mass halos and (only at 1σ) a smaller number of massive ones, that is, a steeper fall with a cutoff at $\sim 10^{14} M_{\odot}/h$.

Two-parameter ST fit

Table 5.7 shows the marginalized statistical results from the fourth MCMC run, that is, the two-parameter ST fit in a tomographic analysis, whereas Figure 5.19 (in purple) shows the full corner plots (one for each redshift bin), compared with the corresponding tomographic results from the three-parameter ST fit (in black). For visual purposes, the part concerning just the HMF parameters is

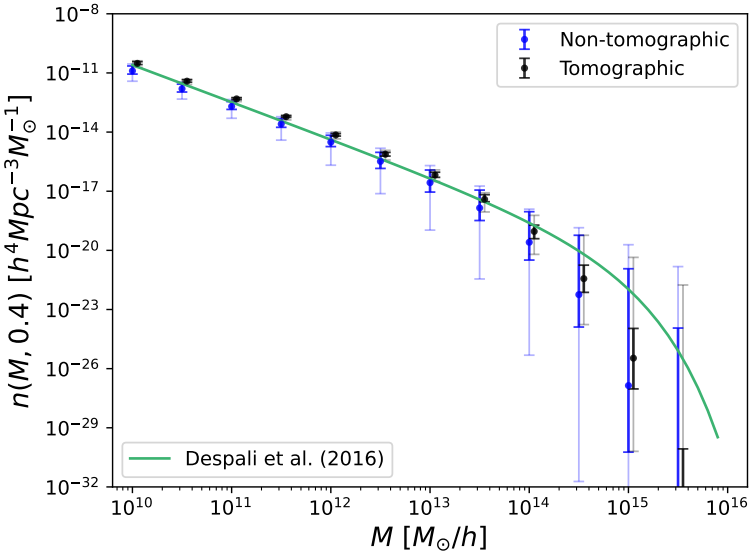

Figure 5.18: Credible intervals (68% in bold and 95% in faint colors) for the $z = 0.4$ HMF at different mass values when the full posterior distribution is sampled for the first (in blue) and third (in black) MCMC runs of 5.2: a non-tomographic and a tomographic analysis with a three-parameter ST fit. The ST best fit by Despali et al. (2016) is drawn with a green line.

shown in Figure 5.20.

The most remarkable result is the fact that the constraints on the HMF parameters (a and p) are even tighter with respect to the three-parameter fit and that the difficulties that were present in the single-bin study can be surmounted by the tomographic setup. In particular, mean values of $a = 2.47^{+0.43}_{-0.46}$ and $p = -1.29^{+0.36}_{-0.37}$ are found at 68% credibility, with a median of 2.45 and -1.30, respectively.

Regarding the HOD, Figure 5.19 shows once again the expected redshift evolution of the average minimum halo mass in bins 1, 2 and 4, with mean values of $\log M_{\mathrm{min}_1} = 10.75^{+1.08}_{-0.60}$, $\log M_{\mathrm{min}_2} = 11.38^{+1.27}_{-0.59}$ and $\log M_{\mathrm{min}_4} = 13.12^{+0.97}_{-0.06}$. However, the one-dimensional posterior distribution of M_{min_3} displays an unusually large probability region for low masses, which appears to break this tendency.

Table 5.7: Parameter prior distributions and summarized marginalized statistical results for the fourth MCMC run of 5.2: a tomographic analysis with a two-parameter ST fit..

Parameter	Prior	Median	Mean	68% CI	95% CI
a	$\mathcal{U}[0,10]$	2.45	2.47	$[2.01, 2.90]$	$[1.59, 3.44]$
p	$\mathcal{U}[-10, 0.5]$	-1.30	-1.29	$[-1.66, -0.93]$	$[-2.05, -0.56]$
α_1	$\mathcal{U}[0, 1.5]$	0.57	0.61	$[0.00, 0.79]$	$[0.00, 1.50]$
$\log M_{min_1}$	$\mathcal{U}[9, 16]$	11.02	10.77	$[10.37, 11.72]$	$[9.00, 11.61]$
$\log M_{1_1}$	$\mathcal{U}[9,16]$	13.39	13.29	$[12.44, 16.00]$	$[10.56, 16.00]$
α_2	$\mathcal{U}[0, 1.5]$	0.94	0.87	$[0.66, 1.39]$	$[0.00, 1.50]$
$\log M_{min_2}$	$\mathcal{U}[9, 16]$	12.09	11.64	$[11.42, 12.67]$	$[9.00, 12.41]$
$\log M_{1_2}$	$\mathcal{U}[9,16]$	13.14	13.17	$[11.68, 15.19]$	$[10.48, 16.00]$
α_3	$\mathcal{U}[0, 1.5]$	1.39	1.33	$[1.35, 1.50]$	$[0.94, 1.50]$
$\log M_{min_3}$	$\mathcal{U}[9, 16]$	10.50	10.72	$[9.00, 11.31]$	$[9.00, 12.81]$
$\log M_{1_3}$	$\mathcal{U}[9,16]$	11.03	11.21	$[9.70, 12.10]$	$[9.05, 13.35]$
α_4	$\mathcal{U}[0, 1.5]$	0.72	0.74	$[0.00, 1.50]$	$[0.00, 1.50]$
$\log M_{min_4}$	$\mathcal{U}[9, 16]$	13.64	13.63	$[13.56, 13.73]$	$[13.46, 13.80]$
$\log M_{1_4}$	$\mathcal{U}[9,16]$	14.64	14.71	$[14.21, 16.00]$	$[13.69, 16.00]$

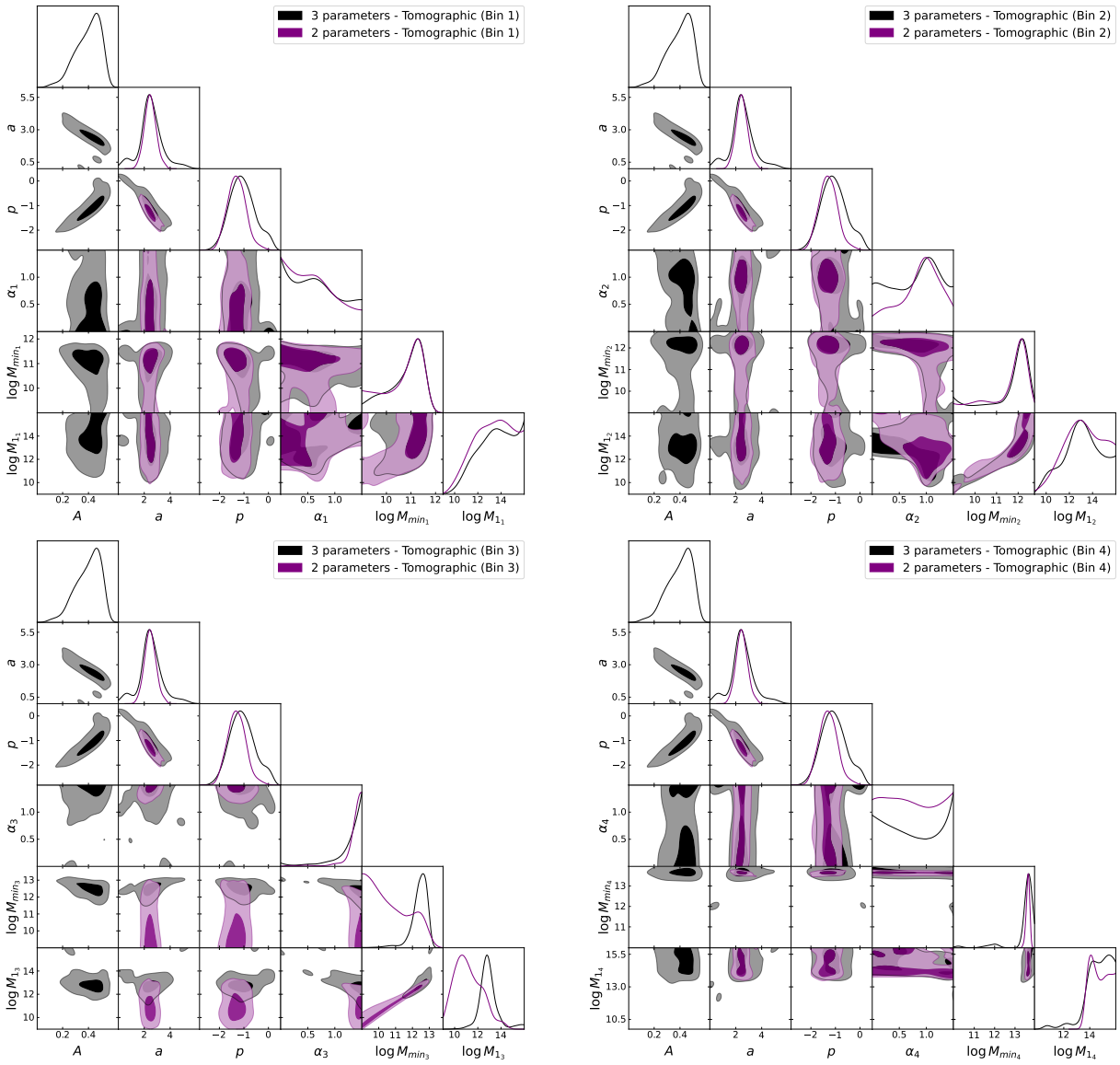

Figure 5.19: Marginalized posterior distributions and contour plots for the third and fourth MCMC runs of 5.2: a tomographic analysis with a three- (in black) and a two-parameter (in purple) ST fit.

We think that this effect is caused by the seemingly preferred low values of M_{1_3}, which in turn forces $M_{\min_3}$ to be smaller. Regardless of this aspect, the marginalized posterior distributions of both a and p converge much more quickly than the rest of parameters, thus proving the robustness of the observational constraints on the HMF with respect to the HOD, which is furthermore the main goal of this paper.

The outstanding reduction of the uncertainties in the determination of the HMF parameters with respect to the non-tomographic setup is highlighted in Figure 5.21, where median values and credible intervals for the $z = 0.4$ HMF are shown (in purple) and compared to those associated with the three-parameter fit (in black) and the ST best-fit values found by Despali et al. (2016). As stressed by this figure, the disagreement between our results and the traditional simulation-based ST values is increased for the two-parameter fit at more than 3σ, even at large masses and we can safely conclude that we predict a higher number density of halos for masses below $\sim 10^{13} M_\odot/h$ and a lower number density for masses above $\sim 10^{14} M_\odot/h$.

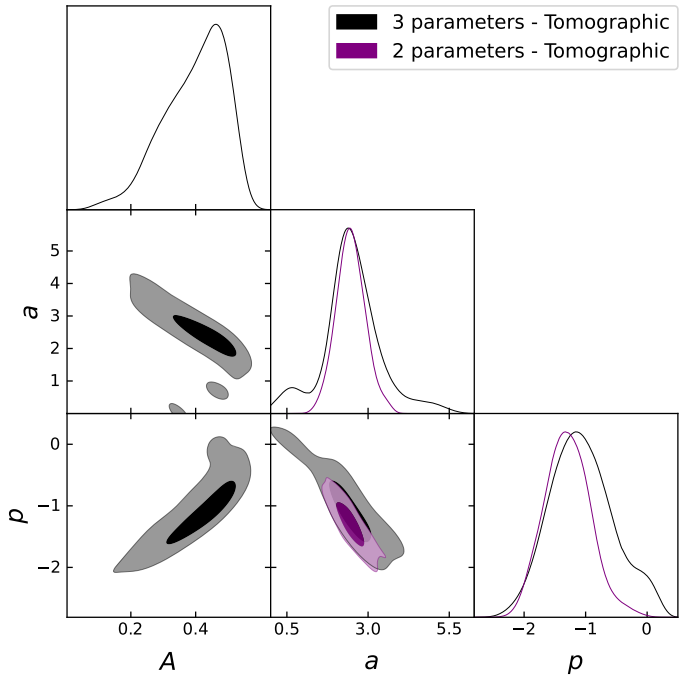

Figure 5.20: Marginalized posterior distributions and contour plots for the HMF parameters for the third and fourth MCMC runs of 5.2: a tomographic analysis with a three- (in black) and a two-parameter (in purple) ST fit.

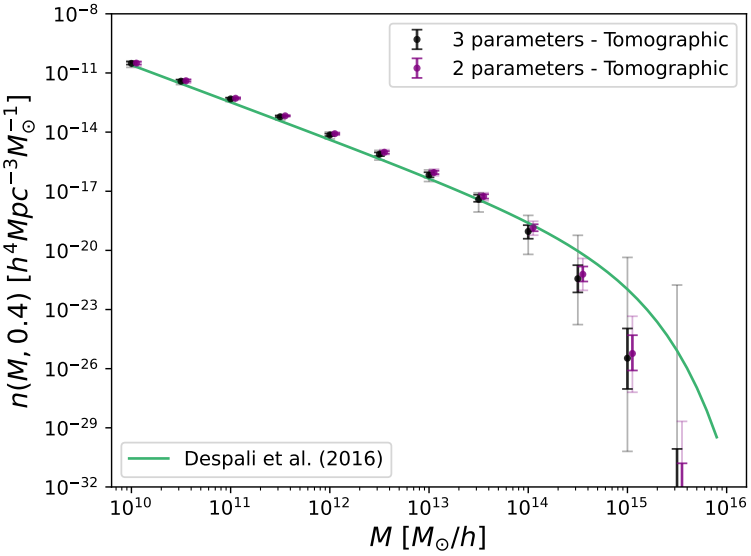

Figure 5.21: Credible intervals (68% in bold and 95% in faint colors) for the $z = 0.4$ HMF at different mass values when the full posterior distribution is sampled for the third (in black) and fourth (in purple) MCMC runs of 5.2: a tomographic analysis with a three- and a two-parameter ST fit. The ST best fit by Despali et al. (2016) is drawn with a green line.

5.2.3 A step toward testing universality

The marginalized statistical results from the fifth, sixth, seventh and eighth non-tomographic runs, that is, the three-parameter ST fit in each redshift bin independently, are shown in Table 5.8. The marginalized posterior distributions of the parameters in bins 1 to 4 are depicted in Figure 5.22 in red, cyan, orange and black, respectively, whereas the full corner plots are shown in Figure 5.23.

Focusing first on the HOD parameters, their marginalized posterior distributions show the same

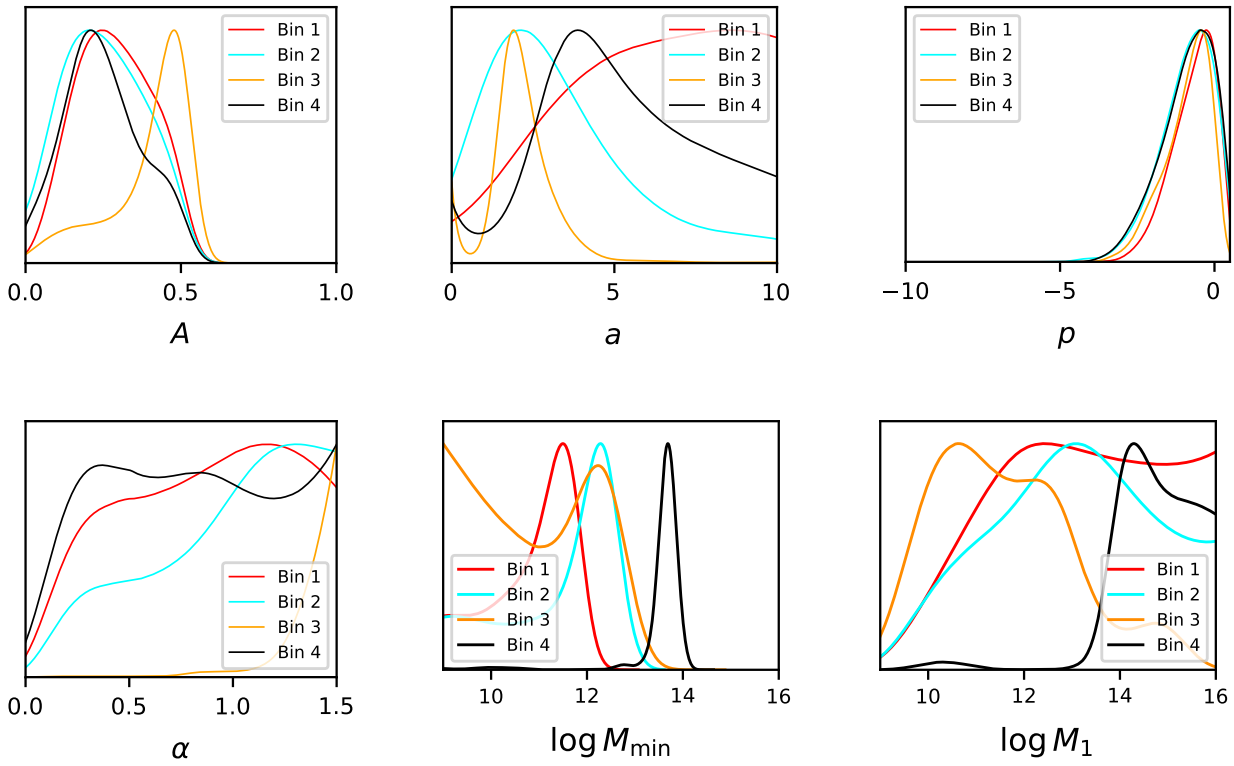

Figure 5.22: Marginalized posterior distributions for the A_i, a_i, p_i, α_i, $M_{\rm min_i}$ and M_{1_i} parameters in the fifth, sixth, seventh and eighth MCMC runs of 5.2: a non-tomographic analysis on bins 1 to 4 (in red, cyan, orange and black, respectively).

distinctive features as the two-parameter ST fit in the tomographic case. Indeed, both $M_{\rm min_3}$ and M_{1_3} take unusually low values with a very high probability. Although the reason behind this is unclear as of now, we stress that the main objective of this work is the constraining of the HMF, not finding tight bound for the HOD. Regarding the rest of redshift bins, the behavior is qualitatively similar to that of the tomographic case, with an expected larger dispersion, namely for bins 1 and 2.

