UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE POSGRADO EN ASTRONOMÍA

INSTITUTO DE ASTRONOMÍA

LA CONEXIÓN GALAXIA-HALO Y SU EVOLUCIÓN

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS (ASTRONOMÍA)

PRESENTA:
ALDO ARMANDO RODRÍGUEZ PUEBLA

TUTOR:
DR. VLADIMIR ANTÓN AVILA REESE, INSTITUTO DE ASTRONOMÍA-UNAM

MÉXICO, D.F. SEPTIEMBRE 2013

GALAXY-HALO CONNECTION AND ITS EVOLUTION

ALDO A. RODRÍGUEZ-PUEBLA

The noblest pleasure is the joy of understanding.

— Leonardo da Vinci

ABSTRACT

Understanding the galaxy-dark matter halo connection is essential to identify the driving processes of galaxy evolution, and to constrain models and simulations of galaxies in the cosmological context. In this direction, semi-empirical approaches, such as the abundance matching technique, the halo occupational distribution model, and the conditional mass function formalism, have emerged as a simple yet powerful statistical approach for connecting observable properties of galaxies to their host CDM halos. This Thesis presents a new approach that combines and extends the main ideas behind these frameworks. It uses observational data as constrains, such as the galaxy stellar mass function and/or its decomposition into centrals and satellites, and the spatial clustering of galaxies. This approach is applied to infer separately the central galaxy stellar-to-halo mass (CHMR) and the satellite galaxy stellar-to-subhalo mass relations (SSMR) and their intrinsic scatters, which are found to be small. We discuss on the robustness and uncertainties of the CHMR and SSMR for local galaxies by considering several combinations of observational data, and emphasize on the differences between the CHMR and the SSMR, usually ignored in the literature. Our approach allows to extend observations to any central/satellite occupational distribution and to masses below the completeness of current surveys by extrapolating the galaxy-dark matter halo connection at dwarf scales. We study the distribution of the most massive satellite-to-central mass ratio in groups/clusters, and find that the centrals are statistically different from the satellites. We study also the massive satellite population ($m_* > 4 \times 10^7$ $M_\odot$) of Milky Way-sized galaxies. Our results suggest that there is not a massive satellite missing problem in the ΛCDM scenario. However, we confirm that the maximum circular velocity of satellites in dark-matter-only simulations with $m_* < 10^8$ $M_\odot$ are systematically larger than those obtained from observational inferences from current analysis of the Milky-Way bright dwarfs. We also infer the galaxy stellar mass function and galaxy clustering from the COSMOS survey and use them to determine the redshift evolution of CHMR and SSMR from $z = 0.2$ to $z = 1$. We report evidence for little evolution of these relations. We further generalize our approach to obtain separately the SHMR (as well as the baryon-to-halo mass relation, BHMR) both for blue and red central galaxies. The results suggest that the SHMR does not segregate significantly by color. Finally, we use static models of disks galaxies seeded in dark matter halos in order to probe the obtained BHMR and SHMR for blue galaxies in the light of local observed scaling and dynamical relations of disk galaxies.

This Thesis builds an statistical framework for constraining the galaxy-(sub)halo connection, and ultimately, theories about the evolution of central and satellite galaxies in the light of the current paradigm of structure formation and evolution, the ΛCDM model.

PUBLICATIONS

Some of the Chapters presented in this Thesis have been published in different journals. Below, the list of them:

"On the stellar and baryonic mass fractions of central blue and red galaxies", Rodríguez-Puebla, Avila-Reese, Colín & Firmani 2011, RevMexAA, 47, 235.

"The Stellar-Subhalo Mass Relation of Satellite Galaxies", Rodríguez-Puebla, Drory & Avila-Reese 2012, ApJ 756, 2.

"The Galaxy-Halo/Subhalo Connection: Mass Relations and Implications for Some Satellite Occupational Distributions", Rodríguez-Puebla Avila-Reese & Drory 2013, ApJ 767, 92.

"The Massive Satellite Population of Milky-Way Sized Galaxies", Rodríguez-Puebla Avila-Reese & Drory 2013, ApJ, in press.

"Seeding the local disk galaxy population", Rodríguez-Puebla, Avila-Reese, Colín & Firmani 2011, RevMexAA (CS), vol. 40, p. 84.

In this age of specialization men who thoroughly know one field are often incompetent to discuss another.

— Richard Feynman

ACKNOWLEDGEMENTS

I would like to express my gratitude to the following people for their support and assistance: first and foremost, thanks to Vladimir Avila Reese without whom this Thesis would not exist. His patience has been extraordinary. Also thanks to Niv Drory for his support and all the facilities to the COSMOS data. Many thanks to everybody who made comments and suggestions that helped to improve the content of this thesis. And finally, this thesis is entirely dedicated to my parents, Graciela & Alejandro. Their love and endless support made this Thesis possible.

CONTENTS

I PREAMBULO 1

1 INTRODUCCIÓN 3
- 1.1 Antecedentes y motivación . . . 3
  - 1.1.1 Más allá de una relación M_*–M_h universal . . . 6
  - 1.1.2 Física y evolución galáctica a partir de la conexión galaxia–halo oscuro . . . 7
- 1.2 Objetivos . . . 9
- 1.3 Metodología y observaciones . . . 11
- 1.4 Contenido . . . 12

II A STATISTICAL APPROACH FOR LINKING GALAXIES TO HALOS 15

2 THE SATELLITE–SUBHALO MASS RELATION 17
- 2.1 Introduction . . . 17
- 2.2 The Abundance Matching Technique . . . 19
  - 2.2.1 Modeling the central & satellite GSMFs . . . 19
  - 2.2.2 The relation to standard abundance matching . . . 20
  - 2.2.3 Inputs for matching abundances . . . 21
  - 2.2.4 Procedure and uncertainties . . . 22
- 2.3 Results . . . 23
  - 2.3.1 The satellite GSMF and the SSMR . . . 23
  - 2.3.2 The satellite CSMF . . . 27
  - 2.3.3 Abundance matching and clustering . . . 29
- 2.4 Summary and Discussion . . . 30
  - 2.4.1 On the inference of the SSMR and its implications for the average mass relation . . . 31
  - 2.4.2 On the intrinsic scatter in the SSMR . . . 34
  - 2.4.3 Implications for satellite/subhalo evolution . . . 34

3 GALAXY–HALO/SUBHALO CONNECTION 37
- 3.1 Introduction . . . 37
- 3.2 Methodology . . . 39
  - 3.2.1 Connecting galaxies to halos and subhalos . . . 40
  - 3.2.2 The two-point correlation function . . . 44
  - 3.2.3 Parameters in the model . . . 45
  - 3.2.4 Observational data sets and strategy . . . 46
- 3.3 The analysis . . . 47
  - 3.3.1 The GSMFs & projected 2PCFs . . . 47
  - 3.3.2 Mass relations . . . 50
  - 3.3.3 Constraining the intrinsic scatter of the stellar-to-(sub)halo mass relations . . . 52
  - 3.3.4 The conditional stellar mass functions . . . 54

3.4 Discussion . . . 55
3.4.1 Robustness and model uncertainties of the stellar-to-(sub)halo mass relations . . . 55
3.4.2 The satellite-subhalo mass relation at the low-mass end . . . 58
3.4.3 Interpreting the bump of the GSMF . . . 59
3.5 Conclusions . . . 59

III APPLICATIONS FOR SOME SATELLITE OCCUPATIONAL DISTRIBUTIONS AT $z \sim 0$ 63

4 PROBABILITY DISTRIBUTIONS OF SATELLITES 65
4.1 Introduction . . . 65
4.2 Probability distributions of satellites . . . 66
4.3 Results . . . 67
4.3.1 The most massive satellite mass distribution . . . 67
4.3.2 The probability of Milky Way–Magellanic Clouds systems . . . 68
4.4 Conclusions . . . 69

5 THE MASSIVE SATELLITE POPULATION OF MILKY WAY-SIZED GALAXIES 73
5.1 Introduction . . . 74
5.2 The method . . . 76
5.2.1 The galaxy group mock catalog . . . 76
5.3 Results and comparison to observations . . . 77
5.3.1 How common is the Milky-Way system? . . . 81
5.3.2 The halo masses of MW-like systems . . . 82
5.4 Satellite vs ΛCDM subhalo populations . . . 84
5.5 Summary & Discussion . . . 87

IV THE STELLAR-(SUB)HALO MASS RELATIONS UP TO $z \sim 1$ 91

6 STELLAR-TO-HALO MASS RELATIONS FROM THE COSMOS SURVEY 93
6.1 Introduction . . . 93
6.2 The galaxy-(sub)halo connection model . . . 95
6.2.1 Theoretical assumptions . . . 96
6.2.2 Parameters in the model . . . 98
6.3 The Data . . . 98
6.3.1 The galaxy stellar mass functions . . . 99
6.3.2 The Angular Correlation function . . . 100
6.4 Procedure & uncertainty treatment . . . 102
6.4.1 The fitting procedure . . . 102
6.4.2 The uncertainties . . . 104
6.5 Results . . . 105
6.5.1 The stellar-to-(sub)halo mass relationships . . . 105
6.5.2 The GSMF of central & satellite galaxies . . . 108
6.5.3 The satellite fraction . . . 110
6.6 Discussion . . . 110
6.6.1 Comparison with previous works . . . 110
6.6.2 On the lack of evolution of the stellar-to-(sub)halo mass relations . 113

6.7 Conclusions 115

V THE GALAXY-TO-HALO CONNECTION OF BLUE AND RED GALAXIES AT $z \sim 0$ 117

7 STELLAR/BARYONIC MASS FRACTIONS OF BLUE AND RED GALAXIES 119
7.1 Introduction 119
7.2 The method 121
7.2.1 Galaxy and Baryonic Stellar Mass Functions 122
7.2.2 Halo and sub-halo mass functions 126
7.3 Results 130
7.3.1 The overall, central, and satellite stellar–halo mass relations 130
7.3.2 The stellar-halo mass relations for central blue and red galaxies . . 132
7.3.3 The baryonic-halo mass relations for central blue and red galaxies 138
7.4 Discussion 139
7.4.1 Comparison with other works 139
7.4.2 Interpretations and consistency of the results 143
7.5 Summary and Conclusions 144

8 STELLAR-TO-HALO MASS RELATIONS AND CLUSTERING OF BLUE/RED GALAXIES 147
8.1 Introduction 147
8.2 The model 149
8.2.1 The fractions of halos hosting blue/red galaxies as a function of mass 149
8.2.2 The fractions of blue/red satellite galaxies as function of halo mass 152
8.2.3 The spatial clustering of galaxies 153
8.2.4 Summary of the model 154
8.3 Observational data and fitting procedure 156
8.3.1 The data 156
8.3.2 Fitting procedure and uncertainties 156
8.4 Results 159
8.4.1 Best fit model and the fraction of halos hosting blue/red central galaxies 159
8.4.2 The stellar-to-halo mass relations of blue and red central galaxies 160
8.4.3 Occupational numbers 162
8.5 Discussion 162
8.5.1 The stellar-to-halo mass relations of blue and red galaxies 162
8.5.2 The host halos of blue and red galaxies 165
8.6 Conclusions 166

VI IMPLICATIONS OF THE BARYON-TO-HALO MASS RELATION ON THE STRUCTURAL AND DYNAMICAL PROPERTIES OF LOCAL GALAXIES 167

9 THE BARYON FRACTION AND THE STRUCTURAL/DYNAMICAL DISK CORRELATIONS I 169
9.1 Introduction 169
9.2 The analytical approach 171

9.2.1 The main equations . . . 173
9.2.2 The Tully-Fisher relation . . . 177
9.2.3 The radius–mass relation . . . 178
9.2.4 Disk-to-total velocity ratio at $2.2R_d$ vs mass . . . 179
9.3 Results: Implications for galaxy scaling relations . . . 179
9.4 Conclusions . . . 182
10 THE BARYON FRACTION AND THE STRUCTURAL/DYNAMICAL DISK CORRELATIONS II 185
10.1 Introduction . . . 185
10.2 Seeding disk galaxies in CDM halos: the static model . . . 187
10.2.1 Adiabatic Contraction . . . 187
10.2.2 Star Formation and stellar quantities . . . 188
10.2.3 The procedure . . . 189
10.3 Results . . . 190
10.3.1 Scaling relations . . . 190
10.3.2 The baryonic-to-dark matter ratio in disk galaxies . . . 193
10.4 Conclusions . . . 195

VII CONCLUSIONS 199
11 CONCLUSIONS 201

VIII APPENDIX 207
A THE SPATIAL CLUSTERING OF GALAXIES IN THE HOD MODEL 209
B THE MARKOV CHAIN MONTE CARLO METHOD 211
B.1 The Metropolis-Hasting algorithm . . . 211
B.2 The MCMC method in multidimensions . . . 212
C THE FITTING PROCEDURE 213
D THE HALO OF BLUE AND RED CENTRALS 215
E THE MEAN COLOR-TO-STELLAR MASS RELATION OF BLUE AND RED CENTRAL GALAXIES 217

BIBLIOGRAPHY 220

ACRONYMS

MOF-Λ Materia Oscura Fría con Λ

FC2P Función de correlación de dos puntos

FMEG Función de masa estelar de galaxias

FMH Función de masa de halos

TCA Técnica de la correlación de Abundacias

TFE Tasa de formación estelar

TFEe Tasa de formación estelar especifica

HAM Historia de agregación de masa

SDSS Sloan Digital Sky Survey

MaNGA Mapping Near Galaxies at APO

CALIFA Calar Alto Legacy Integral Field Spectroscopy Area

AMT Abundance matching Technique

MR Mass relations

GSMF Galaxy stellar mass function

Λ-CMD Λ-Cold dark matter

CSMF Conditional stellar mass function

SHMR Stellar-to-halo mass relation

SSMR Satellite stellar-to-subhalo mass relation

IMF Initial mass function

HOD Halo Occupation Distribution

subhCMF Subhalo conditional mass function

HMF Halo mass function

subHMF Subhalo mass function

MW Milky-Way

2PCF Two-point Correlation function

CHMR Central-halo mass relation

NFW Navarro-Frenk-White

MCMC Markov Chain Monte Carlo

dSph Dwarf spheroidal

MC Magellanic Cloud

LMC Large Magellanic Cloud

SMC Small Magellanic Cloud

Sgr Saggitarius

For Saggitarius

SNe Supernovae

Λ-WMD Λ-warm dark matter

TF Tully-Fisher

COSMOS Cosmic Evolution Survey

SFR Star formation rate

SSFR Specific star formation rate

BG Blue galaxy

RG Red galaxy

GBMF Galaxy baryon mass function

SPS Stellar population synthesis

NYU-VAGC New York University value added galaxy catalog

FOF Friend-of-friends

MAH Mass aggregation history

GHET Galaxy Hybrid Evolutionary Track

bTFR Baryonic Tully-Fisher relation

TFR Tully-Fisher relation

FCME Función condicional de masa estelar

VL Vía Lactea

NM Nube de Magallanes

NMaM Nube Mayor de Magallanes

NMeM Nube Menor de Magallanes

TFE Tully-Fisher estelar

TFB Tully-Fisher bariónica

MHa Metropolis-Hasting algorithm

Mc Markov chain

Part I

PREAMBULO

Se presentan los antecedentes, motivación, objetivos y metodología de la Tesis. La estructura de la misma está basada principalmente en cuatro artículos ya publicados/en prensa, dos más en preparación, y parcialmente hace uso de resultados de otro artículo sometido a arbitraje. En el Capítulo introductorio se busca integrar el contenido de dichos trabajos en una descripción general de la línea de investigación que se siguió durante el desarrollo de la Tesis doctoral.

1 INTRODUCCIÓN

1.1 ANTECEDENTES Y MOTIVACIÓN

El entendimiento de la evolución de las galaxias en el contexto cosmológico es uno de los grandes retos de la astronomía y cosmología de la presente década. Dicho entendimiento se centra principalmente en determinar cómo se agrupan las galaxias y cómo adquieren sus diferentes propiedades (masa baríonica, masa estelar, tipo morfológico, tamaño, color, tasa de formación estelar, metalicidad, etc.) en el contexto del paradigma cosmológico actual de formación de estructuras.

De acuerdo a este paradigma, las estructuras actuales son el producto de la *evolución gravitacional* en un universo en expansión de un campo primigenio de tenues perturbaciones en densidad; dichas perturbaciones a su vez se propone fueron sembradas en la remota época inflacionaria, cuando las fluctuaciones cuánticas del vacío se desconectaron causalmente para convertirse en perturbaciones a la métrica del espacio-tiempo hasta que, con el transcurrir del tiempo cósmico, se vuelven a conectar causalmente convirtiéndose en verdaderas perturbaciones en densidad (reseñas del paradigma cosmológico actual, conocido como el modelo de *Materia Oscura Fría con* Λ (MOF-Λ), se pueden encontrar p. ej. en Avila-Reese [5], Longair [128], Mo, van den Bosch & White [143]).

En el contexto del paradigma MOF-Λ actual, la materia oscura domina sobre la baríonica por lo que la evolución de las perturbaciones que dan origen a las estructuras de gran escala es básicamente gravitacional. Los bariones son atrapados posteriormente en los potenciales gravitacionales de las estructuras de MOF. Cuando estas estructuras alcanzan su máxima expansión y colapsan gravitacionalmente, se convierten en regiones autogravitantes en equilibrio virial separadas ya de la expansión del Universo. Éstas regiones autogravitantes se les denomina como *halos de materia oscura*. El proceso de ensamblaje de masa de los halos se da en una fase altamente no lineal del campo de perturbaciones donde diferentes escalas del campo interactúan unas con otras. Como resultado, los halos oscuros tienen un proceso extendido de ensamblaje, incorporando material constantemente en forma difusa y a través de acreción (fusiones) de estructuras más pequeñas; las cuales mientras éstas sobreviven en el interior del halo, son denominadas como *subhalos*.

Actualmente, la amplia evidencia observacional apunta a que las galaxias se han formado y evolucionado en los halos de materia oscura (y en los subhalos previo a su acreción), donde múltiples mecanismos del tipo (g)astrofísico se mezclan para dar origen a las propiedades que observamos de las galaxias. Naturalmente, esto ha llevado a suponer que las propiedades observadas de las galaxias deberían estar íntimamente relacionadas con las propiedades y evolución de los halos de materia oscura.

Desde un punto de vista teórico, las propiedades, evolución y acumulamiento de los halos de MOF se encuentran bien entendidos mediante el uso de simulaciones numéri-

cas y/o a través de modelos analíticos. Desde esta perspectiva, estudiar las propiedades globales de las galaxias en términos de la conexión galaxia–halo es de gran conveniencia para entender y constreñir los fenómenos astrofísicos que han jugado un rol critico en la formación y evolución de las galaxias. En esta dirección, los *enfoques semiempíricos* son una alternativa sencilla pero a la vez poderosa para lograr la conexión galaxia–halo oscuro.

La esencia de los enfoques semiempíricos es la conexión galaxia–halo utilizando como base observacional las descripciones estadísticas de la población de galaxias, tales como las funciones (por unidad de volumen comóvil) de: luminosidad, masa estelar M_*, velocidad máxima de rotación V_m, dispersión de velocidades σ_V y radio efectivo $r_{\rm eff}$; la función de correlación de dos puntos (FC2P) global y/o separada en intervalos de M_*; etc. La base teórica son las funciones de masa de los halos y subhalos de MOF. Como resultado de conectar estadísticamente las poblaciones teóricas de halos y observadas de galaxias, se obtiene la relación M_*–M_h (así como eventualmente V_m–M_h, $r_{\rm eff}$–M_h, etc.) en las épocas y rangos de masas que puedan proveer las observaciones (v.gr. Vale & Ostriker [216], Kravtsov et al. [115], Conroy, Wechsler & Kravtsov [58], Shankar et al. [187], Weinberg et al. [234], Baldry, Glazebrook & Driver [16], Blanton, Geha & West [31], Vale & Ostriker [217], Conroy & Wechsler [57], Drory et al. [75], Moster et al. [148], Behroozi, Conroy & Wechsler [19], Chae [49], Guo et al. [102], Rodríguez-Puebla et al. [177], Rodríguez-Puebla, Drory & Avila-Reese [178], Papastergis et al. [159], Behroozi, Wechsler & Conroy [20], Reddick et al. [170], por mencionar algunos). El esquema mostrado en la Fig. 2 ilustra la esencia de los enfoques semiempíricos.

La relación M_*–M_h (o $f_s \equiv M_*/M_h - M_h$) es el resultado de la compleja gastrofísica de los bariones atrapados (o perdidos) por los halos de MOF y refleja la *eficiencia de crecimiento de la masa estelar de la galaxia en función de la masa del halo*. El principal objetivo de los modelos/simulaciones de formación y evolución de galaxias deductivos (ab initio) es obtener y explicar "productos finales" como la relación M_*–M_h a cada época. En este sentido, el contar con determinaciones semi-empíricas de la relación M_*–M_h (y otras relaciones como V_m–M_h y $r_{\rm eff}$–M_h) a $z \sim 0$ y otros z's *es de gran utilidad para constreñir los modelos y simulaciones de formación y evolución de galaxias*.

Actualmente existen maneras directas de inferir la relación M_*–M_h a $z \sim 0$, por ejemplo, con la técnica de lentes débiles galaxia-galaxia (Mandelbaum et al. [133], Mandelbaum, Seljak & Hirata [132], Schulz, Mandelbaum & Padmanabhan [184]), la cinemática de satélites (Conroy et al. [56], More et al. [145, 146], Wojtak & Mamon [241]) y estudios en cumulos/grupos de galaxias (Lin & Mohr [126], Yang et al. [244, 249], Hansen et al. [103], Yang et al. [250]). Desafortunadamente estos métodos directos tienen un intervalo dinámico limitado y requieren, por el bajo cociente señal a ruido, del método de apilamiento de muchas galaxias[1] lo cual puede introducir sesgos. Por otro lado, están las técnicas indirectas o semi-empiícas para inferir la relación M_*–M_h, del tipo de las mencionadas anteriormente, mismas que permiten llegar a intervalos dinámicos más amplios al conectar distribuciones estadísticas observadas de las galaxias con las predichas para los halos oscuros.

La técnica más sencilla de los métodos semiempíricos y que menos suposiciones introduce, es la de empatar las funciones acumulativas de masa estelar (FMEG) y de masa

1 En ingles está técnica es conocida como "stacking".

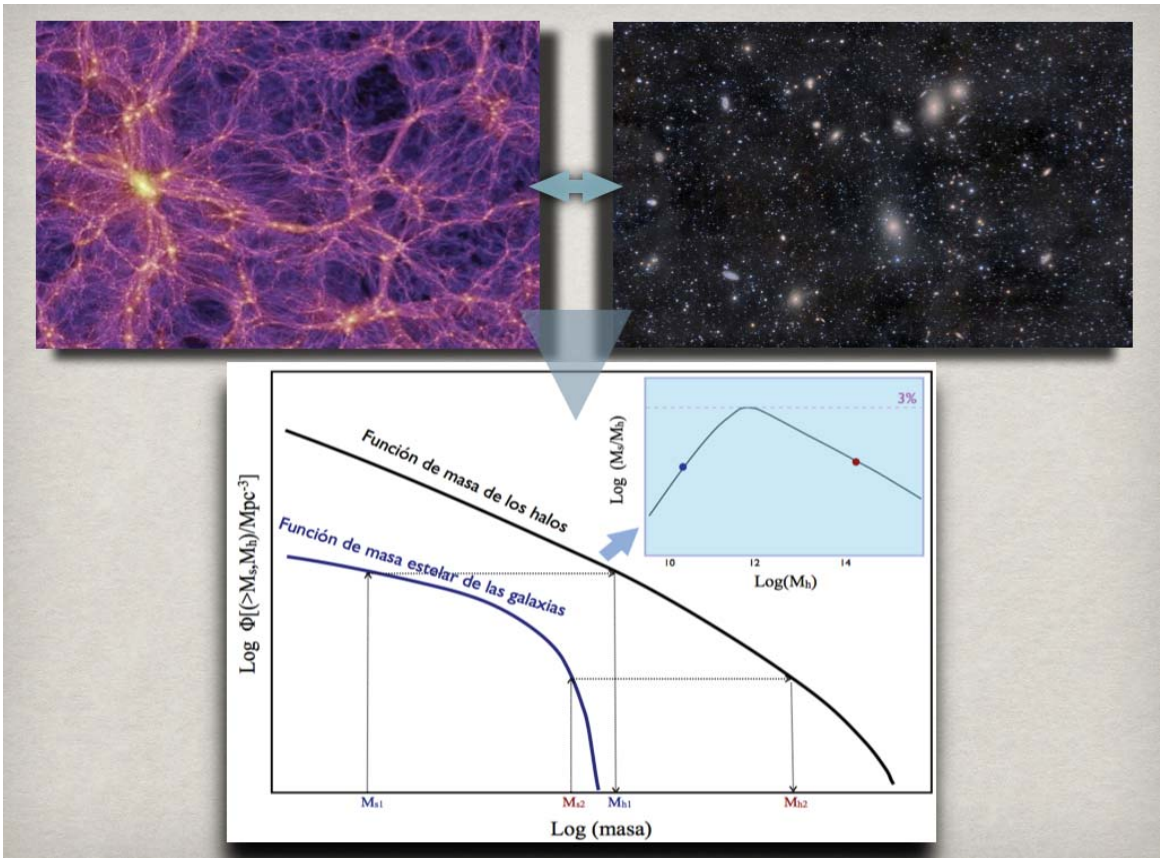

Figure 2: Conexión semi-empíricia galaxia–halo: la población de halos de MOF obtenida en simulaciones numéricas cosmológicas es confrontada estadísticamente con la población observada de galaxias bajo la suposición de que en cada halo/subhalo existe una galaxia central/satélite. Por ejemplo, empatando las funciones de masa acumulativas de galaxias y halos se infiere la masa de halo M_h que le corresponde a una dada masa estelar de galaxia M_* y se puede calcular el cociente M_*/M_h (eficiencia de crecimiento de la masa estelar) en función de la masa del halo; como se aprecia en la figura, este cociente a $z \sim 0$ tiene un pico de máxima eficiencia alrededor de $M_h \sim 10^{12}$ $M_\odot$ (que coincidentemente corresponde a la masa del halo de la Vía Lactea) e incluso en el pico el valor es mucho menor que la fracción bariónica universal, que es 15–17%. La conexión galaxia–halo se puede afinar más introduciendo modelos de ocupación de satélites en los halos y de función condicional de M_*, para lo cual se requiere más restrictores observacionales, como sería la FC2P global o en diferentes intervalos de M_*.

de halos (FMH) para lograr asignar así una M_h a cada M_* (*Técnica de la Correspondencia de Abundancias*, ver 2; [115, 216, 19] y ver más referencias ahí). Esta técnica se ha revelado como muy práctica y poderosa, dando resultados similares a los métodos directos mencionados arriba en el intervalo de M_* y a los z muy locales en los que se pueden comparar. Otros métodos semi-empiícos más sofisticados para lograr la conexión galaxia–halo pero que introducen más suposiciones y parámetros libres, son los así llamados *Modelo de Ocupación de Halos* [25, 60, 259, 1, 208, 258, 229, 228, 223] y *Modelo de la Probabilidad Condicional de la Función de Luminosidad* [245, 244, 249, 247, 250, 148, 121, 122]. En este tipo de modelos, se requiere también información observacional extra como la función de correlación de dos puntos y/o la Función Condicional de Luminosidad (o M_*). En ésta Tesis se presenta un enfoque que integra todas estas técnicas y modelos a fin de investigar más allá de una simple relación M_*–M_h universal y usar los resultados para sondear la validez del escenario cosmológico MOF-Λ, así como constreñir procesos y parámetros astrofśicos de la evolución galáctica.

1.1.1 *Más allá de una relación M_*–M_h universal*

Las inferencias semi-empíricas, por simplicidad, conectan masas de halo a masas estelares (u otras propiedades de escala) sin distinguir si la galaxia es central o satélite o alguna otra propiedad (relación universal). Se entiende por *galaxia central* aquella asociada al halo principal (no contenido en un halo más grande) y generalmente formada en su centro, mientras que una *galaxia satélite* es aquella asociada a un subhalo, es decir un halo contenido en uno más grande; desde un punto de vista observacional, una galaxia satélite es un miembro de un grupo menos masivo (luminoso) que la galaxia central del grupo[2]. La división de galaxias en centrales y satélites en este sentido, es una primera aproximación de lo que es el medio ambiente que rodea una galaxia determinada. Una galaxia de una dada masa estelar, la Vía Láctea por ejemplo, no tiene las mismas propiedades si es central en su halo o es satélite en un cúmulo de galaxias. Por lo tanto, *es crucial obtener la relación M_*–M_h al menos separada en galaxias centrales y satélites.* La diferencia entre ambas relaciones refleja cuánto el medio ambiente del halo anfitrión influye sobre la ulterior eficiencia de crecimiento de masa estelar de las galaxias y sobre la evolución del subhalo.

En las inferencias semi-empíricas comunes se supone también implícitamente que la eficiencia de crecimiento de M_* es función sólo de la masa del halo, por lo que la relación M_*–M_h se considera universal, aunque se introduce una dispersión intrínseca que no se especifica con qué propiedades galácticas correlacionaría. Es bien sabido en la astronomía extragaláctica que la masa estelar o luminosidad no determinan de manera dominante el resto de las propiedades de las galaxias, como ocurre con las estrellas. A paridad de M_* las galaxias presentan una amplia distribución en propiedades intensivas como índice de color, tipo morfológico (T), tasa de formación estelar específica

2 En esta caso nos referimos a "grupos" basados en halos, es decir el conjunto de todas las galaxias ligadas gravitacionalmente en el interior del radio virial de un halo anfitrión (ver Yang et al. [249], Yang, Mo & van den Bosch [247]). Esta definición de grupo es algo diferente a la que históricamente se ha usado en la astronomía observacional. Un grupo basado en halo puede ser un cúmulo o grupo de galaxias pero también una galaxia tipo Vía Láctea con sus satélites.

(TFEe≡TFE/M_*), etc.; ver una reseña en Blanton & Moustakas [32]. En realidad las distribuciones de la mayoría de las propiedades suelen ser bimodales, correlacionando los modos con M_*. En términos muy generales, se habla de dos secuencias o nubes en la distribución de las propiedades galácticas: *(a)* la nube de galaxias dominadas por un disco soportado por rotación ($T \geq 1$ o índice de Sèrsic $n \leq 2.4$), azules, con TFEe altas (activas) y localizadas generalmente en ambientes de densidad promedio o baja; *(b)* la secuencia de galaxias dominadas por un esferoide soportado por dispersión de velocidades ($T < 1$ o índice de Sèrsic $n > 2.4$), rojas, con TFEe bajas (pasivas) y localizadas generalmente en ambientes densos (v. gr., Norberg et al. [155, 154], Zehavi et al. [259], Li et al. [123], Zehavi et al. [258], Weinmann et al. [236], Poggianti et al. [164], Martínez, O'Mill & Lambas [134], Blanton & Berlind [30], Hansen et al. [103], Yang, Mo & van den Bosch [247], Blanton & Moustakas [32]). Mientras que las galaxias del primer modo tienden a ser de baja masa, las del segundo modo tienden a ser de masas más altas, no obstante hay una fuerte dispersión en esta tendencia.

Por tanto, una cuestión de relevancia a estudiar es si la marcada bimodalidad en la distribución de propiedades de las galaxias se revela en la relación M_*–M_h, es decir si las galaxias se segregan por color, TFEe, morfología, etc. en dicha relación. De ser así, significaría que la masa del halo influye no sólo en la eficiencia de la formación de la masa estelar de las galaxias sino que también en sus propiedades intensivas. Si no es así, entonces es más probable que las propiedades intensivas, mismas que parcialmente correlacionan con la masa estelar, son determinadas en gran parte por procesos internos de la galaxia o por otras propiedades del halo aparte de su escala, por ej. su historia de agregación de masa (HAM), la presencia o no de fusiones mayores, su momento angular expresado a través del parámetro de giro λ. En conclusión, es de gran relevancia determinar la relación M_*–M_h *por separado al menos para los dos principales modos de la distribución de propiedades intensivas de las galaxias, por ejemplo, para las galaxias divididas en rojas y azules*.

1.1.2 *Física y evolución galáctica a partir de la conexión galaxia–halo oscuro*

Hasta ahora se habló de la relación M_*–M_h, siendo ésta el "producto final" de los procesos de evolución de galaxias en función de la masa del halo, pero también es de gran interés poder determinar la relación masa bariónica (M_b=M_*+ M_g)–M_h, donde M_g es la masa en gas frío[3] en la galaxia. Dicha relación, comparada con la relación M_*–M_h, da información sobre la eficiencia de transformación del gas en estrellas en función de la escala. Como se verá a lo largo de esta Tesis, la razón f_s=M_*/M_h tiene un pico (máxima eficiencia de crecimiento de M_*) a $M_h \sim 10^{12}$ $M_\odot$ (ver Fig. 2). A masas mayores, decrece con M_h y a masas menores decrece aún más rápido mientras menor es M_h. Este comportamiento con la escala se interpreta principalmente de la siguiente manera:

- En halos con $M_h > 10^{12}$ $M_\odot$, a mayor masa, el tiempo de enfriamiento radiativo del gas que se virializa a través de choques durante el colapso gravitacional del

3 Estrictamente el gas frío de una galaxia en su mayoría está constituido de hidrógeno atómico y molecular en menor medida de otros componentes como molculas de CO, etc. Por tanto, a lo largo de está Tesis nos referiremos al gas frío como aquel constituido de hidrógeno atómico y molecular.

halos es más largo, de tal manera que sólo una fracción pequeña de los bariones se incorpora a la galaxia en épocas muy tempranas a través de procesos violentos de fusiones; además estas galaxias tienen hoyos negros supermasivos que al ser alimentados producen Núcleos Galácticos Activos capaces de eyectar y calentar gas evitando que éste caiga ulteriormente en la galaxia, dejándola con una fracción f_s pequeña.

- En halos con $M_h < 10^{12}$ $M_\odot$, mientras menos masivos son, con más facilidad se pierde gas de la galaxia, y eventualmente del halo, debido a la retroalimentación de la TFE (principalmente por la energía inyectada por las supernovas); para halos de masas muy pequeñas, incluso su potencial es tan débil que casi no se logra atrapar el gas intergaláctico ionizado originado por la primera generación de estrellas y/o galaxias, las cuales produjeron un fondo de radiación en el ultravioleta calentando el gas y así volviendo ineficiente la formación de estelar y entonces el crecimiento de las galaxias (Hoeft et al. [107]).

No obstante, aparte de los procesos globales de captura y pérdida del gas, el mismo proceso de transformación del gas en estrellas podría ser función de la masa. En efecto, mientras menos masivas son las galaxias, mayores fracciones de gas frío suelen tener, lo cual sugiere que fueron menos eficientes en transformar su gas en estrellas. *Por lo tanto, la restricción semi-empírica de la relación M_b–M_h es de gran utilidad para constreñir el rol de los procesos de eyección de gas por retroalimentación y de eficiencia de formación de gas molecular y ulterior formación estelar, en función de la masa del halo*. Además es importante, como en el caso de la relación M_*–M_h, contar con esta determinación separada en galaxias centrales/satélites y rojas/azules.

La razón M_b/M_h (llamada fracción galáctica bariónica, f_b) en función de M_h es un ingrediente clave en modelos de galaxias tanto del tipo estático como del tipo evolutivo, en especial para galaxias con morfología dominda por disco. Dicha fracción es un parámetro utilizado en este tipo de modelos para calcular las propiedades de discos en equilibrio centrífugo "cargados" en halos oscuros (v.gr., Mo, Mao & White [144], Zavala et al. [255], Dutton et al. [80], Gnedin et al. [96], Dutton & van den Bosch [79]) o para calcular la evolución completa de éstos discos utilizando como esqueleto la evolución de los halos oscuros, e incluyendo la formación estelar y la evolución secular de los discos (v.gr., Firmani & Avila-Reese [82, 83], Firmani, Avila-Reese & Rodríguez-Puebla [85], van den Bosch [218]). Usando este tipo de modelos, es interesante explorar los efectos de la relación f_b–M_h obtenida para galaxias de disco (digamos azules) sobre las relaciones de escala y otras correlaciones, tales como la relación de Tully-Fisher estelar y radio–M_*, la correlación entre el cociente M_* a masa dinámica (o velocidad al máximo del disco a velocidad total) con M_h, M_*, la densidad superficial, etc. (ver por ej. Zavala et al. [255], Dutton et al. [80], Gnedin et al. [96], Avila-Reese et al. [12], Reyes et al. [171]). Este tipo de análisis permite sondear la autoconsitencia de la conexión galaxia–halo de materia oscura a nivel de las propiedades internas y observables de las galaxias de disco y permite identificar nuevos posibles procesos astrofísicos relevantes en la evolución de las galaxias de disco. En miras a la futura generación de catálogos y muestras extensas de galaxias locales con propiedades estructurales, dinámicas, de poblaciones estelares

y de metalicidad determinadas de manera uniforme[4], los resultados de estos análisis podrán ser confrontados con observaciones y usados para interpretar las mismas.

Un sondeo más empírico de la autoconsistencia entre la conexión galaxia–halo de materia oscura y las propiedades internas de las galaxias se puede lograr mapeando dicha conexión obtenida con los métodos estadísticos sobre planos de correlaciones entre propiedades de las galaxias y viceversa (ver intentos preliminares en esta dirección en Blanton, Geha & West [31], Chae [49]), en este proceso de mapéo se revelan aspectos claves de la gastrofísica implicada en la formación de las galaxias y/o del marco teórico subyacente, el escenario de MOF-Λ. Dicho mapéo se puede realizar tanto a $z \sim 0$ como a otros z's en los que exista la información observacional necesaria. También, dependiendo de la disponibilidad de los datos observacionales, este mapéo se puede realizar por separado para distintos tipos de galaxias ya sea por color, tipo morfológico, TFEe, etc.

La determinación de las FMEG y funciones de correlación a diferentes corrimientos al rojo son cada vez más rutinarias, haciendo uso de catastros como el COSMOS [185], AEGIS [65], CANDELS [100], etc. Esto genera entonces la posibilidad de obtener la conexión galaxia–halo a diferentes épocas. Una importante consecuencia de determinar las relaciones entre propiedades extensivas de las galaxias y la masa del (sub)halo a diferentes épocas es que se puede hacer uso de las historias individuales de agregación de masa de los halos para obtener entonces la evolución individualizada (trazas evolutivas) de las propiedades de las galaxias, por ejemplo de la M_* [57, 84]. En otras palabras, al lograr la conexión semi-empírica entre, p. ej., M_* y M_h y al usar la evolución teórica de M_h, ¡ se logra también inferir la evolución "individual" de M_*! Este enfoque iniciado por Conroy & Wechsler [57] y Firmani & Avila-Reese [84] ha cobrado popularidad reciéntemente (v. gr. Yang et al. [250], Moster, Naab & White [147], Behroozi, Wechsler & Conroy [20]).

El enfoque semiempírico, basado en en la información estadística que proveen los grandes catastros modernos, se ha consolidado definitivamente como una herramienta poderosa para constreñir los procesos astrofísicos y evolutivos de las galaxias y para poner a prueba el paradigma cosmológico subyacente. Los trabajos realizados en esta Tesis constituyen una contribución a esta línea de investigación reciente.

1.2 OBJETIVOS

La presente Tesis se enfoca en una serie de objetivos en torno a la cuestión general de la conexión semi-empírica galaxia–halo de MOF a través de un modelo estadístico, así como al uso de los resultados obtenidos para constreñir propiedades poblacionales a $z \sim 0$ y aspectos evolutivos de las galaxias desde $z \sim 1$. El enfoque desarrollado permite "extender" las limitaciones de las observaciones al nivel estadístico más completo requerido, así como extrapolar las observaciones a masas donde las observaciones todavía no son completas. Los objetivos particulares más relevantes de esta Tesis son:

4 Estos catálogos se están ya generando o se generarán con proyectos como "Mapping Near Galaxies at APO" (MaNGA) en el SDSS-IV y "Calar Alto Legacy Integral Field Spectroscopy Area" (CALIFA). Haciendo uso de la técnica de espectroscopía de campo integral, se obtienen los espectros en absorpción y emisión a lo largo y ancho de las galaxias que cuentan ya con fotometría bidimensional multibanda de tal manera que todas las propiedades mencionadas, tanto locales como globales, pueden ser determinadas

- Explorar la viabilidad o no de la suposición común de que la relación M_*–M_h es la misma para galaxias centrales y satélites. Esto, tomando en cuenta las dos definiciones de masa del subhalo, al momento de su acreción y al momento de observación.

- Desarrollar un modelo semiempírico estadístico para inferir las relaciones M_*–M_h locales de galaxias centrales y satélites utilizando no sólo la correspondencia de abundancias sino además información sobre la función de correlación de dos puntos, misma que permite constreñir las distribuciones ocupacionales de satélites en halos. En base a esto, probar el comportamiento general de la relación M_*–M_h separada para galaxias centrales y satélites. Constreñir la dispersión de estas relaciones y predecir la Función de Masa Estelar Condicional en función de la masa del halo anfitrión.

- Explorar la auto-consistencia del enfoque semiempírico y sus aciertos o no con las predicciones que hace. En base a esto, sondear el escenario jerárquico de MOF-Λ subyacente y evaluar si nuestros resultados pueden constreñir procesos/parámetros astrofísicos de este escenario.

- Aplicar el modelo semiempírico desarrollado para extender y extrapolar a escalas menores las observaciones usadas como base del modelo con el fin de predecir diferentes estadísticas de ocupación de halos. Por ejemplo, las distribuciones y dinámica interna de satélites en galaxias del tipo Vía Láctea y las distribuciones del cociente masa del satélite más masivo entre masa de la central (brecha) en función de la masa. Usar los resultados al nivel de observaciones para hacer inferencias astrofísicas de diversa índole.

- Extender el modelo estadístico desarrollado para (1) inferir la relación M_*–M_h separada no sólo en centrales y satélites, sino también en galaxias rojas y azules; y (2) para inferir las distribuciones ocupacionales de halos para galaxias centrales rojas/azules y sus fracciones correspondientes de satélites rojos/azules.

- Estudiar si la relación M_*–M_h se segrega por color o si la dispersión de esta relación correlaciona con el color. Constreñir las fracciones de halos anfitriones de galaxias azules y rojas en función de M_h. Explorar en general qué puede estar a la base de la dispersión en la estrecha relación M_*–M_h desde halos galácticos hasta halos de cúmulos de galaxias. Explorar si esta dispersión depende de la riqueza del número del grupo.

- Generalizar el análisis con el modelo estadístico a otras épocas, haciendo uso en particular de las observaciones del catastro COSMOS (hasta $z \sim 1$) e inferir así una conexión evolutiva entre la población galáctica local y las de alto z. En particular, explorar si la relación M_*–M_h de galaxias centrales y satélites cambia significativamente o no con el corrimiento al rojo y determinar la evolución de la fracción de satélites en función de la masa.

- Generalizar el modelo estadístico semiempírico para determinar las relaciones M_b–M_h de galaxias rojas/azules, es decir incluir la información del contenido de

gas frío en las galaxias. Estudiar las implicaciones para la evolución galáctica de la fracción bariónica de galaxias rojas y azules en función de la masa obtenida (f_b=M_b/M_h vs M_h).

- Sondear las consecuencias de la fracción bariónica $f_b(M_h)$ obtenida para galaxias azules (tardías) en lo que concierne a las propiedades estructurales y dinámicas de las galaxias de disco. Desarrollar para esto un modelo "estático" de sembrado de galaxias de disco en equilibrio centrífugo en el interior de halos de MOF contraídos adiabáticamente por la presencia del disco, donde la fracción bariónica es un parámetro base. Explorar si las relaciones de escala y otras correlaciones a nivel estelar y bariónico son consistentes con las observaciones.

1.3 METODOLOGÍA Y OBSERVACIONES

Para lograr los objetivos planteados en esta Tesis (Sección 1.2) se requiere de modelos estadísticos capaces de conectar las poblaciones teóricas de halos y subhalos de MOF con las poblaciones observadas de galaxias centrales y satélites. En este sentido, como se describió en la Sección 1.1, varias técnicas y modelos semiempíricos fueron ya desarrollados en los últimos años. La principal herramienta metodológica de esta Tesis será un nuevo modelo semiempírico más completo que los previos. El modelo desarrollado aquí combina la Técnica de la Correspondencia de Abundancias con los modelos de Ocupación de Halos y de Probabilidad Condicional de la Función de Luminosidad (de masa estelar en nuestro caso).

El modelo requiere de un método estadísitico de ajuste paramétrico potente. En algunos casos se usará la minimización de cuadrados con el método de conjunto de direcciones de Powell en multidimensiones. Sin embargo, para nuestro modelo completo, se aplicará el método más poderoso de Cadenas de Markov Monte Carlo con implementaciones desarrolladas para nuestro caso particular.

En cuanto a la base observacional para el modelo estadístico semiempírico, la principal muestra local que se usará es el catálogo de grupos basados en halos de Yang et al. [249] y Yang, Mo & van den Bosch [247]. Este catálogo está construido de una muestra completa de galaxias desde $M_* \sim 7 \times 10^8$ $M_\odot$ del DR4 del Sloan Digital Sky Survey (SDSS). Con este catálogo se construye la FMEG separada en galaxias centrales y satélites y para cada caso, en rojas y azules. También se hará uso de FMEG obtenidas por otros autores en base a muestras obtenidas de otro DRs del SDSS, (v. gr., Baldry, Glazebrook & Driver [16]). En este caso, sólo se cuenta con la FMEG total (no separada en centrales/satélites y/o rojas/azules) pero se llega a masas más pequeñas, debido a correcciones por completez que incluyen el brillo superficial.

Para la función de correlación de dos puntos (FC2P) local se usarán los resultados de Yang et al. [250] obtenida a partir del DR7 del SDSS. Estos autores presentan la FC2P proyectada de las galaxias en diferentes intervalos de masa estelar, lo cual nos provee una base observacional más completa para constreñir los parámetros de nuestro modelo. Li et al. [123] presenta las FC2P proyectadas por separado para las galaxias rojas y azules, lo cual nos sirve para comparar con las predicciones de nuestro modelo.

Para los datos observacionales a otras épocas, se usarán en nuestras exploraciones varias muestras y recopilaciones de la FMEG. Sin embargo, los resultados presentados en esta Tesis se refieren principalmente a los datos del catastro COSMOS, Scoville et al. [185] y Drory et al. [75]. Estos datos fueron cedidos amablemente por el Dr. Niv Drory para nuestro uso. El catastro COSMOS consta de $\sim 300,000$ galaxias con z fotométrico determinado hasta $z \sim 1$; con el mismo se pueden obtener las FMEG promedio al menos en cuatro intervalos de z. Las galaxias tienen definida su masa estelar, color y TFE. Debido a la pequeña área del cielo que cubre el COSMOS, las FMEG sufren del problema de varianza cósmica; en particular entre $z \sim 0.5$ y 0.9 parece ser que la muestra se encuentra en regiones algo sobredensas. Por lo tanto, buscaremos introducir algunas correcciones para resolver este problema.

En cuanto a la FC2P, en base a los datos del COSMOS, hemos usado un método adecuado para extraer la FC2P angular en los mismos intervalos de z en que obtuvimos las FMEG. Se aplicarán también una serie de correcciones por efectos de borde y por varianza cósmica.

Finalmente, para la modelación de galaxias sembradas en halos de MOF en una época dada (modelos "estáticos"), se implementará un modelo iterativo tipo Mo, Mao & White [144]. Con la generalización que se hará de dicho modelo, se podrán "sembrar" discos o esferoides en halos de MOF con un dado perfil de densidad y momento angular, y resolver así la distribución de masas estelar, bariónica y oscura de todo el sistema acoplado. Se usará también un código seminumérico de evolución de galaxias de disco (Firmani & Avila-Reese [82], Firmani, Avila-Reese & Rodríguez-Puebla [85]) para seguir la evolución completa de galaxias bajo las restricciones de fracción bariónica que se infieran del estudio semiempírico.

1.4 CONTENIDO

En lo que sigue, la Tesis presenta diez Capítulos más, agrupados en seis Partes.

- En la Parte II, se aborda la cuestión sobre cuánto afecta a la inferencia de la relación M_*–M_h el considerar o no la separación de las galaxias en centrales y satélites y de los halos en principales y subhalos (Capítulo 2). Luego se presenta nuestro modelo estadístico completo y se explora la firmeza de las inferencias de la relación M_*–M_h separada para galaxias centrales y satélites, su dispersión y predicciones en cuanto a las distribuciones ocupacionales de centrales y satélites en los halos (Capítulo 3). El contenido del ambos Capítulos fue parte de dos artículos publicados: Rodríguez-Puebla, Drory & Avila-Reese 2012, ApJ 756, 2; Rodríguez-Puebla Avila-Reese & Drory 2013, ApJ 767, 92.

- En la Parte III se aplican las distribuciones ocupacionales obtenidas para explorar la dependencia con M_h del cociente masa del satélite más masivo a masa de la central (brecha) y las probabilidades de encontrar galaxias con masas similares a tipo las Nubes de Magallanes en halos de distintas masas (Capítulo 4). En el Capítulo 5, en base a un catálogo sintético construido con el modelo semiempírico completo, se estudian una serie de distribuciones (configuraciones) de satélites más masivos que Fornax ($M_* \sim 4 \times 10^7$ M$_\odot$) en galaxias con masas estelares similares a la de la

Vía Láctea con el fin de encontrar la probabilidad de un sistema como el de la Vía Láctea, encontrar las masas de halo más probables para este sistema y constreñir la dinámica interna de los satélites en galaxias tipo Vía Láctea; este último objetivo se pretende usar como prueba al escenario MOF-Λ a escalas de galaxias enanas. El contenido del Capítulo 4 corresponde a la segunda parte del artículo publicado en el ApJ mencionado en la Parte II (Rodríguez-Puebla, Avila-Reese & Drory 2013, ApJ 767, 92, mientras que el Capítulo 5 corresponde al contenido de otro artículo en el ApJ: Rodríguez-Puebla, Avila-Reese & Drory 2013, ApJ (en prensa).

- La extensión de nuestro modelo estadístico semiempírico a otros corrimientos al rojo (hasta $z \sim 1$) es presentada en la Parte IV. Se obtienen resultados relacionados a la relación M_*–M_h separada en centrales y satélites usando los datos del catastro COSMOS; se predice también la evolución de la fracción de satélites en función de la masa. El único Capítulo de esta parte (6), esta reflejado de un artículo en preparación.

- En la Parte V, usando la Técnica de la Correspondencia de Abundancias se obtienen las relaciones locales M_*–M_h y M_b–M_h de galaxias centrales separadas en rojas y azules y se discuten las implicaciones (Capítulo 7). Para lograr estas inferencias, se tuvo que hacer algunas suposiciones acerca de los halos que albergan galaxias rojas y azules. En el Capítulo 8 se presenta un enfoque más general, basado en nuestro modelo semi-empŕico completo, para constreñir la relación M_*–M_h de galaxias centrales y satélites separadas en rojas y azules, tomando en cuenta información relacionada la FC2P. Tambýen encontramos estás relaciones de manera local al aplicar el mismo modelo y restringimos (en vez de suponer a priori) las funcionen de masa de los halos anfitriones de galaxias rojas y azules. El Capťulo 7 corresponde a un artículo publicado: Rodríguez-Puebla, Avila-Reese, Colín &T Firmani 2011, RevMexAA, 47, 235; mientras que el Capítulo 8 aborda algunas cuestiones en el articulo remitido, Lacerna, Rodríguez-Puebla, Avila-Reese, Hernández-Toledo 2013, ademas de ser un artículo en preparación.

- Finalmente, en la Parte VI se exploran las consecuencias de la fracción bariónica en función de M_h obtenidas para galaxias azules (tardías en una primera aproximación) sobre las propiedades estructurales y dinámicas de galaxias de disco. En el Capťulo 9 se presenta un enfoque analítico capaz de predecir las tendencias promedio de las relaciones de escala bariónicas, mientras que en el Capítulo 10 se usa un modelo completo para sembrar discos en halos de MOF tomando en cuenta la contracción adiabática que sufre el halo y las diferentes dispersiones en los parámetros del modelo.

Las conclusiones más generales de la Tesis son presentadas en la Parte VII. Las conclusiones más especificas se presentan al final de cada Capítulo.

Part II

A STATISTICAL APPROACH FOR LINKING GALAXIES TO HALOS

En esta Parte II, se estudia la conexión galaxia-(sub)halo enfatizando la correspondencia natural entre galaxia central y halo, y galaxia satélite y subhalo. En particular, se exploran las relaciones masa estelar galaxia central-masa halo, M_*-M_h, y masa estelar galaxia satélite-masa subhalo, m_*-$m_{\rm sub}$, tomando en cuenta dos definiciones de la masa de los subhalos: al tiempo de su acreción, $m_{\rm sub}^{\rm acc}$, y al tiempo de su observación, $m_{\rm sub}^{\rm obs}$. En el Capitulo 2 se generaliza la técnica de la correlación de abundancias para inferir las relaciones M_*-M_h y m_*-$m_{\rm sub}$ por separado. De esta manera se explora la viabilidad de la suposición comúnmente adoptada en la literatura de que la relaciones M_*-M_h y m_*-$m_{\rm sub}$ son idénticas, tanto para $m_{\rm sub}^{\rm acc}$ como para $m_{\rm sub}^{\rm obs}$. El resultado es que esto no es correcto. En el Capitulo 3 se presenta un modelo estadístico más completo, el cual relaciona la función de masa estelar de galaxias centrales y satétiles, la función de masa de halos/subhalos de materia oscura fría, las funciones condicionales de masa estelar de satélites y las funciones de correlación de dos puntos en diferentes intervalos masa. Este nuevo modelo permite inferir las relaciones M_*-M_h y m_*-$m_{\rm sub}$ por separado de una manera auto-consistente con todo el cuerpo de observaciones. Mostramos que la relación m_*-$m_{\rm sub}^{\rm obs}$ es diferente a la relación m_*-$m_{\rm sub}^{\rm acc}$, y a su vez ambas son diferentes a la relación M_*-M_h de galaxias centrales. Se muestra que estas relaciones obtenidas a través de nuestro modelo son firmes y con dispersiones intrínsecas pequeñas. Se demuestra que la correlación de abundancias es equivalente a la correlación de los números ocupacionales en halos (relacionados con el acumulamiento espacial), razón por la cual la conexión galaxia-halo obtenida con la técnica de la correlación de abundancias predice funciones de correlación cercanas a las observadas. No obstante, como se demuestra, los resultados son más firmes y exactos cuando se toma en cuenta de manera auto-consistente tanto el empate la correlación de abundancias como el acumulamiento espacial.

El Capítulo 2 corresponde al artículo publicado: *"The Stellar-Subhalo Mass Relation of Satellite Galaxies", Rodríguez-Puebla, Drory & Avila-Reese 2012, ApJ 756, 2*. El Capítulo 3 corresponde a una parte del artículo publicado: *"The Galaxy-Halo/Subhalo Connection: Mass Relations and Implications for Some Satellite Occupational Distributions", Rodríguez-Puebla Avila-Reese & Drory 2013, ApJ 767, 92*

2

THE STELLAR–SUBHALO MASS RELATION OF SATELLITE GALAXIES AT $z \sim 0$

This Chapter was published as: Rodríguez-Puebla A.; Drory N.; Avila-Reese V., 2012, ApJ, 756, 2.

ABSTRACT

We extend the abundance matching technique (AMT) to infer the satellite–subhalo and central–halo mass relations (MRs) of local galaxies, as well as the corresponding satellite conditional mass functions. We use the observed galaxy stellar mass function (GSMF) decomposed into centrals and satellites and the Λ-CDM distinct halo and subhalo mass functions as inputs. We explore the effects of defining the subhalo mass, $m_{\rm sub}$, at the time of (sub)halo accretion ($m_{\rm sub}^{\rm acc}$) versus defining it at the time of observation ($m_{\rm sub}^{\rm obs}$); and we test the standard assumption that centrals and satellites follow the same MRs. We show that this assumption leads to predictions in disagreement with observations, specially when $m_{\rm sub}^{\rm obs}$ is used. Instead, we find that when the satellite–subhalo MRs are constrained by the satellite GSMF, they are always different from the central–halo MR: the smaller the stellar mass, the less massive is the subhalo of satellites as compared to the halo of centrals of the same stellar mass. This difference is more dramatic when $m_{\rm sub}^{\rm obs}$ is used instead of $m_{\rm sub}^{\rm acc}$. On average, for stellar masses lower than $\sim 2 \times 10^{11}$ $M_\odot$, the dark mass of satellites decreased by $60 - 65\%$ with respect to their masses at accretion time. We find that MRs for both definitions of subhalo mass yield satellite conditional mass functions (CSMF) in agreement with observations. Also, when these MRs are used in a halo occupation model, the predicted two–point correlation functions at different stellar mass bins agree with observations. The average stellar–halo MR is close to the MR of central galaxies alone, and conceptually this average MR is equivalent to abundance matching the cumulative total GSMF to the halo + subhalo mass function (the standard AMT). We show that the use of $m_{\rm sub}^{\rm obs}$ leads to less uncertain MRs than $m_{\rm sub}^{\rm acc}$, and discuss some implications of the obtained satellite–subhalo MR. For example, we show that the tension between abundance and dynamics of Milky-Way satellites in the Λ-CDM cosmogony disspaers if the faint-end slope of the GSMF upturns to a value of ~ -1.6.

2.1 INTRODUCTION

In recent years the abundance matching technique (AMT) has emerged as a simple yet powerful statistical approach for connecting galaxies to halos without requiring knowledge of the underlying physics [e.g., 216, 115, 58, 187, 16, 57, 75, 19, and references therein].

Briefly, the AMT assumes a one-to-one monotonic relationship between stellar and halo masses which can be constrained by matching the cumulative observed galaxy stel-

lar mass function (GSMF) to the theoretical halo plus subhalo cumulative mass function. Interestingly enough, this simple approach successfully reproduces the observed spatial clustering of galaxies [e.g., 58, 148]. The AMT allows to probe the average galaxy stellar–halo mass relation, $M_*(M_\mathrm{h})$ (hereafter SHMR), delivering very useful information for constraining models of galaxy evolution [e.g., 102, 85, 6].

The above has motivated several authors to use the AMT extensively. For example, with the advent of large galaxy surveys at different redshifts, the AMT has been applied for constraining the evolution of the *average* SHMR [e.g., 58, 75, 57, 148, 19]. As a natural extension, these studies have been combined with predicted average halo mass aggregation histories in order to infer *average* galaxy M_* growth histories as a function of mass [57, 84, see for a review Avila-Reese & Firmani [9], and references therein]. By including observational information on the gas content of galaxies, the AMT has been also used to constrain the baryon mass to M_h relation of galaxies [16, 177]. Finally, variants of the AMT, where instead of mass functions, circular velocity functions or functions of any other galaxy/halo global property are employed, have been explored, too [e.g., 58, 31, 214].

The AMT has been commonly applied to the total (central plus satellite galaxies) GSMF matched against the total (distinct plus satellite) halo population. This approach has been criticized, because quite different average SHMRs are obtained for different proposed forms of the satellite stellar-subhalo mass relation (SSMR, $m_*(m_\mathrm{sub})$) and the central SHMR ($M_*(M_\mathrm{h})$[1]; 152).

A common (questionable) assumption is that the SSMR is identical to the central SHMR. Under this assumption, it is also common to define subhalo mass at the time of *accretion* ($m_\mathrm{sub}^\mathrm{acc}$) rather than at the time of *observation* ($m_\mathrm{sub}^\mathrm{obs}$), when subhalos have lost a significant fraction of mass due to tidal stripping. The use of $m_\mathrm{sub}^\mathrm{acc}$ has been justified because the question of subhalo mass loss is avoided in this way, and regarding the satellite m_*, it is expected that it remains almost constant since its infall into the host halo. The projected two-point correlation function of galaxies is reproduced under these assumptions [58, 148]. It should also be said that while the (local) SHMR for central galaxies has been determined [e.g., 133, 146], the stellar–subhalo mass relation for satellites/subhalos, SSMR, has not been yet discussed in detail in the literature.

In view of the above, some important questions arise. Why does using $m_\mathrm{sub}^\mathrm{acc}$ instead of $m_\mathrm{sub}^\mathrm{obs}$ lead to the correct clustering of galaxies? Does the $m_\mathrm{sub}^\mathrm{acc}$–$M_*$ relation reproduce the observed satellite GSMF, the conditional stellar mass function, and spatial clustering of galaxies at the same time? Even more fundamentally, if is not assumed that the SSMR is identical to the central SHMR, then, what follows for the SSMR, either using $m_\mathrm{sub}^\mathrm{obs}$ or $m_\mathrm{sub}^\mathrm{acc}$? Does it deviate from the central SHMR?

In this paper we extend the common AMT to constrain both the central SHMR and the SSMR separately, as well as the average (total) SHMR. By construction, this formalism also allows to predict the mean satellite conditional mass function (CSMF), i.e., the probability that satellites of a given stellar mass reside in distinct host halos of a given mass. We will (i) test whether the SSMR and the central SHMR have the same shape; (ii) discuss the consequences of defining the subhalo mass at accretion time vs. at observed

1 In order to make the distinction explicit, we shall use upper-case letters for the central galaxy and the distinct halo masses and lower-case letters for the satellite galaxy and subhalo masses.

(present) time; and (iii) check the self-consistency of our predicted present-day central SHMR and SSMR by comparison with the observed satellite CSMF and the spatial clustering of galaxies.

This Chapter is laid out as follows. In Section 2.2 we present the AMT, focusing on the details of our extended abundance matching. In Section 2.3 we present the predicted stellar-halo mass relations (2.3.1) and satellite CSMFs (2.3.2) for cases when the SSMR is assumed equal to the central SHMR, and when both mass relations are independently constrained. In 2.3.3, a Halo Occupation Distribution (HOD) model is used to explore whether the predicted central SHMR's and SSMR's are consistent with the observed spatial clustering of galaxies. Section 2.4 is devoted to our conclusions and a discussion of the results and their implications.

All our calculations are based on a flat ΛCDM cosmology with $\Omega_\Lambda = 0.73$, $h = 0.7$, and $\sigma_8 = 0.84$, close to those reported by WMAP 7.

2.2 THE ABUNDANCE MATCHING TECHNIQUE

In this section we describe the technique of matching abundances between central galaxies and halos and satellite galaxies and subhalos, separetely, which we present here as an extension to the standard AMT.

2.2.1 *Modeling the central & satellite GSMFs*

To model the central GSMF, let $P_{\rm cen}(M_*|M_{\rm h})$ denote the probability distribution function that a distinct halo of mass $M_{\rm h}$ hosts a central galaxy of stellar mass M_*. Then the number density of central galaxies with stellar masses between M_* and $M_* + dM_*$ is given by

$$\phi_{\rm cen}(M_*)dM_* = dM_* \int_0^\infty P_{\rm cen}(M_*|M_{\rm h})\phi_{\rm h}(M_{\rm h})dM_{\rm h}. \quad (1)$$

For the population of satellite galaxies in individual subhalos, let $P_{\rm sat}(m_*|m_{\rm sub})$ be the probability distribution function that a subhalo $m_{\rm sub}$[2] hosts a satellite galaxy of stellar mass m_*. Thus the average satellite CSMF (the number of satellite galaxies of stellar mass between m_* and $m_* + dm_*$ that reside in distinct host halos of mass $M_{\rm h}$, e.g., 248) is

$$\Phi_s(m_*|M_{\rm h})dm_* = dm_* \int_0^\infty P_{\rm sat}(m_*|m_{\rm sub})\Phi_{\rm sub}(m_{\rm sub}|M_{\rm h})dm_{\rm sub}, \quad (2)$$

where $\Phi_{\rm sub}(m_{\rm sub}|M_{\rm h})$ is the subhalo conditional mass function [subhCMF, i.e., the number of subhalos of mass between $m_{\rm sub}$ and $m_{\rm sub} + dm_{\rm sub}$ residing in host halos of mass $M_{\rm h}$; e.g., 41]. A natural link between the satellite GSMF, $\phi_{\rm sat}$, and the distinct halo mass function (HMF, $\phi_{\rm h}$) arises once the satellite CSMF is given:

$$\phi_{\rm sat}(m_*)dm_* = dm_* \int_0^\infty \Phi_s(m_*|M_{\rm h})\phi_{\rm h}(M_{\rm h})dM_{\rm h}. \quad (3)$$

2 Whenever we use $m_{\rm sub}$ we refer to subhalo mass generically. In practice, that can either be the mass at accretion time, $m_{\rm sub}^{\rm acc}$, or at observation (present-day) time, $m_{\rm sub}^{\rm obs}$.

Inserting equation (2) into equation (3) and rearranging terms, the satellite GSMF can be rewritten in terms of $P_{\rm sat}(m_*|m_{\rm sub})$:

$$\phi_{\rm sat}(m_*)dm_* = dm_* \int_0^\infty P_{\rm sat}(m_*|m_{\rm sub})\phi_{\rm sub}(m_{\rm sub})dm_{\rm sub}, \tag{4}$$

where the subhalo mass function (subHMF) is given by

$$\phi_{\rm sub}(m_{\rm sub})dm_{\rm sub} = dm_{\rm sub} \int_0^\infty \Phi_{\rm sub}(m_{\rm sub}|M_{\rm h})\phi_{\rm h}(M_{\rm h})dM_{\rm h}. \tag{5}$$

Equations (1) and (4) describe the abundance matching in its differential form for the central-halo and satellite-subhalo populations, respectively. The distribution probability $P_{\rm cen}(M_*|M_{\rm h})$ is defined by the mean $M_*(M_{\rm h})$ relation and a scatter around it of σ_c, while the distribution probability $P_{\rm sat}(m_*|m_{\rm sub})$, assumed to be independent of host halo mass (as is commonly adopted in the AMT), is defined by the mean $m_*(m_{\rm sub})$ relation and a scatter around it of σ_s. Observe that once $P_{\rm sat}(m_*|m_{\rm sub})$ is given, the satellite CSMF is a prediction according to equation (2).

Here, $P_{\rm cen}(M_*|M_{\rm h})$ and $P_{\rm sat}(m_*|m_{\rm sub})$ are modeled as lognormal distributions with a width (scatter around the stellar mass) assumed to be constant and the same for both centrals and satellites, $\sigma_c = \sigma_s = 0.173$ dex. Such a value was inferred for central galaxies from the analysis of general large group catalogs (YMB09) and it is supported by recent studies on the kinematics of satellite galaxies [146]. Regarding the intrinsic scatter of the satellite-subhalo relation, it has not been discussed in detail in the literature. While the exploration of this scatter is beyond the scope of the present paper, our conclusions will not depend critically on the assumed value for it or even if it is allowed to depend on host halo mass. We will further discuss this question in Section 2.4.2. Both $m_*(m_{\rm sub})$ and $M_*(M_{\rm h})$ are parametrized by the same modified two-power-law form proposed in Behroozi, Conroy & Wechsler [19]. This five-parameters function is quite general and, in the context of the AMT, has been shown to reproduce the main features of a Schechter-like GSMF.

2.2.2 *The relation to standard abundance matching*

In the standard AMT the cumulative halo+subhalo mass function and the total observed cumulative GSMF are matched to determine the mass relation between halos and galaxies, which is assumed to be monotonic. In this context, no intrinsic scatter in the stellar mass at a given halo is assumed. In our approach, where the galaxy and halo populations are separated into centrals/satellites and distinct halo/subhalos, the latter entails that the probability distribution functions of centrals and satellites take the particular forms: $P_{\rm cen}(\mathcal{M}|M_{\rm h}) = \delta(\mathcal{M} - M_*(M_{\rm h}))$ and $P_{\rm sat}(\mathcal{M}|m_{\rm sub}) = \delta(\mathcal{M} - m_*(m_{\rm sub}))$, where $M_*(M_{\rm h})$ and $m_*(m_{\rm sub})$ are the mean central-halo and satellite-subhalo mass relations, and δ is for the δ-Dirac function. The above "no scatter" probability distribution function for centrals applied in Eq. (1) would lead us to conclude that the cumulative central GSMF, $n_{\rm cen}(> M_*)$, should match the cumulative distinct halo mass function, $n_{\rm h}(> M_{\rm h}(M_*))$. The same reasoning applies for satellites/subhalos. Therefore, we arrive to the standard AMT formulation:

$$n_g(> M_*) = n_h(> M_{\rm h}) + n_{\rm sub}(> M_{\rm h}), \tag{6}$$

where $n_g(>M_*) \equiv n_{\rm cen}(>M_*) + n_{\rm sat}(>M_*)$ is the total GSMF.

Since the abundance matching can be applied to centrals/halos and satellites/subhalos separately, let us analyze now only the latter. Under the assumption that the $m_*(m_{\rm sub})$ relation is independent of the host halo mass, it is clear that using either the abundance matching of all satellites and all subhalos, $n_{\rm sat}(>m_*) = n_{\rm sub}(>m_{\rm sub})$, or the matching of their corresponding mean occupational numbers, one may find exactly the same $m_*(m_{\rm sub})$ relation. In this sense, we state that matching abundances is equivalent to matching occupational numbers:

$$\langle N_s(>m_*|M_{\rm h})\rangle = \langle N_{\rm sub}(>m_{\rm sub}|M_{\rm h})\rangle \Longleftrightarrow n_{\rm sat}(>m_*) = n_{\rm sub}(>m_{\rm sub}). \tag{7}$$

In the case that the probability distribution function $P_{\rm sat}(m_*|m_{\rm sub})$ includes scatter around the mean SSMR, as we consider here, the above conclusion remains the same as long as $P_{\rm sat}(m_*|m_{\rm sub})$ is assumed to be independent on halo mass. In general, the inclusion of constant scatter in the galaxy-halo mass relations is not a conceptual problem for the AMT, but it slightly modifies the shape of the mass relations at the high mass end [see 19]. Finally, note that if $P_{\rm sat}(m_*|m_{\rm sub})$ depends on $M_{\rm h}$, then $\phi_{\rm sat}$ may not be directly related to $\phi_{\rm sub}$ (see Eq. 8) and using either the matching of satellites and subhalo abundances or the matching of their corresponding occupational numbers would not lead to find exactly the same $m_*(m_{\rm sub})$.

2.2.3 *Inputs for matching abundances*

The inputs required for the procedure described above are the subhCMF, the distinct HMF, and the observed satellite and central GSMFs.

For the subhCMF, we use the results obtained in Boylan-Kolchin et al. [41, BK10] based on the analysis of the Millennium-II Simulation. This simulation assumes the same cosmological parameters and the same particle number as the Millennium Simulation, but in a smaller volume increasing up to 125 times the mass resolution. It consists of 2160^3 particles, each of mass $m = 6.885 \times 10^6 h^{-1}{\rm M}_\odot$ in a periodic cube of length $L = 100 h^{-1}$Mpc. Observe that this mass particle resolution is around four orders of magnitude below the subhalo masses required ($\sim 10^{10} h^{-1}{\rm M}_\odot$) to match the lower stellar mass limit in the YMB09 GSMF. The fitting formula for the cumulative subhCMF reported in BK10 at the $[10^{12}, 10^{12.5}]h^{-1}{\rm M}_\odot$ mass interval is:

$$\langle N_{\rm sub}(>m_{\rm sub}|M_{\rm h})\rangle = \mu_0 \left(\frac{\mu}{\mu_1}\right)^a \exp\left[-\left(\frac{\mu}{\mu_{\rm cut}}\right)^b\right], \tag{8}$$

where $\mu = m_{\rm sub}/M_{\rm h}$. For $m_{\rm sub} = m_{\rm sub}^{\rm acc}$, $(\mu_0, \mu_1, \mu_{\rm cut}, a, b,) = (1, 0.038, 0.225, -0.935, 0.75)$, while for $m_{\rm sub} = m_{\rm sub}^{\rm obs}$, $= (1.15^{(\log M_{\rm h} - 12.25)}, 0.01, 0.096, -0.935, 1.29)$. According to BK10, the shape of the $m_{\rm sub}^{\rm obs}$ subhCMF remains the same for other halo masses but its normalization, μ_0, systematically increases with $M_{\rm h}$, roughly by 15% per dex in $M_{\rm h}$. Such a behavior has been reported in an analysis of the Millennium simulations by Gao et al. [90, for closely related results, with small differences in the amplitude, see also 91, 220, 260, 3, 93]. We introduce the dependence $\mu_0 = 1.15^{(\log M_{\rm h} - 12.25)}$, where $\mu_0 = 1$ at $\log M_{\rm h} = 12.25$. In the case of the $m_{\rm sub}^{\rm acc}$ subhCMF, the normalization factor is nearly

independent of $M_{\rm h}$, i.e., $\mu_0 = 1$ (BK10; see also 94). The subhCMF is given by $\Phi_{\rm sub} = dN_{\rm sub}/dm_{\rm sub}$.

In order to construct the $m_{\rm sub}^{\rm acc}$ subhCMF, BK10 traced each subhalo back in time so that they were able to find the point at which its bound mass reached a maximum, i.e., the time the halo became a subhalo. The latter guarantees that we are working with the surviving population of accreted halos and no further assumptions on the merging process are necessary.