The situation is much more interesting for the HMF parameters. With respect to p_i, Figure 5.22 highlights the fact that its posterior distribution barely evolves with redshift. Regarding the normalization parameters, A_i, the situation is again different in bin 3. While the posterior distributions in the rest of redshift bins remain practically unchanged, that of bin 3 is notoriously displaced to larger values with a low-probability tail at the mode of the rest of distributions, which is nonetheless translated into a smaller spread. We are confident that this behavior is related to that of the HOD in the same bin, but its analysis is beyond the scope of this work. This becomes clearer in the marginalized distributions of the a_i parameters, among which that of a_3 is the narrowest. The rest show a somehow erratic behavior, probably indicative of the non-tomographic character of the MCMC runs.

As is evident from the very large overlap in the distributions (due to the sizeable spread), universality appears to hold at first glance. To test this in a quantitative manner, we sampled the function $f(\nu, z)$ using the posterior distribution from all redshift bins, whose corresponding 68% probability bands are shown in Figure 5.24 (in red, cyan, orange and black for bins 1 to 4) along with the ST best fit (with a dotted line) found by Despali et al. (2016). The most interesting aspect is that the contour for bin 3 is extremely narrow when compared to the rest, which suffer from the long high-probability tails of the marginalized distribution of a_i, the parameter that sources most of the uncertainty in the

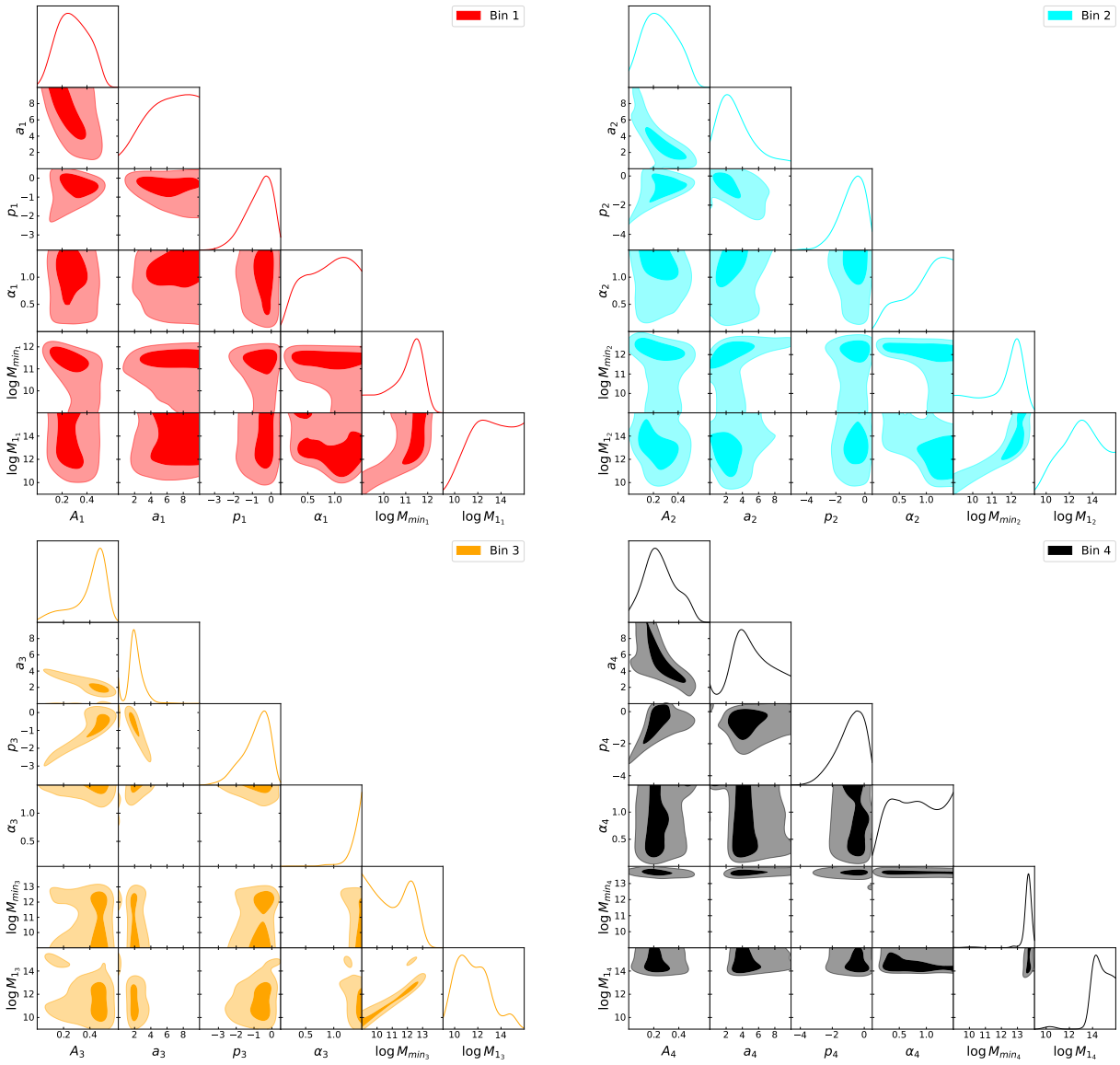

Figure 5.23: Marginalized posterior distributions and contour plots for fifth, sixth, seventh and eighth MCMC runs of 5.2: a non-tomographic analysis on bins 1 to 4 (in red, cyan, orange and black, respectively).

HMF. Therefore, the contours from bins 1, 2 and 4 overlap at every value of ν, but those from bins 1 and 3 do not (at 68% credibility). However, this is a very introductory analysis and conclusions should be drawn with caution. Indeed, it is well-known that universality in the HMF strongly depends on the mass definition of halos (Courtin et al., 2011; Despali et al., 2016; Ondaro-Mallea et al., 2022), which means that our current results, given the large uncertainties, can only be taken as preliminary. This should be more thoroughly studied by varying the halo mass definition and performing parallel tomographic analyses in different sets of redshift bins.

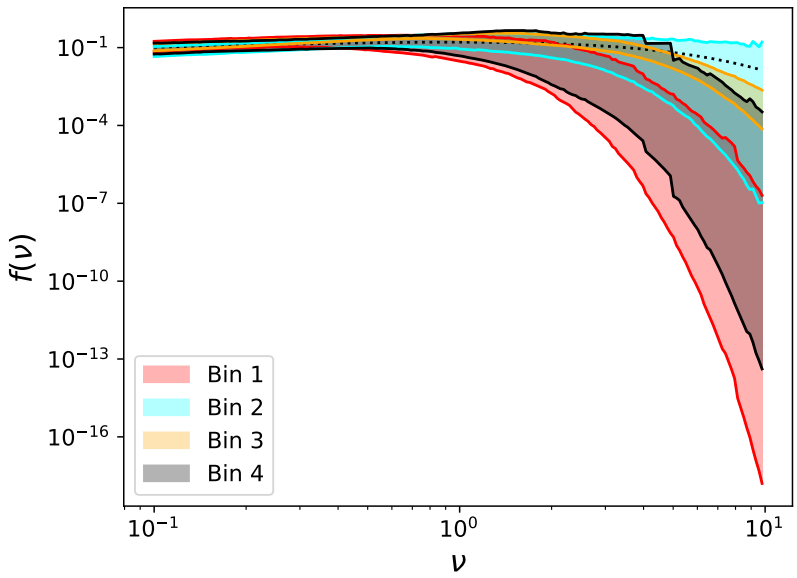

Figure 5.24: Posterior-sampled $f(\nu, z)$ function at different redshift bins. The 68% probability contours are depicted in red, cyan, orange, and black for bins 1 to 4, respectively. The ST best fit by Despali et al. (2016) is plotted with a black dotted line.

Table 5.8: Parameter prior distributions and summarized marginalized statistical results for the the fifth, sixth, seventh and eighth MCMC runs of 5.2: a non-tomographic analysis on bins 1 to 4.

Parameter	Prior	Median	Mean	68% CI	95% CI
A_1	$\mathcal{U}[0,1]$	0.28	0.29	$[0.15, 0.41]$	$[0.07, 0.52]$
a_1	$\mathcal{U}[0,10]$	6.11	5.91	$[4.61, 10.00]$	$[0.00, 10.00]$
p_1	$\mathcal{U}[-10, 0.5]$	-0.57	-0.70	$[-1.13, 0.23]$	$[-2.05, 0.49]$
α_1	$\mathcal{U}[0, 1.5]$	0.88	0.85	$[0.63, 1.50]$	$[0.19, 1.50]$
$\log M_{min_1}$	$\mathcal{U}[9, 16]$	11.22	10.97	$[10.53, 11.99]$	$[9.07, 12.09]$
$\log M_{1_1}$	$\mathcal{U}[9,16]$	13.11	13.09	$[12.12, 16.00]$	$[10.23, 16.00]$
A_2	$\mathcal{U}[0,1]$	0.25	0.26	$[0.10, 0.38]$	$[0.02, 0.50]$
a_2	$\mathcal{U}[0,10]$	2.89	3.37	$[0.48, 4.37]$	$[0.00, 8.07]$
p_2	$\mathcal{U}[-10, 0.5]$	-0.80	$-0,93$	$[-1.45, 0.18]$	$[-2.54, 0.50]$
α_2	$\mathcal{U}[0, 1.5]$	1.04	0.95	$[0.80, 1.50]$	$[0.23, 1.50]$
$\log M_{min_2}$	$\mathcal{U}[9, 16]$	12.10	11.60	$[11.13, 12.90]$	$[9.00, 12.63]$
$\log M_{1_2}$	$\mathcal{U}[9,16]$	13.02	12.96	$[11.35, 15.00]$	$[10.10, 16.00]$
A_3	$\mathcal{U}[0,1]$	0.45	0.40	$[0.36, 0.55]$	$[0.10, 0.58]$
a_3	$\mathcal{U}[0,10]$	1.99	2.14	$[1.31, 2.79]$	$[0.00, 3.84]$
p_3	$\mathcal{U}[-10, 0.5]$	-0.75	-0.89	$[-1.33, 0.07]$	$[-2.45, 0.33]$
α_3	$\mathcal{U}[0, 1.5]$	1.41	1.35	$[1.35, 1.50]$	$[0.96, 1.50]$
$\log M_{min_3}$	$\mathcal{U}[9, 16]$	11.05	11.03	$[9.00, 12.00]$	$[9.00, 12.84]$
$\log M_{1_3}$	$\mathcal{U}[9,16]$	11.45	11.64	$[9.88, 12.72]$	$[9.08, 14.99]$
A_4	$\mathcal{U}[0,1]$	0.24	0.26	$[0.10, 0.36]$	$[0.03, 0.50]$
a_4	$\mathcal{U}[0,10]$	4.86	5.17	$[2.45, 7.30]$	$[0.00, 10.00]$
p_4	$\mathcal{U}[-10, 0.5]$	-0.77	-0.91	$[-1.40, 0.26]$	$[-2.55, 0.50]$
α_4	$\mathcal{U}[0, 1.5]$	0.79	0.79	$[0.53, 1.50]$	$[0.00, 1.50]$
$\log M_{min_4}$	$\mathcal{U}[9, 16]$	13.69	13.58	$[13.49, 13.89]$	$[12.75, 14.18]$
$\log M_{1_4}$	$\mathcal{U}[9,16]$	14.64	14.63	$[13.95, 15.54]$	$[13.69, 16.00]$

6. Ongoing work and future prospects

Although the main results of this PhD thesis have already been described, I believe it useful to dedicate a brief chapter to discussing some preliminary additional findings as well as to highlighting the open future lines of work for the submillimeter galaxy magnification bias.

6.1 The h constant and the normalization of the power spectrum

One of the first things that may have caught the reader's eye in light of the results of this thesis is the apparent impossibility to obtain two-sided constraints on the adimensional Hubble constant, h. However, when the sensitivity of the cross-correlation function with respect to this parameter is analyzed, it comes as no surprise. However, as with any sensitivity analysis, one should keep track of the rest of parameters that are kept constant. Indeed, if one decides to characterize the six-parameter ΛCDM model with $\tau, n_s, \sigma_8, \Omega_m, \Omega_b$ and h, then the variation of the angular cross-correlation function with respect to h is minimal and only takes place at the largest angular scales (as shown by the left panel of Figure 6.1), where the error bars are more sizeable. However, there is freedom to choose a different set of parameters to describe the cosmological model. For instance, σ_8 is used to fix the normalization of the linear matter power spectrum,

$$P^{\text{lin}}_{\text{dm-dm}}(k, z=0) = \mathcal{N}\, T^2(k) \frac{D^2(z=0)}{D^2(z_i)} k^{n_s}, \tag{6.1}$$

according to 2.90, but what if we used the normalization constant $\mathcal{N}$ itself as a free parameter? The answer is shown in the right panel of Figure 6.1, where the variation of the cross-correlation function with respect to h is shown while keeping n_s, Ω_m, Ω_b and $\mathcal{N}$ constant. The sensitivity to h is in this case substantial at all angular scales, including where the data are much better constrained. It should be emphasized that this does not mean that running an MCMC analysis on the second set of parameters will necessarily yield tighter constraints on h than using, as we have, the first set of parameters, since the $\mathcal{N}$ parameter could absorb the uncertainties or induce new degeneracies. What it does seem to mean is that an MCMC run with a fixed value of $\mathcal{N}$ is likely to produce better results than one with a

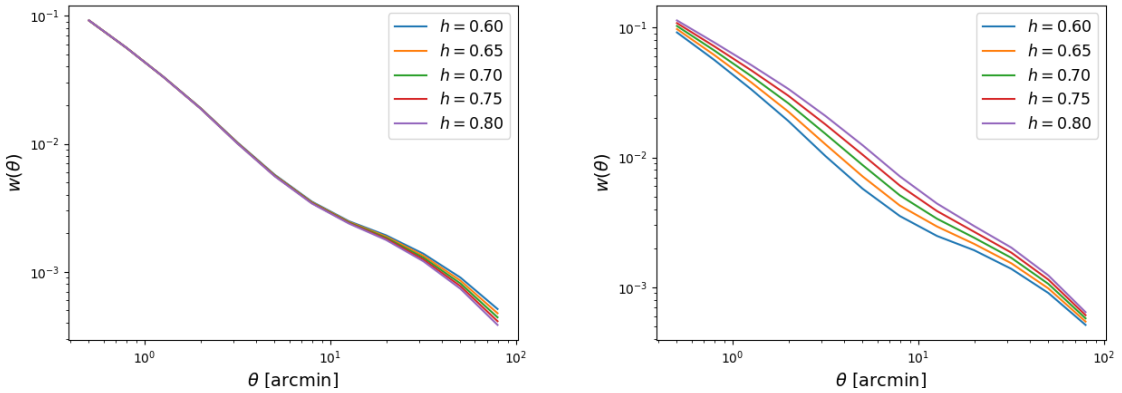

Figure 6.1: Sensitivity of the angular cross-correlation function to the h parameter. The left panel depicts the variation with fixed values of n_s, Ω_m, Ω_b and σ_8, whereas the right panel keeps n_s, Ω_m, Ω_b and the normalization parameter $\mathcal{N}$ fixed.

fixed value of σ_8. In other words, losing one (otherwise well-constrained) parameter could help toward determining another.

On this note, we ran several MCMC runs for the H-ATLAS/GAMA cross-correlation function in a non-tomographic setup. Assuming a flat ΛCDM cosmology and uniform prior distributions for the HOD (with the same ranges as in 4.2), we performed two MCMC analyses on the HOD, Ω_m and h, a first one keeping a fixed value of σ_8 and a second one doing the same for $\mathcal{N}$[1]. The corresponding contour plots and marginalized distributions for these two cases are shown in Figure 6.2 in gray and red, respectively.

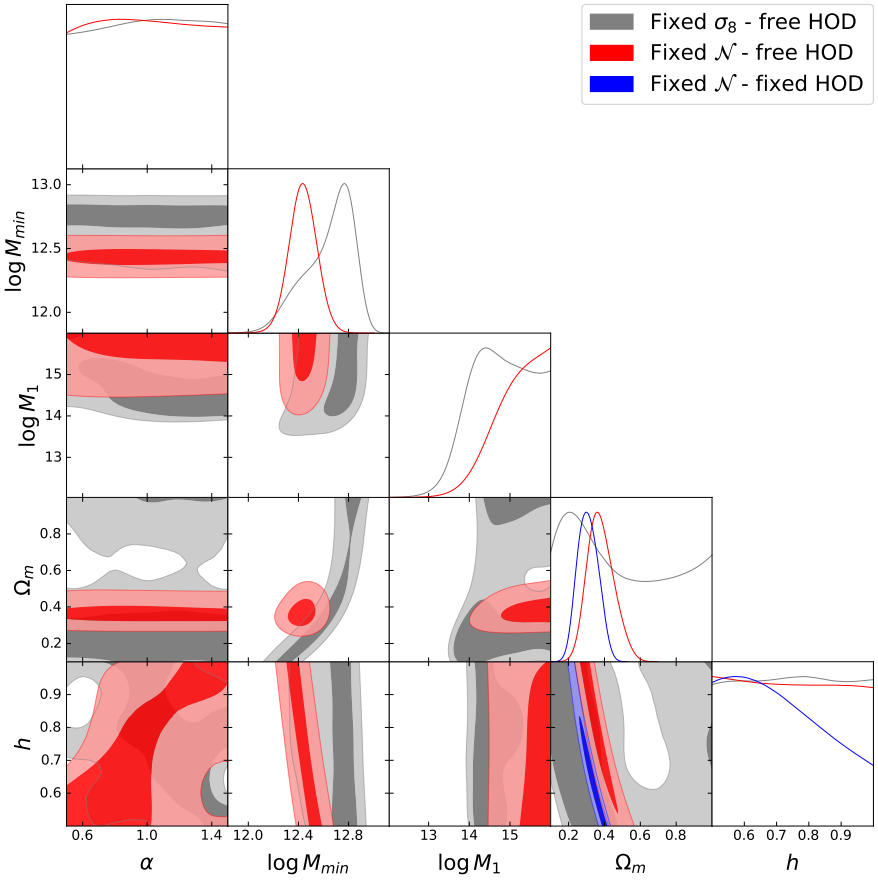

Figure 6.2: Marginalized posterior distributions and contour plots for the HOD parameters, Ω_m and h for a non-tomographic analysis with a fixed value of σ_8 (in gray) and $\mathcal{N}$ (in red). An additional case with a fixed HOD is shown in blue.