The difference between the Millennium-II simulation cosmology and ours leads to differences in the resulting abundances of subhalos of roughly a few percent in the amplitude of the subhalo mass function (BK10). This is also supported by previous works that explored the impact of changing cosmological parameters on the subhalo occupational statistics (e.g., 261). Additionally, to be consistent with the same cosmology for which the subhalo subhCMF was inferred, we repeated all the analysis to be showed below but using the WMAP1 cosmology. We find that all our results are practically the same.

For the distinct HMF, we will use the formula given by Sheth & Tormen [192]. This formula provides a reasonable fit to the the virial mass[3] function at $z \sim 0$ measured in large cosmological N-body simulations [e.g., 113, 64].

For our purpose, the decomposition of the GSMF and the CSMFs into centrals and satellites galaxies is necessary. Using a large general group catalog [249] based on the data from the SDSS, YMB09 constructed and studied the decomposition of the GSMF and the CSMFs into centrals and satellites galaxies. In that paper, a central galaxy was defined as the most massive galaxy in a group and the remaining galaxies as satellites. For the mass completeness limit in the GSMF, they adopted the value as function of redshift proposed in van den Bosch et al. [219]. They have also taken into account incompleteness in the group mass by considering an empirical halo-mass completeness limit (for details we refer the reader to YMB09).

Where necessary, halo masses are converted to match our definition of virial mass, and stellar masses are converted to the Chabrier [48] IMF. In particular, YMB09 defined halo masses with the average density 180 times the background density, according to Giocoli et al. [93, see their appendix B] these halos are $\sim 11\%$ larger than our definition of virial mass.

2.2.4 *Procedure and uncertainties*

We constrain the parameters of the functions proposed to describe the central SHMR and SSMR by means of Eqs. (1) and (3), and by using the Powell's directions set method in multi-dimensions for the minimization [167]. Note that in our analysis the reported statistical errors in the GSMFs, as well as the intrinsic scatter in the mass relations are taken into account. However, we will not analyze rigorously here the effects of uncertainties on the mass relations as well as their errors. Instead, we remit the reader to previous works [148, 19, 177].

3 The mass enclosed within the radius at which, according to the spherical collapse model, the overdensity of a sphere is $\Delta_{\rm vir}$ times larger than the *matter* critical density of the used cosmological model; for the cosmology assumed here, $\Delta_{\rm vir}(z=0) = 97$.

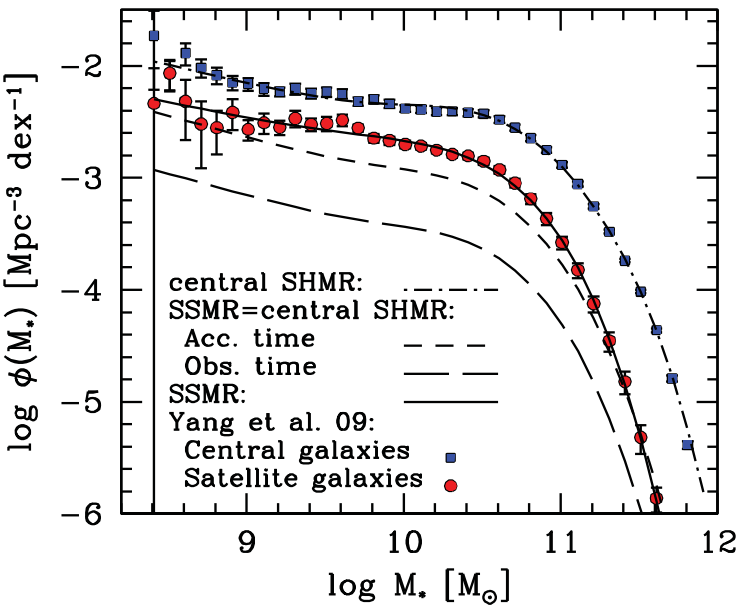

Figure 3: Satellite GSMFs calculated under the assumption that $P_{\rm sat}(m_*|m_{\rm sub}) = P_{\rm cen}(M_*|M_{\rm h})$ and for the cases $m_{\rm sub}^{\rm acc}$ (long-dashed line) and $m_{\rm sub}^{\rm obs}$ (short-dashed line) were used for the subhalo mass definition. Filled circles and squares with error bars show the YMB09 central and satellite GSMFs, respectively. The solid line is for the case when $P_{\rm sat}(m_*|m_{\rm sub})$ was determined using the YMB09 satellite GSMF as a constraint, i.e., is the best model fit to this function. The dotted dashed lines shows the central SHMR.

Behroozi, Conroy & Wechsler [19] studied in detail the uncertainties and effects on the average SHMR due to different sources of error like those in the observed GSMFs, including stellar mass estimates; in the halo mass functions; in the uncertainty of the cosmological parameters; and in linking galaxies to halos, including the intrinsic scatter in this connection. These authors have found that the largest uncertainty by far in the SHMR is due to the systematic shifts in the stellar estimates. The second important source of uncertainty is due to the intrinsic scatter, that we take into account in our analysis. Other statistical and sample variance errors have negligible effects, at least for local galaxies. According to the Behroozi, Conroy & Wechsler [19] study, the statistical and systematical uncertainties account for 1σ errors in the SHMR of approximately 0.25 dex at all masses, which is almost totally due to the uncertainty in stellar mass estimates. We have explored here also the effects of the subhalo CMF uncertainty on the SSMR. By using the 25% per dex in $M_{\rm h}$ variation reported by Giocoli et al. [93] (instead of 15%), we find that the SSMR shifts in m_* by only ≈ 0.04 dex.

2.3 RESULTS

2.3.1 *The satellite GSMF and the SSMR*

By means of the procedure described in Section 2, we calculate first the satellite GSMF (Fig. 3) when the SSMR and the central SHMR are assumed to be the same, i.e., $m_*(m_{\rm sub}) =$

$M_*(M_h)$. This is equivalent to assume that $P_{sat}(m_*|m_{sub}) = P_{cen}(M_*|M_h)$ if the intrinsic scatter of both relations is the same. We obtain the central SHMR by matching abundances of YMB09 central galaxies to distinct halos. This relation and the subhalo mass function obtained from the theoretical subhCMF (eq. 8), are used to infer the satellite GSMF (eq. 4). The satellite GSMF is presented for the two cases of subhalo mass definition: $\phi_{sat,acc}$ when m_{sub}^{acc} is used (long-dashed line), and $\phi_{sat,obs}$ when m_{sub}^{obs} is used (short-dashed line). The observational results of YMB09 are plotted as well.

Under the assumption that $m_*(m_{sub}) = M_*(M_h)$, the predicted number density of satellites at masses below the knee is underestimated on the average by a factor of ~ 2 when using m_{sub}^{acc}, and ~ 5 when using m_{sub}^{obs}. Note that the former is closer to the YMB09 data. The reason is simply because the normalization of the m_{sub}^{acc} subHMF is higher and closer to the distinct HMF than the normalization of the m_{sub}^{obs} subHMF. Therefore, satellites of equal m_* are expected to have a higher number density when using the accreted-time (m_{sub}^{acc}) subHMF compared to using the observed-time (m_{sub}^{obs}; present-day) subHMF.

However, neither m_{sub}^{acc} nor m_{sub}^{obs} are able to reproduce the observed satellite GSMF, and the discrepancy is due to the basic assumption of a common stellar mass–(sub)halo mass relation for centrals and satellites. In the case that $m_{sub} = m_{sub}^{acc}$, this is equivalent to assume that the SSMR is independent of redshift. But in fact this cannot be the case since the satellite mass m_* hardly will remain the same since it was accreted to the present epoch. On the other hand, when using $m_{sub} = m_{sub}^{obs}$, that the SSMR is equal to the central SHMR implies that both have evolved, on average, identically. This cannot be the case because it is evident that the population of subhalos evolved differently to distinct halos, mainly by losing mass due to tidal striping [e.g., 115, 220].

The next step in our analysis is to allow the SSMR and central SHMR to be different, i.e., $m_*(m_{sub}) \neq M_*(M_h)$. In this case, $P_{sat}(m_*|m_{sub})$ is determined by means of Eq. (4) using the YMB09 satellite GSMF as a constraint. For the subhCMF, we again use both definitions of subhalo mass, m_{sub}^{acc} and m_{sub}^{obs}. For illustrative purpose, we present the resulting satellite GSMF for the case when m_{sub}^{obs} was used (solid line in Fig 3; an almost identical GSMF is obtained when m_{sub}^{acc} was used).

As shown in Fig. 4, the SSMRs obtained by using m_{sub}^{acc} (long-dashed line) and m_{sub}^{obs} (solid line) are quite different. The central SHMR (dot-dashed line) is the same for both cases. The error bar in the left panel shows a 1σ uncertainty of 0.25 dex in the normalization of the mass relations. This is roughly the uncertainty estimated by Behroozi, Conroy & Wechsler [19] taking into account all the systematical and statistical sources of errors (see 2.2.4).

When using the accretion-time subhalo mass, m_{sub}^{acc}, we note that the resulting SSMR at $\log(m_*/M_\odot) < 11$ systematically lies above the central SHMR, with differences in the stellar-mass axis (halo-mass axis) of $\sim$ 0.5 dex (0.2 dex) at the smallest masses. For $\log(M_*/M_\odot) > 11$ this trend is inverted, but the differences between central and satellites are very small. However, the relation obtained this way should be taken with caution. By construction, each m_{sub}^{acc} is itself a cumulative distribution of all the objects accreted in a time interval Δz. Therefore such a SSMR entails that all accreted objects of mass m_{sub}^{acc} would evolve, on average, to host the same m_* despite having been accreted at *different times*. We discuss this in 2.4.1.

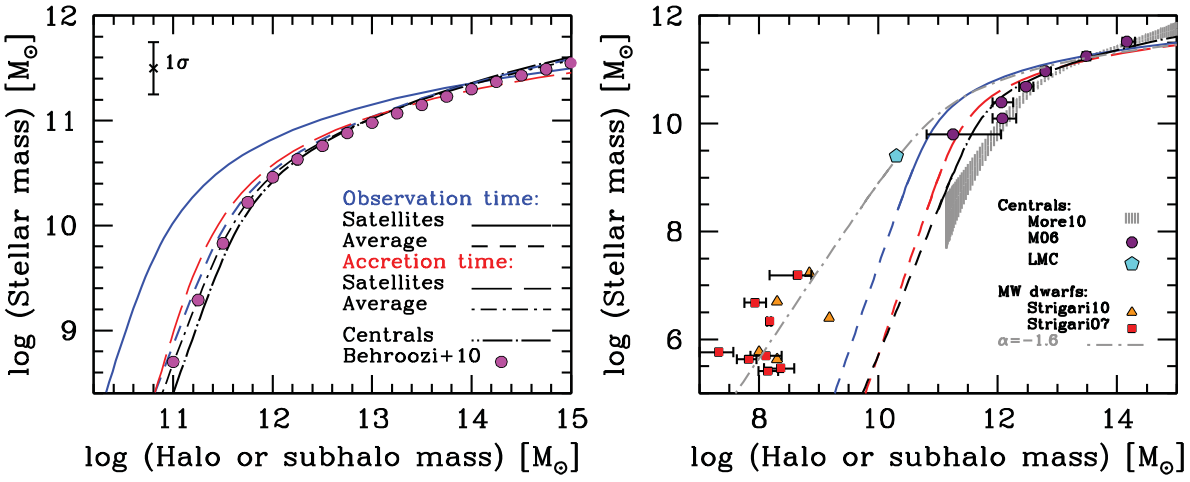

Figure 4: *Left panel:* Inferred mass relations for satellite galaxies when the subhalo mass is defined as $m_{\rm sub}^{\rm obs}$ (blue solid line) and as $m_{\rm sub}^{\rm acc}$ (red long-dashed-line), and for central galaxies (black dot-long-dashed line). The density-weighted average relation when using subhalo mass $m_{\rm sub}^{\rm obs}$ is plotted with a blue short-dashed line, while when using subhalo mass $m_{\rm sub}^{\rm acc}$ is plotted with a black short-dotted line. For comparison, the (average) mass relation obtained in Behroozi, Conroy & Wechsler [19] is also plotted (filled circles). *Right panel:* A comparison of the mass relations of satellite and central galaxies with direct observational inferences (the same line code of left panel for models is used; dashed lines indicate extrapolations to lower masses). Filled circles with error bars correspond to the mass relation of central galaxies from the analysis of staked weak-lensing in Mandelbaum et al. [133]. Dashed area indicates the 68% of confidence in the mass relation of central galaxies using the kinematics of satellites [146]. The inferred total mass at the tidal radii for the brightest dwarf galaxies obtained in Strigari et al. [200] and Strigari, Frenk & White [201] are plotted with filled squares and triangles, respectively. Filled pentagon shows the mass at the tidal radius for the Large-Magallanic Cloud [235]. The gray dotted-dashed curve is the $m_{\rm sub}^{\rm obs}$ SSMR assuming a faint-end slope in the satellite GSMF of $\alpha = -1.6$.

When using the observation-time (present-day) subhalo mass, $m_{\rm sub}^{\rm obs}$, the SSMR (solid blue line) and central SHMR are very different, though they show the same trend as when using $m_{\rm sub}^{\rm acc}$. For example, on average, a satellite with $\log(M_*/{\rm M}_\odot) = 10$ resides in a subhalo a factor of ~ 4 less massive than the halo of a central galaxy with the same stellar mass. Notice that $m_{\rm sub}^{\rm obs}(m_*) < m_{\rm sub}^{\rm acc}(m_*)$ and that the difference increases the lower the mass is. This is consistent with the picture that most massive subhalos, on average, fell into larger halos just very recently and they have not had time to lose significant amounts of mass due to tidal striping, in contrast to the lowest mass subhalos.

This also suggests that the SSMR for both definitions of subhalo mass should tend to the central SHMR at the high-mass end, but this is not the case as seen in Fig. 4 where small differences remain. The possible reasons are that, firstly, the intrinsic scatter around the stellar–(sub)halo mass relations is actually lower for the former than for the latter (here we assumed it to be the same for satellites and centrals, see 2.4.2). Secondly, that the YMB09 satellite GSMF may underestimate the true satellite mass function at large masses [see also 194].

Fiber collisions could introduce some systematic error that may affect the YMB09 group catalog. To study the impact of this possible systematic error, YMB09 divided their group catalog into two sample: one that uses galaxies with known redshifts, and another that includes galaxies that lack redshifts due to fiber collisions. When compared the corresponding satellite CSMFs from both samples (see their Fig. 6), they found that the sample for which the correction for fiber collisions has been taking into account, has a higher amplitude of the CSMFs than when this correction has not been applied, specially in low mass halos. However, the difference is very marginal and well within the error bars. We conclude that fiber collisions in the YMB09 group catalog are not a serious source of systematics and should not to affect our conclusions. Regarding completeness and contamination of their group catalog [for details see 249], 80% have a completeness greater than 0.6, while 85% have a contamination lower than 0.5. In terms of purity, their halo-based group finder is consistent with the ideal situation.

Finally, we note that the mass relation usually obtained by matching abundances between the *total* GSMF and the halo plus subhalo mass function, in the light of the decomposition into centrals and satellites, could be interpreted as a *density–weighted average* SHMR:

$$\langle M_*(M)\rangle_\phi = \frac{\phi_{\rm sub}(M)}{\phi_{\rm DM}(M)} m_*(M) + \frac{\phi_{\rm h}(M)}{\phi_{\rm DM}(M)} M_*(M), \tag{9}$$

where $\phi_{\rm DM}(M) = \phi_{\rm sub}(M) + \phi_{\rm h}(M)$, $m_*(M)$ is the mean SSMR and $M_*(M)$ is the mean central SHMR. This relation is plotted in Fig. 4 with short-dashed-dot and short-dashed lines when using $m_{\rm sub}^{\rm acc}$ and $m_{\rm sub}^{\rm obs}$, respectively. Since most galaxies in the YMB09 catalog are centrals, the central SHMR is very close to the density-weighted average SHMR. For comparison, we plotted the Behroozi, Conroy & Wechsler [19] average mass relation (filled circles), which is in excellent agreement with our density-weighted average SHMR when using the accreted-time subhalo mass, $m_{\rm sub}^{\rm acc}$.

Observe that differences between the satellite and the average (total) mass relations are small when $m_{\rm sub}^{\rm acc}$ is used, while differences become dramatic when $m_{\rm sub}^{\rm obs}$ is used. The above explains why under the assumption that $m_*(m_{\rm sub}^{\rm acc}) = M_*(M_{\rm h}) = \langle M_*(M)\rangle_\phi$, the resulting satellite GSMF are closer to observations. On the other hand, since the $m_{\rm sub}^{\rm acc}$ subHMF has a higher normalization than the $m_{\rm sub}^{\rm obs}$ subHMF, the above shows that when assuming $m_*(m_{\rm sub}^{\rm obs}) = M_*(M_{\rm h}) = \langle M_*(M)\rangle_\phi$, we should expect that the resulting satellite GSMF is significantly below the observed satellite GSMF.

2.3.1.1 *Comparison with other observational inferences*

In the right panel of Fig. 4, we plot some observational inferences of halo and subhalo masses as a function of stellar mass. The inferred $\langle M_{\rm h}\rangle(M_*)$ of central galaxies from staked weak-lensing studies using the SDSS [133] are shown as filled circles with error bars. Mandelbaum et al. [133] reported the data actually for blue and red galaxies separately. We estimated the average mass relation for central galaxies as: $\langle M_{\rm h}\rangle(M_*) = f_b(M_*)\langle M_{\rm h}\rangle_b(M_*) + f_r(M_*)\langle M_{\rm h}\rangle_r(M_*)$, where $f_b(M_*)$ and $f_r(M_*)$ are the blue and red galaxy fractions in the sample, and $\langle M_{\rm h}\rangle_b$ and $\langle M_{\rm h}\rangle_r$ are the corresponding blue and red mass relations. The inferred $\langle \log(M_*)\rangle(M_{\rm h})$ for central galaxies from

staked kinematics of satellites [146] are plotted as the dashed area indicating the 68% of confidence.

Our inferred central SHMR (dotted-dashed curve) is consistent with the weak-lensing inferences at all masses, and with the satellite kinematics inferences at masses $M_* \gtrsim 10^{11}$ $M_\odot$; for smaller masses, our halo masses are a factor up to ~ 2 smaller than the satellite kinematics inferences. In fact, it was already noted that using the kinematics of satellite galaxies yields halo masses around low mass galaxies that are systematically larger than most other methods, specially for red central galaxies [146, 194, 177].

Regarding satellites, unfortunately, there are not direct inferences of their subhalo masses. Some model-dependent estimates based on dynamical observations of Milky-Way (MW) satellites were presented in the literature. For example, using the line-of-sight velocity dispersions measured for the brightest spheroidal dwarf galaxies, Strigari et al. [200] and Strigari, Frenk & White [201] determined their masses within their tidal radii. These dynamical masses, plotted in Fig. 4 (filled squares and triangles, respectively), are expected to be of the order of $m_{\rm sub}^{\rm obs}$. We also plot an estimate of the mass at the tidal radius for the Large-Magallanic Cloud [filled pentagon, 235]. The SSMRs constrained here do not extend to the small masses of MW satellites but we plot their extrapolations to these masses (dashed curves). The gray dotted-dashed curve will be discussed in section 2.4.3.

2.3.2 *The satellite CSMF*

From the approach described in Section 2, another statistical quantity that deserves to be subject of study is the satellite CSMF (Eq. 2). We calculate the mean halo–density–weighted CSMF at the $[M_{\rm h_1}, M_{\rm h_2}]$ bin as:

$$\langle \Phi_s \rangle = \frac{\int_{M_{\rm h_1}}^{M_{\rm h_2}} \Phi_s(m_*|M_{\rm h})\phi_{\rm h}(M_{\rm h})dM_{\rm h}}{\int_{M_{\rm h_1}}^{M_{\rm h_2}} \phi_{\rm h}(M_{\rm h})dM_{\rm h}}. \tag{10}$$

This quantitiy has been inferred from observations by YMB09, again using their SDSS galaxy catalog (filled circles with error bars in Fig. 5).

First, we consider again the case assuming $m_*(m_{\rm sub}) = M_*(M_{\rm h})$. When $m_{\rm sub}$ is defined at the observation time, the resulting CSMFs are lower than the YMB09 CSMFs by a factor of ~ 5 in the power-law regime (roughly the same factor by which $\phi_{\rm sat,obs}$ is lower than the YMB09 observed satellite GSMF). Similarly, when $m_{\rm sub}$ is defined at the accretion time, the predicted CSMFs in the power-law regime are below the YMB09 CSMFs by nearly the same factor, ~ 2, that $\phi_{\rm sat,acc}$ lies below the satellite GSMF. The normalization of the CSMF increases faster with $M_{\rm h}$ when $m_{\rm sub}^{\rm obs}$ is used instead of $m_{\rm sub}^{\rm acc}$. This is because we allow the $m_{\rm sub}^{\rm obs}$ subhCMF normalization to vary with host halo mass, while the $m_{\rm sub}^{\rm acc}$ subhCMF normalization is independent of host halo mass.

The black continuous ($m_{\rm sub}^{\rm obs}$) and blue long-dashed ($m_{\rm sub}^{\rm acc}$) lines in Fig. 5 (almost indistinguishable one from other) are the predictions when $P_{\rm sat}(m_*|m_{\rm sub})$ has been constrained by means of the observed satellite GSMF. The agreement of the predicted satellite CSMF's, for both $m_{\rm sub}^{\rm obs}$ and $m_{\rm sub}^{\rm acc}$ with the YMB09 CSMF's is now remarkable at all

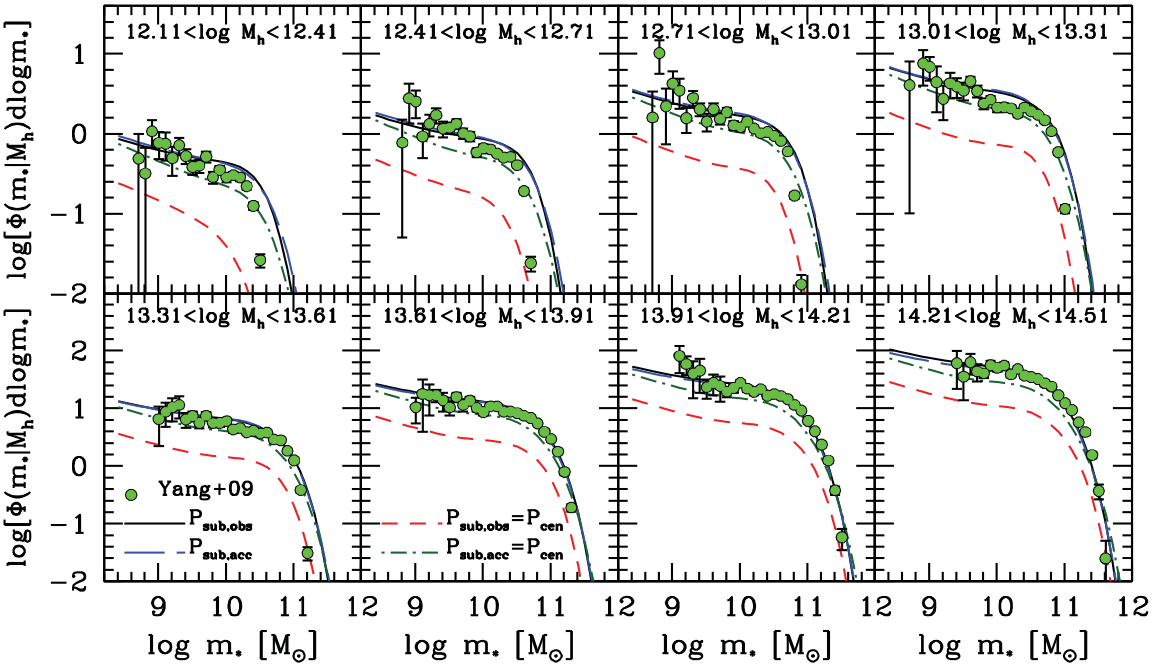

Figure 5: Density–weighted average satellite CSMF in eight halo mass intervals. Red short-dashed and green short-dashed-dot lines are for the cases when the central SHMR and the SSMR were assumed to be equal and $m_{\rm sub}^{\rm obs}$ and $m_{\rm sub}^{\rm acc}$ were used, respectively. The black solid and blue long-dashed lines are again for $m_{\rm sub}^{\rm obs}$ and $m_{\rm sub}^{\rm acc}$, respectively, but in the case the central and satellite mass functions were independently constrained by means of our extended AMT (they overlap most of time). Filled circles with error bars show the CSMFs inferred from observations by YMB09. Note that their halo masses were converted to match our virial definition.

halo mass bins for low/intermedium stellar masses. Although, as above, the normalization of the CSMF's increases faster when $m_{\rm sub} = m_{\rm sub}^{\rm obs}$ than when $m_{\rm sub} = m_{\rm sub}^{\rm acc}$, the differences between both cases at any mass are less than 0.05 dex, within the error bars of the observational data.

Despite the overall agreement, for halo mass bins lower than $\sim 10^{13}$ $M_\odot$, the number of massive satellite galaxies is overestimated, specially at the lowest $M_{\rm h}$ bins. A possible reason for this is the assumption that the scatter in $P_{\rm sat}(m_*|m_{\rm sub})$ is constant while in reality it could depend on $M_{\rm h}$ as well as on $m_{\rm sub}$. However, the probability of finding massive satellite galaxies in halos less massive than $\sim 10^{13}$ $M_\odot$ is low and they do not contribute significantly to the mean total density of satellite galaxies. Therefore, this assumption does not change our conclusions, see also 2.4.

Our analysis shows that *assuming* $P_{\rm sat}(m_*|m_{\rm sub}) = P_{\rm cen}(M_*|M_{\rm h})$ *the resulting satellite CSMFs are not consistent with observations*. Instead, when $P_{\rm sat}(m_*|m_{\rm sub})$ is independently constrained using the observed satellite GSMF, there is a clear agreement, no matter what definition of $m_{\rm sub}$ was employed for the subhCMF.

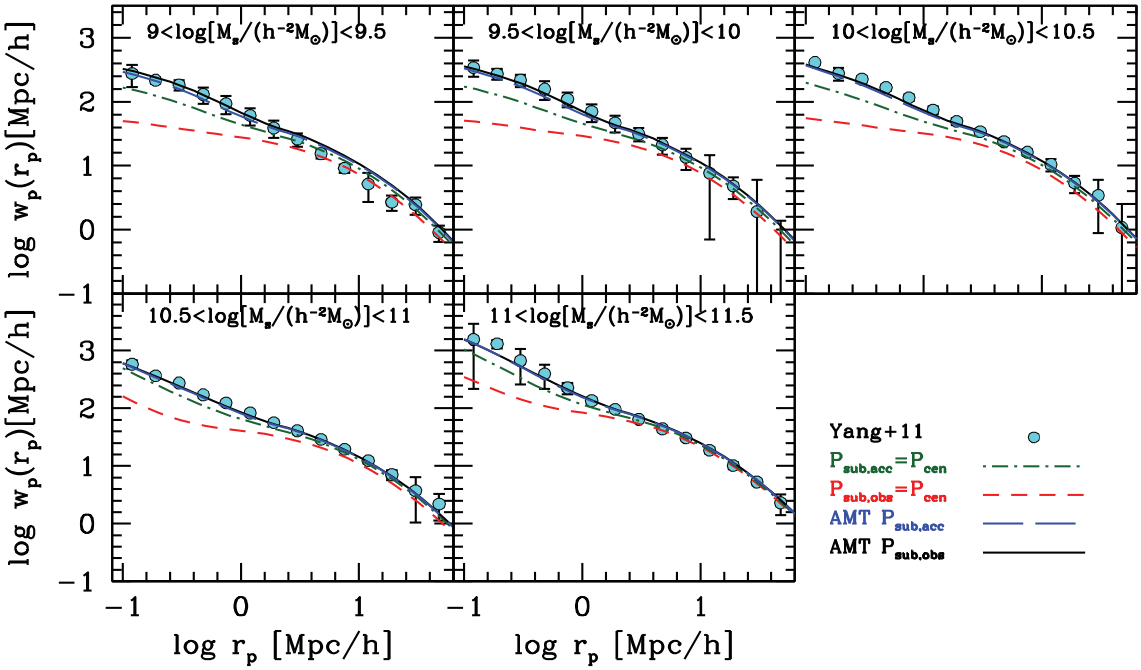

Figure 6: Projected two–point correlation functions of galaxies in five stellar mass intervals obtained with the HOD model using different galaxy-halo mass relations obtained with our AMT. Gray short-dashed and green short-dashed-dot lines are for the cases when the central SHMR and the SSMR were assumed to be equal and $m_{\rm sub}^{\rm obs}$ and $m_{\rm sub}^{\rm acc}$ were used, respectively. The black solid and blue long-dashed lines are again for $m_{\rm sub}^{\rm obs}$ and $m_{\rm sub}^{\rm acc}$, respectively, but in the case the central and satellite mass functions were independently constrained (they overlap most of time). The observed projected correlation functions reported in Yang et al. [250] are shown by filled circles with error bars.

2.3.3 *Abundance matching and clustering*

It has been noted in the literature that the average (total) SHMR obtained with the standard AMT is consistent with the observed spatial clustering of galaxies [58, 148]. We will test now whether this is the case for the mass relations of central and satellite galaxies obtained here with our extended AMT. We will compute the galaxy projected correlation function by means of a HOD model for each of the mass relation obtained in 2.3.

A HOD model is a statistical tool mainly used to describe the clustering of galaxies (e.g., 25, 60, 245, 259, 258, 121, 122, 250, and more references therein). In contrast to the AMT, which is a quasi empirical tool, a HOD employs modeling motivated by results of cosmological N-body [e.g., 115] and hydrodynamical [e.g., 262] simulations.

In short, a HOD model describes the probability that a halo of mass $M_{\rm h}$ hosts a number of N galaxies with stellar masses greater than M_*. Once the occupational numbers are defined, the two-point correlation function can be computed assuming that the total mean number of galaxy pairs is the contribution of all pairs coming from galaxies in the same halo (one-halo term, corresponding to smaller separations) and pairs from different halos (two-halo term, corresponding to larger separations). For a detailed description for the HOD model we employ here, see Appendix A.

First, consider the case when $P_{\rm sat}(m_*|m_{\rm sub}) = P_{\rm cen}(M_*|M_{\rm h})$. The short-dashed curves in Fig. 6 show the projected correlation functions in five stellar mass bins for the case when the $m_{\rm sub}^{\rm obs}$ subhCMF was used. The Yang et al. [250] galaxy projected correlation functions from the DR7 SDSS are plotted as filled circles with error bars. The resulting correlation functions are clearly below observations, mainly in the one-halo term. This is because using $m_{\rm sub}^{\rm obs}$ underestimates the satellite GSMF and CSMF, resulting in a strong deficit of satellite galaxies. Observe that if $N_s \sim 0$, then $N \sim N_c$ and therefore, $b_g(M_*) \sim \langle b(M_{\rm h})\rangle$ where $\langle b(M_{\rm h})\rangle$ is the mean weighted halo bias function, see Eq. 138.

When using the subhCMF for $m_{\rm sub}^{\rm acc}$ instead of $m_{\rm sub}^{\rm obs}$ (dot-short-dashed curves), the agreement with the observed correlation functions is better, though at scales where the one-halo term dominates, the predictions are still below observations. This is, again, because the satellite GSMF and CSMF are underestimated in this case. We remark that using the (average or total) SHMR obtained with the standard AMT in the HOD model by matching the total GSMF to the total halo+subhalo mass function (in the case of $m_{\rm sub}^{\rm acc}$), leads to excellent agreement with the observed correlation functions, a result that is well known. However in this case the SSMR is not constrained, instead it is implicitly assumed to be equal to the central SHMR (for $m_{\rm sub}^{\rm acc}$). With our extended AMT, we can explicitly separate both mass relations. When they are assumed to be equal and the central SHMR is constrained with the central GSMF, then we obtain the predictions already shown, in particular the correlation functions. The fact that the predicted correlation functions, when $m_{\rm sub}^{\rm acc}$ is used, are close to those predicted in the standard AMT (and to the observed ones) is because the central and average SHMR are indeed close, as we discussed in §§3.1, see Fig. 4.

Thus, under the assumption that $P_{\rm sat}(m_*|m_{\rm sub}) = P_{\rm cen}(M_*|M_{\rm h})$, the observed clustering of galaxies is better reproduced when the subhalo mass in abundance matching is defined as $m_{\rm sub}^{\rm acc}$ rather than $m_{\rm sub}^{\rm obs}$. Nevertheless, even in the former case, the agreement with observations is only marginal.

We now turn the analysis to the cases where the SSMR is not assumed to be equal to the central SHMR. The black solid and blue long-dashed lines in Fig. 6 show the predicted correlation functions in the cases where either $m_{\rm sub}$ or $m_{\rm sub}^{\rm acc}$ were used. Both cases lead to very similar results and agree very well with observations.

Therefore, *the HOD model combined with the central and satellite mass relations independently constrained with the extended AMT, is able to reproduce the observed correlation functions, no matter if* $m_{\rm sub}^{\rm obs}$ *or* $m_{\rm sub}^{\rm acc}$ *are used*. This successful prediction is a consequence of the good agreement obtained between our predicted satellite CSMFs and those inferred from observations (2.3.2 and 5).

2.4 SUMMARY AND DISCUSSION

In this Chapter, we extend the AMT in order to constrain both the central stellar–halo and the satellite–subhalo mass relations separately, using as an input (i) the distinct halo and subhalo mass functions, and (ii) the observed central and satellite GSMFs. Our formalism, by construction, predicts also the satellite CSMFs as a function of host halo mass, and when applied to a HOD model, allows to predict the spatial correlation functions. We present results for the cases when the SSMR is assumed to be equal to the central

SHMR, $P_{\rm sat}(m_*|m_{\rm sub})$=$P_{\rm cen}(M_*|M_{\rm h})$, and when both mass relations are constrained independently (i.e., it is not assumed that $P_{\rm sat}(m_*|m_{\rm sub})$=$P_{\rm cen}(M_*|M_{\rm h})$). All our analysis is carried out for subhalo masses defined at accretion time, $m_{\rm sub}^{\rm acc}$, and at the observed time (present day), $m_{\rm sub}^{\rm obs}$. The main results and conclusions are as follows:

- Assuming that the mass relation between satellites and subhalos is identical to the mass relation between centrals and distinct halos (including their intrinsic scatters), $P_{\rm sat}(m_*|m_{\rm sub})$=$P_{\rm cen}(M_*|M_{\rm h})$, the predicted satellite GSMF, CSMFs and projected two–point correlation functions lie below those obtained from observations for both definitions of $m_{\rm sub}$, though the disagreements are small when $m_{\rm sub}$=$m_{\rm sub}^{\rm acc}$ (Figs. 3, 5, 6). We conclude that assuming $P_{\rm sat}(m_*|m_{\rm sub})$=$P_{\rm cen}(M_*|M_{\rm h})$ leads to predictions in disagreement with observations, specially when $m_{\rm sub}^{\rm obs}$ is used.

- When the SSMR is no longer assumed to be equal to the central SHMR and instead is constrained by means of the observed satellite GSMF, the predicted satellite CSMFs and projected correlation functions agree in general with observations, both for $m_{\rm sub}^{\rm obs}$ and $m_{\rm sub}^{\rm acc}$. However, for halo masses lower than $\sim 10^{13}$ $M_\odot$, the number of very massive (rare) satellites is over-predicted.

- The resulting $m_{\rm sub}$–m_* relations when using either $m_{\rm sub}^{\rm obs}$ or $m_{\rm sub}^{\rm acc}$ are quite different from each other, and in each case are different from the central SHMR (Fig. 4). For a given stellar mass, the satellite subhalo mass is smaller than central halo mass, and the mass difference is increasing the lower the mass is. These differences are dramatic when $m_{\rm sub}^{\rm obs}$ is used.

- Our density-weighted average (centrals + satellites) SHMRs are close to the central SHMR when either $m_{\rm sub}^{\rm obs}$ or $m_{\rm sub}^{\rm acc}$ is used (central galaxies dominate in the YMB09 catalog). Such an average SHMR coincides conceptually with the one inferred from matching the total (centrals+ satellites) cumulative GSMF and the halo + subhalo cumulative mass function (standard AMT).

2.4.1 *On the inference of the SSMR and its implications for the average mass relation*

The conclusions listed above can be well understood by examining the basic ideas behind the extended AMT, as we show in Section 2.2. Essentially, matching abundances of satellite galaxies to subhalos is equivalent to matching their corresponding occupational numbers, that is:

$$\langle N_{\rm s}(>m_*|M_{\rm h})\rangle = \langle N_{\rm sub}(>m_{\rm sub}|M_{\rm h})\rangle. \tag{11}$$

The opposite is also true: matching their corresponding occupational numbers is equivalent to matching their abundances. This is an important result because it shows that once $P_{\rm sat}(m_*|m_{\rm sub})$ (and $P_{\rm cen}(M_*|M_{\rm h})$) is properly constrained, we will obtain the correct conditional mass functions and consequently the correct spatial clustering for galaxies.

The above means that there is a unique $m_{\rm sub}(m_*)$ relationship for each definition of $m_{\rm sub}$, which depends solely on $\langle N_{\rm sub}(>m_{\rm sub}|M_{\rm h})\rangle$. Because of this uniqueness, it follows that the $m_{\rm sub}^{\rm obs}(m_*)$ and $m_{\rm sub}^{\rm acc}(m_*)$ relations should be different, and any incorrect

assumption on each one of these relations will lead to inconsistencies in the conditional mass functions and spatial clustering of galaxies, as for example those that we have found here when $P_{\rm sat}(m_*|m_{\rm sub})$ was assumed to be equal to $P_{\rm cen}(M_*|M_{\rm h})$. Under this assumption, when $m_{\rm sub}^{\rm acc}$ was used, the inconsistencies were actually small. This is because in this case the "incorrect" assumption for the SSMR is actually not too far from the "correct" result obtained when $P_{\rm sat}(m_*|m_{\rm sub}^{\rm obs})$ is independently constrained (compare dot-dashed and solid green curves in Fig. 4), contrary to what happens when $m_{\rm sub}^{\rm obs}$ is used.

It is important to remark that in the standard AMT, only the average SHMR is constrained (using the total GSMF), leaving unconstrained the SSMR, something that on its own introduces a large uncertainty in the average SHMR [see 152]. We have shown that such average SHMR is conceptually equal to the density–weighted average mass relation obtained here from the observationally constrained central SHMR and SSMR. Therefore, our resulting average mass relation is expected to be less uncertain than previous determinations. On the other hand, this average mass relation is expected to be close to the central SHMR because most of the galaxies in the used observational catalog are centrals.

We conclude that in order to properly infer the SSMR and the central SHMR at the same time, and in this way reduce the uncertainty in the average SHMR, additional observational constraints than the total GSMF are necessary. The most obvious and direct is the GSMF decomposed into central and satellite galaxies, something that was provided by YMB09. However, observe that, according to eq. (11), the satellite CSMFs or the clustering of galaxies, modulo the observational errors, provide observational constraints that lead to similar inferences of the SSMR, because of the uniqueness of this relation for a given well defined $\langle N_{\rm sub}(>m_{\rm sub}|M_{\rm h})\rangle$ (see above).

Finally, we note that obtaining the SSMR for the subhalo mass defined at the accretion time introduces uncertainties due to our ignorance about evolutionary processes of the stellar mass since accretion . This does not happen when the SSMR is obtained for both the satellite and subhalo masses defined at the same epoch, for instance the present time. When matching abundances for the $m_{\rm sub}^{\rm acc}$ case, the fact that (1) $m_{\rm sub}^{\rm acc}$ is itself a cumulative distribution of all objects accreted over a period of time, and that (2) m_* may have changed between accretion and observation, are not taken into account. In other words, it is implicitly assumed that the satellite stellar mass stops evolving soon after accretion. In reality the situation is actually quite complex in the sense that, depending on the accretion time and the orbit of the satellites, the evolution of their stellar masses is diverse, with some of them early quenched and others actively evolving, perhaps in some cases as the central ones of the same mass [see e.g., 237, for semi-empirical inferences on such a complexity of galaxy evolution in groups]. This diversity introduces an intrinsic uncertainty on the results. Such an uncertainty might be accounted for the probability distribution functions: $P(m_*|m_{*,\rm acc},z)$, which gives the probability that a satellite accreted at epoch z evolves, on average, to the observed satellite m_*, and $P(m_{*,\rm acc}|m_{\rm sub}^{\rm acc},z)$, which gives the probability that a subhalo $m_{\rm sub}^{\rm acc}$ hosts a galaxy of mass $m_{*,\rm acc}$ at the time of accretion. Now, the satellite CSMF (Eq. 2) can be written as [143]:

$$\Phi_s(m_*|M_{\rm h}) = \int\int\int P(m_*|m_{*,\rm acc},z)P(m_{*,\rm acc}|m_{\rm sub}^{\rm acc},z)\times\Phi(m_{\rm sub}^{\rm acc}|M_{\rm h},z)dm_{*,\rm acc}dm_{\rm sub}^{\rm acc}dz.$$

(12)

Note that in our analysis in §§3.1, we implicitly assume that the stellar mass of satellite galaxies does not change once they become satellites, i.e. $P(m_*|m_{*,\rm acc},z) = \delta(m_* - m_{*,\rm acc},z)$, and that $P(m_{*,\rm acc}|m_{\rm sub}^{\rm acc},z)$ is independent of redshift. Thus, the application of the AMT to infer the satellite CSMF and the m_*–$m_{\rm sub}$ relation for subhalo mass defined at its accretion time formally requires more observational constraints at higher redshifts. This is a problem already faced by previous authors [e.g., 250].

The above is not the only way to formally write the satellite CSMF; it can be written in a way that instead of implying knowledge of the change of m_* from accretion to observation, implies just knowledge on the change of the subhCMF between these two epochs. Let us consider the distribution function, $P_{\rm acc}(m_{\rm sub}^{\rm obs}|m_{\rm sub}^{\rm acc},z)$, giving the probability that halos accreted at epoch z evolve, on average, to the observed (present-day) subhalos $m_{\rm sub}^{\rm obs}$, and the probability distribution function of these subhalos of hosting a galaxy of mass m_*, $P_{\rm sat}(m_*|m_{\rm sub}^{\rm obs})$. In this case, the satellite CSMF (Eq. 2) is written as

$$\Phi_s(m_*|M_{\rm h}) = \int\int\int P_{\rm sat}(m_*|m_{\rm sub}^{\rm obs})P_{\rm acc}(m_{\rm sub}^{\rm obs}|m_{\rm sub}^{\rm acc},z) \times \Phi(m_{\rm sub}^{\rm acc}|M_{\rm h},z)dm_{\rm sub}^{\rm obs}dm_{\rm sub}^{\rm acc}dz, \tag{13}$$

and therefore the satellite GSMF, Eq. 4, is given by

$$\phi_{\rm sat}(m_*) = \int\int\int P_{\rm sat}(m_*|m_{\rm sub}^{\rm obs})P_{\rm acc}(m_{\rm sub}^{\rm obs}|m_{\rm sub}^{\rm acc},z) \times \phi_{\rm sub}(m_{\rm sub}^{\rm acc},z)dm_{\rm sub}^{\rm obs}dm_{\rm sub}^{\rm acc}dz. \tag{14}$$

Since the $m_{\rm sub}^{\rm acc}$ subHMF would evolve into the $m_{\rm sub}^{\rm obs}$ subHMF, we write

$$\phi_{\rm sub}(m_{\rm sub}^{\rm obs}) = \int\int P_{\rm acc}(m_{\rm sub}^{\rm obs}|m_{\rm sub}^{\rm acc},z)\phi_{\rm sub}(m_{\rm sub}^{\rm acc},z)dm_{\rm sub}^{\rm acc}dz. \tag{15}$$

This last equation is the abundance matching of accreted subhalos to present-day subhalos. Therefore,

$$\phi_{\rm sat}(m_*) = \int P_{\rm sat}(m_*|m_{\rm sub}^{\rm obs})\phi_{\rm sub}(m_{\rm sub}^{\rm obs})dm_{\rm sub}^{\rm obs}. \tag{16}$$

This equation is nothing but abundance matching satellite galaxies to subhalos *at the time they are observed*. Hence, the reason that the satellite GSMF matches the subHMF in a more direct way for subhalo masses defined at the observation time (eq. 16) than at the accretion time (eq. 12), is that in the latter case the unknown $P(m_*|m_{*,\rm acc},z)$ and $P(m_{*,\rm acc}|m_{\rm sub}^{\rm acc},z)$ "evolutionary" functions have to be introduced. However, we acknowledge that for the former case, our ignorance on the scatter around the SSMR is a also potential source of uncertainty. All our calculations are under the assumption that this scatter is the same as the scatter of the central SHMR. In any case, even if these scatters are different, note that including scatter affects the stellar-to-(sub)halo mass relation only at its high-mass end, where on average satellites are expected to be accreted recently, hence their SSMR and scatter are yet similar to those of centrals/halos.

2.4.2 *On the intrinsic scatter in the SSMR*

A possible source of systematic errors in our analysis is the assumption that the intrinsic scatter around the SSMR, σ_s is constant and equal to the scatter around the central SHMR. To probe the impact of this assumption we repeated all our analysis but this time assuming $\sigma_s = 0$. When comparing the results using $\sigma_s = 0$ to those obtained based on $\sigma_s = 0.173$ dex, we find that they are consistent with each other, and therefore with the satellite CSMFs and with the galaxy spatial clustering measured from the YMB09 catalog. In more detail, we find that the resulting CSMF's reproduce observations for $\sigma_s = 0$ slightly better than for $\sigma_s = 0.173$ dex, especially at the massive end. This is because when the intrinsic scatter is not taken into account ($\sigma_s = 0$), the shape of the SSMRs steepens at the massive end (see also 19). Consequently, for a given m_*, the subhalo mass is larger, and the abundances of larger (sub)halos is lower in general than those of smaller halos. Therefore, the number density of satellites at the massive end is lower. However, the projected correlation functions remain almost the same because the probability of finding a massive satellite galaxy in host halos less massive than $\log M_{\rm h} \sim 13$ is very low. They do not contribute significantly to the mean total density of galaxies. Although better models are needed in order to give a realistic form for σ_s, our main conclusions seem to be robust to variations in the adopted value for σ_s.

2.4.3 *Implications for satellite/subhalo evolution*

The local SSMR obtained for both definitions of the subhalo mass, $m_{\rm sub}^{\rm acc}$ and $m_{\rm sub}^{\rm obs}$, are such that at halo masses smaller than $2 - 10 \times 10^{13}$ $M_\odot$ and at a given galaxy stellar mass, the corresponding subhalo mass is smaller on average than the halo mass of centrals (Fig. 4). This difference increases the smaller the mass is, and much more for the subhalo mass defined at the observed time (present-day). In the case of $m_{\rm sub}^{\rm acc}$, the differences might be because the halo mass at the epoch it became a subhalo (accretion time) is smaller than its present-day counterpart at a given stellar mass and/or because the satellite stellar mass increased faster than the central one for a given halo mass. In fact, it is difficult to make any inference in this case because the abundance matching is between *local* galaxies and (sub)halos at *different past* epochs. In any case, the fact that the inferred mass relations for satellites and centrals when $m_{\rm sub}^{\rm acc}$ is used are not too different, suggests that the central galaxy–distinct halo mass relation does not change too much with time, at least since the epoch at which most of the subhalos were accreted.

When $m_{\rm sub}^{\rm obs}$ is used, both abundances of satellites and subhalos are matched at the same epoch, the observation (present-day) time. In this case the strong difference between the satellite and central mass relations can be interpreted mainly as the result of subhalo mass loss due to tidal stripping. Besides, the smaller the subhalo, the larger is the mass loss on average, though there is a large scatter and highly dependent of the orbit. Probably, the different evolution in stellar mass between central and satellite galaxies could also play a role for the differences but not as significant a role that the one related to halo and subhalo mass evolution.

From Fig. 4 one sees that for a given m_*, the $m_{\rm sub}^{\rm obs}$–to–$m_{\rm sub}^{\rm acc}$ ratio is 0.35–0.40 for the smallest masses up to $m_* \sim 2 \times 10^{11}$ $M_\odot$. At larger masses, this ratio rapidly tends

to 1. Therefore, the subhalos of satellites galaxies less massive than $m_* \sim 2 \times 10^{11}$ $M_\odot$ have lost, on average (for all host halo masses[4]), 60–65% of their masses since they were accreted. It should be noted that this is a rough approximation and the evolution of the stellar mass since the satellite was accreted should be taken into account, see §4.1. This result shows us that the galaxy-(sub)halo connection for satellite galaxies is far from direct; present-day satellites of masses $m_* \sim 7 \times 10^8$ $M_\odot$ and larger have formed in subhalos that at the time they were accreted onto galaxy sized halos were on average a factor 2.5–3 larger than at the present epoch. This has severe implications for studies aimed to constrain the Λ-CDM scenario at the level of subhalo/satellite distributions.

For example, it has been discussed that seeding the subhalos in simulations of MW-like halos by using an extrapolation to low masses of the stellar–halo mass relation obtained by means of the AMT, predicts a MW dwarf spheroidal (dSph) luminosity function in agreement with the observed one. However, the circular velocities at the maximum (or the masses at the infall) of the subhalos associated to the dSphs are significantly larger than inferences from observed kinematics [37].

In the right panel of Fig. 4 we have plotted the extrapolation to low masses of our SSMRs, both for subhalo masses defined at the present day (red line) and at the infall time (blue line). The observational points in the panel are for MW satellites, which subhalo masses were estimated at their truncation radii. Thus, if we assume that these masses are roughly equal to the present-day subhalo masses in the Λ-CDM simulations, then the simulated subhalo masses, $m_{\rm sub}^{\rm obs}$, are up to $\approx 10-30$ times larger than those associated to dSphs. If the comparison is done with the extrapolation of the average (or central) SHMR, then the differences increases by a factor of ~ 3 more [see also Fig. 7 in 37].

Our extrapolated results show that the discrepancy in subhalo mass between MW bright dSphs and Λ-CDM simulations is smaller than previously reported but is still significant. Note that for the extrapolation, we have used the same slope of the YMB09 satellite GSMF at the low mass end, $\alpha = -1.25$ (Fig. 3). If this slope steepens for smaller masses, for example to a value of $\alpha = -1.6$, then our extended AMT predict the $m_{\rm sub}^{\rm obs}$ SSMR plotted as the gray dotted-dashed curve in Fig. 4, which is already consistent with the dynamical estimates.

The GSMF at low masses may be significantly incomplete because of missing low-surface brightness galaxies. By taking into account the bivariate distribution of stellar mass versus surface brightness, Baldry, Glazebrook & Driver [16] have found evidence for an upturn in the faint-end GSMF slope ($\alpha \approx -1.6$) for a subsample of field SDSS galaxies. More recently, using the GAMA survey, a slope of $\alpha \approx -1.47$ has been reported [14]. Steep faint-end slopes have been also found at higher redshifts. For instance, using the COSMOS field, Drory et al. [75] have measured slopes of $\alpha \sim -1.7$ at all redshifts $z \leq 1$ There are also pieces of evidence that the faint-end slope of the GSMF (or luminosity function) changes with the environment: in clusters of galaxies it steepens significantly [for a discussion see 16, and the references therein]. The cluster GSMF is actually related to the satellite GSMF, through the satellite CSMF.

We conclude that using a correct AMT for connecting satellite galaxies to their present-day subhalos and assuming a steep faint-end slope in the satellite GSMF ($\alpha \sim -1.6$), the

4 The dependence of the satellite subhalo mass loss on host halo mass will be explored elsewhere.

predicted SSMR for the Λ-CDM cosmogony would be consistent with the dynamics of MW satellites.

ACKNOWLEDGMENTS

We thank Alexie Leathaud, Ramin Skibba, Surhud More, Xiaohu Yang, and the anonymous Referee for useful comments, and suggestions. A. R-P acknowledges a graduate student fellowship provided by CONACyT. N. D. and V. A. acknowledge to CONACyT grant 128556 (Ciencia Básica). V.A acknowledges PAPIIT-UNAM grant IN114509. We are grateful to X. Yang for providing us in electronic form their data for the CSMFs.

3

THE GALAXY–HALO/SUBHALO CONNECTION AT $z \sim 0$

Part of this Chapter was published as: Rodríguez-Puebla A.; Avila-Reese V.; Drory N., 2013, ApJ 767, 92.

ABSTRACT

We infer the local stellar-to-halo/subhalo mass relations (MRs) for central and satellite galaxies *separately*. Our statistical method is extending the abundance matching, halo occupation distribution, and conditional stellar mass function formalisms. We constrain the model using several combinations of observational data, consisting of the total galaxy stellar mass function (GSMF), its decomposition into centrals and satellites, and the projected two-point correlation functions (2PCFs) measured in different stellar mass (M_*) bins. In addition, we use the ΛCDM halo and subhalo mass functions. The differences among the resulting MRs are within the model-fit uncertainties (which are very small, smaller than the intrinsic scatter between galaxy and halo mass), no matter what combination of data are used. This shows that matching abundances or occupational numbers is equivalent, and that the GSMFs and 2PCFs are tightly connected. We also constrain the values of the intrinsic scatter around the central-halo (CH) and satellite-subhalo (SS) MRs assuming them to be constant: $\sigma_c = 0.168 \pm 0.051$ dex and $\sigma_s = 0.172 \pm 0.057$ dex, respectively. The CH and SS MRs are actually different, in particular when we take the subhalo mass at the present-day epoch instead of at their accretion time. When using the MRs for studying the satellite population (e.g., in the Milky Way, MW), the SS MR should be chosen instead of the average one. Our model allows one to calculate several population statistics.

3.1 INTRODUCTION

The statistical description of the galaxy population is a valuable tool for understanding the properties of galaxies and the way they cluster, as well as the role that mass and environment play in shaping these properties. Moreover, statistical descriptors such as the luminosity function, the galaxy stellar mass function (GSMF), and the two-point correlation function (2PCF) has allowed us to probe galaxy evolution and its connection to the cosmological initial conditions of structure formation [e.g., 162, 252]. Such a connection is of vital importance in studies devoted to the development of the current ΛCDM cosmological paradigm. A key ingredient in these studies is the link between galaxy and dark matter halo properties. Such a link allows to project the theoretical dark matter halo population onto the observable galaxies.

Recently, progress towards connecting galaxies and halos has been made through the development of several techniques for observationally estimating the dark halo masses of luminous galaxies, such as weak lensing [133, 132, 184], kinematics of satellite galaxies

[56, 145, 146, 241], and galaxy clusters [249, 247, 103, 250]. However, these direct probes of halo mass still have large uncertainties.

Consequentially, semi-empirical approaches that link the galaxy and dark matter halo distributions statistically are of great importance. For example, the Halo Occupation Distribution (HOD) formalism, which describes the probability for finding N galaxies in halos of mass $M_{\rm h}$, has been used successfully to understand the non-linear relation between the distribution of galaxies and matter, for instance, at the level of the power spectra [186, 161, 60, 252], or the two-point correlation functions [25, 60, 259, 1, 208, 258, 229, 228, 223, and references therein].

However, the HOD model provides only information on the total number of galaxies above some luminosity or stellar mass threshold per halo, and constrains only the halo mass of the central galaxy. In order to describe the detailed halo occupation and mass distribution of central and satellite galaxies, Yang, Mo & van den Bosch [245] introduced the conditional luminosity (or stellar mass) function (CSMF) in the HOD model (see also e.g., 247, hereafter YMB09, 148, 121, 122, 250). The CSMF is defined as the average number of galaxies with stellar masses between $M_* \pm dM_*/2$ occupying a halo of a given mass $M_{\rm h}$. Nevertheless, both the HOD model and the CSMF formalism assume a parametric description for the satellite population distributions which is constrained using observations.

In order to avoid an arbitrary parametric description for the satellite population, the above models can be generalized with the abundance matching technique (hereafter, AMT; e.g., 216, 115, 58, 187, 234, 16, 217, 57, 75, 148, 19, 102, 20, 170, 159). Under the hypothesis that there exists a one-to-one monotonic relation between stellar mass and halo plus subhalo mass, matching the total galaxy and halo plus subhalo abundances yields a *global* (average) relation between M_* and $M_{\rm h}$. Note that in this simple procedure, the central-to-halo and satellite-to-subhalo mass relations (hereafter CHMR and SSMR, respectively) are not differentiated. Recently, Simha et al. [193] have found in their cosmological N-body/hydrodynamics simulations that both mass relations are nearly identical if the subhalo masses, $m_{\rm sub}$, are defined at their accretion times. Additionally, previous studies have shown that when the AMT results are applied to the HOD model with $m_{\rm sub}$ defined at the accretion epoch, then the spatial clustering of galaxies is mostly recovered (e.g.,58, 148). Similar results are expected when $m_{\rm sub}$ is defined at the observation time but a global offset is applied to account for the average effect of subhalo mass loss due to tidal stripping [216, 234].

On the other hand, there is no reason to assume a priori the SSMR to be identical to the CHMR (152, 178, hereafter RDA12). For accretion-time $m_{\rm sub}$, such an assumption implies that the change of the stellar masses of satellites after their accretion will be such that they would occupy the $z = 0$ central-to-halo mass relation or, more generally, that the CHMR almost does not change with time. Recent studies based on large halo-based group catalogs [e.g., 237] or on the predicted bulge-to-total mass ratio of central galaxies[254] have shown that once satellite galaxies are accreted, they evolve roughly as a central galaxy at least for several Gyrs. This could imply that the SSMR with $m_{\rm sub}$ defined at accretion time may not be equal to the $z = 0$ CHMR. Nevertheless, in the cosmological simulations of Simha et al. [193], despite the fact that satellites continue to

grow after accretion, both mass relations end up similar. It is therefore likely that the growth in mass as well as the change of the CHMR with time are very small.

In the previous Chapter, see also RDA12, we extended the AMT to determine the CHMR and SSMR separately, using the observed decomposition of the GSMFs into centrals and satellites. We have found that indeed the SSMR is not equal to the CHMR, and that applying them to the HOD + CSMF model leads to satellite CSMF and correlation functions in excellent agreement with observational data. Actually, when $m_{\rm sub}$ is defined at the accretion time, the $z = 0$ mass relations become close but not equal (see Fig. 2 in RDA12). Additionally, RDA12 show that the uncertainty in the AMT related to the satellite stellar mass growth can be avoided if subhalo masses are defined at the time of observation rather than at the time of accretion. RDA12 also suggest that the central-halo and satellite-subhalo mass relations can be determined simultaneously using the correlation functions as observational input, instead of the GSMF decomposed into satellites and centrals. This is presumably because *matching abundances of satellite to subhalos is essentially equivalent to matching their corresponding occupational numbers (and vice versa)*.

In the present Chapter, we aim to test the above statements. We will also probe how robust the determinations of the central-to-halo and satellite-to-subhalo mass relations through our extended AMT and HOD+CSMF combined model are. We will explore whether these mass relations vary significantly depending on the combinations of observational data being used; in particular, we will explore whether the uncertainties in the model parameters that describe the mass relations shrink significantly when more observational constraints are added.

In Section 3.2 we describe our extended AMT, already presented in the previous Chapter, combined with the HOD+CSMF model, and present the different combinations of data to be used to constrain the model parameters. The results of our model for the different data sets are presented in Section 3.3. In particular, we compare the central-halo and satellite-subhalo mass relations obtained using different data sets. We also constrain the intrinsic scatter around the mean central-halo and satellite-subhalo mass relations. At the end of this section we discuss the halo occupational statistics related to the halo mass dependence of the satellite CSMF. Section 3.4 is devoted to discuss the robustness of the obtained mass relations and their model uncertainties, as well as the implications of extrapolating our obtained SSMR to masses as small as the MW dwarf spheroidal galaxies. Finally, we present our conclusions in Section 3.5.

We adopt cosmological parameter values close to WMAP 7: $\Omega_\Lambda = 0.73, \Omega_{\rm M} = 0.27, h = 0.70, n_s = 0.98$ and $\sigma_8 = 0.84$.

3.2 METHODOLOGY

In the following we present our model connecting galaxies to halos and subhalos via their occupational numbers. This is done under the assumption that on average the central-to-halo and satellite-to-subhalo relations are monotonic. The model relates in a self-consistent way the GSMF decomposed into centrals and satellites, the ΛCDM halo/subhalo mass functions, the satellite CSMFs, and the galaxy projected 2PCFs. As a result, it constrains both the CHMR and the SSMR, and predicts the satellite CSMF and

several other occupational statistics. Unlike previous models of this kind [e.g., 148], the CHMR and SSMR are treated separately.

3.2.1 *Connecting galaxies to halos and subhalos*

The *total* GSMF is decomposed into satellites and central galaxies,

$$\phi_g(M_*) = \phi_{g,\mathrm{cen}}(M_*) + \phi_{g,\mathrm{sat}}(M_*), \tag{17}$$

which after integration yields the mean cumulative number density of galaxies with stellar masses greater than M_*,

$$\int_{M_*}^{\infty} \phi_g dM_*{}' = \int_{M_*}^{\infty} \phi_{g,\mathrm{cen}} dM_*{}' + \int_{M_*}^{\infty} \phi_{g,\mathrm{sat}} dM_*{}', \tag{18}$$

or, in short,

$$n_g(>M_*) = n_{g,\mathrm{cen}}(>M_*) + n_{g,\mathrm{sat}}(>M_*). \tag{19}$$

3.2.1.1 *Central galaxies*

For constructing the central GSMF, we will use the conditional probability that a given halo of mass M_h is inhabited by a central galaxy with stellar mass between $M_* \pm dM_*/2$, $P_\mathrm{cen}(M_*|M_\mathrm{h})dM_*$, and assume this distribution to be log-normal:

$$P_\mathrm{cen}(M_*|M_\mathrm{h})dM_* = \frac{dM_*}{\sqrt{2\pi\sigma_c^2} M_* \ln(10)} \times \exp\left[-\frac{\log^2(M_*/M_{*,\mathrm{c}}(M_\mathrm{h}))}{2\sigma_c^2}\right], \tag{20}$$

with σ_c being the intrinsic scatter (width), expressed in dex units, around $\log M_{*,\mathrm{c}}(M_\mathrm{h})$, the mean stellar-to-halo mass relation of *central galaxies* (CHMR). Formally, $P_\mathrm{cen}(M_*|M_\mathrm{h})$ maps the HMF onto the central GSMF, thereby encoding all the physical processes involved in galaxy formation inside the halos. We parametrize $\log M_{*,\mathrm{c}}(M_\mathrm{h})$ using the functional form proposed by Behroozi, Wechsler & Conroy [20],

$$\log M_{*,\mathrm{c}}(M_\mathrm{h}) = \log(\epsilon_c M_{1,c}) + f(\log(M_\mathrm{h}/M_{1,c})) - f(0), \tag{21}$$

where

$$f(x) = \delta_c \frac{(\log(1+e^x))^{\gamma_c}}{1+e^{10^{-x}}} - \log(10^{\alpha_c x}+1). \tag{22}$$

This function behaves as power law with slope α at masses much smaller than $M_{1,c}$, and as a sub-power law with slope γ_c at larger masses. This parametrization maps the ΛCDM HMF to a Schechter-like GSMF [183].

The mean number density of central galaxies with stellar masses between $M_* \pm dM_*/2$, (i.e., the central GSMF) is given by

$$\phi_{g,\mathrm{cen}}(M_*)dM_* = dM_* \int_0^{\infty} P_\mathrm{cen}(M_*|M_\mathrm{h})\phi_\mathrm{h}(M_\mathrm{h})dM_\mathrm{h}, \tag{23}$$

where $\phi_{\rm h}(M_{\rm h})$ is the *distinct* HMF. We use the fitted results to the *distinct* HMF from cosmological simulations carried out in Tinker et al. [207] as reported in their Appendix B. Here we define halo masses at the virial radius, i.e. the halo radius where the spherical overdensity is $\Delta_{\rm vir}$ times the mean matter density, with $\Delta_{\rm vir} = (18\pi^2 + 82x - 39x^2)/\Omega(z)$, and $\Omega(z) = \rho_m(z)/\rho_{\rm crit}$ and $x = \Omega(z) - 1$ [43].

Having defined $P_{\rm cen}(M_*|M_{\rm h})$, the cumulative probability that a halo of mass $M_{\rm h}$ hosts a central galaxy with a stellar mass greater than M_* is simply

$$\int_{M_*}^{\infty} P_{\rm cen}(M_*|M_{\rm h}) dM_*, \tag{24}$$

which coincides with the definition of the mean occupational number of central galaxies, $\langle N_c(>M_*|M_{\rm h})\rangle$. Finally, we are able to infer the mean number density of galaxies with stellar mass greater than M_*, that is, $n_{g,{\rm cen}}(>M_*) = \int_0^\infty \langle N_c(>M_*|M_{\rm h})\rangle \phi_{\rm h}(M_{\rm h}) dM_{\rm h}$.

3.2.1.2 *Satellite galaxies*

Since satellites are expected to reside in subhalos, we will use a similar approach to centrals, i.e. we will establish a link between the properties of satellite galaxies to those of the subhalos. However, in this case, one should take into account that (i) before becoming a satellite they occupy a distinct halo, and (ii) the subhalo mass, $m_{\rm sub}$, can be defined at the observation time (their present-day mass in our case) or at the accretion time (the epoch when a distinct halo became a subhalo).

Item (ii) is discussed in RDA12. First, RDA12 show that once the subhalo mass function is provided for any definition of subhalo mass, the satellite-to-subhalo mass relation, SSMR, can be constrained consistently with the observed satellite GSMF. Therefore, the use of one or another is subject to practical criteria. On one hand, with the accretion-epoch definition, the central and satellite mass relations are almost the same, as observations (RDA12) and simulations (193) show, and the obtained SSMR for this case is free of a potential dependence on host halo mass. Besides, the accretion-time $m_{\rm sub}$ definition is less sensitive to the specifics of the halo finding algorithm than the observed-time definition. On the other hand, the SSMR for the subhalo mass defined at accretion time is actually a nominal relation, where the abundance matching is carried out for the satellite GSMF at the *present epoch* but for a subhalo mass function constructed for subhalos accreted at *different previous epochs*. The physical interpretation of this nominal relation requires assumptions about the evolution of galaxies. Instead, when matching present-day satellite abundances with present-day subhalo abundances, the connection is direct and no assumptions about evolution are necessary (see Chapter 2 and also RDA12,§§4.1, for an extensive discussion).

Here, our constraints for the SSMR refer to $m_{\rm sub}$ defined at the same epoch that the observational input is provided for, that is the present time. However, some results will be presented also for the accretion-time $m_{\rm sub}$.

For the subhalo abundance, given as the subhalo conditional mass function, we use the results obtained in Boylan-Kolchin et al. [41] based on the Millennium-II simulation. It is worth noting that the lowest subhalo masses we probe in this work ($\sim 10^{10} - 10^{11}{\rm M}_\odot$, depending on the GSMF used) are around 3–4 orders of magnitude above the mass resolution of this simulation. The present-day subhalo mass is the mass enclosed within a

truncation radius, which is defined as the radius where the spherically-averaged density profile starts to flatten or to increase with radius. The fitting formula for the mean cumulative number of subhalos with present-day (observed) mass $m_{\rm sub}$ given a host halo mass $M_{\rm h}$ is:

$$\langle N_{\rm sub}(>m_{\rm sub}|M_{\rm h})\rangle = \mu_0 \left(\frac{\mu}{\mu_1}\right)^a \exp\left[-\left(\frac{\mu}{\mu_{\rm cut}}\right)^b\right], \tag{25}$$

where $\mu = m_{\rm sub}/M_{\rm h}$ and $\{\mu_0, \mu_1, \mu_{\rm cut}, a, b\} = \{1.15^{(\log M_{\rm h}-12.25)}, 0.010, 0.096, -0.935, 1.29\}$. Then, the number of subhalos of mass between $m_{\rm sub} \pm dm_{\rm sub}/2$ residing in host halos of mass $M_{\rm h}$ (the SubhCMF), is simply

$$\Phi_{\rm sub}(m_{\rm sub}|M_{\rm h})dm_{\rm sub} = d\langle N_{\rm sub}(>m_{\rm sub}|M_{\rm h})\rangle. \tag{26}$$

The average cumulative number of subhalos reported in Boylan-Kolchin et al. [41], Eq. (25), was actually obtained for MW-sized halos. However, as the authors discuss, the normalization factor, μ_0, has been found to vary with $M_{\rm h}$, roughly 15% per dex in $M_{\rm h}$. For this reason we introduce the quantity $\mu_0 = 1.15^{(\log M_{\rm h}-12.25)}$ [see also 90].

The difference in cosmology between the Millennium-II simulation and ours leads to differences in the resulting abundances of subhalos of roughly a few per cent in the amplitude of the subhalo mass function (41), and it has little effects on our results (see RDA12). In any case, we introduce a correction to first order, taking advantage of the fact that Tinker et al. [207] provides the distinct HMF as a function of the relevant cosmological parameters. First, the subhalo mass function is calculated from Eqs. (25) and (26) and the Tinker et al. [207] HMF defined for the Millenium cosmology. Then, the "Millenium-cosmology" subhalo-to-halo mass function ratio is calculated, $T(M) = \phi_{\rm sub,MII}(M)/\phi_{\rm h,MII}(M)$. This ratio is now used to recalculate the subhalo mass function for our cosmology as $\phi_{\rm sub}(M) = T(M)\phi_{\rm h}(M)$, where $\phi_{\rm h}(M)$ is the Tinker et al. [207] HMF for our cosmology. Finally, assuming the same functional form for the subhalo conditional mass function (Eq. 26), with the same μ_0, we obtain the new parameters for our cosmology from χ^2 fitting $\{\mu_1, \mu_{\rm cut}, a, b\} = \{0.011, 0.096, -0.935, 1.342\}$. These are actually very close to what is reported in Boylan-Kolchin et al. [41].

Analogously to centrals, to construct the satellite GSMF we introduce the probability, $P_{\rm sat}(M_*|m_{\rm sub})dM_*$, that a subhalo of mass $m_{\rm sub}$ hosts a satellite galaxy with stellar mass between $M_* \pm dM_*/2$. In general there is no reason for assuming $P_{\rm cen}(M_*|M_{\rm h}) = P_{\rm sat}(M_*|m_{\rm sub})$[1]. We again adopt a log-normal form,

$$P_{\rm sat}(M_*|m_{\rm sub})dM_* = \frac{dM_*}{\sqrt{2\pi\sigma_s^2}M_*\ln(10)} \times \exp\left[-\frac{\log^2(M_*/M_{*,\rm s}(m_{\rm sub}))}{2\sigma_s^2}\right], \tag{27}$$

1 This assumption may actually lead to inconsistent results, even for the accretion-time $m_{\rm sub}$ definition, as shown in RDA12. For $m_{\rm sub}$ defined at the present time, tidal stripping affects the masses of the subhalos producing a systematic offset between the galaxy-halo and satellite-subhalo mass relations, which is sometimes incorporated as an assumed global offset in the AMT analyses [e.g., 216, 234]. For $m_{\rm sub}$ defined at the accretion epoch, the two relations become actually close according to the extended AMT analysis of RDA12 or to the results of cosmological simulations [193], but there may be still offsets and differences in scatter because of the uncertain evolution of the satellites after accretion (see Fig. 2 in RDA12 and Figs. 3 and 7 below).

where σ_s is the scatter (width) around the logarithm in base 10 of $M_{*,s}(m_{\rm sub})$, the mean satellite-subhalo mass relation (SSMR). Similarly to centrals, we parametrize $\log M_{*,s}(m_{\rm sub})$ using Eq. (21). The reason is because, as observations suggest, the shape of the satellite GSMF is also a Schechter-like function (e.g., YMB09; 250), which is easily reproduced from the halo or subhalo mass function using the parametrization given by Eq. (21).