The results are bittersweet. On the one hand, fixing the normalization parameter $\mathcal{N}$ yields a well-constrained Ω_m parameter, with a mean value of $\Omega_m = 0.38^{+0.06}_{-0.08}$, to be compared with the much poorer constraints of 4.1 and 4.3. The improvement is even more remarkable if we compare it to what we obtained for the case with a fixed σ_8, which involves a bimodal distribution for Ω_m and highlights the problems of the model in this case. On the other hand, the posterior distribution for h is completely flat for both cases, which was not at all expected. A third MCMC run was carried out, this time with a fixed HOD, to shed more light on the issue, the results for which are shown in Figure 6.2 (in blue). Although the spread of the Ω_m distribution is barely affected (the shift only depends on the selected values for the HOD) and some large values of h seem to be disfavored, the issue is not resolved.

Further efforts are needed to understand this unexpected effect, including a more thorough analysis

[1]The $\mathcal{N}$ parameter was fixed to its corresponding value for a *Planck* cosmology.

of the MCMC procedure that follows the path of the walkers and studies the almost vertical degeneracy direction between Ω_m and h. Moreover, a tomographic analysis could be helpful, as it was in 5.2, when the performance of the two-parameter ST fit proved remarkable with respect to the non-tomographic setup.

6.2 Observational constraints on neutrino masses

Regardless of the potential progress made regarding the previous section, fixing the normalization parameter $\mathcal{N}$ has another clear advantage. Indeed, an original idea for this PhD thesis was to assess the power of the submillimeter galaxy magnification bias to constrain the total mass of neutrinos, but it was initially discarded given the little influence it had on our observable. Accommodating neutrino masses into the cosmological framework requires an extension of the usual six-parameter ΛCDM model. Since, to a very good approximation, cosmological observables are chiefly sensitive to the sum of neutrino masses[2] (Lesgourgues and Pastor, 2006), defined as

$$\sum m_\nu \equiv \sum_{i=1}^{3} m_i, \tag{6.2}$$

the minimal extension involves the sum of neutrino masses as an additional free parameter. Indeed, the Friedmann equation for a flat Universe in presence of massive neutrinos reads

$$\frac{H^2(z)}{H_0^2} = \left[(\Omega_{\rm dm} + \Omega_b)(1+z)^3 + \Omega_r(1+z)^4 + \Omega_\Lambda + \frac{\rho_\nu(z)}{\rho_{\rm crit}(z=0)}\right], \tag{6.3}$$

It should be noticed that the barotropic index w of the neutrino equation of state is not constant, hence the impossibility of writing a simple redshift scaling.

A careful consideration of the thermal history of neutrinos (taking into account that their decoupling from the cosmic plasma is not instantaneous given to the proximity in time to e^+e^- annihilation) yields (Lesgourgues and Pastor, 2006)

$$\Omega_\nu h^2 = \frac{\sum m_\nu}{93.14 \text{ eV}}, \tag{6.4}$$

where $\Omega_\nu \equiv \rho_\nu(z=0)/\rho_{\rm crit}(z=0)$, which can be safely applied regardless of their degeneracy in mass, since the data from neutrino oscillations show that, even if non-degenerate, only the lightest neutrino state could still be relativistic today (Bilenky, 2016), thus negligibly contributing to the energy density. For our redshifts of interest, the conclusions remain unchanged and the Friedmann equation can be written as

$$\frac{H^2(z)}{H_0^2} = \left[(\Omega_c + \Omega_b + \Omega_\nu)(1+z)^3 + \Omega_\gamma(1+z)^4 + \Omega_\Lambda\right]. \tag{6.5}$$

The effect of massive neutrinos in the process of structure formation, including their introduction into the halo model, has been extensively studied in a series of papers like Costanzi et al. (2013), Castorina et al. (2014), LoVerde (2014) or Massara et al. (2014), as well as the excellent review by Lesgourgues and Pastor (2006) and will not be discussed here, since it lies beyond the scope of this thesis. The main idea, however, is that neutrino masses introduce a free-streaming scale which leaves an imprint on the matter power spectrum and a scale-dependent growth function for density perturbations. This

[2]The neutrino flavor eigenstates, ν_e, ν_μ and ν_τ need to be clearly differentiated from the neutrino mass eigenstates, ν_1, ν_2 and ν_3, which have physical masses associated with them. Analogously to quark mixing, these two sets of states are related by the so-called Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix.

can be taken into account through the use of the Boltzmann solver CLASS for the precise computation of the transfer functions and by making the necessary modifications to the framework of the halo model following the aforementioned references. The left panel of Figure 6.3 shows the variation of the corresponding cross-correlation function with the Ω_ν parameter, while keeping Ω_m, Ω_b, n_s, h and σ_8 constant. Given that our observable is only (and barely) sensitive to this parameter at the largest angular scales, this justifies the initial decision of discarding any attempts to observationally constrain neutrino masses.

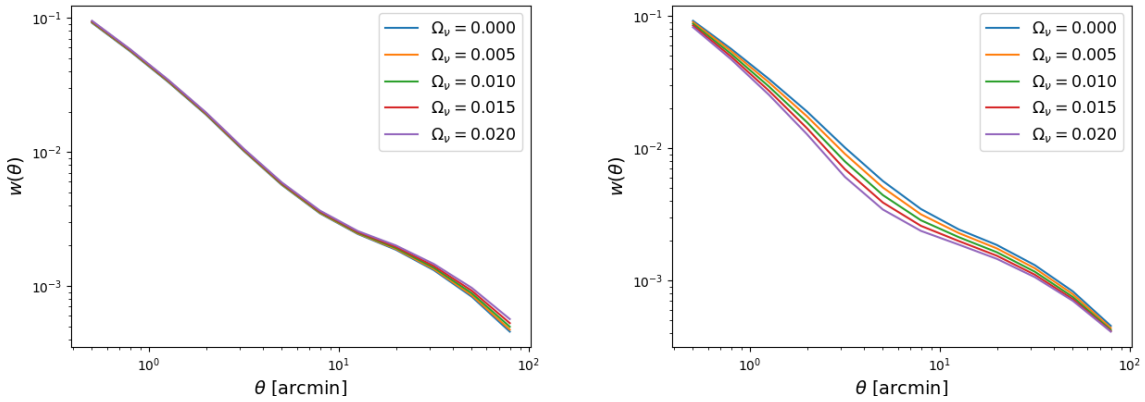

Figure 6.3: Sensitivity of the angular cross-correlation function to the Ω_ν parameter. The left panel depicts the variation with fixed values of n_s, Ω_m, Ω_b, h and σ_8, whereas the right panel keeps n_s, Ω_m, Ω_b, h and the normalization parameter $\mathcal{N}$ fixed.

However, the advantage that I mentioned at the beginning comes to life in the right panel of Figure 6.3, which shows the sensitivity of the cross-correlation function to the Ω_ν parameter, this time while Ω_m, Ω_b, n_s, h and $\mathcal{N}$ are kept constant. The difference is, once again, remarkable, since a non-negligible variation is observed at all angular scales and this could imply a promising method

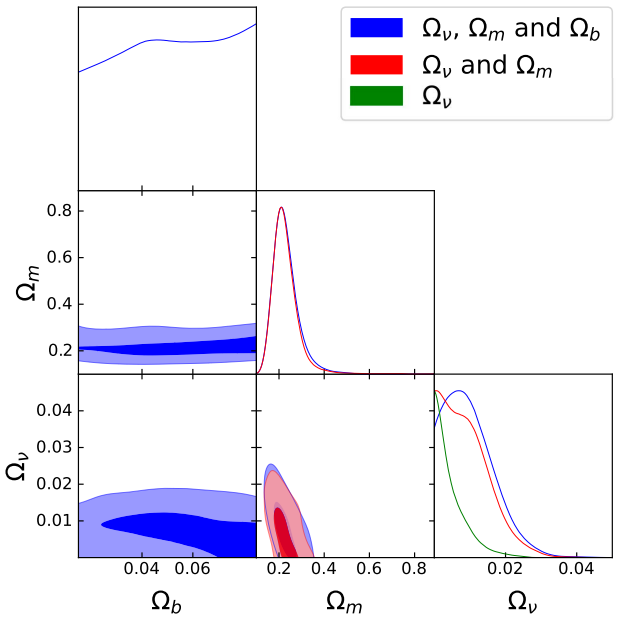

Figure 6.4: Marginalized posterior distributions and contour plots for Ω_ν, Ω_m and Ω_b for the three MCMC runs of increasing complexity as described in the text (in green, red and blue, respectively).

to constrain neutrino masses. This idea was put to the test before letting the pessimistic views of the author be biased by the results of the previous section. Three MCMC runs of increasing complexity were carried out, with the inclusion of only Ω_ν for the first one, Ω_ν and Ω_m for the second one and Ω_ν, Ω_m and Ω_b for the last one. The corresponding contour plots and marginalized distributions of these runs are shown in Figure 6.4 in green, red and blue, respectively.

The results are quite remarkable in terms of uncertainty. For the first and second MCMC runs (and as usual with cosmological constraints of neutrino masses), we only obtain upper bounds of $\Omega_\nu < 0.006$ (< 0.01) and $\Omega_\nu < 0.01$ (0.02) at 68% (95%) credibility. The third and most general run provided a significant two-sided 68% credible interval of $[0.0001, 0.0128]$ and a 95% upper bound of $\Omega_\nu < 0.02$. At this point, these constraints on the neutrino density parameter are comparable in terms of uncertainty to those from *Planck* (Planck Collaboration et al., 2020a), even when combined with other cosmological probes, which can be seen by translating our results from Ω_ν to $\sum m_\nu$ via (6.4). As regards the matter density parameter, the distribution is remarkably well-constrained for the two runs including it and there is not even a slight variation between them, with a mean value of $\Omega_m = 0.23^{+0.03}_{-0.06}$. Therefore, the inclusion of the baryon density parameter, which nonetheless cannot be constrained, does not spoil the results from the rest of parameters. Although both the Hubble constant and the HOD were kept fixed in this analysis, the results are promising enough to be hopeful that future efforts along this line can prove relevant.

6.3 Further studies regarding the theoretical modeling

Let us now discuss some of the possible studies or improvements that can be carried out on the theory side. As concerns the halo model itself, it has been shown to suffer from a number of shortcomings that could become relevant enough for precise cosmological analyses. These issues are mainly related to the modeling of the transition between the 1-halo and 2-halo regimes and/or baryonic physics (Mead et al., 2015, 2020; Philcox et al., 2020; Acuto et al., 2021) and taking them into account might contribute to a better performance of the physical framework. Moreover, the limitations of the basic HOD should be studied, as well as the feasibility of more complex analytic or simulation-based models.

The line of work that regards the HMF is itself filled with numerous possibilities. The first of them is the clear generalization to include cosmology in the MCMC analysis, whose effects on the observational determination of the HMF are expected to be small under the paradigm of universality. The potential variations in the derived number density of halos from using different HMF fits, including theory-based models like the one from Lapi and Danese (2020) should also be studied, with the assurance that a tomographic analysis would be almost certain to surmount problems regarding parameter prior ranges. Furthermore, the relevance of the halo mass definition (and perhaps the mass-concentration relation) should be examined, both for the observational constraints themselves and for any attempt at proving the redshift dependence of the HMF parameters.

An ambitious last project would entail making progress toward the characterization of the nature of dark matter, which could be pursued via the inclusion of the corresponding transfer functions (computed, for instance, by adequate Boltzmann solvers) into the halo model and the assessment of the additional changes to the physical framework. A second method, which probes smaller physical scales, would exploit the stacking technique of Bonavera et al. (2019) and Fernandez et al. (2022) and would require a thorough analysis of how the physical properties of dark matter translate into the shape of their density profile and their concentration.

6.4 Improvements on the methodology and data samples

Lastly, the uncertainty and precision of our results are expected to benefit from a refinement of the methodology and the use of foreground and background catalogs with a larger area and/or number density of sources. Indeed, throughout this thesis we have used a background sample of submillimeter galaxies extracted from the H-ATLAS catalog, but *Herschel* has observed several other regions that could be used to increase the background statistics. Although the information is publicly available in catalogs like HSPSC or HELP, an effort toward building a homogeneous sample would be needed. Furthermore, the TOLTEC camera on the Large Millimeter Telescope (Pope et al., 2019) or the mid/near-infrared James Webb Space Telescope (Gardner et al., 2006) will certainly increase the area and/or number density of background candidates. The foreground sample must of course be enlarged accordingly, for which data from Viking, DES, JPAS or the future *Euclid* mission could be exploited after the corresponding homogenization process.

Regarding the methodology, the relevance of cosmic variance needs to be seriously assessed with an independent analysis of each region in the sky and a final combination of the corresponding results. Moreover, the number and associated ranges of the redshift bins for a tomographic analysis need to be optimized to understand whether the bulk of the uncertainties arise from any particular choice. Lastly, the angular scales at which the cross-correlation function is measured need to be re-examined, since the use of an equally-spaced binning might not be the optimal choice, especially considering the importance of the transition region from the 1-halo to the 2-halo regime.

7. Summary and conclusions

The ever-increasing amount of available astronomical data makes it an exciting time to scrutinize the secrets of the *cosmos*, however titanic a task it may seem. In this thesis, I have presented the submillimeter galaxy magnification bias as a novel and independent cosmological probe that is complementary to weak lensing shear-based observables and has the potential to bring us closer to the understanding of the Universe as a whole. Let us summarize in this last chapter the main results and ideas that I have presented and discussed in this dissertation.

In a flux-limited sample, the weak lensing phenomenon of magnification bias induces a variation in the number density of background sources around a set of foreground lenses with respect to the absence of the latter. If the two samples of objects have non-overlapping redshift distributions, this effect can be quantified via a non-negligible cross-correlation signal, which probes the connection between lens galaxies and the dark matter distribution they trace. Although this idea had been known for some decades, it was abandoned long ago in favor of shear given the largely fluctuating logarithmic slope of the number counts of quasars. However, the work of González-Nuevo et al. (2014) and González-Nuevo et al. (2017), which motivated this PhD thesis, laid the foundations for its revival as a cosmological probe thanks to the optimal physical properties of submillimeter galaxies for this kind of analysis: high-redshift distribution, steep number counts and low emission in the optical band. Because of this, throughout this PhD thesis, a background sample of submillimeter galaxies from the H-ATLAS survey has been used, along with both spectroscopic and photometric samples of foreground galaxies from GAMA and SDSS, respectively.

A first preliminary non-tomographic analysis of the H-ATLAS/GAMA cross-correlation (4.1) allowed us to conclude that the binning scheme for the estimation of our observable was not relevant as far as a constant shift was concerned, except for potential lower-than-average data points, which could induce small deviations in the results. With uniform prior distributions on the halo occupation distribution (HOD), and as far as a semi-free flat ΛCDM cosmology was concerned, we found a mean value of $\sigma_8 = 0.78^{+0.07}_{-0.15}$ at 68% credibility and only a lower bound of $\Omega_m > 0.24$ at 95% credibility, whereas no meaningful information could be extracted regarding the adimensional Hubble constant, h. These results were robust against the use of more restricting (Gaussian) priors on the HOD, with only a slight reduction of the uncertainty in the σ_8 parameter. Concerning the HOD, the average minimum halo mass was the only well-constrained parameter, with a mean value $\log M_{\rm min} = 12.53^{+0.29}_{-0.16}$ for the uniform-prior case. Although not competitive, these findings had the interesting feature of not showing the typical $\Omega_m - \sigma_8$ degeneracy characterizing cosmic shear results.

In parallel with this analysis, the submillimeter galaxy magnification bias was applied to the observational determination of the halo mass function (HMF) under a fixed flat ΛCDM cosmology (4.2). Assuming both a Sheth-Tormen (ST) and a Tinker model for the number density of dark matter halos, we found that, under the assumption that all HMF parameters were positive, the ST fit could only fully explain the cross-correlation measurements by forcing the mean number of satellite galaxies in a halo to increase substantially from its prior mean value. The Tinker fit, on the other hand, provided a robust description of the data with respect to changes in the HOD parameters, but with some dependence of the statistical conclusions on the allowed prior ranges for two of its parameters. A further analysis with the former model, which did not suffer from this arbitrariness problem, allowed us to conclude that two-sided constraints could be found for all HMF parameters of the ST fit if not all of

them were forced to be positive (as it appears common in the best-fit analyses of the literature) and if the normalization was left as a free parameter. All cases were perfectly compatible with traditional simulation-based values of both the ST and Tinker models given the large uncertainties. However, there was a 1σ discrepancy for the most general "free" ST fit, with a tendency toward a larger number density of dark matter halos for masses of $10^{11.5}\,M_\odot/h < M < 10^{14.5}\,M_\odot/h$.