The next step is to link satellites to subhalos. The most natural way to do this is via their occupational numbers [e.g., 248]. Let $\Phi_{\rm sat}(M_*|M_{\rm h})$ be the CSMF giving the mean number of satellites of stellar mass $M_* \pm M_*/2$ residing in a host halo of mass $M_{\rm h}$:

$$\Phi_{\rm sat}(M_*|M_{\rm h})dM_* = dM_* \int_0^\infty P_{\rm sat}(M_*|m_{\rm sub})\Phi_{\rm sub}(m_{\rm sub}|M_{\rm h})dm_{\rm sub}. \tag{28}$$

The similarity with Eq. (23) is not a coincidence, since this is actually the AMT in its differential form but at the level of CSMFs. Integrating this over stellar mass gives the mean occupation of satellite galaxies in individual halos:

$$\langle N_s(>M_*|M_{\rm h})\rangle = \int_{M_*}^\infty \Phi_{\rm sat}(M_*|M_{\rm h})dM_*. \tag{29}$$

At this point we are in a position to compute the satellite GSMF:

$$\phi_{g,{\rm sat}}(M_*)dM_* = dM_* \int_0^\infty \Phi_{\rm sat}(M_*|M_{\rm h})\phi_{\rm h}(M_{\rm h})dM_{\rm h}, \tag{30}$$

and in the case that σ_s is a constant,

$$\phi_{g,{\rm sat}}(M_*)dM_* = dM_* \int_0^\infty P_{\rm sat}(M_*|m_{\rm sub})\phi_{\rm sub}(m_{\rm sub})dm_{\rm sub}, \tag{31}$$

which is the matching of satellite galaxies to subhalos. The mean number density of satellite galaxies with stellar mass greater than M_* is given by:

$$n_{g,{\rm sat}}(>M_*) = \int_0^\infty \langle N_s(>M_*|M_{\rm h})\rangle \phi_{\rm h}(M_{\rm h})dM_{\rm h}. \tag{32}$$

Finally, note that the relation between $P_{\rm cen}(M_*|M_{\rm h})$ and $P_{\rm sat}(M_*|m_{\rm sub})$ with the distribution $P(M_*|M)$ used in the standard AMT (e.g., 217, 19) is given by

$$P(M_*|M) = \frac{\phi_{\rm sub}(M)}{\phi_{\rm DM}(M)}P_{\rm sat}(M_*|M) + \frac{\phi_{\rm h}(M)}{\phi_{\rm DM}(M)}P_{\rm cen}(M_*|M), \tag{33}$$

where $\phi_{\rm DM}(M) = \phi_{\rm sub}(M) + \phi_{\rm h}(M)$ and M applies either to the distinct halo or subhalo masses. Then, the above equation relates the mass relation commonly obtained through the standard AMT with those obtained in this Chapter,

$$\langle \log M_*(M)\rangle = \frac{\phi_{\rm sub}(M)}{\phi_{\rm DM}(M)}\langle \log M_{*,s}(M)\rangle + \frac{\phi_{\rm h}(M)}{\phi_{\rm DM}(M)}\langle \log M_{*,c}(M)\rangle, \tag{34}$$

where $M_{*,s}(M)$ and $M_{*,c}(M)$ are the SSMR and CHMR, respectively. It is worth noting that the standard AMT is recovered if both $P_{\rm cen}(M_*|M_{\rm h})$ and $P_{\rm sat}(M_*|m_{\rm sub})$ are assumed to be $\delta-$functions. Then $n_{g,{\rm cen}}(>M_*) + n_{g,{\rm sat}}(>M_*) = n_{\rm sub}(>M_{\rm h}) + n_{\rm h}(>M_{\rm h})$. For a detailed discussion see RDA12.

3.2.2 *The two-point correlation function*

So far, the galaxy-(sub)halo link is based on an extended AMT. However, having modeled the occupational numbers for central and satellite galaxies, we can now introduce information related to the spatial clustering. For convenience, we will write $\langle N \rangle \equiv \langle N(> M_*|M_{\rm h})\rangle$, $\langle N_c \rangle \equiv \langle N_c(> M_*|M_{\rm h})\rangle$ and $\langle N_s \rangle \equiv \langle N_s(> M_*|M_{\rm h})\rangle$.

As usual, the two-point correlation function is decomposed into two parts,

$$1 + \xi_{\rm gg}(r) = [1 + \xi^{\rm 1h}_{\rm gg}(r)] + [1 + \xi^{\rm 2h}_{\rm gg}(r)], \tag{35}$$

where $\xi^{\rm 1h}_{\rm gg}(r)$ describes pairs within the same halo (one-halo term), while $\xi^{\rm 2h}_{\rm gg}(r)$ describes pairs occupying different haloes (two-halo term). Notice that the terms $1 + \xi^{\rm 1h}_{\rm gg}(r)$ and $1 + \xi^{\rm 2h}_{\rm gg}(r)$ indicates that we are adding the pairs directly to the term $1 + \xi_{\rm gg}(r)$.

To compute the one-halo term, we need to count all galaxy pairs $\langle N(N-1)\rangle/2$ separated by a distance $r \pm dr/2$ within individual halos of mass $M_{\rm h}$, following a pair distribution $\lambda(r)dr$ weighted by the abundance of distinct halos, $\phi_{\rm h}$, and normalized by the mean galaxy number density n_g,

$$1 + \xi^{\rm 1h}_{\rm gg}(r) = \frac{1}{2\pi r^2 n_g^2} \int_0^\infty \frac{\langle N(N-1)\rangle}{2} \lambda(r) \phi_{\rm h}(M_{\rm h}) dM_{\rm h}. \tag{36}$$

The contribution to the mean number of galaxy pairs from central-satellite pairs and satellite-satellite pairs is given by

$$\frac{\langle N(N-1)\rangle}{2}\lambda(r)dr = \langle N_c\rangle\langle N_s\rangle\lambda_{c,s}(r)dr + \frac{\langle N_s(N_s-1)\rangle}{2}\lambda_{s,s}(r)dr. \tag{37}$$

We assume that central-satellite pairs follow a pair distribution function $\lambda_{c,s}(r)dr = 4\pi\tilde{\rho}_{\rm NFW}(M_{\rm h}, r)r^2dr$, where $\tilde{\rho}_{\rm NFW}(M_{\rm h}, r)$ is the normalized NFW halo density profile. The satellite-satellite pair distribution, $\lambda_{s,s}(r)dr$, is then the normalized density profile convolved with itself, that is, $\lambda_{s,s}(r)dr = 4\pi\lambda_{\rm NFW}(M_{\rm h}, r)r^2dr$, where $\lambda_{\rm NFW}$ is the NFW profile convolved with itself. An analytic expression for $\lambda_{\rm NFW}(M_{\rm h}, r)$ is given by Sheth et al. [190]. Both $\tilde{\rho}_{\rm NFW}$ and $\lambda_{\rm NFW}$ depend on the halo concentration parameter, $c_{\rm NFW}$. N-body numerical simulations show that this parameter weakly anti-correlates with mass, $c_{\rm NFW} = a - b\times\log M_{\rm h}$, though with a large scatter.

Based on results of N-body [115] and hydrodynamic [262] simulations, we will assume that the number of satellite-satellite pairs follow a Poisson distribution with mean $\langle N_s\rangle^2 = \langle N_s(N_s-1)\rangle$. This is also supported by the analysis based on a large catalog of galaxy groups by Yang, Mo & van den Bosch [246].

For the two-halo term, where $r > 2R_h(M_{\rm h})$, all pairs must come from galaxies in separate halos. We compute the two-halo term from the non-linear matter correlation function, $\xi_m(r)$ following [196]:

$$\xi^{\rm 2h}_{\rm gg}(r) = b_g^2\zeta^2(r)\xi_m(r), \tag{38}$$

where $\zeta(r)$ is the scale dependence of dark matter halo bias [209, see their Eq. B7], and,

$$b_g = \frac{1}{n_g}\int_0^\infty b(M_{\rm h})\langle N(> M_*|M_{\rm h})\rangle\phi_{\rm h}(M_{\rm h})dM_{\rm h}, \tag{39}$$

Table 1: Constraints

Constraints:					
Data set	Satellite GSMF	Central GSMF	Total GSMF	2PCF	χ^2
A	YMB09	YMB09	—	✗	$\chi^2_{\phi_{\rm sat}} + \chi^2_{\phi_{\rm cen}}$
B	✗	✗	YMB09	Y12	$\chi^2_{\phi_{\rm all}} + \chi^2_{w_{\rm p}}$
C	YMB09	YMB09	—	Y12	$\chi^2_{\phi_{\rm sat}} + \chi^2_{\phi_{\rm cen}} + \chi^2_{w_{\rm p}}$
B1	✗	✗	BGD08	Y12	$\chi^2_{\phi_{\rm all}} + \chi^2_{w_{\rm p,bin}}$
Predictions:					
Data set	Satellite GSMF	Central GSMF	Total GSMF	2PCF	sat. CSMF
A	✗	✗	—	✓	✗
B	✓	✓	✗	✗	✓
C	✗	✗	✗	✗	✓
B1	✓	✓	✗	✗	✓

is the galaxy bias with $b(M_{\rm h})$ being the halo bias function [191].

Once we have calculated $\xi_{\rm gg}(r)$, we relate it to the projected two-point correlation function (2PCF), $w_{\rm p}(r_{\rm p})$, by

$$w_{\rm p}(r_{\rm p}) = 2\int_0^\infty \xi_{\rm gg}(\sqrt{r_{\rm p}^2 + x^2})dx. \tag{40}$$

In this model, Eqs. (20–40) relate the observed 2PCF to the central and satellite occupational number distributions, which on their own are related to the central and satellite–(sub)halo mass relations. In consequence, the correlation function is related to the total GSMF and its decomposition into centrals and satellites. Therefore, since the GSMFs and PCFs are tightly connected, any combination of these observational constraints is not expected to provide independent constrains on the mass relations and the occupational number distributions. However, we expect that the uncertainties in the determinations of these functions are reduced as more observational constraints are introduced. We will explore this question in more detail in §§4.1.

3.2.3 *Parameters in the model*

Ultimately, our model, which in total consists of ten free parameters –if σ_c, σ_s, and the $c_{\rm NFW} - M_{\rm h}$ relation are fixed– constrains the central and satellite stellar-to-(sub)halo mass relations. Five parameters are used to model the CHMR (Eq. 21): $M_{1,c}$, ϵ_c, α_c, δ_c, and γ_c; and five more to model the SSMR (and therefore the satellite occupational numbers): $m_{1,s}$, ϵ_s, α_s, δ_s, and γ_s. Note that the success of our model relies on the ability to choose a parametric description of the M_*–$M_{\rm h}$ and M_*–$m_{\rm sub}$ relations (Eq. 21), such that the observed total GSMF and its decomposition into centrals and satellites are well-reproduced. As discussed previously, the main motivation for the functional forms chosen here is that they are able to reproduce Schechter-like GSMFs accurately.

Using a SDSS halo-based group catalog, YMB09 found that the intrinsic scatter around the CHMR is approximately independent of halo mass and log-normally distributed, with a mean width of $\sigma_c(\log M_h)$=0.173 dex. This result is also supported by studies of satellite-galaxy kinematics [145, 146] and analysis using HOD models [245, 59, 45, 122]. Additionally, AMT results are able to reproduce the GSMF and the spatial clustering of galaxies simultaneously when using σ_c =const. (e.g., 148; RDA12; 170). On the other hand, the scatter σ_s around the SSMR has not yet been discussed in the literature. In RDA12 it is assumed to be the same as for central galaxies, giving results consistent with the observed projected 2PCFs and satellite CSMFs. Having said that, we assume the intrinsic scatters σ_c and σ_s to be independent of halo mass and equal to 0.173 dex. Nevertheless, as we have discussed, the constraints provided by the GSMF decomposed into centrals and satellites and the projected 2PCFs are not independent but rather they are complementary. Therefore, when using all these constraints, it may be possible to leave σ_c and/or σ_s as free parameters. We will perform this exercise in Section 3.3. Finally, for the relation of the concentration parameter $c_{\rm NFW}$ with mass, we use the fit to numerical simulations by Muñoz-Cuartas et al. [150].

3.2.4 *Observational data sets and strategy*

A combination of the total, central, and satellite GSMFs, and the projected 2PCF for different M_* bins are necessary to constrain our model. In the following, we will experiment with different combinations of these data. We wish to understand how the stellar–(sub)halo mass relations vary depending on the combination of observational data used to constrain them. In particular, we would like to explore whether the uncertainty in the model parameters drops significantly by introducing more observational constraints.

The different observational data to be used for constraining the model parameters are as follows:

- The YMB09 GSMF decomposed into central and satellite galaxies. These data were obtained from a large halo-based galaxy group catalog constructed in Yang et al. [249] from the SDSS DR4 (they define a central galaxy as the most massive galaxy in a group with the remaining galaxies being satellites). Both the central and the satellite GSMF are well-described by Schechter functions, with central galaxies being the more abundant population at all masses, at least above the low-mass limit of the sample, $\log(M_*/{\rm M}_\odot)$ = 8.4.

- The total BGD08 GSMF, which is well described by a double Schechter function. This GSMF is steeper at the low-mass end than the YMB09 GSMF (See also 75). BGD08 have actually extended the GSMF to a lower limit, $\log(M_*/{\rm M}_\odot)$ = 7.4, by introducing a surface-brightness completeness correction.

- The projected 2PCFs determined in five M_* bins by Yang et al. [250, hereafter Y12] based on the SDSS DR7.

The combinations explored to constrain the model parameters consists of four data sets:

Set A consists of the YMB09 central and satellite GSMFs, and is used to constrain the model parameters of our extended AMT; in this case, the projected 2PCFs in various mass bins are predicted. Set B consists of the *total* YMB09 GSMF and the Y12 projected 2PCFs, and this is used to constrain our full combined model; the GSMF decomposed into centrals and satellites is a prediction. Set B1 is similar to set B but instead of the YMB09 GSMF, the Baldry, Glazebrook & Driver [16, hereafter BGD08] GSMF is used. Set C consists of all the available data: the YMB09 GSMF *decomposed into centrals/satellites* and the Y12 projected 2PCF determined in different M_* bins; this data set over-constrains our full combined model.

We notice that fiber collisions in the SDSS data underlying the group catalog may introduce an extra source of uncertainty when using the satellite GSMF for constraining the parameters in sets A and C. However, this seems to be a small effect at most since YMB09 show that satellite CSMFs with a correction for fiber collisions are only marginally different. It is also important to highlight that the authors report only the diagonal elements of the covariance matrix for the projected 2PCFs. We expect that the full covariance matrix would reduce possible systematic errors and extra uncertainties in some of the constrained parameters. As discussed in §3.1, the lack of the covariance matrix seems to affect the results for the abundance of satellite galaxies, however, these effects are of minor importance.

Table 1 summarizes the different data sets presented above and specifies where the observables are used as constraints and in where they are predicted by the model.

We use Markov Chain Monte Carlo (MCMC, see Appendix B) methods for sampling the best fit parameters that maximize the likelihood function $\mathcal{L} \propto \exp(-\chi^2/2)$. Each MCMC chain consist of 1.5×10^6 elements. See Appendix C for details on the full procedure.

3.3 THE ANALYSIS

In our model the central/satellite GSMFs are tightly connected to the projected 2PCFs in such a way that given one the other can be inferred (§§3.1). This connection passes through the underlying halo/subhalo statistics and the stellar-to-(sub)halo mass relations. Therefore, the latter, together with the satellite CSMF, are predictions in all cases (§§3.2 and §4, respectively).

3.3.1 *The GSMFs & projected 2PCFs*

Figure 7 shows the model results for the central, satellite, and total GSMFs, while Fig. 8 shows the projected 2PCFs in different mass bins. The observational data are also plotted in these figures (symbols with error bars). The shaded regions in the figures are the resulting model-fit standard deviations calculated from the 1.5×10^6 MCMC models for sets B and B1 (the upper and lower panels of Fig. 7, respectively), and for set A (Fig. 8). These standard deviations are associated with the uncertainties in the model parameters, and are produced partially by the uncertainties in the observations used to constrain the model. For a discussion on the model scatter see Sect. 5.1.

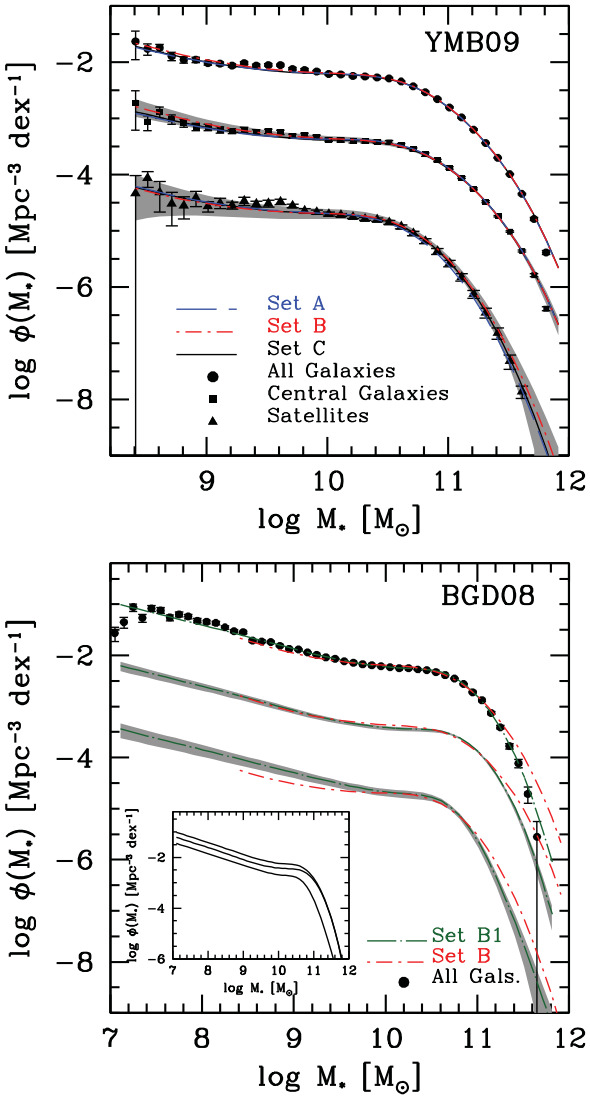

Figure 7: The GSMF for all, central, and satellite galaxies. For clarity, in each panel the central GSMF was shifted down by 1 dex, and the satellite GSMF by 2 dex. *Upper panel:* Model results for sets A (blue long dashed line), B (red dot-short-dashed line), and C (solid line), compared to the observed YMB09 GSMFs for all (filled circles), central (filled squares), and satellite galaxies (filled triangles). For sets A and C the curves are just the best joint fit to the data, while for set B are model predictions. The shaded areas correspond to the standard deviation of the 1.5×10^6 MCMC models for set B. *Lower panel:* Same as upper panel but for the set B1 (green dot-long-dashed line and shaded areas). The predictions for set B are repeated in this panel (red dot-dashed line). The corresponding observational total BGD08 GSMF is showed with solid circles and error bars. The inset shows how the central and the satellite GSMFs add up to give the total GSMF in the case of set B1.

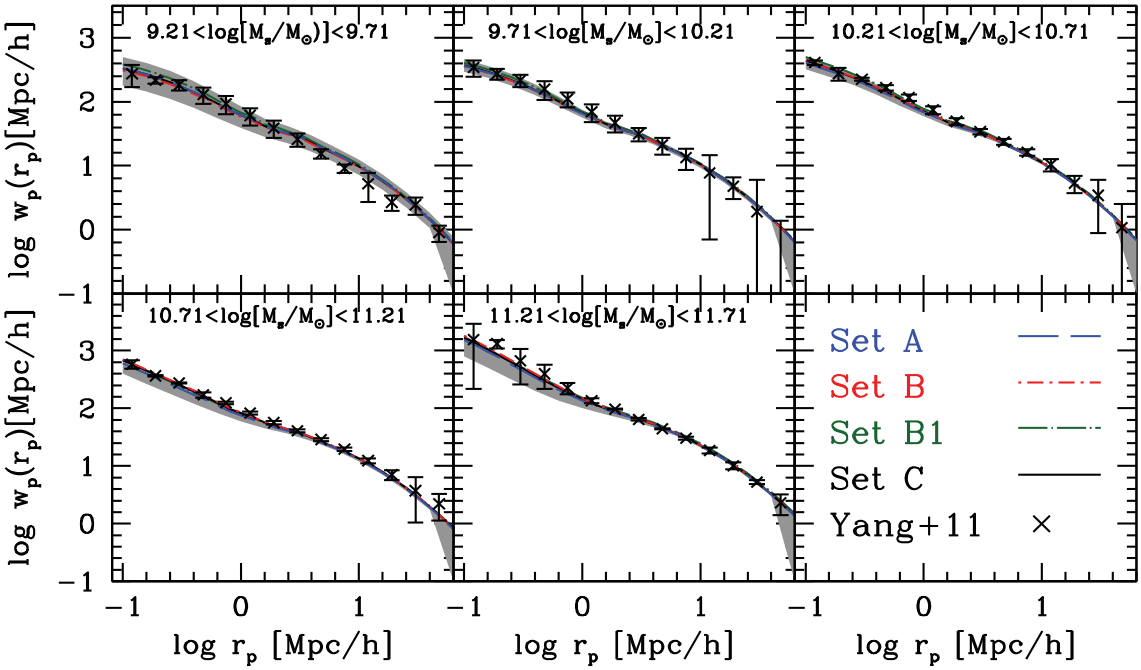

Figure 8: Projected 2PCFs in five stellar mass bins corresponding to the data sets A, B, C, and B1 (see Table 1), and to the Y12 observational determinations (crosses with error bars). The lines corresponding to each set are indicated in the last panel. For set A, the plotted 2PCFs are predictions, while for the rest of the sets are just the joint best fits. The shaded areas show the standard deviation for the 1.5×10^6 MCMC models.

For set A, which is constrained by the YMB09 satellite and central GSMFs (solid symbols in Fig. 7; red dot-dashed curves are just the joint best fits to data), the model predicts the projected 2PCFs (red dot-dashed curves with shaded areas in Fig. 8). Both the amplitude and the shape of the predicted projected 2PCFs are in excellent agreement with observations (crosses with error bars, Y12) at each stellar mass bin plotted in Fig. 8. This result is not surprising as shown in RDA12. What is interesting, however, is that the standard deviations are consistent with the errors reported in the observations. Note that the 1-halo term is the zone with the largest uncertainty, which arises directly from the uncertainty in the satellite GSMF. For set B, which is constrained by the total YMB09 GSMF and the Y12 projected 2PCFs (black curves are just the joint best fits to data), the model predicts the central and satellite GSMFs (black curves and gray shaded regions in the upper panel of Fig. 7). The model predictions agree very well with observations, and therefore with set A. These results show that *the central/satellite GSMFs and the 2PCFs are tightly connected in such a way that given one, the other can be inferred through our model.* Observe that in Fig. 7, the standard deviations are consistent with the error bars both for the satellite and the central GSMF. The former has the largest uncertainties. Therefore, the lack of information from the projected 2PCF covariance matrix seems to affect mostly the abundance of satellite galaxies or equivalently, the 1-halo term in the projected 2PCF.

Set B1 (lower panel of Fig. 7) is similar to set B, but the BGD08 total GSMF is used as a constraint instead of YMB09 data. Therefore, the model total GSMF and 2PCFs (green dot-long-dashed curves in Figs. 7 and 8, respectively) are just the joint best fits

Table 2: Fit parameters

Central galaxies:										
Data set	$\log M_{1,c}$	stdev	$\log \epsilon_c$	stdev	α_c	stdev	δ_c	stdev	γ_c	stdev
A	11.477	0.073	-1.582	0.050	2.252	0.461	3.558	0.206	0.485	0.044
B	11.480	0.066	-1.580	0.038	1.982	0.338	3.530	0.198	0.491	0.040
C	11.493	0.068	-1.600	0.047	2.138	0.417	3.572	0.202	0.487	0.043
B1	11.676	0.056	-1.475	0.027	2.056	0.110	2.454	0.183	0.514	0.047
Satellite galaxies:										
Data set	$\log m_{1,s}$	stdev	$\log \epsilon_s$	stdev	α_s	stdev	δ_s	stdev	γ_s	stdev
A	10.761	0.069	-0.992	0.063	2.469	0.710	3.616	0.260	0.435	0.077
B	10.773	0.088	-0.951	0.052	2.670	0.792	3.612	0.255	0.437	0.075
C	10.775	0.064	-0.957	0.052	2.474	0.657	3.586	0.260	0.423	0.071
B1	11.017	0.90	-0.709	0.044	2.322	0.191	1.667	0.225	0.993	0.133

to the data, but the central and satellite GSMFs are predicted (green dot-long-dashed curves with gray shaded areas in the lower panel of Fig. 7). Note that the BGD08 GSMF extends to lower stellar masses. The resulting slope of the satellite GSMF at the faint end ($\log(M_*/\mathrm{M}_\odot) \lesssim 9.6$), is $\alpha \sim -1.5$, which is steeper than set B, $\alpha \sim -1.2$, and the bump around $\mathcal{M}^* \sim 8 \times 10^{10}\mathrm{M}_\odot$ in the central GSMF is more pronounced. A steeper total GSMF implies a major contribution of satellite galaxies to the total GSMF at low masses.

Finally, for set C, neither the central/satellite GSMFs nor the projected 2PCFs at different mass bins are predicted but rather employed to constrain the model parameters. Therefore, the blue long-dashed curves shown for set C in Figs. 7 and 8 are just the joint best fits to the observations; they are not predictions. For this set, the predictions are the constraints on the stellar-to-(sub)halo mass relations. The question now is how different can these relations and their uncertainties be from those inferred using the other data sets.

3.3.2 *Mass relations*

In the upper panels of Fig. 9 we plot the central, satellite, and average stellar-to-(sub)halo mass ratios (stellar mass fractions, f_*) as a function of the (sub)halo mass obtained for each data set listed in Table 1. The stellar mass fractions are obtained directly from the corresponding mass relations. Table 6.2.2 lists the best fit MCMC model parameters of these relations for each of the sets. We use Eq. (34) to compute the average f_*. Recall that the average relation, $\langle \log M_*(M) \rangle$, is conceptually what is commonly obtained with the standard AMT. However, in the latter case it is not possible to distinguish the mass

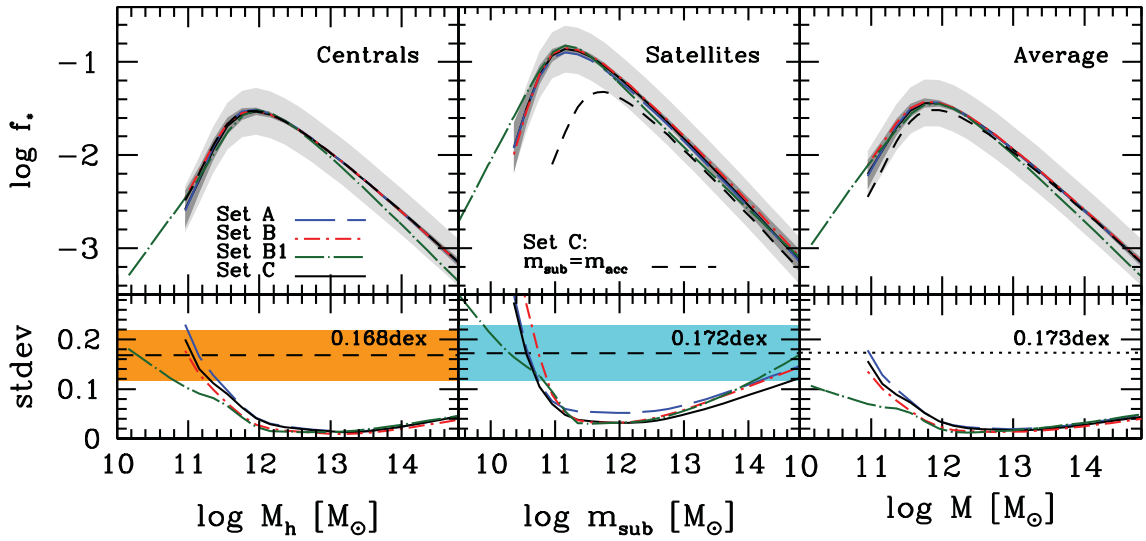

Figure 9: *Upper panels:* From left to right, the constrained stellar mass fractions of central and satellites, and of the number-density average (Eq. 34) of both. The lines corresponding to each set are indicated inside the panels. Short-dashed curves in the second and third panels are the constrained mass relations when the subhalo mass is defined at its accretion time. The systematic uncertainty due to stellar mass determination (0.25 dex) is shown with the light-gray shaded areas. Gray light dashed areas indicate the MCMC model-fit standard deviation in the case of set C. *Lower panels:* The MCMC model-fit standard deviations for each data set. The short dashed lines in the right and middle lower panels are the intrinsic scatters, σ_c and σ_s, constrained for the set C assuming them to be constant. The color shaded area show the standard deviations of these values.

relations for centrals and satellites, and it is common practice to assume them equal. As shown in RDA12, this assumption is not correct.

In general, we find that the shape of the stellar fractions for both the centrals/halos and satellites/subhalos rises steeply at low masses, reaching a maximum and then declines roughly as a power law towards higher masses. We do not find significant differences among the stellar mass fractions obtained for sets A, B and C. Observe how all of them lie well within the 1σ uncertainty which is dominated by the systematic uncertainty in the stellar mass determination ($\sim$ 0.25 dex, light shaded area in Fig. 9; see 19). All these relations even lie well within the standard deviation of the MCMC models, shown as a dark gray shaded area in set C (the others being very similar). In the lower panels of Fig. 9 we plot the the standard deviations as a function of (sub)halo mass for each set, which we discuss in detail in §5.1.

We arrive at two important implications: (1) the very small standard deviations obtained for the stellar-to-(sub)halo mass relations implies that the assumption that on average there is a monotonic relation between galaxy and (sub)halo masses is consistent with the data; (2) the result that set A and set B lead to very similar mass relations confirms that *matching abundances is equivalent to matching occupational numbers and vice versa*, as suggested in RDA12. Therefore, constraining the model parameters with all the observational information, as in set C, should lead again to the same stellar-to-(sub)halo mass relations as in sets A and B. Indeed, this is what we obtain.

The central f_*–M_h relations for sets A–C at the low (high) mass end scale roughly as $f_* \propto M_h^{1.5}$ ($f_* \propto M_h^{-0.7}$). The average f_*–M_h relations are such that they lie above but closer to centrals, simply because they are the dominant population. Instead, the satellite f_*–m_sub relations are quite different to centrals, both in the amplitude and in the location of the maximum of f_*. The maximum shifts from $\log(M_\mathrm{h}/\mathrm{M}_\odot) \approx 11.9$ to $\log(m_\mathrm{sub}/\mathrm{M}_\odot) \approx 11.2$. These differences are basically due to the fact that subhalos lose mass due to tidal striping (on average 60–65% of the mass since accretion for subhalos hosting satellites less massive than $\sim 2 \times 10^{11}$ $\mathrm{M}_\odot$; RDA12; see also 216, 234, 230). However, even when the subhalo mass at accretion time is used, some differences remain, showing that the assumptions about evolution made in order to construct the nominal SSMR for this case are roughly but not exactly obeyed.

In Fig. 9 we plot the model results for subhalo mass defined at accretion time for the set C (black dashed curve)[2]. The f_*–m_sub (or SSMR) relation now lies close to the central f_*–M_h (or CHMR) relation. Recall that for connecting the present-day observed satellite M_* to (sub)halo masses at their different accretion epochs, one implicitly assumes that the satellite stellar masses change in a way that at $z = 0$ the SSMR is equal to the CHMR. Our ignorance about how the satellite masses evolve introduces an extra uncertainty in the determination of the SSMR when the subhalo mass at the accretion time is used (Y12; RDA12). In any case, as extensively discussed in RDA12, for one or another definition of subhalo mass, there is a unique but different average SSMR for which the satellite GSMF and CSMF, and the correlation functions are in agreement with observations. Nevertheless, when the SSMR is assumed to be equal to the CHMR, the predicted satellite GSMF and CSMF, and correlation functions depart from observations. They do so more strongly for the observation-time definition, m_obs, and less strongly for the accretion-time definition, m_acc(see Figs. 1, 3, and 4 in RDA12).

The mass relations for set B1 (green dot-dashed curves) are somewhat different to those of set B: at high masses the central (and the average) f_*–M_h relation is steeper than in set B, and at low masses the satellite f_*–m_sub relation is shallower. Recall that the massive-end of the total BGD08 GSMF (set B1) decays faster than in the YMB09 GSMF (set B; see Fig. 7). Consequently, at a fixed M_h the CHMR for set B1 is systematically lower than for set B. At low masses, the GSMF in set B1 is steeper than in set B, causing this a steeper satellite GSMF and therefore a shallower decay of M_* as m_sub decreases as compared to set B.

3.3.3 *Constraining the intrinsic scatter of the stellar-to-(sub)halo mass relations*

The results presented above are obtained under the assumption of lognormal intrinsic scatter around the CHMR and the SSMR with constant 1σ widths of 0.173 dex for both relations. Our model is over-constrained by observations in set C. Therefore, we may leave one or both of the intrinsic scatters as free parameters, but keep the assumption that they are constant, that is, independent of (sub)halo mass.

The results of leaving only one of σ_c or σ_s free are very similar to leaving both free at the same time. The MCMC algorithm in the latter case constrains the intrinsic scatters

2 The SubhCMF for subhalos defined at the accretion time given by Boylan-Kolchin et al. [41, see also Giocoli, Tormen & van den Bosch [94]] has been used (see RDA12 for details).

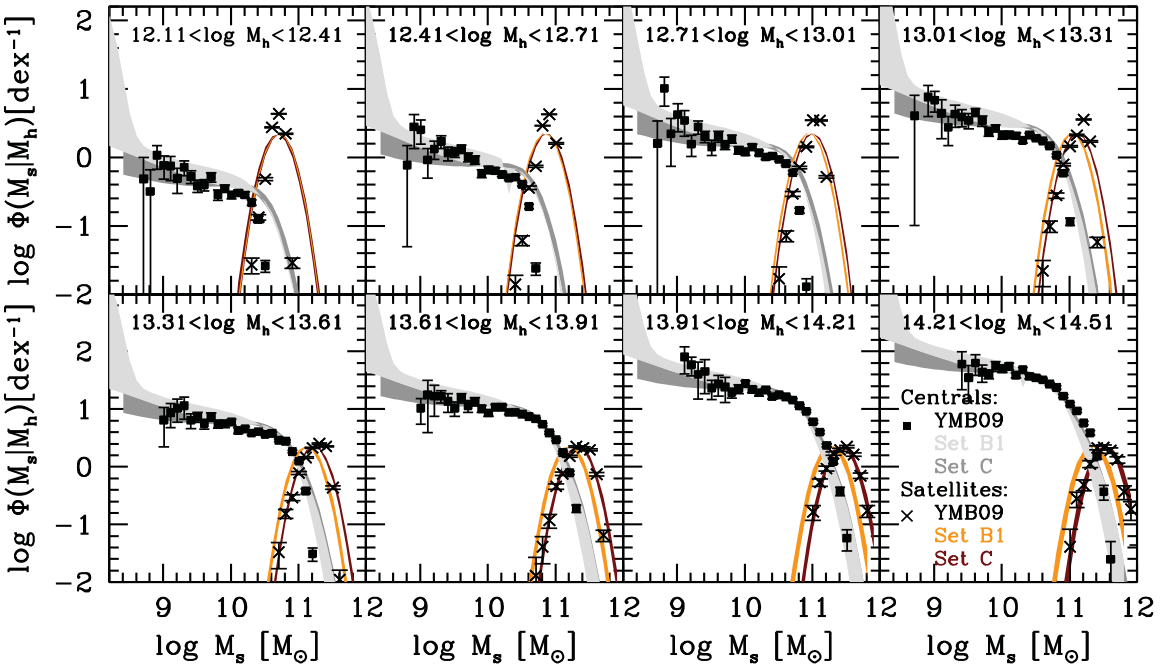

Figure 10: Central mass probability distributions and satellite CSMFs, $P_{\rm cen}$ and $\Phi_{\rm sat}$, in eight halo mass bins for set C (dark red and dark gray shaded areas, respectively) and set B1 (orange and light gray shaded areas, respectively). The shaded areas correspond to the standard deviation of the MCMC model-fits for each set, which for $P_{\rm cen}$ are actually very thin. The observational inferences by YMB09 are plotted with crosses for $P_{\rm cen}$, and with filled squares for the satellite CSMFs. Their halo masses were converted to match our virial definition.

to be $\sigma_c = 0.168 \pm 0.051$ dex and $\sigma_s = 0.172 \pm 0.057$ dex. These values are surprisingly close to those we have assumed. These values are plotted in the lower panels of Fig. 9. The constrained mass relation parameters also remain almost the same. So, under the assumption of lognormal distributed and constant intrinsic scatters, our results confirm previously estimated values of the scatter for central galaxies, and predict similar values for the scatter around the mass relation of satellite galaxies.

3.3.4 *The conditional stellar mass functions*

In Fig. 10, we plot the resulting central galaxy mass probability distributions, $P_{\rm cen}$, and the satellite CSMFs, $\Phi_{\rm sat}$, in eight halo mass bins both for set C (dark red and dark gray areas, respectively) and set B1(orange and light gray areas, respectively). Because the predictions for sets A and B are very similar to those of set C we do not plot them separately. In fact, what is plotted in Fig. 10 are the standard deviations (scatters) of the MCMC models for each set, which for $P_{\rm cen}$ are actually very small. $P_{\rm cen}$ is the probability distribution for a halo of a fixed mass to host a central galaxy of a given stellar mass (eq. 20), while $\Phi_{\rm sat}$ refers to the mean number of satellite galaxies residing in a host halo of a fixed mass (eq. 28). We compute $P_{\rm cen}$ averaged in each $[M_{\rm h_1}, M_{\rm h_2}]$ bin as:

$$\langle P_{\rm cen} \rangle = \frac{\int_{M_{\rm h_1}}^{M_{\rm h_2}} P_{\rm cen}(M_*|M_{\rm h})\phi_{\rm h}(M_{\rm h})dM_{\rm h}}{\int_{M_{\rm h_1}}^{M_{\rm h_2}} \phi_{\rm h}(M_{\rm h})dM_{\rm h}}, \tag{41}$$

while for satellites, the averaged CSMF is given by:

$$\langle \Phi_{\rm sat} \rangle = \frac{\int_{M_{\rm h_1}}^{M_{\rm h_2}} \Phi_{\rm sat}(M_*|M_{\rm h})\phi_{\rm h}(M_{\rm h})dM_{\rm h}}{\int_{M_{\rm h_1}}^{M_{\rm h_2}} \phi_{\rm h}(M_{\rm h})dM_{\rm h}}. \tag{42}$$

As seen in Fig. 10, the smaller the halo mass, the larger is the stellar mass gap between the most common central galaxy and the most abundant satellites. In other words, on average, as smaller is the halo, the larger is the ratio of the central galaxy mass to the masses of the satellite population.

In Fig. 10 we also show the corresponding observational results by YMB09 for centrals (crosses) and satellites (filled squares). The agreement between the model predictions for set C and the observational data is remarkable. However, some marginal differences are observed. As the halo mass decreases, the width of the central probability distribution is systematically somewhat broader, and therefore its amplitude is lower compared with the YMB09 data. This could be due to the assumption that the intrinsic scatter σ_c is independent of $M_{\rm h}$. There are some pieces of evidence that σ_c slightly depends on $M_{\rm h}$ as discussed in Yang, Mo & van den Bosch [247, see also 263].

Regarding the satellite CSMFs, the abundance of very massive satellites is slightly, but systematically, overestimated for $M_{\rm h} \lesssim 10^{13}{\rm M}_\odot$. This was noted already in RDA12, which suggest that a possible reason is due to the assumption that the intrinsic scatter σ_s is independent of $M_{\rm h}$. We have explored this possibility, and found that as $M_{\rm h}$ decreases, the scatter σ_s does tend to zero in order to reproduce the observations. Such a behavior

is not expected at all. Another possibility, and the most likely, is that the YMB09 satellite GSMF is underestimated (e.g., 194; RDA12), although the results from set B indicate that the obtained satellite GSMF is consistent with observations. In any case, the excess of massive satellites in low mass halos does not contribute significantly to the total mean density of galaxies.

For set B1, we observe that $P_{\rm cen}$ is shifted to slightly lower values of M_* as the halo mass increases when comparing with observations (and set C). This is a consequence of the observed trends of the M_*–$M_{\rm h}$ relations between set B1 and C (see Fig. 9), and it is ultimately related to the fact that the high-mass end the BGD08 GSMF decreases faster than the YMB09 GSMF. On the other hand, the satellite CSMFs for set B1 are slightly steeper at low stellar masses to those of set C. This is a consequence of the BGD08 GSMF being steeper than that of YMB09 at low masses. Note that the uncertainty in the CSMFs for set B1 dramatically increases at the lowest masses. This is because at these masses there is no information on the 2PCF, so that the total GSMF alone poorly constraints the CSMFs. Set B1 also overestimates the abundance of massive satellites in low mass halos.

3.4 DISCUSSION

3.4.1 *Robustness and model uncertainties of the stellar-to-(sub)halo mass relations*

The main result from Section 3.3 is that both the inferred central-halo and satellite-subhalo mass relations do not change for all the combinations of data sets we explored. In other words, these relations seem to be determined robustly, no matter whether only the central/satellite GSMFs (set A) or whether only the total GSMF and the 2PCFs (set B) or whether all of these data (set C) are used. The results confirm what is expected: the GSMFs of central and satellites galaxies are well connected with the 2PCFs and both are part of a general statistical description of the galaxy population. However, it could be that the uncertainties around the mass relations depend on the set of observables used. In particular, we expect that the uncertainties should be smaller when all the observational data are used to constrain the model.

From the results of the MCMC search over 1.5×10^6 models we can identify at each (sub)halo mass the average M_* and its standard deviation. The average stellar masses for a given (sub)halo mass are indistinguishable to those given by the average stellar-to-(sub)halo mass relations constructed with the best fit parameters obtained with the MCMC method. The standard deviations can be interpreted then as the 1σ model-fit uncertainty around these relations [see also 146]. This uncertainty is due to (i) the inability of the proposed stellar-to-(sub)halo mass relations (Eqs. 21) to reproduce jointly the observational data, and (ii) the observational errors in these data. The dark gray areas in Figs. 9 and 11 correspond to the standard deviations for the set C; the much wider light gray areas show the scatter of 0.25 dex attributed to the systematic uncertainty in the determination of the stellar mass [19].

How do the model-fit uncertainties in the stellar-to-(sub)halo mass relations for set C compare with the other sets? To our surprise, the uncertainties in sets A, B, and B1 are as small as for set C. Actually, the uncertainties are smaller than the intrinsic scatter between galaxy and halo mass at least for halo masses larger than $\sim 10^{11}$ $M_\odot$. The

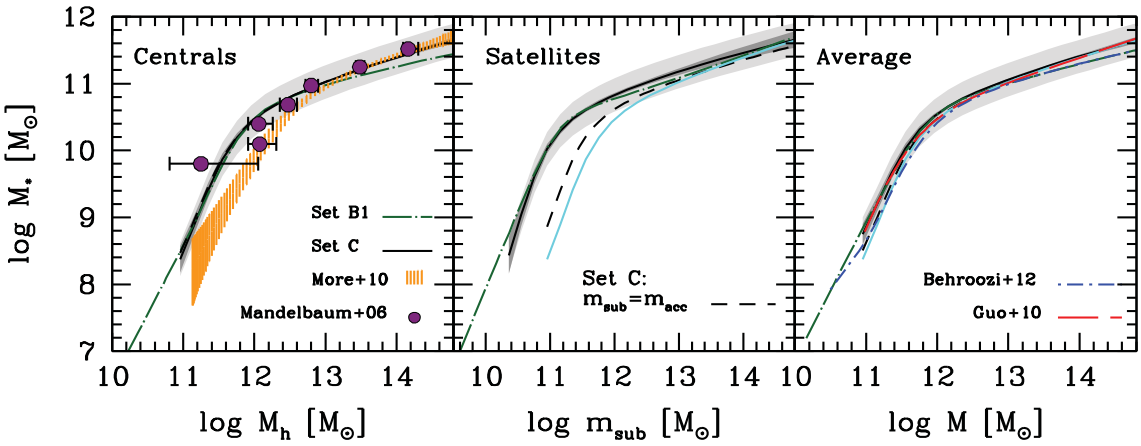

Figure 11: Central-halo, satellite-subhalo, and average mass relations for sets C (solid line) and B1 (green dot-dashed line). Short-dashed curves in the medium and right panels are the mass relations when subhalo mass is defined at its accretion time. The systematic uncertainty due to stellar mass determinations is show with the light-gray shaded area. Gray dashed area indicates the standard deviation of the MCMC model fits in for set C. Filled circles with error bars correspond to the mass relation of central galaxies from the analysis of staked weak-lensing by Mandelbaum et al. [133]. Orange dashed area indicates the 68% of confidence in the mass relation of central galaxies using the kinematics of satellites [146]. Abundance matching results reported in Behroozi, Wechsler & Conroy [20], and Guo et al. [102] are plotted with the blue dot-short-dashed and the red long-dashed lines. For comparison we have plotted with the cyan solid lines in the middle and right panels the central-halo mass relations for set C. Observe how in the middle panel the SSMR for the subhalo mass defined at the accretion time lies above the CHMR by a factor of ~ 3, while in the right panel the nominal average mass relation at the accretion time is a factor of ~ 1.25 higher than the CHMR.

lower panels of Fig. 9 show the MCMC standard deviations as a function of (sub)halo mass for all the sets studied here. If any, the major differences are for the uncertainties in the SSMR: they become large at large subhalo masses and are larger for set A (and B1) and smaller for set C. At the smallest (sub)halo masses the model uncertainties for all the sets increase significantly (but yet below the systematic uncertainty of 0.25 dex). This is related to the larger observational errors at smaller stellar masses both for the GSMFs, in particular the one for satellites, and the projected 2PCFs. The small model-fit uncertainties obtained in the determination of the central-halo and satellite-subhalo mass relations through our model again lead us to conclude that these determinations are robust.

How do the stellar-to-(sub)halo mass relations compare with previous works? The most direct (but highly uncertain) methods to infer the halo masses of galaxies for large samples of objects are through weak lensing and satellite kinematics. In the left panel of Fig. 11, we reproduce the results for *central* galaxies of $\langle \log(M_*) \rangle$ as a function of $M_{\rm h}$ by using stacked satellite kinematics [146, Dr. S. More kindly provided us the data in electronic form] and of $\langle M_{\rm h} \rangle$ as a function of M_* by using stacked weak lensing analysis[3] [133]. The latter authors inferred the CHMR separated into late- and early-type galaxies, $\langle M_{\rm h} \rangle_l(M_*)$ and $\langle M_{\rm h} \rangle_e(M_*)$, respectively. We compute the average $\langle M_{\rm h} \rangle(M_*)$ relation for all central galaxies as

$$\langle M_{\rm h} \rangle(M_*) = f_l(M_*) \langle M_{\rm h} \rangle_l(M_*) + f_e(M_*) \langle M_{\rm h} \rangle_e(M_*), \tag{43}$$

where $f_l(M_*)$ and $f_e(M_*)$ are the fraction of late- and early-type galaxies of stellar masses M_* in the sample. Note also that for the More et al. [146] and Mandelbaum et al. [133] results, small corrections in $M_{\rm h}$ were applied for consistency to our definition of virial mass, as well as in M_* to be consistent with the Chabrier [48] IMF adopted here.

As seen in Fig. 11, the Mandelbaum et al. [133], weak-lensing determinations are consistent with our CHMR. However if one takes into account that their dependence of $\langle M_{\rm h} \rangle$ on M_* would be flatter at high masses in case it is deduced from the inverse relation (see footnote), then our determination for set C would be steeper. Instead, the results for set B1 would probably be in better agreement with Mandelbaum et al. [133] at high masses. The More et al. [146] satellite-kinematics determinations are consistent with our results for masses larger than $M_{\rm h} \sim 4 \times 10^{12}$ $M_\odot$, but at smaller masses their amplitude can be 2-3 times lower. This discrepancy between satellite kinematics and other methods has been noted previously (e.g., 146, 194, 177).

In the right panels of Fig. 11 we compare the average stellar-to-halo mass relation (Eq. 34) obtained for set C with those of Guo et al. [102] (red long-dashed curve) and by Behroozi, Wechsler & Conroy [20] (blue dot-dashed curve). These authors obtained their relations by matching the abundances of all galaxies to abundances of halos plus subhalos. In general, our average mass relation is consistent with these previous global AMT results, though a direct comparison might not be fair, because we do not assume

3 As widely discussed in Behroozi, Conroy & Wechsler [19], due to the scatter, the inferences are slightly different depending on whether M_* is constrained as a function of $M_{\rm h}$ or viceversa. As these authors show, the main difference is at the high-mass end, in the direction that the CHMR being flatter if is inferred as a function of $M_{\rm h}$.

that the mass relations for centrals–halos and satellites–subhalos are the same. Also, the definition of subhalo mass used here (at the time of observation) is different to the one used in the above papers, who define it at the time of accretion. Hence, we also plot our SSMR for the subhalo mass defined at the accretion time in Fig. 11 (dashed curve in the central panel, set C) and the corresponding average mass relation (dashed curve in the right panel). The nominal SSMR is close but not equal to the present-day CHMR. We found that the nominal SSMR lies above the CHMR at most by a factor of ~ 3, while the nominal average mass relation is a factor of ~ 1.25 higher than the CHMR. To establish the former relation, one should know how the CHMR (at the accretion time, the satellite is still a central galaxy and the subhalo is a distinct halo) changes with time, and how the satellite mass evolved since accretion. Assuming that the CHMR is the same at all epochs leads to the nominal SSMR to be equal to the CHMR (RDA12). The fact that we find both relations to be close (but not equal) implies then that the galaxy–halo connection changes only little with time. This seems to be also the situation in cosmological simulations (193; 71).

For set B1, which uses the BGD08 GSMF, the CHMR changes slightly its slope at low masses, while the SSMR becomes systematically shallower than in the case of set C. This is because the BGD08 GSMF becomes steeper at lower masses. However, when the density-weighted average is calculated, the slope change seen for the centrals is almost smeared out. For the Behroozi, Wechsler & Conroy [20] total (average) mass relation the slope change is present, presumably because the contribution of the satellite-subhalo mass relation is not taken into account. On the other hand, if we use the subhalo mass at the accretion time instead that at the observation time, then the smearing-out of the slope change is less evident.

3.4.2 *The satellite-subhalo mass relation at the low-mass end*

An interesting question is how to extend the GSMFs and stellar-to-(sub)halo mass relations towards low masses, since most potential issues with the ΛCDM scenario are happening at small scales. As recently discussed by Boylan-Kolchin, Bullock & Kaplinghat [38], the ΛCDM scenario can be compatible with the overall abundance of MW satellites, but it predicts subhalos that are too massive (or too concentrated) compared to dynamical observations of the brightest dwarf spheroidal (dSph) satellites. This can be visualized using the dSph stellar mass vs. subhalo maximum circular velocity (or mass) diagram, comparing the observations for the bright MW dSphs with extrapolations of *total* (centrals+satellites) abundance matching results to low masses. For a given M_*, the MW dSphs have subhalo circular velocities (or masses) much larger (by $\sim 1.5-2$ dex in mass) than the extrapolated AMT results.

Our model has the advantage that it allows to constrain the CHMR and the SSMR separately (Fig. 9). The extrapolation of the latter only is what actually should be used for comparison with the MW satellites. RDA12 show that if the faint-end extrapolation of the GSMF is as steep as -1.6 (BGD08; 75) and is completely dominated by satellite galaxies, then the ΛCDM subhalo masses are consistent with the subhalo masses of the observed MW dSphs. Here, masses are defined at the estimated tidal radii of the dwarf satellites (see Fig. 2 in RDA12 and references therein). By using our model, we are able

to decompose the BGD08 GSMF into satellites and centrals (set B1; Fig. 7). The faint-end slope of the satellite GSMF (down to $\sim 2.5 \times 10^7$ $M_\odot$) indeed resembles the total mass function, *but satellites do not dominate over centrals*. Therefore, the inferred SSMR gives still too large subhalo masses (by 0.3–0.4 dex) as compared with the tidal masses of the MW dwarfs. We should note that in set B1, the Y12 projected 2PCFs are used, and for stellar masses smaller than reported in Y12 ($\sim 1 \times 10^9$ $M_\odot$), no projected 2PCF constraints are applied. There are some hints that the projected 2PCFs of galaxies at small distances (one-halo term, where satellites dominate) are steeper than those measured in YMB09, especially for the smallest galaxies [124]. If this is the case, then we can easily show that the satellites become more abundant in the GSMF and the SSMR is flatter at low masses, leading to a better agreement with the inferred tidal (subhalo) masses for the MW dSph satellites.

3.4.3 *Interpreting the bump of the GSMF*

Several interpretations of the shape of the total GSMF have been offered in the literature [16, 75, 124, 34, 165]. In this section we will focus on interpreting the shape of the GSMF using arguments based on the occupation statistics of galaxies within halos. Looking at Fig. 10 (see also Fig. 12 in Chapter 4), it becomes apparent that as the halo mass increases, the likelihood of finding at least one satellite with a stellar mass *similar* to that of the central galaxy increases rapidly. Also, the stellar mass range covered by the satellite population is narrower and closer to that of the central as the halo mass increases. Assuming that these features of the satellite population mass distribution are robust and have been in place since the assembly of the central, it follows that the central's probability of growing by accreting large (compared to itself) satellites was largest in high-mass halos that today occupy the bump and high-mass end of the mass function.

3.5 CONCLUSIONS

An statistical model that combines the AMT with the HOD and CSMF formalisms is presented. The model allows to constrain the central-halo and satellite-subhalo mass relations (CHMR and SSMR) separately, as well as the satellite CSMFs inside the halos. The ΛCDM halo mass function and subhalo conditional mass functions were used as input. From the observational point of view, the model works with the total GSMF and its decomposition into centrals and satellites, and the 2PCFs. Therefore, the observations used to constrain the model can be different combinations of data: either the central/satellite GSMFs (from YMB09; set A), or the total GSMF (from YMB09 or BGD08) and the Y12 projected 2PCFs (sets B and B1), or all the data, i.e., the GSMFs of centrals and satellites and the projected 2PCFs (set C). Our aim was to explore how sensitive are the determinations of the mass relations and their uncertainties to the different data set used to constrain the model, as well as to test the overall consistency of the observations with the ΛCDM halo/subhalo mass functions. Related to the latter, we explored model predictions regarding some satellite number distributions. The main conclusions are:

• The constrained parameters of the CHMR and SSMR are almost identical for all sets of data, showing that these relations (and therefore, also the satellite CSMFs) are robust with respect to different combinations of the data used to constrain the model. To our surprise, even the model-fit uncertainties in the constrained stellar-to-(sub)halo mass relations are very similar for the different combinations of data sets, including the one where all the data are used (set C). These uncertainties are smaller than the assumed intrinsic scatters (0.173 dex) for $M_{\rm h} \gtrsim 10^{11}$ $M_\odot$, and of that order for smaller masses where the observational determinations of the GSMFs and projected 2PCFs have larger errors.

• For set A, the projected 2PCFs are predictions, while for set B (and B1), the GSMF decomposition into centrals and satellites are predictions. In each case, these predictions agree very well with the observations. This shows that matching central/satellite and (sub)halo *abundances* (set A) is equivalent to matching central/satellite and (sub)halo *occupational numbers*, in which case the 2PCFs are necessary (sets B, B1), and vice versa. In both cases, the CHMR and SSMR are intermediate relations. The key novelty in our model is that both relations are constrained separately instead of being assumed equal. Our results show also that the satellite/central GSMF is tightly connected to the spatial clustering of the population, both at the level of the one- and two-halo terms, as well as with the satellite mass functions inside the halos.

• For set C, neither the projected 2PCFs nor the GSMF decomposition are predictions, instead observational determinations of these functions are used to constrain the model. This allows us to leave the widths of the intrinsic scatter around the CHMR and SSMR (assumed independent of mass and log-normally distributed) as free parameters. We obtain $\sigma_c = 0.168 \pm 0.051$ dex and $\sigma_s = 0.172 \pm 0.057$ dex. For centrals, our result confirms previous estimates, and for satellites we find that the intrinsic scatter is almost the same as for centrals.

• The satellite-subhalo mass relation, where subhalo masses are defined at the observation time, is not equal to the central-halo relation. For the former, the stellar mass scales as ${M_{\rm h}}^{2.5}$ at the low mass-end and as ${M_{\rm h}}^{1.7}$ at the high-mass end (set C), while for the latter, these scalings go as $M_* \propto {M_{\rm h}}^{2.9}$ and ${M_{\rm h}}^{1.7}$, respectively. This difference is mainly due to the fact that subhalos lose mass (60-65%) due to tidal striping. When $m_{\rm sub}$ is defined at the accretion time, the nominal SSMR is actually close to the CHMR but again not equal. The SSMR lies above the CHMR at most by a factor of ~ 3, while the average mass relation is a factor of ~ 1.25 higher than the CHMR, implying that the CHMR likely changes little with time.

• In set B1, we use the BGD08 total GSMF, which extends to masses as lower as $\log(M_*/{\rm M}_\odot)$=7.4. This function is steeper at the low-M_* end and decays faster at the highest masses than the YMB09 GSMF. Therefore, the CHMR and SSMR are slightly different to those in set B. In particular, the lowest masses show a slight flattening as compared to results of set B. For the satellites, if extrapolated to even lower masses, this implies smaller subhalo masses for a given stellar mass than usually obtained from the standard AMT. This is diminishing the potential problem of too massive ΛCDM subhalos for the bright MW dSphs.

Our model allows us to infer in a natural way any statistical distribution for the central and satellite galaxy populations, for example the satellite CSMF and the mass

distributions and probabilities of particular subpopulations of satellites as a function of halo mass. The obtained satellite CSMFs in different halo mass bins agree very well with those inferred from the SDSS halo-based galaxy groups in YMB09.

We conclude that the semi-empirical results we obtain here, both for the central-halo and satellite-subhalo mass relations and their intrinsic scatters, are quite robust and imply full consistency of the ΛCDM halo and subhalo populations with several statistical distributions of the observed populations of central and satellite galaxies down to $M_* \sim 10^9$ $M_\odot$.

ACKNOWLEDGMENTS

We thank the Referee, David Weinberg, for a constructive report that helped to improve the paper. A. R-P acknowledges a graduate student fellowship provided by CONACyT. N. D. and V. A. acknowledge CONACyT grants 128556 and 167332.

Part III

APPLICATIONS FOR SOME SATELLITE OCCUPATIONAL DISTRIBUTIONS AT $z \sim 0$

En esta Parte III, exploramos diversas distribuciones ocupacionales y probabilísticas de las galaxias en función de su masa de halo, $M_{\rm h}$. En el Capítulo 4 se infiere la distribución de masa del satélite más masivo como función de $M_{\rm h}$, utilizando las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}$ previamente constreñidas en el Capítulo 3. Al comparar la distribución de masa del satélite más masivo con aquella de las galaxias centrales, mostraremos que las galaxias centrales son galaxias estadísticamente especiales en el halo, es decir no son una realización estadística de la muestra de miembros del grupo/cúmulo, por lo menos hasta halos tan masivos como 10^{15} $M_\odot$. Mientras menor la masa, mayor es la brecha entre las masas de la central y el satélite más masivo. En el Capítulo 5 se construye un catalogo sintético en base al modelo presentado en los Capítulos 2 y 3, con el fin de estudiar una serie de distribuciones de satélites (más masivos que Fornax, $m_* \sim 4 \times 10^7 M_\odot$) alrededor de galaxias tipo Vía Láctea (VL). Nuestro catálogo se puede entender como una prolongación de las observaciones actuales que permite estudiar estadísticamente las poblaciones satelitales de decenas de miles de galaxias tipo VL hasta masas tan pequeñas como Fornax. Se muestra que la VL no es un sistema común al tener dos satélites tan masivos como las Nubes de Magallanes, pero tampoco es un caso atípico; la función de masa de sus satélites está dentro del 1-σ del promedio. Por la configuración satelital de la VL, su masa de halo estadísticamente no es menor a $\sim 1.4 \times 10^{12}$ $M_\odot$ al nivel de 1-σ. Nuestros resultados sugieren que en la cosmología ΛCDM no existe un problema de satélites masivos perdidos a nivel de abundancias. Sin embargo, a nivel de dínamica interna, los satélites sumidos en subhalos de materia oscura fría, a partir de $m_* \sim 10^8$ $M_\odot$ y para masas menores, parecen no estar en buen acuerdo (demasiado concentrados) con las pocas inferencias que hay de velocidad circular al máximo de galaxias enanas, representando un potencial problema para la cosmología ΛCDM o para nuestro entendimiento de la física de galaxias enanas satélites.

El Capítulo 4 corresponde a una parte del artículo publicado: *"The Galaxy-Halo/Subhalo Connection: Mass Relations and Implications for Some Satellite Occupational Distributions", Rodríguez-Puebla, Avila-Reese & Drory 2013, ApJ 767, 92*. El Capítulo 5 corresponde al artículo en prensa: *"The Massive Satellite Population of Milky-Way Sized Galaxies", Rodríguez-Puebla Avila-Reese & Drory 2013, ApJ, en prensa.*

4

PROBABILITY DISTRIBUTIONS OF SATELLITES AS A FUNCTION OF HALO MASS

Part of this Chapter was published as: Rodríguez-Puebla A.; Avila-Reese V.; Drory N., 2013, ApJ 767, 92.

ABSTRACT

Using the central stellar-to-halo mass relations and the satellite stellar-to-halo mass relations constrained in Chapter 3, we predict the mass distribution both for the most massive satellite and the central galaxy as a function of halo mass. Additionally, we also predict the probabilities of having N satellites in the mass range of the Magellanic-Clouds (MCs) for a range of possible MW-halo masses. We find that the central galaxy M_* is not on average within the mass distribution of the most-massive satellite, even for cluster-sized halos, i.e., centrals are not a mere realization of the high-end of the satellite mass function; however for $> 3 \times 10^{13}$ $M_\odot$ halos, $\sim$ 15% of centrals could be. We also find that the probabilities of Milky Way-sized halos of having N MCs-sized satellites agree well with observational measures; for a halo mass of 2×10^{12} $M_\odot$, the probability to have 2 MCs is 5.4%, but if we exclude those systems with satellites larger than the MCs, then the probability decreases to $< 2.2\%$.

4.1 INTRODUCTION

The statistical model for linking galaxies to (sub)halos presented in Chapter 3, allows to infer in a natural way any statistical distribution for the central and satellite galaxy populations. As an example, in that Chapter we have presented the predicted central and satellite conditional stellar mass functions, CSMF, as a function of halo mass. The obtained central and satellite CSMFs in different halo mass bins agree very well with those inferred from the SDSS halo-based galaxy groups in YMB09. The above result give the confidence to explore more and different mass distributions and probabilities of particular subpopulations of satellite galaxies as a function of halo mass. The main aim of this Chapter is to use the mass relations of Chapter 3 and explore their implications for some satellite occupational distributions. In particular, we will explor two interesting statistics related to well-posed astronomical problems. The first one, is a longstanding question and refers whether the brightest cluster galaxies are a mere statistical extreme of the luminosity function in clusters or they form a different class [e.g., 213]. To study this problem, we will compute the distribution of the stellar mass gap between the central and the most-massive satellite galaxy as a function of halo mass. The second one, refers to the probabilities for MW-like halos to have $N_{\rm MC}$ MC-sized satellites.

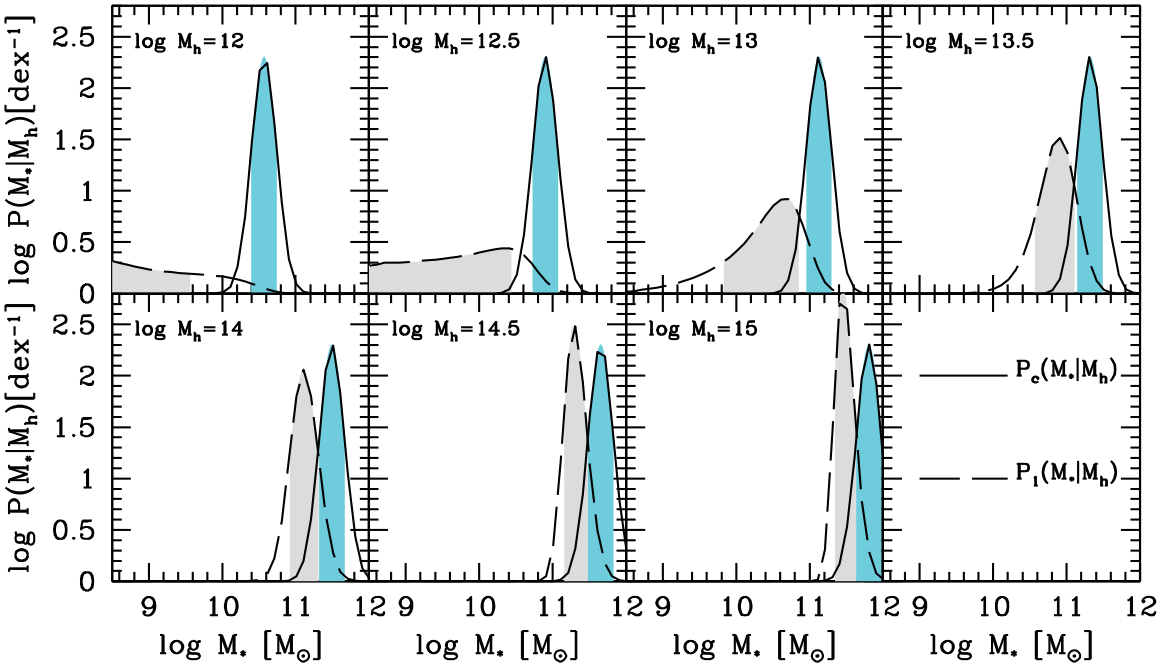

Figure 12: Stellar mass distributions of centrals (solid line) and the most massive satellites (dashed line) in seven different host halo masses. Shaded areas indicate the 68% of the corresponding distributions. The central galaxy masses on average are not part of the 1σ most-massive satellite distributions. For masses above $\sim 3 \times 10^{13}$ $M_\odot$, only approximately 15% of both distributions overlap. For smaller masses, this fraction rapidly decreases down to 3% at $M_h = 10^{12}$ $M_\odot$. The fraction of cases with overlapping distributions are expected to correspond to those cases where the central galaxy mass is a statistical realization of the most-massive satellite distribution.

4.2 PROBABILITY DISTRIBUTIONS OF SATELLITES

Once we have constrained the distribution of satellite galaxies in Chapter 3 by means of the central stellar-to-halo mass relations and the satellite stellar-to-halo mass relations, we can now predict the probabilities of having N satellites of a fixed M_* or in a particular M_* range as a function of M_h. It is assumed in our model that the second moment of the satellite distribution follows a Poissonian distribution (see section 3.2.2); the second moment is necessary to estimate chance probabilities for any given number of satellites. The probability of finding N satellites with stellar mass above M_* in a host halo of mass M_h is then given by:

$$P(N) = \frac{N_s^N e^{-N_s}}{N!}, \tag{44}$$

where for convenience we redefined $N_s = \langle N_s(> M_*|M_h)\rangle$.

4.3 RESULTS

4.3.1 *The most massive satellite mass distribution*

We can use the satellite CSMF and Eq. (44) to compute the mass probability distribution of the most massive satellite in halos of different masses. This is given by the following expression [e.g., 142, 217]:

$$\mathcal{P}_1(M_*|M_{\rm h})dM_* = \frac{\partial N_s}{\partial M_*} \times e^{-N_s} dM_*, \tag{45}$$

Note that $\int_{M_*}^{\infty} \mathcal{P}_1(M_*|M_{\rm h})dM_* = P(\geq 1)$, where $P(\geq 1)$ is the probability of finding at least one satellite galaxy more massive than M_*, $P(\geq 1) = 1 - P(0)$.

The results are shown in Fig. 12, where dashed and solid lines are for the most massive satellite and central galaxy mass distributions, respectively. The latter is by assumption a lognormal function of width $\sigma_c = 0.173$ dex. The shaded areas indicate the 68% width of the corresponding distributions. As seen in Fig. 12, the mass distribution of the most massive satellite changes with $M_{\rm h}$: in massive halos, it becomes closer to the distribution of the central galaxy, while in lower mass halos it tends towards small satellite masses compared to the central. This difference in masses, expressed in magnitudes, is referred in the literature of galaxy groups/clusters as the magnitude gap. The behavior seen in Fig. 12 is just a consequence of the satellite CSMFs showed in Fig. 10.

For halos larger than $\sim 1 - 3 \times 10^{13}$ $M_\odot$, the mean and standard deviation of the most massive satellite mass distribution slightly increase and decrease with $M_{\rm h}$, respectively, while for smaller masses, the mean value of $\mathcal{P}_1(M_*|M_{\rm h})$ strongly decreases as $M_{\rm h}$ decreases (faster than the central galaxy mass does) and the standard deviation increases. This transition is just at the mass corresponding to small classical galaxy groups. Therefore, our result seems to be a consequence of the fact that in groups, the larger the system's mass is, the smaller is the collision cross sections for big galaxies of close masses so that more of them survive. Instead, in galaxy-sized halos, due to their smaller velocity dispersions, the galaxy collision cross sections are large in such a way that the largest galaxies probably merged into one dominant central. Besides, the smaller the halo, the earlier most of its mass assembled on average; hence, the (wet) mergers of the most massive galaxies in the halo would have happened early. However, a fraction of the galaxy-sized halos, while on average dynamically old, can accrete massive satellites late. This could partially explain the wide distribution of masses of the second most massive satellite in MW-sized halos. For example, as seen in Fig. 12, the probability for these halos to have the most massive satellite ~ 5 times larger than the LMC is close to the probability to have this satellite as massive as the LMC.

From Fig. 10 we see that the mass of the central galaxy in the largest halos could be the statistical extreme of the satellite CSMF. According to Fig. 12, the mass distributions of the most massive satellite and central galaxy become closer as $M_{\rm h}$ increases. However, quantitatively, we see that the mean of $\mathcal{P}_1(M_*|M_{\rm h})$ lies outside of the 1σ of the central galaxy mass distribution even for a 10^{15} $M_\odot$ halo, i.e., the central and the most massive satellite galaxy, on average, are not expected to be drawn from the same exponentially

decaying mass function; this criterion is similar to the observational one introduced by Tremaine & Richstone [213].

We can estimate the fraction of systems where both mass distributions overlap, and consider that this fraction corresponds to the cases where the most massive satellite and the central galaxy are drawn from the same distribution. For masses above $\sim 3 \times 10^{13}$ $M_\odot$, approximately 15% of halos would have central galaxies that are not statistically peculiar with respect to the satellites. For smaller masses, this fraction rapidly decreases down to 3% at $M_h = 10^{12}$ $M_\odot$. In conclusion, most of centrals seem to form a statistically different class of galaxies with respect to the satellites at all halo masses, with a small fraction of cases, up to $\sim$ 15% in cluster-sized halos, being the exception, that is to say the centrals in these cases could be a statistical realization of the high-mass end of the satellite CSMF.

In order to compare our population statistics in detail with observations, the systems should be selected by the central galaxy M_* and/or group richness instead of the halo mass. We will will carry out this exercise elsewhere by using a mock catalog based on the the distributions constrained with our model.

4.3.2 *The probability of Milky Way–Magellanic Clouds systems*

Our model results and Eq. (45) can be used to compute the probability of having one, two, or N Magellanic Clouds (MCs) satellites in MW-sized halos. We calculate these probabilities for a range of possible MW-halo masses discussed in the literature: $(0.7, 1, 2, 3) \times 10^{12}$ $M_\odot$. We use $M_{LMC} = 2.3 \times 10^9 M_\odot$ and $M_{SMC} = 5.3 \times 10^8 M_\odot$ [109] for the stellar masses of the MCs.

Firstly, we are interested in calculating statistics that can be compared with observations. From a large SDSS sample, Liu et al. [127] have estimated the fraction of isolated galaxies with MW-like luminosities that do not have ($N_{MC} = 0$) and that have $N_{MC} = 1, 2, 3, 4, 5$, or 6 MC-sized satellites. We calculate similar probabilities for each of the halo masses mentioned above. In order to compare with Liu et al. [127], we do not exclude systems with satellites more massive than the LMC. The results from Liu et al. [127], for a search of MC-sized satellites up to 150 kpc around the primary, are plotted with crosses in Fig. 13 (from their Table 1)[1]. Note that in our case satellites are counted inside the host virial radius ($\sim 200 - 300$ kpc). Liu et al. [127] plot in their Fig. 8 the probabilities with a search radius up to 250 kpc only for $N_{MC} = 0, 1, 2, 3$. We reproduce these measurements in Fig. 13 with gray symbols and error bars.

Our predicted probabilities for set C are plotted in Fig. 13. The probabilities of MW-like halos hosting MC-sized luminous satellites (but including possible larger satellites) increase with M_h. Recall that in the case of Liu et al. [127] the central galaxy luminosity is fixed. In this sense, our results suggest that this luminosity ($M_{0.1r} = -21.2 \pm 0.2$ mag) can be associated to halos of different masses: for those galaxies with 1 or 2 MC-sized satellites, the preferred masses are $\approx 1 - 2 \times 10^{12}$ $M_\odot$, while for those rarer systems with

1 The selection criteria and observational corrections for searching for MC-like satellites are actually quite diverse. Liu et al. [127, see also 44] explored the sensitivity of the probabilities to changes in various selection parameters and found that their results can be slightly different, with the largest sensitivity being the the satellite search radius around the primary.

3 to 6 MC-sized satellites, the preferred masses are $> 2 \times 10^{12}$ $M_\odot$. Interestingly enough, from the inverse of the CHMR (set C), taking into account the intrinsic scatter around this relation, the halo masses corresponding to the MW stellar mass, $\log(M_*/M_\odot)$ = 10.74±0.1, are $\log(M_h/M_\odot) = 12.31 \pm 0.22$. Therefore, the rare halos that host 1 or 2 MC-sized satellites are those on the low-mass side of the halo mass distribution given the MW central stellar mass, while the much more rarer halos hosting 3 to 6 MC-sized satellites are those in the high-mass end of such a distribution.

The probability of the concrete case of two MC-sized satellites (N_{MC}=2; but not excluding the possibility of satellites larger than the LMC) in a MW-like halo of 2×10^{12} $M_\odot$ is 5.4% for set C (see Fig. 12). If we exclude now from our model predictions the possibility of having satellites larger than the LMC (as it happens in the MW system), then the probability drops to a upper limit of 2.2%.

The statistics of finding MW-sized galaxies with satellites in the concrete mass range of the MCs is limited. This statistics is actually part of the more general cumulative conditional satellite mass function. Having this function for galaxies we may then ask, for instance, whether the MW is rare because it has two too massive satellites or because it has a deficiency of massive (larger than LMC) satellites with respect to the average. We will report results related to these questions elsewhere by using a mock galaxy catalog generated with the distributions constrained here. The mock catalog will allow us also to infer several statistics given the central galaxy stellar mass in which we are interested in (e.g., the MW one) instead of exploring a range of possible halo masses as was done here.

Finally, the agreement between the predicted and observationally determined probabilities is reasonable within the uncertainties. Such an agreement indicates that the model is self-consistent as well as consistent with the underlying ΛCDM scenario. Note that this self-consistency has been proven down to the scales of the MC galaxies and at the level of satellite population distributions. Similar probabilities were found also using large N-body cosmological simulations and looking for MW-sized halos with subhalos that have dynamical properties similar to the MCs [41, 44].

4.4 CONCLUSIONS

The model presented in Chapter 3 allows to infer any statistical distribution for the central and satellite galaxy populations. In this Chapter we have used the central stellar-to-halo mass relations and the satellite stellar-to-halo mass relations constrained in Chapter 3 in order to explore; (1) the distribution of the stellar mass gap between the central and the most-massive satellite galaxy as a function of halo mass, and (2) the probabilities for MW-like halos to have N_{MC} MC-sized satellites. Our conclusions regarding these questions are:

- With decreasing halo mass, the mass distribution of the most massive satellite as compared to the the distribution of the central galaxy become more different and shifted to lower masses. This shows that the central is a statistically exceptional galaxy in the halo (group). For masses larger than $M_h \sim 3 \times 10^{13}$ $M_\odot$, the differences become smaller but even in this case only $\sim$ 15% of halos seem to have the

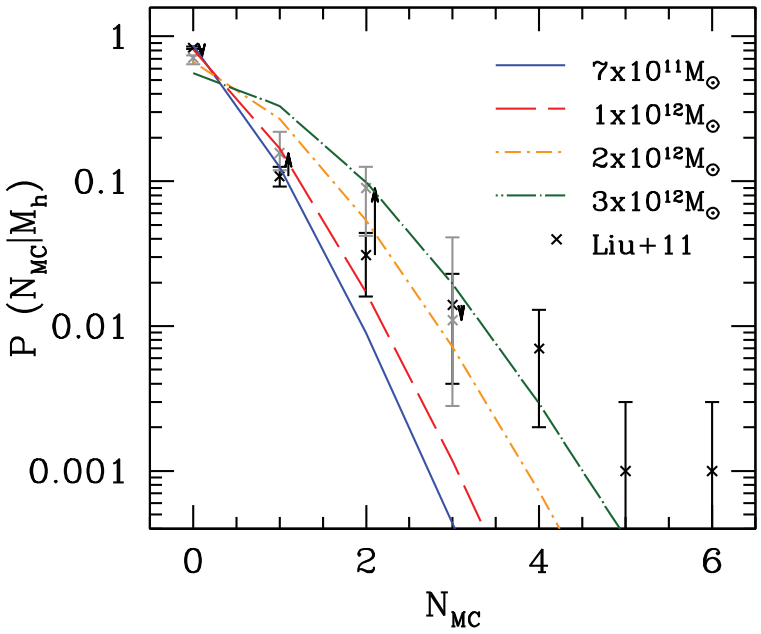

Figure 13: Probability of occurrence of $N_{\rm MC}$ MC-sized satellites in a range of possible MW-sized host halos (different lines are for the different masses indicated in the legends) based on the results for set C. Observational determinations by [127] for a large sample of SDSS galaxies are shown with black (gray) crosses for distances from the host up to 150 kpc (250 kpc). The black arrows show how the the occurrence of MC-sized satellites change when the search radii goes from 150 kpc to 250 kpc from the host.

most massive satellite statistically indistinguishable from the central one, which implies that the latter could be a mere statistical realization of the massive-end of the satellite CSMF instead of realization of a different galaxy.

- For the range of halo masses in question for the MW, we find that the probabilities to have $N_{\rm MC}$ MC-sized satellites are in good agreement with the observational determinations by Liu et al. [127]. A MW-halo mass of $\lesssim 2 \times 10^{12}$ $M_\odot$ would agree better with the observational determinations for two MC-sized satellites ($N_{\rm MC} = 2$). When excluding the cases that satellites are larger than the LMC, the probabilities become even lower: $< 2.2\%$ for $M_{\rm h} = 2 \times 10^{12}$ $M_\odot$.

5

THE MASSIVE SATELLITE POPULATION OF MILKY WAY-SIZED GALAXIES

This Chapter will be published as: Rodríguez-Puebla A; Avila-Reese V.; Drory N., 2013, in press at ApJ.

ABSTRACT

Several occupational distributions for satellite galaxies more massive than $m_* \approx 4 \times 10^7$ $M_\odot$ around Milky-Way (MW)-sized hosts are presented and used to predict the internal dynamics of these satellites as a function of m_*. For the analysis, a large galaxy group mock catalog is constructed on the basis of (sub)halo-to-stellar mass relations fully constrained with currently available observations, namely the galaxy stellar mass function decomposed into centrals and satellites, and the two-point correlation functions at different masses. We find that 6.6% of MW-sized galaxies host 2 satellites in the mass range of the Small and Large Magellanic Clouds (SMC and LMC, respectively). The probabilities of the MW-sized galaxies to have 1 satellite equal or larger than the LMC or 2 satellites equal or larger than the SMC or 3 satellites equal or larger than Saggitarius (Sgr) are $\approx 0.26, 0.14$, and 0.14, respectively. The cumulative satellite mass function of the MW, $N_s(\geq m_*)$, down to the mass of the Fornax dwarf is within the 1σ distribution of all the MW-sized galaxies. We find that MW-sized hosts with 3 satellites more massive than Sgr (as the MW) are among the most common cases. However, the most and second most massive satellites in these systems are smaller than the LMC and SMC by roughly 0.7 and 0.8 dex, respectively. We conclude that the distribution $N_s(\geq m_*)$ for MW-sized galaxies is quite broad, the particular case of the MW being of low frequency but not an outlier. The halo mass of MW-sized galaxies correlates only weakly with $N_s(\geq m_*)$. Then, it is not possible to accurately determine the MW halo mass by means of its $N_s(\geq m_*)$; from our catalog we constrain a lower limit of 1.38×10^{12} $M_\odot$ at the 1σ level. Our analysis strongly suggests that the abundance of massive subhalos should agree with the abundance of massive satellites in all MW-sized hosts, i.e. there is not a (massive) satellite missing problem for the ΛCDM cosmology. However, we confirm that the maximum circular velocity, $v_{\rm max}$, of the subhalos of satellites smaller than $m_* \sim 10^8$ $M_\odot$ is systematically larger than the $v_{\rm max}$ inferred from current observational studies of the MW bright dwarf satellites; differently from previous works, this conclusion is based on an analysis of the overall population of MW-sized galaxies. Some pieces of evidence suggest that the issue could refer only to satellite dwarfs but not to central dwarfs; then, environmental processes associated to dwarfs inside host halos combined with SN-driven core expansion should be at the basis of the lowering of $v_{\rm max}$.

5.1 INTRODUCTION

According to the current paradigm of cosmic structure formation and evolution, galaxies form inside Cold Dark Matter (CDM) halos, which grow both by diffuse mass accretion and by incorporation of smaller halos that become subhalos. Inside the subhalos (at least inside the more massive ones) galaxies should also have formed prior to their halo's infall, becoming satellite galaxies. Therefore, the present-day population of satellites around central galaxies is the product of the halo/subhalo assembly and the survival/destruction history of the the galaxies inside the subhalos. N-body simulations within the context of the ΛCDM cosmological scenario provide us with the subhalo conditional mass function (subHCMF) as a function of host halo mass $M_{\rm h}$ [see for recent results, e.g., 197, 94, 41, 90, 20]. Using this function and statistical models constrained by observations, the central/satellite–halo/subhalo mass connection can be established [e.g., 44, 176, hereafter RAD13]. In this way, the abundances of the galaxy satellite population as a function of $M_{\rm h}$ can be calculated (satellite conditional stellar mass function, CSMF). In this paper, our interest is focused on these abundances for systems with a central galaxy of Milky Way (MW) stellar mass, $M_{*,\rm MW}$.

With the advent of large galaxy surveys, some observational statistical studies of the satellite abundance of central galaxies, in particular those of MW luminosity or mass, have beed published. Several statistical distributions have been determined this way, for instance, the fractions of MW-sized galaxies with a given number N_s of satellites in the mass range of the Magellanic Clouds (MC) or with masses equal or larger than the LMC or the SMC [109, 211, 127, 44, 174]. A natural question is *whether the ΛCDM scenario makes predictions in agreement with these statistical results related to scales smaller than previously probed*.

The works mentioned above conclude that the MW is a rare case with significantly more massive (MC-sized) satellites than other galaxies of similar luminosity or mass. Other studies determine the average luminosity distribution of bright satellites around centrals [120, 101, 226, 202, 110, 181]. The distribution of the MW bright satellites seems to lie above the average found for MW-sized galaxies. In spite of all of these studies, it is not yet clear whether the satellite luminosity (mass) distribution of the MW is rare in a statistically significant sense. It could be that the MW-sized galaxies have a broad range of satellite luminosity distributions, the MC-like case being not particularly frequent but not an outlier.

The question on how typical is the MW satellite mass distribution has acquired relevance recently. This distribution, being the best studied one, is used to compare with subhalo distributions predicted in the context of the ΛCDM and alternative cosmological scenarios in order to test these scenarios at the smallest scales [c.f. 37, 38, 129, 222, 257]. However, such a comparison relies (i) on the hope that the MW satellite CSMF is not atypical and (ii) on the assumed halo mass for the MW (the subhalo abundance strongly depends on $M_{\rm h}$, e.g., 90, 224). For example, Boylan-Kolchin, Bullock & Kaplinghat [37, 38] have shown that for a few ΛCDM halos of $\sim 10^{12}$ $M_\odot$ resimulated at very high resolution, there is a significant excess of subhalos with too high masses or maximum circular velocities ($v_{\rm max} > 25$ km/s) with respect to what is inferred for the MW satellite population (the so-called "too big to fail" problem). By means of an analytical

model for generating a large sample of ΛCDM halos with their corresponding subhalo populations, Purcell & Zentner [169] propose that the large variation in the latter among different host halos ameliorates the "too big to fail" problem: at least $\sim$ 10% of their MW-sized halos host subhalo populations in agreement with the MW dwarf satellite kinematics. Wang et al. [224] suggest that the problem is ameliorated if the MW halo mass is simply less massive than is commonly thought, $M_{\rm h} \lesssim 10^{12}$ M$_\odot$.

In all of these works, the main caveats are the way the MW satellite population is put into the statistical context, and the way the populations of the predicted subhalos and of the observed MW satellites are matched. Here, we attempt to overcome these caveats by using a large mock catalog of MW-sized galaxies, constructed on the basis of (sub)halo-to-stellar mass relations fully constrained with currently available observations, namely the galaxy stellar mass function, GSMF, decomposed into centrals and satellites, and the projected two-point correlation functions, 2PCFs, measured at different stellar mass bins (for references see Section 2). While these observations are complete only down to $\approx 2 \times 10^8$ M$_\odot$, the occupational procedure used to construct the catalog allows one to "extrapolate" observations down to the stellar masses that match the minimum halo/subhalo masses considered here. In RDA13 [see also 44], a preliminary attempt of studying the massive satellite population of MW-sized galaxies has been presented; however, in that paper the results are given as a function of $M_{\rm h}$ instead of $M_{*,\rm MW}$, which introduces freedom to choose the right $M_{\rm h}$ to be used for the MW.

Our main result from analyzing the mock catalog is that the ΛCDM scenario is statistically consistent with observations regarding the abundances and internal dynamics of satellites in MW-sized galaxies down to satellite stellar masses $m_* \sim 10^8$ M$_\odot$. At lower masses, down to the limit of our study ($m_* \sim 10^7$ M$_\odot$), the abundances continue being consistent but the internal dynamics of observed dwarf satellites suggest that their subhalos have $v_{\rm max}$ values smaller than those of the ΛCDM subhalos, under the assumption that the $v_{\rm max}$ of the latter remain the same after galaxy formation and evolution. Our conclusions are not affected by uncertainties on the matching of subhalo-satellite abundances, on the statistical interpretation of the MW nor on the halo MW mass. Regarding the latter, we instead find the $M_{\rm h}$ distribution of the MW analogs [see also 44].

The layout of this Chapter is as follows. In Section 5.2 we briefly describe the semi-empirical occupational approach for linking galaxies to halos and subhalos and how, by using the results of this approach, we construct a mock catalog of 2 million central galaxies, each one with its satellite population down to $m_* \sim 10^7$ M$_\odot$. From this catalog, we select a subsample of about 41000 central galaxies with MW-like stellar masses. In Section 3, we present different statistical distributions for the massive satellite population of the MW-sized galaxies and compare them to some observational studies. We investigate the question of how common the MW satellite mass distribution is in 5.3.1, while in 5.3.2 we present the halo mass distribution of the MW analogs. In Section 5.4 we present $v_{\rm max}$ vs. stellar mass for the mock galaxy (both satellites and centrals) and compare with observations. Our conclusions and a discussion are given in Section **??**. We adopt cosmological parameter values close to WMAP 7: $\Omega_\Lambda = 0.73$, $\Omega_{\rm M} = 0.27$, $h = 0.70$, $n_s = 0.98$, and $\sigma_8 = 0.84$.

5.2 THE METHOD

In what follows, we briefly review the semi-empirical approach we use for connecting galaxies to halos and subhalos of different masses. For an extensive presentation of this approach, see Chapter 3 and also Section 2 of RAD13. The approach relies on the assumption that the central-to-halo and satellite-to-subhalo mass relations (CHMR and SSMR, respectively) are monotonic. By parametrizing these mass relations, with their intrinsic scatter included, one can use the predicted ΛCDM distinct halo and conditional subhalo mass functions (HMF and subHCMF) to generate the halo/subhalo occupational distributions both for *central and satellite* galaxies. Therefore, this method encapsulates the main ideas behind the abundance matching technique, the halo occupation distribution model, and the conditional stellar mass function formalism (RAD13; see also 178). The advantage of the approach is that all the relevant observed statistical distributions of *central and satellite* galaxies (the GSMF decomposed into centrals and satellites, the CSMFs, and the 2PCFs) are consistently related to each other and with the predicted halo/subhalo statistical distributions (the HMF and subHCMF).

The outputs of this approach are the CHMR and SSMR, including their intrinsic scatters, and the satellite CSMFs as *a function of halo mass* $M_{\rm h}$. Here we will use the best constrained CHMR and SSMR obtained in RAD13. These relations were (over)constrained by making use of all the available observational information (data set C in RAD13): the central and satellite GSMFs determined by Yang, Mo & van den Bosch [247] down to 2.5×10^8 $M_\odot$ and the projected 2PCFs determined by Yang et al. [250] in five stellar mass bins. For the distinct HMF and subHCMF, the Tinker et al. [207] and Boylan-Kolchin et al. [41, see also Gao et al. [90]] fits to cosmological simulations were used, respectively.

5.2.1 *The galaxy group mock catalog*

Instead of using the analytical CHMR and SSMR directly, we apply these functions and their scatters to generate a mock galaxy group catalog. With this catalog we will explore several statistical satellite distributions that can be compared with some direct observational determinations given as a function of the central stellar mass. The catalog is generated as follows:

- From a minimum halo mass of $M_{\rm h,min} = 10^{10.5} M_\odot$, a population of 2×10^6 halos is sampled from the distinct HMF. Each halo is randomly picked from this function by generating a random number U uniformly distributed between 0 and 1 and finding the value for $M_{\rm h}$ that solves the equation $n_h(M_{\rm h})/n_h(M_{\rm h,min}) = U$. Here n_h is the cumulative distinct HMF.

- To each halo a central galaxy with stellar mass M_* is assigned randomly from the probability distribution $P(M_*|M_{\rm h})$, i.e. the mean M_*–$M_{\rm h}$ relation and its intrinsic scatter which is assumed to be lognormal distributed with a width of 0.173 dex (see RAD13).

- To each halo defined by its mass $M_{\rm h}$ a subhalo population above $m_{\rm sub,min} = 10^9$ $M_\odot$ is assigned randomly by assuming a Poisson distribution [115, 41]. First, the total

number of subhalos, $\mathcal{N}_{\rm sub}$, above $m_{\rm sub,min}$ is specified by choosing an integer from a Poisson distribution with mean $\langle N_{\rm sub}(> m_{\rm sub,min}|M_{\rm h})\rangle$, where this mean number is taken from the subHCMF for the given $M_{\rm h}$. Then, the mass $m_{\rm sub}$ for each subhalo is assigned by solving the equation $\langle N_{\rm sub}(> m_{\rm sub}|M_{\rm h})\rangle / \langle N_{\rm sub}(> m_{\rm sub,min}|M_{\rm h})\rangle = u$, where u again is a random number uniformly distributed between 0 and 1. Note that this last step should be repeated $\mathcal{N}_{\rm sub}$ times in order to assign subhalo masses to each one of the $\mathcal{N}_{\rm sub}$ subhalos. The lower limit in subhalo mass is enough to sample satellite galaxies with stellar masses larger than $m_* \approx 10^7 {\rm M}_\odot$, see Fig. 7 of RAD13.

- To each subhalo we assign a satellite galaxy with stellar mass m_*, taken from the probability distribution $P(m_*|m_{\rm sub})$, i.e. the mean m_*–$m_{\rm sub}$ relation and an intrinsic scatter, assumed to be lognormal distributed with a width equal to the central/halo case (the latter assumption seems to be reasonable, see RAD13).

The mock catalog generated in this way reproduces the observational statistical functions used to constrain the CHMR and SSMR, namely the GSMF separated into central and satellite galaxies and the 2PCFs in several mass bins. However, the catalog contains much more information, which can be thought as an "extension" as well as an extrapolation to lower masses of the observations. In particular, we can find the overall satellite number distributions down to $m_* = 10^7$ M$_\odot$ around galaxies of a given stellar mass M_*.

Fig. 14 illustrates the mean CHMR (solid line) and its 1σ scatter (0.173 dex; gray shaded area) for the data set C as reported in RAD13. The 2 million mock central galaxies sample this distribution by construction. The short dashed line indicates the mass of central galaxies with $\log M_* = 10.74$ while the dotted lines are 0.1 dex above and below defining a subsample of galaxies with stellar masses in the bin $\log(M_{*,\rm MW}/{\rm M}_\odot) \in [10.64, 10.84]$, which corresponds to the stellar mass estimate for the MW [86]. The 40694 realizations out of the 2 millions that fall within this narrow M_* range are represented using black dots (*MW-sized galaxies*). The shape of the resulting distribution of this subsample of central galaxies as a function of $M_{\rm h}$ is shown in the bottom panel of the figure. The mean and the standard deviation for this distribution are $\log(M_{\rm h}/{\rm M}_\odot) = 12.312 \pm 0.277$.

5.3 RESULTS AND COMPARISON TO OBSERVATIONS

In the previous Section, we generated a mock catalog of central galaxies corresponding to the stellar mass estimates for the MW. These galaxies have halos in a broad range of masses (see Fig. 14). From this sample, we can then establish the fractions (probabilities) of systems with N_s satellites within a (stellar) mass range or above a given mass; this mass can not be smaller than $m_* = 10^7$ M$_\odot$, the minimal mass used to construct the mock catalog (see §§5.2.1). Therefore, our results will be restricted to the population of the largest satellites.

For the statistical calculations, we will assume that the stellar masses of the LMC and SMC satellite galaxies are $m_{\rm LMC} = 2.3 \times 10^9$ M$_\odot$ and $m_{\rm SMC} = 5.3 \times 10^8$ M$_\odot$ [109]. We will also consider that the third most massive MW satellite is Sagittarius (Sgr). For a $V-$band, the absolute magnitude of -13.63 mag and a stellar mass-to-luminosity ratio of 2 for

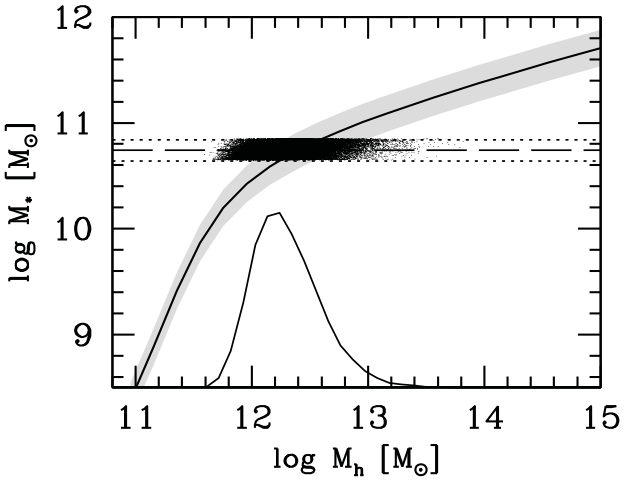

Figure 14: Stellar-to-halo mass relation for central galaxies. The solid line indicates the CHMR reported in RAD13, while the gray shaded area shows the 1σ scatter around the mean, assumed to be 0.173 dex. Galaxies that are identified with MW-sized galaxies are those lying in the bin $\log M_* \in [10.64, 10.84]$ indicated with the dotted lines. The dashed line indicates the mean of this bin. The black dots are the 40694 realizations of MW-sized galaxies. The resulting distribution as a function of halo mass for MW-sized galaxies is showed below the CHMR. The mean and the standard deviation for this distribution are $\log M_{\rm h} = 12.312$ and 0.277 dex, respectively.

Sgr[1] imply a stellar mass of $m_{\rm Sgr} = 5 \times 10^7$ M$_\odot$. Sgr is a tidally-stripped dwarf. Based on observations of its tidal tails, a total magnitude (core + tails) of ≈ -15 mag is obtained [153]. Then, a rough estimate of the core + tails stellar mass of Sgr is $m_{\rm Sgr+t} = 1.5 \times 10^8$ M$_\odot$. The fourth most massive MW satellite is Fornax (For), with a V-band absolute magnitude of -13.3 mag. An estimate of its stellar mass is $m_{\rm For} = 4.3 \times 10^7$ M$_\odot$, de Boer et al. [66].

Our mock catalog was constructed based on observational constraints, so the different satellite population statistics should be consistent with those of real galaxies; we expect that this consistency is preserved for the extrapolations to lower masses using this catalog. In what follows, we compare the results from the mock catalog with observational distributions of MW-sized galaxies and their population of massive satellites. It is important to remark that we *do not assume a particular halo mass* for the studied MW-sized galaxies.

From a large SDSS sample, Liu et al. [127] estimate the fraction of MW-sized isolated galaxies that do not have any ($N_{\rm MC} = 0$) and that have $N_{\rm MC} = 1, 2, 3, 4, 5$ or 6 MC-sized satellites. In the same way, we find in our mock catalog the different fractions of MW-sized galaxies with $N_{\rm MC} = 0, 1...6$ satellites in the stellar mass range $m_{\rm SMC} - m_{\rm LMC}$. Figure 15 shows the predicted probabilities (long-dashed line). The probability of MW-sized galaxies having two MC-sized satellites is 6.6%. If, from the subsample of MW-like galaxies having $N_{\rm MC}$ MC-sized satellites we exclude those with satellites larger than the

1 We assume for Sgr a stellar population with average metallicity [Fe/H]≈ -0.5 dex [50, 51] and average age of 8 Gyr [22].

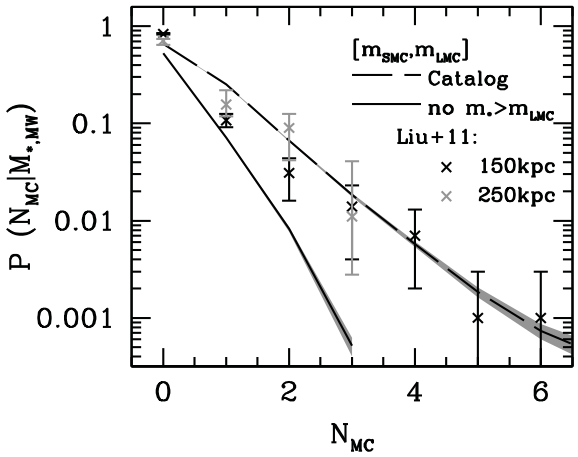

Figure 15: Probability of occurrence of $N_{\rm MC}$ satellites in the MC mass range around MW-sized galaxies (long-dashed line; solid line is for the extra condition of no satellites larger than the LMC). The shaded areas are the respective Poissonian errors from the counting. Direct observational results from [127] are plotted with black (separation from the host up to 150 kpc) and gray (separations up to 250 kpc) skeletal symbols.

LMC, then the probabilities decrease even further (solid line). For $N_{\rm MC} = 2$ and no satellites larger than LMC, the probability is now only 0.08%. Note that this implies that by far most of those MW-like systems that have $N_{\rm MC} = 2$ should have at least one satellite more massive than the LMC; the MW system does not have such a satellite.

The results from Liu et al. [127], for a search of MC-sized satellites (not excluding systems with satellites larger than the LMC) up to 150 kpc around the primary, are plotted as crosses in Fig. 15. Note that in our case satellites are counted inside the host virial radius ($\sim 200 - 300$ kpc). Based on Fig. 8 in Liu et al. [127], we also plot the probabilities when the search radius is increased up to 250 kpc (data are provided only for $N_{\rm MC} = 0, 1, 2, 3$). It should be said that the selection criteria and observational corrections to correct for MC-sized satellites are quite diverse. Liu et al. [127, see also Busha et al. [44]] explored the sensitivity of the probabilities to changes in various selection parameters and found that their results can slightly change; the results are particularly sensitive to the satellite search radius around the primary.

The agreement between the probabilities in our mock catalog and the Liu et al. [127] observations is good within the uncertainties. It is encouraging that the mock catalog predicts the statistics of very rare events, as those systems with $N_{\rm MC} \geq 3$, in good agreement with observations. Regarding the more common events, in the catalog there is a $\sim 66\%$ chance of MW-sized galaxies without MC-sized satellites, while Liu et al. [127] report 71% of such galaxies (for radii up to 250 kpc); this is because we also have slightly more galaxies with $N_{\rm MC} = 1$ than in Liu et al. [127] (the probabilities of systems with more MC-sized satellites are even lower and do not contribute significantly). These small differences can be explained due to the fact that the search radius for satellites in Liu et al. [127] is up to 250 kpc, while in our case there is a fraction of MW-sized galaxies with massive halos, whose virial radii are larger than 250 kpc. If the satellite

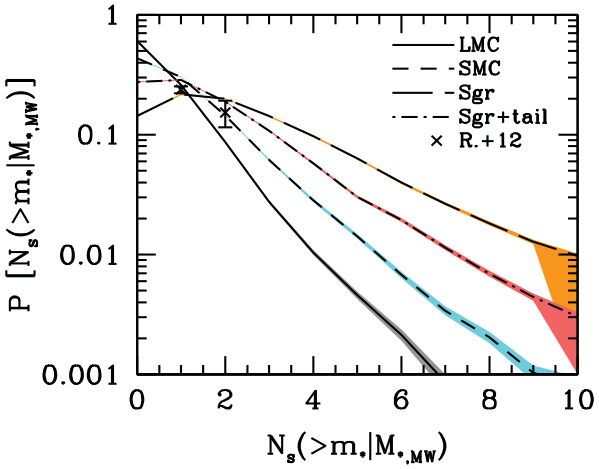

Figure 16: Probability of occurrence of N_s satellites around MW-sized galaxies with stellar masses equal or larger than the LMC, SMC, Sgr+tail, and Sgr (solid, short-dashed, dot-short-dashed, and dashed lines, respectively). The skeletal symbols are the observational inferences by Robotham et al. [174], corrected for a search radius up to 250 kpc, for satellites equal or more massive than the LMC and the SMC.

search radius would be increased in Liu et al. [127], then the fraction of MW-sized galaxies without MC-sized satellites would decrease.

An alternative statistical study of MW analog systems was presented in Robotham et al. [174]. Based on a sample of MW-sized galaxies from the new GAMA survey [74], they have found the fractions of objects in this sample with at least one satellite as massive as the LMC or with two satellites at least as massive as the SMC. From our mock catalog we can calculate the fractions of MW-sized galaxies with any number of satellites equal or larger than a given stellar mass m_*, $P[N_s(\geq m_*|M_{*,\rm MW})]$. Figure 16 shows these probabilities for $m_* \geq m_{\rm LMC}$(solid line), $m_* \geq m_{\rm SMC}$ (dashed line), and $m_* \geq m_{\rm Sgr}$ (long-dashed line; the dot-dashed line is for the case when the tails of Sgr are included in its mass). The colored contours around the lines are the corresponding Poissonian errors from counting. The probabilities of finding one satellite equal or more massive than the LMC and two satellites equal or more massive than the SMC are 26% and 14.5%, respectively. In the case of Robotham et al. [174] these probabilities are 11.9% (11.2%–12.8%) and 3.4% (2.7% –4.5%). However, in Robotham et al. [174] the satellite search radius was fixed to only 70 kpc. From Liu et al. [127], we roughly estimate the factors by which these fractions could increase if the search radius were to be extended to to 250 kpc; the factors are at least 2 and 4.5 for $N_{\rm MC} = 1$ and $N_{\rm MC} = 2$, respectively (they could be larger because Liu et al. [127] limit the search to only satellites in the $m_{\rm LMC} - m_{\rm SMC}$ mass range). Taking these correction factors into account, the agreement between the predicted probabilities and those determined by Robotham et al. [174] becomes quite good.

Recently, several authors have measured the complete (bright) satellite abundances around bright centrals, in particular those with luminosities close to the MW and M31, by using adequate samples from the SDSS [120, 101, 226, 202, 181] and from the Canada-

France-Hawai Telescope Legacy Survey [110]. In each one of these studies, different criteria for the sample selection, different searching and correction methodologies, various radii for the satellite search, etc. were applied. Therefore, the results are not easy to compare.

In general, these works find that the conditional bright satellite luminosity function of MW/M31-sized galaxies is described by a relatively steep power law, and a normalization such that down to ~ 6 magnitudes fainter than the central there is on average a factor of $1.5-2$ fewer satellites than the average of the MW and M31. The MW satellite CSMF measured in our mock catalog agrees in general with these studies, but it seems to be slightly overabundant above the mass (or luminosity) corresponding to the SMC, in particular with respect to Wang & White [226] and Strigari & Wechsler [202]. In the case of Jiang, Jing & Li [110, and in a less extent for Guo et al. [101]], a slight flattening at the high-end of the luminosity function is seen, which is similar to our case. We recall that the direct observational searches of satellites are for a fixed radius around the central, which is 250 or 300 kpc typically (the exception is Jiang, Jing & Li [110] who use the virial radius determined by the Yang et al. [249] group finding algorithm). In the mock catalog we count the satellites inside the virial radius, which for a non-negligible fraction of galaxies, is larger than 300 kpc. Therefore, it is expected that the number of satellites counted in the direct observational studies (in special of the most massive ones, which are more probable to be at larger radii) should be slightly lower than in our mock catalog.

We conclude that the population of the largest satellites around MW-sized central galaxies in our mock catalog agrees in general with several direct observational determinations, which present different and limited satellite population statistics. The advantage of our mock catalog, constrained by observations, is that allows calculate any satellite occupational statistics, and to extend the satellite mass limit to masses lower than current direct observational studies. In this way, one may explore in more detail how are the satellite populations of MW-sized galaxies and how particular is the MW system.

5.3.1 *How common is the Milky-Way system?*

According to Fig. 16, the MW is less common than similar sized galaxies in the sense that it has one satellite as massive as the LMC or two satellites equal or more massive than the SMC; there are more MW-sized galaxies that do not have satellites of mass $m_* \geq m_{\rm LMC}$ (60.6% vs 26.1% for those with one satellite) or have less than two satellites more massive than $m_{\rm SMC}$ (85.5% vs 14.5% for those with two satellites). However, the MW can be considered a common galaxy in the sense that it has three satellites more massive than $m_{\rm Sgr}$. In general, what we learn from Fig. 16 is that the satellite number distributions are relatively wide and there is not a strongly preferred number of satellites above a given mass. For example, the probabilities of having 0, 1, 2, 3, or 4 satellites with $m_* \geq m_{\rm Sgr}$ are within a factor of less than two from each other.

The fact that the satellite number distributions of MW-sized galaxies are broad can also be seen in the plot of the cumulative number of satellites above a given mass m_* as a function of m_*, $N_s(\geq m_*|M_{*,\rm MW})$, which is related to the satellite CSMF discussed

above. Figure 17 shows the average (solid line) and the 1σ scatter (gray shaded area) of $N_s(\geq m_*|M_{*,\rm MW})$ from the mock catalog. The latter is quite broad. The cyan line corresponds to the MW (the red line is for the case the mass of Sgr includes the tidal tails). The MW massive satellite population is within 1σ of the number distribution of satellites as a function of mass of all MW-sized galaxies, being above the average by less than a factor of 2 at the MC satellite masses, and very close to the average regarding its three (four) satellites equal or more massive than Sgr (For). By means of direct observational determinations Guo et al. [101], Strigari & Wechsler [202], and Jiang, Jing & Li [110] arrived to a similar conclusion. From a frequency point of view, we find that the MW-sized galaxies with one satellite $\geq m_{\rm LMC}$ (two satellites $\geq m_{\rm SMC}$), as the MW, happen only 1/0.6=1.68 (2/1.02 =1.92) times less frequently than the average (see also Fig. 16).

In fact, given that the (massive) satellite number distribution as a function of mass of MW-sized galaxies is relatively broad, several kind of "configurations" have close probabilities and all are relatively low. Besides, as more constraints are imposed on the configuration (as for example, to have two satellites in the SMC–LMC mass range but not larger than the LMC, see Fig. 15), the lower will be the frequency of occurrence. However, *this does not imply that systems with a particular configuration are outliers, i.e., out of* 1σ.

In Fig. 17, we also show the mean satellite cumulative mass function of the subsamples of MW-sized galaxies constrained to have $N_s \geq 3$ (short dashed line), $N_s = 3$ (long dashed line), and $N_s \leq 3$ (dot-dashed line) satellites more massive than Sgr. It is interesting to see that *galaxies with exactly 3 satellites more massive than Sgr are close to the average for MW-sized centrals*, but they have typically the most and second most massive satellites smaller than the LMC and SMC by roughly 0.7 and 0.8 dex, respectively. The subsample of galaxies with $N_s \geq 3$ satellites more massive than Sgr describes better the satellite mass function of the MW down to the SMC or to Sgr when including its tails. Finally, we see that the MW definitively does not belong statistically to the subsample of MW-sized centrals with $N_s \leq 3$ satellites more massive than Sgr, contrary to what is assumed in Wang et al. [224].

The analysis presented above for the MW system can be applied also to M31. Recent observational results show that M31 has at least twice as many satellites as the MW [251]. Specifically, it has six satellites brighter than the luminosity of Sgr, making M31 an outlier according to Fig. 17. However, the stellar mass of M31 is a factor of ~ 2 larger than the MW [e.g., 205]. Therefore, it is expected that the M31 halo is more massive than the MW one, hence the M31 halo should host more satellites. This question will be analyzed in detail elsewhere.

5.3.2 *The halo masses of MW-like systems*

The host halo mass distribution of the MW-sized galaxies in the mock catalog is plotted in Fig. 14. The distribution is broad, with mean and median values of 2.05×10^{12} and 1.91×10^{12} $M_\odot$. It is known that for clusters and groups of galaxies the total dynamical mass of the system correlates with the richness (number of members above a given mass, see e.g., 172). Is this also the case for MW-like systems? Could we constrain statistically

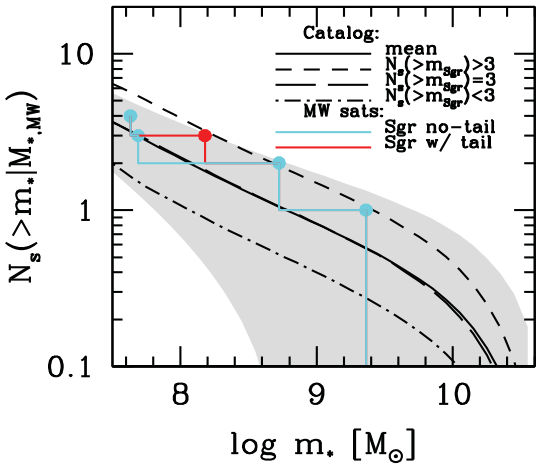

Figure 17: Cumulative satellite mass function of MW-sized galaxies in the mock catalog (solid line) and its 1σ scatter (gray shaded area). Subsamples of MW-sized galaxies constrained to have $N_s \geq 3$, $N_s = 3$, and $N_s \leq 3$ satellites more massive than Sgr are indicated with the short-dashed, long-dashed, and dot-dashed lines, respectively. The cyan staggered line corresponds to the MW satellite galaxies, while the red line is for the case the mass of Sgr includes their tidal tails.

the MW halo mass by its number of satellites above a given mass or in between a given mass range?

In Fig. 18 we plot the mean host $M_{\rm h}$ of the mock MW-systems with N_s satellites with a mass larger or equal than the LMC (solid line), SMC (short-dashed line), and Sgr (long-dashed line), and with masses in between the SMC and LMC (dot-dashed blue line). The statistical scatter in all the cases is roughly ~ 0.24 dex in $\log M_{\rm h}$. For clarity, we plot the scatter (vertical lines) only for the cases corresponding to the MW, i.e., $N_s = 1$ for the solid line, $N_s = 2$ for the short-dashed line, $N_s = 3$ for the long-dashed line, and $N_s = 2$ for the dot-dashed blue line (slightly shifted horizontally). Figure 18 shows that, in general, there is a correlation of $M_{\rm h}$ with N_s but it is weak. The scatter of $M_{\rm h}$ around a given M_* does not depend significantly on N_s for galaxies below the knee in the M_*–$M_{\rm h}$ relation (see Fig. 14).

From Fig. 18 we can say that *at the* 1σ *level, the halo mass of MW-like systems is not smaller than* $1.38 \times 10^{12}\ M_\odot$. This limit is for the case of 3 satellites with $m_* \geq m_{\rm Sgr}$ (the mean $M_{\rm h}$ for this case is $\log(M_{\rm h}/{\rm M}_\odot)$=12.33). Interestingly enough, most of the observational estimates of the MW halo mass give values above 10^{12} M$_\odot$. For example, the most recent work, based on the proper motion of the Leo I dwarf galaxy in combination with numerical simulations, favors a value of $1.6 \times 10^{12}{\rm M}_\odot$ [39, and references therein]. For restrictions related to the number of MC-sized satellites, the typical halo masses are slightly larger as seen in Fig. 18; for example, $\log(M_{\rm h}/{\rm M}_\odot) = 12.430 \pm 0.232$ for the case of 2 satellites more massive than $m_{\rm SMC}$. This estimate is somewhat larger than the one obtained by Busha et al. [44], who used the Bolshoi N-body cosmological simulation [113] to look for MW-sized halos with two subhalos with maximum circular velocity, $v_{\rm max}$, larger than 50 km/s (according to our $v_{\rm max}$–m_* relation, this mass corresponds to

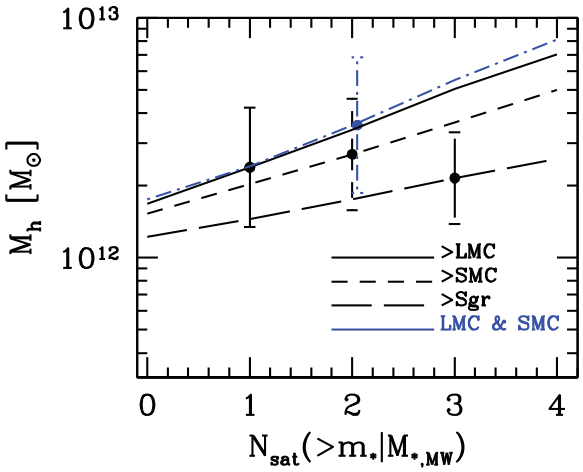

Figure 18: Mean host halo mass of the MW-sized galaxies in the mock catalog as a function of the number of satellites with masses larger or equal than the LMC, SMC, Sgr + tails, and Sgr (see the corresponding lines inside the plot). The case for the interval between the LMC and SMC is also included (blue dot-dashed line). Observe how the halo mass of MW-sized galaxies correlates weakly with the number of satellite galaxies. Nevertheless, from all the showed cases, the MW halo mass is not smaller than 1.38×10^{12} $M_\odot$ at 1σ level.

a smaller m_* than the one used here for the SMC, see Fig. 19 below; therefore, the host M_h estimated in Busha et al. [44] would be larger, in better agreement with our study, if they had used the v_{max} corresponding to m_{Sgr}). The orbital information of the MC-sized subhalos in N-body simulations has been also used for improving the statistical determinations of the MW halo mass [36, 44], finding that the typical masses should be above $\log(M_h/M_\odot) = 12.2\text{-}12.3$.

5.4 SATELLITE VS ΛCDM SUBHALO POPULATIONS

The statistical method used to construct our mock catalog allows for a connection between satellite and subhalo masses to be made. This connection is constrained by the observed satellite GSMF and the projected correlation functions at different mass bins (see RAD13), and it can be extrapolated to stellar masses lower than the completeness limit of the observational samples. In papers such as Boylan-Kolchin, Bullock & Kaplinghat [37, 38] and Lovell et al. [129], the satellite population of the MW is used to discuss the consistency of the predicted subhalo population in the ΛCDM or ΛWDM scenarios, but an uncertainty remains about whether the MW and their satellites are a typical system and what the halo mass of the MW is [e.g., 169, 224]. With our observationally-based catalog, we do not face such a problem since we account for a large population of MW-sized systems (centrals + satellites), with their corresponding host halo masses.

Our mock catalog offers a statistically complete sample of MW-sized galaxies with their satellite populations, for which we can "measure" the subhalo masses associated to the satellites. By using the tight correlation between maximum circular velocity, v_{max},

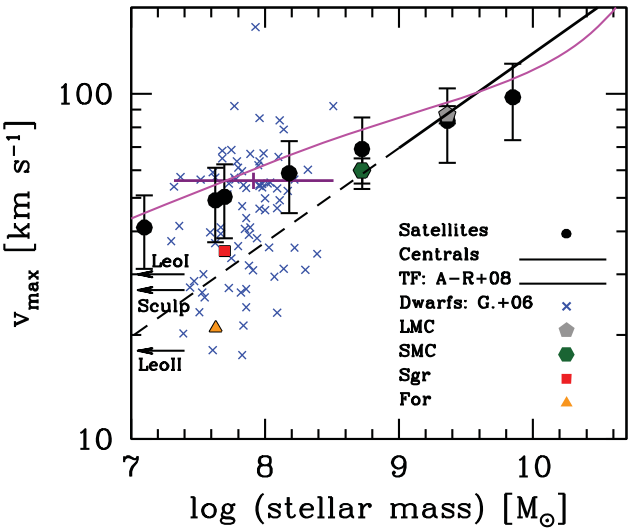

Figure 19: The internal dynamics of dwarf galaxies as a function of stellar mass. Solid circles with error bars show the mean and the standard deviation of the mock catalog subhalo v_{max} centered at different satellite stellar masses (m_*= $1.2 \times 10^7 M_\odot$, $m_{*,For}$, $m_{*,Sgr}$, $m_{*,SMC}$, $m_{*,LMC}$ and $7.1 \times 10^9 M_\odot$. The magenta long-dashed line indicates the mean v_{max}–M_* relation for the mock central galaxies. Observational estimates for the LMC [157], SMC [198], Sgr [153] and For [201] are plotted with the color filled symbols. The inferred values of v_{max} by Peñarrubia, McConnachie & Navarro [160] for the next three smaller MW dwarfs, Leo I, Sculptor, and Leo II, are indicated with arrows; their stellar masses are smaller than 10^7 $M_\odot$. The dashed line is an extrapolation to lower masses of the stellar (inverse) TF relation of field disk galaxies reported in Avila-Reese et al. [12] down to $\sim 10^9$ $M_\odot$. Individual measurements of V_{rot} and stellar mass for both central and satellite dwarfs by Geha et al. [92] are plotted with crosses. While the error bars in the Geha et al. [92] sample are different in each galaxy, on average they are of the order of $\sim \pm 5$ km/s. For a subsample of isolated dwarfs, Blanton, Geha & West [31] report a median $V_{rot} = 56 \pm 3$ km/s (violet cross).

and mass of the subhalos measured in the Millenium-II Simulation [41, taking into account the scatter around this correlation], the m_*–v_{max} relation and its scatter can be predicted. Note that implicity we are assuming that the subhalo v_{max} is not altered by baryonic effects. Therefore, in our case, the question is not about a consistency between the number of ΛCDM subhalos (above a given m_{sub} or v_{max}) and the number of satellites (above the m_* corresponding to m_{sub} or v_{max}) –this consistency was stablished by construction in the method– but about whether the predicted m_*–m_{sub} or m_*–v_{max} relations agree with direct observations.

Figure 19 shows the mean and standard deviation of the v_{max} vs. m_* relation for all the satellites above $m_* = 10^7$ $M_\odot$ around MW-sized galaxies in our mock catalog. The scatter is due to the dispersions in host halo masses, in the m_*–m_{sub} relation, and in the m_*–v_{max} relation. The dashed line is an extrapolation to lower masses of the stellar (inverse) Tully-Fisher (TF) relation of field disk galaxies as determined from a suitable catalog in Avila-Reese et al. [12, the stellar mass was shifted by -0.1 dex to convert from the diet-Salpeter to the Chabrier initial mass function]. The v_{max}–m_* relation of the satellites seems to bend towards the low-mass side of the TF relation of larger galaxies. In fact, a close trend is followed by central galaxies; the solid line in Fig. 19 shows the mean of the v_{max} vs. M_* relation for central galaxies in the mock catalog. Such a trend is in agreement with some direct observational studies of the TF relation of dwarf galaxies [c.f. 138, 70, 2]. The scatter of the v_{max}–m_* (as well as the v_{max}–M_*) relation increases towards lower masses, also in agreement with direct observational studies. The bend of the stellar TF relation at velocities below ~ 100 km/s and the increase of its scatter is also observed in cosmological numerical simulations [72] and it is explained by the strong loss of baryons due to SN-driven feedback in low-amplitude gravitational potentials.

In Fig. 19 we also plot the individual measurements of the maximum rotation velocity V_{rot} and stellar mass for (central and satellite) dwarf galaxies by Geha et al. [92, crosses][2]. The scatter is high, and down to stellar masses $\sim 10^8$ $M_\odot$ most of dwarfs are close to those from our catalog and above the extrapolated TF relation. There are some hints that those dwarfs in the high-V_{rot} side in the Geha et al. [92] sample tend to be centrals. For example, Blanton, Geha & West [31] select the subsample of very isolated dwarfs from Geha et al. [92]; these are certainly central galaxies. They report a median $V_{rot} = 56 \pm 3$ km/s for this subsample which spans almost all the mass range of the total sample. This value is also plotted in Fig. 19 (violet cross) and it agrees well with the velocities of our central galaxy sample.

For masses smaller than 10^8 $M_\odot$, there is a significant fraction of observed dwarfs with lower v_{max} than the mock dwarfs, although the scatter is high. We also plot the values of v_{max} and m_* inferred observationally for the MW satellites LMC, SMC, Sgr, and For (color filled symbols; for the sources see the figure caption), as well as the inferred values of v_{max} by Peñarrubia, McConnachie & Navarro [160] for the next three smaller MW dwarfs, Leo I, Sculptor, and Leo II (indicated with arrows; their stellar masses are certainly smaller than 10^7 $M_\odot$). While the LMC and SMC fall close to the

2 Note that (i) in several cases the HI line widths used to estimate V_{rot} underestimates the real maximum velocity that could be at a radius larger than that where gas is observed; and (ii) the galaxy+subhalo v_{max} after baryon matter is included in the numerical simulations may be lower than in the pure dark matter subhalo v_{max} (see for a discussion Section 5).

mock satellites, the observational inferences of v_{max} for Sgr and For are smaller than the mean of the mock satellites; a similar difference is expected for the next smaller dwarfs, Leo I, Sculptor, and Leo II. Even the lower-1σ scatter, given mainly by those systems in low-mass host halos, is higher than the For v_{max}.

For a large sample of galaxies (and not only for the MW galaxy satellites), the results shown above confirm *a potential problem for the small dwarf galaxies (stellar masses* $\lesssim 10^8$ $M_\odot$*): they seem to be associated to significantly less concentrated (smaller* v_{max}*) systems than those the ΛCDM scenario predicts* [37, 38]. However, the question that remains open is *whether this problem refers to both central and satellites galaxies or only to the latter.* According to the above, it could be that those dwarfs in the Geha et al. [92] sample that are in the low-V_{rot} side are satellites, while those in the high-V_{rot} side are centrals, and as can be appreciated in Fig. 19, they are consistent with the mock central dwarfs.

5.5 SUMMARY & DISCUSSION

??

By means of a statistical approach that observationally constrains the galaxy-(sub)halo connection for central and satellite galaxies, we generate a realization of 2×10^6 central galaxies and their populations of satellites. Each galaxy is characterized by its stellar and (sub)halo mass and, by construction, the catalog reproduces (i) the observed central/satellite GSMFs and projected two-point correlation functions in several stellar mass bins down to their completeness limits ($m_* \sim 2.5 \times 10^8$ $M_\odot$, though we extrapolate it down to $\sim 10^7$ $M_\odot$); (ii) the ΛCDM distinct halo mass function [207] as well as the conditional subhalo mass functions [41]. From this catalog we picked all the central galaxies with MW stellar mass, $\log M_* = 10.74 \pm 0.1$ dex (40694 objects), and studied the (massive) satellite occupational distributions of these galaxies. The main results from the "observation" of the MW-like systems in the mock catalog are:

- The fractions (probabilities) of MW centrals with N_{MC} satellites within the MCs stellar mass range or above the SMC or LMC masses are in general agreement with direct observational studies [127, 44, 174] after correcting for the satellite search radius, which in our case is the virial radius of the host halo (see Figs. 15 and 16). For example, we obtain that the probability of finding 2 satellites in the MC mass range is $\sim 6.6\%$ (or $\sim 0.08\%$ if we add the condition of having no satellites more massive than the LMC); the probabilities of having 1 satellite with $m_* \geq m_{LMC}$, 2 satellites with $m_* \geq m_{SMC}$ or 3 with $m_* \geq m_{Sgr}$ are 26.1%, 14.5%, and 14.3%, respectively. We also find that the average (massive) satellite mass function of the mock MW-sized galaxies is consistent with recent direct observational determinations of the (bright) conditional satellite luminosity function.

- Having the two most massive satellites be as massive as the MCs makes the MW less common, but it is not a rare case in the sense of an outlier. In our catalog, MW-sized galaxies with one satellite $\geq m_{LMC}$ (two satellites $\geq m_{SMC}$), as the MW, happen only 1/0.6=1.68 (2/1.02 =1.92) times less frequently than the average. The cumulative satellite mass function of the MW down to the mass of For is within the 1σ distribution of all the MW-sized galaxies, lying above the average by less

than a factor of two at the MCs masses and close to the average at the For and Sgr masses (Fig. 17). MW-sized centrals with exactly 3 satellites more massive than Sgr are among the most common ones, but they have typically the most and second most massive satellites smaller than the LMC and SMC by roughly 0.7 and 0.8 dex, respectively. In general, we find that the satellite number distributions of MW-sized galaxies are relatively broad.

- As opposed to clusters and groups of galaxies, the halo mass $M_{\rm h}$ of MW-sized galaxies correlates weakly with the number of satellites above a given mass (Fig. 18). The mean $\log M_{\rm h}$ and its standard deviation for galaxies with 3 satellites equal or more massive than Sgr is 12.33 ± 0.19. For 2 satellites with $m_* \geq m_{\rm SMC}$ or m_* in the SMC-LMC mass range, the mean and standard deviation are $\log M_{\rm h} = 12.430 \pm 0.232$ and $\log M_{\rm h} = 12.552 \pm 0.283$, respectively. Therefore, it is not possible to constrain the halo mass of MW-sized galaxies with appreciable accuracy with the satellite population abundance of the MW, but one can say that, at the 1σ level, this mass is not smaller than 1.38×10^{12} $M_\odot$, consistent with recent claims based on the combination of high numerical simulations with the proper motion of Leo I [39].

- In our catalog of MW-sized galaxies, the number of ΛCDM subhalos (above a given $m_{\rm sub}$ or $v_{\rm max}$) is consistent with the number of satellites (above the m_* corresponding to $m_{\rm sub}$ or $v_{\rm max}$) by construction and, since the satellite abundances of the mock galaxies are in agreement with different direct observational studies, one may conclude that there is not a (massive) satellite missing problem for the ΛCDM model. However, we find *an internal dynamics problem*: the $v_{\rm max}$ of the subhalos of satellites smaller than $\sim 10^8$ $M_\odot$ seems to be systematically larger than the $v_{\rm max}$ inferred from current observational studies of dwarf satellites, by factors $\sim 1.3 - 2$ at the masses of Sgr and For (Fig. 19). There are some hints that this issue could refer only to the satellite dwarfs but not to the central ones.

We conclude that the general agreement of our satellite abundance statistics with direct observations signals towards a self-consistency in the construction of the mock catalog, and it shows that *the underlaying ΛCDM (sub)halo abundances and internal dynamics are consistent with observations down to the scales of the MC galaxies.* For smaller masses, our results point out to a possible issue in the internal dynamics of the ΛCDM (sub)halos as compared with the observed satellite dwarfs. These results confirm the conclusions by Boylan-Kolchin, Bullock & Kaplinghat [37, 38] for satellite spheroidal dwarfs *but with the difference that in our case the results refer to the overall population of MW-sized galaxies*. Therefore, our conclusions are free of uncertainties intrinsic to the analysis in Boylan-Kolchin, Bullock & Kaplinghat [37, 38] about whether the MW system is atypical or not and and also in regards to the MW halo mass to be used [e.g., 169, 224]. However, before arriving to any conclusion, several aspects of these results should be carefully discussed (for an extensive discussion see Boylan-Kolchin, Bullock & Kaplinghat [37, 38]). Here we highlight two observational caveats.

(i) Our prediction refers to the maximum circular velocity of the pristine subhalo, $v_{\rm max}$. Observations refer to the maximum or last-point measured galaxy rotation velocity, $V_{\rm rot}$,

or to a model-dependent $v_{\rm max}$ constrained by measurements of the stellar velocity dispersion under several assumptions. Because dwarf galaxies are dark matter dominated, in the context of the ΛCDM model it is expected that $v_{\rm max}$ is attained at a radius much larger than the optical one, where the observational tracers are not available. Then, it could be that the current observational inferences are underestimating the actual values of $v_{\rm max}$. Regarding the dispersion-supported dwarf spheroidals, their unknown stellar velocity anisotropy and/or halo shape make the inference of their mass distributions ambiguous [e.g., 200, 104, 242, and more references therein].

(ii) The mock catalog was constructed using both the m_*–$m_{\rm sub}$ and the M_*–$M_{\rm h}$ relations constrained with the Yang, Mo & van den Bosch [247] central and satellite GSMFs down to $\sim 2\times10^8{\rm M}_\odot$, as well as with observed projected correlation functions reported in Yang et al. [250]. For smaller masses, we use just an extrapolation of this relation and its scatter. If the satellite GSMF at smaller masses would be steeper than the Yang, Mo & van den Bosch [247] faint-end or the scatter larger than the assigned by us (due, for example, to highly stochastic star formation and tidal effects in the satellite dwarfs), then the relation would be shallower and more scattered, which implies lower subhalo masses (or $v_{\rm max}$) on average at a given m_* and higher scatter in these quantities. In Rodríguez-Puebla, Drory & Avila-Reese [178], by using a low-mass slope of -1.6 for the satellite GSMF, we obtained subhalo masses for $m_* = 10^7 - 10^8$ M$_\odot$ dwarfs as small as the tidal masses (close to the subhalo masses) inferred for some MW satellites. This slope is given by Baldry, Glazebrook & Driver [16] for the GSMF, which goes down to $\sim 2.5\times10^7$ M$_\odot$ after applying a correction for surface brightness incompleteness. However, the GSMF in this case refers to all galaxies. In RAD13 we decomposed the Baldry, Glazebrook & Driver [16] GSMF into centrals and satellites, and obtaining a m_*–$m_{\rm sub}$ relation that implies subhalo masses larger than the tidal masses by roughly 0.3-0.4 dex. Future samples, complete down to the smallest masses and decomposed into central and satellite galaxies, should tell us whether the satellite GSMF towards very small masses is steep enough or not as to imply subhalo masses (or $v_{\rm max}$) in better agreement with current dynamical studies.

Finally, if the observations regarding the faint-end of the satellite GSMF and the internal dynamics of the dwarf satellites remain roughly as those discussed here, then our results could be an indication that baryonic physics significantly affects the inner structure of the very small subhalos that host dwarf satellites. A possible physical mechanism for explaining the decrease of the inner concentration, and therefore of $v_{\rm max}$, in low-mass ΛCDM (sub)halos could be feedback-driven gas outflows. By means of N-body/Hydrodynamics cosmological simulations, some authors have shown that repeated strong outflows during the halo/galaxy growth are able to drag with them the inner dark matter decreasing the inner gravitational potential [136, 130, 98, 264, 156], though it seems difficult that such an effect would be able to lower $v_{\rm max}$ to the required values [38, 73]. However, some numerical simulations show that *in the case of satellites*, the combination of this effect with the stronger tidal effects of the halo when a central baryonic galaxy is included, as well as the lowered baryon fractions of the dwarf satellites, work in the direction of reducing the circular velocities of the simulated MW satellite dwarf spheroidals to the levels required by the results of Boylan-Kolchin, Bullock & Kaplinghat [38] or our ones [42, 4, 99].

We have found some hints that the apparent problem of too low-circular velocities of dwarfs smaller than $m_* \sim 10^8$ $M_\odot$ refers mostly to satellite galaxies but not to central ones. If this is the case, then such a problem is explained by the plausible physical mechanisms mentioned above. However, if the *problem remains for isolated dwarfs*, then this could the necessity of a modification in the cosmological scenario, for example, by introducing warm or self-interacting dark matter [129, 222].

ACKNOWLEDGMENTS

We are grateful to Dr. J. Zavala for thoughtful comments on the draft of this paper. We also thank to the anonymous referee for a constructive report that helped to improve this paper. A. R-P. acknowledges a graduate student fellowship provided by CONACyT. V. A-R. and N. D. acknowledge CONACyT (ciencia básica) grants 167332 and 128556.

Part IV

THE STELLAR-(SUB)HALO MASS RELATIONS UP TO $z \sim 1$

En esta Parte IV, se estudian las relaciones masa estelar galaxia central-masa halo, M_*-$M_{\rm h}$, y masa estelar galaxia satélite-masa subhalo, m_*-$m_{\rm sub}$, por separado, en cuatro diferentes intervalos de corrimiento al rojo, hasta $z \sim 1$. En el Capítulo 6 se extiende el modelo estadístico presentado en el Capítulo 3 a otras épocas, utilizando la función de masa estelar de galaxias y la función de correlación angular de dos puntos del catastro COSMOS hasta $z \sim 1$. Se muestra que la evolución de las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm acc}$ es débil. Los principales cambios son: el decrecimiento ligero con el tiempo de la masa $M_{\rm h}$ al pico del cociente $f_* = M_*/M_{\rm h}$ ("downsizing" en masa); una ligera caída y empinamiento hacia z's más altos de la parte de bajas masas de la relación M_*-$M_{\rm h}$, así como una ligera subida de la parte de masas altas de la relación m_*-$m_{\rm acc}$. El pico de f_*, la eficiencia de crecimiento de masa estelar, para galaxias centrales es aproximadamente constante, $f_* \approx 0.03$, mientras que para galaxias satélites decrece ligeramente con z, siendo en promedio $f_* \approx 0.04$. Se encuentra también que la fracción de galaxias satélites con relación al total a una dada masa estelar es siempre menor a ~ 0.35, decreciendo ligeramente hasta $M_* \sim 2 \times 10^{10}$ $M_\odot$ y luego fuertemente a masas mayores donde la gran mayoría de las galaxias son centrales; este comportamiento se observa en todas las épocas pero las fracciones en sí disminuyen hacia altos corrimientos al rojo, es decir existían menos galaxias satélites en el pasado. Se discuten ampliamente las implicaciones de nuestros resultados en lo que concierne al ensamblaje de la masa estelar de las galaxias en función de la escala.

6

THE EVOLUTION OF THE STELLAR-TO(SUB)HALO MASS RELATIONS OF CENTRAL/SATELLITE GALAXIES AND OF THE SATELLITE FRACTION FROM THE COSMOS SURVEY

ABSTRACT

By means of a statistical model to connect galaxies to cold dark matter halos, we infer the redshift evolution of stellar-to-(sub)halo mass relations separately for central and satellite galaxies, CHMR and SSMR, respectively. The observational data are the galaxy stellar mass function, GSMF, and the angular two-point correlation function from $z = 0.2$ to $z = 1$ extracted from the COSMOS survey. Our main result shows that the CHMR and SSMR change little with redshift. At most, the low-mass end of the former slightly decreases with z, and the high-mass end of the latter, slightly increases with z. For central galaxies, the halo mass at which the peak of the stellar mass growth efficiency is reached, $f_* = M_*/M_{\rm h}$, shifts weakly towards lower masses at lower z (downsizing in mass). This efficiency is roughly the same since $z \sim 1$ with a peak value of $f_* \sim 0.03$. For satellite galaxies, f_* has a slightly different evolution and slightly larger values, explained by the fact that the subhalo mass at the accretion time is on average smaller than the current halo masses of central galaxies. The fraction of satellite galaxies decreases with stellar mass, strongly for masses $> 2 \times 10^{10}$ $M_\odot$, at all redshifts. With time, the fractions at each mass increases; the maximum value of this fraction is at the lowest redshift and mass bins, ~ 0.35. We find that a toy model, where the CHMR is the same at all z, describes reasonable well our main results regarding the central/satellite GSMFs, the satellite fractions, and the satellite CSMFs at different z. However, this model implies a constant specific star formation rate with M_* at low masses, in tension with direct determinations of the SSFR vs M_* of local and high-redshift galaxies.

6.1 INTRODUCTION

In recent years, the galaxy-halo connection has been routinely inferred by using an ample variety of approaches, from those able to probe directly the masses of the dark matter halos (by using galaxy weak lensing, [133, 132, 184], kinematics of satellite galaxies, [56, 145, 146, 241], and galaxy clusters, [249, 247, 103, 250]) to those that indirectly infer halo masses. The latter are based on statistical analysis, for instance the Halo Occupation Distribution (HOD) model [25, 60, 259], the Conditional Stellar Mass Function (CSMF) formalism [245, 249, 148, 121, 122, 250], and the Abundance Matching Technique [AMT 216, 115, 58, 187, 234, 16, 217, 57, 75, 148, 19, 102, 20, 170, 159]. While improvements on direct techniques have been developed in recent years, they are still limited due to their need to stack large ensembles in order to attain a acceptable signal to noise. In addition, the dynamical range they can probe is limited. Therefore, indirect approaches have received much attention due to their relative simplicity and the ample dynamical range

for connecting galaxies to halos, not only for local galaxies but at different redshifts, [e.g., 122, 250].

Indirect approaches can be divided into two main approaches. The first approach, which includes the HOD model, and the closely related CSMF formalism, describes the probability that a halo of mass $M_{\rm h}$ host a N number of galaxies above some luminosity/stellar mass (M_*) threshold which can be constrained by using the observed spatial clustering of galaxies. The second approach, the AMT, assumes a one-to-one monotonic relationship between luminosity or M_* and $M_{\rm h}$, and it is constrained by matching the observed cumulative galaxy luminosity or stellar mass function to the theoretical halo plus subhalo cumulative mass function. The indirect techniques assume that at the center of each dark matter halo a central galaxy is located, with satellite galaxies orbiting within the virial radius of the central galaxy. These satellite galaxies are thought to be formed in dark matter halos prior to become subhalos. As a result, a central stellar-to-halo mass and satellite stellar-to-subhalo mass relations are stablished. In the case of the HOD model and the CSMF formalism, the satellite-to-subhalo connection is not implemented. In the case of the AMT, this connection is implemented but under the assumption that it is identical to the one of the central galaxies. Consequently, little it is known about the satellite-to-subhalo connection both for local and high redshift galaxies (but see Rodríguez-Puebla, Drory & Avila-Reese [178], Rodríguez-Puebla, Avila-Reese & Drory [176], Neistein et al. [152] and Part II for local galaxies, and Watson & Conroy [231] for a first approach at higher redshifts). Understanding the details of the satellite-to-subhalo connection is of vital importance. It provides information about how the stellar mass of satellite galaxies have evolved since their accretion, and shed light on the physical process that may affect their evolution.

In Chapters 2 and 3 we have developed an approach that combines the AMT, the HOD model, and the CSMF formalism in order to determine *separately* the central stellar-to-halo mass and the satellite stellar-to-subhalo mass relations. In this approach the subhalo mass is defined either at the time of accretion or at the current (observed) time. We have applied the approach to obtain the local ($z \sim 0$) central stellar-to-halo and satellite-to-subhalo mass relations. In the present Chapter, we extend this approach to higher redshifts. To that end, we utilize the Cosmic Evolution Survey [COSMOS, 185], which allows to estimate *both* the galaxy stellar mass function (GSMF) and the galaxy spatial clustering at different redshift, up to $z \sim 1$. Unfortunately, the COSMOS survey is affected by cosmic variance. In order to overcome this issue, we will introduce some corrections.

The inferences of the stellar-to-(sub)halo mass relations at different epochs provide relevant constraints to galaxy evolution as a function of mass. For example, if the mass relation at masses below the knee ($M_{\rm h} \lesssim 10^{12}$ M$_\odot$) decreases with z, then a fast late stellar mass assembly for low-mass galaxies is expected. Instead, a non-evolving mass relation, implies little M_* growth and therefore, low specific star formation rates for these galaxies. By connecting the mass relations at different redshifts with the predicted mean halo mass assembly histories in $N-$body simulations, the respective mean galaxy M_* histories can be inferred as has been down preliminary in the pioneering works by Conroy & Wechsler [57] and Firmani & Avila-Reese [84].

In recent years, several determinations of the global stellar-to-halo mass relation up to high redshifts appeared in the literature [19, 20, 148, 147, 61, 122, 250]. Nevertheless, whether this mass relation changes or not with redshift it is yet a matter of great debate. By means of our approach, we will be able to obtain accurate constrains because, contrary to previous works, we handle separately the central-halo and the satellite-subhalo connection. Furthermore, our determinations will be carried out in narrower redshift bins than most of previous works.

During the preparation of this work, Watson & Conroy [231] submitted a paper in which they also constrain the mass relation of satellite galaxies. However, these authors do not constrain both the central and satellite mass relations in a self-consistent way.

This Chapter is organized as follows, In section 6.2, we describe the extended AMT and HOD+CSMF model. In Section 6.3, we present the data employed to constrain the model parameters, consisting of the total GSMF and the angular projected correlation function from the COSMOS survey from redshift $z = 0.2$ to $z = 1$. Section 6.4, presents the procedure and the uncertainty treatment to constrain the mass relations of centrals and satellites. In section 6.5, we present our results about the redshift evolution of the central-to-halo mass relation and the satellite mass relation. In section 6.6, we discuss on the possible consequence of the trends obtained for the mass relations. Finally, in section 6.7, we present our conclusions.

We adopt cosmological parameter values close to WMAP 7: $\Omega_\Lambda = 0.73, \Omega_M = 0.27, h = 0.70, n_s = 0.98$ and $\sigma_8 = 0.84$.

6.2 THE GALAXY-(SUB)HALO CONNECTION MODEL

In the following, we describe briefly the semi-empirical approach used to connect the stellar masses of central and satellite galaxies with their respective masses of halos and subhalos. The basic idea behind this approach relies on the assumption that the central stellar-to-halo mass and the satellite stellar-to-subhalo mass relations (CHMR and SSMR, respectively) are monotonic. Clearly, if the dark matter halos are the hosts of central galaxies and the subhalos are the hosts of satellite galaxies, the CHMR and SSMR provide a direct link to the mass function of central and satellite galaxies, respectively. Furthermore, given the conditional subhalo mass function, or equivalently, the subhalo occupational numbers (usually obtained from $N-$body cosmological simulations [see e.g., 41, , and references therein]), the SSMR offers the most practical solution to the conditional stellar mass function of satellite galaxies, or equivalently, the galaxy occupational numbers. Therefore, this approach naturally encapsulates the main ideas behind the abundance matching technique, the halo occupation distribution model, and conditional stellar mass function formalism. A detailed description of the approach and its application for constraining the CHMR and the SSMR at $z \sim 0$ have been presented in Part II or in RDA12 and RAD13. *Here we will apply the approach to constrain the CHMR and the SSMR at higher redshifts, up to $z \sim 1$, by using the data provided by COSMOS.*

Conceptually, the main steps in our statistical approach at each redshift can be summarized as follows:

1. Propose parametric functions for the mean CHMR and SSMR. In general, these parametric functions are not necessarily the same. Assume that both central and satellite galaxies are drawn from lognormal distributions with mean CHMR and CHMR and scatters σ_c and σ_s.

2. Define the theoretical distinct halo and the conditional subhalo mass functions (HMF and subHMF, respectively); for the latter, the subhalo mass can be defined at the observation time, $m_{\rm obs}$, or at the time when the subhalo has been accreted into the distinct halo, $m_{\rm acc}$.

3. Select the observed GSMF; if available, use the GSMF decomposed into central and satellite galaxies, and/or the observed spatial clustering of galaxies; otherwise, combine the information of the global GSMF and the observed spatial clustering of galaxies (the latter is the case for samples at high redshifts, where the decomposition into central and satellite galaxies is difficult).

4. Assume that the spatial distribution of subhalos, and hence of satellite galaxies, follow the halo mass density profile[1]. Assume also that the number distribution of subhalos inside a distinct halo follows a Poissonian distribution (this assumption is based on results from N-body cosmological simulations, e.g., Kravtsov et al. [115])

5. By means of a fitting procedure of the model to the selected observational data, constrain the parameters of the CHMR and SSMR and their scatters.

The advantage of this approach is that self-consistently relates the GSMF decomposed into centrals and satellites, the HMF/subHMF, the satellite conditional stellar mass relations (CSMFs), and the galaxy two-point correlation functions (2PCFs). Recall that unlike previous models of this kind, the CHMR and the SSMR are handled separately. Recently, Watson & Conroy [231], presented a similar approach that also allows to constrain the SSMR, but using as input the previously determined total CHMR.

As we have concluded in RAD13 (Chapter 3), the constrained mass relations and their uncertainties are similar using either the observed GSMF decomposed into centrals and satellites plus the projected 2PCFs (set C in RAD13) or only the total GSMF plus the 2PCFs (set B in RAD13). Note that in the latter case, the GSMF decomposition into centrals and satellites is a prediction of the model. Similarly to our set B in RAD13, in this work we constraint the mass relations by using the total GSMF and the spatial clustering of galaxies obtained from the COSMOS survey, then the corresponding central and satellite GSMFs are also predictions of the model.

In the following, we briefly describe our theoretical assumptions, and in the next Section we describe the observational data used to constrain the overall model.

6.2.1 *Theoretical assumptions*

As it is commonly assumed in the literature (e.g., 250, RAD13, and more references therein), we model the distribution of central galaxies at a fixed halo mass, $P_{\rm cen}(M_*|M_{\rm h})$,

1 Of course this is only an approximation since detailed studies in $N-$body simulations have shown that this may not be the case specially near the host halo centre, (e.g. Springel et al. [197])

as a lognormal distribution, with the parameters $\log M_*(M_h)$ and σ_c denoting, respectively, the mean logarithmic stellar mass and the standard deviation in units of dex. For satellites galaxies, we assume also a lognormal distribution for $P_{sat}(M_*|m_{sub})$, with a standard deviation σ_s in units of dex. In essence, the probability distribution functions $P_{cen}(M_*|M_h)$ and $P_{sat}(M_*|m_{sub})$ encode all the physical processes involved in galaxy evolution.