Aware of the impact that large-scale measurement biases could have on our results, a careful treatment of the cross-correlation estimation procedure was carried out in 4.3. The differences in the large-scale data due to choice of both the tiling scheme and the foreground sample were highlighted and corrected by computing the integral constraint value and by taking account of surface density variations on both the foreground and background samples. The first conclusion was that the "minitiles" tiling scheme was the best option, since it did not suffer from arbitrariness in the correction procedure and provided a large sample of minimal areas to reduce the effect of cosmic variance. Furthermore, the much higher number density of objects of a foreground SDSS sample did not appear to compensate for the effect of its (less precise) photometric redshifts when compared to the GAMA spectroscopic sample. When the large-scale biases were corrected, we found two-sided constraints on both Ω_m and σ_8, with mean values of $\Omega_m = 0.45^{+0.13}_{-0.12}$ and $\sigma_8 = 0.84^{+0.11}_{-0.18}$ for the GAMA sample and $\Omega_m = 0.46^{+0.11}_{-0.18}$ and $\sigma_8 = 0.99^{+0.12}_{-0.11}$ for the SDSS one, all at 68% credibility. The HOD and the h parameter show values that are similar to those found in 4.1. An additional joint analysis of both foreground samples, which served as a starting point for a tomographic study, improved the constraints on σ_8 to a mean value of $\sigma_8 = 0.75^{+0.07}_{-0.10}$ and made it clear that further studies could reduce the uncertainties on the $\Omega_m - \sigma_8$ plane.

With the corrected cross-correlation data and motivated by the previous results, a tomographic analysis of the H-ATLAS/GAMA cross-correlation function was carried out in 5.1, where both cosmology and the HOD were included in the MCMC analysis but only the latter was allowed to evolve between foreground redshift bins. In general terms, and for all three cosmological models that were studied (semi free flat ΛCDM, w_0CDM and w_0w_aCDM), a clear redshift evolution of the minimum halo mass was found, in agreement with the results of González-Nuevo et al. (2017), but the larger error bars of bins 1 and 2 made it difficult to extend this conclusion to the mean halo mass to host a single satellite galaxy. Nevertheless, the posterior distributions for the HOD were found to be very robust with respect to the considered cosmological model, with only slight isolated variations. As far as the cosmological parameters were concerned, there was a remarkable improvement with respect to the non-tomographic analysis. Indeed, mean values of $\Omega_m = 0.33^{+0.08}_{-0.16}$, $0.38^{+0.09}_{-0.13}$ and $0.34^{+0.07}_{-0.23}$ were found at 68% credibility for the ΛCDM, w_0CDM and w_0w_aCDM models, respectively, to be compared with only the lower bound of 4.1. Although the Hubble constant could only be assigned upper bounds that were not very significant, the σ_8 parameter also benefited from a tomographic setup and displayed a remarkable consistency among cosmological models, with mean values of $\sigma_8 = 0.87^{+0.13}_{-0.12}$, $0.87^{+0.11}_{-0.14}$ and $0.88^{+0.13}_{-0.16}$, respectively. Regarding the dark energy equation of state, our results were in agreement with a cosmological constant, with mean values of $w_0 = -1.00^{+0.56}_{-0.56}$ and $-1.09^{+0.43}_{-0.63}$ for the w_0CDM and w_0w_aCDM models and a mean value of $w_a = -0.19^{+1.67}_{-1.69}$. It should be noticed that the uncertainties in these two parameters were found to be comparable to those coming from non-combined results from other experiments. Lastly, the typical degeneracy on the $\Omega_m - \sigma_8$ plane characterizing cosmic shear results was once again not present, which was a further sign of the potential of our observable.

In parallel with this work, the natural tomographic extension of the observational determination of the HMF was carried out and described in 5.2. On the grounds of the conclusions from 4.2, only the

ST model was used in this analysis, but two different cases were studied, namely the "free" or three-parameter ST fit and a two-parameter ST fit obtained by requiring that all the mass be bound up into dark matter halos, an assumption which was more in keeping with the underlying halo model. The improvement with respect to the non-tomographic analysis was remarkable in both cases, even with the additional freedom of setting uniform priors on the evolving HOD. Regarding the HOD, we found a clear redshift evolution of the average halo mass, confirming the findings of 5.1. However, the most interesting findings involved the three-parameter ST fit yielding results that were not compatible at 2σ with traditional simulation-based values of the HMF, predicting a larger number density for low-mass halos and (although only at 1σ) a smaller number of massive ones. The situation was more pronounced for the two-parameter ST fit, which increased the disagreement with the usual values to more than 3σ, concluding that we predicted a higher number density of halos for masses below $\sim 10^{13}\, M_{\odot}/h$ and a lower number density for masses above $\sim 10^{14}\, M_{\odot}/h$.

This PhD thesis has presented both preliminary and follow-up improved results on the submillimeter galaxy magnification bias as a cosmological probe. The road ahead is certainly exciting and full of possibilities. Indeed, interesting cosmological constraints have been found, as well as a number of discrepancies with traditional values found in the literature, which motivates the need of further improvements both in the theoretical modeling and the estimation methodology. Future analyses will undoubtedly continue to unveil the potential of this *underexploited cosmological probe*.

It becomes essential at this point to take stock of the time and effort one has devoted to the elaboration of their PhD thesis. In my humble view, I have fulfilled both the professional and personal objectives I had set out to accomplish and I look forward to continuing down this path. As emphasized in the previous chapter, the road ahead is filled with possibilities and, although I do not share Bertrand Russel's pessimistic view on the three passions governing his life, I do believe the following quote adds an accurate finishing touch to this thesis by gathering a state of mind to which the fellow scientist is certain to have related to at some point in their life:

> *"With equal passion I have sought knowledge. I have wished to understand the hearts of men, I have wished to know why the stars shine. And I have tried to apprehend the Pythagorean power by which number holds sway above the flux. A little of this, but not much, I have achieved".*
>
> Bertrand Russel - Prologue to *Autobiography*

Resumen y conclusiones

La creciente cantidad de información astronómica disponible hace del momento actual una época emocionante para escudriñar los secretos del *cosmos*, por titánica que parezca esta tarea. En esta tesis, he presentado el sesgo de magnificación de las galaxias submilimétricas como una sonda cosmológica novedosa e independiente que es complementaria a observables basados en el *shear* del efecto de lente gravitacional débil y que tiene el potencial de acercarnos a comprender el Universo como un todo. Resumamos finalmente los resultados principales y las ideas que he presentado y discutido en este trabajo.

En una muestra de flujo limitado, el fenómeno del sesgo de magnificación, basado en el efecto de lente gravitacional débil, induce una variación en la densidad numérica de galaxias de fondo alrededor de un conjunto de lentes con respecto a la ausencia de las últimas. Si las dos muestras de objetos tienen distribuciones de *redshift* que no se solapan, este efecto puede cuantificarse a través de una señal no despreciable de correlación cruzada, la cual muestrea la conexión entre las galaxias lente y la distribución de materia oscura que trazan. Aunque esta idea ya era conocida desde hacía décadas, se abandonó hace tiempo en favor del *shear* dadas las grandes fluctuaciones en la pendiente logarítmica del número de cuentas de los cuásares. Sin embargo, los trabajos de González-Nuevo et al. (2014) y González-Nuevo et al. (2017), que motivaron la presente tesis, sentaron las bases para su renacimiento como sonda cosmológica gracias a las óptimas propiedades físicas de las galaxias submilimétricas para este tipo de análisis: distribuciones de alto *redshift*, pendiente acuciada del número de cuentas y baja emisión en el óptico. Debido a esto, a lo largo de esta tesis doctoral, se ha utilizado una muestra de fondo de galaxias submilimétricas de la misión H-ATLAS, junto con muestras de galaxias lente espectroscópicas y fotométricas, provenientes de GAMA y SDSS, respectivamente.

Un primer análisis no tomográfico de la correlación cruzada H-ATLAS/GAMA (4.1) nos permitió concluir que la estrategia de *binning* para la estimación de nuestro observable no era relevante siempre que se tratara de una traslación constante, excepto en el caso de puntos de señal más baja que la media, que podrían inducir pequeñas desviaciones en los resultados. Con distribuciones *a priori* uniformes en la distribución de ocupación de halos (DOH), y en lo que respecta a una cosmología ΛCDM plana y semilibre, hemos encontrado un valor medio de $\sigma_8 = 0.78^{+0.07}_{-0.15}$ al 68% de credibilidad y solo una cota inferior de $\Omega_m > 0.24$ al 95%, mientras que no se ha podido extraer información relevante sobre la constante de Hubble adimensional, h. Se ha demostrado la robustez de estos resultados respecto al uso de distribuciones *a priori* más restrictivas (Gaussianas) para la DOH, con solamente una pequeña reducción de la incertidumbre de σ_8. Respecto a la DOH, la masa mínima promedio ha sido el único parámetro bien restringido, con un valor medio de $\log M_{\rm min} = 12.53^{+0.29}_{-0.16}$ para el caso de distribución *a priori* uniforme. Aunque no son competitivos, estos resultados preliminares no muestran la degeneración $\Omega_m - \sigma_8$ característica de los resultados del *shear* cósmico.

En paralelo a este análisis, se aplicó el sesgo de magnificación de las galaxias submilimétricas a la determinación observacional de la función de masa de halos (FMH) bajo una cosmología ΛCDM plana fija (4.2). Suponiendo tanto un modelo Sheth-Tormen (ST) como uno de Tinker para la densidad numérica de halos de materia oscura, hemos encontrado que, bajo la suposición de que todos los parámetros FMH eran positivos, el *fit* de ST solo podía explicar las medidas de correlación cruzada forzando un incremento sustancial del número promedio de galaxias satélite en un halo respecto a su valor medio *a priori*. El *fit* de Tinker, sin embargo, ha proporcionado una descripción de los datos

robusta respecto a cambios en los parámetros HOD, pero con cierta dependencia de las conclusiones estadísticas con los rangos de las distribuciones *a priori* de dos de sus parámetros. Un análisis adicional con el modelo de ST, que no padecía de este problema de arbitrariedad, nos ha permitido concluir que se pueden encontrar restricciones bilaterales para todos los parámetros FMH del *fit* de ST si no se fuerza la positividad de todos ellos (como parece común en los análisis de literatura) y si se deja la normalización como parámetro libre. Se ha demostrado que todos los casos son perfectamente compatibles con los valores basados en simulaciones, tanto del modelo de ST como el de Tinker, dadas las grandes indeterminaciones. Sin embargo, hemos encontrado una discrepancia a 1σ para el *fit* ST más general, con una tendencia hacia una mayor densidad numérica de halos para masas de $10^{11.5} M_\odot/h < M < 10^{14.5} M_\odot/h$.

Conscientes del impacto que podían tener los sesgos de medida a gran escala, en 4.3 se llevó a cabo un tratamiento cuidadoso del procedimiento de estimación de la correlación cruzada. Tras destacar las diferencias en los datos debidas a la elección de la estrategia de *tiling* y de la muestra lente, se corrigieron mediante el cálculo del valor del *integral constraint* y teniendo en cuenta las variaciones en la densidad superficial de las muestras de fondo y lente. La primera conclusión ha sido que la estrategia de "*minitiles*" era la mejor opción, ya que no padecía de arbitrariedad en el proceso de correción y proporcionaba un tamaño de muestra grande en cuanto al número de áreas mínimas para reducir el efecto de la varianza cósmica. Además, la mayor densidad numérica de objetos lente de la muestra SDSS no ha parecido compensar el efecto de los *redshift* fotométricos (menos precisos) en comparación con la muestra espectroscópica GAMA. Con la correción de los sesgos a gran escala, hemos encontrado restricciones bilaterales tanto en Ω_m como en σ_8, con valores medios de $\Omega_m = 0.45^{+0.13}_{-0.12}$ y $\sigma_8 = 0.84^{+0.11}_{-0.18}$ para la muestra GAMA y $\Omega_m = 0.46^{+0.11}_{-0.18}$ y $\sigma_8 = 0.99^{+0.12}_{-0.11}$ para la muestra SDSS, ambos al 68%. En cuanto a la DOH y al parámetro h, se han obtenido valores similares a los encontrados en 4.1. Un análisis adicional conjunto de las dos muestras lente, que ha servido como punto de inicio para un estudio tomográfico, ha permitido mejorar las restricciones en σ_8, con un valor medio de $\sigma_8 = 0.75^{+0.07}_{-0.10}$ y ha clarificado que estudios posteriores podrían reducir las incertidumbres en el plano $\Omega_m - \sigma_8$.

Con los datos de correlación cruzada corregidos y dados los resultados anteriores, se llevó a cabo un análisis tomográfico de la correlación cruzada H-ATLAS/GAMA en 5.1, donde se incluyó tanto la cosmología como la DOH en el algoritmo MCMC, aunque solo se permitió a la última variar entre *bins* de *redshift*. En líneas generales, y para los tres modelos cosmológicos estudiados (ΛCDM, w_0CDM y $w_0 w_a$CDM planos y semilibres), se ha encontrado una clara evolución de la masa mínima promedio con el *redshift*, de acuerdo con los resultados de González-Nuevo et al. (2017), aunque las grandes barras de error de los *bins* 1 y 2 han dificultado extender las conclusiones a la masa promedio para acoger una sola galaxia satélite. Sin embargo, se ha demostrado que las distribuciones *a posteriori* de la DOH son robustas respecto al modelo cosmológico considerado, con solo pequeñas variaciones aisladas. En cuanto a los parámetros cosmológicos, se ha visto una mejora excepcional en relación al análisis no tomográfico. En efecto, se han encontrado valores promedio de $\Omega_m = 0.33^{+0.08}_{-0.16}$, $0.38^{+0.09}_{-0.13}$ y $0.34^{+0.07}_{-0.23}$ al 68% de credibilidad para los modelos ΛCDM, w_0CDM y $w_0 w_a$CDM, respectivamente, a comparar con la cota inferior de 4.1. Aunque solamente se han podido asignar cotas superiores a la constante de Hubble con poca significancia estadística, el parámetro σ_8 también se ha beneficiado del marco tomográfico y ha mostrado una consistencia excepcional entre modelos cosmológicos, con valores promedio de $\sigma_8 = 0.87^{+0.13}_{-0.12}$, $0.87^{+0.11}_{-0.14}$ y $0.88^{+0.13}_{-0.16}$, respectivamente. En cuanto a la ecuación de estado de la energía oscura, se ha demostrado que nuestros resultados son compatibles con una constante

cosmológica, con valores medios de $w_0 = -1.00^{+0.56}_{-0.56}$ y $w_0 = -1.09^{+0.43}_{-0.63}$ para los modelos w_0CDM y w_0w_aCDM y un valor medio de $w_a = -0.19^{+1.67}_{-1.69}$. Es importante notar que las incertidumbres en estos dos parámetros son comparables a las obtenidas mediante resultados no combinados de otros experimentos. Finalmente, y de nuevo, no se ha encontrado la degeneración $\Omega_m - \sigma_8$ típica de los resultados del *shear* cósmico, lo que demuestra el potencial de nuestro observable.

En paralelo a este trabajo, la extensión tomográfica natural de la determinación observacional de la FMH se llevó a cabo en 5.2. Debido a las conclusiones de 4.2, solamente se utilizó el modelo ST en este análisis, aunque se estudiaron dos casos diferentes: el *fit* ST de tres parámetros y el de dos, obtenido exigiendo que toda la masa se encuentre en halos de materia oscura, una suposición más acorde al modelo de halos. Se ha observado una mejora excepcional con respecto al estudio no tomográfico en ambos casos, incluso con la libertad adicional proporcionada por distribuciones *a priori* uniformes en la DOH. En cuanto a la DOH, hemos encontrado una evolución clara de la masa mínima promedio del halo con el *redshift*, confirmando así las conclusiones de 5.1. Sin embargo, los resultados más importantes han sido la no compatibilidad a 2σ del modelo ST de tres parámetros con valores tradicionales basados en simulaciones, prediciendo una densidad numérica mayor para halos de masa baja y (aunque solo a 1σ) una menor para halos masivos. Esta discrepancia se volvió más acuciada para el *fit* de dos parámetros, aumentando hasta más de 3σ, concluyendo con la predicción de una densidad numérica de halos mayor para masas por debajo de $\sim 10^{13} M_\odot/h$ y menor para masas por encima de $\sim 10^{14} M_\odot/h$.

Bibliography

T. M. C. Abbott, F. B. Abdalla, A. Alarcon, et al. Dark Energy Survey year 1 results: Cosmological constraints from galaxy clustering and weak lensing. *Physical Review D*, 98(4):043526, August 2018. doi: 10.1103/PhysRevD.98.043526.

T. M. C. Abbott, M. Aguena, A. Alarcon, et al. Dark Energy Survey Year 1 Results: Cosmological constraints from cluster abundances and weak lensing. *Physical Review D*, 102(2):023509, July 2020. doi: 10.1103/PhysRevD.102.023509.

T. M. C. Abbott, M. Aguena, A. Alarcon, et al. Dark energy survey year 3 results: Cosmological constraints from galaxy clustering and weak lensing. *Physical Review D*, 105:023520, Jan 2022. doi: 10.1103/PhysRevD.105.023520.

A. Acuto, I. G. McCarthy, J. Kwan, et al. The BAHAMAS project: evaluating the accuracy of the halo model in predicting the non-linear matter power spectrum. *Monthly Notices of the Royal Astronomical Society*, 508(3):3519–3534, December 2021. doi: 10.1093/mnras/stab2834.

K. L. Adelberger, C. C. Steidel, M. Pettini, et al. The Spatial Clustering of Star-forming Galaxies at Redshifts $1.4 < z < 3.5$. *The Astrophysical Journal*, 619(2):697–713, February 2005. doi: 10.1086/426580.

R. Ahumada, C. A. Prieto, A. Almeida, et al. The 16th Data Release of the Sloan Digital Sky Surveys: First Release from the APOGEE-2 Southern Survey and Full Release of eBOSS Spectra. *The Astrophysical Journal Supplement Series*, 249(1):3, July 2020. doi: 10.3847/1538-4365/ab929e.

H. Aihara, N. Arimoto, R. Armstrong, et al. The Hyper Suprime-Cam SSP Survey: Overview and survey design. *Publications of the Astronomical Society of Japan*, 70:S4, January 2018. doi: 10.1093/pasj/psx066.

S. Alam, M. Ata, S. Bailey, et al. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: cosmological analysis of the DR12 galaxy sample. *Monthly Notices of the Royal Astronomical Society*, 470(3):2617–2652, September 2017. doi: 10.1093/mnras/stx721.

S. W. Allen, D. A. Rapetti, R. W. Schmidt, et al. Improved constraints on dark energy from Chandra X-ray observations of the largest relaxed galaxy clusters. *Monthly Notices of the Royal Astronomical Society*, 383(3):879–896, January 2008. doi: 10.1111/j.1365-2966.2007.12610.x.

O. Almaini, J. S. Dunlop, C. J. Willott, et al. Correlations between bright submillimetre sources and low-redshift galaxies. *Monthly Notices of the Royal Astronomical Society*, 358(3):875–882, April 2005. doi: 10.1111/j.1365-2966.2005.08790.x.

A. Amblard, A. Cooray, P. Serra, et al. Herschel-ATLAS: Dust temperature and redshift distribution of SPIRE and PACS detected sources using submillimetre colours. *Astronomy and Astrophysics*, 518:L9, July 2010. doi: 10.1051/0004-6361/201014586.

A. Amon, D. Gruen, M. A. Troxel, et al. Dark Energy Survey Year 3 results: Cosmology from cosmic shear and robustness to data calibration. *Physical Review D*, 105(2):023514, January 2022. doi: 10.1103/PhysRevD.105.023514.

A. Amvrosiadis, E. Valiante, J. Gonzalez-Nuevo, et al. Herschel-ATLAS : the spatial clustering of low- and high-redshift submillimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 483 (4):4649–4664, March 2019. doi: 10.1093/mnras/sty3013.

R. E. Angulo, O. Hahn, and T. Abel. How closely do baryons follow dark matter on large scales? *Monthly Notices of the Royal Astronomical Society*, 434(2):1756–1764, 07 2013. ISSN 0035-8711. doi: 10.1093/mnras/stt1135.

Halton Arp. *Quasars, redshifts, and controversies*. 1987.

M. W. Auger, T. Treu, A. S. Bolton, et al. The Sloan Lens ACS Survey. IX. Colors, Lensing, and Stellar Masses of Early-Type Galaxies. *The Astrophysical Journal*, 705(2):1099–1115, November 2009. doi: 10.1088/0004-637X/705/2/1099.

D. J. Bacon, A. R. Refregier, and R. S. Ellis. Detection of weak gravitational lensing by large-scale structure. *Monthly Notices of the Royal Astronomical Society*, 318(2):625–640, October 2000. doi: 10.1046/j.1365-8711.2000.03851.x.

T. J. L. C. Bakx, S. Eales, and A. Amvrosiadis. A search for the lenses in the Herschel Bright Sources (HerBS) sample. *Monthly Notices of the Royal Astronomical Society*, 493(3):4276–4293, April 2020. doi: 10.1093/mnras/staa506.

I. K. Baldry, A. S. G. Robotham, D. T. Hill, et al. Galaxy And Mass Assembly (GAMA): the input catalogue and star-galaxy separation. *Monthly Notices of the Royal Astronomical Society*, 404(1): 86–100, May 2010. doi: 10.1111/j.1365-2966.2010.16282.x.

I. K. Baldry, M. Alpaslan, A. E. Bauer, et al. Galaxy And Mass Assembly (GAMA): AUTOZ spectral redshift measurements, confidence and errors. *Monthly Notices of the Royal Astronomical Society*, 441(3):2440–2451, July 2014. doi: 10.1093/mnras/stu727.

J. M. Bardeen, J. R. Bond, N. Kaiser, et al. The Statistics of Peaks of Gaussian Random Fields. *The Astrophysical Journal*, 304:15, May 1986. doi: 10.1086/164143.

A. J. Barger, L. L. Cowie, D. B. Sanders, et al. Submillimetre-wavelength detection of dusty star-forming galaxies at high redshift. *Nature*, 394(6690):248–251, July 1998. doi: 10.1038/28338.

M. Bartelmann and P. Schneider. Weak gravitational lensing. *Physics Reports*, 340(4-5):291–472, January 2001. doi: 10.1016/S0370-1573(00)00082-X.

E. Baxter, J. Clampitt, T. Giannantonio, et al. Joint measurement of lensing-galaxy correlations using SPT and DES SV data. *Monthly Notices of the Royal Astronomical Society*, 461(4):4099–4114, October 2016a. doi: 10.1093/mnras/stw1584.

E. Baxter, J. Clampitt, T. Giannantonio, et al. Joint measurement of lensing-galaxy correlations using SPT and DES SV data. *Monthly Notices of the Royal Astronomical Society*, 461(4):4099–4114, October 2016b. doi: 10.1093/mnras/stw1584.

G. J. Bendo, R. D. Joseph, M. Wells, et al. Dust Temperatures in the Infrared Space Observatory Atlas of Bright Spiral Galaxies. *The Astronomical Journal*, 125(5):2361–2372, May 2003. doi: 10.1086/374361.

N. Benitez, R. Dupke, M. Moles, et al. J-PAS: The Javalambre-Physics of the Accelerated Universe Astrophysical Survey. *arXiv e-prints*, art. arXiv:1403.5237, March 2014.

A. A. Berlind, D. H. Weinberg, A. J. Benson, et al. The Halo Occupation Distribution and the Physics of Galaxy Formation. *The Astrophysical Journal*, 593(1):1–25, August 2003. doi: 10.1086/376517.

M. Béthermin, H. Dole, M. Cousin, et al. Submillimeter number counts at 250 μm, 350 μm and 500 μm in BLAST data. *Astronomy and Astrophycis*, 516:A43, June 2010. doi: 10.1051/0004-6361/200913910.

F. Bianchini, A. Lapi, M. Calabrese, et al. Toward a Tomographic Analysis of the Cross-Correlation between Planck CMB Lensing and H-ATLAS Galaxies. *The Astrophysical Journal*, 825(1):24, July 2016. doi: 10.3847/0004-637X/825/1/24.

S. Bilenky. Neutrino oscillations: From a historical perspective to the present status. *Nuclear Physics B*, 908:2–13, 2016. ISSN 0550-3213. doi: https://doi.org/10.1016/j.nuclphysb.2016.01.025. Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015.

A. W. Blain and M. S. Longair. Submillimetre cosmology. *Monthly Notices of the Royal Astronomical Society*, 264(2):509–521, 09 1993. ISSN 0035-8711. doi: 10.1093/mnras/264.2.509.

A. W. Blain, I. Smail, R. J. Ivison, et al. The history of star formation in dusty galaxies. *Monthly Notices of the Royal Astronomical Society*, 302(4):632–648, February 1999. doi: 10.1046/j.1365-8711.1999.02178.x.

Andrew Blain, Ian Smail, R Ivison, et al. Submillimeter galaxies. *physrep*, 369:111–176, 10 2002.

C. Blake and J. Wall. Quantifying angular clustering in wide-area radio surveys. *Monthly Notices of the Royal Astronomical Society*, 337(3):993–1003, December 2002. doi: 10.1046/j.1365-8711.2002.05979.x.

C. Blake, A. Pope, D. Scott, et al. On the cross-correlation of sub-mm sources and optically selected galaxies. *Monthly Notices of the Royal Astronomical Society*, 368(2):732–740, May 2006. doi: 10.1111/j.1365-2966.2006.10158.x.

M. R. Blanton, M. A. Bershady, B. Abolfathi, et al. Sloan Digital Sky Survey IV: Mapping the Milky Way, Nearby Galaxies, and the Distant Universe. *The Astronomical Journal*, 154(1):28, July 2017. doi: 10.3847/1538-3881/aa7567.

S. Bocquet, A. Saro, K. Dolag, et al. Halo mass function: baryon impact, fitting formulae, and implications for cluster cosmology. *Monthly Notices of the Royal Astronomical Society*, 456(3):2361–2373, 12 2015. ISSN 0035-8711. doi: 10.1093/mnras/stv2657. URL `https://doi.org/10.1093/mnras/stv2657`.

N. W. Boggess, J. C. Mather, R. Weiss, et al. The COBE Mission: Its Design and Performance Two Years after Launch. *The Astrophysical Journal*, 397:420, October 1992. doi: 10.1086/171797.

L. Bonavera, J. González-Nuevo, S. L. Suárez Gómez, et al. QSOs sigposting cluster size halos as gravitational lenses: halo mass, projected mass density profile and concentration at z~0.7. *Journal of Cosmology and Astroparticle Physics*, 2019(9):021, September 2019. doi: 10.1088/1475-7516/2019/09/021.

L. Bonavera, J. González-Nuevo, M. M. Cueli, et al. Cosmology with the submillimetre galaxies magnification bias: Proof of concept. *Astronomy and Astrophysics*, 639:A128, July 2020. doi: 10.1051/0004-6361/202038050.

L. Bonavera, M. M. Cueli, J. González-Nuevo, et al. Cosmology with the submillimetre galaxies magnification bias. Tomographic analysis. *Astronomy and Astrophysics*, 656:A99, December 2021. doi: 10.1051/0004-6361/202141521.

J. R. Bond, S. Cole, G. Efstathiou, et al. Excursion Set Mass Functions for Hierarchical Gaussian Fluctuations. *The Astrophysical Journal*, 379:440, October 1991. doi: 10.1086/170520.

C. Borys, S. Chapman, M. Halpern, et al. The Hubble Deep Field North SCUBA Super-map—I. Submillimetre maps, sources and number counts. *Monthly Notices of the Royal Astronomical Society*, 344(2):385–398, 09 2003. ISSN 0035-8711. doi: 10.1046/j.1365-8711.2003.06818.x.

N. Bourne, L. Dunne, S. J. Maddox, et al. The Herschel-ATLAS Data Release 1 - II. Multi-wavelength counterparts to submillimetre sources. *Monthly Notices of the Royal Astronomical Society*, 462(2): 1714–1734, October 2016. doi: 10.1093/mnras/stw1654.

T. G. Brainerd, R. D. Blandford, and I. Smail. Weak Gravitational Lensing by Galaxies. *The Astrophysical Journal*, 466:623, August 1996. doi: 10.1086/177537.

J. N. Bregman, B. A. Snider, L. Grego, et al. Far-Infrared Emission From E and E/S0 Galaxies. *The Astrophysical Journal*, 499(2):670–676, May 1998. doi: 10.1086/305653.

T. Budavári, A. J. Connolly, A. S. Szalay, et al. Angular Clustering with Photometric Redshifts in the Sloan Digital Sky Survey: Bimodality in the Clustering Properties of Galaxies. *The Astrophysical Journal*, 595(1):59–70, September 2003. doi: 10.1086/377168.

J. S. Bullock, T. S. Kolatt, Y. Sigad, et al. Profiles of dark haloes: evolution, scatter and environment. *Monthly Notices of the Royal Astronomical Society*, 321(3):559–575, March 2001. doi: 10.1046/j.1365-8711.2001.04068.x.

R. S. Bussmann, M. A. Gurwell, H. Fu, et al. A Detailed Gravitational Lens Model Based on Submillimeter Array and Keck Adaptive Optics Imaging of a Herschel-ATLAS Submillimeter Galaxy at z = 4.243. *The Astrophysical Journal*, 756(2):134, September 2012. doi: 10.1088/0004-637X/756/2/134.

R. S. Bussmann, I. Pérez-Fournon, S. Amber, et al. Gravitational Lens Models Based on Submillimeter Array Imaging of Herschel-selected Strongly Lensed Sub-millimeter Galaxies at z > 1.5. *The Astrophysical Journal*, 779(1):25, December 2013. doi: 10.1088/0004-637X/779/1/25.

M. Cacciato, F. C. van den Bosch, S. More, et al. Galaxy clustering and galaxy–galaxy lensing: a promising union to constrain cosmological parameters. *Monthly Notices of the Royal Astronomical Society*, 394(2):929–946, 03 2009. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2008.14362.x.

M. Cacciato, F. C. van den Bosch, S. More, et al. Cosmological constraints from a combination of galaxy clustering and lensing – III. Application to SDSS data. *Monthly Notices of the Royal Astronomical Society*, 430(2):767–786, 02 2013. ISSN 0035-8711. doi: 10.1093/mnras/sts525.

J. A. Calanog, H. Fu, A. Cooray, et al. Lens Models of Herschel-selected Galaxies from High-resolution Near-IR Observations. *The Astrophysical Journal*, 797(2):138, December 2014. doi: 10.1088/0004-637X/797/2/138.

S. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*. Benjamin Cummings, 2003. ISBN 0805387323.

C. Casey, D. Narayanan, and A. Cooray. Dusty star-forming galaxies at high redshift. *Physics Reports*, 541, 02 2014. doi: 10.1016/j.physrep.2014.02.009.

E. Castorina, E. Sefusatti, R. K. Sheth, et al. Cosmology with massive neutrinos II: on the universality of the halo mass function and bias. *Journal of Cosmology and Astroparticle Physics*, 2014(2):049, February 2014. doi: 10.1088/1475-7516/2014/02/049.

T. Castro, S. Borgani, K. Dolag, et al. On the impact of baryons on the halo mass function, bias, and cluster cosmology. *Monthly Notices of the Royal Astronomical Society*, 500(2):2316–2335, 11 2020. ISSN 0035-8711. doi: 10.1093/mnras/staa3473.

J. L. Cervantes-Cota, S. Galindo-Uribarri, and G. F. Smoot. The Legacy of Einstein's Eclipse, Gravitational Lensing. *Universe*, 6(1):9, December 2019. doi: 10.3390/universe6010009.

E. L. Chapin, S. C. Chapman, K. E. Coppin, et al. A joint analysis of BLAST 250-500 μm and LABOCA 870 μm observations in the Extended Chandra Deep Field-South. *Monthly Notices of the Royal Astronomical Society*, 411(1):505–549, February 2011. doi: 10.1111/j.1365-2966.2010.17697.x.

S. C. Chapman, I. Smail, A. W. Blain, et al. A Population of Hot, Dusty Ultraluminous Galaxies at z ~2. *The Astrophysical Journal*, 614(2):671–678, October 2004. doi: 10.1086/423833.

S. C. Chapman, A. W. Blain, I. Smail, et al. A Redshift Survey of the Submillimeter Galaxy Population. *The Astrophysical Journal*, 622(2):772–796, April 2005. doi: 10.1086/428082.

R. Chary and D. Elbaz. Interpreting the Cosmic Infrared Background: Constraints on the Evolution of the Dust-enshrouded Star Formation Rate. *The Astrophysical Journal*, 556(2):562–581, August 2001. doi: 10.1086/321609.

M. Chevallier and D. Polarski. Accelerating Universes with Scaling Dark Matter. *International Journal of Modern Physics D*, 10(2):213–223, January 2001. doi: 10.1142/S0218271801000822.

A. Cimatti, P. Andreani, H. Röttgering, et al. Vigorous star formation hidden by dust in a galaxy at a redshift of 1.4. *Nature*, 392(6679):895–897, April 1998. doi: 10.1038/31872.

D. L. Clements, E. Rigby, S. Maddox, et al. Herschel-ATLAS: Extragalactic number counts from 250 to 500 microns. *Astronomy and Astrophysics*, 518:L8, July 2010. doi: 10.1051/0004-6361/201014581.

S. Cole and N. Kaiser. Biased clustering in the cold dark matter cosmogony. *Monthly Notices of the Royal Astronomical Society*, 237:1127–1146, April 1989. doi: 10.1093/mnras/237.4.1127.

P. P. Coles and P. F. Lucchin. *Cosmology: The Origin and Evolution of Cosmic Structure*. Wiley, 2003. ISBN 9780470852996.

M. Colless. First results from the 2dF Galaxy Redshift Survey. *Philosophical Transactions of the Royal Society of London Series A*, 357(1750):105, January 1999. doi: 10.1098/rsta.1999.0317.

A. Cooray and R. K. Sheth. Halo models of large scale structure. *Physics Reports*, 372(1):1–129, December 2002. doi: 10.1016/S0370-1573(02)00276-4.

A. Cooray and R. K. Sheth. Halo models of large scale structure. *Physics Reports*, 372(1):1–129, 2002. ISSN 0370-1573. doi: https://doi.org/10.1016/S0370-1573(02)00276-4.

K. Coppin, E. L. Chapin, A. M. J. Mortier, et al. The SCUBA Half-Degree Extragalactic Survey - II. Submillimetre maps, catalogue and number counts. *Monthly Notices of the Royal Astronomical Society*, 372(4):1621–1652, November 2006. doi: 10.1111/j.1365-2966.2006.10961.x.

K. Coppin, M. Halpern, D. Scott, et al. The SCUBA HAlf Degree Extragalactic Survey – VI. 350-μm mapping of submillimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 384(4): 1597–1610, 02 2008. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2007.12808.x.

M. Costanzi, F. Villaescusa-Navarro, M. Viel, et al. Cosmology with massive neutrinos III: the halo mass function and an application to galaxy clusters. *Journal of Cosmology and Astroparticle Physics*, 2013(12):012, December 2013. doi: 10.1088/1475-7516/2013/12/012.

J. Courtin, Y. Rasera, J. M. Alimi, et al. Imprints of dark energy on cosmic structure formation - II. Non-universality of the halo mass function. *Monthly Notices of the Royal Astronomical Society*, 410 (3):1911–1931, 2011. doi: 10.1111/j.1365-2966.2010.17573.x.

L. L. Cowie, A. J. Barger, E. B. Fomalont, et al. The Evolution of the Ultraluminous Infrared Galaxy Population from Redshift 0 to 1.5. *The Astrophysical Journal*, 603(2):L69–L72, March 2004. doi: 10.1086/383198.

P. Cox, M. Krips, R. Neri, et al. Gas and Dust in a Submillimeter Galaxy at z = 4.24 from the Herschel Atlas. *The Astrophysical Journal*, 740(2):63, October 2011. doi: 10.1088/0004-637X/740/2/63.

M. Crocce, P. Fosalba, F. J. Castander, et al. Simulating the Universe with MICE: the abundance of massive clusters. *Monthly Notices of the Royal Astronomical Society*, 403(3):1353–1367, 2010. doi: 10.1111/j.1365-2966.2009.16194.x6.