For both $M_*(M_h)$ and $M_*(m_{sub})$ at a given redshift, we adopt the functional form proposed in Behroozi, Wechsler & Conroy [20],

$$\log M_* = \log(\epsilon M_1) + f(\log(M/M_1)) - f(0), \tag{46}$$

where

$$f(x) = \delta_c \frac{(\log(1+e^x))^{\gamma_c}}{1+e^{10^{-x}}} - \log(10^{\alpha_c x}+1). \tag{47}$$

Here M applies either to distinct halo or subhalo mass. This function behaves as power law with slope α at masses much smaller than $M_{1,c}$, and as a sub-power law with slope γ_c at larger masses. The main motivation for this functional form is its ability to map Schechter-like GSMFs from power-law halo or subhalo mass distributions; in the mass ranges of interest, the ΛCDM HMF and subHMF are indeed power laws.

For the the distinct HMF, we use the accurate fits to ΛCDM cosmological N-body simulations presented in Tinker et al. [207]. Here we define the halo masses at the virial radius, i.e. the radius where the spherical overdensity is Δ_{vir} times the mean matter density, with $\Delta_{vir} = (18\pi^2 + 82x - 39x^2)/\Omega(z)$, and $\Omega(z) = \rho_m(z)/\rho_{crit}$ and $x = \Omega(z) - 1$. [43].

For the conditional subHMF, we use the results obtained in Boylan-Kolchin et al. [41] based on the Millennium-II simulation at redshift $z = 0$. Previous studies based on either the analysis of $N-$body simulations or using extended Press-Shechter formalism have shown that (1) the conditional subhalo mass function is self-similar, i.e., independent of halo mass [91, 69, 115, 210, 20] and (2) it evolves little with redshift [220, 94]. Based on these results, for all the redshifts studied here, we used the fitting form obtained in Boylan-Kolchin et al. [41] at $z = 0$ but re-scaled to our cosmology (see section 2 of RAD13 or Chapter 3). Our method allows to use the subhalo mass defined either at the accretion time (m_{sub}=m_{acc}) or at the current (observation) time (m_{sub}=m_{obs}). In this work, we use the first definition in order to compare with some previous works based on the AMT (see Chapters 2 and 3).

For the dark matter halo density profile, used to describe the spatial distribution of subhalos, we assume the Navarro-Frenk-White profile, which can be specified by M_h and the concentration parameter c_{NFW}. For the values of c_{NFW}, we use the main $c_{NFW}-M_h$ relation at different epochs as reported from large N-body cosmological simulations in Muñoz-Cuartas et al. [150].

From the COSMOS survey we measure the angular 2PCF at a given redshift bin $[z_{min}, z_{max}]$. In order to connect it to the 3D 2PCF, ξ_{gg}[2], we use the Limber's equation:

$$\omega(\theta) = \frac{2}{c}\int_{z_{min}}^{z_{max}} H(z)\left(\frac{dn}{dz}\right)^2 dz \int_0^\infty \xi_{gg}(\sqrt{u^2 + x^2(\bar{z})\theta^2})du, \tag{48}$$

2 The clustering of galaxies is computed using the HOD model [25, 60]. The HOD model describes the probability that a halo of mass M_h hosts a number of N galaxies with stellar masses greater than M_*. Once

where c is the speed of the light, $H(z)$ is the Hubble parameter at the redshift z, dn/dz is the normalized redshift distribution of the galaxies in the sample and $x(\bar{z})$ is the comoving distance to the median redshift. For the calculation of the angular clustering, we consider dn/dz to be constant at each redshift bin.

6.2.2 *Parameters in the model*

Our model consists in total of ten free parameters for a given redshift bin. Five parameters are to describe the CHMR: $M_{1,c}$, ϵ_c, α_c, δ_c, and γ_c; and five more, to describe the SSMR (and therefore the satellite occupational numbers): $m_{1,s}$, ϵ_s, α_s, δ_s, and γ_s.

While in principle, both σ_c and σ_s, the intrinsic scatters of the CHMR and SSMR, could be considered as two more free parameters in the model, we opt to leave them fixed, independent of halo mass, as well as of redshift, and with a value of 0.173 dex. The lack of information about the central/satellite GSMFs at higher redshifts makes uncertain the restriction of these parameters. The assumption is well justified at $z \sim 0$ by results based on large halo-based group catalogs (YMB09), analysis of satellite-galaxy kinematics [145, 146], and analysis based on the galaxy spatial clustering [245, 59, 45, 122, see also RAD13]. While there is no systematic studies of these scatters at higher redshifts, a few attempts to measure them have been performed using the galaxy spatial clustering [61], and in combination with galaxy-galaxy weak lensing [see e.g., 122]. These studies show that the scatter around the CHMR is nearly constant and independent of $M_{\rm h}$ and z. In addition, the value of σ_c has been found consistent with the value obtained for local galaxies. On the other hand, the scatter σ_s around the SSMR has not yet been discussed in detail in the literature. In RDA12 (Chapter 2), at $z \sim 0$, it was assumed to be the same as for central galaxies, giving results consistent with the observed projected 2PCFs and satellite CSMFs. In RAD13 (Chapter 3), using the decomposition of the GSMF into centrals and satellites, together with the projected 2PCFs, we have shown that when σ_c and σ_s are left as free parameters, the constrained values are actually very similar and close to 0.173 dex. In conclusion, we assume that both σ_c and σ_s are the same at all redshifts.

In Section 6.4 we describe how the model parameters are constrained by by the observational data (presented in the next Section)

6.3 THE DATA

In order to constrain the galaxy-to-(sub)halo mass relations at different epochs both for central and satellite galaxies, one requires a large and deep galaxy survey, from which we can measure the GSMFs (at least the total one) and the galaxy clustering at different epochs. This is the case of the COSMOS survey [185]. This survey covers an area of $2\mathrm{deg}^2$, containing approximately 60000 galaxies with photometric redshifts derived from 30 broad and medium bands in the redshift interval $0.2 < z < 1$ [e.g., 108, 46].

the occupational numbers are defined, $\xi_{\rm gg}$ can be computed by assuming that the total mean number of galaxy pairs is the contribution of all pairs coming from galaxies in the same halo (one-halo term) and pairs from different halos (two-halo term). For a detailed description for the HOD model employed here, see RAD13 or Chapter 3.

In the following, we summarize the observational data utilized to constrain the model parameters. Firstly, we briefly describe the GSMFs obtained in Drory et al. [75] from the COSMOS survey. We then describe how we measure the angular correlation functions for various mass thresholds in the COSMOS survey within the same redshifts bins.

6.3.1 *The galaxy stellar mass functions*

We use the GSMFs from the COSMOS survey obtained in Drory et al. [75] in four redshift bins: $0.2 < z < 0.4$, $0.4 < z < 0.6$, $0.6 < z < 0.8$ and $0.8 < z < 1$. Drory et al. [75] obtained the stellar masses through the full multi-band photometry spectral energy distribution fitting using stellar population models with a Chabrier [48] IMF.

Unfortunately, the volume the coverage of the COSMOS survey is relatively small in such a way that the sample suffers of cosmic variance issues. Several galaxy clustering analysis of this survey have reported a significant excess clustering at large scales, i.e., a large fraction of galaxies seem to lie in over-dense structures [140, 114, 67]. This excess, which affects galaxies in some redshift intervals within $0.6 < z < 1.0$, could be a potential source of uncertainty since the GSMF has some dependence on environment. In the following we describe how we attempt to correct the Drory et al. [75] GSMFs at a first order due to this environment effect. For the excess clustering correction see next subsection.

The strongly clustered nature of the COSMOS survey has been identified at different redshifts but seems to be more important at redshifts around $z = 0.9$ Kovač et al. [114], de la Torre et al. [67]. In essence, one expects that the galaxy number density in high density environments is higher compared to the galaxy number density in the field, see e.g., Tempel et al. [206]. In order to correct to first order the COSMOS GSMF at $z \sim 0.9$, we have renormalized the GSMF by a factor of ~ 0.8. We obtain this correction by comparing with different GSMFs published in literature coming from different surveys [e.g., 149]. On the other hand, at redshift $z \sim 0.5$, apparently there is a large underdensity which results in a lower GSMF by a factor of ~ 2 [see 75]. In comparison with other published GSMFs at this redshift we opted to correct the GSMF by a factor of ~ 1.6.

Figure 20 shows the corrected COSMOS total GSMFs at the four redshift bins (dots with error bars). The stellar masses at which the sample in each redshift bin is complete increases from $\sim 5 \times 10^8$ $M_\odot$ at $z \sim 0.3$ to $\sim 3 \times 10^9$ $M_\odot$ at $z \sim 0.9$. The solid lines are the best model fit to the observations (the uncertainties of the fits are not shown to avoid confusion in the plot. Nevertheless, not that these are of the order of the observed ones). The shaded areas show the corresponding 1σ uncertainty regions of the model (predicted) central and satellite GSMFs. The solid lines correspond to the total GSMF and its decomposition into centrals and satellites, calculated under the assumption that the stellar-to-(sub)halo mass relation is constant in time (see Sect. 6.5 below).

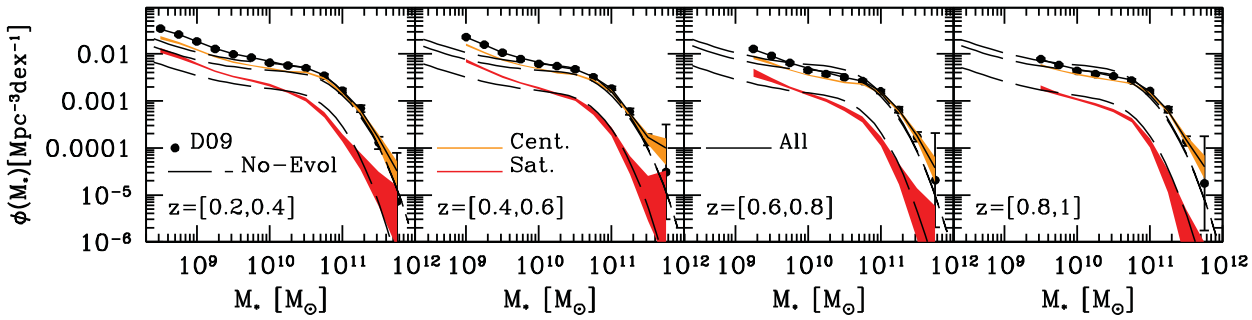

Figure 20: The GSMFs for all, central and satellite galaxies at four redshift bins. The best fit model to the total GSMF is indicated with the solid line, while the resulting central and satellite GSMFs are indicated respectively with the orange and red shaded areas which represent the 1σ confidence level. The model predictions for no evolution in the CHMR and SSMR are indicated with the long dashed lines. Filled circles indicate the total Drory et al. [75] GSMF from the COSMOS field.

6.3.2 *The Angular Correlation function*

In this Section we present the measurements of the angular 2PCF, $\omega(\theta)$, in the same redshift bins where the GSMFs were reported. In addition, for each z bin, the angular correlation function is computed for several stellar mass thresholds, spaced equally in bins of 0.5 dex in $\log M_*$. Notice that the lowest mass threshold in each z bin is defined to correspond to the lowest M_* limit of the GSMFs. Our galaxy sample is based on the COSMOS galaxy survey, version 1.8, dated July 2010.

To estimate the galaxy clustering at each mass threshold, we use the Landy & Szalay [119] estimator,

$$\omega_{LS}(\theta) = 1 + \frac{DD(\theta)}{RR(\theta)}\frac{N_R^2}{N_D^2} + 2\frac{DR(\theta)}{RR(\theta)}\frac{N_R}{N_D} \tag{49}$$

where $DD(\theta)$, $RR(\theta)$ and $DR(\theta)$ are the data-data, random-random and the data-random pairs, respectively. The number of galaxies in the data and random catalogs are denoted by N_D and N_R, respectively. For each correlation function measured, we create a random catalog with 1.5×10^5 galaxies covering the same area as the original sample.

To estimate the errors in the galaxy clustering for each mass threshold sample we use the bootstrap method. In this method one forms a set of N_{rs} resampling of the original sample and of the same size by randomly picked galaxies from the original sample and replacing them with the original sample. Then, the covariance matrix is given by,

$$C(\omega_i, \omega_j) = \frac{1}{N_{rs} - 1}\sum_{k=1}^{N_{rs}}(\omega_j^k - \omega_{LS,j})(\omega_i^k - \omega_{LS,i}) \tag{50}$$

where ω_j^k is the value of the two point angular correlation function at the jth bin in the kth resampling and we have adopted $N_{rs} = 5$.

6.3.2.1 *The integral constraint*

In small sample areas, as in the COSMOS survey, measurements of the clustering at large angles are biased, this is because one cannot estimate the true $\omega(\theta)$ when the angles are comparable to the survey area. The impact of this effect is that the estimation of the angular clustering is negatively offset from the true $\omega(\theta)$. This offset, C, can be estimated by carrying out the following integration over the field area Ω:

$$C = \frac{1}{\Omega^2} \int \int \omega(\theta) d\Omega_1 d\Omega_2, \tag{51}$$

where C is known as the integral constraint. In practice, one can use the random-random pairs to carry out this integral numerically,

$$C = \frac{\sum \omega(\theta) RR(\theta)}{\sum RR(\theta)}. \tag{52}$$

Traditionally, the estimation for C is made by an iterative process for which one requires a prior model for the true $\omega(\theta)$. In the literature, the model for the true $\omega(\theta)$ is often assumed to be a simple power-law, i.e., $\omega(\theta) = A_\omega \theta^{\gamma-1}$, where A_ω is the amplitude and the slope γ has the value of ~ 1.8, (see e.g., Roche et al. [175]). However, assuming a power-law for $\omega(\theta)$ could lead us to the introduction of an extra source of error. We therefore decide to use an alternative method to determine C, through which we obtain a self-consistent determination of $\omega(\theta)$ instead of assuming it as a power law.

As mentioned earlier, when using the mass relation between galaxies and (sub)halos obtained via the AMT, the clustering properties of the galaxies, such as the amplitude and the shape of $\omega(\theta)$, are well characterized within the context of the HOD model (e.g.,58, 148). This is true when the subhalo mass is defined at the accretion epoch, $m_{\rm sub} = m_{\rm acc}$. To estimate C we take the advantage of this result. At each redshift bin, we match the abundances between galaxies and halos plus subhalos by assuming that $P_{\rm cen}(M_*|M_{\rm h}) = P_{\rm sat}(M_*|m_{\rm sub})$; under this hypothesis, we compute a first model for $\omega(\theta)$ (see e.g., Chapter 2 and 3, and see also 58, 148). Using equation (52) we obtain a first estimate for C. For the next step, we use the halo model. Assuming that $P_{\rm cen}(M_*|M_{\rm h}) \neq P_{\rm sat}(M_*|m_{\rm sub})$ and using as constraints the overall GSMF and the corrected $\omega(\theta) = \omega_{LS}(\theta) + C$, we obtain a new set of mass relations for centrals and satellites, separately. Then C is recalculated according to Eq. (52) . If the new C is substantially different than the old one, then we repeat the process until some convergence is reached. Observe that the resulting mass relation constrained under this scheme is not what we report along this paper. These mass relations were used only to estimate C accurately. In this case, the mass relations were constrained by finding the model parameters that maximize the likelihood function $\mathcal{L} \propto \exp(-\chi^2/2)$ by using the Powell's directions set method in multi-dimensions, Press et al. [167].

As mentioned above, in the COSMOS survey there is a significant excess of clustering at large scales which might potentially affect our measurements of the galaxy correlation function, particularly at $z \sim 0.9$. de la Torre et al. [67] have investigated how the observed clustering of galaxies depends on the local environment, in particular in the zCOSMOS survey, which is a subsample of the COSMOS field. Based on mock galaxy catalogs that mimic the zCOSMOS survey, de la Torre et al. [67] found that when the top ten precent

tail of the galaxies in the high density environments are excluded, the galaxy clustering reassembles the $\Lambda-$CDM correlation function at large scales. To correct at a first order our measurements of the angular galaxy clustering at $z \sim 0.9$, we have found that the following renormalization for the last mass threshold sample: $\omega_{\rm corr}(\theta|z \sim 0.9, 11 < \log M_*) = \omega(\theta|z \sim 0.9, 11 < \log M_*) - 0.01$, roughly mimics the correction proposed in de la Torre et al. [67].

Figure 21 shows the resulting measured angular correlation functions for the four redshift bins and for all mass threshold samples (skeletal symbols with error bars). The redshift for each mass threshold sample is indicated in each panel. At a fixed redshift, the strength of the clustering increases with mass. At a fixed stellar mass, low redshifts galaxies are more clustered. This is consistent with previous reports of the galaxy clustering between $z = 0.2$ to $z = 1$ (e.g., 61).

6.4 PROCEDURE & UNCERTAINTY TREATMENT

We now describe the procedure for constraining the ten free parameters of the model (§§6.2.2) by using the available observational data (§§6.3) and taking into account different kind of uncertainties. By constraining these parameters, the model allows to predict at each redshift bin the central and satellite galaxy-to-(sub)halo mass relations, the satellite CSMFs, and the decomposition of the GSMFs into centrals and satellites.

6.4.1 *The fitting procedure*

To sample the best fit parameters that maximize the likelihood function $\mathcal{L} \propto \exp(-\chi^2/2)$, we use a Markov Chain Monte Carlo (MCMC) method. Each MCMC chain consist of 5×10^5 elements for each mass threshold sample. We define the function χ^2 as,

$$\chi^2 = \chi^2_{\phi_g} + \sum_{i,j} \chi^2_{\omega^i_j(\theta)}, \tag{53}$$

where for the GSMF, the χ^2's are defined as:

$$\chi^2_{\phi_g} = \sum_{i=1}^{N_{\rm bin}} \left(\frac{{\phi_g}^i_{,\rm model} - {\phi_g}^i_{,\rm obs}}{\sigma^i_{\rm obs}} \right)^2, \tag{54}$$

$N_{\rm bin}$ is the number of bins in the total GSMF with an ith value of ${\phi_g}^i_{,\rm obs}$ and an error of $\sigma^i_{\rm obs}$. The ith value of the total GSMF computed in the model is denoted as ${\phi_g}^i_{,\rm model}$.

For the angular 2PCF, the χ^2 is defined as:

$$\chi^2_{\omega^i_j(\theta)} = \sum_{i=1}^{N_{*,\rm bin}} \sum_{j=1}^{N_{\theta,\rm bin}} \left(\frac{\omega^{i,j}_{\rm model} - \omega^{i,j}_{\rm obs}}{\sigma^{i,j}_{\rm obs}} \right)^2, \tag{55}$$

where $N_{*,\rm bin}$ is the number of stellar mass bins, $N_{\theta,\rm bin}$ denotes the number of bins in the two point angular correlation function, $\omega^{i,j}_{\rm model}(\omega^{i,j}_{\rm obs})$ is the amplitude of the observed (modeled) galaxy clustering at the j-th angular distance bin of the i-th stellar mass threshold bin.

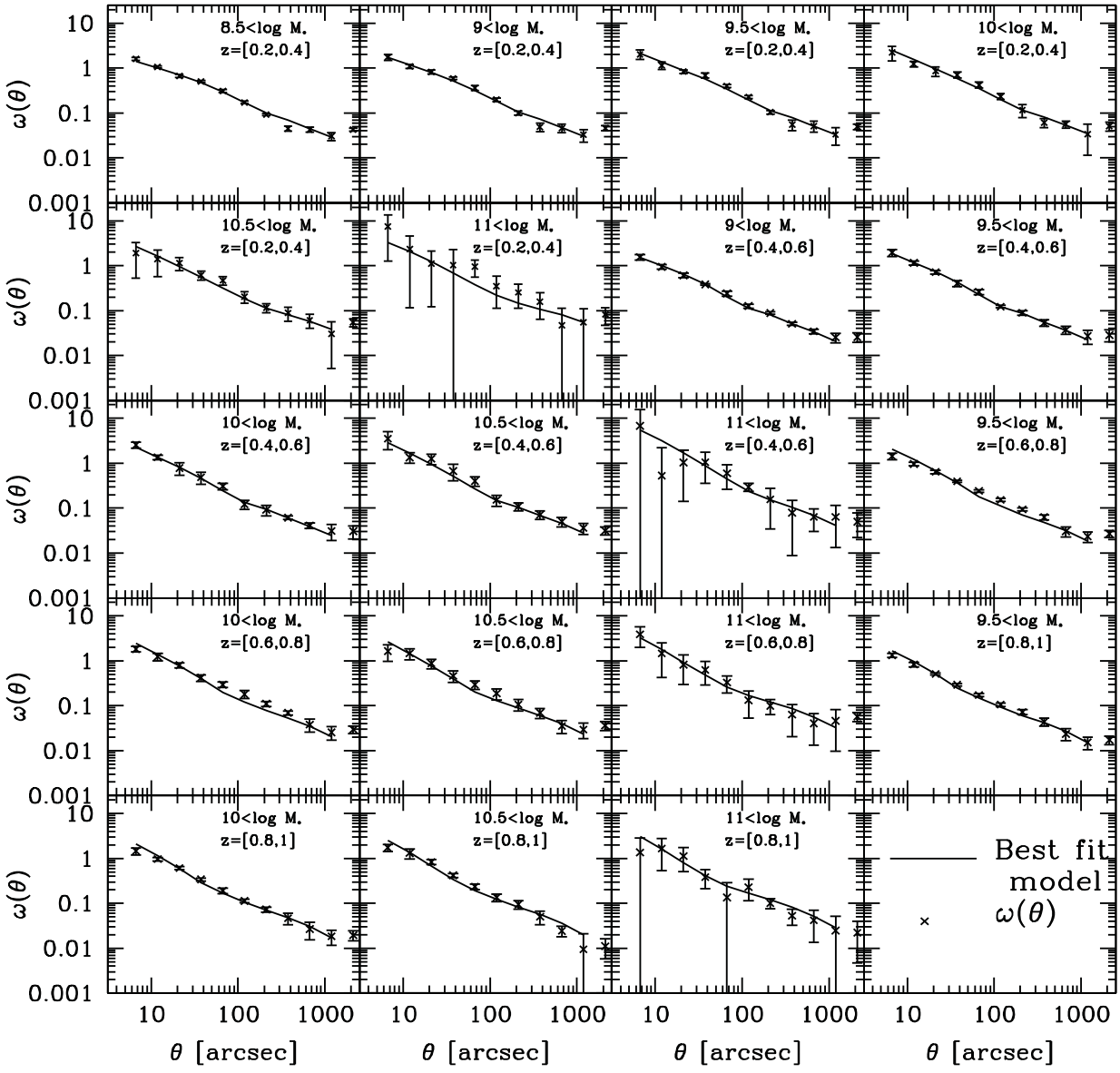

Figure 21: Angular two point angular correlation functions at different redshift bins and stellar mass thresholds. The skeletal symbols with error bars, show our estimation of the galaxy clustering in the COSMOS field, while the solid lines indicate our best fit models. Note that the redshift bins and the stellar mass thresholds are indicated in each panel.

In the case of $\chi^2_{\omega_j(\theta)}$ we use only the diagonal of the covariance matrix. As noted in RAD13 the lack of the full covariance matrix seems to affect the results for the abundance of satellite galaxies. Nevertheless, the error introduced under this scheme is within the errors of the observed satellite GSMF. Additionally, Zheng, Coil & Zehavi [263] found that the resulting marginalized distributions of the HOD parameters when using the diagonal of the covariance matrix were closely similar to those when the full covariance matrix is considered. Therefore, we conclude that the errors introduce in the posterior distributions of the model parameters are of minor importance.

We begin by finding the set of ten parameters, $\mathbf{a} = \{M_{1,c}, \epsilon_c, \alpha_c, \delta_c, \gamma_c, M_{1,s}, \epsilon_s, \alpha_s, \delta_s, \gamma_s\}$, that minimizes χ^2 using the Powell's directions set method in multi-dimensions, Press et al. [167]. This first set of parameters are used as the starting point to sample the parameter space with the MCMC method. We also need to specify for each parameter an initial proposal distribution, which generates the candidate for sampling the parameter space. We assume that each proposal distribution is Gaussian distributed. The standard deviation for each parameter, σ_i, is calculated from the covariance matrix. The covariance matrix or error matrix of $\mathbf{a}$ is defined as the inverse of the $n \times n$ matrix $\alpha = \epsilon^{-1}$, computed according to

$$\alpha_{kl} = \frac{1}{2}\frac{\partial^2 \chi^2(\mathbf{a})}{\partial a_k \partial a_l}. \tag{56}$$

Therefore, the standard deviations in the parameters correspond to the square roots of the terms in its diagonal, i.e., $\sigma(a_i) = \sqrt{\epsilon_{ii}}$. We use the complete covariance matrix as our best initial guess for the covariance matrix of the model parameters. Then the covariance matrix for the proposed matrix generates candidates from a previous step according to: $a_i = a_{i,\mathrm{curr}} + \epsilon^{\mathrm{p}}_{ij} x_j$ where $\epsilon^{\mathrm{p}}_{ij}$ is given by $\epsilon^{\mathrm{p}}_{ij} = \epsilon_{ij}/n^2$, with $n = 10$ the number of parameters to be fitted and x_j is a vector consisting of $j = 1, ..., 10$ standard normal deviates.

Using these results, we sample a first chain with 100,000 models, from which we compute the complete covariance matrix, $\epsilon^{\mathrm{c}}_{ij}$. If the ratio of each prior, $\sqrt{\epsilon_{ii}}$, to each element of the resulting diagonal covariance matrix, $\sqrt{\epsilon^{\mathrm{c}}_{ii}}$, lies in the range $0.8 \leq \sqrt{\epsilon_{ii}/\epsilon^{\mathrm{c}}_{ii}} \leq 1.2$, then we initialize a second chain with 5×10^5 elements for the model analysis; otherwise, we repeat the procedure $j-$times until the ratio of the covariances of the previous chain with the last one reachs the condition $0.8 \leq \sqrt{\epsilon^{j-1}_{ii}/\epsilon^{j}_{ii}} \leq 1.2$, that is, until there is not a sufficiently significant improvement in the standard deviations of the model parameters. The $j-$covariance matrix for the proposed distribution is given by $\epsilon^{\mathrm{p},j}_{il} = \epsilon^{j}_{il}/n^2$. Then, we run a last chain with 5×10^5 elements for the model analysis. This procedure usually takes one or two iterations.

6.4.2 *The uncertainties*

When computing the mass relations, there are several sources of uncertainties that should be take into account. Among the most relevant are those in the observed GSMF, including stellar mass estimates; in the halo mass functions; in the uncertainty of the cosmological parameters; and in linking galaxies to halos, including the intrinsic scatter

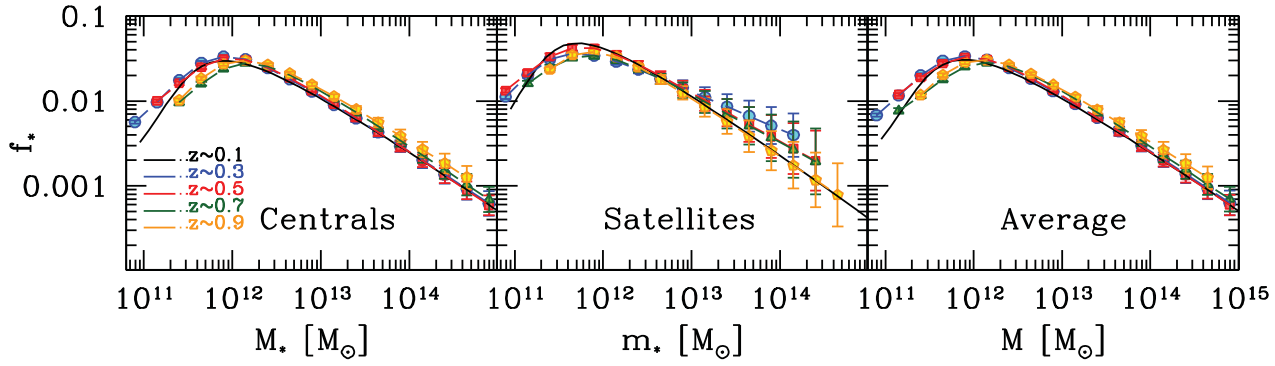

Figure 22: From left to right, the constrained stellar mass fraction of central and satellites, and the number-density average mass relation. The filled blue circles, red filled squares, green filled triangles and the yellow filled pentagons correspond to the redshifts centered at $z \sim 0.3$, 0.5, 0.7 and 0.9. Note the weak redshift evolution of the mass relations, specially for high masses.

in this connection. In Behroozi, Conroy & Wechsler [19] the impact of these sources of uncertainties has been studied in detail. These authors have found that the largest uncertainty by far in the CHMR is due to systematic shifts in the stellar mass estimates, roughly ~ 0.25 dex, which is the value that we adopt. The second important source of uncertainty is due to the intrinsic scatter of the stellar mass at a fixed halo mass. We account for this error in our analysis by introducing the lognormal distributions $P_{\rm cen}(M_*|M_{\rm h})$ and $P_{\rm sat}(M_*|m_{\rm sub})$. The third important source of uncertainty becomes relevant with redshift. This is the error introduced when the stellar masses are estimated with limited photometric information. For this source of uncertainty, galaxies have an intrinsic scatter relative to the true galaxy stellar mass. In our analysis, we account for this source of uncertainty by introducing a Gaussian distribution, $G(M_*|m_*)$, in such a way the observed GSMF, $\phi_g(M_*)$, and the true GSMF, $\phi_{g,\rm true}(M_*)$, are related via the following convolution,

$$\phi_g(M_*, z) = \int_0^\infty G(M_*|m_*, z)\phi_{g,\rm true}(m_*, z)dm_*. \tag{57}$$

In sum, the impact of this source of uncertainty is that the observed (convolved) GSMF is shallower than the true GSMF at the high mass end. According to Behroozi, Wechsler & Conroy [20], $G(M_*|m_*)$ has a mean of zero and its scatter depends on redshift according to $\sigma(z) = \sigma_0 + \sigma_z z$, with $\sigma_0 = 0.07$ and $\sigma_z = 0.04$. This is consistent with previous results in the literature [e.g., 55, 112, 163] Finally, we assume that other statistical and sample variance errors have negligible effects, at least for redshifts $z < 1$.

6.5 RESULTS

6.5.1 *The stellar-to-(sub)halo mass relationships*

We now present the resulting COSMOS mass relations after sample the posterior distributions of the model parameters by using the MCMC method described above.

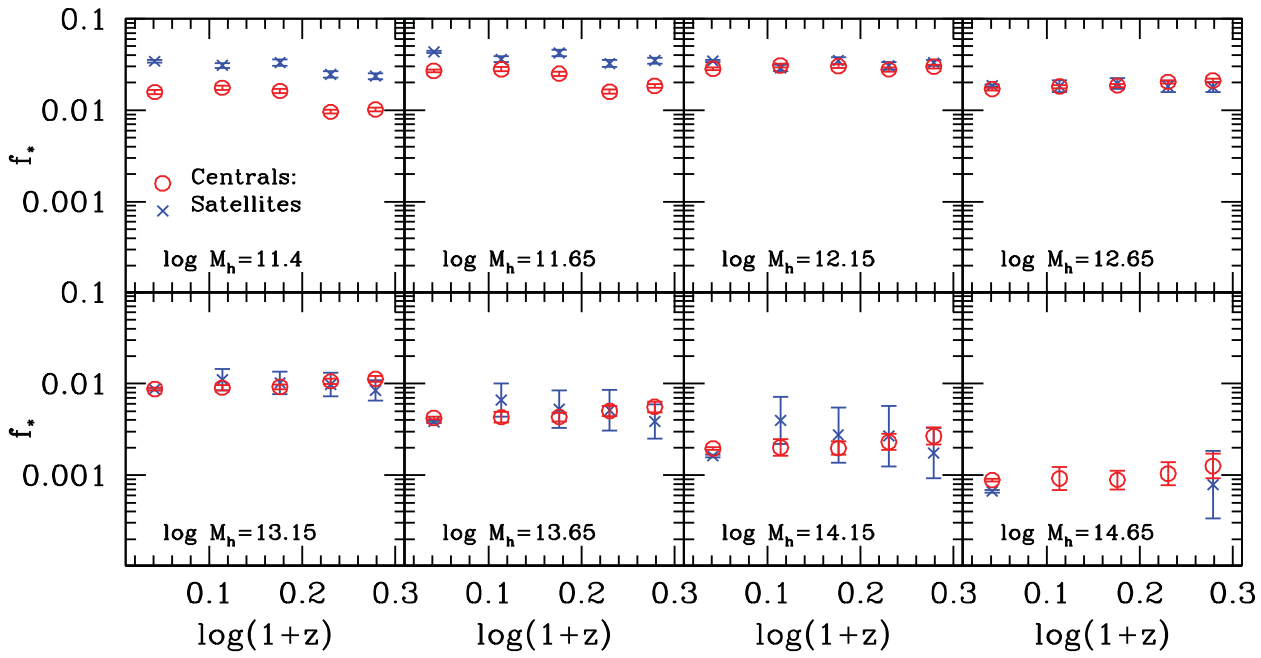

Figure 23: Redshift evolution of the stellar-to-halo mass ratio, f_*, at a fixed (sub)halo mass, $f_*(z|M_h)$, in various (sub)halo masses. The values of central galaxies are denoted by the red circles with error bars, while satellite galaxies are plotted with the skeletal symbols with error bars.

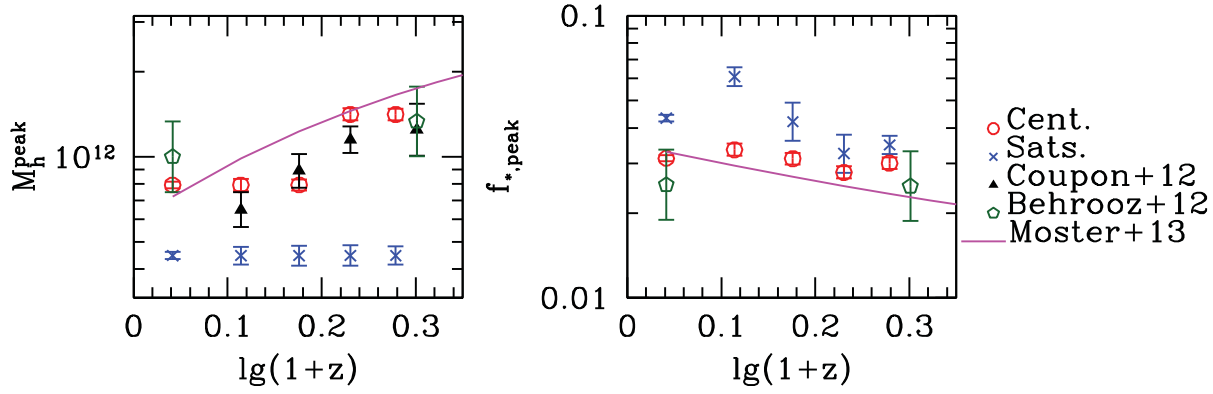

Figure 24: *Left Panel:* Redshift evolution of the (sub)halo mass at which the peak of the stellar-to-(sub)halo mass ratio reaches its peak value, M_h^{peak}. Model predictions for centrals (open red circles with error bars) and satellites (blue skeletal symbols with error bars), compared to the galaxy clustering analysis inferences of Coupon et al. [61] (black filled triangles with error bars) and the AMT results from Behroozi, Wechsler & Conroy [20] (open green pentagons) and Moster, Naab & White [147] (magenta solid line) *Right Panel:* Redshift evolution of the maximum value of the stellar-to-(sub)halo mass ratio.

Figure 22 shows, the central, satellite and average stellar-to-(sub)halo mass fractions, $f_* = M_*/M$, where $M = M_{\rm h}$ or $= m_{\rm acc}$, vs M (left, middle and right panels, respectively) corresponding to each one of the four redshift bins analyzed here. In general the $f_* - M$ relations change little with redshift. In the case of central galaxies, f_* slightly decreases with z for masses below the peak, and it remains almost the same for larger masses. In the case of satellite galaxies, the most notable difference in the evolution is at the largest masses: f_* decreases at higher redshifts.

To study in more detail the possible changes of the $f_* - M$ relations, in Fig. 23 we plot the redshift evolution of f_* at a fixed (sub)halo mass, $f_*(z|M)$, in eight different host halo/subhalo masses. We begin by examining $f_*(z|M_{\rm h})$ for central galaxies (red circles with error bars). The value of $f_*(z|M_{\rm h})$ for halos smaller than $M_{\rm h} \sim 10^{12}$ M$_\odot$, increases slightly for low z, i.e., for the low-mass halos, the efficiency of stellar mass growth increases towards lower redshifts. For $M_{\rm h} = 10^{11.4}$ M$_\odot$, if one extrapolates the trend down to $z \sim 0$, then f_* changes from $z \sim 0$ to ~ 1 by a factor of 2, while for $M_{\rm h} = 10^{12.15}$ M$_\odot$, there is already no change. For larger masses, f_* starts to decrease very slowly as z is lower, i.e., galaxies in very massive halos were slightly more efficient in growing their stellar masses at higher redshifts.

In the case of satellite galaxies, analogously to centrals, $f_*(z|m_{\rm acc})$ changes little with z (blue crosses with error bars in Fig. 23. For subhalo masses at the accretion below $m_{\rm acc} \sim 10^{13}$ M$_\odot$, f_* does not change systematically with z, while for larger masses a slight systematical increase of f_* at low z is observed. In other words, massive halos at the moment that they were accreted into a larger halo they formed stars more efficiently than those that are accreted in later epochs.

From Fig. 22, we see that the halo/subhalo mass corresponding to the peak values of f_*, shifts to lower values with time for central galaxies, while it remains roughly constant for satellite galaxies. In Fig. 24, we plot the values of these peak masses and stellar-to-(sub)halo mass ratios at the different redshifts analyzed here. For central galaxies, the peak halo mass is $M_{\rm h}^{\rm peak} \approx 8 \times 10^{11}M_\odot$ at $z \sim 0.3$ and $M_{\rm h}^{\rm peak} \sim 1.5 \times 10^{12}M_\odot$ at $z \sim 0.9$ (solid red circles with error bars in the left panel of Fig. 24). The difference at these two redshifts is a factor of ~ 2 and suggests a downsizing effect in the efficiency of the galaxy stellar mass growth. Instead, for the satellite galaxies, the peak subhalo mass, $m_{\rm acc}^{\rm peak}$ (blue crosses with error bars), remains almost the same at all redshifts, with a value of $m_{\rm acc}^{\rm peak} \sim 5 \times 10^{11}M_\odot$, i.e., it is lower at all times than the peak halo mass of central galaxies. This result is related to the definition of the (sub)halo mass at the epoch of its accretion, rather than to a question of efficiency: $m_{\rm acc}$ traces on average the masses of halos as they were in the past (at their accretion time) and since the halo mass is smaller on average in the past, then $m_{\rm acc} < M_{\rm h}$ in general. Therefore, the overall satellite stellar-to-subhalo mass relations, in particular their peaks, are shifted on average to lower masses with respect to the central mass relations.

On the right panel of Fig. 24 we plot the corresponding $f_*^{\rm peak}$ values as a function of redshift both for centrals and satellite galaxies. The value of $f_*^{\rm peak}$ for centrals does not change systematically since $z \sim 1$; in the case of satellites, $f_*^{\rm peak}$ slightly increases at low z. The values of centrals are slightly lower than those of satellites, but not because the latter are more efficient in their stellar mass growth than the former, but because the

halo masses at the accretion time were on average smaller than those of the halos at the current epoch. This effect increases at low z.

In the same Fig. 24, we compare our result with recent measurements of $M_{\rm h}^{\rm peak}$ reported in the literature. Coupon et al. [61] (black triangles), based on the galaxy clustering analysis of luminosity threshold samples from the CFHTL, infer the redshift evolution of $M_{\rm h}^{\rm peak}$ at the same redshifts bins than our measurements. The inferences based on the AMT by Behroozi, Wechsler & Conroy [20] and Moster, Naab & White [147] are also plotted: open pentagons for the two redshifts reported in the former paper, and the magenta solid line for the fit reported in the latter paper. Our results are qualitatively consistent with these previous inferences of $M_{\rm h}^{\rm peak}$.

For completeness, in the right panel of Fig. 22 we plot the average (global) stellar-to-(sub)halo mass fractions as a function of mass. Conceptually, this relation is the one that is commonly obtained in the standard AMT (see Eq. 17 in RAD13). Recall that in that case the central and satellite relations can not be differentiated. The average $f_* - M_{\rm h}$ relation is a density-weighted average relation between the respective central and satellite relations. We see that these relations at different z are similar to those of the central galaxies; this is because central galaxies are more common in the observational sample than satellites.

Since we find strikingly little evolution of the mass relations for centrals, as well as for satellite galaxies, the average mass relation almost also do not evolve, as seen in Fig. 22. Recently Behroozi, Wechsler & Conroy [20] arrived to a similar result. Presumably, this is a consequence of the little redshift evolution of the global GSMF, and possibly, of the central GSMF. In order to test this hypothesis, we compare the model predictions for the central and satellite GSMFs with those inferred assuming a non-evolutionary model for the CHMR and the SSMR.

6.5.2 *The GSMF of central & satellite galaxies*

The resulting predictions for the central (orange filled area) and satellite (red filled area) GSMFs from $z = 0.2$ to $z = 1$ are plotted in Fig. 20 above. The filled areas represent the confidence at the 1-σ level. At higher redshifts the average (total) GSMF is more dominated by the central GSMF component than at lower redshifts. With cosmic time, the abundance of satellite galaxies increases, specially from intermediate to low masses. The high-mass end remains almost the same at all redshifts. Note that the low-mass end slope of the GSMFs of centrals and satellites are similar to those of the average (global) GSMFs at all redshifts; this means that the number ratio of satellite to central galaxies is nearly constant at these masses.

The long-dashed black solid lines in each panel indicate the resulting central, satellite and total GSMFs from a model, where it is assumed that the CHMR and SSMR do not change with redshift (non-evolutionary model). The constant CHMR and SSMR for this model are fixed to be those constrained locally in RAD13 (Chapter 3), where the data are richer and more accurate than at higher redshifts. These mass relations were accurately constrained by using the observational information of the GSMF decomposed

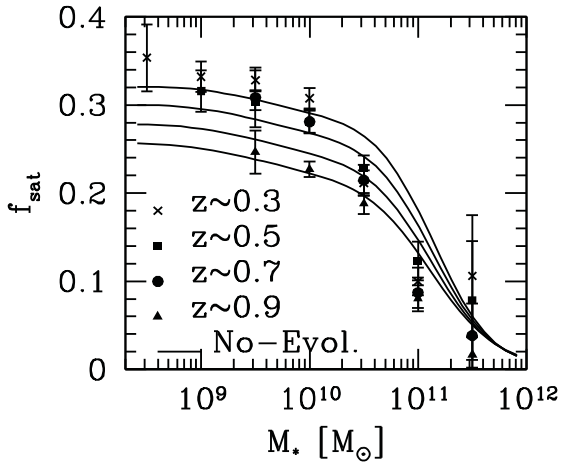

Figure 25: Satellite fraction as a function of the stellar mass for the redshift bins centered at $z = 0.3, 0.5, 0.7$, and 0.9 (crosses, filled circles, triangles and squares, respectively). Solid lines indicates the fraction obtained from the non-evolutionary model, see the text for details.

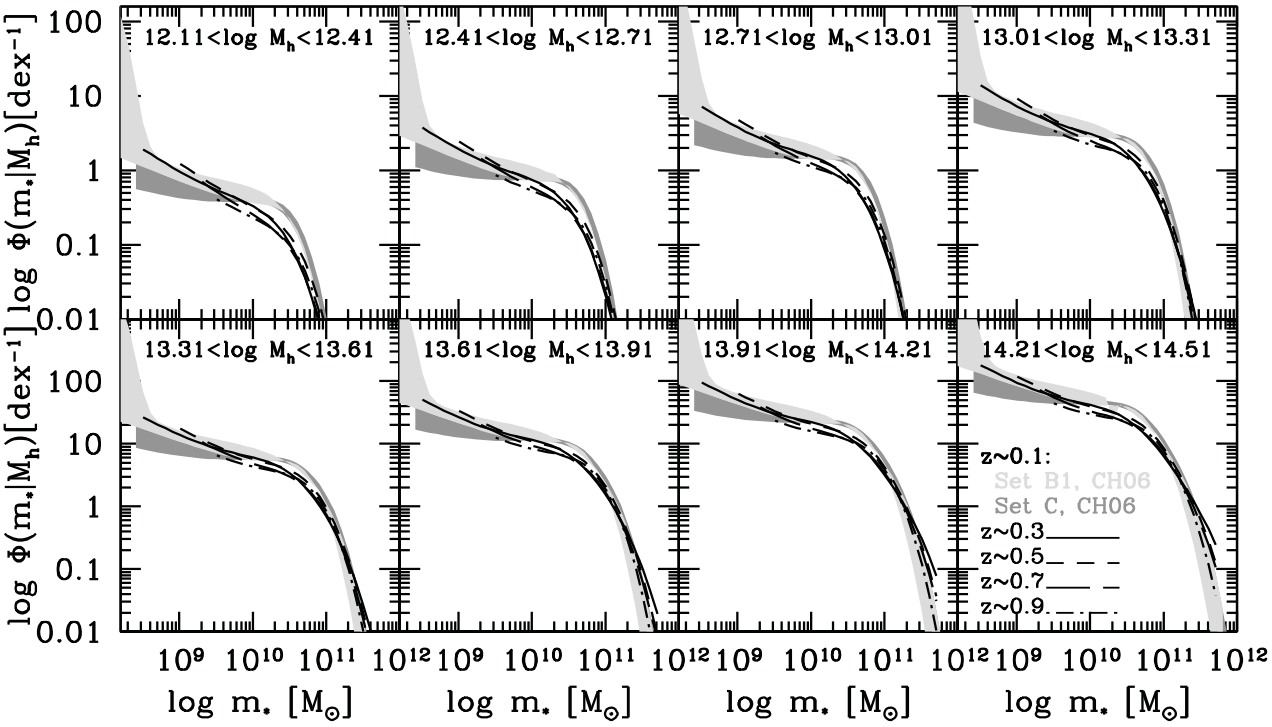

Figure 26: Satellite CSMFs in eight halo mass bins for the redshift bins centered at $z = 0.3, 0.5, 0.7$, and 0.9 (solid, short-dashed, long-dashed and dot-short dashed lines, respectively). For comparison we present the Satellite CSMFs constrained in Chapter 3 using set B and set C, light gray and dark gray areas.

into centrals and satellites and the galaxy correlation function decomposed into several mass bins (set C).

We observe that the non-evolutionary model roughly reproduces the abundances of centrals and satellites up to $z = 1$. This means that the little redshift evolution of the observed GSMF seen up to $z = 1$ is a direct consequence of the (almost) lack of evolution in the mass relations or, i.e., in the efficiency of stellar mass growth of galaxies as a function of halo mass. Nevertheless, this model cannot reproduce the (weak) redshift evolution of the GSMF at its high mass end, mainly the one corresponding to satellites.

6.5.3 *The satellite fraction*

Figure 25 shows the fraction of those galaxies that are satellites, $f_{\rm sat}$, as a function of stellar mass for the redshift bins centered at $z = 0.3, 0.5, 0.7$, and 0.9 (crosses, filled circles, triangles and squares, respectively). For comparison, we show the resulting fractions obtained from the non-evolutionary model (solid black lines). In general, at all redshifts $f_{\rm sat}$ slowly decreases with M_* up to $M_* \sim 10^{10}$ $M_\odot$, and for larger masses, it decreases strongly. This fraction for most of the masses increases at low z. At $M_* \sim 10^{10} M_\odot$, $f_{\rm sat}$ increased by factor of ~ 1.5 from $z = 0.9$ to $z = 0.3$. The non-evolutionary model is qualitatively similar to our results but it shows slightly lower (higher) values of $f_{\rm sat}$ at smaller (larger) masses at all redshifts.

Finally, since the SSMR determines the halo occupation distribution of satellite galaxies at a fixed halo mass, the slight change of SSMR with redshift implies that the satellite population follows a redshift independent occupational distribution at a fixed halo mass [e.g., 61]. Therefore the satellite CSMF as a function of $M_{\rm h}$ should be also approximately independent of redshift. This is indeed the case for our model results. Figure 26 shows the predicted satellite CSMFs in eight halo mass bins for each one of the four redshift bins. Recall that the HOD of subhalos as a function of $M_{\rm h}$ used here is independent of redshift according to analysis of $N-$body cosmological simulations [91, 69, 115, 210, 20].

6.6 DISCUSSION

6.6.1 *Comparison with previous works*

The aim of this Chapter was to constrain the redshift evolution of the mass relations for both centrals and satellites, CHMR and SSMR, respectively, in the COSMOS survey, and to estimate the decomposition of the GSMFs into centrals and satellites at different z's . There are only a few previous attempts to infer the SSMR separately for local galaxies [178, 176, 152], and only one at higher redshifts [231]. Our determinations are the most complete (4 redshift bins) up to $z \sim 1$. In the following, in order to make a direct comparison with previous works, we have divided the discussion between those that infer the mass relations of central galaxies only and those that infer the average mass relation.

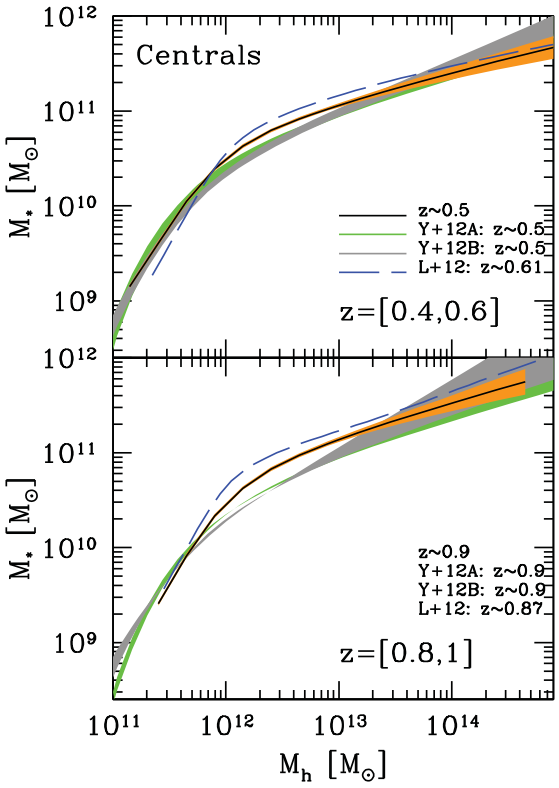

Figure 27: Stellar-to-halo mass relations for central galaxies at the redshift bins $z = [0.4, 0.6]$ and [0.8,1]. Model results, solid lines bracketed by the orange shaded areas, compared to the galaxy clustering+weak lensing analysis from Leauthaud et al. [122] and the results based on CSMF formalism from Yang et al. [250].

Figure 27 presents a comparison of the CHMRs at the redshift bins centered in $z = 0.5$ and $z = 0.9$ (solid lines bracketed by orange dashed regions) with those reported in Yang et al. [250] and Leauthaud et al. [122].

The observational constraints used by Yang et al. [250] were the total GSMFs up to $z \sim 4$, the conditional stellar mass function at $z \sim 0$ in various halo mass bins, and the galaxy clustering for $z \sim 0$ galaxies. They experimented with two different sets for the GSMF up to $z \sim 4$: one reported in Pérez-González et al. [163] and based on the Spitzer data (Y+12A), and another one reported in Drory et al. [76] and based on the deep multicolor data in the FORS deep field and the Great Observatories Origins Deep Survey-south/Chandra deep field-south region (Y+12B). Their corresponding CHMRs are plotted in Fig. 27 as green and grey shaded areas, respectively (they represent the 1σ confidence level of their inferences). In the case of Leauthaud et al. [122], they combined staked galaxy-galaxy weak lensing data and galaxy clustering at various mass thresholds from the COSMOS data. The resulting mean halo masses as a function of stellar mass are also plotted (dashed blue lines). Note that for the latter the comparison is not direct since Leauthaud et al. [122] inferred the mean $M_{\rm h}$ as a function of M_*; we actually plot the inverse of their reported relation. Note also that we have applied small corrections to $M_{\rm h}$ in the Yang et al. [250] results in order to convert their halo mass definition to our virial mass, as well as in M_* in order to be consistent with the Chabrier [48] initial mass function adopted here.

In general, our results are roughly consistent with both the Yang et al. [250] and Leauthaud et al. [122] results, in particular with the latter. This is not surprising since Leauthaud et al. [122] actually used similar data to those used here. For the $z \sim 0.9$ bin, our results and even more those of Leauthaud et al. [122], show stellar masses larger than those of Yang et al. [250] for halos in the mass range $10^{12} - 10^{13}$ $M_\odot$. This discrepancy with the COSMOS data has been noted before (see 250) and it is probably due to the overdensity in the COSMOS field at $z \sim 0.9$. It seems that our attempt to correct for the overdensity, both in the GSMF and the angular 2PCF, ameliorated the problem as compared with the Leauthaud et al. [122] results.

In Figure 28 we present a comparison of the average stellar-to-halo mass relation at the redshift bins centered in $z = 0.5$ and $z = 0.9$ (solid lines surrounded by orange dashed regions) with previous works, namely Behroozi, Conroy & Wechsler [19], Behroozi, Wechsler & Conroy [20] and Moster, Naab & White [147]. Note that these authors used methods closely related to the abundance matching technique.

Behroozi, Conroy & Wechsler [19] (violet long dashed line) used as observational input the total Pérez-González et al. [163] GSMF in a wide range of redshifts, and considered several sources of uncertainty. Behroozi, Wechsler & Conroy [20] used the combined data sets of Baldry, Glazebrook & Driver [16], Moustakas et al. [149] and Pérez-González et al. [163]. We present their results at $z \sim 0.9$ with the red long dashed line. Moster, Naab & White [147] used the GSMFs by Pérez-González et al. [163] and Santini et al. [182] GSMFs (green solid line).

Our inferred average mass relations are consistent with previous AMT results, specially at $z \sim 0.5$. At redshift ~ 0.9 we observe a discrepancy between our results with those obtained in Behroozi, Conroy & Wechsler [19], the reason is not clear. However, the comparison between our results with those from Behroozi, Wechsler & Conroy [20]

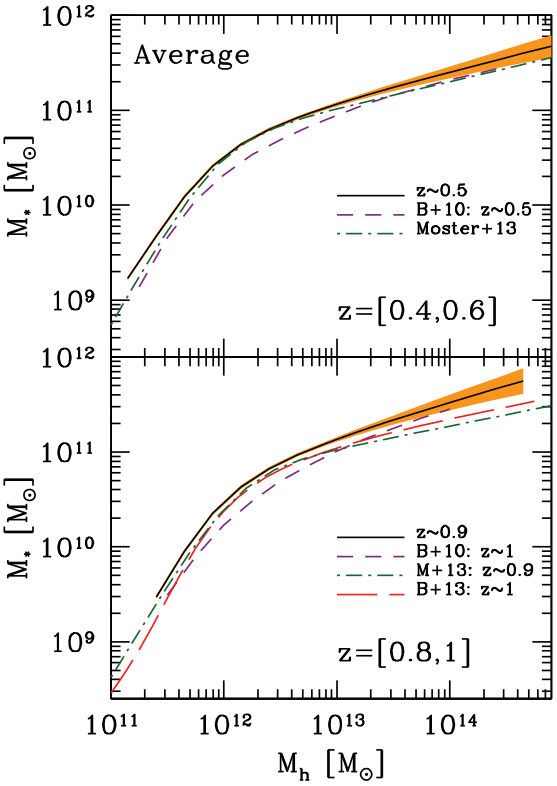

Figure 28: Similar to Fig. 27 but for the average stellar-to-halo mass relation. Comparison with the results based on the AMT from Behroozi, Conroy & Wechsler [19], Behroozi, Wechsler & Conroy [20] and Moustakas et al. [149] are shown with the short-dashed, long-dashed and short dotted-lines, respectively.

are consistent within the error bars. Note that Behroozi, Wechsler & Conroy [20] used similar data than Behroozi, Conroy & Wechsler [19] for redshifts $z < 1$, but the former have include also the Moustakas et al. [149] GSMF as a constraint.

6.6.2 *On the lack of evolution of the stellar-to-(sub)halo mass relations*

The main result of this Chapter is the slight change with redshift obtained for both the CHMR and SSMR (Figs. 22 and 23), which implies that the efficiency of stellar mass growth of both central and satellite galaxies for a given host halo mass is roughly the same since $z \sim 1$. However, for central galaxies, we found a shift of the halo mass at the peak efficiency towards lower masses at low z (downsizing; Fig. 24). These findings are probably mainly due to the slight evolution of the total GSMF since $z \sim 1$. To test this hypothesis, we have applied a model in which the CHMR and the SSMR do not evolve with redshift and obtained a GSMF change with z consistent with the one observed in the COSMOS GSMF, althought at the high-mass end the prediction evolves less than the observation. The latter is associated to the (slight) change that the peak efficiency scale seems to suffer (downsizing).

The almost lack of evolution of the CHMR since $z \sim 1$ offers a first order description of the stellar mass growth of central galaxies. For the case of a no evolving CHMR, the stellar mass growth of a central galaxy is determine by the shape of this relation and the growth rate of its host halo, that is, $M_*(z) \propto M_{\rm h}^{\beta(M_{\rm h};z)}(z)$, where $\beta(M_{\rm h};z)$ is the slope of the CHMR around the mass in question and at the epoch z; for a non-evolving CHMR, β at a given $M_{\rm h}$ is the same at all redshifts. The local CHMR (see Chapter 3) is such that β acquires roughly the following values:

$$\begin{aligned} \beta &\approx 2.5 \quad \text{for} \quad M_{\rm h} < 10^{12}{\rm M}_\odot, \\ \beta &\approx 1 \quad \text{for} \quad M_{\rm h} \sim 10^{12}{\rm M}_\odot, \\ \beta &\approx 0.3 \quad \text{for} \quad M_{\rm h} > 10^{12}{\rm M}_\odot. \end{aligned} \tag{58}$$

Using the above equations, we can infer a very crude model for the stellar mass growth rate of galaxies, $\dot{M}_*$. By differentiating in time $M_*(z) \propto M_{\rm h}^{\beta}(z)$ (omiting by the moment the dependence of β on $M_{\rm h}$), one obtains $\dot{M}_* \propto \beta \times M_{\rm h}^{\beta} \times \dot{M}_{\rm h}$. From N-body cosmological simulations, Fakhouri & Ma (2010) obtained that $\dot{M}_{\rm h} \propto f(z) \times M_{\rm h}^{1.1}$, where $f(z)$ is a function that strongly increases with z. Hence, $\dot{M}_* \propto \beta \times M_{\rm h}^{\beta+0.1} \times f(z) \propto \beta \times M_*^{1+0.1/\beta} \times f(z)$. The specific stellar mass growth rate is then: $(\dot{M}_*/M_*) \propto \beta \times M_*^{0.1/\beta} \times f(z)$. Therefore, for a given z, the normalization of $(\dot{M}_*/M_*)$, given by β, decreases almost and order of magnitude from low to high masses (see eq. 58), i.e., low-mass galaxies are more actively growing than massive galaxies. For any mass, $(\dot{M}_*/M_*)$ is higher at higher z according to the dependence on $f(z)$. For galaxies living in halos much smaller than the peak mass, $M_{\rm h} \sim 10^{12}$ M$_\odot$, $\beta = 2.5$, and therefore, $(\dot{M}_*/M_*) \propto \beta \times M_*^{0.04} \times f(z)$, i.e., their specific M_* growth rate is almost independent on mass. For low-mass galaxies the specific M_* growth rate is practically equal to the in situ specific star formation rate (SSFR). For large masses, along with the SSFR, dry mergers (low gas content) can significantly contribute to the M_* growth of galaxies.

The non-evolution model makes predictions that at a qualitatively level agree with the empirical picture of stellar mass growth of galaxies as a function of their masses (see e.g., Firmani & Avila-Reese [84], Behroozi, Wechsler & Conroy [20]). However, in more detail, we remark a possible shortcoming. At small masses, a large body of observational inferences show that the SSFR increases on average for small M_* for local and higher redshift galaxy populations (downsizing in SSFR; see for a review 87, 9, 17). The non-evolution model predicts a constant SSFR; in order for the SSFR to increase for low masses, the f_*–$M_{\rm h}$ relation at low masses should decreases and steepen with z. Our constrained CHMR shows this behavior, though very incipiently. Thus, a moderate evolution of the low-mass side of the CHMR seems to be necessary in order to be in agreement with observations of the SSFR of galaxies.

At large masses, the observational pieces of evidence suggest that galaxies are passive since long time ago, i.e. they are almost not forming stars anymore since high redshifts (e.g., Pérez-González et al. [163]). This is roughly our result obtained for large masses for a non-evolving CHMR, but there is yet room for a small growth by dry mergers. It is still a matter of debate how much massive passive galaxies grow by dry mergers since $z \sim 1$. In the case of our constrained CHMRs, they slightly increase with z at the high-mass end, which implies even less stellar mass growth than the non-evolving case.

Thus, our results disfavor the role of dry mergers in the late mass growth of massive passive galaxies.

Whereas at a first order the redshift evolution of CHMR seems to be constant, in detail we observe that the halo mass where the peak of f_*, $M_h^{\rm peak}$, is reached increases with redshift. This means that high mass halos were more efficient in growing stars during the past. Nevertheless, we observe that this efficiency is roughly the same since $z \sim 1$ with a critical value of $f_* \sim 0.03$. For satellite galaxies, the position of the peak of f_* stays also roughly constant but higher compared to centrals, $f_{\rm peak} \sim 0.04$ vs. ~ 0.03. Similarly, when one looks at the position of $M_{\rm acc}^{\rm peak}$, it is almost constant, while the $M_h^{\rm peak}$ for centrals decreases from $z \sim 0.9$ to ~ 0.3 by a factor of 2, being all the time larger than $M_{\rm acc}^{\rm peak}$ (see Fig. 24). These differences can be explained mostly by the fact that since $m_{\rm acc}$ is the mass of a halo accreted in the past, it is systematically lower than the current halo masses of central galaxies. It is also possible that the difference between the CHMR and the SSMR gives quantitative information about how much, on average, satellite galaxies have grown after their accretion.

Finally, if the CHMR and the SSMR are independent on redshift and roughly similar, i.e., CHMR $\sim$ SSMR, then the distribution functions $P_{\rm cen}(M_*|M_{\rm h}) \sim P_{\rm sat}(M_*|m_{\rm sub})$, which is one of the main hypothesis of the abundance matching technique. Is possible that this the main reason for the success of the AMT [see also, 231],

6.7 CONCLUSIONS

By means of a statistical model that combines the AMT, the HOD model, and the CSMF formalism, we have inferred the stellar-to-(sub)halo mass relations for central and satellite galaxies at different redshifts up to $z \sim 1$ for the COSMOS data. This deep survey offers the possibility to calculate the necessary information for constraining the parameters of our model, namely the GSMF and the angular 2PCFs at different redshift bins, from $z = 0.3$ to $z = 0.9$ in our case. We have corrected the GSMF and angular 2PCFs at the redshifts were the COSMOS survey is apparently affected by the large-scale overdensities and underdensities within its relatively small area. Our main results and conclusions are as follow.

- The CHMR and SSMR change slight with redshift. The most remarkable changes are at the low-mass end of the former which slightly decreases with z, and at the high-mass end of the latter, which slightly increases with z. The halo mass corresponding to the peak of the stellar-to-halo mass ratio f_* of central galaxies, $M_h^{\rm peak}$, decreases a factor of 2 from $z = 0.9$ to $z = 0.3$ (downsizing in mass), while the corresponding peak mass for the satellite galaxies, $m_{\rm acc}^{\rm peak}$, remains the same and it is ever smaller than $M_h^{\rm peak}$. The values of f_* are roughly 0.03 and 0.04 for central and satellite galaxies, respectively, at all redshifts. These small differences between the central and satellite mass relations can be explained mainly by the fact that the halo masses at the past accretion time (used for assigning subhalo masses to the satellites) are systematically smaller than the current halos masses of centrals because the halo mass increases with time on average.

- Our model allows to decompose the GSMF at each z into centrals and satellites. The fraction of satellite galaxies at a given stellar mass, $f_{\rm sat}$, slightly decreases with M_* up to $\sim 10^{10}{\rm M}_\odot$, and for larger masses strongly decreases. This behavior is the same at the four redshift bins; however, the values of $f_{\rm sat}$ increase with time for almost all masses. For instance, for $M_* \sim 10^{10}{\rm M}_\odot$, $f_{\rm sat} \approx 0.32$ at $z = 0.3$ and it increased by a factor of ~ 1.5 since $z = 0.9$. The satellite CSMFs as a function of host halo mass remain roughly the same at different redshifts.

In spite of the cosmic variance issues of the COSMOS sample, we were able to obtain robust constrains on the CHMR and SSMR of central and satellite galaxies at four redshift bins since $z \sim 1$, as well as on the fractions of satellite galaxies as a function of M_* at each redshift. We find that the change of the mass relations with z, in particular the one corresponding to central galaxies, is small. Therefore, a model that assumes no evolution of the CHMR describes reasonable well our main results regarding the GSMFs of central/satellites, the satellite fractions, and the satellite CSMFs at different epochs. However, we have seen that a complete lack of evolution of the CHMR implies no downsizing in SSFR, i.e., a constant SSFR with M_* at low masses, in potential conflict with direct determinations of the SSFR vs M_* of local and high-redshift galaxies. Since our constrained stellar-to-halo mass relation for centrals actually slightly decreases and steepens with z at low masses, some downsizing in SSFR is expected. For large masses, our CHMR slightly decreases with z, which implies almost no growth of M_* for massive galaxies, and hence a minimal possibility for dry mergers.

ACKNOWLEDGMENTS

We are grateful to X. Yang for providing us in electronic form their data for the CHMR. We also thank to J. Coupon for sending us in electronic form his measurements on $M_{\rm h}^{peak}$

Part V

THE GALAXY-TO-HALO CONNECTION OF BLUE AND RED GALAXIES AT $z \sim 0$

En esta Parte V nos enfocamos en estudiar la conexión local galaxia central-halo separada en galaxias rojas y azules con el objetivo de dilucidar si la eficiencia de crecimiento de masa estelar de las galaxias ($f_* = M_*/M_{\rm h}$), aparte de la masa de halo, $M_{\rm h}$, depende del color de la galaxia. En el Capitulo 7 se utiliza la técnica del empate de las abundancias para obtener las relaciones masa estelar galaxia central-masa halo, M_*-$M_{\rm h}$, y masa bariónica galaxia central-masa halo, M_b-$M_{\rm h}$, para galaxias centrales rojas y azules, para lo cual se hacen suposiciones sobre los halos que albergan galaxias rojas y azules basadas principalmente en criterios de fusiones mayores. Se encuentra que las relaciones M_*-$M_{\rm h}$ para galaxias centrales rojas y azules difieren poco, siendo la eficiencia f_* ligeramente mayor para galaxias rojas para masas mayores a $M_{\rm h} \sim 3 \times 10^{11}$ M$_\odot$. En el caso de las relaciones bariónicas, M_b-$M_{\rm h}$, las diferencias entre rojas y azules son aún menores, a excepción de las masas bajas, $M_{\rm h} \lesssim 3 \times 10^{11}$ M$_\odot$, donde la relación para azules se hace más plana. En el Capitulo 8 se presenta un modelo estadístico más completo que el simple empate de abundancias, mismo que relaciona de manera auto-consistente la función de masa de galaxias centrales y satétiles separadas en rojas y azules, las funciones condicionales de masa estelar de satélites y la función de correlación de dos puntos separadas en rojas y azules, así como la función de masa de (sub)halos de materia oscura fría. Además en este modelo, la fracción de halos que albergan galaxias rojas y azules se constriñe usando las distribuciones color-M_* observadas. Como resultado de la auto-consistencia que se logra entre todas las observaciones mencionadas y las funciones de masa de halos/subhalos, se obtiene la relación M_*-$M_{\rm h}$ por separado en rojas y azules, así como también la función de masa de los halos que albergan galaxias rojas y azules y las funciones de masa condicional de satélites rojos y azules en función de $M_{\rm h}$. Se muestra que las relaciones M_*-$M_{\rm h}$ de azules y rojas no difieren significativamente entre sí, y por lo tanto no se segregan sistemáticamente de la relación M_*-$M_{\rm h}$ total de galaxias centrales. Las suposiciones hechas en el Capítulo 8 sobre los halos que albergan galaxias rojas y azules, basadas en criterios de fusiones mayores, parecen ser consistentes con los resultados obtenidos del modelo, incluyendo la función de correlación observada de galaxias rojas y azules. No obstante, algunos refinamientos a esas suposiciones son necesarios. Se concluye que la eficiencia de crecimiento de M_* de las galaxias centrales es fuertemente controlada por la masa del halo (potencial gravitacional) debido a los procesos astrofísicos de eyección de gas por retro-alimentación de SNs y AGNs y de enfriamiento radiativo del gas, mismos que son principalmente dependientes de la escala; posibles dependencias de la historia de agregación de masa de los halos o su concentración que podrían segregar a las galaxias por color en la relación M_*-$M_{\rm h}$ son borradas.

El Capítulo 7 corresponde al artículo publicado: *"On the stellar and baryonic mass fractions of central blue and red galaxies", Rodríguez-Puebla, Avila-Reese, Colín & Firmani 2011, RevMexAA, 47, 235*. Algunas cuestiones que se discuten en el Capítulo 8 han sido abordadas en el artículo remitido *"Central galaxies in different local environments at z ∼ 0: do they have similar properties?", Lacerna, Rodríguez-Puebla, Avila-Reese, Hernández-Toledo 2013*

7

ON THE STELLAR AND BARYONIC MASS FRACTIONS OF CENTRAL BLUE AND RED GALAXIES

This Chapter was published as: Rodríguez-Puebla A; Avila-Reese V.; Colín P.; Firmani C., 2011, RevMexAA, 47, 235.

ABSTRACT

Using the abundance matching technique, we infer the local stellar and baryonic mass–halo mass (M_s-M_h and M_b-M_h) relations separately for *central* blue and red galaxies (BGs and RGs). The observational inputs are the SDSS central BG and RG Stellar Mass Functions and the measured gas mass-M_s relations. For halos associated to central BGs, the distinct ΛCDM Halo Mass Function is used and set up to exclude: (i) the observed group/cluster mass function and (ii) halos with a central major merger at resdshifts $z \leq 0.8$. For central RGs, the complement of this mass function to the total one is used. At $M_h > 10^{11.5}M_\odot$, the M_s of RGs tend to be higher than those of BGs for a given M_h, the difference not being larger than 1.7. At $M_h < 10^{11.5}M_\odot$, this trend is inverted. For BGs (RGs): (a) the maximum value of $f_s = M_s/M_h$ is $0.021^{+0.016}_{-0.009}$ ($0.034^{+0.026}_{-0.015}$) and it is attained at $\log(M_h/M_\odot)$= 12.0 (=11.9); (b) $f_s \propto M_h$ ($f_s \propto {M_h}^3$) at the low-mass end while at the high-mass end, $f_s \propto {M_h}^{-0.4}$ ($f_s \propto {M_h}^{-0.6}$). The baryon mass fractions, f_b=M_b/M_h, of BGs and RGs reach maximum values of $f_b = 0.028^{+0.018}_{-0.011}$ and $f_b = 0.034^{+0.025}_{-0.014}$, respectively. At $M_h < 10^{11.3}M_\odot$, the dependence of f_b on M_h is much steeper for RGs than for BGs. We discuss the differences found in the f_s-M_h and f_b-M_h relations between BGs and RGs in the light of semi-empirical galaxy evolution inferences.

7.1 INTRODUCTION

The galaxy stellar and baryonic mass functions (*GSMF* and *GBMF*, respectively), inferred from the observed luminosity function and gas fraction–stellar mass (f_g–M_s) relation, contain key statistical information to understand the physical processes of galaxy formation and evolution. Within the context of the popular Λ Cold Dark Matter (ΛCDM) hierarchical scenario, dark matter halos are the sites where galaxies form and evolve [239, 238]. Hence, a connection between *GBMF* or *GSMF* and the halo mass function (*HMF*) is expected. The result of such a connection is the galaxy stellar and baryonic mass–halo mass relations, M_s-M_h and M_b–M_h, and their intrinsic scatters, both set by complex dynamical and astrophysical processes intervening in galaxy formation and evolution [see for recent reviews, 18, 5, 23]. In this sense, the M_b/M_h and M_s/M_h ratios quantify the efficiency at which galaxy and star formation proceeds within a halo of mass M_h. Therefore, the empirical or semi-empirical inference of the M_b–M_h and M_s–M_h relations and their scatters (locally and at other epochs) is nowadays a challenge of great relevance in astronomy.

For simplicity, in statistical studies like those related to the *GSMF*, galaxies are labelled by their mass alone. However, by their observed properties, correlations, and evolution, galaxies show a very different nature, at least for the two major groups in which they are classified: the rotationally-supported disk star-forming (late-type) and the pressure-supported spheroid quiescent (early-type). In the same way, the evolution of galaxies is expected to differ if they are centrals or satellites. The main intrinsic processes of galaxy evolution are associated to central galaxies, while satellite galaxies undergo several *extra* astrophysical processes because of the influence of the environment of the central galaxy/halo system in which they were accreted. Hence, if the $M_{\rm b}$–$M_{\rm h}$ or $M_{\rm s}$–$M_{\rm h}$ relations are used for constraining galaxy formation and evolution processes, these relations are required separately for at least the two main families of late- and early-type galaxies and taking into account whether the galaxy is central or satellite. Fortunately, in the last years several studies appeared, in which a decomposition of complete *GSMF*s by color, concentration or other easily measurable indicators of the galaxy type was carried out [e.g., 21, 188, 26]. Evenmore, in a recent work Yang, Mo & van den Bosch [247, hereafter YMB09] used the Sloan Digital Sky Survey (SDSS) data for obtaining the *GSMF*s of both central and central + satellite galaxies separated in each case into blue and red objects.

With the coming of large galaxy surveys, a big effort has been done in constraining the $z \sim 0$ *total* $M_{\rm s}$–$M_{\rm h}$ relation (i) *directly* by estimating halo masses with galaxy-galaxy weak lensing, with kinematics of satellite galaxies or with X-ray studies; and (ii) *indirectly* by linking observed statistical galaxy properties (e.g., the galaxy stellar mass function *GSMF*, the two-point correlation function, galaxy group catalogs) to the theoretical *HMF* [see for recent reviews and more references, 148, 146, 19, hereafter BCW10]. While the latter approach does not imply a measure-based determination of halo masses, it is simpler from a practical point of view, as it allows to cover larger mass ranges, and can be extended to higher redshifts than the former approach [see recent results in, 57, 148, 225, and BCW10]. Besides, both the weak lensing and satellite kinematics methods in practice are (still) statistical in the sense that one needs to stack large number of galaxies in order to get sufficient signal-to-noise. This introduces a significant statistical uncertainty in the inferred halo masses.

The indirect approach for linking galaxy and halo masses spans a large variety of methods, among them the Halo Occupation Distribution [161, 25, 115] and the Conditional Luminosity Function formalisms [245, 244]. These formalisms introduce a priori functional forms with several parameters that should be constrained by the observations. Therefore, the final inferred $M_{\rm s}$–$M_{\rm h}$ relation is actually model-dependent and yet sometimes poorly constrained due to degeneracies in the large number of parameters. A simpler and more empirical method –in the sense that it uses only the *GSMF* (or luminosity function) as input and does not require to introduce any model– has been found to give reasonable results. This indirect method, called the abundance matching technique (hereafter, AMT; e.g., 216, 115, 58, 187, 234, 16, 217, 57, 75, 148, 19, 102, 20, 170, 159), is based on the assumption of a monotonic correspondence between $M_{\rm s}$ and $M_{\rm h}$; in the limit of zero scatter in the $M_{\rm s}$–$M_{\rm h}$ relation, the halo mass $M_{\rm h}$ corresponding to a galaxy of stellar mass $M_{\rm s}$, is found by matching the observed cumulative *GSMF* to the theoretical cumulative *HMF*.

In this Chapter we apply the AMT in order to infer the local M_s–M_h relation for *central blue and red* galaxies separately, which requires as input *both* the observed central blue and red *GSMF*s, taken here from YMB09. Note that in order to infer the M_s–M_h relation of galaxy subpopulations (e.g., blue/red or central/satellite ones) solely from the overall *GSMF*, models for each subpopulation should be introduced, which largely increases the uncertainty in the result. Regarding the *HMF*s to be matched with the corresponding observed central *GSMF*s, the theoretical *HMF* is decomposed into two functions –associated to halos hosting blue and red galaxies– based on empirical facts: blue galaxies are rare as central objects in groups/clusters of galaxies, and they should not have undergone late major mergers because of the dynamical fragility of disk (blue) galaxies. Nowadays, it is not clear whether the M_s–M_h relation varies significantly or not with galaxy color or type. Previous studies that discussed this question were based on direct methods: the weak lensing [133] and satellite kinematics [146] techniques. The uncertainties in the results of these studies are yet large, and can be subject to biases intrinsic to the sample selection and to effects of environment.

We also estimate here the galaxy baryon mass-halo mass relations, M_b–M_h[1], where $M_b = M_s + M_g$, by using the *GSMF*s combined with average observational estimates of the galaxy gas mass, M_g, as a function of M_s. The galaxy baryonic mass fraction, $f_b = M_b/M_h$, and its dependence on mass is important for constraining models and simulations of galaxy evolution, and is also a key input for some approaches, implemented for modelling the most generic population of galaxies, namely *isolated (central) disk galaxies* [e.g., 144, 82, 218, 203, 80, 96, 79]. In these and other studies, it was shown that several disk galaxy properties, correlations and their scatters depend (or are constrained) by f_b. In a similar way, the f_b-M_h dependence is expected to play some role in the results of structural and dynamical models of spheroid-dominated galaxies.

In Section 7.2 we describe the method and the data input. The stellar/baryon mass–halo mass relations for the total, blue and red (sub)samples are presented in Section 7.3. In Section 7.4 we compare our results with other observational works, and discuss whether they are consistent or not with expectations of semi-empirical inferences. The summary and our conclusions are given in Section 7.5.

7.2 THE METHOD

The AM statistical technique is based on the hypothesis of a one-to-one monotonic increasing relationship between M_s (or M_b) and M_h. Therefore, by matching the *cumulative* galaxy stellar and halo mass functions, for a given M_s a unique M_h is assigned:

$$\int_{M_h}^{\infty} \phi_h(M'_h)dM'_h = \int_{M_s}^{\infty} \phi_s(M'_s)dM'_s, \tag{59}$$

where ϕ_h is the overall *HMF* (distinct + subhalos) and ϕ_s is the overall *GSMF*; *distinct* halos are those not contained inside more massive halos. It is reasonable to link central galaxies with distinct halos. Therefore, in the case of using the *GSMF* for only central galaxies, the distinct *HMF* should be used for the matching. Since the main purpose of this Chapter is the inference of the M_s–M_h (and the corresponding M_b–M_h) relation for

1 We assume that the galaxy baryonic mass is included in the halo (virial) mass M_h.

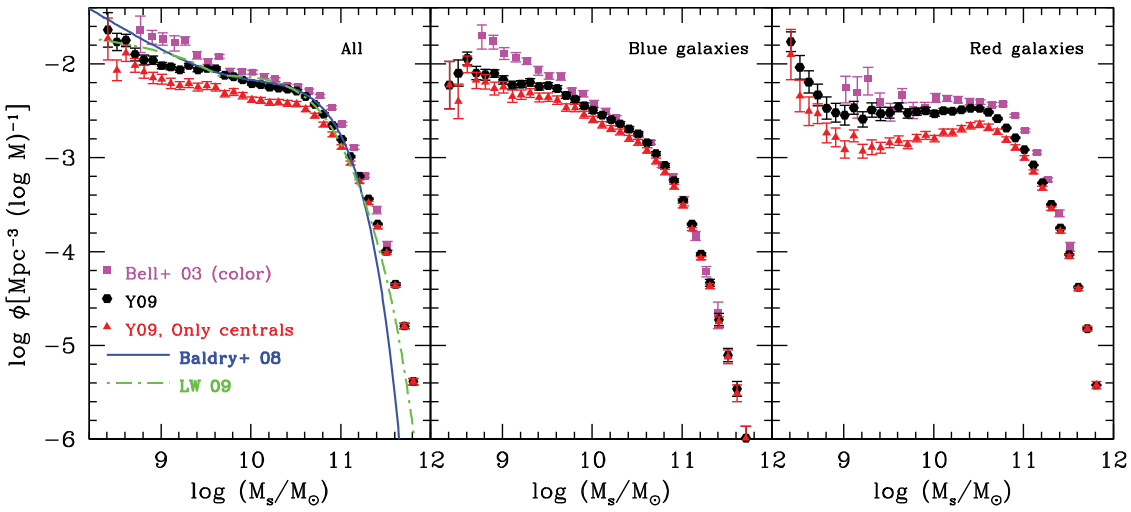

Figure 29: *Left panel:* Different local *GSMF*s for *all* galaxies. The reported data in Bell et al. [21, pink squares] and YMB09 (black hexagons) are plotted directly, while for Baldry, Glazebrook & Driver [16, blue solid line] and Li & White [124, dot-dashed green line], the best fits these authors find to their samples are plotted. Red triangles show the data from YMB09 corresponding to the *GSMF* of central-only galaxies. *Middle and right panels:* Data corresponding to the decomposition of the *GSMF* into blue and red galaxies, respectively, from Bell et al. [21] and for all and central-only galaxies from YMB09.

blue (red) galaxies, (i) a *GSMF* that separates galaxies by color is necessary (the data to be used here are discussed in 7.2.1), and (ii) a criterion to select the halos that will likely host blue (red) galaxies shall be introduced (see 7.2.2.1).

In this Chapter we will not carry out an exhaustive analysis of uncertainties in the inference of the M_s–M_h relation with the AMT. This was extensively done in BCW10 [see also, 148]. In BCW10 the uncertainty sources are separated into three classes: (i) in the observational inference of *GSMF*, (ii) in the dark matter HMF, which includes uncertainties in the cosmological parameters, and (iii) in the matching process arising primarily from the intrinsic scatter between M_s and M_h.

7.2.1 *Galaxy and Baryonic Stellar Mass Functions*

In the last years, complete galaxy luminosity functions (and therefore, *GSMF*s) were determined for local samples covering a large range of luminosities (masses). The stellar mass is inferred from (multi)photometric and/or spectral data (i) by using average stellar mass-to-light ratios, depending only on color (inferred from application of stellar population synthesis –SPS– models to galaxy samples with independent mass estimates, e.g. Bell et al. [21]), or (ii) by applying directly the SPS technique to each sample galaxy, when extensive multi-wavelength and/or spectral information is available.

In both cases, a large uncertainty is introduced in the inference of M_s due to the uncertainties in the IMF, stellar evolution, stellar spectral libraries, dust extinction, metallicity,

etc. Bell et al. [21] estimated a scatter of ≈ 0.1 dex in their M_s/L ratios in infrared bands. Conroy, Gunn & White [55] carried out a deep analysis of propagation of uncertainties in SPS modelling and concluded that M_s at $z \sim 0$ carry errors up to ~ 0.3 dex (but see Gallazzi & Bell [89]). Here, we will consider an overall systematical uncertainty of 0.25 dex in the M_s determination (see BCW10).

Most of the current local *GSMF*s were inferred from the 2dF Galaxy Redshift Survey, Two Micron All-Sky Survey (2MASS) and SDSS [e.g., 52, 21, 13]. The low-mass completeness limit due to missing of low surface brightness galaxies is at $\sim 10^{8.5}$ M_s [16]. An upturn of the *GSMF* close to this end (below $M_s \sim 10^9$ $M_\odot$) was confirmed in several recent works (16; YMB09; 124). Due to this upturn, a better fit to the *GSMF*s is obtained by using a double or even triple Schechter function. Since the low-mass end of the *GSMF* is dominated by late-type galaxies, this upturn plays an important role in the M_s–M_h relation of late-type galaxies at low masses.

For our purposes, observational works where the *GSMF* is decomposed into late- and early-types galaxies are required. Such a decomposition has been done, for example, in Bell et al. [21], who combined 22679 SDSS Early Data Release and 2MASS galaxies, and used two different criteria, color and concentration, to split the sample into two types of galaxies. A much larger sample taken from the NYU-VAGC based on the SDSS DR4 has been used by YMB09 (see also Yang, Mo & van den Bosch [246]), who split the sample into *blue and red* subsamples according to a criterion in the ${}^{0.1}(g-r) - M_r$ diagram. In both works, M_s is calculated from the $r-$band magnitude by using the corresponding color-dependent M_s/L_r ratio given in Bell et al. [21]. In YMB09 each color subsample is in turn separated into central and satellite galaxies according to their memberships in the constructed groups, where the central galaxy is defined as the most massive one in the group and the rest as satellite galaxies.

In Figure 29, the Bell et al. [21] and YMB09 *GSMF*s are reproduced by using the data sets reported in these papers. In the left panel, the full sample from each work (solid squares and solid hexagons, respectively) are plotted, as well as the case of central-only galaxies from YMB09 (solid triangles); both *GSMF*s and the other ones plotted in this figure are normalised to $h = 0.7$ and to a Chabrier [48] IMF. In the central and right panels, the corresponding blue (late-type) and red (early-type) sub-samples are plotted with the same symbols on the left panel. For the Bell et al. [21] sub-samples, only those separated by their color criterion are plotted. Both *GSMF*s corresponding to the full and blue sub-samples are in good agreement for $M_s \gtrsim 10^{9.5}$ $M_\odot$. For lower masses, the Bell et al. [21] *GSMF*'s are higher. On one hand, the Bell et al. [21] sample is much smaller than the YMB09 one (therefore the effect of cosmic variance is more significant). On the other hand, the redshift completeness and M_s limit in YMB09 is treated with updated criteria.

In Figure 29, we also plot fits to the overall *GSMF* presented in Baldry, Glazebrook & Driver [16, double Schechter function, solid blue line] and in Li & White [124, triple Schechter function, dashed green line] for new SDSS releases and by using directly SPS models to estimate M_s for each galaxy. These fits agree well with the YMB09 data in the mass range $9.2 \lesssim \log(M_s/M_\odot) \lesssim 11.2$. For smaller masses, the Baldry, Glazebrook & Driver [16] fit tends to be steeper while the Li & White [124] fit tends to be shallower than the YMB09 data. For larger masses, both fits decrease faster with M_s than the YMB09

data. All these (small) differences are due to the different methods used to estimate M_s, as well as the different volumes and limit corrections of the samples (see,16; YMB09; and 124 for discussions).

The split into two colors of the sample used by YMB09 is a rough approximation to the two main families of disk- and spheroid-dominated galaxies. It is well known that the morphological type correlates with the galaxy color, though with a large scatter. There is for example a non-negligible fraction of galaxies (mostly highly inclined) that are red but of disk-like type (e.g., Bernardi et al. [26]). However, given that here we consider a partition of the overall sample just in two groups, we believe that is reasonable to assume in a first approximation that the color criterion for the partition will also serve at this level as a morphological criterion.

For the YMB09 sample, the fractions of blue and red galaxies are $\approx 55\%$ and $\approx 45\%$, respectively, for $M_s \gtrsim 3 \times 10^8$ $M_\odot$. Red galaxies dominate the total *GSMF* at large masses. At $M_s \approx 2 \times 10^{10}$ $M_\odot$ the abundances of red and blue galaxies are similar and at lower masses the latter are increasingly more abundant than the former. For $M_s \lesssim 10^9$ $M_\odot$, the abundance of red galaxies, mainly central ones, steeply increases towards smaller masses. The existence of this peculiar population of faint central red galaxies is discussed in YMB09. Wang et al. [227] suggested that these galaxies are hosted by small halos that have passed through their massive neighbors, and the same environmental effects that cause satellite galaxies to become red are also responsible for the red colors of such galaxies. However, as these authors showed, even if the environmental effects work, there are in any case over 30% of small halos that are completely isolated in such a way that these effects can not be invoked for them.

In the YMB09 sample, around 70% of the galaxies are central. As mentioned in the Introduction, the inference of the M_s–M_h relation for central-only galaxies is important for studies aimed to constrain galaxy formation and evolution in general; satellite galaxies are interesting on its own but they lack generality because their evolution and properties are affected by extra environmental processes.

In what follows, the YMB09 *GSMF* provided in tabular form and split into blue/red and central/satellite galaxies will be used to apply the AMT. Our main goal is to infer the M_s–M_h relation for central blue (late-type) and red (early-type) galaxies.

We will infer also the corresponding M_b–M_h (baryonic) relations. The blue and red *GBMF*s are estimated from the blue and red *GSMF*s, respectively, where in order to pass from M_s to M_b, the cool (atomic and molecular) gas mass, M_g, corresponding on average to a given M_s is taken from the empirical blue and red M_g–M_s relations. In Figure 30, a compilation of observational estimates is plotted in the M_s–M_g plane for a sample of disk galaxies that includes low surface brightness galaxies from Avila-Reese et al. [12, blue dots with error bars; they added H_2 mass contribution by using an estimate for the H_2-to-HI mass ratio as a function of galaxy type], and for another galaxy sample from McGaugh [137, blue crosses; no H_2 contribution is considered and their dwarf galaxies were excluded]. An orthogonal linear doubly-weighted regression to the data from Avila-Reese et al. [12] gives:

$$\frac{M_g}{10^{10}\mathrm{M}_\odot} = 0.43 \times \left(\frac{M_s}{10^{10}\mathrm{M}_\odot}\right)^{0.62}. \tag{60}$$

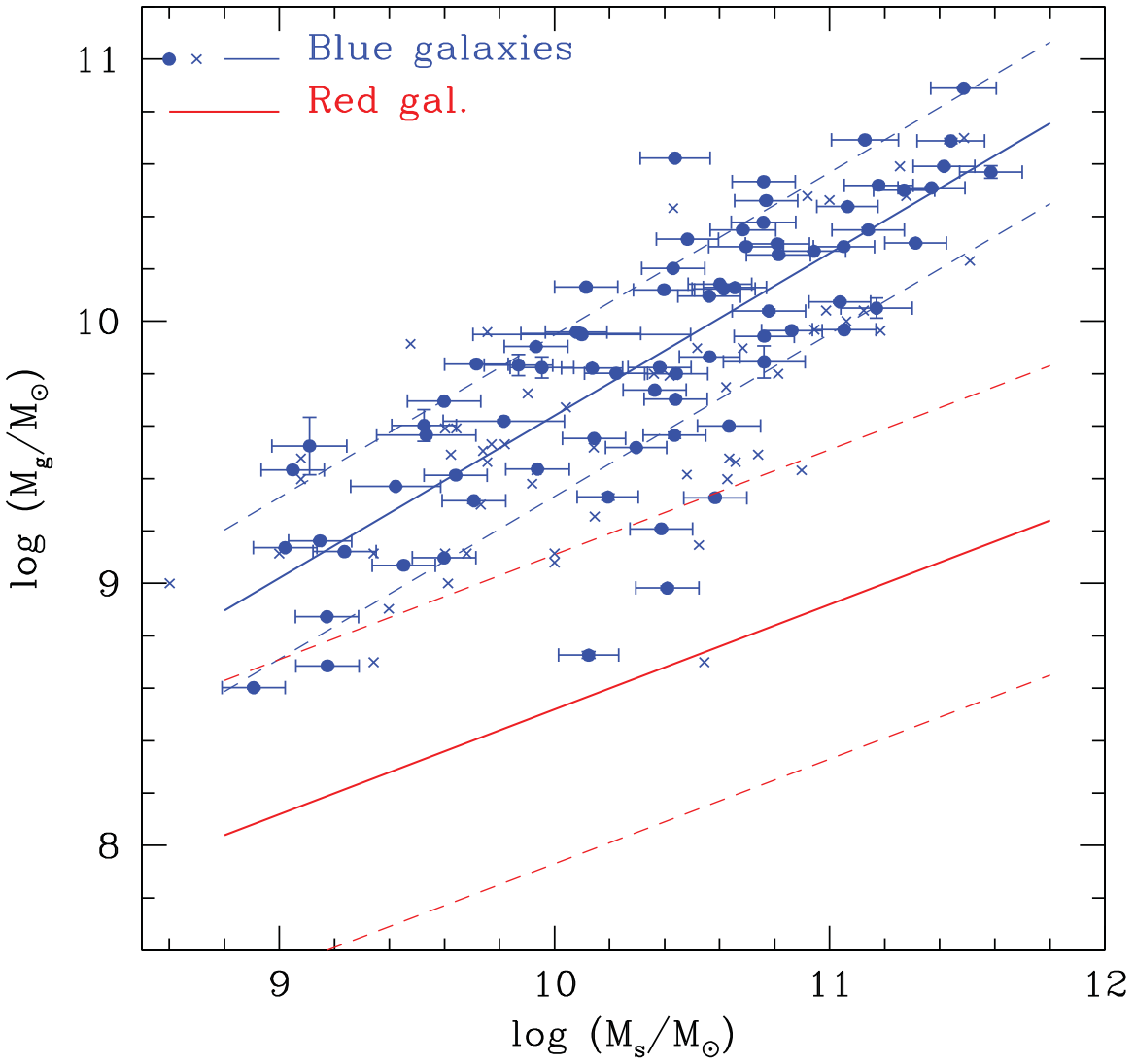

Figure 30: Gas mass vs stellar mass for a sample of disk high and low surface brightness galaxies collected and homogenised by Avila-Reese et al. [12, blue dots with error bars] and for a sample of disk galaxies presented by McGaugh [137, blue crosses]. The solid blue line is the orthogonal linear doubly-weighted regression to the data from the former authors and the dashed lines show an estimate of the intrinsic scatter around the fit. The solid red line is an estimate of the M_g–M_s correlation for red galaxies by using our fit to blue galaxies and the ratio of blue-to-red atomic gas fraction determined in Wei et al. [233, see text].

This fit is plotted in Figure 30 with its corresponding estimated scatter (≈ 0.3 dex; blue solid and dashed lines). This is the relation and its scatter used to calculate $M_{\rm b}$ and the blue *GBMF*. A similar relation has been inferred by Stewart et al. [199]. The gas fractions in red galaxies are much smaller than in blue galaxies. For sub-samples of blue and red galaxies, Wei et al. [233] reported the atomic gas fractions versus $M_{\rm s}$ (molecular gas was not included). The ratio of both of their fits as a function of $M_{\rm s}$ is used here to estimate from eq. (60, blue galaxies) the corresponding average $M_{\rm g}$ for red galaxies as a function of $M_{\rm s}$ (red solid line). As an approximation to the scatter (short-dashed lines), the average scatter reported for red galaxies in Wei et al. [233] is adopted here.

7.2.2 *Halo and sub-halo mass functions*

A great effort has been done in the last decade to determine the *HMF* at $z = 0$ and at higher redshifts in N-body cosmological simulations. A good fit to the results, at least for low redshifts, is the universal function derived from a Press-Schechter formalism [166] generalized to the elliptical gravitational collapse [192, hereafter S-T]. In fact, Tinker et al. [207] have shown that at the level of high precision, the *HMF* changes for different cosmological models and halo mass definitions as well as a function of z. For our purposes and for the cosmology used here, the S-T approximation provides a good description of the $z = 0$ *HMF* of distinct halos:

$$\phi_h(M_h)dM_h = A\left(1+\frac{1}{\nu^{2q}}\right)\sqrt{\frac{2}{\pi}}\frac{\bar{\rho}_M\nu}{M_h^2}\left|\frac{d\ln\sigma}{d\ln M_h}\right|\exp\left[-\frac{\nu^2}{2}\right]dM_h \qquad (61)$$

where $A = 0.322$, $q = 0.3$, $\nu^2 = a(\delta_c/D(z)\sigma(M_{\rm h}))$ with $a = 0.707$, $\delta_c = 1.686\Omega_m^{0.0055}$ is the linear threshold in the case for spherical collapse in a flat universe with cosmological constant, $D(z)$ is the growth factor and $\sigma(M_{\rm h})$ is the mass power spectrum variance of fluctuations linearly extrapolated to $z = 0$. The halo (virial) mass, $M_{\rm h}$ is defined in this Chapter as the mass enclosed within the radius where the overdensity is $\bar{\rho}_{\rm vir} = \Delta$ times the *mean matter* density, $\bar{\rho}_M$; $\Delta \approx 340$ according to the spherical collapse model for the cosmology used here. The cosmological parameters assumed here are close to those of WMAP5 (Komatsu et al. 2009): $\Omega_M = 0.27, \Omega_\Lambda = 1 - \Omega_m = 0.73, h = 0.70, \sigma_8 = 0.8$.

The distinct *HMF* should be corrected when a *GSMF* corresponding to *all* galaxies is used in the AMT. In this case, satellite galaxies are included in the *GSMF*. Therefore, subhalos should be taken into account in the *HMF*. The subhalo fraction is no more than $\approx 20\%$ of all the halos at $z = 0$ [e.g., 187, 58, 93, see also BCW10]. When necessary, we correct the S-T *HMF* for (present-day) subhalo population by using the fitting formula to numerical results given in Giocoli et al. [93]:

$$\frac{dn(m_{sub})}{d\ln m_{sub}} = A_0 m_{sub}^{\eta-1}\exp\left[-\left(\frac{m_{sub}}{m_0}\right)^{\gamma}\right], \qquad (62)$$

with $\eta = 0.07930$, $\log A_0 = 7.812$, $\log(m_0/{\rm M}_\odot) = 13.10$ and $\gamma = 0.407$.

The upper panel of Figure 31 shows the (distinct) S-T *HMF* (solid line), the subhalo *HMF* (short-long-dashed line), and the distinct+subhalo *HMF* (dash-dotted line). The correction by sub-halos in the abundance is small at low masses and negligible

at high masses. When the *GSMF* refers only to central galaxies –which is the case in this Chapter–, then it is adequate to use the distinct *HMF* for the AMT, i.e. *the subhalo abundance correction is not necessary*.

7.2.2.1 *Haloes hosting blue and red galaxies*

In the AMT, the cumulative *GSMF* and *HMF* are matched in order to link a given M_s to M_h. When a subsample of the total *GSMF* is used –as is the case for inferring the M_s–M_h relation of only late- or early-type galaxies– it would not be correct to use the total *HMF* for the matching. This function, in the ignorance of which is the mass function of halos hosting blue (red) galaxies, at least should be re-normalised (decreased uniformly) by the same fraction corresponding to the decrease of the sub-sample *GSMF* with respect to the total *GSMF*. In YMB09, $\approx$ 55% ($\approx$ 45%) of the galaxies are in the blue (red) sub-samples for $M_s \gtrsim 3 \times 10^8$ $M_\odot$. We may go one step further by proposing general observational/physical conditions for halos to be the host of blue (late-type) or red (early-type) galaxies. Note that the division we do here of galaxies is quite broad –just in two groups–, therefore very general conditions are enough.

Haloes that host central blue and red galaxies are expected to have (i) a different environment, and (ii) a different merger history. We take into account these two factors in order to roughly estimate the *HMF* of those halos that will host today central blue and red galaxies.

Environment.- Blue (late-type) galaxies are rare in the centers of groups and clusters of galaxies [high-density environments; e.g., 155, 259, 123, 68, 158, 32, and more references therein]. For example, in the SDSS YMB09 sample that we use here (see also Weinmann et al. [236]), among the groups with 3 or more members, the fraction of those with a central blue galaxy is only $\approx$ 20%, and most of these central galaxies are actually of low masses. Therefore, cluster- and group-sized halos (more massive than a given mass) can not be associated to central blue galaxies when using the AMT. This means that the halo mass function of groups/clusters of galaxies should be excluded from the theoretical *HMF* (Shankar et al. [187]).

Heinämäki et al. [105] have determined the *HMF* of groups with 3 or more members and with a number density enhancement $\delta n/n \geq 80$ from the Las Campanas Redshift Survey. The authors estimated the corresponding group virial mass on the basis of the line-of-sight velocity and harmonic radius of the group, in such a way that this mass is defined at the radius where $\delta n/n$=80. The observational galaxy overdensity $\delta n/n$ is related to the mass overdensity $\delta\rho/\rho$ roughly through the bias parameter b: $\delta\rho/\rho = (1/b)\times\delta N/N$, where $b \approx 1/\sigma_8$ [135]. Hence, for $\sigma_8 = 0.8$, $\delta\rho/\rho \approx 64$; since the group selection was carried out in [215], where an Einstein-de Sitter cosmology was used, then $\rho = \rho_{crit}$ in this case. In our case, the halo virial mass is defined at the radius where $\delta\rho/\rho \approx 340$ (see §§2.2); in terms of ρ_{crit}, our overdensity is $340 \times \Omega_M = 92$. Therefore, the halo virial masses in Heinämäki et al. [105] should be slightly larger than those used here. For the NFW halos of masses larger than $\sim 10^{13}$ $M_\odot$, the differences are estimated to be a factor of 1.10-1.20. We correct the group masses of Heinämäki et al. [105] by 15%. In the upper panel of Fig. 31, the corrected group (halo) mass function is reproduced (solid dots) and an eye-fit to them is plotted (dot-dashed cyan line).

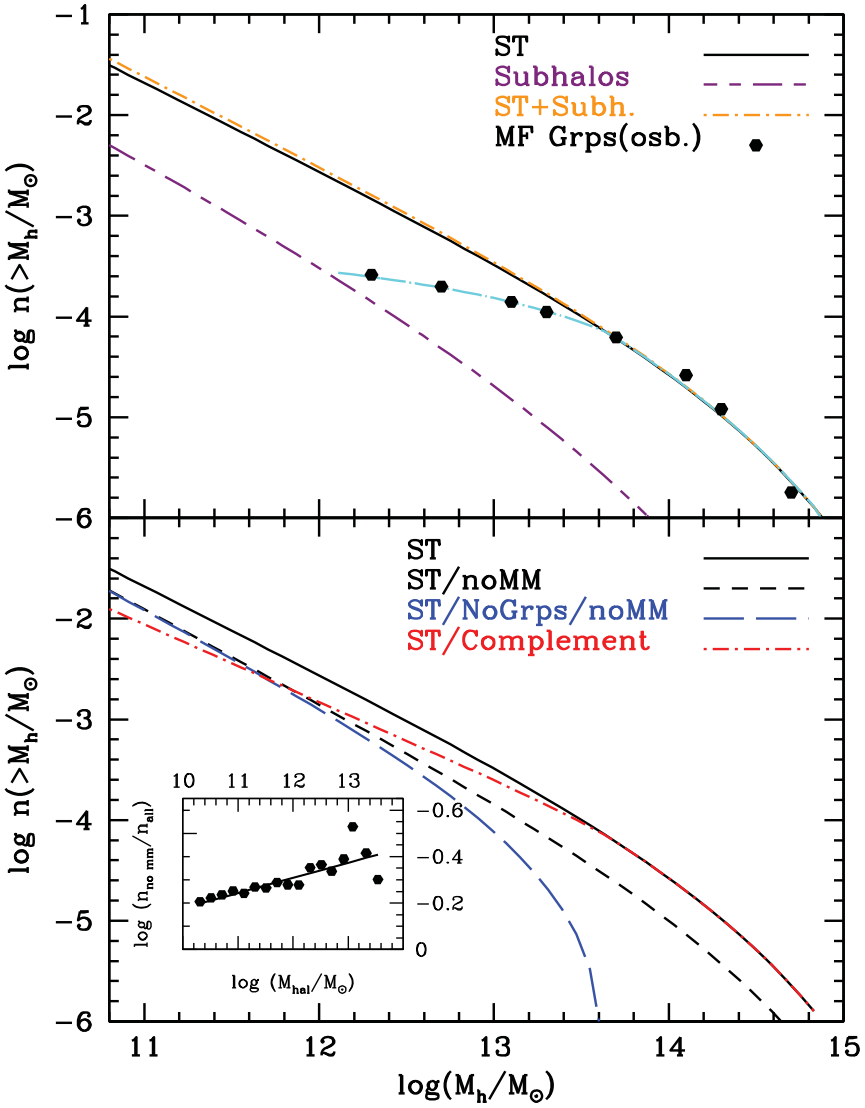

Figure 31: *Upper panel:* Distinct S-T *HMF* for the cosmology adopted in this paper (solid black line), sub-halo mass function at $z = 0$ according to Giocoli et al. [93, short-long-dashed purple line], and the sum of both (dot-dashed orange line). The solid dots are measures of the group/cluster mass function according to Heinämäki et al. [105] and adequately corrected to our definition of virial halo mass; the dot-long-dashed cyan line is an eye-fit to the data. *Lower panel:* The same distinct S-T *HMF* (solid black line) shown in the upper panel but (i) excluding halos that suffered late major mergers –since $z = 0.8$– (short-dashed black line) and (ii) excluding these halos and those of observed groups/clusters (long-dashed blue line). The latter is the *HMF* to be assigned to the sub-sample of central blue galaxies. The complement of this function to the total (S-T) (dot-dashed red line) is the *HMF* to be assigned to the sub-sample of central red galaxies. The inset shows the ratio of number densities of halos that did not suffer major mergers since $z = 0.8$ to all the (distinct) halos according a cosmological N-body simulation (Colín et al. 2011, see text). The fit to this ratio (solid line in the inset) is what has been used to correct the S-T *HMF* for halos that did not suffer late major mergers.

Merger history.- Disk (blue, late-type) galaxies are dynamically fragile systems and thus they are not expected to survive strong perturbations such as those produced in major mergers or close interactions. However, as several theoretical studies have shown [e.g., 173, 97], when the mergers are gas-rich ('wet') and/or at early epochs (in fact, both facts are expected to be correlated), it is highly probable that a gaseous disk is regenerated or formed again with the late accreted gas. Therefore, a reasonable restriction for halos that will host disk galaxies is that they did not undergo *central* major mergers since a given epoch (at earlier epochs, while the central major merger may destroy the disk, a new gaseous disk can be formed later on). Based on numerical simulations, Governato, Mayer & Brook [97] suggested that a 'wet' major merger of disk galaxies at $z \sim 0.8$ has yet a non-negligible probability of rebuilding a significant disk by $z \sim 0$. We will assume here that halos for which their *centers* have a major merger at $z < 0.8$ will not host a disk galaxy.

In Colín et al. (2011; in prep.) the present-day abundance fraction of halos with no *central* major merger since $z = 0.8$ was measured as a function of $M_{\rm h}$ from an N-body ΛCDM cosmological high–resolution simulation with $\Omega_m = 0.24$, $\Omega_\Lambda = 0.76$, and $\sigma_8 = 0.75$ (box size and mass per particle of 64 h^{-1}Mpc and $1.64 \times 10^7 h^{-1}\mathrm{M}_\odot$, respectively). The friends-of-friends (FOF) method with a linking-length parameter of 0.17 was applied for identifying halos. The mass ratio to define a major merger was $q = M_{\rm h,2}/M_{\rm h,1} > 0.2$ and the merger epoch was estimated as that one when the center of the accreted halo arrived to the center of the larger halo by dynamical friction; this epoch is calculated as the cosmic time when both FOF halos have "touched" plus the respective dynamical friction (merging) time as given by the approximation of Boylan-Kolchin, Ma & Quataert [40]. The fraction of halos that did not suffer a major merger since $z = 0.8$ with respect to all the halos as a function of $M_{\rm h}$ measured in Colín et al. (2011) is used here to correct our distinct S-T *HMF*. This measured fraction is showed in the inset in the lower panel of Fig. 31; the solid line is a linear fit by eye in the log-log plot: $\log(n_{\rm noMM}/n_{\rm all}) = 0.472 - 0.065\log(M_{\rm h}/\mathrm{M}_\odot)$. As it is seen, the fraction slightly decreases with mass, which is consistent with the idea that larger mass halos are assembling later with a significant fraction of their masses being acquired in late major mergers. After the correction mentioned above, we get the mass function of halos that did not suffer a central major merger ($q > 0.2$) since $z = 0.8$ (short-dashed black line in the lower panel of Fig. 31).

The final corrected HMFs.- The function obtained after (i) subtracting from the distinct S-T *HMF* the group mass function and (ii) excluding halos that did not suffer a late central major merger is plotted in Fig. 31 (blue long-dashed line). This mass function is proposed here to correspond to halos that host blue galaxies today. The overall number fraction of these halos with respect to the distinct ones (described by the S-T *HMF*) is $\sim$ 58%, which is roughly consistent with the fraction of blue galaxies in the YMB09 sample. The *HMF* corresponding to the complement is plotted in Fig. 31 as the red dot-dashed curve. By exclusion, this *HMF* will be associated to the *GSMF* of the red central galaxy sub-sample for deriving the $M_{\rm s}$–$M_{\rm h}$ relation of red galaxies.

7.3 RESULTS

7.3.1 *The overall, central, and satellite stellar–halo mass relations*

In Fig. 32, the M_s–M_h relation obtained by using the Li & White [124] *GSMF* (see 7.2.1 and Fig. 29) and the S-T *HMF* corrected to include sub-halos is plotted (long-dashed blue line). The relation given by BCW10, who also used as input the Li & White [124] *GSMF*, is shown (short-dashed red line). Both curves are almost indistinguishable, indicating an excellent consistency in spite of the differences in some of the methodological aspects.

Further, we plot in Fig. 32 the M_s–M_h relation but using now the total YMB09 *GSMF* (dot-dashed pink line). This relation is similar to the one inferred using the Li & White [124] *GSMF*. For $\log(M_h/M_\odot) \gtrsim 12$, the former shifts with mass slightly to higher values of M_s for a given M_h than the latter (at $\log(M_h/M_\odot) = 13.5$ the difference is not larger than 0.08 dex in $\log M_s$). Such a shift is explained by the (small) systematical difference between the YMB09 and Li & White [124] *GSMF*s at masses larger than $\log(M_s/M_\odot) \sim 11$ (see 7.2.1 and Fig. 29).

In Fig. 32, the M_s–M_h relations given in Baldry, Glazebrook & Driver [16, dot-dashed orange line], Moster et al. [148, short-long-dashed line] and Guo et al. [102, dotted green line] are also plotted. When it was necessary, we have corrected the stellar masses to the Chabrier IMF, and the halo masses to the definition of virial mass used here (see 7.2.2.1). As mentioned above, Baldry, Glazebrook & Driver [16] corrected their *HMF* to exclude groups/clusters of galaxies (something that we do but only for the central blue galaxies, see 7.2.1 and the result below). As seen in Fig. 32, their correction produces a steeper M_s–M_h relation at the high-mass side than in our case. Moster et al. [148] and Guo et al. [102] constrained the M_s–M_h relation by assigning stellar masses to the halos and subhalos of an N-body cosmological simulation in such a way that the total *GSMF* is reproduced. Therefore, by construction, their M_s–M_h relations take into account the group/cluster halo masses issue. The M_s–M_h relations in both works are also slightly steeper than ours at high masses but shallower on average than that one of Baldry, Glazebrook & Driver [16]. Note that in BCW10 the scatter in M_s at fixed M_h was taken into account but the group/cluster halo masses issue was not.

The M_s–M_h relation using the YMB09 *GSMF* only for central galaxies and the distinct (S-T) *HMF* is plotted in the lower panel of Fig. 32 (solid black line). At large masses, this relation is quite similar to the one for all galaxies/satellites and halos/sub-halos (dot-dashed pink line). This is because at large masses the great majority of galaxies are centrals and the correction by sub-halos is negligible (see Figs. 29 and 31). At lower masses, the exclusion of satellites and sub-halos implies a lower M_s for a given M_h. This is because the *GSMF* decreases more than the *HMF* at lower masses when passing from the total (galaxy and halo) samples to the central-only galaxy/distinct halo samples. The physical interpretation of this result could be that satellite galaxies of a given M_s have less massive halos than central galaxies due to tidal stripping. The M_s–M_h relation derived only for the satellites YMB09 *GSMF* and the Giocoli et al. [93] $z = 0$ sub-halo *HMF* is plotted in the lower panel of Fig. 32 (short-long-dashed cyan line).

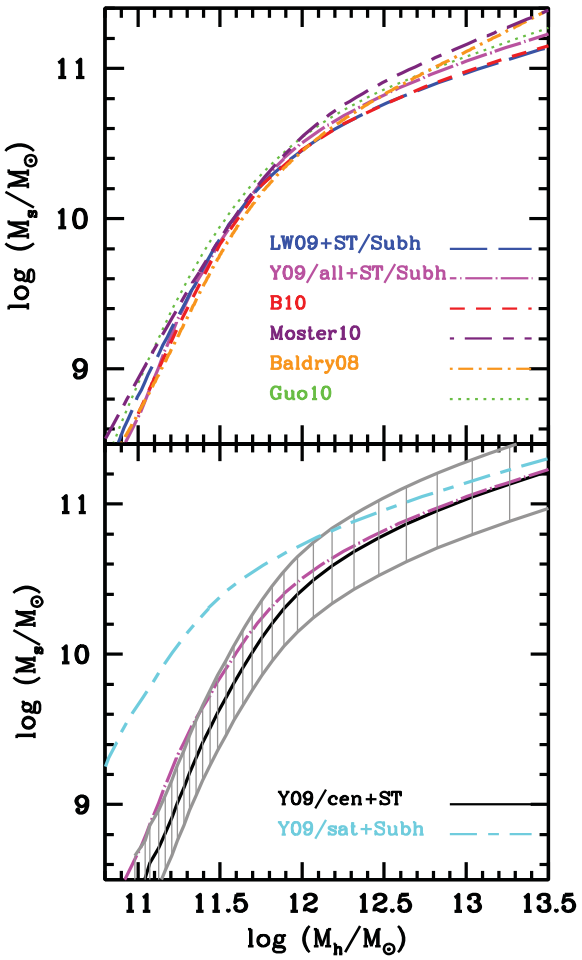

Figure 32: *Upper panel:* Stellar mass vs halo mass as inferred here by using the Li & White [124] overall *GSMF* and the S-T *HMF* increased by the subhalo population (long-dashed blue line) to be compared with the BCW10 inference, who used the same *GSMF* (short-dashed red line). The dot-dashed pink line shows the same M_s vs M_h inference but using the overall YMB09 *GSMF*. Different estimates of the overall M_s–M_h relation by other authors (indicated in the panel), who took into account in different ways the issue of group/cluster masses (see text) are also plotted *Lower panel:* Same M_s–M_h relation as in the upper panel (dot-dashed pink line) but for the central-only YMB09 *GSMF* and the S-T (distinct) *HMF* (solid line). The grey curves connected by vertical lines show the estimated 1σ uncertainty for the latter case. The M_s–M_h relation inferred for the only satellite YMB09 *GSMF* and the Giocoli et al. [93] $z = 0$ sub-halo mass function is plotted with the short-long-dashed cyan line.

7.3.1.1 *Uncertainties*

The uncertainty (standard deviation) in the M_s–M_h relation obtained by using the YMB09 central *GSMF* and the distinct S-T *HMF* (solid line), is plotted in Fig. 32 (grey curves connected by vertical lines). As remarked in 7.2, we did not take into account all possible uncertainty sources in the M_s–M_h relation but have just considered the two following ones:

(i) The systematic uncertainty in stellar mass estimates, which is an uncertainty in the *GSMF*. We assume for this uncertainty a scatter of 0.25 dex (Gaussian distributed) independent of mass, and propagate it to the M_s–M_h relation (by far it results the dominant source of error in the relation obtained with the AMT, see below and BCW10).

(ii) The intrinsic scatter in stellar mass at a fixed halo mass, which is an uncertainty in the process of matching abundances. To take into account this scatter in M_s at fixed M_h a probability density distribution should be assumed. The convolution of this distribution with the true or intrinsic *GSMF* gives the measured *GSMF*. The cumulative true *GSMF* is then the one used for the AMT (BCW10). The observational data allow to estimate the scatter in luminosity (or M_s) which appears to be independent of M_h [145, 146, 247]. In BCW10 a log-normal mass-independent scatter in M_s of 0.16 ± 0.04 is assumed. Here, we follow the overall procedure of BCW10 for taking into account this scatter.

We also explored the effect of *(iii)* the statistical uncertainty in the number density of the *GSMF* (as given in YMB09), but we have found that the effect is negligible compared to the one produced by item (i) (see also BCW10, their §§4.3.1). The effect of the intrinsic scatter in M_s for a given M_h is also very small in the overall scatter of the M_s–M_h relation but it affects the high mass end of the calculated M_s–M_h relation, where both the *GSMF* and *HMF* decay exponentially, since there are more low mass galaxies that are scattered upward than high mass galaxies that are scattered downward (BCW10). For instance, at $M_h = 10^{13.5}$ $M_\odot$, the stellar mass after including this scatter is 1.2 times smaller. The contribution from all other sources of error, including uncertainties in the cosmological model, is much smaller ranging from 0.02 to 0.12 dex at $z = 0$.

From Fig. 32 we see that the 1σ uncertainty in the M_s–M_h relation is approximately 0.25 dex in $\log M_s$ without any systematic dependence on M_h, in good agreement with previous results [19, 148]. This uncertainty is larger than the differences between the M_s–M_h average relations found by different authors, including those that use the indirect AMT but with different *GSMF*s, methodologies, and corrections, and those who use more sophisticated formalisms (see BCW10 and More et al. [146]). On one hand, this shows that most methods and recent studies aimed at relating halo masses to observed galaxies as a function of their stellar masses are converging to a relatively robust determination. On the other hand, this result suggests that attaining a higher precision in estimating M_s from observations is a crucial task for lowering the uncertainty in the inference of the M_s–M_h relation.

7.3.2 *The stellar-halo mass relations for central blue and red galaxies*

The upper and lower left panels of Figure 33 show the mean M_s–M_h and f_s–M_h relations for: all central galaxies (solid line, as in Fig. 32), central blue (short-dashed line), and

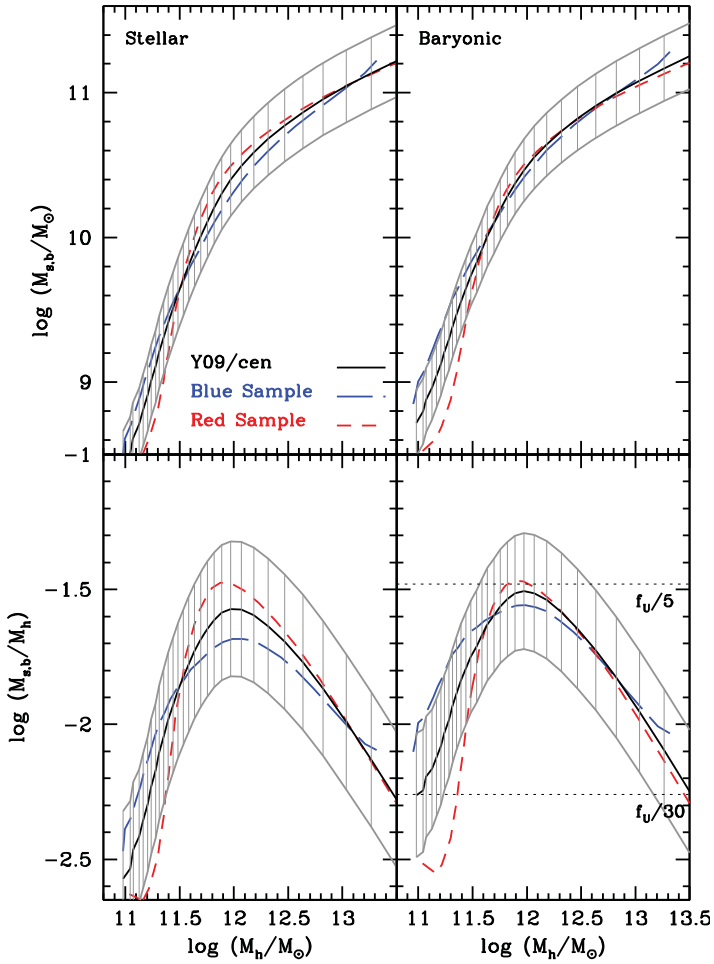

Figure 33: *Left panels:* Mean M_s–M_h (top) and f_s–M_h (down) relations of all central (solid black line), blue central (long-dashed blue line), and red central (short-dashed red line) galaxies as inferred here by using the YMB09 data. The grey curves connected by vertical lines show the 1σ uncertainty when all galaxies are considered; similar uncertainty bands around the main relations are found for the blue and red sub-samples (see Fig. 36). *Right panels:* Same as in left panels but for M_b instead of M_s. Dotted lines: $f_b = f_U/5$ and $f_U/30$, where $f_U = 0.167$ is the universal baryon fraction.

central red (long dashed line) galaxies. In order to infer these relations for blue galaxies, we use the central blue YMB09 *GSMF* and the distinct (S-T) *HMF* corrected for excluding halos (i) associated to observed groups/clusters of galaxies and (ii) that suffered central major mergers since $z = 0.8$ (see 7.2.2.1). In the case of red galaxies, the central red YMB09 *GSMF* and the *HMF* complementary to the one associated to blue galaxies were used.

The shaded area in Figure 33 is the same 1σ uncertainty showed in Fig. 32 for the overall central sample. The uncertainties corresponding to the M_s–M_h and f_s–M_h relations for the blue and red galaxy sub-samples would be close to the one of the total sample in case the corrections made to the *HMF* do not introduce an extra uncertainty. In fact this is not true, in particular for the group/cluster mass function introduced to correct the *HMF* associated to blue galaxies. Unfortunately, the work used for this correction (Heinämäki et al. [105]) does not report uncertainties. Hence, the uncertainties calculated here for the blue and red samples (shown explicitly in Fig. 36 below) could be underestimated, specially at large masses.

In the mass range $11.5 \lesssim \log(M_h/M_\odot) \lesssim 13.0$, the M_s–M_h and f_s–M_h relations for central blue (red) galaxies lie slightly below (above) the relations corresponding to the overall sample. For masses below these ranges, the trends invert. The f_s–M_h curves for blue and red sub-samples peak at $\log(M_h/M_\odot) = 11.98$ and 11.87, with values of $f_s = 0.021^{+0.016}_{-0.009}$ and $f_s = 0.034^{+0.026}_{-0.015}$, respectively. The corresponding stellar masses at these peaks are $\log(M_s/M_\odot) = 10.30 \pm 0.25$ for blue galaxies and $\log(M_s/M_\odot) = 10.40 \pm 0.25$ for red galaxies. These masses are around a factor of 0.23 and of 0.30 the characteristic stellar mass $M^\star \approx 10^{10.93} M_\odot$ of the overall YMB09 *GSMF*, respectively. The maximum difference between the blue and red mean M_s–M_h relations is attained at $\log(M_h/M_\odot) \approx 11.9$; at this mass, the f_s value of the former is 1.7 times smaller than the f_s of the latter. For larger masses this difference decreases.

At the low-mass end, roughly $f_s \propto M_h$ ($\propto M_s^{0.5}$) and $f_s \propto M_h^{3.0}$ ($\propto M_s^{0.8}$) for the blue and red samples, respectively, while at the high-mass end, $f_s \propto M_h^{-0.4}$ ($\propto M_s^{-0.7}$) and $f_s \propto M_h^{-0.6}$ ($\propto M_s^{-1.5}$), respectively.

It is important to note that the differences between blue and red M_s–M_h relations at almost all masses are within the 1σ uncertainty of our inferences. We conclude that the M_s–M_h (f_s–M_h) relation does not depend significantly on galaxy color (type). If any, the mean f_s–M_h relation of red galaxies is narrower and more peaked than the one of blue galaxies. In the mass range where the abundances of blue and red galaxies are closer ($10.0 < \log(M_s/M_\odot) < 10.7$), the intrinsic scatter around the M_s–M_h relation would slightly correlate with color in the sense that the redder (bluer) the galaxy, the larger (smaller) its M_s is for a fixed M_h, with a maximum average deviation from the mean due to color no larger than ~ 0.1 dex. For masses smaller than $M_s \approx 10^{9.7}\ M_\odot$, the correlation of the scatter with color would invert.

The (slight) differences between blue and red M_s–M_h (f_s–M_h) relations can be understood basically by the differences in the respective cumulative *GSMF*s and, at a minor level, by the differences of the corresponding *HMF*s for each case. The sharp peak in the red f_s–M_h relation is associated to the turn-over at $M_s \sim 10^{10.5}\ M_\odot$ in the *GSMF* of red galaxies (see Fig. 29).

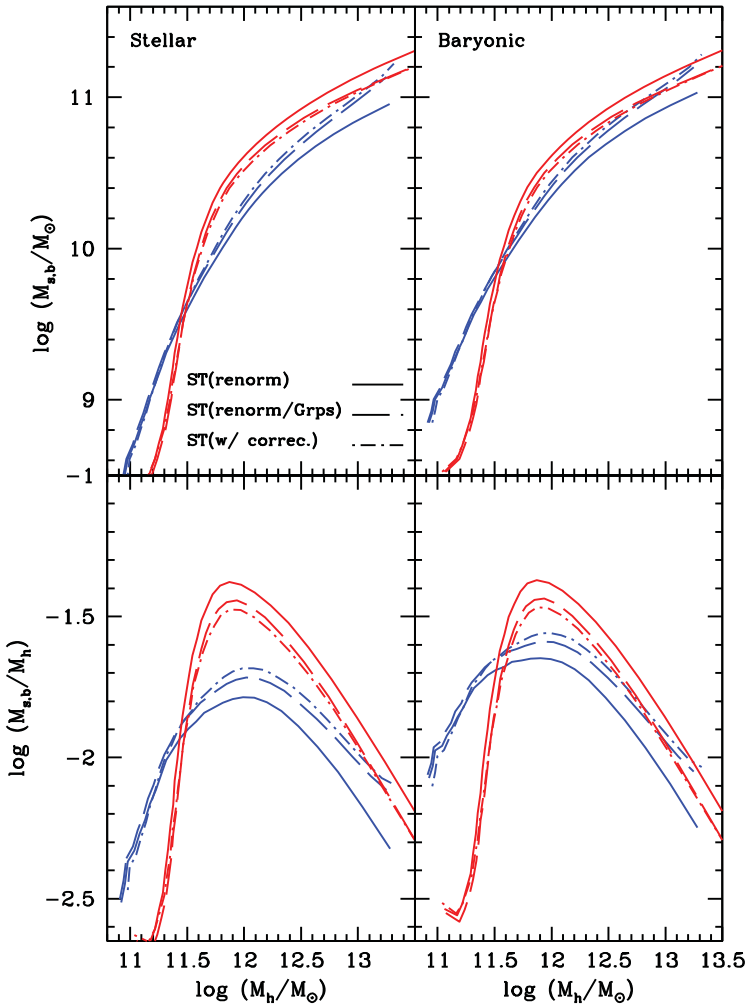

Figure 34: *Left panels:* Mean M_s–M_h (top) and f_s–M_h (bottom) relations of central blue (blue lines) and red (red lines) galaxies when (i) no systematical corrections to the corresponding "blue" and "red" *HMF*s were applied apart from re-normalisations in the global abundance (see text, solid lines), (ii) the *HMF*s were corrected by group/cluster abundances and re-normalised (long-dashed lines), and (iii) the *HMF*s were corrected both by group/cluster abundances and late major mergers (as in Fig. 32, dot-dashed lines). *Right panels:* Same as in left panels but for M_b instead of M_s.

Table 3: Fit parameters

Parameter	All	Blue	Red
$\log M_{0,h}$	11.97	11.99	11.87
$\log M_s^*$	10.40	10.30	10.40
β	0.34	0.37	0.18
α	1.45	0.90	1.50
γ	0.90	0.90	0.90
$a\ (M_s < M_s^*)$	0.000	0.125	0.000
$a\ (M_s > M_s^*)$	0.095	0.125	0.093

In order to estimate the influence of the corrections introduced to the *HMF* for blue (red) galaxies, we have redone the analysis by using the original distinct (S-T) *HMF* without any correction but re-normalised to obtain the same fraction of halos as the fraction implied by the *GSMF* of blue (red) galaxies with respect to the total *GSMF*. The results are shown in Fig. 34 with solid curves of blue color (blue galaxies) and red color (red galaxies). For comparison, the corresponding relations plotted in Fig. 33 are reproduced here (dot-dashed blue and red lines, respectively). One sees that the corrections to the *HMF* we have introduced for associating halos to the blue and red galaxy sub-samples act in the direction of reducing the differences among them in the M_s–M_h (f_s–M_h) relations, specially for larger masses. The group/cluster mass function correction to the *HMF* hosting central blue galaxies is the dominant one. The dashed blue and red curves show such a case, when only this correction (and a small re-normalisation) is applied.

7.3.2.1 *Analytical fits to the stellar–halo mass relations*

From the comparison of the *GSMF* and *HMF* it is easy to deduce that high- and low-mass galaxies have significantly different M_s–M_h scalings, a fact attributed to the different feedback/gas accretion mechanisms dominating in large and small systems (see e.g., Benson et al. [24]). The transition point between the low- and high-mass scalings defines a characteristic halo mass $M_{0,h}$ and an associated stellar mass M_s^*. Therefore, it was common to describe the M_s–M_h relation as a double-power law with the turnover point at $M_{0,h}$. However, BCW10 have argued recently that a power-law at the high-mass side is conceptually a bad description for the M_s–M_h relation and proposed a modification to it. Our results show indeed that a power-law is not enough to describe the high-mass side of the M_s–M_h relations.

We have found that a good analytical description to the overall, blue, and red mean M_s–M_h relations can be obtained for the inverse of the relations (M_h as a function of M_s, as in BCW10) by proposing a power-law dependence for low masses and a sub-

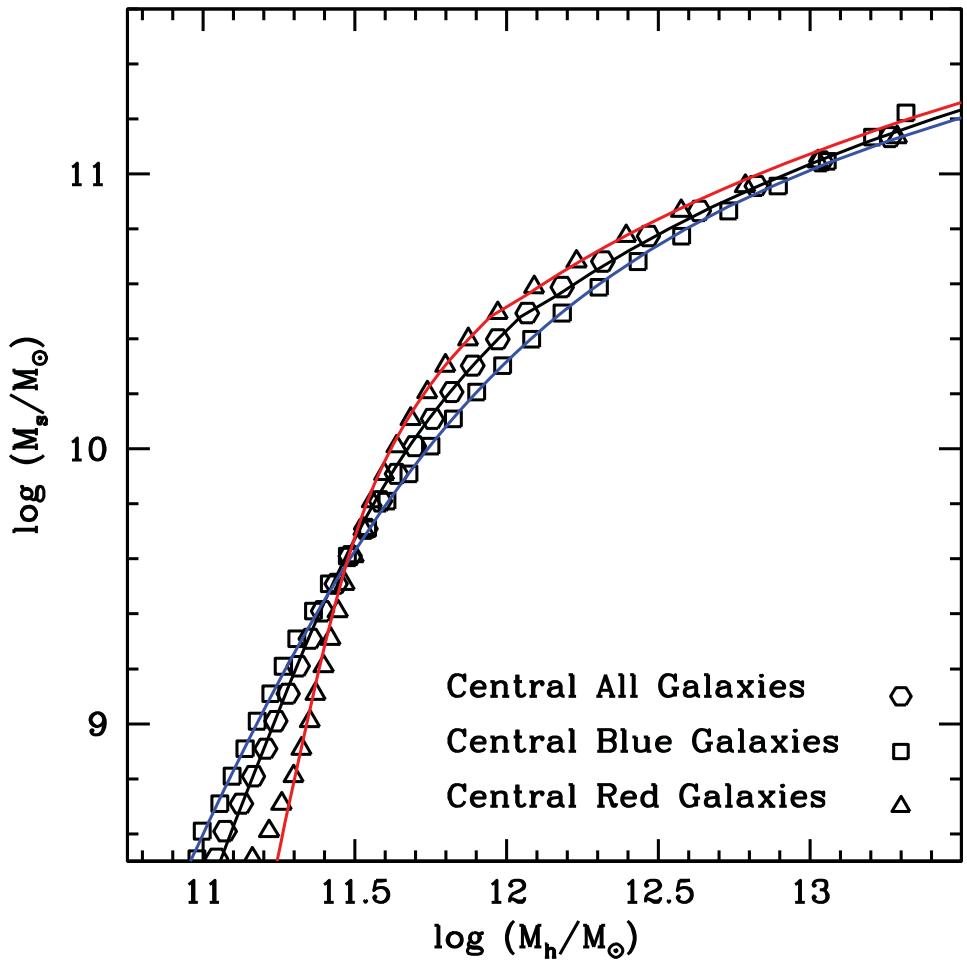

Figure 35: Analytical fits given by eq. (63) and Table 1 compared to the mean M_s–M_h relation for all central galaxies (black solid line) and the central blue (blue solid line) and red (red solid line) galaxy sub-samples.

exponential law for high masses (see BCW10). The functional form that fits well the three M_h–M_s relations is:

$$M_h = \frac{M_{0,h}}{2^\gamma}\left[\left(\frac{M_s}{M_s^*}\right)^{\beta/\gamma} + \left(\frac{M_s}{M_s^*}\right)^{\alpha/\gamma}\right]^\gamma 10^{a(M_s/M_s^*-1)} \quad (63)$$

where β regulates the behaviour of the relation at masses $M_s < M_s^*$, α together with the sub-exponential term ($a < 1$) regulate the behaviour at masses $M_s > M_s^*$, and γ regulates the transition of the relation around M_s^*. In Table 1 we present the values of all the parameters that best fit our results for the (central) overall, blue, and red M_s-M_h relations. Note that a assumes two different values depending on whether the mass is smaller or larger than M_s^*.

Figure 35 shows the three mean M_s–M_h relations obtained here and the functional form given in eq. (63) with the corresponding parameters reported in Table 1. The functional form is an excellent fit to the overall and blue M_s–M_h relations at all masses and to the red M_s–M_h relation for masses larger than $M_h \approx 10^{11.3}$ $M_\odot$.

7.3.3 *The baryonic-halo mass relations for central blue and red galaxies*

The right upper and lower panels of Fig. 33 show the mean M_b–M_h and f_b–M_h relations, for all central galaxies (solid line), central blue (long-dashed blue line), and central red (short-dashed red line) galaxies. The blue and red *GBMF*s were calculated from the corresponding *GSMF*s adding to M_s the respective gas mass, M_g (see **??**). The total *GBMF* is the sum of both of them. The error in M_b was calculated as the sum in quadratures of the errors in M_s and M_g. This error, together with the intrinsic scatter in M_s (see §§2.2), both propagated to the M_b–M_h relation, account for an uncertainty (standard deviation) of ~ 0.23 dex in $\log M_b$ at all masses (grey curves connected by vertical lines in Fig. 33).

The baryonic mass fraction, f_b, for blue galaxies is larger than the corresponding stellar one, f_s, in particular at smaller halo masses. At $M_h \approx 10^{11}$ $M_\odot$, f_b is a factor of 2.4 higher than f_s, while the peak of $f_b = 0.028^{+0.018}_{-0.011}$ (at $M_h = 10^{12.0}$ $M_\odot$) is only 1.3 times larger than the peak of f_s (at $M_h = 10^{12.0}$ $M_\odot$). For larger masses, the difference between f_b and f_s decreases, while for smaller masses, f_b is increasingly larger than f_s. For red galaxies, f_s and f_b are very similar, some differences being observed only at the lowest masses.

For masses larger (smaller) than $M_h \approx 10^{11.6}$ $M_\odot$, the differences between the M_b–M_h (f_b–M_h) relations of blue and red galaxies become smaller (larger) than in the case of stellar masses (left panels of Fig. 33). In general, the f_b bell-shaped curve for red galaxies is more peaked and narrower than the one for blue galaxies.

For blue galaxies, roughly $f_b \propto {M_h}^{0.7}$ (${M_b}^{0.4}$) at the low-mass end, and $f_b \propto {M_h}^{-0.5}$ (${M_b}^{-0.8}$) at the high-mass end. For red galaxies, roughly $f_b \propto {M_h}^{2.9}$ (${M_b}^{0.8}$) at the low-mass end, and $f_b \propto {M_h}^{-0.6}$ (${M_b}^{-1.5}$) at the high-mass end. For halos of masses $M_h \approx 10^{11.0}$ $M_\odot$ and $M_h \approx 10^{13.2}$ $M_\odot$, the baryon fraction for blue (red) galaxies decreases to values $f_b \approx 0.004$ and 0.0085 ($f_b \approx 0.0031$ and 0.0071), respectively. Therefore, for all masses, $f_b << f_U$, where $f_U \equiv \Omega_b/\Omega_M$ is the universal baryon mass fraction; for the cosmology used here, $f_U = 0.167$.

7.4 DISCUSSION

7.4.1 *Comparison with other works*

As discussed previously (see Fig. 32), our inference of the local overall M_s–M_h relation is in general in good agreement with several recent works that make use of the AMT [e.g., 16, 102, 148, 19]. The aim in this Chapter was to estimate the M_s–M_h and M_b–M_h relations for blue (late-type) and red (early-type) central galaxies separately. We have found that the differences between the means of the obtained relations for blue and red galaxies are within the 1σ uncertainty (see Fig. 33). In more detail, the mean stellar and baryonic mass fractions (f_s and f_b) as a function of M_h for red galaxies are narrower and more peaked than those for blue galaxies in such a way that in a given mass range (11.5–13.0 and 11.5–12.5 in $\log(M_h/M_\odot)$ for the stellar and baryonic cases, respectively) the former are higher than the latter and outside these ranges, the trend is inverted, specially at the low–mass side.

There have been only a few previous attempts to infer the halo masses of central galaxies as a function of mass (luminosity) *and* galaxy type [133, 146]. These works use direct techniques (see 7.1), which are, however, limited by low signal-to-noise ratios, specially for less massive systems in such a way that halo mass estimates are reliable only for galaxies with $M_s \gtrsim 10^{10}\ M_\odot$. These techniques are galaxy-galaxy weak lensing and kinematics of satellite galaxies around central galaxies. In order to overcome the issue of low signal-to-noise ratios, large samples of galaxies are stacked together in bins of similar properties (e.g., luminosity, M_s, galaxy type) to obtain higher signals of the corresponding measures (the tangential shear in the case of lensing and the weighted satellite velocity dispersion in the case of satellite kinematics). Moreover, estimates of M_h with these sophisticated techniques are subject to several assumptions, among them, those related to the internal halo mass distribution. It is usual to assume the Navarro, Frenk & White [151] density profile with the mean concentration for a given mass as measured in N-body cosmological simulations.

It is not easy to achieve a fair comparison of the results obtained with the AMT formalism and those with the direct methods. We have inferred the mean (and scatter) of $\log M_s$ as a function of M_h, while the weak lensing and satellite kinematics techniques constrain M_h as a function of M_s [see e.g., 146]; besides, the former calculates the mean of M_h (and its scatter) in a linear scale instead of a logarithmic one. These different ways of defining the relationship between stellar and halo masses, depending on the shapes and scatters of the corresponding relations, diverge more or less among them. In BCW10 (see their Fig. 10), it was shown that at low masses ($\log(M_h/M_\odot) \lesssim 12$, $\log(M_s/M_\odot) \lesssim 10.5$), averaging $\log M_s$ as a function of M_h or $\log M_h$ as a function of M_s give equivalent results for the AMT, but at high masses, where the M_s–M_h relation becomes much shallower, this relation becomes steeper (higher stellar mass at a fixed halo mass) for the latter case with respect to the former one.

In Fig. 36, the results from Mandelbaum et al. [133] are reproduced, left panels for central late-type galaxies and right panels for central early-type galaxies (solid squares with error bars). The error bars are 95 percent confidence intervals (statistical). Mandelbaum et al. [133] have used the (de Vaucouleours/exponential) bulge-to-total ratio,

frac_deV, given in the SDSS PHOTO pipeline as a criterion for late- (frac_deV<0.5) and early-type (frac_deV≥ 0.5) separation. This criterion of course is not the same as the color used in YMB09, but there is a correlation between both of them in such a way that a comparison between our results and those of Mandelbaum et al. [133] is possible at a qualitative level. Note that we have diminished the halo masses of Mandelbaum et al. [133] by $\approx 15\%$ to convert to our definition of halo virial mass. In more recent works, Mandelbaum, Seljak & Hirata [132] and Schulz, Mandelbaum & Padmanabhan [184] reported a new weak lensing analysis for the massive central early-type galaxies using the seventh SDSS data release (DR7) and a more sophisticated criteria for selecting the early-type lens population. Their results are plotted in the right panel of Fig. 36 with solid triangles and open squares, respectively.

In the case of the satellite kinematics estimates of $M_{\rm h}$ by More et al. [146], the same SDSS sample and similar recipes as in YMB09 for calculating $M_{\rm s}$, classifying galaxies into blue and red, and finding central and satellites galaxies were used. More et al. [146] applied their analysis to constrain the mean $\log M_{\rm h}$ as a function of $M_{\rm s}$, but also present the constraints of their model for the mean of $\log M_{\rm s}$ as a function of $M_{\rm h}$. Their results for the latter case, kindly made available to us in electronic form by Dr. S. More, are reproduced in Fig. 36 as the shaded (orange) regions which represent the 68% confidence intervals. Transforming from their definitions to ours of halo mass and IMF, their $M_{\rm h}$ and $M_{\rm s}$ were diminished by $\approx 15\%$ and $\approx 25\%$, respectively. The dotted horizontal lines in each panel show the approximate range in $M_{\rm s}$, where the estimates are reliable according to More et al. [146, see their Fig. 11].

More et al. [146] also reported results for the average $M_{\rm h}$ as a function of $M_{\rm s}$ split in central blue and red galaxies according to the galaxy group analysis by Yang et al. [249]. The solid (cyan) curves in Fig. 36 reproduce these results.

Finally, the standard $\pm 1\sigma$ deviation intervals that we have obtained from the AMT are reproduced in Fig. 36 for central blue and red galaxies (solid blue and red curves connected by vertical lines, respectively). Note that in the estimates with direct methods, the systematic uncertainty in $M_{\rm s}$, which is the main source of error in the AMT, was not taken into account.

Our inference for early-type (red) galaxies is consistent (within the uncertainties, errors, and different ways of presenting the constraints) with the weak lensing results of Mandelbaum et al. [133], Mandelbaum, Seljak & Hirata [132] and Schulz, Mandelbaum & Padmanabhan [184], and with the galaxy group analysis of Yang et al. [249] as reported in More et al. [146]. With respect to the satellite kinematics analysis by More et al. [146], their mean halo masses for $M_{\rm s} \sim 5\times 10^9 - 10^{11}$ $M_\odot$, are larger than ours (and those of Mandelbaum et al. [133] by factors around 2). For larger masses, all determinations agree roughly with our results. In fact, there is some indication that satellite kinematics yields halo masses around low mass central galaxies that are systematically larger than most other methods, specially for red central galaxies Skibba et al. [194, but see More et al. [146] for a discussion].

For late-type (blue) galaxies, our results are in reasonable agreement with those of Mandelbaum et al. [133] for masses $M_{\rm s} \lesssim 10^{10.8}$ $M_\odot$. At higher masses, their results imply halo masses for a given $M_{\rm s}$ smaller than ours, with the difference increasing towards higher stellar mass. The discrepancy would be weaker if one considers that the

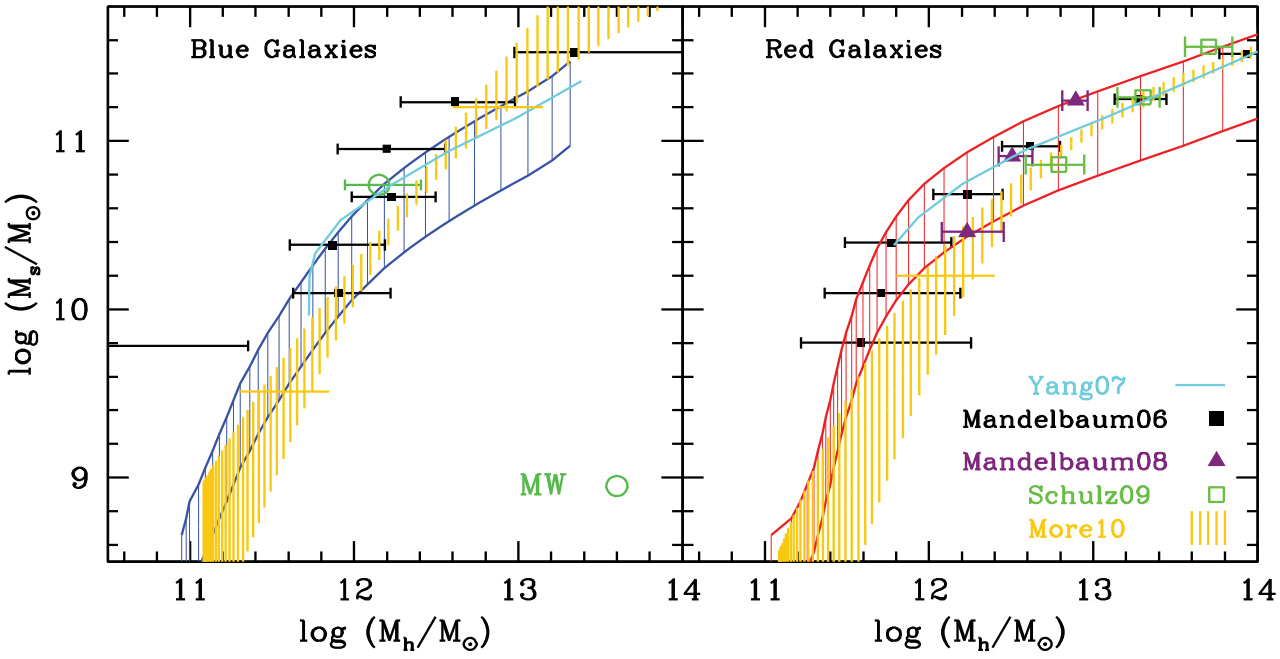

Figure 36: Comparison with other observational inferences. *Left panel*: $M_s - M_h$ relation for blue (late-type) galaxies. The blue curves connected by vertical lines encompass the $\pm 1\sigma$ interval inferred here. We also reproduce the inferences using galaxy-galaxy weak lensing by Mandelbaum et al. [133, black squares], galaxy groups Yang et al. [249, cyan solid line], and satellite kinematics More et al. [146, orange vertical lines]. Estimates for the Milky Way are shown (open circle with error bar). *Right panel*: $M_s - M_h$ relation for red (early-type) galaxies. The red curves connected by vertical lines encompass the $\pm 1\sigma$ interval inferred here. Other determinations as in the left panel but for early-type galaxies are shown. More recent inferences with the weak lensing technique by Mandelbaum, Seljak & Hirata [132, filled violet triangles] and by Schulz, Mandelbaum & Padmanabhan [184, open green squares] are also plotted.

mean M_s–M_h relation in our case becomes steeper when calculating M_h as a function of M_s. On the other hand, it must be said that the number statistics becomes poor for massive late-type galaxies, resulting in a stacked weak lensing analysis with large error bars. For example, in the two most massive bins in the Mandelbaum et al. [133] sample (the two uppermost points in Fig. 36), only 5 and 11 percent of the galaxies are classified as late types. Future weak lensing studies should confirm whether high-mass late-type galaxies have such relatively small halos as found in Mandelbaum et al. [133]. Regarding the comparison with the satellite kinematics inferences of More et al. [146], the agreement is reasonable at least up to $M_s \approx 10^{11}$ $M_\odot$, although the relation inferred by these authors is shallower than ours. For larger masses, these authors caution that their results become very uncertain, as in the weak lensing case, because of poor statistics of massive blue galaxies. The galaxy groups [249] inference, in the mass range allowed by this technique, gives halo masses slightly smaller than the means of our inference for a given M_s.