M. M. Cueli, L. Bonavera, J. González-Nuevo, et al. A direct and robust method to observationally constrain the halo mass function via the submillimeter magnification bias: Proof of concept. *Astronomy and Astrophysics*, 645:A126, January 2021. doi: 10.1051/0004-6361/202039326.

M. M. Cueli, L. Bonavera, J. González-Nuevo, et al. Tomography-based observational measurements of the halo mass function via the submillimeter magnification bias. *Astronomy and Astrophysics*, 662:A44, June 2022. doi: 10.1051/0004-6361/202142949.

W. Cui, M. Baldi, and S. Borgani. The halo mass function in interacting dark energy models. *Monthly Notices of the Royal Astronomical Society*, 424(2):993–1005, August 2012. doi: 10.1111/j.1365-2966.2012.21267.x.

W. Cui, S. Borgani, and G. Murante. The effect of active galactic nuclei feedback on the halo mass function. *Monthly Notices of the Royal Astronomical Society*, 441(2):1769–1782, 05 2014. ISSN 0035-8711. doi: 10.1093/mnras/stu673.

Dark Energy Survey Collaboration, T. Abbott, F. B. Abdalla, et al. The Dark Energy Survey: more than dark energy - an overview. *Monthly Notices of the Royal Astronomical Society*, 460(2):1270–1299, August 2016. doi: 10.1093/mnras/stw641.

M. Davis, G. Efstathiou, C. S. Frenk, et al. The evolution of large-scale structure in a universe dominated by cold dark matter. *The Astrophysical Journal*, 292:371–394, 1985. doi: 10.1086/1631689.

T. de Graauw, F. P. Helmich, T. G. Phillips, et al. The Herschel-Heterodyne Instrument for the Far-Infrared (HIFI). *Astronomy and Astrophysics*, 518:L6, July 2010. doi: 10.1051/0004-6361/201014698.

J. T. A. de Jong, G. A. Verdoes Kleijn, K. H. Kuijken, et al. The Kilo-Degree Survey. *Experimental Astronomy*, 35(1-2):25–44, January 2013. doi: 10.1007/s10686-012-9306-1.

T. de Jong, P. E. Clegg, B. T. Soifer, et al. IRAS observations of Shapley-Ames galaxies. *Astrophysical Journal*, 278:L67–L70, March 1984. doi: 10.1086/184225.

V. Desjacques, D. Jeong, and F. Schmidt. Large-scale galaxy bias. *Physics Reports*, 733:1–193, February 2018. doi: 10.1016/j.physrep.2017.12.002.

G. Despali, C. Giocoli, R. E. Angulo, et al. The universality of the virial halo mass function and models for non-universality of other halo definitions. *Monthly Notices of the Royal Astronomical Society*, 456(3):2486–2504, 2016. doi: 10.1093/mnras/stv2842.

N. A. Devereux and J. S. Young. The Gas/Dust Ratio in Spiral Galaxies. *The Astrophysical Journal*, 359:42, August 1990. doi: 10.1086/169031.

M. J. Devlin, P. A. R. Ade, I. Aretxaga, et al. Over half of the far-infrared background light comes from galaxies at z>=1.2. *Nature*, 458(7239):737–739, April 2009. doi: 10.1038/nature07918.

M. Dickinson, D. Stern, M. Giavalisco, et al. Color-selected Galaxies at z ~6 in the Great Observatories Origins Deep Survey. *The Astrophysical Journal*, 600(2):L99–L102, January 2004. doi: 10.1086/381119.

S. Dodelson. *Modern Cosmology*. 2003.

S. Dodelson. *Gravitational Lensing*. Cambridge University Press, 2017. doi: 10.1017/9781316424254.

H. Dole, G. Lagache, J. L. Puget, et al. The cosmic infrared background resolved by Spitzer. Contributions of mid-infrared galaxies to the far-infrared background. *Astronomy and Astrophysics*, 451 (2):417–429, May 2006. doi: 10.1051/0004-6361:20054446.

S. P. Driver, P. Norberg, I. K. Baldry, et al. GAMA: towards a physical understanding of galaxy formation. *Astronomy and Geophysics*, 50(5):5.12–5.19, October 2009. doi: 10.1111/j.1468-4004.2009.50512.x.

S. P. Driver, D. T. Hill, L. S. Kelvin, et al. Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release. *Monthly Notices of the Royal Astronomical Society*, 413(2):971–995, May 2011. doi: 10.1111/j.1365-2966.2010.18188.x.

J. S. Dunlop, P. A. R. Ade, J. J. Bock, et al. The BLAST 250 μm-selected galaxy population in GOODS-South. *Monthly Notices of the Royal Astronomical Society*, 408(4):2022–2050, November 2010. doi: 10.1111/j.1365-2966.2010.17278.x.

E. Dwek, R. G. Arendt, M. G. Hauser, et al. The COBE Diffuse Infrared Background Experiment Search for the Cosmic Infrared Background. IV. Cosmological Implications. *The Astrophysical Journal*, 508(1):106–122, November 1998. doi: 10.1086/306382.

S. Eales, L. Dunne, D. Clements, et al. The Herschel ATLAS. *Publications of the Astronomical Society of the Pacific,*, 122(891):499, May 2010. doi: 10.1086/653086.

A. S. Eddington. *Space, time and gravitation. an outline of the general relativity theory.* 1920.

A. Edge, W. Sutherland, K. Kuijken, et al. The VISTA Kilo-degree Infrared Galaxy (VIKING) Survey: Bridging the Gap between Low and High Redshift. *The Messenger*, 154:32–34, December 2013.

A. Einstein. Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin*, pages 844–847, January 1915.

D. J. Eisenstein and W. Hu. Baryonic Features in the Matter Transfer Function. *The Astrophysical Journal*, 496(2):605–614, March 1998. doi: 10.1086/305424.

D. J. Eisenstein and W. Hu. Power Spectra for Cold Dark Matter and Its Variants. *The Astrophysical Journal*, 511(1):5–15, January 1999. doi: 10.1086/306640.

R. Ellis. Gravitational lensing: A unique probe of dark matter and dark energy. *Philosophical transactions. Series A, Mathematical, physical, and engineering sciences*, 368:967–87, 03 2010. doi: 10.1098/rsta.2009.0209.

L. Fernandez, M. M. Cueli, J. González-Nuevo, et al. Galaxy cluster mass density profile derived using the submillimetre galaxies magnification bias. *Astronomy and Astrophysics*, 658:A19, February 2022. doi: 10.1051/0004-6361/202141905.

D. J. Fixsen, E. Dwek, J. C. Mather, et al. The Spectrum of the Extragalactic Far-Infrared Background from the COBE FIRAS Observations. *The Astrophysical Journal*, 508(1):123–128, November 1998. doi: 10.1086/306383.

D. Foreman-Mackey, D. W. Hogg, D. Lang, et al. emcee: The MCMC Hammer. *Publications of the Astronomical Society of the Pacific*, 125(925):306, 2013. doi: 10.1086/670067.

J. Freundlich, F. Jiang, A. Dekel, et al. The Dekel-Zhao profile: a mass-dependent dark-matter density profile with flexible inner slope and analytic potential, velocity dispersion, and lensing properties. *Monthly Notices of the Royal Astronomical Society*, 499(2):2912–2933, December 2020. doi: 10.1093/mnras/staa2790.

H. Fu, E. Jullo, A. Cooray, et al. A Comprehensive View of a Strongly Lensed Planck-Associated Submillimeter Galaxy. *The Astrophysical Journal*, 753(2):134, July 2012. doi: 10.1088/0004-637X/753/2/134.

W. Fugmann. Statistical gravitational lensing and the Lick catalogue of galaxies. *Astronomy and Astrophysics*, 240(1):11–21, December 1990.

J. P. Gardner, J. C. Mather, M. Clampin, et al. The James Webb Space Telescope. *Space Science Reviews*, 123(4):485–606, April 2006. doi: 10.1007/s11214-006-8315-7.

J. W. Gibbs. *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundations of Thermodynamics*. Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics. C. Scribner's sons, 1902.

C. Giocoli, F. Marulli, M. Baldi, et al. Characterizing dark interactions with the halo mass accretion history and structural properties. *Monthly Notices of the Royal Astronomical Society*, 434(4):2982–2998, October 2013. doi: 10.1093/mnras/stt1218.

J. Glenn, A. Conley, M. Béthermin, et al. HerMES: deep galaxy number counts from a P(D) fluctuation analysis of SPIRE Science Demonstration Phase observations. *Monthly Notices of the Royal Astronomical Society*, 409(1):109–121, November 2010. doi: 10.1111/j.1365-2966.2010.17781.x.

J. González-Nuevo, A. Lapi, S. Fleuren, et al. Herschel-ATLAS: Toward a Sample of ~1000 Strongly Lensed Galaxies. *The Astrophysical Journal*, 749(1):65, April 2012. doi: 10.1088/0004-637X/749/1/65.

J. González-Nuevo, A. Lapi, M. Negrello, et al. Herschel-ATLAS/GAMA: SDSS cross-correlation induced by weak lensing. *Monthly Notices of the Royal Astronomical Society*, 442(3):2680–2690, August 2014. doi: 10.1093/mnras/stu1041.

J. González-Nuevo, A. Lapi, L. Bonavera, et al. H-ATLAS/GAMA: magnification bias tomography. Astrophysical constraints above ~1 arcmin. *Journal of Cosmology and Astroparticle Physics*, 2017 (10):024, October 2017. doi: 10.1088/1475-7516/2017/10/024.

J. González-Nuevo, S. L. Suárez Gómez, L. Bonavera, et al. SHALOS: Statistical Herschel-ATLAS lensed objects selection. *Astronomy and Astrophysics*, 627:A31, July 2019. doi: 10.1051/0004-6361/201935475.

J. González-Nuevo, M. M. Cueli, L. Bonavera, et al. Cosmological constraints on the magnification bias on sub-millimetre galaxies after large-scale bias corrections. *Astronomy and Astrophysics*, 646:A152, February 2021. doi: 10.1051/0004-6361/202039043.

J. Goodman and J. Weare. Ensemble samplers with affine invariances. *Communications in Applied Mathematics and Computational Science*, 5(1):65–80, 2010. doi: 10.2140/camcos.2010.5.65.

G. L. Granato, G. De Zotti, L. Silva, et al. A Physical Model for the Coevolution of QSOs and Their Spheroidal Hosts. *The Astrophysical Journal*, 600(2):580–594, January 2004. doi: 10.1086/379875.

M. J. Griffin, A. Abergel, A. Abreu, et al. The Herschel-SPIRE instrument and its in-flight performance. *Astronomy and Astrophysics*, 518:L3, July 2010. doi: 10.1051/0004-6361/201014519.

J. E. Gunn and J. R. Gott. On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *The Astrophysical Journal*, 176:1, August 1972. doi: 10.1086/151605.

T. Hamana, M. Shirasaki, S. Miyazaki, et al. Cosmological constraints from cosmic shear two-point correlation functions with HSC survey first-year data. *Publications of the Astronomical Society of Japan*, 72(1):16, February 2020. doi: 10.1093/pasj/psz138.

E. Harrison. *Cosmology: The Science of the Universe*. Cambridge University Press, 2 edition, 2000. doi: 10.1017/CBO9780511804540.

D. Herranz. Foreground-Background Galaxy Correlations in the Hubble Deep Fields. In J. Tran Thanh Van, Yannick Mellier, and Marc Moniez, editors, *Cosmological Physics with Gravitational Lensing*, page 197, January 2001.

C. Hikage, M. Oguri, T. Hamana, et al. Cosmology from cosmic shear power spectra with Subaru Hyper Suprime-Cam first-year data. *Publications of the Astronomical Society of Japan*, 71(2):43, April 2019. doi: 10.1093/pasj/psz010.

H. Hildebrandt, M. Viola, C. Heymans, et al. KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing. *Monthly Notices of the Royal Astronomical Society*, 465(2): 1454–1498, February 2017. doi: 10.1093/mnras/stw2805.

H. Hildebrandt, F. Köhlinger, J. L. van den Busch, et al. KiDS+VIKING-450: Cosmic shear tomography with optical and infrared data. *Astronomy and Astrophysics*, 633:A69, January 2020. doi: 10.1051/0004-6361/201834878.

J. A. Hodge and E. da Cunha. High-redshift star formation in the Atacama large millimetre/submillimetre array era. *Royal Society Open Science*, 7(12):200556, December 2020. doi: 10.1098/rsos.200556.

D. W. Hogg and D. Foreman-Mackey. Data Analysis Recipes: Using Markov Chain Monte Carlo. *The Astrophysical Journal Supplement Series*, 236(1):11, May 2018. doi: 10.3847/1538-4365/aab76e.

W. S. Holland, E. I. Robson, W. K. Gear, et al. SCUBA: a common-user submillimetre camera operating on the James Clerk Maxwell Telescope. *Monthly Notices of the Royal Astronomical Society*, 303(4):659–672, 03 1999. ISSN 0035-8711. doi: 10.1046/j.1365-8711.1999.02111.x.

W. S. Holland, D. Bintley, E. L. Chapin, et al. SCUBA-2: the 10 000 pixel bolometer camera on the James Clerk Maxwell Telescope. *Monthly Notices of the Royal Astronomical Society*, 430(4): 2513–2533, April 2013. doi: 10.1093/mnras/sts612.

J. R. Houck, D. P. Schneider, G. E. Danielson, et al. Unidentified IRAS sources : ultra high-luminosity galaxies. *The Astrophysical Journal l*, 290:L5–L8, March 1985. doi: 10.1086/184431.

D. H. Hughes, S. Serjeant, J. Dunlop, et al. High-redshift star formation in the Hubble Deep Field revealed by a submillimetre-wavelength survey. *Nature*, 394(6690):241–247, July 1998. doi: 10.1038/28328.

E. Ibar, R. J. Ivison, A. Cava, et al. H-ATLAS: PACS imaging for the Science Demonstration Phase. *Monthly Notices of the Royal Astronomical Society*, 409(1):38–47, November 2010. doi: 10.1111/j.1365-2966.2010.17620.x.

L. Infante. Correlation analysis of faint galaxies. *Astronomy and Astrophysics*, 282(2):353–362, February 1994.

R. J. Ivison, J. S. Dunlop, D. H. Hughes, et al. Dust, Gas, and the Evolutionary Status of the Radio Galaxy 8C 1435+635 at z = 4.25. *The Astrophysical Journal*, 494(1):211–217, February 1998a. doi: 10.1086/305208.

R. J. Ivison, I. Smail, J.-F. Le Borgne, et al. A hyperluminous galaxy at z = 2.8 found in a deep submillimetre survey. *Monthly Notices of the Royal Astronomical Society*, 298(2):583–593, 08 1998b. ISSN 0035-8711. doi: 10.1046/j.1365-8711.1998.01677.x.

R. J. Ivison, I. Smail, A. J. Barger, et al. The diversity of SCUBA-selected galaxies. *Monthly Notices of the Royal Astronomical Society*, 315(2):209–222, June 2000. doi: 10.1046/j.1365-8711.2000.03376.x.

R. J. Ivison, T. R. Greve, I. Smail, et al. Deep radio imaging of the SCUBA 8-mJy survey fields: submillimetre source identifications and redshift distribution. *Monthly Notices of the Royal Astronomical Society*, 337(1):1–25, November 2002. doi: 10.1046/j.1365-8711.2002.05900.x.

R. J. Ivison, T. R. Greve, J. S. Dunlop, et al. The SCUBA HAlf Degree Extragalactic Survey – III. Identification of radio and mid-infrared counterparts to submillimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 380(1):199–228, 08 2007. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2007.12044.x.

R. J. Ivison, A. M. Swinbank, B. Swinyard, et al. Herschel and SCUBA-2 imaging and spectroscopy of a bright, lensed submillimetre galaxy at z = 2.3. *Astronomy and Astrophysics*, 518:L35, July 2010. doi: 10.1051/0004-6361/201014548.

R. J. Ivison, A. J. R. Lewis, A. Weiss, et al. The Space Density of Luminous Dusty Star-forming Galaxies at z > 4: SCUBA-2 and LABOCA Imaging of Ultrared Galaxies from Herschel-ATLAS. *The Astrophysical Journal*, 832(1):78, November 2016. doi: 10.3847/0004-637X/832/1/78.

A. Jenkins, C. S. Frenk, S. D. M. White, et al. The mass function of dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 321(2):372–384, 2001. doi: 10.1046/j.1365-8711.2001.04029.x.

R. D. Joseph and G. S. Wright. Recent star formation in interacting galaxies - II. Super starbursts in merging galaxies. *Monthly Notices of the Royal Astronomical Society*, 214:87–95, May 1985. doi: 10.1093/mnras/214.2.87.

S. Joudaki, C. Blake, C. Heymans, et al. CFHTLenS revisited: assessing concordance with Planck including astrophysical systematics. *Monthly Notices of the Royal Astronomical Society*, 465(2):2033–2052, February 2017. doi: 10.1093/mnras/stw2665.

S. Joudaki, C. Blake, A. Johnson, et al. KiDS-450 + 2dFLenS: Cosmological parameter constraints from weak gravitational lensing tomography and overlapping redshift-space galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, 474(4):4894–4924, March 2018. doi: 10.1093/mnras/stx2820.

N. Kaiser. On the spatial correlations of Abell clusters. *The Astrophysical Journal*, 284:L9–L12, September 1984. doi: 10.1086/184341.

G. Kauffmann, J. M. Colberg, A. Diaferio, et al. Clustering of galaxies in a hierarchical universe - I. Methods and results at z=0. *Monthly Notices of the Royal Astronomical Society*, 303(1):188–206, February 1999. doi: 10.1046/j.1365-8711.1999.02202.x.

Jr. Kennicutt, R. C. The Global Schmidt Law in Star-forming Galaxies. *The Astrophysical Journal*, 498(2):541–552, May 1998. doi: 10.1086/305588.

M. Kilbinger. Cosmology with cosmic shear observations: a review. *Reports on Progress in Physics*, 78(8):086901, July 2015. doi: 10.1088/0034-4885/78/8/086901.