In general, most techniques for inferring the relationship between stellar and halo masses of galaxies agree among them within factors up to 2–3 in M_h [19, 146, 78]. This seems to be also the case for samples partitioned into late- and early-type galaxies, as shown here. However, beyond the detailed comparison between our results and those obtained with direct techniques, it seems that there is a systematic qualitative difference: in our case, at a given halo mass (for $10^{11.5}$ $M_\odot \lesssim M_h \lesssim 10^{13.0}$ $M_\odot$), blue centrals, on average, have lower stellar masses than red centrals, while in the case of determinations with direct techniques, the opposite applies at least for masses larger than $M_h \sim 10^{12}$ $M_\odot$ [133, 146, see also Figs. 33 and 36].

A partial source of bias contributing to this difference could be that in the weak lensing and satellite kinematics techniques the same concentration for halos hosting late- and early-type galaxies is assumed. If halos of late- (early-)type galaxies are less (more) concentrated than the corresponding average, then for the same measure (shear or satellite velocity dispersion), the halo masses are expected to be higher (lower) than the obtained ones. Therefore, the differences found [133, 146] in the mass halos of late- and early-type galaxies of a given M_s would decrease or even invert their sing.

While it is difficult to make any robust statement about possible systematics in a given technique regarding late and early types, we ask ourselves what should be modified in our assumptions in order to invert the behaviour of the M_s–M_h relations with galaxy type (color) obtained here. We have shown in Fig. 34 that our corrections to the *HMF* had the effect of making closer the M_s–M_h relations of blue and red galaxies at large masses. One possibility to invert the relation this is to make even steeper (shallower) the *HMF* corresponding to blue (red) galaxies, mainly at the high-mass end (see Fig. 31, lower panel). This would imply, for instance, a higher correction to the *HMF* due to groups than that made by us. The group/cluster mass function we used (Heinämäki et al. [105]) is one of the most general ones found in the literature; it includes all kinds of groups/clusters with 3 or more members and $\delta N/N \geq 80$. The authors note that their sample is complete down to a dynamical mass roughly equivalent to $M_h = 5 \times 10^{13}$ $M_\odot$. It could be that the abundance of groups of lower masses is larger than that given in Heinämäki et al. [105], although it is difficult to accept that blue galaxies are completely absent in the centers of small and loose groups of a few (> 2) members.

Last but not least, in Fig. 36 we include observational estimates for our Galaxy (open circle). The uncertainties in the estimates of M_h for the Milky Way are still large but better than most of the estimates for other individual galaxies. For recent reviews on different results see Guo et al. [102] and Dutton et al. [78]. In Fig. 36 we plot a recent estimate of M_h based on observations of 16 high velocity stars (Smith et al. [195]). These authors find $M_h = 1.42^{+1.14}_{-0.54} \times 10^{12} M_\odot$, which is in good agreement with several previous works [e.g., 240, 180, 125], though results from Xue et al. [243] suggest lower values (but see a recent revision by Przybilla et al. [168]). For its M_h, the M_s of Milky Way seems to be in the extreme of blue galaxies, close to values typical of red galaxies. It should be said that it is an open question whether of the Milky Way is an average galaxy or not. In the stellar Tully-Fisher and radius–M_s relations (e.g., Avila-Reese et al. [12]), the Milky Way is shifted from the average to the high-velocity and low-radius sides, respectively.

7.4.2 *Interpretations and consistency of the results*

Although our main result is that the differences between the M_s–M_h and M_b–M_h relations for central blue and red galaxies are marginal (within the uncertainties of our estimates), we will explore whether such differences are expected or not. For this it is important to approach the problem from an evolutionary point of view.

In Firmani & Avila-Reese [84, hereafter FA10], the estimates of the M_s–M_h relation for all galaxies at different redshifts, up to $z = 4$ (BCW10), and the average ΛCDM individual halo mass aggregation histories (MAHs) were used to determine the individual *average* M_s growth of galaxies in general as a function of mass (Galaxian Hybrid Evolutionary Tracks, GHETs). It was found that the more massive the galaxies, the earlier they transit from the active (star-forming, blue) regime to a passive (red) phase (population 'downsizing'), while their corresponding halos continue growing, more efficiently at later epochs as more massive they are ('upsizing'). The inferred trend for the transition stellar mass is $\log(M_{tran}/M_\odot) \approx 10.30 + 0.55z$. Therefore, galaxies of mass $M_s \approx 10^{10.3}$ $M_\odot$ are becoming passive (red) today. For $M_s \gtrsim M_{tran}$, the larger the mass, the redder on average will be the galaxy. The opposite applies for $M_s \lesssim M_{tran}$, the smaller the mass, the bluer will be the galaxy. Interestingly, $M_s \approx 10^{10.3}$ $M_\odot$ is roughly the mass where the overall YMB09 blue and red *GSMF*s cross: for masses larger than this crossing mass, M_{cross}, redder galaxies become more and more abundant than bluer ones and the inverse happens at smaller masses (see Fig. 29).

Galaxies that are transiting from active to passive at $z \sim 0$ (those around $M_{tran} \approx 10^{10.3}$ $M_\odot$) have probably been subjected recently to a process that induced an *efficient* transformation of the available gas into stars in such a way that their stellar populations started to redden passively. Hence, for a given M_h, they are expected to have a higher M_s (or f_s) than those galaxies of similar mass that did not suffer (yet?) the process mentioned above (bluer ones). The relatively small difference in f_s for blue and red galaxies we have found here (whose maximum is attained around $M_{tran} \sim M_{cross}$, Fig. 33) would imply that the scatter around M_{tran} is moderate.

Galaxies more massive than M_{tran} (or M_{cross}), according to the evolutionary analysis by FA10, the process of efficient gas consumption into stars (and the further cessation of M_s growth) earlie as more massive is the galaxy, while their halos continue growing.

Therefore, one expects that the more massive the galaxy, the redder and the lower its stellar (and baryonic) mass fraction f_s will be on average. The few blue massive galaxies may have slightly smaller stellar masses (lower f_s) than the corresponding red ones because they should have transformed gas into stars less efficiently in the past. Therefore, by including gas, i.e. when passing to f_b the difference between blue and red massive galaxies at large masses should become negligible. This is what indeed happens (see Fig. 33).

Galaxies less massive than M_{tran} (or M_{cross}) at $z \sim 0$, according to FA10, are in general more actively assembling their stellar masses as smaller they are ('downsizing in specific SFR), while their dark halo mass growth is already very slow. This implies the existence of relatively larger reservoirs of cold gas in low mass galaxies (gas not related to the halo-driven infall) because the SF has been delayed in the disk and/or cold gas is being lately (re)accreted into the galaxy. However, if for some reason the gas reservoir in these galaxies is lost, then the galaxy will redden and its baryonic and stellar mass fractions will be smaller than of the galaxies that were able to keep their gas reservoir (the majority), in agreement with our inferences here (Fig. 33).

7.5 SUMMARY AND CONCLUSIONS

By means of the AM technique and using the central blue and red *GSMF*s, constructed from the local SDSS sample by YMB09, we have inferred the local M_s–M_h (or f_s–M_h) relations for *central* galaxies and for the sub-samples of blue and red galaxies. To derive the relations for the sample of blue galaxies, (i) the mass function of observed groups/-clusters of galaxies is subtracted from the distinct (S-T) *HMF* (blue, late-type galaxies are not observed in the centers of groups and clusters), and (ii) halos that suffered a major merger since $z = 0.8$ are excluded. For red galaxies, the *HMF* is assumed to be the complement of the "blue" one, with respect to the overall (distinct) *HMF*. We only consider as sources of uncertainty the systematical error in assigning stellar masses to galaxies (0.25 dex) and the intrinsic statistical scatter in stellar mass at a fixed halo mass (0.16 dex). By using the observational M_g–M_s relation and its scatter, we transited from M_s to M_b (=M_s+ M_g) in the *GSMF* and estimated the overall blue and red *GBMF*s, which are used to obtain the corresponding baryonic M_b–M_h (or f_b–M_h) relations using the AM technique.

The M_s–M_h relation obtained here agrees rather well with previous studies (see Fig. 32). The small differences found in this work can be explained mainly in terms of the different *GSMF*s used in each study, and to a less extent by variations in the methodology. The 1σ uncertainty in the M_s–M_h relation is ≈ 0.25 dex in $\log M_s$. The M_s–M_h relation of central galaxies lies below (lower M_s for a given M_h) the overall one by a factor ~ 1.6 at $M_h = 10^{11}$ $M_\odot$ and less than 5% for $M_h > 10^{13}$ $M_\odot$.

Our main result refers to the calculation of the *central* M_s–M_h and M_b–M_h relations for the two broad populations into which the galaxy sample can be divided: blue (late-type) and red (early-type) galaxies. We highlight the following results from our analysis:

- At $M_h \gtrsim 10^{11.3}$ $M_\odot$ the mean stellar mass fraction f_s of blue galaxies is lower than the one of red galaxies, the maximum difference being attained at $M_h \approx 10^{11.7}$ $M_\odot$; at this mass, the f_s of red galaxies is 1.7 times the one of blue galaxies (see Fig. 33). At

larger masses, the difference decreases until it disappears. At $M_{\rm h} \lesssim 10^{11.3}$ M$_\odot$ the trend is reversed as blue galaxies tend to have higher values of $f_{\rm s}$ than red ones. In the case of the baryonic mass fractions, $f_{\rm b}$, the same trends remain but at $M_{\rm h} \gtrsim 10^{11.3}$ M$_\odot$ the difference in $f_{\rm b}$ between blue and red galaxies is small, while for smaller masses, the difference increases.

• The $M_{\rm s}$–$M_{\rm h}$ and $M_{\rm b}$–$M_{\rm h}$ (or $f_{\rm s}$–$M_{\rm h}$ and $f_{\rm b}$–$M_{\rm h}$) relations of central blue and red sub-samples do not differ significantly from the respective relations of the overall central sample, and these differences are within the 1σ uncertainty (Fig. 33). For blue (red) galaxies, the maximum value of $f_{\rm s}$ is $0.021^{+0.016}_{-0.009}$ ($0.034^{+0.026}_{-0.015}$) and is attained in halos of mass $M_{\rm h} = 10^{11.98}$ M$_\odot$ ($M_{\rm h} = 10^{11.87}$ M$_\odot$); the corresponding stellar mass is $M_{\rm s} = 10^{10.30\pm0.25}$ M$_\odot$ ($M_{\rm s} = 10^{10.40\pm0.25}$ M$_\odot$), which is around 0.23 (0.30) times $M^\star$, the Schechter fit characteristic mass of the overall *GSMF* of YMB09. For smaller and larger masses, $f_{\rm s}$ significantly decreases.

• We have compared our results with the few observational inferences of the $M_{\rm s}$–$M_{\rm h}$ relation for blue (late-type) and red (early-type) galaxies that exist in the literature. Although these studies estimate halo masses using direct techniques (weak lensing and galaxy satellite kinematics), they are still limited by the stacking approach they need to apply (due to the low signal-to-noise ratio of individual galaxies) and by the large uncertainties owing to unknown systematics. The overall differences among the different studies (including ours) amount up to factors 2-4 at a given mass (these factors being much smaller at other masses) for most methods (Fig. 36). For blue galaxies, all methods agree reasonably well for low masses ($M_{\rm h} \lesssim 3\times10^{12}$ M$_\odot$), but at higher masses, our inference implies larger halos for a given $M_{\rm s}$. For red galaxies, at high masses ($M_{\rm h} \gtrsim 3\times10^{12}$ M$_\odot$), all methods agree reasonably well, but at lower masses, the satellite kinematics technique produces halo masses, for a given $M_{\rm s}$, larger than those obtained by other methods.

• According to our results, for $M_{\rm h} \lesssim 10^{11.3}$ M$_\odot$, the intrinsic scatter of the $M_{\rm s}$–$M_{\rm h}$ relation should slightly anti-correlate with galaxy color (for a fixed $M_{\rm h}$, the bluer the galaxy, the higher its $M_{\rm s}$), while for more massive systems, the correlation should be direct (for a fixed $M_{\rm h}$, the redder the galaxy, the higher its $M_{\rm s}$). For massive blue galaxies in order to have higher $f_{\rm s}$ values than the red ones, as the results from direct techniques suggest, the *HMF* halos hosting blue (red) galaxies should be even steeper (shallower) than what we have proposed here; this seems unlikely.

• The maximum baryon mass fraction of blue and red galaxies are $f_{\rm b} = 0.028^{+0.018}_{-0.011}$ and $f_{\rm b} = 0.034^{+0.025}_{-0.014}$, respectively, much smaller than $f_U = 0.167$ in both cases, and these maxima are attained at $M_{\rm h} \approx 10^{12}$ M$_\odot$. At large masses $f_{\rm b}$ decreases approximately as $f_{\rm b} \propto {M_{\rm h}}^{-0.5}({M_{\rm b}}^{-0.8})$ for blue galaxies and as $f_{\rm b} \propto {M_{\rm h}}^{-0.6}({M_{\rm b}}^{-1.5})$ for red galaxies, in such a way that from $M_{\rm h} \approx 5\times10^{12}$ M$_\odot$, blue galaxies have on average slightly larger values of $f_{\rm b}$ than red ones. At low masses, the $f_{\rm b}$ of red galaxies strongly decreases as the mass is smaller $f_{\rm b} \propto {M_{\rm h}}^{2.9}({M_{\rm b}}^{0.8})$, while for blue galaxies, due to the increasing gas fractions towards smaller masses, $f_{\rm b}$ decreases slower than $f_{\rm s}$, as $f_{\rm b} \propto {M_{\rm h}}^{0.7}$ ($\propto {M_{\rm b}}^{0.4}$).

The AM technique has been revealed as a relatively simple but powerful method for connecting empirically galaxies to dark halos. Here we extended this technique towards inferences for the blue and red galaxy sub-populations separately. By introducing a minimum of assumptions –otherwise the method becomes to similar to a semi-analytical

model– we have found that the stellar and baryon mass–halo mass relations of blue and red galaxies do not differ significantly among them and from the overall ones. The maximum differences are around the peak of these relations, $M_{\rm h} \approx 10^{12}$ $M_{\odot}$, and are consistent qualitatively with the inference that the galaxies in these halos are transiting from an active to a quiescent regime (FA10). Those that transited recently is because they had an efficient process of gas consumption into stars and further cessation of $M_{\rm s}$ growth; therefore, they should be redder and with higher $f_{\rm s}$ values than those that still did not transit. For larger and lower masses than $M_{\rm h} \approx 10^{12}$ $M_{\odot}$, the differences decrease and even invert, something that is also consistent with the inferences by FA10, based on semi-empirical estimates of the evolution of the overall $M_{\rm s}$–$M_{\rm h}$ relation.

ACKNOWLEDGMENTS

We thank the Referee for his/her useful comments and suggestions. We are grateful to Dr. S. More for sending us in electronic form their data plotted in Fig. 36. A.R-P. acknowledges a graduate student fellowship provided by CONACyT. We thank PAPIIT-UNAM grant IN114509 and CONACyT grant 60354 for partial funding, as well as a bilateral DFG-CONACyT grant through which we got access to results from N-body numerical simulations performed by Dr. S. Gottloeber.

8

ARE THE STELLAR-TO-HALO MASS RELATIONS OF BLUE AND RED GALAXIES CONSISTENT WITH THEIR SPATIAL CLUSTERING?

Using a statistical model that combines the abundance matching technique, the halo occupation distribution (HOD) model, and the conditional stellar mass function formalism, we infer the stellar-to-halo mass relations, SHMR, of blue and red central galaxies, the fraction of halos hosting blue central galaxies, and the HOD statistics of blue and red satellite galaxies as a function of halo mass, $M_{\rm h}$. We use as input the observed color distributions as a function of stellar mass M_*, blue/red galaxy stellar mass functions, and projected two point correlation functions decomposed into blue/red galaxies. The fraction of halos hosting blue central galaxies decreases with $M_{\rm h}$ and it has a maximum value of $f_{\rm blue} \sim 0.8$ attained at $M_{\rm h} \approx 10^{11.4}{\rm M}_\odot$. We find that this result agrees at a qualitatively level with the proposal in Chapter 7 that blue galaxies avoid halos that suffered major mergers or that form rich group/cluster systems. The SHMR of central blue and red galaxies do not differ significantly between them. This implies that the efficiency of galaxy stellar mass growth, $f_* = M_*/M_{\rm h}$, is nearly independent of the central galaxy color, although it strongly changes with $M_{\rm h}$. The almost lack of segregation of the M_*–$M_{\rm h}$ relation of central galaxies by color points out to a large dominion of the scale of the halos over other properties, e.g., its concentration, connected to the efficiency of galaxy stellar mass growth.

8.1 INTRODUCTION

So far, most of current efforts for connecting galaxies to halos have focused on the galaxy population as a whole, with the stellar mass, M_*, as the only galaxy property to link with the cold dark matter (CDM) halos (see Part II). However, it is well know that mass is not the only key parameter of galaxies, as in the case of stars. From the most general point of view, galaxies can be accommodated into two very broad classes [see for a review, 32]: red objects dominated by a pressure-suported component with typically old stellar populations and inefficient star formation (early-type galaxies); and blue objects dominated by a rotationally-supported disk with old and young stellar populations and active star formation (late-type galaxies). Late- and early-type galaxies can have any mass, though statistically the former tend to be smaller and the latter larger. Among the intensive (scale independent) properties that characterize these two galaxy sequences, the most easy to obtain in large surveys is the color index. Therefore, *a next step in the galaxy-halo connection is to establish it for at least galaxies separated into two groups, blue and red ones*. Constraining the role that color plays as a second parameter in the galaxy-halo connection sheds certainly light on our understanding of galaxy evolution.

The distribution of galaxies as a function of color is bimodal [e.g., 15]. The shape of the bimodality changes systematically with luminosity or M_*: the blue peak decreases in

favor of the red one as the luminosity or M_* increases. The bimodal distribution and its dependence on M_* is quantitatively different for central and satellite galaxies. In a more general context, the denser the environment, the redder the galaxies are on average, and the red galaxies are systematically more clustered spatially than the blue ones [see, e.g., 123, 258]

In the previous Chapter [177], we have made a first attempt for inferring the stellar-to-halo mass relations (SHMRs) for blue and red galaxies separately by using the abundance matching technique (AMT). Motivated by theoretical and observational facts, the CDM distinct halo mass function (HMF) used as a match to the galaxy stellar mass function (GSMF) of central blue galaxies, was corrected to correspond to halos hosting blue galaxies by *assuming* that (1) cluster- and group-sized halos are less likely to host blue central galaxies, and (2) blue central galaxies are not expected to live in halos that suffered major mergers since $z = 0.8$. The corresponding HMF of red central galaxies, used to abundance match the central red GSMF, is the complement to the total distinct HMF. Under the such assumptions, the SHMRs of blue and red centrals were obtained. The corresponding halo masses of both groups agree roughly with direct inferences of halo masses of blue and red central galaxies at the masses where the comparison is possible [e.g., 146].

The next important question to answer is; *whether the obtained SHMRs provide the correct link of blue and red galaxies to the CDM halos, in the sense that the spatial clustering of linked blue and red galaxies is consistent with the corresponding observed spatial clustering.*

In this Chapter, instead of assuming corrections to the HMF to assign host halos to blue and red galaxies, we will introduce an empirical approach that makes use of the observed color distributions as a function of M_* and the observed projected correlation functions of blue and red galaxies in different M_* bins. In this way, by construction, the obtained blue and red SHMRs are consistent with the observed color distributions and spatial clustering of galaxies. The main assumption behind our approach is that the color of a galaxy depends on the halo mass *only* through the empirical dependence of color on M_*. The latter dependence, as already mentioned, is actually weak in the sense that for a given M_*, galaxies may have a broad (bimodal) distribution of color. Thus, with our semi-empirical approach, we will be able to constrain the fractions of halos at each halo mass that are linked to blue and red galaxies, the fractions of blue and red satellites occupying halos of a given mass, and the SHMRs of blue and red galaxies.

The fraction f_s=M_*/M_h gives information about the efficiency of stellar mass growth inside halos. The semi-empirically inferred strong correlation of f_s on M_h, with a low scatter (see Part II), indicates that this efficiency is determined mainly by the halo mass. We would like to know if this efficiency is different for today blue and red galaxies, i.e., whether the f_s–M_h (or M_*–M_h) relation segregates significantly by color or not.

This Chapter is organized as follows. In Section 8.2, we describe the approach for inferring the fraction of halos from the distinct halo mass functions that are associated to blue galaxies, as well as the our statistical model for constraining the blue/red SHMRs. In Section 8.4, we present our results. Section 8.5 is devoted to a discussion of the implications of some of our results. Section 8.6 presents the main conclusions of the work.

8.2 THE MODEL

In this Section we describe the semi-empirical approach used for assigning blue and red central galaxies to their dark matter halos, as well as the satellite occupational numbers of blue and red satellite galaxies. The statistical nature of this model allows us to relate in a self-consistent way the blue and red GSMF decomposed into centrals and satellites galaxies, the projected two-point correlation function (2PCF) of blue and red galaxies and the ΛCDM halo/subhalo mass functions. The model outputs are: the mean stellar-to-halo mass relations (SHMR) of blue and red central galaxies, the fraction of halos hosting blue central galaxies and the fraction of blue satellite galaxies as a function of halo mass. Our main goal here is to constrain these set of model outputs by using observations of the color distributions as a function of stellar mass, the blue and red GSMFs, and the projected 2PCF decomposed into blue and red galaxies.

The statistical approach developed here is based on the AMT, HOD model, and CSMF formalism (see Part II) but in order to include the blue and red galaxy populations it requires of two additions.

1. For connecting blue and red central galaxies to their dark matter halos (§8.2.1), we introduce the conditional probability distribution function that a distinct halo of mass $M_{\rm h}$ host a blue (red) central galaxy of stellar mass M_*, denoted by $P_{\rm B,c}(M_*|M_{\rm h})$ ($P_{\rm R,c}(M_*|M_{\rm h})$). As a result, these distributions contain information about the mean SHMR of blue and red central galaxies, $M_{*,\rm B}(M_{\rm h})$ and $M_{*,\rm R}(M_{\rm h})$, respectively. However, as was outlined in Chapter 6, in order to infer both $M_{*,\rm B}(M_{\rm h})$ and $M_{*,\rm R}(M_{\rm h})$, the unknown fraction of halos hosting blue and red central galaxies is a required input of the model. We overcome this issue by assuming that the color of a central galaxy mainly depends on the mass of the galaxy rather than on the halo mass. Based on this assumption and having previously constrained the global central stellar-to-halo mass relation and its scatter ($P_{\rm c}(M_*|M_{\rm h})$; see §§8.2.1), it is possible to obtain a simple empirical model for the fraction of halos hosting blue central galaxies.

2. In agreement with studies of galaxy groups and clusters, and those based on galaxy clustering, we introduce a model for the fractions of blue/red satellite galaxies as a function of halo mass. These fractions are necessary to use the observed blue/red correlation functions as constrains to the model.

8.2.1 *The fractions of halos hosting blue/red galaxies as a function of mass*

In the context of the AMT, the connection between the central GSMF, $\phi_{\rm g,c}(M_*)$, and the distinct halo mass function, $\phi_{\rm h}(M_{\rm h})$, arises naturally by assuming a probability distribution function, denoted by $P_{\rm c}(M_*|M_{\rm h})$, that a distinct halo of mass $M_{\rm h}$ hosts a central galaxy of stellar mass M_* [216, 115, 58, 187, 234, 16, 217, 57, 75, 148, 19, 102, 20, 170, 159, see also Chapters 2 and 3]. As a result of this connection, one obtains information about the mean SHMR of central galaxies, $M_*(M_{\rm h})$. Similarly, in the case that the total central GSMF is expressed as the sum of the blue and the red central GSMFs, one can simply define the conditional probability distribution functions $P_{\rm B,c}(M_*|M_{\rm h})$ and $P_{\rm R,c}(M_*|M_{\rm h})$

(see above) to establish the connection between the "blue" and "red" halo mass functions and the GSMFs of blue and red centrals. Therefore, the GSMF of blue and red centrals are given by:

$$\phi_{\rm B,c}(M_*) = \int P_{\rm B,c}(M_*|M_{\rm h})\phi_{\rm h,B}(M_{\rm h})dM_{\rm h}, \tag{64}$$

and

$$\phi_{\rm R,c}(M_*) = \int P_{\rm R,c}(M_*|M_{\rm h})\phi_{\rm h,R}(M_{\rm h})dM_{\rm h}, \tag{65}$$

respectively. Here $\phi_{\rm h,B}$ and $\phi_{\rm h,R}$ are the mass functions of dark matter halos hosting blue and red central galaxies, respectively. These equations describe the abundance matching formalism in its differential form for the blue central/halo and the red central/halo populations.

Formally, the relation between $P_{\rm c}(M_*|M_{\rm h})$ and the distribution functions $P_{\rm B,c}(M_*|M_{\rm h})$ and $P_{\rm R,c}(M_*|M_{\rm h})$ is given by

$$P_{\rm c}(M_*|M_{\rm h}) = P_{\rm B,c}(M_*|M_{\rm h})f_{\rm B}(M_{\rm h}) + P_{\rm R,c}(M_*|M_{\rm h})f_{\rm R}(M_{\rm h}). \tag{66}$$

where $f_{\rm B}(M_{\rm h})$ denotes the fraction of halos hosting blue centrals, and $f_{\rm R}(M_{\rm h}) = 1 - f_{\rm B}(M_{\rm h})$ denotes the fraction of halos hosting red centrals. Note that $f_{\rm B} = \phi_{\rm h,B}/\phi_{\rm h}$ and $f_{\rm R} = \phi_{\rm h,R}/\phi_{\rm h}$.

Ideally, when one knows either $f_{\rm B}(M_{\rm h})$ or $f_{\rm R}(M_{\rm h})$[1] the problem of constraining the parameters in the distribution functions $P_{\rm B,c}(M_*|M_{\rm h})$ and $P_{\rm R,c}(M_*|M_{\rm h})$ is relatively easy by solving Eqs. 64 and 65[2]. However, $f_{\rm B}(M_{\rm h})$ is actually an unknown function and one needs to *assume* a model for $f_{\rm B}(M_{\rm h})$. In what follows, we describe a simple method to estimate $f_{\rm B}$ empirically under the assumption that the color of a central galaxy is not directly related to the halo mass but to the stellar mass. Note that we are not claiming that there is not a correlation between the color of the central galaxy and halo mass. In the context of our model, this correlation exists but through the correlation of the stellar mass with the halo mass. In what follows, we describe the empirical model used to estimate $f_{\rm B}$.

In order to assign correctly colors to central galaxies in our approach, we must connect statistically all the quantities of interest, i.e., stellar mass, color, and halo mass. For that reason, we have introduced the distribution function $P(C_{\rm gr}, M_*|M_{\rm h})$, denoting the probability that a distinct halo of mass $M_{\rm h}$ hosts a central galaxy of stellar mass M_* and color $C_{\rm gr}$. This distribution function can be decomposed into a "red sequence", denoted by the function R, and a "blue cloud", denoted by the function B;

$$P(C_{\rm gr}, M_*|M_{\rm h}) = B(C_{\rm gr}, M_*|M_{\rm h})f_{\rm B}(M_{\rm h}) + R(C_{\rm gr}, M_*|M_{\rm h})f_{\rm R}(M_{\rm h}), \tag{67}$$

Notice that marginalizing over $C_{\rm gr}$, Eq. (66) is recovered. For the moment, let us focus on the right hand side of Eq. (67). Applying the Bayes' Theorem, we can write:

$$B(C_{\rm gr}, M_*|M_{\rm h}) = \mathcal{B}(C_{\rm gr}|M_*, M_{\rm h})P_{\rm B,c}(M_*|M_{\rm h}), \tag{68}$$

1 Because the fractions $f_{\rm R}(M_{\rm h})$ and $f_{\rm B}(M_{\rm h})$ are mutually exclusive, hereafter we only refer to $f_{\rm B}(M_{\rm h})$.

2 Note that in Chapter 7 we used these equations to constrain the mean relations $M_{*,\rm B}(M_{\rm h})$ and $M_{*,\rm R}(M_{\rm h})$, provided that $\phi_{\rm h,B}(M_{\rm h})$ was obtained by theoretical and observational facts about disk galaxies. In this Chapter we again use these equation in order to constraint both $M_{*,\rm B}(M_{\rm h})$ and $M_{*,\rm R}(M_{\rm h})$, but this time by proposing an empirical model for $\phi_{\rm h,B}(M_{\rm h})$, and hence, for $f_{\rm B}$.

and

$$R(C_{\rm gr}, M_*|M_{\rm h}) = \mathcal{R}(C_{\rm gr}|M_*, M_{\rm h})P_{\rm R,c}(M_*|M_{\rm h}). \tag{69}$$

Under the assumption that the color distribution of central galaxies at a fixed stellar mass is independent of halo mass, the distributions $\mathcal{B}(C_{\rm gr}|M_*, M_{\rm h})$ and $\mathcal{R}(C_{\rm gr}|M_*, M_{\rm h})$ are simply the conditional distributions $\mathcal{B}(C_{\rm gr}|M_*)$ and $\mathcal{R}(C_{\rm gr}|M_*)$, respectively[3]. Therefore,

$$B(C_{\rm gr}, M_*|M_{\rm h}) \approx \mathcal{B}(C_{\rm gr}|M_*)P_{\rm B,c}(M_*|M_{\rm h}), \tag{70}$$

and

$$R(C_{\rm gr}, M_*|M_{\rm h}) \approx \mathcal{R}(C_{\rm gr}|M_*)P_{\rm R,c}(M_*|M_{\rm h}). \tag{71}$$

These equations show that the correlation between the color of the central galaxy and $M_{\rm h}$ comes from the correlation that exist between the stellar and halo. This correlation can be obtained as follows:

$$C_{\rm gr}(M_{\rm h}) \equiv \int C_{\rm gr}P(C_{\rm gr}, M_*|M_{\rm h})dC_{\rm gr}dM_* = C^{\rm B}_{\rm gr}(M_{\rm h})f_{\rm B}(M_{\rm h}) + C^{\rm R}_{\rm gr}(M_{\rm h})f_{\rm R}(M_{\rm h}), \tag{72}$$

where the first equality is the definition of $C_{\rm gr}(M_{\rm h})$, and the mean blue and red color-to-halo mass relation are obtained from Eqs. (70) and (71),

$$C^{\rm B}_{\rm gr}(M_{\rm h}) = \int C^{\rm B}_{\rm gr}(M_*)P_{\rm B,c}(M_*|M_{\rm h})dM_*, \tag{73}$$

and

$$C^{\rm R}_{\rm gr}(M_{\rm h}) = \int C^{\rm R}_{\rm gr}(M_*)P_{\rm R,c}(M_*|M_{\rm h})dM_*, \tag{74}$$

respectively. From Eq. (72), it follows that the fraction of halos $f_{\rm B}(M_{\rm h})$ hosting blue central galaxies as a function of halo mass is,

$$f_{\rm B}(M_{\rm h}) = \frac{C_{\rm gr}(M_{\rm h}) - C^{\rm R}_{\rm gr}(M_{\rm h})}{C^{\rm B}_{\rm gr}(M_{\rm h}) - C^{\rm R}_{\rm gr}(M_{\rm h})}. \tag{75}$$

In practice, we do not evaluate the integrals in Eqs. (72), (73) and (74). Instead, as shown in Appendix D, the following expression is a reasonable approximation,

$$f_{\rm B}(M_{\rm h}) \approx \frac{C_{\rm gr}(M_*(M_{\rm h})) - C^{\rm R}_{\rm gr}(M_{*,\rm R}(M_{\rm h}))}{C^{\rm B}_{\rm gr}(M_{*,\rm B}(M_{\rm h})) - C^{\rm R}_{\rm gr}(M_{*,\rm R}(M_{\rm h}))}. \tag{76}$$

The importance of the above result is that $f_{\rm B}$ can be constrained from the observed mean color-to-stellar mass relation, $C_{\rm gr}(M_*)$, and the mean blue and red color-to-stellar

3 Observe that if the scatter of the mean $M_*(M_{\rm h})$ relation is zero, then $\mathcal{B}(C_{\rm gr}|M_*, M_{\rm h}) = \mathcal{B}(C_{\rm gr}|M_*(M_{\rm h}))$ and $\mathcal{R}(C_{\rm gr}|M_*, M_{\rm h}) = \mathcal{R}(C_{\rm gr}|M_*(M_{\rm h}))$, i.e., we obtain the same result but avoiding any assumption on the relation between the color and the halo mass. Actually, there are several pieces of evidence based on the analysis of the clustering of galaxies [61, 250, 176], staked galaxy weak-lensing [122], satellite kinematics [145, 146] and statistical properties of galaxies groups [249, 247, 250, 118], that the scatter around the mean SHMR is very tight.

mass relations, $C^{\rm B}_{\rm gr}(M_*)$ and $C^{\rm R}_{\rm gr}(M_*)$, and their scatters. In Appendix E, we describe the data employed and how we obtain these relations, see also section 8.2.4. Note that the calculation of $f_{\rm B}$ involves the mean SHMR of all central galaxies and the mean SHMR of blue and red central galaxies. This directly entails that once we are constraining the mass relations of blue and red centrals, we are also constraining $f_{\rm B}$. Finally, note that our model assumes that in general $M_*(M_{\rm h}) \neq M_{*,\rm B}(M_{\rm h}) \neq M_{*,\rm R}(M_{\rm h})$.

8.2.2 *The fractions of blue/red satellite galaxies as function of halo mass*

The conditional stellar mass function of satellite galaxies, $\Phi_{\rm sat}(M_*|M_{\rm h})(M_*)$, denotes the number of satellites of stellar mass between M_* and $M_* \pm dM_*$ residing in distinct host halos of mass $M_{\rm h}$. Under this definition, the satellite GSMF, $\phi_{\rm sat}(M_*)$, is directly related to the HMF, $\phi_{\rm h}(M_{\rm h})$, according to,

$$\phi_{\rm sat}(M_*) = \int \Phi_{\rm sat}(M_*|M_{\rm h})\phi_{\rm h}(M_{\rm h})dM_{\rm h}. \tag{77}$$

which after integration over stellar mass yields to

$$n_{\rm sat}(M_*) = \int \langle N_s(>M_*|M_{\rm h})\rangle \phi_{\rm h}(M_{\rm h})dM_{\rm h}, \tag{78}$$

where $\langle N_s(>M_*|M_{\rm h})\rangle$ is the number of satellite galaxies residing in a distinct host halo above stellar mass M_*.

In order to separate satellite galaxies in colors we introduce the function $\mathcal{F}^{\rm sat}_{\rm b}$, which specifies the fraction of blue satellite galaxies above some stellar mass threshold M_* residing in halos of mass $M_{\rm h}$. In its more general form, $\mathcal{F}^{\rm sat}_{\rm b}(M_*|M_{\rm h})$ can be thought as a function depending both on M_* and $M_{\rm h}$.

Studies on the spatial clustering of galaxies have shown its dependence on luminosity, stellar mass and color [155, 154, 259, 123, 258]. In particular, when galaxies are divided by color, red galaxies appear to be more strongly clustered. Therefore, a qualitative interpretation is that red galaxies preferentially reside in groups and clusters, that is, high-density environments. On the other hand, studies on galaxy groups have shown that the relative fraction of red satellites increases with cluster mass [236, 164, 134, 30, 103, 247]. Based on these facts and after experimenting with different functional forms, we define the fraction of blue satellite galaxies above the stellar mass threshold M_* as a function of halo mass $M_{\rm h}$ as:

$$\mathcal{F}^{\rm sat}_{\rm b}(M_*|M_{\rm h}) = \min[1, \max[0, \mathcal{F}_{\rm b}(M_*|M_{\rm h})]], \tag{79}$$

where

$$\mathcal{F}_{\rm b}(M_*|M_{\rm h}) = \mathcal{F}_0 \left(\frac{M_{\rm h}}{M_{\rm cut,b}}\right)^{\alpha_b} \left(\frac{M_*}{\sigma_b M_*(M_{\rm h})}\right)^{\mu_b}. \tag{80}$$

Note that the function $\mathcal{F}^{\rm sat}_{\rm b}(M_*|M_{\rm h})$ is always constrained to be between zero and one. We assume that $\mathcal{F}_0$, $M_{\rm cut,b}$, α_b, μ_b and σ_b are free parameters. For the function $\mathcal{F}_{\rm b}(M_*|M_{\rm h})$ the slopes α_b and μ_b control the behavior of the fraction as a function of halo and stellar mass, respectively, and $M_{\rm cut,b}$ is a characteristic mass. The inclusion of the term $M_*(M_{\rm h})$

as a characteristic mass is motivated by the results reported on the satellite CSMF in YMB09. The amplitude of this function is controled by $\mathcal{F}_0$.

Finally the satellite GSMF of blue and red galaxies are given by:

$$\phi_{\rm sat,blue}(M_*) = \int \frac{\partial \langle N_{s,\rm b}(>M_*|M_{\rm h})\rangle}{\partial M_*}\phi_{\rm h}(M_{\rm h})dM_{\rm h}, \tag{81}$$

and

$$\phi_{\rm sat,red}(M_*) = \phi_{\rm sat}(M_*) - \phi_{\rm sat,blue}(M_*), \tag{82}$$

respectively, where $\langle N_{s,\rm b}(>M_*|M_{\rm h})\rangle = \mathcal{F}_{\rm b}^{\rm sat}(M_*|M_{\rm h}) \times \langle N_s(>M_*|M_{\rm h})\rangle$.

8.2.3 *The spatial clustering of galaxies*

Our approach includes the so-called HOD model, which connects the occupational statistics of galaxies as a function of halo mass with their spatial clustering [e.g., 259].

The HOD model characterize the probability that a halo of mass $M_{\rm h}$ hosts a number of N galaxies with stellar masses greater than M_*. Within this formulation, the 2PCF is obtained by assuming that the total mean number of galaxy pairs is the contribution of all pairs coming from galaxies in the same halo (one-halo term) and pairs from different halos (two-halo term). When dividing the population of galaxies into its blue and red components, the generalization of the HOD model is relatively easy. In this case, the one-halo regime gives information about the spatial clustering of galaxy pairs within a common halo and lying in the same color sequence. For the two-halo regime, galaxy pairs drawn from the same color sequence are counted from different halos. For the details of the HOD model used in this paper we refer to the reader to Rodríguez-Puebla, Drory & Avila-Reese [178] and Rodríguez-Puebla, Avila-Reese & Drory [176] (Part II). Here we outline only the main features of the model.

As it is common for the one-halo term, we assume that the number of central-satellite pairs follow the normalized mass density halo profile, taken to have the Navarro, Frenk & White [151] shape and the number of satellite-satellite pairs related to this normalized density profile convolved with itself. This profile is characterized in terms of the total halo mass and the concentration parameter.

The two-halo term, which dominates at large scales, is obtained as the weighted average of the halo bias function [191] times the non-linear matter 2PCF [196]. The weight of the halo bias will be proportional to the mean occupational number of galaxies times the halo mass function.

Finally, we relate the correlation function, $\xi_{\rm gg}^{\rm C}(r)$, to the projected correlation function, $w_{\rm p}^{\rm C}(r_{\rm p})$, by integration over the line of sight,

$$w_{\rm p}^{\rm C}(r_{\rm p}) = 2\int_0^\infty \xi_{\rm gg}^{\rm C}(\sqrt{r_{\rm p}^2 + x^2})dx, \tag{83}$$

where C denotes blue or red.

8.2.4 *Summary of the model*

Our main goal in this Chapter is to constrain the mean SHMR of blue and red central galaxies, the fraction of halos hosting blue central galaxies, and the fraction of blue/red satellite galaxies as a function of halo mass. To do so, we employ the blue and red GSMF of all galaxies and the projected 2PCFs decomposed into blue and red galaxies as observational constraints. Below, we summarize the main equations that relate our model with these observational constraints.

1. The GSMF of blue and red centrals are given by

$$\phi_{\rm B,c}(M_*) = \int P_{\rm B,c}(M_*|M_{\rm h})\phi_{\rm h,B}(M_{\rm h})dM_{\rm h}, \tag{84}$$

and

$$\phi_{\rm R,c}(M_*) = \int P_{\rm R,c}(M_*|M_{\rm h})\phi_{\rm h,R}(M_{\rm h})dM_{\rm h}, \tag{85}$$

Here the halo mass functions hosting blue and red central galaxies are given by $\phi_{\rm h,B} = f_{\rm B}(M_{\rm h})\phi_{\rm h}(M_{\rm h})$ and $\phi_{\rm h,R}(M_{\rm h}) = \phi_{\rm h}(M_{\rm h}) - \phi_{\rm h,B}(M_{\rm h})$.

2. The fraction of halos hosting blue galaxies is given by

$$f_{\rm B}(M_{\rm h}) \approx \frac{C_{\rm gr}(M_*(M_{\rm h})) - C^{\rm R}_{\rm gr}(M_{*,\rm R}(M_{\rm h}))}{C^{\rm B}_{\rm gr}(M_{*,\rm B}(M_{\rm h})) - C^{\rm R}_{\rm gr}(M_{*,\rm R}(M_{\rm h}))}. \tag{86}$$

3. The satellite GSMF of blue and red galaxies are given by

$$\phi_{\rm sat,blue}(M_*) = \int \frac{\partial\langle N_{s,\rm b}(>M_*|M_{\rm h})\rangle}{\partial M_*}\phi_{\rm h}(M_{\rm h})dM_{\rm h}, \tag{87}$$

and

$$\phi_{\rm sat,red}(M_*) = \phi_{\rm sat}(M_*) - \phi_{\rm sat,blue}(M_*), \tag{88}$$

respectively, where $\langle N_{s,\rm b}(>M_*|M_{\rm h})\rangle$ is the number of blue satellite galaxies above some stellar mass threshold M_* residing in halos of mass $M_{\rm h}$.

4. We compute $\langle N_{s,\rm b}(>M_*|M_{\rm h})\rangle = \mathcal{F}^{\rm sat}_{\rm b}(M_*|M_{\rm h}) \times \langle N_s(>M_*|M_{\rm h})\rangle$, where $\mathcal{F}^{\rm sat}_{\rm b}(M_*|M_{\rm h})$ is the fraction of blue satellite galaxies above some stellar mass threshold M_* residing in halos of mass $M_{\rm h}$; this fraction is given by a parametric function motivated by observational studies.

5. The total blue and red GSMFs are given by

$$\phi_{\rm blue}(M_*) = \phi_{\rm B,c}(M_*) + \phi_{\rm sat,blue}(M_*), \tag{89}$$

and

$$\phi_{\rm red}(M_*) = \phi_{\rm R,c}(M_*) + \phi_{\rm sat,red}(M_*), \tag{90}$$

6. We compute the projected 2PCF of blue and red galaxies using the HOD model.

In item (1), we assume lognormal distributions both for $P_{\rm B,c}(M_*|M_{\rm h})$ and $P_{\rm R,c}(M_*|M_{\rm h})$. Both lognormal distributions are characterized by their mean, $M_{*,\rm B}(M_{\rm h})$ and $M_{*,\rm R}(M_{\rm h})$, and their scatters σ_b and σ_r. Our adopted parameterization for the mean SHMR of blue central galaxies, $M_{*,\rm B}(M_{\rm h})$, is motivated by the functional form proposed in Behroozi, Wechsler & Conroy [20]. This function consist of five parameters and in the context of the AMT reproduces the main features of a Schechter-like function from a power-law halo mass distribution. For red central galaxies, by using Eq. (66), one finds that the relation between the mean SHMR of all central galaxies, $M_*(M_{\rm h})$, the mean SHMR of blue central galaxies, $M_{*,\rm B}(M_{\rm h})$, and the mean SHMR of red central galaxies, $M_{*,\rm R}(M_{\rm h})$, is given by,

$$\log(M_{*,\rm R}(M_{\rm h})) = \frac{\log(M_{*,\rm B}(M_{\rm h}))f_{\rm B}(M_{\rm h}) - \log(M_*(M_{\rm h}))}{1 - f_{\rm B}(M_{\rm h})}. \tag{91}$$

where we have use the fact that $f_{\rm R}(M_{\rm h}) = 1 - f_{\rm B}(M_{\rm h})$. Therefore, the red mean $M_{*,\rm R}(M_{\rm h})$ relation is a function that does not need to be independetly constrained. For the SHMR of all central galaxies, we use the set C constrained in Chapter 3. This SHMR and its scatter was inferred by using the decomposition of the central and satellite GSMF and the projected 2PCFs in different M_* bins.

For the distinct HMF, we use the accurate fit to the abundance of dark matter halos measured in large cosmological N-body simulations by Tinker et al. [207].

In item (2), we use the observed color-stellar mass distributions. As detailed in Appendix E, by using the Yang et al. [249] group catalog, based on the SDSS DR4, we constrain the mean and the width of the blue and red distribution of central galaxies as a function of stellar mass. To do so, we assume that the distribution of ${}^{0.1}(g-r)$ colors as a function of M_* of central galaxies are drawn from bi-modal distribution functions that depends on M_*. The distribution of each mode is represented by Gaussian functions denoted by $\mathcal{R}(C_{\rm gr}|M_*)$ and $\mathcal{B}(C_{\rm gr}|M_*)$, respectively. The parameters $C^{\rm B}_{\rm gr}(M_*)$ and $C^{\rm R}_{\rm gr}(M_*)$, and $\sigma_B(M_*)$ and $\sigma_R(M_*)$ denote the mean and the standard deviation for the blue-cloud and red-sequence modes, respectively.

In items (3) and (4), the main ingredient is the fraction of blue satellite galaxies above the stellar mass threshold M_* residing in halos of mass $M_{\rm h}$. The parameterization used here is given by Eq. (80).

In item (6), we use the relation between concentration parameter $c_{\rm NFW}$ as a function of halo mass obtained by Muñoz-Cuartas et al. [150] from fits to N-body simulations. Additionally, based on the results of high-resolution N-body [115] and hydrodynamic simulations of galaxy formation [262], we assume that the second moment of satellite galaxies, $\langle N_s(N_s-1)\rangle$, follows a Poisson distribution with mean $\langle N_s\rangle^2 = \langle N_s(N_s-1)\rangle$.

Finally, in total our model consist of ten parameters. Five parameters to model the SHMR for blue (and therefore red central galaxies, see above), and five more for the occupational numbers of blue and red satellites galaxies. We recall that the global SHMR of all central galaxies is taken from a previous determinations (by using the same Yang et al. [249] data used here).

8.3 OBSERVATIONAL DATA AND FITTING PROCEDURE

8.3.1 *The data*

In order to constrain the model parameters in our model, the combination of the total blue and red GSMFs and the projected 2PCFs of blue and red galaxies for various mass bins are necessary.

Here we use the YMB09 GSMF of blue and red galaxies. These data were obtained from the SDSS DR4 and within the redshift range $0.01 < z < 0.2$. Stellar masses were obtained from the fitted relation between the stellar-to-light mass ratio and color reported in Bell et al. [21]. To be consistent with the Chabrier [48] initial mass function adopted here, we subtract 0.1 dex to the stellar masses of YMB09 galaxies.

For the correlation functions, we use the Li et al. [123] measurements of $w_{\rm p}(r_{\rm p})$ in five different stellar mass bins for a sample of $\sim 200,000$ galaxies based on the SDSS. They also measured the clustering of galaxies spitted in blue and red ones. Note that Li et al. [123] use a very similar color cut to ours to define blue galaxies.

8.3.2 *Fitting procedure and uncertainties*

To constraint the set of ten parameter introduced in our model, we use the Powell's directions set method in multi-dimensions for the minimization [167]. We infer the best fitting model parameters by minimizing the merit function

$$\chi^2 = \chi^2(\phi^{\rm Y09}_{\rm tot,blue}) + \chi^2(\phi^{\rm Y09}_{\rm tot,red}) + \chi^2(w^{\rm L06}_{\rm p,blue}) + \chi^2(w^{\rm L06}_{\rm p,red}), \tag{92}$$

where,

$$\chi^2(\phi^{\rm Y09}_{\rm tot,color}) = \frac{1}{N_{\rm bin}} \sum_{i=1}^{N_{\rm bin}} \left(\frac{\phi_{g,{\rm model}}^{i,{\rm tot,color}} - \phi_{g,{\rm obs}}^{i,{\rm tot,color}}}{\sigma_{\rm obs}^{i,{\rm tot,color}}} \right)^2 . \tag{93}$$

Here, $\phi_{g,{\rm obs}}^{i,{\rm tot,color}}$ are the observed values of the total blue and red GSMFs with errors $\sigma_{\rm obs}^{i,{\rm sat/cen,color}}$, while $\phi_{g,{\rm model}}^{\rm color}$ denotes the same but for the model values.

For the spatial clustering we assume that:

$$\chi^2(w^{\rm L06}_{\rm p,color}) = \frac{1}{N_{\rm s,bin} N_{\rm r,bin}} \sum_{i=1}^{N_{\rm s,bin}} \sum_{j=1}^{N_{\rm r,bin}} \left(\frac{w_{\rm p,model}^{i,j} - w_{\rm p,obs}^{i,j,{\rm color}}}{\sigma_{\rm obs}^{i,j,{\rm color}}} \right)^2 . \tag{94}$$

where $N_{\rm s,bin}$ is the number of stellar mass bins; $N_{\rm r,bin}$ denotes the number of bins in the projected 2PCF; $w_{\rm p,obs}^{i,j,{\rm color}}(w_{\rm p,model}^{i,j,{\rm color}})$ is the amplitude of the observed (modeled) projected 2PCF at the jth projected distance bin of the ith stellar mass bin, for blue and red galaxies.

In Behroozi, Conroy & Wechsler [19], the uncertainties that affect the SHMR have been discussed in great detail. In that paper the authors conclude that the largest uncertainty by far in the SHMRs, at least for local galaxies, is due to systematics shifts in the stellar mass estimates. Here we assume that this error is of the order of 0.25 dex. The second

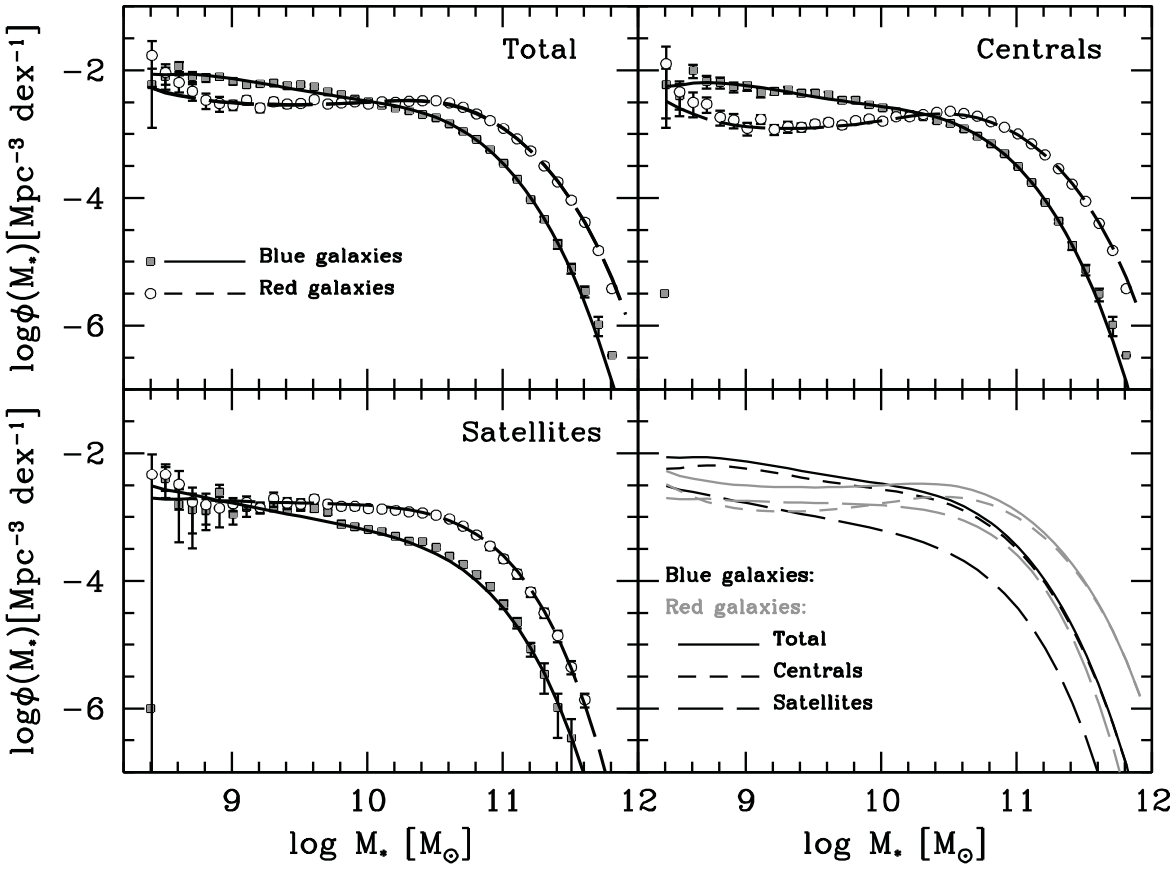

Figure 37: Galaxy stellar mass function for all, central and satellite blue and red galaxies. In all the panels, except bottom right panel, the solid lines indicate the abundance of blue galaxies, while the long dashed lines indicate the abundance of red galaxies. The observational inferences reported in YMB09 for the blue (red) galaxies are plotted with filled squares (empty circles). *Upper left panel:* Best fitting model to the observed total GSMF of blue and red galaxies. *Upper right panel:* Model predictions for the central GSMF of blue and red galaxies. *Bottom left panel:* Model predictions for the satellite GSMF of blue and red galaxies. *Bottom right panel:* For illustrative purpose, we compare the blue and the red GSMFs for all, central, and satellite galaxies.

most important source is the uncertainty in the intrinsic scatter of the mean SHMR. In our analysis, the intrinsic scatter of the stellar mass at a fixed halo mass is modeled (by construction) by using the distribution functions $P_{\rm B,c}(M_*|M_{\rm h})$ and $P_{\rm R,c}(M_*|M_{\rm h})$.

Finally, observe that while the YMB09 GSMF and the Li et al. [123] projected 2PCF in various mass bins are used to constrain the model parameters, the fraction of halos hosting blue central galaxies, the GSMFs decomposed into centrals and satellites, and the CSMFs in different halo mass bins for blue and red galaxies are predictions of our model.

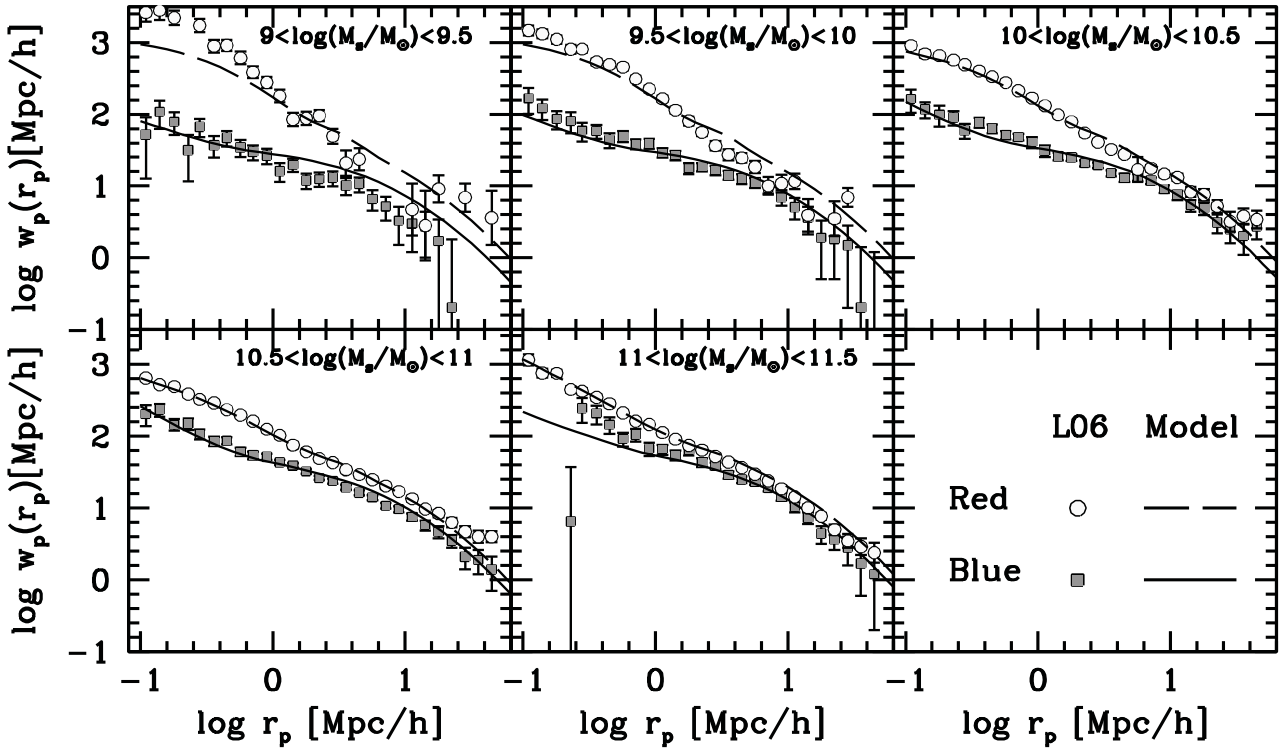

Figure 38: Projected 2PCF in five stellar mass bins for blue and red galaxies. The solid lines indicate the best fitting model for blue galaxies, while long-dashed lines are for red galaxies. The reported projected 2PCF in Li et al. [123] for blue and red galaxies are shown with the filled squares and empty circles, respectively.

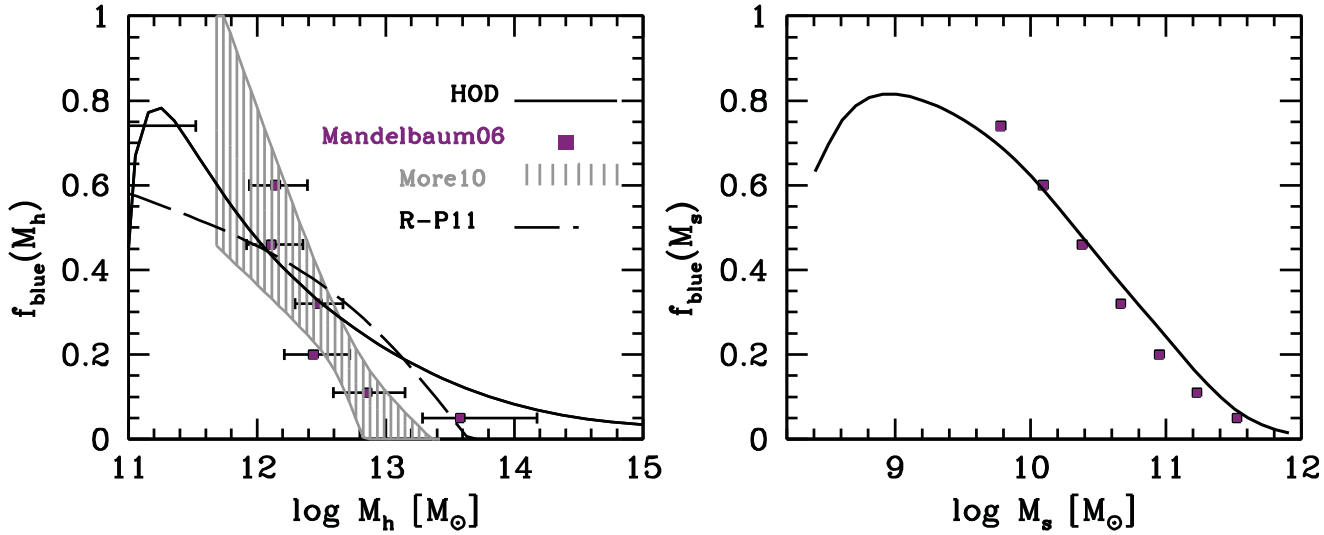

Figure 39: *Left Panel:* Fraction of halos from the distinct halo mass function associated to blue galaxies. Fractions reported from staked weak-lensing [133] and satellite kinematics [146] are plotted with filled squares/error bars and the shaded area, respectively. *Left Panel:* The same as the left panel but as a function of stellar mass.

8.4 RESULTS

8.4.1 *Best fit model and the fraction of halos hosting blue/red central galaxies*

The free parameters of our model (§§8.2.4) are constrained by fitting the model to the observed GSMF of blue and red centrals and the blue/red projected 2PCFs in different M_* bins. Figures 37 (upper right panel) and 38 present the best fits for the GSMFs of central blue and red galaxies, and the projected 2PCFs of blue and red galaxies in five M_* bins. These figures show the accuracy at which the model can be constrained. The model has some difficulty in reproducing the steep slope at the lowest masses of the red central GSMF, as well as the steep 2PCF at the smallest distances for these galaxies. Recall that the decomposition of the GSMFs into centrals and satellites is already a prediction of the model, not a fit[4]. The left upper and lower panels of Fig. 37 compare predictions of our best fitted model with the observations of the total and only satellite blue/red GSMFs from YMB09. The agreement is good, the model predictions are always within the uncertainty of the observations.

The left panel of Fig. 39 presents the resulting fraction of halos hosting blue central galaxies as a function of $M_{\rm h}$, while the right panel shows the same fractions but as a function of M_*. Our best constrained model is indicated with a solid line. The dashed line is the fraction used in the previous Chapter [177] under the assumptions mentioned in the Introduction. For comparison, in the left panel the filled squares with error bars represent the blue fraction as reported from staked weak-lensing [133][5], and the shaded area indicates the results from the analysis of satellite kinematics [146]. In the right panel, the observed fractions of blue centrals as a function of stellar mass reported in Mandelbaum et al. [133] are plotted as well.

We observe that the fraction of halos hosting blue centrals in our constrained model present a maximum of $f_{\rm blue} \approx 0.8$ at $M_{\rm h} = 10^{11.4}{\rm M}_\odot$, and as $M_{\rm h}$ increases, $f_{\rm blue}$ declines slowly. At $\sim 10^{12}{\rm M}_\odot$ the fraction of halos hosting blue and red centrals is equal, i.e., $f_{\rm blue} = f_{\rm red} = 0.5$, and it falls below 0.1 for masses larger than $M_{\rm h} \sim 5 \times 10^{13}{\rm M}_\odot$. The shape of this fraction is basically constrained by the shape of the fraction of blue centrals as a function of stellar mass, which is given by the observed central blue and red GSMFs (reproduced in the right panel of Fig. 39). In this sense, the fast increase of $f_{\rm blue}$ as $M_{\rm h}$ at low halo masses, and then its strong decline, is the result of the flattening (steepening) of the low-mass end shapes of blue (red) central GSMFs, respectively (see Fig. 37).

From the comparison with direct inferences of halo masses, we see that our model predicts a slightly higher fraction of high-mass blue halos. In the analysis of satellite kinematics, practically all halos more massive than $\sim 3 \times 10^{13}{\rm M}_\odot$ hosts only red central galaxies, while our results show that this fraction is about $\sim 85\%$ which is more consistent with weak-lensing results. Note that in the case of the analysis of staked satellite kinematics, blue galaxies were selected by using a stricter color cut than in YMB09 catalog (see Appendix A of [146]). Presumably, this is the reason why there are not ha-

4 In any case, we have explored also the case of fitting the blue and red GSMFs provided by YMB09 instead of predicting them (analogous to set C in Chapter 3) and found that the SHMR and other predictions remain the same.

5 The data was taken using their table 1.

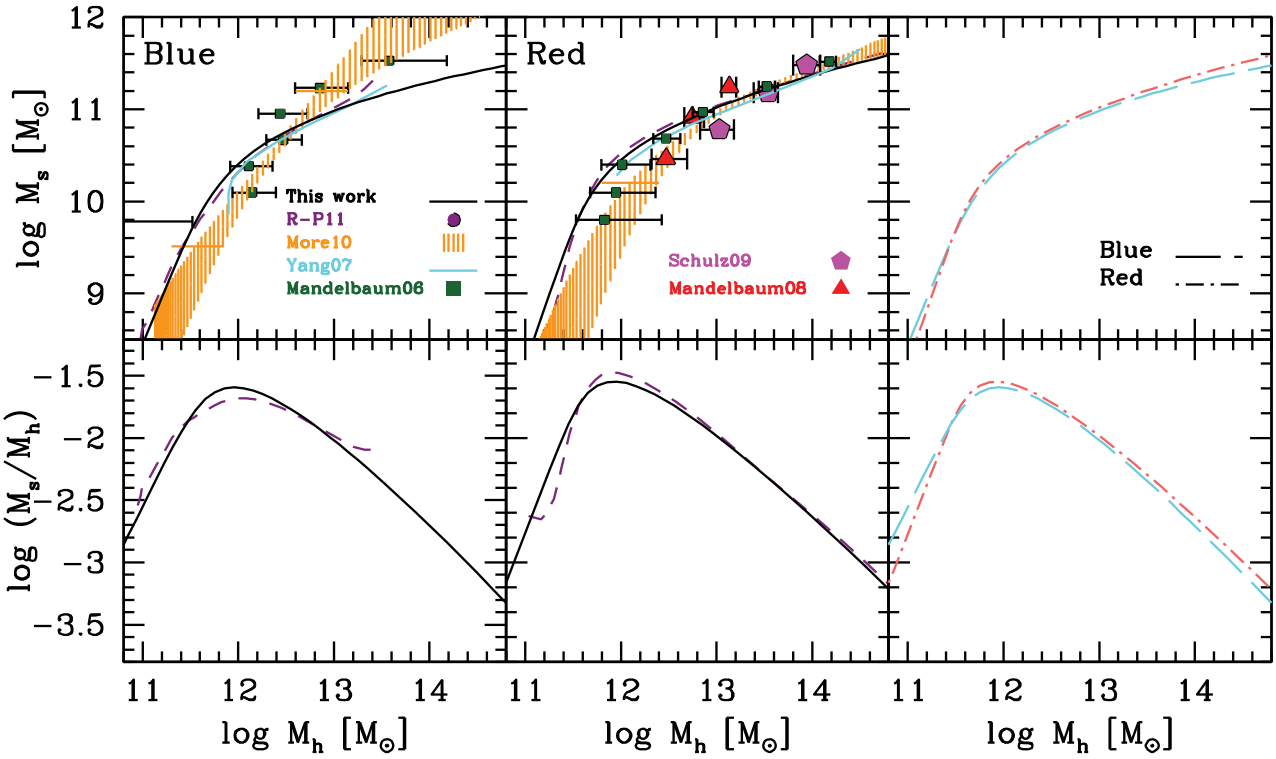

Figure 40: Stellar-to-halo mass relations of blue and red central galaxies. *Upper panels:* From left to right, the SHMR of blue and red galaxies, and a comparison between them. Solid lines are for the results obtained here. Filled squares, pentagons, and triangles with error bars correspond to the results by Mandelbaum et al. [133], Mandelbaum, Seljak & Hirata [132] and Schulz, Mandelbaum & Padmanabhan [184] using galaxy-galaxy weak lensing. The results obtained from group catalogs [249] are shown with cyan solid lines, while indirect inferences using the AMT [177] are shown with short dashed lines. *Lower panels:* From left to right, th f_s–M_h relations obtained here for blue and red central galaxies, and a comparison between them.

los hosting blue centrals galaxies more massive than $\sim 10^{13}M_\odot$. At halo masses below $\sim 10^{13}M_\odot$, our fraction is in good agreement in both cases, within the uncertainties, both for staked weak-lensing and satellite kinematics results.

8.4.2 *The stellar-to-halo mass relations of blue and red central galaxies*

Top panels of Fig. 40 show the constrained local SHMRs for blue and red central galaxies (solid line). In the same figure we plot some direct inferences of halo masses based on staked weak lensing [133, 132, 184][6], and staked satellite kinematics [146]. The indirect statistical inferences by Yang et al. [249] and [177], for both blue and red central galaxies, are also plotted for comparison.

The right panel of Fig. 40 compares the blue and red central SHMRs. We find no significative differences between the SHMRs of both galaxy populations. If any, there is a systematical trend of blue centrals to have slightly lower stellar masses than red

6 Note that the mass relations based on staked weak-lensing measurements are averaged at a fixed stellar mass, i.e, $\langle M_h\rangle(M_*)$, while our mass relation are averaged at a fixed halo mass, $\langle M_*\rangle(M_h)$.

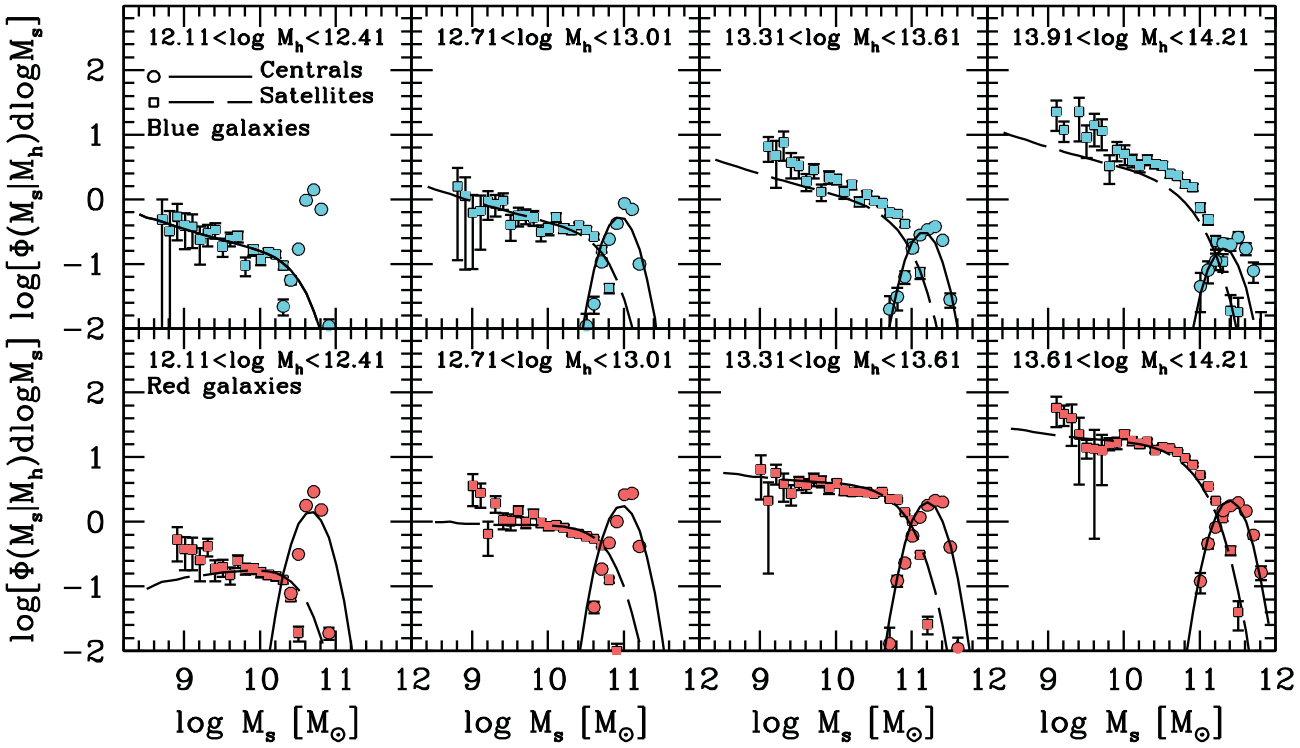

Figure 41: Density-weighted average CSMFs of blue and red galaxies. The solid and long dashed lines show the resulting central and satellite CSMFs. The filled circles and squares with error bars show the corresponding observational inferences by YMB09. *Upper panels:* The CSMFs for blue galaxies. *Lower panels:* The CSMFs for red galaxies.

centrals for a given halo mass above $M_{\rm h} \sim 7 \times 10^{11}$ $M_\odot$. At low masses, the blue (red) SHMR scales as $M_* \propto {M_{\rm h}}^{2.56}$ ($M_* \propto {M_{\rm h}}^{3.03}$). At high masses, the scaling for blue (red) centrals is $M_* \propto {M_{\rm h}}^{0.63}$ ($M_* \propto {M_{\rm h}}^{0.57}$).