T. Kitayama and Y. Suto. Semianalytic Predictions for Statistical Properties of X-Ray Clusters of Galaxies in Cold Dark Matter Universes. *The Astrophysical Journal*, 469:480, October 1996. doi: 10.1086/177797.

A. Kovacs, S. C. Chapman, C. D. Dowell, et al. SHARC-2 350 μm Observations of Distant Submillimeter-selected Galaxies. *The Astrophysical Journal*, 650(2):592–603, October 2006. doi: 10.1086/506341.

A. V. Kravtsov, A. A. Berlind, R. H. Wechsler, et al. The Dark Side of the Halo Occupation Distribution. *The Astrophysical Journal*, 609(1):35–49, July 2004. doi: 10.1086/420959.

E. Kreysa, H. Gemuend, J. Gromke, et al. Bolometer array development at the Max-Planck-Institut fuer Radioastronomie. In Thomas G. Phillips, editor, *Advanced Technology MMW, Radio, and Terahertz Telescopes*, volume 3357 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, pages 319–325, July 1998. doi: 10.1117/12.317367.

H. Kuehr, A. Witzel, I. I. K. Pauliny-Toth, et al. A Catalogue of Extragalactic Radio Sources Having Flux Densities Greater than 1-JY at 5-GHZ. *Astronomy and Astrophysics Supplement*, 45:367, September 1981.

C. Lacey and S. Cole. Merger rates in hierarchical models of galaxy formation. *Monthly Notices of the Royal Astronomical Society*, 262(3):627–649, June 1993. doi: 10.1093/mnras/262.3.627.

C. Lacey and S. Cole. Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-Body Simulations. *Monthly Notices of the Royal Astronomical Society*, 271:676, December 1994. doi: 10.1093/mnras/271.3.676.

C. Lacey and S. Cole. Merger Rates in Hierarchical Models of Galaxy Formation - Part Two - Comparison with N-Body Simulations. *Monthly Notices of the Royal Astronomical Society*, 271(3):676, 1994. doi: 10.1093/mnras/271.3.676.

S. D. Landy and A. S. Szalay. Bias and Variance of Angular Correlation Functions. *Astrophysical Journal*, 412:64, 1993. doi: 10.1086/172900.

J. Lange, F. C. van den Bosch, A. R. Zentner, et al. Maturing satellite kinematics into a competitive probe of the galaxy–halo connection. *Monthly Notices of the Royal Astronomical Society*, 482(4): 4824–4845, 11 2018. ISSN 0035-8711. doi: 10.1093/mnras/sty2950.

A. Lapi and L. Danese. A Stochastic Theory of the Hierarchical Clustering. I. Halo Mass Function. *The Astrophysical Journal*, 903(2):117, November 2020. doi: 10.3847/1538-4357/abb944.

A. Lapi, F. Shankar, J. Mao, et al. Quasar Luminosity Functions from Joint Evolution of Black Holes and Host Galaxies. *The Astrophysical Journal*, 650(1):42–56, October 2006. doi: 10.1086/507122.

A. Lapi, J. González-Nuevo, L. Fan, et al. Herschel-ATLAS Galaxy Counts and High-redshift Luminosity Functions: The Formation of Massive Early-type Galaxies. *The Astrophysical Journal*, 742 (1):24, November 2011. doi: 10.1088/0004-637X/742/1/24.

A. Lapi, M. Negrello, J. González-Nuevo, et al. Effective Models for Statistical Studies of Galaxy-scale Gravitational Lensing. *The Astrophysical Journal*, 755(1):46, August 2012. doi: 10.1088/0004-637X/755/1/46.

R. Laureijs, J. Amiaux, S. Arduini, et al. Euclid Definition Study Report. *arXiv e-prints*, art. arXiv:1110.3193, October 2011.

G. T. Laurent, J. E. Aguirre, J. Glenn, et al. The Bolocam Lockman Hole Millimeter-Wave Galaxy Survey: Galaxy Candidates and Number Counts. *The Astrophysical Journal*, 623(2):742–762, April 2005. doi: 10.1086/428823.

A. Lawrence, S. J. Warren, O. Almaini, et al. The UKIRT Infrared Deep Sky Survey (UKIDSS). *Monthly Notices of the Royal Astronomical Society*, 379(4):1599–1617, August 2007. doi: 10.1111/j.1365-2966.2007.12040.x.

E. Le Floc'h, C. Papovich, H. Dole, et al. Infrared Luminosity Functions from the Chandra Deep Field-South: The Spitzer View on the History of Dusty Star Formation at 0 <~z <~1. *The Astrophysical Journal*, 632(1):169–190, October 2005. doi: 10.1086/432789.

B. Leibundgut and M. Sullivan. Type Ia Supernova Cosmology. *Space Science Reviews*, 214(2):57, March 2018. doi: 10.1007/s11214-018-0491-8.

D. Lemke, U. Klaas, J. Abolins, et al. ISOPHOT - capabilities and performance. *Astronomy and Astrophysics*, 315:L64–L70, November 1996.

J. Lesgourgues. The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview. *arXiv e-prints*, art. arXiv:1104.2932, April 2011.

J. Lesgourgues and S. Pastor. Massive neutrinos and cosmology. *Physics Reports*, 429(6):307–379, July 2006. doi: 10.1016/j.physrep.2006.04.001.

A. Lewis and S. Bridle. Cosmological parameters from CMB and other data: A Monte Carlo approach. *Physical Review D*, 66(10):103511, November 2002. doi: 10.1103/PhysRevD.66.103511.

A. Lewis, A. Challinor, and A. Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *The Astrophysical Journal*, 538(2):473–476, August 2000. doi: 10.1086/309179.

A. R. Liddle and D. H. Lyth. *Cosmological Inflation and Large-Scale Structure*. 2000.

D. N. Limber. The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. *The Astrophysical Journal*, 117:134, January 1953. doi: 10.1086/145672.

Eric V. Linder. Exploring the Expansion History of the Universe. *Physical Review Letters*, 90(9):091301, March 2003. doi: 10.1103/PhysRevLett.90.091301.

J. Liouville. Note sur la théorie de la variation des constantes arbitraires. *Journal de Mathématiques Pures et Appliquées*, pages 342–349, 1838.

J. Liske, I. K. Baldry, S. P. Driver, et al. Galaxy And Mass Assembly (GAMA): end of survey report and data release 2. *Monthly Notices of the Royal Astronomical Society*, 452(2):2087–2126, September 2015. doi: 10.1093/mnras/stv1436.

C. J. Lonsdale, S. E. Persson, and K. Matthews. Infrared observations of interacting/merging galaxies. *Astrophysical Journal*, 287:95–107, December 1984. doi: 10.1086/162666.

C. J. Lonsdale, D. Farrah, and H. E. Smith. Ultraluminous Infrared Galaxies. In John W. Mason, editor, *Astrophysics Update 2*, page 285. 2006. doi: 10.1007/3-540-30313-8_9.

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

LSST Science Collaboration, P. A. Abell, J. Allison, et al. LSST Science Book, Version 2.0. *arXiv e-prints*, art. arXiv:0912.0201, December 2009.

C. Ma and J. N. Fry. Deriving the Nonlinear Cosmological Power Spectrum and Bispectrum from Analytic Dark Matter Halo Profiles and Mass Functions. *The Astrophysical Journal*, 543(2):503–513, November 2000. doi: 10.1086/317146.

A. V. Macciò, A. A. Dutton, F. C. van den Bosch, et al. Concentration, spin and shape of dark matter haloes: scatter and the dependence on mass and environment. *Monthly Notices of the Royal Astronomical Society*, 378(1):55–71, June 2007. doi: 10.1111/j.1365-2966.2007.11720.x.

P. Madau and M. Dickinson. Cosmic Star-Formation History. *Annual Review of Astronomy and Astrophysics*, 52:415–486, August 2014. doi: 10.1146/annurev-astro-081811-125615.

S. J. Maddox and L. Dunne. MADX - a simple technique for source detection and measurement using multiband imaging from the Herschel-ATLAS survey. *Monthly Notices of the Royal Astronomical Society*, 493(2):2363–2372, April 2020. doi: 10.1093/mnras/staa458.

R. Mandelbaum, A. Slosar, T. Baldauf, et al. Cosmological parameter constraints from galaxy-galaxy lensing and galaxy clustering with the SDSS DR7. *Monthly Notices of the Royal Astronomical Society*, 432(2):1544–1575, June 2013. doi: 10.1093/mnras/stt572.

D. C. Martin, J. Fanson, D. Schiminovich, et al. The Galaxy Evolution Explorer: A Space Ultraviolet Survey Mission. *The Astrophysical Journall*, 619(1):L1–L6, January 2005. doi: 10.1086/426387.

E. Massara, F. Villaescusa-Navarro, and M. Viel. The halo model in a massive neutrino cosmology. *Journal of Cosmology and Astroparticle Physics*, 2014(12):053, December 2014. doi: 10.1088/1475-7516/2014/12/053.

P. McDonald, H. Trac, and C. Contaldi. Dependence of the non-linear mass power spectrum on the equation of state of dark energy. *Monthly Notices of the Royal Astronomical Society*, 366(2):547–556, February 2006. doi: 10.1111/j.1365-2966.2005.09881.x.

A. J. Mead, J. A. Peacock, C. Heymans, et al. An accurate halo model for fitting non-linear cosmological power spectra and baryonic feedback models. *Monthly Notices of the Royal Astronomical Society*, 454(2):1958–1975, December 2015. doi: 10.1093/mnras/stv2036.

A. J. Mead, T. Tröster, C. Heymans, et al. A hydrodynamical halo model for weak-lensing cross correlations. *Astronomy and Astrophysics*, 641:A130, September 2020. doi: 10.1051/0004-6361/202038308.

P. Meszaros. The behaviour of point masses in an expanding cosmological substratum. *Astronomy and Astrophysics*, 37(2):225–228, December 1974.

N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, et al. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087–1092, June 1953. doi: 10.1063/1.1699114.

M. Michałowski, J. Hjorth, and D. Watson. Cosmic evolution of submillimeter galaxies and their contribution to stellar mass assembly. *Astronomy and Astrophysics*, 514:A67, May 2010. doi: 10.1051/0004-6361/200913634.

C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W. H. Freeman, San Francisco, 1973. ISBN 978-0-7167-0344-0, 978-0-691-17779-3.

H. Mo, F. C. van den Bosch, and S. White. *Galaxy Formation and Evolution*. 2010.

H. J. Mo and S. D. M. White. An analytic model for the spatial clustering of dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 282(2):347–361, September 1996. doi: 10.1093/mnras/282.2.347.

S. More, F. C. van den Bosch, M. Cacciato, et al. Satellite kinematics – II. The halo mass–luminosity relation of central galaxies in SDSS. *Monthly Notices of the Royal Astronomical Society*, 392(2):801–816, 12 2008. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2008.14095.x.

A. M. J. Mortier, S. Serjeant, J. S. Dunlop, et al. The SCUBA Half-Degree Extragalactic Survey - I. Survey motivation, design and data processing. *Monthly Notices of the Royal Astronomical Society*, 363(2):563–580, October 2005. doi: 10.1111/j.1365-2966.2005.09460.x.

H. Murakami, H. Baba, P. Barthel, et al. The Infrared Astronomical Mission AKARI*. *Publications of the Astronomical Society of Japan*, 59:S369–S376, October 2007. doi: 10.1093/pasj/59.sp2.S369.

S.G. Murray, B. Diemer, Z. Chen, et al. Thehalomod: An online calculator for the halo model. *Astronomy and Computing*, 36:100487, 2021. ISSN 2213-1337. doi: https://doi.org/10.1016/j.ascom.2021.100487.

J. F. Navarro, C. S. Frenk, and S. D. M. White. The Structure of Cold Dark Matter Halos. *The Astrophysical Journal*, 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. *The Astrophysical Journal*, 490(2):493–508, December 1997. doi: 10.1086/304888.

H. Nayyeri, M. Keele, A. Cooray, et al. Candidate Gravitationally Lensed Dusty Star-forming Galaxies in the Herschel Wide Area Surveys. *The Astrophysical Journal*, 823(1):17, May 2016. doi: 10.3847/0004-637X/823/1/17.

M. Negrello, R. Hopwood, G. De Zotti, et al. The Detection of a Population of Submillimeter-Bright, Strongly Lensed Galaxies. *Science*, 330(6005):800, November 2010. doi: 10.1126/science.1193420.

A. F. Neto, L. Gao, P. Bett, et al. The statistics of Λ CDM halo concentrations. *Monthly Notices of the Royal Astronomical Society*, 381(4):1450–1462, November 2007. doi: 10.1111/j.1365-2966.2007.12381.x.

G. Neugebauer, H. J. Habing, R. van Duinen, et al. The Infrared Astronomical Satellite (IRAS) mission. *The Astrophysical Journal l*, 278:L1–L6, March 1984. doi: 10.1086/184209.

I. Newton. *Opticks or, a treatise of the reflections, refractions, inflections and colours of light. The second edition, with additions. By Sir Isaac Newton, Knt.* London, 1718.

J. Neyman and E. L. Scott. A Theory of the Spatial Distribution of Galaxies. *The Astrophysical Journal*, 116:144, July 1952. doi: 10.1086/145599.

S. J. Oliver, L. Wang, A. J. Smith, et al. HerMES: SPIRE galaxy number counts at 250, 350, and 500 μm. *Astronomy and Astrophysics*, 518:L21, July 2010. doi: 10.1051/0004-6361/201014697.

S. J. Oliver, J. Bock, B. Altieri, et al. The Herschel Multi-tiered Extragalactic Survey: HerMES. *Monthly Notices of the Royal Astronomical Society*, 424(3):1614–1635, 08 2012. ISSN 0035-8711. doi: 10.1111/j.1365-2966.2012.20912.x.

L. Ondaro-Mallea, R. E. Angulo, M. Zennaro, et al. Non-universality of the mass function: dependence on the growth rate and power spectrum shape. *Monthly Notices of the Royal Astronomical Society*, 509(4):6077–6090, February 2022. doi: 10.1093/mnras/stab3337.

A. Pais. *Subtle is the lord– : the science and the life of Albert Einstein / Abraham Pais.* Clarendon Press Oxford [Oxfordshire] ; New York, 1982. ISBN 019853907.

N. Palanque-Delabrouille, C. Yèche, J. Baur, et al. Neutrino masses and cosmology with lyman-alpha forest power spectrum. *Journal of Cosmology and Astroparticle Physics*, 2015:011–011, 11 2015. doi: 10.1088/1475-7516/2015/11/011.

L. Pantoni, A. Lapi, M. Massardi, et al. New Analytic Solutions for Galaxy Evolution: Gas, Stars, Metals, and Dust in Local ETGs and Their High-z Star-forming Progenitors. *The Astrophysical Journal*, 880(2):129, August 2019. doi: 10.3847/1538-4357/ab2adc.

E. Pascale, R. Auld, A. Dariush, et al. The first release of data from the Herschel ATLAS: the SPIRE images. *Monthly Notices of the Royal Astronomical Society*, 415(1):911–917, July 2011. doi: 10.1111/j.1365-2966.2011.18756.x.

G. Patanchon, P. A. R. Ade, J. J. Bock, et al. Submillimeter Number Counts from Statistical Analysis of BLAST Maps. *The Astrophysical Journal*, 707(2):1750–1765, December 2009. doi: 10.1088/0004-637X/707/2/1750.

J. A. Peacock and S. J. Dodds. Non-linear evolution of cosmological power spectra. *Monthly Notices of the Royal Astronomical Society*, 280(3):L19–L26, June 1996. doi: 10.1093/mnras/280.3.L19.

J. A. Peacock and R. E. Smith. Halo occupation numbers and galaxy bias. *Monthly Notices of the Royal Astronomical Society*, 318(4):1144–1156, November 2000. doi: 10.1046/j.1365-8711.2000.03779.x.

E. A. Pearson, S. Eales, L. Dunne, et al. H-ATLAS: estimating redshifts of Herschel sources from sub-mm fluxes. *Monthly Notices of the Royal Astronomical Society*, 435(4):2753–2763, 09 2013. ISSN 0035-8711. doi: 10.1093/mnras/stt1369.

A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *The Astrophysical Journal*, 142:419–421, July 1965. doi: 10.1086/148307.

P. G. Pérez-González, G. H. Rieke, E. Egami, et al. Spitzer View on the Evolution of Star-forming Galaxies from z = 0 to z ~3. *The Astrophysical Journal*, 630(1):82–107, September 2005. doi: 10.1086/431894.

A. Petters, H. Levine, and J. Wambsganss. *Singularity theory and gravitational lensing*. 01 2001.

O. H. E. Philcox, D. N. Spergel, and F. Villaescusa-Navarro. Effective halo model: Creating a physical and accurate model of the matter power spectrum and cluster counts. *Physical Review D*, 101(12): 123520, June 2020. doi: 10.1103/PhysRevD.101.123520.

G. L. Pilbratt, J. R. Riedinger, T. Passvogel, et al. Herschel Space Observatory. An ESA facility for far-infrared and submillimetre astronomy. *Astronomy and Astrophysics*, 518:L1, July 2010. doi: 10.1051/0004-6361/201014759.

Planck Collaboration, N. Aghanim, Y. Akrami, et al. Planck 2018 results. VI. Cosmological parameters. *Astronomy and Astrophysics*, 641:A6, September 2020a. doi: 10.1051/0004-6361/201833910.

Planck Collaboration, N. Aghanim, Y. Akrami, et al. Planck 2018 results. VIII. Gravitational lensing. *Astronomy and Astrophysics*, 641:A8, September 2020b. doi: 10.1051/0004-6361/201833886.

A. Poglitsch, C. Waelkens, N. Geis, et al. The Photodetector Array Camera and Spectrometer (PACS) on the Herschel Space Observatory. *Astronomy and Astrophysics*, 518:L2, July 2010. doi: 10.1051/0004-6361/201014535.

M. Polletta, M. Tajer, L. Maraschi, et al. Spectral Energy Distributions of Hard X-Ray Selected Active Galactic Nuclei in the XMM-Newton Medium Deep Survey. *The Astrophysical Journal*, 663 (1):81–102, July 2007. doi: 10.1086/518113.

A. Pope, R. Chary, D. M. Alexander, et al. Mid-Infrared Spectral Diagnosis of Submillimeter Galaxies. *The Astrophysical Journal*, 675(2):1171–1193, March 2008. doi: 10.1086/527030.

A. Pope, I. Aretxaga, D. Hughes, et al. TolTEC/LMT Extragalactic Legacy Surveys: Completing our Census of Dust-Obscured Star Formation. In *American Astronomical Society Meeting Abstracts #233*, volume 233 of *American Astronomical Society Meeting Abstracts*, page 363.20, January 2019.

J. Prat, J. Blazek, C. Sánchez, et al. Dark energy survey year 3 results: High-precision measurement and modeling of galaxy-galaxy lensing. *Physical Review D*, 105(8):083528, April 2022. doi: 10.1103/PhysRevD.105.083528.

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

J. L. Puget, A. Abergel, J. P. Bernard, et al. Tentative detection of a cosmic far-infrared background with COBE. *Astronomy and Astrophysics*, 308:L5, April 1996.

D. Reed, J. Gardner, T Quinn, et al. Evolution of the mass function of dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 346(2):565–572, 2003. doi: 10.1046/j.1365-2966.2003.07113.x.

D. Reed, R. Bower, C. Frenk, et al. The halo mass function from the dark ages through the present day. *Monthly Notices of the Royal Astronomical Society*, 374(1):2–15, 2007. doi: 10.1111/j.1365-2966.2006.11204.x.

J. Rhodes, A. Refregier, and E. J. Groth. Detection of Cosmic Shear with the Hubble Space Telescope Survey Strip. *The Astrophysical Journal*, 552(2):L85–L88, May 2001. doi: 10.1086/320336.

E. E. Rigby, S. J. Maddox, L. Dunne, et al. Herschel-ATLAS: first data release of the Science Demonstration Phase source catalogues. *Monthly Notices of the Royal Astronomical Society*, 415(3):2336–2348, August 2011. doi: 10.1111/j.1365-2966.2011.18864.x.

H. P. Robertson. Kinematics and World-Structure. *The Astrophysical Journal*, 82:284, November 1935. doi: 10.1086/143681.

N. Roche and S. A. Eales. The angular correlation function and hierarchical moments of ~70000 faint galaxies to R=23.5. *Monthly Notices of the Royal Astronomical Society*, 307(3):703–721, August 1999. doi: 10.1046/j.1365-8711.1999.02652.x.

T. Sawala, C. S. Frenk, R. A. Crain, et al. The abundance of (not just) dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 431(2):1366–1382, 03 2013. ISSN 0035-8711. doi: 10.1093/mnras/stt259.

R. J. Scherrer and E. Bertschinger. Statistics of Primordial Density Perturbations from Discrete Seed Masses. *The Astrophysical Journal*, 381:349, November 1991. doi: 10.1086/170658.

P. Schneider, J. Ehlers, and E. E. Falco. *Gravitational Lenses*. 1992. doi: 10.1007/978-3-662-03758-4.

P. Schneider, C. Kochanek, and J. Wambsganss. *Gravitational Lensing: Strong, Weak and Micro*. 01 2006. ISBN 978-3-540-30309-1. doi: 10.1007/978-3-540-30310-7.

B. Schulz, G. Marton, I. Valtchanov, et al. SPIRE Point Source Catalog Explanatory Supplement. *arXiv e-prints*, art. arXiv:1706.00448, June 2017.

R. Scoccimarro, R. K. Sheth, L. Hui, et al. How many galaxies fit in a halo? constraints on galaxy formation efficiency from spatial clustering. *The Astrophysical Journal*, 546:20–34, 2000.

K. S. Scott, M. S. Yun, G. W. Wilson, et al. Deep 1.1mm-wavelength imaging of the GOODS-S field by AzTEC/ASTE - I. Source catalogue and number counts. *Monthly Notices of the Royal Astronomical Society*, 405(4):2260–2278, July 2010. doi: 10.1111/j.1365-2966.2010.16644.x.

R. Scranton, B. Ménard, G. T. Richards, et al. Detection of Cosmic Magnification with the Sloan Digital Sky Survey. *The Astrophysical Journal*, 633(2):589–602, November 2005. doi: 10.1086/431358.

S. Seitz, P. Schneider, and J. Ehlers. Light propagation in arbitrary spacetimes and the gravitational lens approximation. *Classical and Quantum Gravity*, 11(9):2345–2373, September 1994. doi: 10.1088/0264-9381/11/9/016.

M. Seldner, B. Siebers, E. J. Groth, et al. New reduction of the Lick catalog of galaxies. *Astronomical Journal*, 82:249–256, April 1977. doi: 10.1086/112039.

U. Seljak. Analytic model for galaxy and dark matter clustering. *Monthly Notices of the Royal Astronomical Society*, 318(1):203–213, 10 2000. ISSN 0035-8711. doi: 10.1046/j.1365-8711.2000.03715.x.

R. K. Sheth and A. Diaferio. Peculiar velocities of galaxies and clusters. *Monthly Notices of the Royal Astronomical Society*, 322(4):901–917, April 2001. doi: 10.1046/j.1365-8711.2001.04202.x.

R. K. Sheth and G. Tormen. Large scale bias and the peak background split. *Monthly Notices of the Royal Astronomical Society*, 308(1):119–126, 1999. doi: 10.1046/j.1365-8711.1999.02692.x.

R. Shirley, Y. Roehlly, P. D. Hurley, et al. HELP: a catalogue of 170 million objects, selected at 0.36-4.5 μm, from 1270 deg^2 of prime extragalactic fields. *Monthly Notices of the Royal Astronomical Society*, 490(1):634–656, November 2019. doi: 10.1093/mnras/stz2509.

M. Shuntov. *Cosmic Magnification in COSMOS*. PhD thesis, 06 2019.

C. Sifón, M. Cacciato, H. Hoekstra, et al. The masses of satellites in GAMA galaxy groups from 100 square degrees of KiDS weak lensing data. *Monthly Notices of the Royal Astronomical Society*, 454 (4):3938–3951, December 2015. doi: 10.1093/mnras/stv2051.

G. Siringo, E. Kreysa, A. Kovács, et al. The Large APEX BOlometer CAmera LABOCA. *Astronomy and Astrophysics*, 497(3):945–962, April 2009. doi: 10.1051/0004-6361/200811454.

I. Smail, R. J. Ivison, and A. W. Blain. A Deep Sub-millimeter Survey of Lensing Clusters: A New Window on Galaxy Formation and Evolution. *The Astrophysical Journal*, 490(1):L5–L8, November 1997. doi: 10.1086/311017.

I. Smail, R. J. Ivison, F. N. Owen, et al. Radio Constraints on the Identifications and Redshifts of Submillimeter Galaxies. *The Astrophysical Journal*, 528(2):612–616, January 2000. doi: 10.1086/308226.

M. W. L. Smith, E. Ibar, S. J. Maddox, et al. The Herschel-ATLAS Data Release 2, Paper I. Sub-millimeter and Far-infrared Images of the South and North Galactic Poles: The Largest Herschel Survey of the Extragalactic Sky. *The Astrophysical Journal Supplement Series*, 233(2):26, December 2017. doi: 10.3847/1538-4365/aa9b35.

R. E. Smith, J. A. Peacock, A. Jenkins, et al. Stable clustering, the halo model and non-linear cosmological power spectra. *Monthly Notices of the Royal Astronomical Society*, 341(4):1311–1332, June 2003. doi: 10.1046/j.1365-8711.2003.06503.x.

D. Spergel, N. Gehrels, C. Baltay, et al. Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report. *arXiv e-prints*, art. arXiv:1503.03757, March 2015.

N. S. Sugiyama. Using Lagrangian Perturbation Theory for Precision Cosmology. *The Astrophysical Journal*, 788(1):63, June 2014. doi: 10.1088/0004-637X/788/1/63.

N. Suzuki, D. Rubin, C. Lidman, et al. The Hubble Space Telescope Cluster Supernova Survey. V. Improving the Dark-energy Constraints above z $>$ 1 and Building an Early-type-hosted Supernova Sample. *The Astrophysical Journal*, 746(1):85, February 2012. doi: 10.1088/0004-637X/746/1/85.

A. M. Swinbank, I. Smail, S. Longmore, et al. Intense star formation within resolved compact regions in a galaxy at z = 2.3. *Nature*, 464(7289):733–736, April 2010. doi: 10.1038/nature08880.

A. R. Thompson, B. G. Clark, C. M. Wade, et al. The Very Large Array. *The Astrophysical Journal Supplement Series*, 44:151–167, October 1980. doi: 10.1086/190688.

J. Tinker, A. V. Kravtsov, A. Klypin, et al. Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality. *The Astrophysical Journal*, 688(2):709–728, 2008. doi: 10.1086/591439.

G. Tormen, L. Moscardini, and N. Yoshida. Properties of cluster satellites in hydrodynamical simulations. *MNRAS*, 350, 04 2003. doi: 10.1111/j.1365-2966.2004.07736.x.

T. Treu, A. A. Dutton, M. W. Auger, et al. The SWELLS survey - I. A large spectroscopically selected sample of edge-on late-type lens galaxies. *Monthly Notices of the Royal Astronomical Society*, 417 (3):1601–1620, November 2011. doi: 10.1111/j.1365-2966.2011.19378.x.

M. A. Troxel, N. MacCrann, J. Zuntz, et al. Dark Energy Survey Year 1 results: Cosmological constraints from cosmic shear. *Physical Review D*, 98(4):043528, August 2018. doi: 10.1103/PhysRevD.98.043528.

E. Valiante, M. W. L. Smith, S. Eales, et al. The Herschel-ATLAS data release 1 – I. Maps, catalogues and number counts. *Monthly Notices of the Royal Astronomical Society*, 462(3):3146–3179, 07 2016. ISSN 0035-8711. doi: 10.1093/mnras/stw1806.

F. C. van den Bosch, X. Yang, and H.J. Mo. Linking early- and late-type galaxies to their dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 340(3):771–792, 04 2003. ISSN 0035-8711. doi: 10.1046/j.1365-8711.2003.06335.x.

F. C. van den Bosch, S. More, M. Cacciato, et al. Cosmological constraints from a combination of galaxy clustering and lensing – I. Theoretical framework. *Monthly Notices of the Royal Astronomical Society*, 430(2):725–746, 02 2013. ISSN 0035-8711. doi: 10.1093/mnras/sts006.

F. C. van den Bosch, J. U. Lange, and A. R. Zentner. Basilisk: Bayesian hierarchical inference of the galaxy-halo connection using satellite kinematics - I. Method and validation. *Monthly Notices of the Royal Astronomical Society*, 488(4):4984–5013, October 2019. doi: 10.1093/mnras/stz2017.

E. van Uitert, B. Joachimi, S. Joudaki, et al. KiDS+GAMA: cosmology constraints from a joint analysis of cosmic shear, galaxy-galaxy lensing, and angular clustering. *Monthly Notices of the Royal Astronomical Society*, 476(4):4662–4689, June 2018. doi: 10.1093/mnras/sty551.

L. Van Waerbeke, Y. Mellier, T. Erben, et al. Detection of correlated galaxy ellipticities from CFHT data: first evidence for gravitational lensing by large-scale structures. *Astronomy and Astrophysics*, 358:30–44, June 2000.

S. Veilleux, D. C. Kim, and D. B. Sanders. Optical and Near-Infrared Imaging of the IRAS 1 Jy Sample of Ultraluminous Infrared Galaxies. II. The Analysis. *The Astrophysical Journal Supplement Series*, 143(2):315–376, December 2002. doi: 10.1086/343844.

M. Viola, M. Cacciato, M. Brouwer, et al. Dark matter halo properties of GAMA galaxy groups from 100 square degrees of KiDS weak lensing data. *Monthly Notices of the Royal Astronomical Society*, 452(4):3529–3550, October 2015. doi: 10.1093/mnras/stv1447.

M. Vogelsberger, F. Marinacci, P. Torrey, et al. Cosmological simulations of galaxy formation. *Nature Reviews Physics*, 2(1):42–66, January 2020. doi: 10.1038/s42254-019-0127-2.

A. G. Walker. On Milne's Theory of World-Structure. *Proceedings of the London Mathematical Society*, 42:90–127, January 1937. doi: 10.1112/plms/s2-42.1.90.

D. Walsh, R. F. Carswell, and R. J. Weymann. 0957+561 A, B: twin quasistellar objects or gravitational lens? *Nature*, 279:381–384, May 1979. doi: 10.1038/279381a0.

L. Wang, A. Cooray, D. Farrah, et al. HerMES: detection of cosmic magnification of submillimetre galaxies using angular cross-correlation. *Monthly Notices of the Royal Astronomical Society*, 414(1): 596–601, June 2011. doi: 10.1111/j.1365-2966.2011.18417.x.

J. L. Wardlow, A. Cooray, F. De Bernardis, et al. HerMES: Candidate Gravitationally Lensed Galaxies and Lensing Statistics at Submillimeter Wavelengths. *The Astrophysical Journal*, 762(1):59, January 2013. doi: 10.1088/0004-637X/762/1/59.

M. S. Warren, K. Abazajian, D. E. Holz, et al. Precision Determination of the Mass Function of Dark Matter Halos. *The Astrophysical Journal*, 646(2):881–885, 2006. doi: 10.1086/504962.

W. A. Watson, I. T. Iliev, A. D'Aloisio, et al. The halo mass function through the cosmic ages. *Monthly Notices of the Royal Astronomical Society*, 433(2):1230–1245, 2013. doi: 10.1093/mnras/stt791.

R. H. Wechsler, J. S. Bullock, J-R. Primack, et al. Concentrations of Dark Halos from Their Assembly Histories. *The Astrophysical Journal*, 568(1):52–70, March 2002. doi: 10.1086/338765.

N. N. Weinberg and M. Kamionkowski. Constraining dark energy from the abundance of weak gravitational lenses. *Monthly Notices of the Royal Astronomical Society*, 341(1):251–262, May 2003. doi: 10.1046/j.1365-8711.2003.06421.x.

S. Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity*. John Wiley and Sons, New York, 1972. ISBN 978-0-471-92567-5, 978-0-471-92567-5.

A. Weiß, A. Kovács, K. Coppin, et al. The Large Apex Bolometer Camera Survey of the Extended Chandra Deep Field South. *The Astrophysical Journal*, 707(2):1201–1216, December 2009. doi: 10.1088/0004-637X/707/2/1201.

M. W. Werner, T. L. Roellig, F. J. Low, et al. The Spitzer Space Telescope Mission. *The Astrophysical Journal Supplement Series*, 154(1):1–9, September 2004. doi: 10.1086/422992.

M. White, L. Hernquist, and V. Springel. The halo model and numerical simulations. *The Astrophysical Journal*, 550(2):L129–L132, apr 2001. doi: 10.1086/319644.

S. D. M. White and M. J. Rees. Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering. *Monthly Notices of the Royal Astronomical Society*, 183:341–358, May 1978. doi: 10.1093/mnras/183.3.341.

G. W. Wilson, J. E. Austermann, T. A. Perera, et al. The AzTEC mm-wavelength camera. *Monthly Notices of the Royal Astronomical Society*, 386(2):807–818, May 2008. doi: 10.1111/j.1365-2966.2008.12980.x.

D. M. Wittman, J. A. Tyson, D. Kirkman, et al. Detection of weak gravitational lensing distortions of distant galaxies by cosmic dark matter at large scales. *Nature*, 405(6783):143–148, May 2000. doi: 10.1038/35012001.

E. L. Wright, P. R. M. Eisenhardt, A. K. Mainzer, et al. The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. *The Astronomical Journal*, 140(6): 1868–1881, December 2010. doi: 10.1088/0004-6256/140/6/1868.

X. Yang, H. J. Mo, and F. C. van den Bosch. Constraining galaxy formation and cosmology with the conditional luminosity function of galaxies. *Monthly Notices of the Royal Astronomical Society*, 339 (4):1057–1080, 03 2003. ISSN 0035-8711. doi: 10.1046/j.1365-8711.2003.06254.x.

X. Yang, H. J. Mo, and F. C. van den Bosch. Galaxy Groups in the SDSS DR4. II. Halo Occupation Statistics. *The Astrophysical Journal*, 676(1):248–261, March 2008. doi: 10.1086/528954.

D. G. York, J. Adelman, J. E. Anderson, et al. The Sloan Digital Sky Survey: Technical Summary. *The Astronomical Journal*, 120(3):1579–1587, September 2000. doi: 10.1086/301513.

K. Yoshikawa, Y. P. Taruya, A.and Jing, et al. Nonlinear stochastic biasing of galaxies and dark halos in cosmological hydrodynamic simulations. *The Astrophysical Journal*, 558, 04 2001. doi: 10.1086/322445.

I. Zehavi, Z. Zheng, D. H. Weinberg, et al. The luminosity and color dependence of the galaxy correlation function. *The Astrophysical Journal*, 630(1):1–27, sep 2005. doi: 10.1086/431891.

D. H. Zhao, Y. P. Jing, H. J. Mo, et al. Mass and Redshift Dependence of Dark Halo Structure. *The Astrophysical Journal*, 597(1):L9–L12, November 2003a. doi: 10.1086/379734.

D. H. Zhao, H. J. Mo, Y. P. Jing, et al. The growth and structure of dark matter haloes. *Monthly Notices of the Royal Astronomical Society*, 339(1):12–24, February 2003b. doi: 10.1046/j.1365-8711.2003.06135.x.

Z. Zheng, A. A. Berlind, D. H. Weinberg, et al. Theoretical Models of the Halo Occupation Distribution: Separating Central and Satellite Galaxies. *The Astrophysical Journal*, 633(2):791–809, November 2005. doi: 10.1086/466510.