The corresponding $f_{\rm s} = M_*/M_{\rm h}$ fractions vs $M_{\rm h}$ are shown in the bottom panels of Fig. 40. The maximum for blue (red) centrals is $f_{\rm s} \sim 0.025$ ($f_{\rm s} \sim 0.031$) and it occurs at $\log(M_{\rm h}/{\rm M}_\odot) \approx 11.9$ ($\log(M_{\rm h}/{\rm M}_\odot) \approx 11.8$). The $f_{\rm s}$ fraction is interpreted as the efficiency of galaxy stellar mass growth in halos. Our results show then that this efficiency is almost independent of the color type of central galaxies. Red centrals have slightly larger values of $f_{\rm s}$, specially in the $11.7 < \log(M_{\rm h}) < 12.2$ mass range, where the peak efficiency is attained.

Finally, we find that our mass relations are roughly consistent with the staked weak-lensing and satellite kinematics estimates, taking into account their large uncertainties. For blue central galaxies, our results agree well with these inferences over the halo mass range $12 < \log(M_{\rm h}/M_*) < 13$, and for red centrals, the agreement is remarkable for $\log(M_{\rm h}/M_*) > 12.5$. Regarding the halo masses reported in the group catalog of Yang et al. [249], they are in agreement with our results, as well with previous results based on the abundance matching results of Rodríguez-Puebla et al. [177].

8.4.3 *Occupational numbers*

An immediate prediction of the model are the occupational statistics of galaxies in dark matter halos. We compare the model predictions with the observed satellite and central CSMFs for blue and red galaxies in the upper and bottom panels of Figure 41. The filled circles with error bars denote the observed YMB09 central CSMFs, while the filled squares with error bars denote the observed YMB09 satellite CSMFs. Because the observational data are given for counts in different halo mass ranges, we compute the mean halo-density-weighted blue central CSMF in the $[M_h^1, M_h^2]$ bin in order to compare with these observations:

$$\langle \Phi_{c,B} \rangle = \frac{\int_{M_h^1}^{M_h^2} \Phi_{c,b}(M_*|M_{\rm h})\phi_{\rm h,B}(M_{\rm h})dM_{\rm h}}{\int_{M_h^1}^{M_h^2} \phi_{\rm h,B}(M_{\rm h})dM_{\rm h}}, \tag{95}$$

and the mean halo-density-weighted blue satellite CSMF according to:

$$\langle \Phi_{s,B} \rangle = \frac{\int_{M_h^1}^{M_h^2} \Phi_{s,b}(M_*|M_{\rm h})\phi_{\rm h}(M_{\rm h})dM_{\rm h}}{\int_{M_h^1}^{M_h^2} \phi_{\rm h}(M_{\rm h})dM_{\rm h}}. \tag{96}$$

Analogously, we compute the mean halo-density-weighted CSMFs for red central and satellite galaxies.

In general, our central CSMFs match observational expectations fairly well, practically at all masses both for central and satellites CSMFs. Nevertheless the faint-end of the observed red satellite CSMFs in low mass halos, $\log(M_{\rm h}/{\rm M}_\odot) < 12.7$, is steeper than our model prediction, although still within the uncertainties.

Figure 42 shows the mean central and satellite occupational numbers as a function of halo mass for blue and red galaxies, separately, in stellar mass bins (upper panels) and for stellar mass thresholds (bottom panels). The left panels of Fig. 42 shows that the probability of having a blue central galaxy decreases with halo mass and eventually the majority of massive halos will host red central galaxies. This is similar to Fig. 41.

In the middle panels of the same figure we present the occupational numbers of blue an red satellite galaxies. The number of red satellites increases much faster than the number of blue satellites as a function of halo mass. We found that the number of blue satellites increases as a power law with $M_{\rm h}$ with a slope of 0.61, while the red sequences grows as 1.06, independent of the stellar mass. Consistent with our results, previous works have found that power-law slope for blue, l_b, is lower than red galaxies, l_r. For instance, Magliocchetti & Porciani [131] found $l_b = 0.7$ and $l_r = 1.1$, while Collister & Lahav [54] found $l_b = 0.88$ and $l_r = 1.05$.

8.5 DISCUSSION

8.5.1 *The stellar-to-halo mass relations of blue and red galaxies*

It is important to clarify that the inferences of the stellar-to-halo mass relations, SHMRs, of local central blue and red galaxies performed here are not independent. In fact, we are

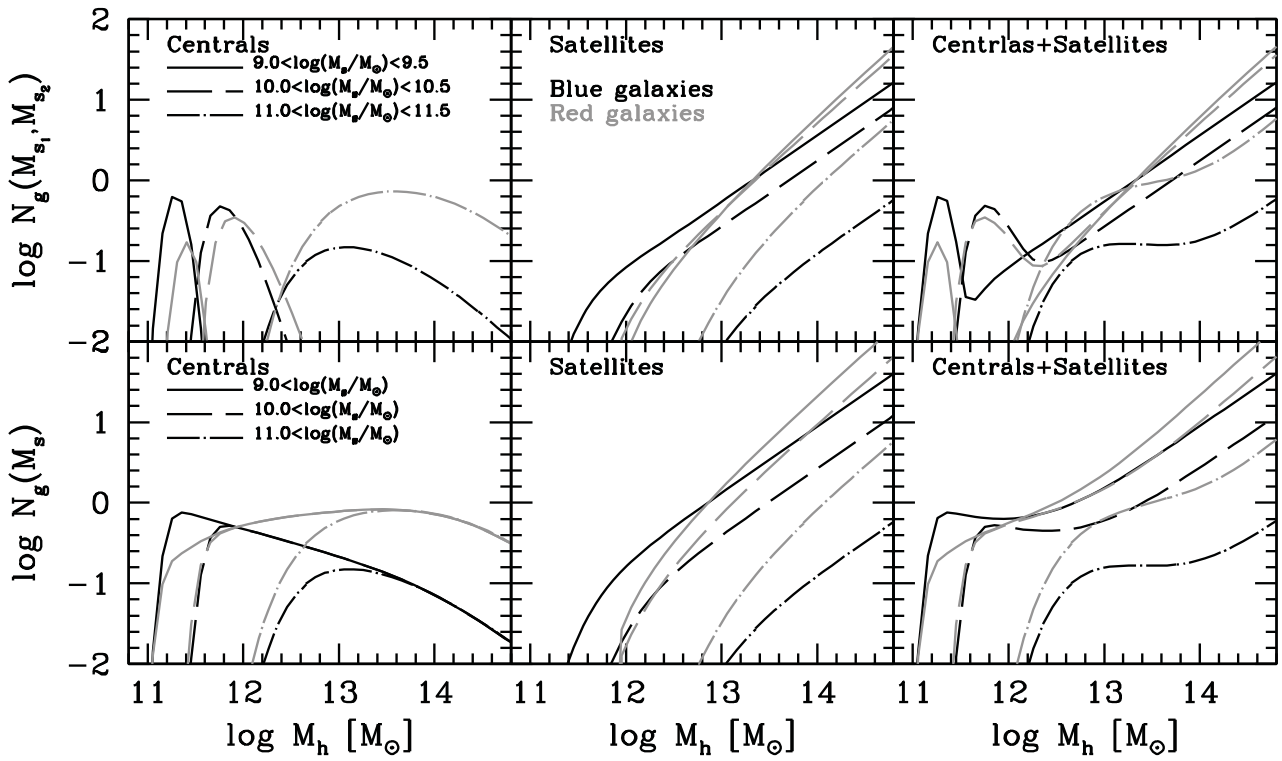

Figure 42: Occupational number statistics for blue and red galaxies (see the text). *Upper panels:* For stellar mass threshold. *lower panels:* For stellar mass bins.

using the previous constrained (Chapter 3) global SHMR, its scatter, and the observed color–M_* distributions, in order to sample the distributions of halos associated to blue and red galaxies. Recall that we do so under the assumptions that for a given M_*, the color does not depend on M_h and that the scatter distributions of the blue/red SHMRs are lognormal and within the scatter of the global SHMR. It is within this scheme and assumptions that, by means of our statistical method, we could constrain the SHMRs of blue and red galaxies. The relevant point to remark is the overall self-consistency among several kind of observations that we attain through our method and with the blue/red central SHMRs constrained through it. We can say that the overall model results (see Section 8.4) are a good description of the observed global GSMF and its decomposition into blue and red components, the observed projected 2PCFs of blue and red galaxies, the color–M_* distributions, and the occupational statistics of blue and red satellites. This encouraging self-consistency of our model and the observations suggests that the underlying blue and red SHMRs are the correct ones.

Our results show that the SHMR of the galaxies separated into two broad groups, blue and red objects, is close to the global one, i.e., galaxies do not segregate significantly by color in the M_*–M_h (or f_s–M_h) relation, implying that the color is not related to the efficiency of galaxy stellar mass growth, f_s. As a first approximation, we could say the color of a galaxy is related to the halo mass aggregation history (MAH): halos that assemble most of their mass early would not provide of baryonic material to the galaxy for its late growth and the galaxy should be redder than anoother formed in a halo with a more extended mass aggregation history [see e.g., 8]. The shape of the halo MAH is

related to its present day concentration [11, 232, 7]. Thus, if the galaxy color correlates with the halo concentration, then a significant segregation of galaxies in the M_*–M_h relation is expected. However, the scatter of the concentration for a given M_h is large, hence a large scatter in the M_*–M_h relation due to color would be obtained. The scatter around the M_*–M_h relation is actually very small, in particular for central galaxies in halos without satellites [118].

Our finding that the f_s–M_h relation of central galaxies does not differ significantly for blue and red galaxies suggests that the physical processes that coin the main f_s–M_h relation largely supersede possible effects of the halo MAH (related to the color) on the final stellar mass.

At masses much smaller than $M_h \sim 10^{12}$ $M_\odot$, the decreasing efficiency of f_s in low mass M_h is determined by the star formation (mainly SN)–driven feedback. The gas mass infall driven by the (cosmological) halo MAH is probably superseded by the local outflow (and eventual re-accretion) processes, in such a way that the variations in M_* for a given M_h for these low-mass systems does not correlate with the halo MAH. Our results show that at these low masses, red galaxies have slightly higher values of f_s than blue galaxies. This could be a consequence of the fact that galaxies formed in halos assembled earlier, had early periods of active star formation and therefore strong SN-driven gas ejection, able to blown up away all the gas around, resulting in today a red galaxy with a stellar mass lower than the case of galaxy formed in the same halo but with a more quite MAH and star formation history.

At masses much larger than $M_h \sim 10^{12}$ $M_\odot$, the decreasing star formation efficiency is larger is determined by the large gas cooling times and the AGN–driven feedback. Again, these processes certainly supersede the gas infall due to the halo MAH, in such a way that the resulting color does not correlate with the final galaxy M_* for a given M_h.

At masses around $M_h \sim 10^{12}$ $M_\odot$, where the efficiency peaks (Fig. 40), our results show slightly higher efficiencies for red galaxies than for the blue ones. The efficiency peaks at these masses, just because these are the scales where (1) the SN-driven outflows are already not important for such large gravitational potentials, and (2) the large gas cooling time and AGN–driven feedback do not yet affect the central galaxy growth. Therefore, at these scales the mentioned above processes are not expected to supersede the effects of the halo MAH on the central galaxy mass growth. Interestingly enough, our results show that around $M_h \sim 10^{12}$ $M_\odot$, red galaxies have slightly larger values of f_s than blue galaxies, in agreement with the idea that they assembled more efficiently in time just because the available baryons were accreted early due to their MAHs; the galaxies of the same present-day halo mass that formed in halos with a more extended MAH, accreted hence later a fraction of the baryons in such a way that they are bluer and with less stars (smaller M_*). In any case, the differences we have found in the efficiencies of blue and red galaxies are very small.

The lack of dependence of the galaxy stellar mass assembly on the MAH of their halos is evidenced also by recent semi-empirical inferences of the SHMR at different redshifts (see Chapter 6 and the references therein). The SHMRs at different redshifts are connected by the average halo MAHs and, this way, the corresponding average galaxy stellar mass aggregation histories are inferred. As early shown in Conroy & Wechsler [57] and Firmani & Avila-Reese [84], the shapes of the average halo and galaxy MAHs are roughly similar only for $M_h \sim 10^{12}$ $M_\odot$. At larger masses, the stellar mass growth

stops at higher redshifts for larger halo masses (downsizing in mass), while the halo MAH shows the opposite behavior, being more active at late times for large halo masses. At masses smaller than $M_{\rm h} \sim 10^{12}$ M$_\odot$, while the halo MAH at late times becomes flatter as smaller is $M_{\rm h}$, the galaxy stellar mass growth becomes actively growing (downsizing in specific star formation rate). Therefore, these semi-empirical inferences are consistent with the idea that the galaxy stellar mass assembly history is detached from the MAH of its host halo.

8.5.2 *The host halos of blue and red galaxies*

According to the results shown in Fig. 39, the fraction of halos in the distinct halo mass function that today are associated to central blue galaxies is maximal at $M_{\rm h} \approx 2 \times 10^{11}M_\odot$, attaining a value of $f_{\rm blue} \sim 0.8$. At smaller masses, $f_{\rm blue}$ decreases, and this is associated to the low-mass end shapes of the blue/red central GSMFs from YMB09, where the abundance of red galaxies increases while the one of blue galaxies remains roughly constant. At larger masses, $f_{\rm blue}$ decreases. Why do the fractions of blue and red central galaxies depend on halo mass?

In Rodríguez-Puebla et al. [177, see also Chapter 7], we have postulated that $f_{\rm blue}$ is associated to the major merger history of the halos, and thus we have imposed two criteria to select halos able to host blue (late-type) galaxies: (1) halos that did not suffer a major merger since $z = 0.8$, and (2) halos that do not host classical groups richer than 3 members (blue centrals in rich group/cluster of galaxies are rarely observed and this could be related to the merger–driven violent assembly of central galaxies in high-density environments).

From Fig. 39, one sees that the assumed $f_{\rm blue}$ in Rodríguez-Puebla et al. [177] roughly agrees with the constrained one here, suggesting that the physics behind the fraction $f_{\rm blue}$ (and its complement, $f_{\rm red}$) is qualitatively the one mentioned above. However, we see also in Fig. 39 that there are some differences at low and high halo masses. The assumption in Rodríguez-Puebla et al. [177] that disk (∼blue) galaxies do not exist in halos that suffered major mergers with ratios larger than 0.2 after $z = 0.8$ seems to be too strong for $M_{\rm h} \lesssim 10^{12}$ M$_\odot$. It could be that for smaller galaxies, the less they are affected by major mergers because small galaxies are more gaseous in such a way that the mergers are actually wet; therefore, the redshift limit can be lowered at smaller masses, increasing the fraction of low-mass halos hosting blue central galaxies. Furthermore, it is important to keep in mind that color is just a first order approach to morphology; some small galaxies are actually seen as blue spheroids in the local universe.

On the other hand, we obtain that a few fraction of massive halos can host blue central galaxies. This fraction is actually constrained by the observed color distribution of massive galaxies in the YMB09 catalog, which shows that there are indeed a few very massive blue centrals (see Fig. 48 in the Appendix E). However, in More et al. [146] the authors applied a stricter color cut in such a way that there are not blue centrals in halos larger than $\sim 10^{13}$M$_\odot$. Therefore, the determination of the fraction of halos at a given mass hosting blue/red central galaxies is actually somewhat subjective, depending on the criterion used to separate galaxies into blues and reds.

8.6 CONCLUSIONS

By means of a statistical model that combines the AMT, the HOD model, and the CSMF formalism, we have constrained the SHMRs of blue and red central galaxies starting from the global central SHMR and its scatter independently inferred first. For constraining the parameters of the model, we use the observed global GSMF and the projected 2PCFs of blue and red galaxies, as well as the color–M_* distributions of central galaxies. Since the model connects all these observational data with the distinct ΛCDM halo mass function, one obtains self-consistently (1) the SHMR of blue and red central galaxies, (2) the fraction of halos hosting blue/red central galaxies, and (3) the fraction of satellite galaxies, that are blue/red as a function of both M_* and $M_{\rm h}$. Our main results are briefly outlined as follows.

- The fraction of halos hosting blue central galaxies significantly decreases with mass from its maximum ($f_{\rm blue} \sim 0.8$) attained at $M_{\rm h} \approx 10^{11.4}{\rm M}_\odot$. This result agrees qualitatively with the proposal in Rodríguez-Puebla et al. [177, Chapter 7] that blue (late-type) galaxies do not live in halos that (1) suffered late major mergers and (2) that host classical groups richer than 3 members. If the shape of the red (blue) GSMF at the low-mass end is confirmed to be as steep (flat) as in YMB09, then $f_{\rm blue} \sim 0.8$ decreases strongly for masses lower than $M_{\rm h} \approx 10^{11.4}{\rm M}_\odot$.
- The SHMR of central blue and red galaxies do not differ significantly between them, suggesting that the efficiency of galaxy stellar mass growth, $f_{\rm s}$, is roughly the same for blue and red galaxies but it strongly changes with $M_{\rm h}$. The constrained blue/red SHMRs agree by construction not only with the observed GSMFs of blue and red galaxies, but also with the observed 2PCFs of blue and red galaxies.
- The obtained CSMFs of blue and red satellites as a function of $M_{\rm h}$ agree with observations.

The obtained SHMRs for central blue and red galaxies rather than independent inferences, are underlying relations of the model, which proved to offer a good description of the observed global GSMF and its decomposition into blue and red components, the observed projected 2PCFs of blue and red galaxies, the color–M_* distributions, and the occupational statistics of blue and red satellites. The almost lack of segregation of the M_*–$M_{\rm h}$ relation of central galaxies by color points out to a large dominion of the scale of the halos over other properties, e.g. the MAH, in determining the efficiency of galaxy stellar mass growth.

ACKNOWLEDGMENTS

We are grateful to S. More for sending us in electronic form their data for the fraction of halos hosting blue central galaxies.

Part VI

IMPLICATIONS OF THE BARYON-TO-HALO MASS RELATION ON THE STRUCTURAL AND DYNAMICAL PROPERTIES OF LOCAL GALAXIES

El estudio de la influencia de la fracción barionica en función de la masa del halo, $f_b(M_h)$, inferida en el Capítulo 7, sobre las relaciones dinámicas y estructurales de las galaxias de disco es abordado en esta Parte VI. En el Capítulo 9, se presenta un enfoque analítico para explorar cómo dependen dichas relaciones de las propiedades del halo (factor "cosmológico") tal como su masa, parámetro de giro λ_h y concentración, y del factor "astrofísico" dado por la eficiencia de crecimiento bariónico de las galaxias en función de M_h, justamente la fracción $f_b(M_h)$. Las conclusiones obtenidas son importantes para interpretar los resultados del siguiente Capítulo. En el Capitulo 10 se presenta un enfoque más completo del modelo estático, en el cual se siembran galaxias de disco en equilibrio centrífugo en halos de materia oscura y se resuelven las propiedades del sistema final, dados los factores "cosmológicos" y "astrofísico", a través de iteraciones. En este caso se toman en cuenta las distribuciones de los parámetros que describen los factores de entrada del modelo, de tal manera que las relaciones dinámico-estructurales obtenidas tienen sus correspondientes dispersiones. Nuestros resultados permiten evaluar las consecuencias de la relación $f_b(M_h)$ recientemente constreñida, en particular para galaxias azules (Rodríguez et al. 2011), identificadas en primera aproximación como galaxias dominadas por discos. Las relaciones de Tully-Fisher estelar y bariónica y sus dispersiones se mantienen cercanas a las observaciones, sin presentar además el problema del punto cero (normalización) reportado en algunos trabajos anteriores. No obstante, las relaciones radio-masa (estelar o bariónica) y cociente de velocidades máximas del disco a total vs. masa resultan ser sensibles a la forma de la relación $f_b(M_h)$. Para estar de acuerdo con la relación radio-masa, las galaxias más masivas deben tener un parámetro de giro menor que el de sus halos. Los discos modelados aumentan el dominio del disco a medida que la densidad superficial se incrementa pero nunca llegan a ser "discos máximos".

El Capítulo 10 contiene material parcialmente presentado en la memoria: *"Seeding the local disk galaxy population", Rodríguez-Puebla, Avila-Reese, Colín & Firmani 2011, RevMexAA (CS), vol. 40, p. 84*

9

INFLUENCE OF THE BARYON MASS FRACTION ON THE DISK GALAXY STRUCTURAL/DYNAMICAL CORRELATIONS. I. AN ANALYTICAL APPROACH

ABSTRACT

By means of an analytical approach, the effects of cosmological/astrophysical initial conditions on the structural-dynamical properties of baryonic disks formed inside CDM halos are explored. Our main goal is to understand the role that the galaxy baryon mass fraction, f_b=M_b/M_h (M_h is the virial halo mass), plays in the main disk galaxy correlations. The recently determined f_b–M_h relation for central blue (disk-dominated) galaxies (Chapter 7) changes strongly with M_h, having a bell shape with a peak at $M_h \sim 10^{12}$ $M_\odot$, with a value much smaller than the universal baryon mass fraction. We find that the f_b–M_h relation almost does not influence the shape and normalization of the $V_{t,max}$–M_b (baryonic Tully-Fisher) relation, where $V_{t,max}$ is the maximum circular velocity of the disk + contracted halo system, but it strongly influences the radius–mass and disk-to-total circular velocity ratio–mass relations. The predicted relations for the ΛCDM cosmology are consistent with observations, with the exception of a deviation to larger radii in the radius–mass relation. This issue suggests that the angular momentum conservation should be relaxed in such a way that for the massive galaxies, some angular momentum transfer result in disks with smaller spin parameters than those of the halos, hence the disks end with smaller radii.

9.1 INTRODUCTION

Several properties and correlations of the disk galaxy population have been studied by means of analytic (static) and semi-numeric (evolutionary) models in the cosmological context [e.g., 144, 10, 82, 218, 255, 80, 96, 203, 79, 85, and reference therein]. Typically, these models use as input parameters he total halo (dark + baryonic) mass, M_h, the baryon-to-halo mass ratio ($f_b \equiv M_b/M_h$, where M_b is the total galaxy mass in baryons), the halo spin parameter λ_h, and the halo NFW concentration parameter, c, or in more detail, the whole halo mass aggregation history, MAH. While λ_h, c or the halo MAH as a function of M_h are provided by cosmological simulations (for the popular Λ Cold Dark Matter cosmology, ΛCDM), the galaxy baryon fraction $f_b(M_h)$ has been treated as a *free parameter*, commonly assumed to be constant or decreasing as M_h decreases.

The actual normalization and shape of the f_b–M_h relation, including its scatter, is a product of the complex physics of baryons in interaction with the growing CDM halos. The processes of gas capture by the halos, its transformation into stars, its ejection due to the stellar-driven (mainly supernova) and AGN-driven feedback and its eventual re-accretion, are expected to depend strongly on halo mass, epoch, and local galaxy properties [see for recent reviews e.g., 143, 23]. These processes are still poorly under-

stood and challenge the current deductive (ab initio) models and numerical simulations of galaxy evolution.

On the other hand, based on large galaxy surveys, direct or indirect determinations of the stellar-to-halo mass ratio ($f_s \equiv M_s/M_h$, where M_s is the galaxy stellar mass) as a function of mass have started to be determined routinely (see previous Chapters of this Thesis and the references therein). The shape of the local f_s–M_h relation is such that it increases strongly from small masses, attaining a maximum at $M_h \sim 10^{12}$ $M_\odot$, and then it decreases with M_h. The value of f_s at the maximum is constrained to be around 0.025–0.035, that is, much lower than the universal baryonic mass fraction, $f_{b,U} \equiv \Omega_b/\Omega_M \approx 0.16$. The f_s fraction is associated to the efficiency of galaxy stellar mass assembly. The shape of f_b as a function of M_h at $z \sim 0$ seems to be qualitatively similar to the one of f_s [16, 177, 159], for masses large than $M_h \sim 5 \times 10^{11}$ $M_\odot$, both f_b and f_s are on average actually close, having approximately the same maximal values, and for smaller masses the former becomes systematically higher as smaller is M_s. This latter behavior is related to the known fact that smaller galaxies have that as the smaller are the galaxies, the higher are their gas higher gas fractions, i.e., lower mass galaxies are less efficient in transforming gas into stars.

Here we are interested in exploring the influence of the "input" galaxy baryon fraction $f_b(M_h)$ on the structural-dynamical correlations of disk galaxies. Therefore, it is important to use a $f_b(M_h)$ function constrained for this kind of galaxies. In Rodríguez-Puebla et al. [177, Chapter 7], by means of the Abundance Matching Technique, the f_s and f_b as a function of M_h were constrained for blue and red galaxies separately, including their scatters. In a first approximation, blue galaxies can be associated mostly to disk-dominated galaxies. In fact, both relations are not too different for blue and red galaxies. In any case, in this work we will use the main f_b–M_h relation obtained in Rodríguez-Puebla et al. [177] for blue galaxies.

For the explorative purpose in which we are interested in this Chapter, we require a simple and transparent approach rather than a more complex model. Here we use a static (non-evolutionary) approach for seeding disks in CDM halos based on Mo, Mao & White [144]. In order to avoid iterative procedures and obtain analytical results for the main correlations (not taking into account scatters), the model presented in this Chapter represents a simplistic version of Mo, Mao & White [144] model. Is important to note that in our analysis refers only to "baryonic" disks (stars+ gas) not affected by mergers or secular processes; that is, we assume that once the exponential disk is formed in centrifugal equilibrium inside the CDM halo, it remains as a disk and no difference between gas and stars is made. In Chapter 10 we will present results for a complete iterative model that takes into account the scatters in the input relations/parameters and that models crudely the fractions of stars and gas. The results obtained here will allow to understand those of the complete iterative model.

This Chapter is organized as follows. In section 9.2 we introduce a simple model in which the approximate analytical formulations for the scaling relations of disk galaxies are fully characterize by the parameters f_b, λ_h and c, by assuming that (1) the total maximum circular velocity of the disk+halo system is reached at the most inner regions, (2) dark matter particles move in circular orbits and (3) the baryonic scaling relations of disk galaxies are a crude representation of the stellar disk scaling relations. Using

these analytical relations, in section 9.3 we explore at a *qualitative level* the impact that the baryon-to-halo mass ratio, f_b, has on the scaling relations of disk galaxies. Finally, in section 9.4 we present our conclusions.

Except when we stated otherwise, we assume a ΛCDM cosmology with cosmological parameters $\Omega_M = 0.27$, $\Omega_M = 0.73$, $h = 0.70$, and a Chabrier [48] IMF.

9.2 THE ANALYTICAL APPROACH

The main trends of the scaling relations and other correlations of disk galaxies can be described in terms of a few cosmological "input" relations and parameters, mainly f_b, λ_h and c as a function of M_h [see e.g., 144, 10, 82, 80, 12, and reference therein]. In the following, an approach, based on Mo, Mao & White [144] but expressed in analytical terms in order to simplify on the dependences, is presented.

Since our goal is to explore how the galaxy baryon mass fraction as a function of M_h influences the disk galaxy main correlations at a qualitative level, we will study three different cases for $f_b(M_h)$:

1. f_b=0.03 (a constant f_b and much smaller than $f_{b,U}$ has been commonly assumed in early static and evolutionary semi-numerical models)

2. $f_b = 0.03(M_h/10^{11.5}h^{-1}M_\odot)^{0.3}$ (the smaller the galaxies, the more dark dominated they are; in order to catch this empirical fact for low-mass galaxies, a baryon mass fraction decreasing with mass has been proposed, see for example Dutton et al. [80]);

3. the mean $f_b(M_h)$ relation constrained by means of the abundance matching technique for blue (disk-dominated) galaxies in Rodríguez-Puebla et al. [177, Chapter 7][1] (for this more realistic case, f_b rapidly increases with M_h up to a maximum and then decreases for larger masses).

Figure 43 shows a comparison between these three cases of galaxy baryon-to-halo mass ratio.

The key aspects of the disk galaxy model in which our analytical approach is based are summarized as follows:

1. Initially, the baryonic gas and the dark matter are uniformly and well mixed in a unique and virialized object described by a Navarro, Frenk & White [151] (NFW) mass density profile, characterized by a total (baryonic + dark) viral mass and radius, M_h and r_{vir}, and a concentration, c.

2. The (baryonic) disk forms in the centre of the halo with a mass that is a fraction f_b of the halo mass, M_b=$f_b M_h$;

1 In that work, we have used the blue and red Galaxy Stellar Mass Functions of central galaxies determined by Yang, Mo & van den Bosch [247] from the SDSS DR5. The Halo Mass Function (HMF) corresponding to blue galaxies was calculated from the total ΛCDM (distinct) HMF but (i) by excluding those halos that suffered major mergers after $z = 0.8$ (theirs disks likely will be destroyed and not regenerated by $z \sim 0$), and (ii) by subtracting the observed HMF of bounded groups with 3 or more members (late-type galaxies are rare as central objects in groups/clusters).

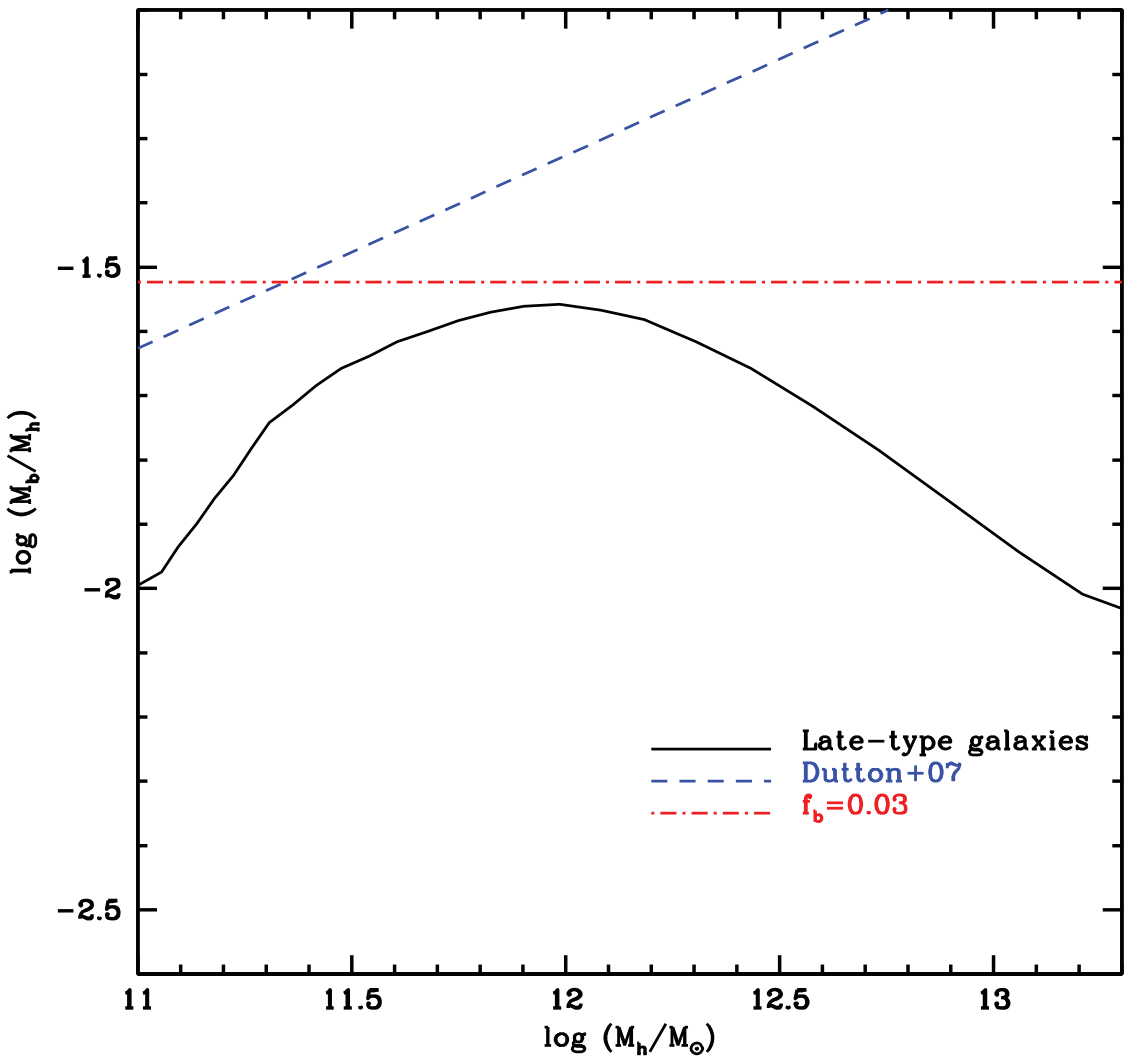

Figure 43: The baryon-to-halo mass ratio as a function of halo mass, $f_{\rm b}(M_{\rm h})$, for the three different cases studied in this Chapter. The red dotted short-dashed shows a constant $f_{\rm b} = 0.03$, while the blue short-dashed line shows a power-law form proposed in Dutton et al. [80]. The $f_{\rm b}$ constrained in Chapter 9 by using the central galaxy stellar mass function from the SDSS DR4 is represented with the black solid line.

3. The angular momentum of the disk is also a fraction of the one of the halo, i.e., $j_d = J_d/J_h$. It follows that the connection between the disk and halo spin parameter is: $\lambda_d = \lambda_h(j_d/f_{\rm b})$. If the detailed angular momentum is conserved, then λ_d=λ_h.

4. The disk is a thin and rotationally supported structure (centrifugal equilibrium) with an exponential density mass profile, $\Sigma_d(r) = \Sigma_0 e^{-r/R_d} = \frac{M_{\rm b}}{2\pi R_d^2} e^{-r/R_d}$.

5. The disk scale radius R_d is determined by the condition of centrifugal equilibrium in the gravitational potential (described by the total circular velocity curve) produced by the disk and the halo, and with the disk angular momentum j_d.

6. The accumulation of baryonic mass in the centre in order to form the disk produces a contraction of the inner halo, in such a way that the centrifugal equilibrium is reached in a gravitational potential different to the simple sum of those produced by the pristine NFW halo and the exponential disk.

In order to keep analytical the full description of the disk/halo mass and circular velocity profiles, the following extra assumptions should be done: (1) the total maximum circular velocity (disk+halo), $V_{\rm t,max}$, is reached at radii much smaller than $r_{\rm vir}$; (2) the dark matter particles in the spherical halo move in circular orbits and the halo contraction proceeds under adiabatic conditions; and (3) detailed angular momentum conservation is obeyed, i.e., λ_d=λ_h.

9.2.1 *The main equations*

In the disk + halo system, the total circular velocity is the sum in quadratures of circular velocities of the disk and the (contracted) halo:

$$V_t^2(r) = V_{\rm d}^2(r) + V_{\rm h}^2(r). \tag{97}$$

In this section our aim is to estimate the maximum circular velocity of the final system (disk + contracted halo), $V_{\rm t,max}$, which will be necessary for calculating the baryonic Tully-Fisher relation (bTFR)[2] in §§9.2.2. Because in the central regions the disk tends to dominate, one expects that the total circular velocity maximum should be also close to $2.2R_{\rm d}$, while for the most dark dominated systems (typically the smaller ones), the maximum is shifted to larger radii. In that direction, let us assume that $V_{\rm t,max} \approx V_t(2.2R_{\rm d})$. Therefore, under our assumption two ingredients are needed, $V_{\rm t,max}^2 \approx V_{\rm t}^2(2.2R_{\rm d}) = V_{\rm d}^2(2.2R_{\rm d}) + V_{\rm h}^2(2.2R_{\rm d})$. For the reminde of this section our strategy will focus in obtaining separately the analytical expression for $V_{\rm h}^2(2.2R_{\rm d})$ and $V_{\rm d}^2(2.2R_{\rm d})$.

2 Historically, the TFR is defined as a relation between the luminosity (or stellar mass) and the maximal rotation velocity, $V_{\rm t,max}$. However, from a theoretical point of view, it is more correct to express $V_{\rm t,max}$ as a function of luminosity (or mass; see for an extensive discussion Avila-Reese et al. [12]). This is called the inverse TFR. In order to avoid this term all the time, we rename here the TFR as the relation between mass (baryonic or stellar) and $V_{\rm t,max}$.

9.2.1.1 *The analytical model*

The *initial* (before contraction) halo, of virial radius $r_{\rm vir}$ and mass $M_{\rm h}$, has a spherical NFW mass density profile characterized by the concentration parameter c:

$$M_{\rm h}(r_i) = M_{\rm h} \times g(c, x), \tag{98}$$

where $x = r_i/r_{\rm vir}$, $c = r_{\rm vir}/r_0$ (r_0 is a scale radius of the NFW density profile), and

$$g(c, x) = \frac{\ln(1+cx) - cx/(1+cx)}{\ln(1+c) - c/(1+c)}. \tag{99}$$

From the spherical collapse model, the virial radius is defined as:

$$r_{\rm vir} = \left(\frac{3M_{\rm h}}{4\pi\Delta\bar{\rho}_M}\right)^{1/3} = \gamma \times {M_{\rm h}}^{1/3}, \tag{100}$$

where $\Delta \approx 340$ for the ΛCDM cosmology assumed here, and $\gamma = (3/4\pi\Delta\rho_m)^{1/3}$. The circular velocity profile of the NFW profile is given by $V_{\rm h}^{\rm in}(r_i) = \sqrt{GM_{\rm h}(r_i)/r_i}$. This profile has a maximum at $r_i = 2.16r_0$ and its value at this radius can be calculated from eqs. (98), (99), and (100):

$$V_{\rm h,max}^{\rm in} = \sqrt{\frac{Gcg(c, 2.16r_0)}{2.16\gamma}}{M_{\rm h}}^{1/3}. \tag{101}$$

The mass and circular velocity profiles of the halo, after the gravitational drag of the disk, $M_{\rm h}(r)$ and $V_h(r)$, can be obtained by mapping the initial mean radius r_i of particles in the NFW halo to the mean radius r at which these particles will end after the disk formation and halo contraction. This mapping can be quantified through the *contraction factor* $\omega = r_i/r$. It is usual to assume adiabatic invariance in the contraction process, then the radial angular momentum of particles is conserved (Blumenthal et al. [33]). From this condition:

$$M_{\rm h}(r_i)r_i = M_{\rm b}(r)r + M_{\rm h}(r_i)(1 - f_{\rm b})r. \tag{102}$$

For the inner regions of the disk + halo system (within a few disk scale radii) and using $r_i = \omega r$, it is reasonable to assume that:

$$c\omega(r/r_{\rm vir}) << 1. \tag{103}$$

For this case:

$$g(c, \omega r/r_{\rm vir}) = \frac{c^2(\omega r/r_{\rm vir})^2}{\ln(1+c) - c/(1+c)} = \sigma(c)(\omega r/r_{\rm vir})^2. \tag{104}$$

Assuming that $r = bR_{\rm d}$ and by means of eqs. (98) and (102), the contraction factor ω is given by:

$$\omega^3 - \omega^2(1 - f_{\rm b}) - \frac{[1 - (1+b)e^{-b}]f_{\rm b}}{\sigma(c)b^2}\left(\frac{r_{\rm vir}}{R_{\rm d}}\right)^2 = 0. \tag{105}$$

By means of the intermediate value theorem we can say that ω has at least one real solution, furthermore, it is easy to show, by using the discriminant of the cubic equation, that ω has one real root and two non-real complex conjugate roots, where the real one is the one of interest in our case. The latter means that the circular velocity of the halo after the gravitational drag can be written as:

$$V_{\rm h}^2(r) = G\frac{M_{\rm h}(r_i \rightarrow r)}{r}(1-f_{\rm b}) = G\frac{M_{\rm h}(\omega r)}{r}(1-f_{\rm b}). \tag{106}$$

As mentioned earlier, the total circular velocity of the disk + halo system is the sum in quadratures of circular velocities of the disk and the *contracted* halo:

$$V_t^2(r) = V_{\rm d}^2(r) + V_{\rm h}^2(r), \tag{107}$$

For a thin exponential disk:

$$V_{\rm d}^2(r) = 4\pi G\Sigma_0 R_{\rm d} y^2[I_0(y)K_0(y) - I_1(y)K_1(y)], \tag{108}$$

where $y = r/2R_{\rm d}$, and $I_{0,1}$ and $K_{0,1}$ are modified Bessel functions of the first and second kind respectively (Freeman [88]). The mapping of r_i to r (which is expressed through the contraction factor $\omega = r_i/r$) is performed iteratively by calculating the scale radius $R_{\rm d}$ of the exponential disk in centrifugal equilibrium in the total gravitational potential field defined by eq. (107). *Since here we are interested in an (approximate) analytical description, the iteration is avoided*. In the approximation that the total circular velocity maximum is close to $2.2R_{\rm d}$, the expression for the disk baryon component is very simple;

$$V_{\rm d}^2(2.2R_{\rm d}) = \alpha\frac{M_{\rm b}}{R_{\rm d}}, \tag{109}$$

where $\alpha = 9.68GB_{0,0}^{1,1}$ and $B_{0,0}^{1,1} = [I_0(1.1)K_0(1.1) - I_1(1.1)K_1(1.1)]$. In contrast, the expression for $V_{\rm h}^2(2.2R_{\rm d})$ of the contracted dark matter halo is more complex;

$$V_{\rm h}^2(2.2R_{\rm d}) = G\frac{M_{\rm h}g(c, 2.2\omega R_{\rm d}/r_{\rm vir})}{2.2R_{\rm d}}(1-f_{\rm b}). \tag{110}$$

Here the mapping from r_i to $r = 2.2R_{\rm d}$ due to halo contraction is contained in ω (see equation (105), for the case $b = 2.2$). Again, in the inner regions of the disk + halo system (within a few disk scale radii), it is reasonable to assume that $c\omega(R_{\rm d}/r_{\rm vir}) << 1$ and as in the case of eq. (104):

$$g(c, 2.2\omega R_{\rm d}/r_{\rm vir}) = \frac{4.84c^2\omega^2(R_{\rm d}/r_{\rm vir})^2}{\ln(1+c) - c/(1+c)}. \tag{111}$$

At this moment our analytical expressions for both $V_{\rm d}^2(2.2R_{\rm d})$ and $V_{\rm h}^2(2.2R_{\rm d})$ depend on scale length radius of the baryon disk, $R_{\rm d}$. Under the assumption of angular momentum conservation and centrifugal equilibrium, this radius can be estimated as follows;

$$R_{\rm d} = \frac{1}{\sqrt{2}}f_R(\lambda_h, c, f_{\rm b})f(c)\lambda_h r_{\rm vir}, \tag{112}$$

where $f_R(\lambda_h, c, f_b; R_d)$ (hereafter f_R) is a function of the input parameters λ_h, c, and f_b, and of the same value of R_d, which takes into account the gravitational effects (including the halo adiabatic contraction) of the disk. Thus, f_R is the one that should be found by an iterative procedure (see Mo, Mao & White [144] and next Chapter for more details). Since f_R, for most of the cases, has values of order unity, we assume it to be a constant close to 1. $f(c)$ is a function of the dark matter halo concentration that appears in the calculation of the total energy of a NFW profile truncated at the viral radius $r_{\rm vir} = cr_0$ (see Mo, Mao & White [144]). By introducing eq. (100) into eq. (112) one obtains:

$$R_{\rm d} = \Gamma f(c) \lambda_h ({f_{\rm b}}^{-1} M_{\rm b})^{1/3}, \tag{113}$$

where $\Gamma = (1/\sqrt{2}) \times f_R \times \gamma$. Thus, *the disk scale radius R_d becomes defined only by the initial cosmological/astrophysical parameters*. Equation (113) will be used in §§9.2.3 for estimating the radius-mass relation. It is easy to find also that the disk central surface density is given by:

$$\Sigma_0 = \frac{M_b^{1/3} f_b^{2/3}}{\pi [\Gamma f(c) \lambda_h]^2}. \tag{114}$$

Having a model to estimate $R_{\rm d}$, we can now obtain that the disk component is given by (see eqs. 109 and 113, and $M_{\rm b}$=$2\pi\Sigma_0 R_s^2$)

$$V_{\rm d}^2(2.2R_{\rm d}) = \alpha \frac{M_{\rm b}}{R_{\rm d}} = \frac{\alpha}{\Gamma f(c)} \frac{{M_{\rm b}}^{2/3} {f_{\rm b}}^{1/3}}{\lambda_h}, \tag{115}$$

while for the contracted dark matter halo (see Eqs. 110, 112 and 111) we get:

$$V_{\rm h}^2(2.2R_{\rm d}) = \frac{\beta}{\Gamma f(c)} \frac{{M_{\rm b}}^{2/3}}{{f_{\rm b}}^{2/3}} \lambda_h (1 - f_{\rm b}), \tag{116}$$

where $\beta = 2.2 G \omega^2 c^2 [\Gamma f(c)/\gamma]^2 / [\ln(1+c) - c/(1+c)]$. Thus, the total circular velocity at $2.2R_{\rm d}$ is given by eqs. (115) and (116):

$$V_t^2(2.2R_{\rm d}) = \frac{\lambda_h}{\Gamma f(c)} \frac{{M_{\rm b}}^{2/3}}{{f_{\rm b}}^{2/3}} \left[\alpha \frac{f_{\rm b}}{\lambda_h^2} + \beta (1 - f_{\rm b}) \right]. \tag{117}$$

We may further compare the maximum final circular velocity ($V_{\rm t,max} \approx V_t(2.2R_{\rm d})$) with the maximum circular velocity of the initial CDM halo (the cosmological one). The ratio of both quantities for a given galaxy is given by eqs. (117) and (101):

$$\frac{V_{\rm t,max}}{V_{\rm h,max}^{\rm in}} \approx \left[\frac{2.16 \alpha \gamma}{\Gamma f(c) G c g(c, 2.16 r_0)} \left(\frac{f_{\rm b}}{\lambda_h} + (1 - f_{\rm b}) \lambda_h \frac{\beta}{\alpha} \right) \right]^{1/2}. \tag{118}$$

The parameter c and the functions f_R, $f(c)$, and $g(c, 2.16r_0)$ are expected to depend very weakly on $M_{\rm h}$ (and hence on $M_{\rm b}$). Therefore, the relevant connection between the final (disk + contracted halo) $V_{\rm t,max}$ and the initial halo $V_{\rm h,max}^{\rm in}$, connection that will be used to estimate the bTFR in §§9.2.2 , is given by:

$$V_{\rm t,max} \approx K \times V_{\rm h,max}^{\rm in} \left(\frac{f_{\rm b}}{\lambda_h} + (1 - f_{\rm b}) \lambda_h \frac{\beta}{\alpha} \right)^{1/2}, \tag{119}$$

where K includes all the constants and terms independent or poorly dependent on mass. The N-body simulations provide us with the "cosmological" TFR, i.e., the relation between $V_{\rm h,max}$ and $M_{\rm h}$, where $M_{\rm h}$=$M_{\rm b}/f_{\rm b}$.

We are also interested in the local disk-to-total circular velocity ratio at $2.2R_{\rm d}$ (see §§9.2.3). From eqs. (115) and (117) we obtain that:

$$\frac{V_{\rm d}(2.2R_{\rm d})}{V_t(2.2R_{\rm d})} = \left[1+\lambda_h^2\frac{(1-f_{\rm b})}{f_{\rm b}}\frac{\beta}{\alpha}\right]^{-1/2}. \tag{120}$$

For a given value of c and λ_h, the term β/α in eqs. (119) and (125) depends mainly on ω the contraction factor of the halo at $2.2R_{\rm d}$ (eq. 105). As mentioned above, the formal procedure is to map r from r_i by an iterative procedure. Here, in order to get the analytical approximate equations, let assume in equation (112) that $f_R = 1$ and write the concentration parameter as $c = c(M_{\rm h}) = c({f_{\rm b}}^{-1}M_{\rm b})$, and then insert the result in eq. (105) (for the functionality of $f(c)$ see Mo, Mao & White [144]). The last step is to fix $b = 2.2$ and then you should be able to solve eq. (105). For radii around 2-3$R_{\rm d}$ and low values of $f_{\rm b}$, the values of ω are $\sim 1.2-3$. Note that when $\omega = 1$ there is no adiabatic contraction, and even more, if $0 < \omega < 1$, then the dark matter halo expands.

9.2.2 *The Tully-Fisher relation*

Using Eq. (118), we can now obtain the relation between $V_{\rm t,max}$ and the baryon mass $M_{\rm b}$, i.e., the bTFR. It is well known that $V^{\rm in}_{\rm h,max}$ correlates with $M_{\rm h}$ (the "cosmological" TFR) as $M_h^{a'}$ (from N-body cosmological simulations, $a' \approx 0.31$ with a small scatter, e.g., Avila-Reese et al. [11, 7]). This is actually close to the simple case of a spherical collapse into an isothermal sphere, $V^{\rm in}_{\rm h,max} \propto {M_{\rm h}}^{1/3}$, where $V_{\rm h}(r)$ is actually constant. Subsituting this into eq. (9.2.2) and taking into account that $M_{\rm h}$=$M_{\rm b}/f_{\rm b}$, we obtain that the bTFR in terms of $f_{\rm b}$ is;

$$V_{\rm t,max} \approx K \times \left[\frac{{f_{\rm b}}^{1/3}}{\lambda_h}+\frac{(1-f_{\rm b})}{{f_{\rm b}}^{2/3}}\lambda_h\frac{\beta}{\alpha}\right]^{1/2}{M_{\rm b}}^{1/3}. \tag{121}$$

Results from numerical simulations show that the halo spin parameter λ_h has a broad lognormal distribution almost independent on halo mass and weakly dependent on environment (e.g., Avila-Reese et al. [7], Bett et al. [29]). Therefore, it is expected that the distribution of λ_h introduces an intrinsic scatter in the bTFR. In this section we are interested only on the average trends with mass, for that reason we will fix $\lambda_h = 0.035$, which is close to the median value found in large numerical simulations (e.g., Bett et al. [29]). Similarly, for the concentration parameter c as a function of $M_{\rm h}$ we use the fit to a large cosmological simulation reported in Muñoz-Cuartas et al. [150].

Most of the analytical and semi-numeric works mentioned on the introduction have shown that the zero-point (normalization) and slope of the TFR relation, $V_{\rm t,max} = A{M_{\rm b}}^a$, are not too sensitive to $f_{\rm b}$, in particular if the adiabatic halo contraction due to disk gravitational drag is taken into account. However, for a strong and peculiar dependence of $f_{\rm b}$ on mass, as the one inferred in Rodríguez-Puebla et al. [177, see fig. 43] for blue galaxies, some curvature in the bTFR is expected: for masses much higher (lower) than

the mass at the peak, $M_{\rm b}(M_{\rm h} \sim 10^{12}{\rm M}_{\odot}) \approx 10^{10.4}{\rm M}_{\odot}$, the slope a should be steeper (shallower) than the average one at intermediate masses. From the toy model the slope of the bTFR is given by

$$a = \frac{d\log V_{\rm t,max}}{d\log M_{\rm b}} \sim \frac{1}{3} - \left(\frac{1}{3} - \frac{1}{2}\eta\right) f_{\rm b}{}', \tag{122}$$

where, $f_{\rm b}{}' = d\log f_{\rm b}/d\log M_{\rm b}$, and

$$\eta = \frac{1 - \frac{\beta}{\alpha}\lambda_h^2}{1 + \frac{\beta}{\alpha}\lambda_h^2 \frac{(1-f_{\rm b})}{f_{\rm b}}}. \tag{123}$$

For simplicity, we assume the ratio β/α as constant. Note that for the typical values of $f_{\rm b}$, c and λ_h, the term inside of the parenthesis in eq. (122) is always positive, while the values of $f_{\rm b}{}'$ could change from positive to negative, see Fig. 43. Immediately, one can see that for a peaked $f_{\rm b}$ relation, at low masses $f_{\rm b}{}' > 0$, hence $a = a_+ < 1/3$; at high masses, $f_{\rm b}{}' < 0$, hence $a = a_- > 1/3$; and at the peak mass, $f_{\rm b}{}' = 0$, hence $a = a_{peak} \sim 1/3$. Therefore, for the $f_{\rm b}(M_{\rm h})$ relation of blue (disk-dominated) galaxies used here, for which $f_{\rm b}{}'(M_{\rm h} < 10^{12}{\rm M}_{\odot}) > f_{\rm b}{}'(M_{\rm h} \sim 10^{12}{\rm M}_{\odot}) > f_{\rm b}{}'(M_{\rm h} > 10^{12}{\rm M}_{\odot})$, one should expect that $a_+ < a_{peak} < a_-$. Therefore, a bend in the bTFR at low and high masses is expected. For a constant baryon fraction, the slope of the bTFR is $\sim 1/3$ at all masses, a reminiscence of the "cosmological" TFR.

9.2.3 *The radius–mass relation*

According to eqs. (112) and (113), the disk scale radius correlates with $M_{\rm h}$ as $R_{\rm d} \propto \lambda_h f(c) r_{\rm vir} \propto \lambda_h f(c) {M_{\rm h}}^{1/3}$, where c depends weakly on $M_{\rm h}$. In fact, the dependence of $f(c)$ on $M_{\rm h}$ is very weak, such that it can be considered as constant. As mentioned, numerical simulations have shown that on average λ_h does not depend on $M_{\rm h}$ [e.g., 29]. Then the radius–mass relation is expected to scale with $f_{\rm b}$ as $R_{\rm d} \propto {M_{\rm h}}^{1/3} \propto (M_{\rm b} {f_{\rm b}}^{-1})^{1/3}$ with a systematical shift that depends on λ_h. The peculiar and relatively pronounced dependence of $R_{\rm d}$ on $f_{\rm b}$ implies some curvature in the $R_{\rm d} - M_{\rm b}$ relation. The slope of this relation is

$$\frac{d\log R_{\rm d}}{d\log M_{\rm b}} \sim \frac{1}{3} - \frac{1}{3} f_{\rm b}{}'. \tag{124}$$

Note that if the slope of the relation $f_{\rm b}$–$M_{\rm h}$ is zero, then $d\log R_{\rm d}/d\log M_{\rm b} = 1/3$. This happens at the maximum of this ratio. For masses below the maximum, $f_{\rm b}{}' > 0$, hence the slope becomes shallower, while for masses above the maximum, $f_{\rm b}{}' < 0$, hence the slope becomes steeper.

9.2.4 *Disk-to-total velocity ratio at $2.2R_{\rm d}$ vs mass*

The ratio of the the disk-to-total circular velocities at $2.2R_{\rm d}$, $V_{\rm d,2.2}/V_{\rm t,2.2}$, characterizes the inner dynamics of disk galaxies and it depends on the baryon-to-dark matter ratio inside optical radii. According to our toy model this is given as:

$$\frac{V_{\rm d}(2.2R_{\rm d})}{V_t(2.2R_{\rm d})} = \left[1+\lambda_h^2\frac{(1-f_{\rm b})}{f_{\rm b}}\frac{\beta}{\alpha}\right]^{-1/2}. \tag{125}$$

Note that the number on brackets is always positive. Moreover, if $f_{\rm b} = f_{\rm b,U} \approx 0.16 < 1$ then $V_{\rm d,2.2}/V_{\rm t,2.2} < 1$. On the order of magnitude, $\lambda_h \sim 10^{-2}$, $f_{\rm b} << 1$ and $\beta/\alpha \sim 10^2$, we then obtain that $\lambda_h^2(1-f_{\rm b})\beta/\alpha \sim \lambda_h$, and

$$\left[1+\lambda_h^2\frac{(1-f_{\rm b})}{f_{\rm b}}\frac{\beta}{\alpha}\right]^{-1/2} \sim (1+\lambda_h/f_{\rm b})^{-1/2}. \tag{126}$$

Therefore, at a first order the slope of the $V_{\rm d,2.2}/V_{\rm t,2.2}$ ratio is,

$$\frac{d\log}{d\log M_{\rm b}}\frac{V_{\rm d,2.2}}{V_{\rm t,2.2}} \propto \frac{\lambda_h/f_{\rm b}}{1+\lambda_h/f_{\rm b}}\frac{df_{\rm b}}{d\log M_{\rm b}} \sim \frac{df_{\rm b}}{d\log M_{\rm b}}, \tag{127}$$

where we use that $f_{\rm b} \sim \lambda_h$.

9.3 RESULTS: IMPLICATIONS FOR GALAXY SCALING RELATIONS

Based on the analytical approach described in the previous Section, we calculate the disk baryonic $V_{\rm t,max}$–$M_{\rm b}$, radius–$M_{\rm b}$, and $V_{\rm d,2.2}/V_{\rm t,2.2}$–$M_{\rm b}$ relations for the three different $f_{\rm b}(M_{\rm h})$ dependences showed in Fig. 43. In all the cases, we use the median values of the CDM halo distributions corresponding to λ_h (=const.=0.035) and $c(M_{\rm h})$. The results are plotted in Fig. 44 for (1) $f_{\rm b}$=0.03 (dot-dashed red line), (2) $f_{\rm b} = 0.03(M_{\rm h}/10^{11.5}h^{-1}M_\odot)^{0.3}$ (dashed blue line), and (3) the Rodríguez-Puebla et al. [177] $f_{\rm b}(M_{\rm h})$ dependence for blue galaxies (black solid line). Additionally, for the case (3), we present results for a lower value of the spin parameter, $\lambda_h = 0.02$ (long-dashed black line).

The panel (a) of Fig. 44 shows the bTFR. The relations for all the cases are relatively similar despite the quite different shapes of $f_{\rm b}(M_{\rm h})$. As expected (see §§9.2.2), for the peaked $f_{\rm b}(M_{\rm h})$ form inferred semi-empirically, the bTFR slope becomes slightly steeper and shallower at masses much higher and lower than the peak mass ($M_{\rm s} \approx 10^{10.4}$ M$_\odot$), respectively (see eq. 122). For intermediate masses, the slope $a \approx 1/3$ as well as the normalization are in good agreement with observational inferences of the bTFR (see e.g., Avila-Reese et al. [12]). For the case of a lower λ_h value (=0.02), there is not an appreciable shift of the normalization neither on the bend of TFR. The trends found for the bTFR are expected to be also valid for the stellar and (infrared) luminous TFRs.

There are some hints that the stellar or infrared TFR has a curvature in the predicted sense at lower masses/luminosities [e.g., 137, 31, 139]. At large masses, the observations are too scarce as to know whether there is or a bend or not.

In panel (b) of Fig. (44), $R_{\rm d}$ vs $M_{\rm b}$ is plotted. We observe that the $R_{\rm d} - M_{\rm b}$ relation is much more dependent on $f_{\rm b}$ than the bTFR. For the semi-empiric $f_{\rm b}(M_{\rm h})$ bell-shaped

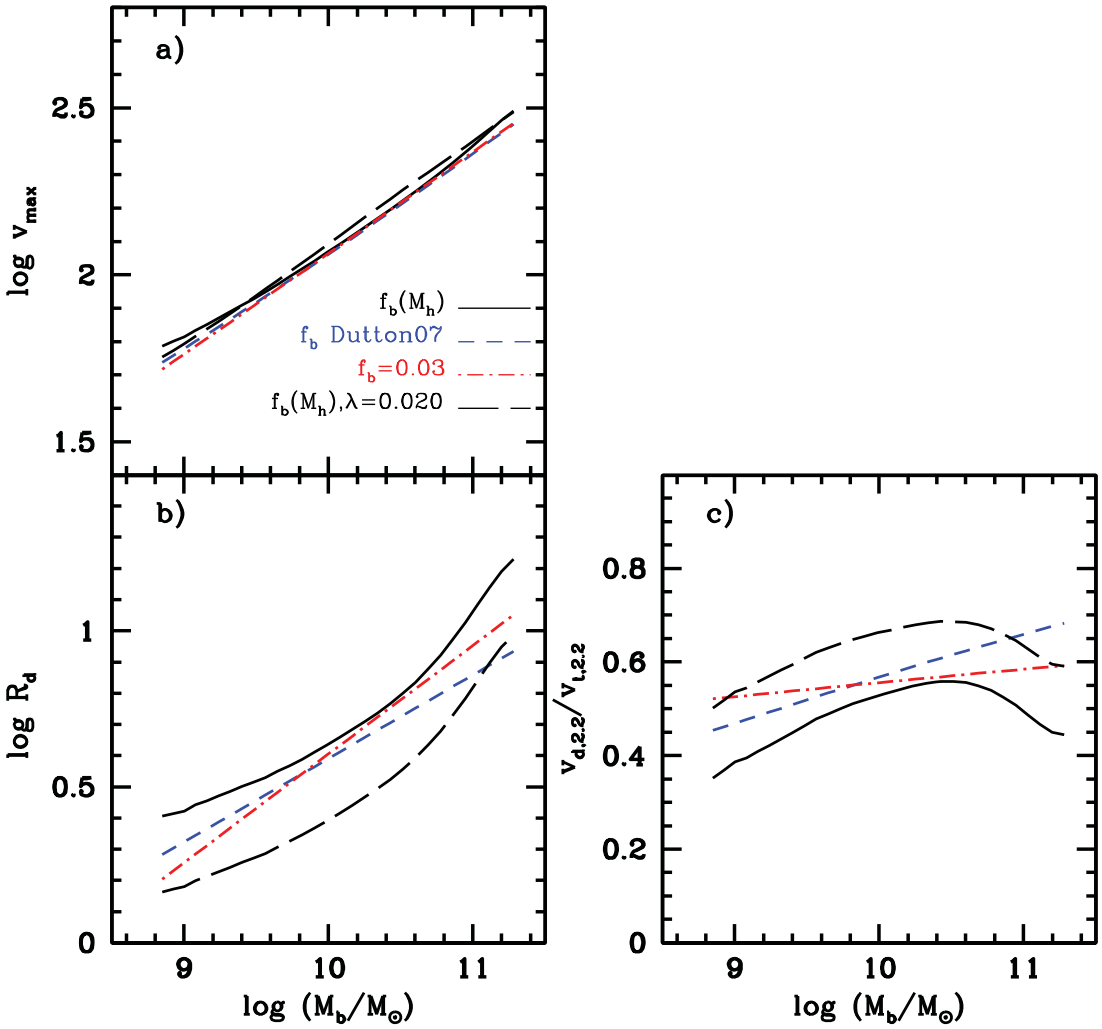

Figure 44: Dynamical and structural scaling relations for disk galaxies. In all the panels the solid, short-dashed and dotted short-dashed lines show the results obtained when using the Chapter 9 $f_{\rm b}(M_{\rm h})$, the power-law form of $f_{\rm b}(M_{\rm h})$ and $f_{\rm b} = 0.03$ and in all cases a parameter spin $\lambda_h = 0.035$. The long-dashed line show the results when using the former case of $f_{\rm b}(M_{\rm h})$ and a lower spin value $\lambda_h = 0.02$. *Panel a)*: The bTFR. This relation is not significantly influenced by the shape of $f_{\rm b}$. Nevertheless, as expected from the peaked Chapter 9 $f_{\rm b}(M_{\rm h})$ relation, the bTFR is slightly steeper at masses below $10^{10.4}{\rm M}_\odot$, above this masses it is shallower. *Panel b)*: The $R_{\rm d}$-$M_{\rm b}$ relation. Observe how the $f_{\rm b}$ significantly influences the shape of the $R_{\rm d}$-$M_{\rm b}$ relation, while the spin parameter λ_h affects only the normalization. *Panel c:* The disk-to-total circular velocity ratio at $2.2R_{\rm d}$. Apparently, the shape of this ratio reassembles the shape of $f_{\rm b}$ for all the cases probed here. All galaxies tend to be submaximum disks even when the spin parameter is lowered to a value of $\lambda_h = 0.020$.

dependence, the $R_{\rm d} - M_{\rm b}$ (average) relation is significantly curved, with a shallower slope at low masses and much steeper slope at the largest masses, something expected according to eq. (124). Therefore, the shape of the $f_{\rm b}(M_{\rm h})$ relation significantly influences the shape of the $R_{\rm d} - M_{\rm b}$ relation of disk galaxies. A strong effect of λ_h on the normalization of this relation is also seen (it seems that the overall shape of the $R_{\rm d}$–$M_{\rm b}$ relation does not change significantly with λ_h). Since the disk central surface density, Σ_0, for a given $M_{\rm b}$, strongly anti-correlates with λ_h (see eq. 114), then the scatter around the normalization of the mean $R_{\rm d} - M_{\rm b}$ relation correlates strongly with Σ_d, in such a way that high (low) surface brightness galaxies (low and high values of λ_h, respectively) lie below (above) the mean relation. Observations confirm this segregation of galaxies in the $R_{\rm d} - M_{\rm b}$ relation [12].

Is the radius–mass relation of observed disk galaxies as curved in Fig. 44? The observational studies refer mainly to the stellar or luminous radius–mass(luminosity) relations. It is expected that from the baryonic to stellar or infrared luminosity radius–mass(luminosity) relations there are not significant changes in the shape. For large samples from the SDSS, it was reported that the stellar radius–mass relation is indeed shallower at low masses and steeper at high masses Shen et al. [189], Dutton et al. [78]. However, the change in the slope at high masses is not as pronounced as plotted in Fig. 44 for the $f_{\rm b}(M_{\rm b})$ bell-shaped relation. Since for lower values of λ_h, the radius is smaller, it could be that the disk spin parameter, λ_h, departs from the halo one towards lower values as more massive are the galaxies. If this is the case, then the slope of the radius–mass relation at large masses would become shallower than in the case of $\lambda_d = \lambda_h$ independent of mass. This opens the possibility that the detailed angular momentum conservation in massive disk galaxies is not obeyed, $\lesssim\lambda_h$.

In panel (c) of Fig. 44 we plot the ratio of the the disk-to-total circular velocity ratio at $2.2R_{\rm d}$, $V_{\rm d,2.2}/V_{\rm t,2.2}$, as a function of $M_{\rm b}$. As can be seen, the $V_{\rm d,2.2}/V_{\rm t,2.2}$ ratio is strongly dependent on $f_{\rm b}$. For the case $f_{\rm b}$=const., this ratio almost does not depend on mass (dot-dashed red line). For $f_{\rm b}$ increasing monotonically with mass, $V_{\rm d,2.2}/V_{\rm t,2.2}$ increases with $M_{\rm b}$ (dashed blue line). However, for the semi-empirical $f_{\rm b}(M_{\rm h})$ relation inferred for blue central galaxies, this ratio becomes a strong function of $M_{\rm b}$ (solid black line), showing a peak at $M_{\rm b} \approx 10^{10.4}$ $M_\odot$. Therefore, the shape of the $f_{\rm b}$–$M_{\rm h}$ relation remains imprinted in the dependence of the disk-to-total circular velocity ratio on mass.

For $\lambda_h = 0.035$, the predicted disks are significantly sub-maximumal, i.e., strongly dark-matter dominated, even those with masses at the peak. For $\lambda_h = 0.020$, the velocity ratio at the peak barely implies the maximum-disk case ($V_{\rm d,2.2}/V_{\rm t,2.2} > 0.75$). There are several pieces of evidence that low-mass disks are deeply sub-maximum (strongly dark-matter dominated), and as the mass increases or the Hubble type is earlier, they can become of maximum-disk type (e.g., Zavala et al. [255], Herrmann & Ciardullo [106], for a review see van der Kruit & Freeman [221], and more references therein). Some authors find that all galaxies, even the massive ones, are in the sub-maximum regime with values of $V_{\rm d,2.2}/V_{\rm t,2.2} = 0.63 \pm 0.10$, Bottema [35], $= 0.57 \pm 0.22$, [116] or $= 0.53 \pm 0.15$, [117], but see Sackett [179] where $V_{\rm d,2.2}/V_{\rm t,2.2} = 0.85 \pm 0.10$ for massive galaxies. More recently, by means of integral field stellar and gas kinematics, Bershady et al. [27] find that the disk-to-total rotation velocity ratio increases with $V_{\rm t,max}$ (from $\sim$ 100 to 250 km/s) but does not overcome $\sim$ 0.7, i.e. galaxies are sub-maximal. The method used by these

authors breaks the disk-halo degeneracy by obtaining independent measures of the total dynamical mass and dynamical disk-mass surface density.

9.4 CONCLUSIONS

A static model of disk + halo galaxies based on Mo, Mao & White [144] has been used to infer analytic relations for the main structural and dynamical correlations of disk galaxies. These analytic relations are expressed in terms of the initial "cosmological" parameters, namely the virial halo mass, $M_{\rm h}$, its concentration c and spin parameter λ_h, and the astrophysical parameter $f_{\rm b}$ (the galaxy baryon-to-virial halo mass ratio). The obtained correlations for the (disk + contracted halo) systems refer actually to disk baryon quantities (mass, radius, etc.) and for exponential disks in which the effects of mergers or secular evolution on the structural/dynamical properties are neglected.

The aim of our analytical approach is to understand in a transparent way the effects of the initial conditions on the main structural/dynamical properties of disk galaxies, in the context of the ΛCDM cosmology. We were particularly interested in exploring the effects of the baryonic mass fraction and its dependence on mass, $f_{\rm b}(M_{\rm h})$. Recent semi-empirical determinations of this dependence for blue (disk-dominated) galaxies by Rodríguez-Puebla et al. [177, Chapter 6] show that $f_{\rm b}(M_{\rm h})$ is far from being constant or a simple power law (Fig. 43).

From the results of our simple analytical approach (resumed in Fig. 44), we have seen that:

- The bTFR ($V_{\rm t,max}$ vs $M_{\rm b}$) is a robust relation that is weakly affected by the cosmological input parameters, as well as by the shape and normalization of $f_{\rm b}(M_{\rm h})$. If any, the bell shape of the semi-empirical $f_{\rm b}$–$M_{\rm h}$ relation imprints a weak bend in the bTFR at low masses (slope shallower than $a \sim 1/3$) and a very weak bend at high masses (slope steeper than $a \sim 1/3$); the slope of $a \sim 1/3$ is a direct imprint of the "cosmological" TFR ($V^{\rm in}_{\rm h,max} - M_{\rm h}$). The normalization of the predicted bTFR for the ΛCDM cosmology is consistent with the observed bTFR; there is not a "zero-point TFR" issue.

- The baryonic $R_{\rm d}$–$M_{\rm b}$ relation depends significantly on $f_{\rm b}(M_{\rm h})$ and its scatter is determined mainly by the variation of λ_h; the latter implies that galaxies are segregated systematically in the radius–mass relation by their surface brightness. The bell shape of the semi-empirically inferred $f_{\rm b}$–$M_{\rm h}$ relation, translates into a strong curvature in the $R_{\rm d}$–$M_{\rm b}$ relation: at masses close to those corresponding to the maximum of $f_{\rm b}(M_{\rm h})$, the slope of the relation is $\sim 1/3$; at lower masses, the slope becomes systematically shallower, while at higher masses, it becomes systematically steeper. The predicted $R_{\rm d}$–$M_{\rm b}$ relation is actually consistent with observations, with the exception of the steepening at high masses; this suggests that the detailed angular momentum conservation should be relaxed at high masses in such a way that the disk spin parameter, λ_d, becomes systematically smaller than λ_h as more massive the galaxy is.

- The disk-to-total circular velocity ratio at $2.2R_{\rm d}$ as a function of $M_{\rm b}$ also depends on the shape and normalization of the $f_{\rm b}$–$M_{\rm h}$ relation. The $V_{\rm d,2.2}/V_{\rm t,2.2}$ ratio is a measure of the luminous-to-dark matter content inside the galaxies. Our results show that for the ΛCDM halos and the low values inferred for $f_{\rm b}$, galaxies are sub-maximal. Only for the cases when the spin parameter is very small, the disk can be concentrated enough as to be close to $V_{\rm d,2.2}/V_{\rm t,2.2} = 0.75$, the value above which a galaxy becomes maximum disk. The dependence of the $V_{\rm d,2.2}/V_{\rm t,2.2}$ ratio on $M_{\rm b}$ for the semi-empirical $f_{\rm b}$–$M_{\rm b}$ relation and $\lambda_h = 0.035$, is such that it decreases in low mass galaxies; a maximum value ($\approx 0.55 - 0.60$) is attained at the peak of the $f_{\rm b}$–$M_{\rm b}$ relation, and at larger masses it decreases again. If the angular momentum conservation is relaxed and λ_d systematically decreases as the mass increases, then, the $V_{\rm d,2.2}/V_{\rm t,2.2}$ ratio would continue increasing with $M_{\rm b}$.

Our analytical results allow to understand in a clear way how the initial conditions influence the disk galaxy properties. We have found that, in particular, the effects of the recently constrained $f_{\rm b}$–$M_{\rm b}$ relation for blue galaxies, are weak in the bTFR but significant in the $R_{\rm d}$–$M_{\rm b}$ and $V_{\rm d,2.2}/V_{\rm t,2.2}$–$M_{\rm b}$ relations. The results encourage us to use the complete static models with the iterative procedure and taking into account the intrinsic scatters of the initial parameters; $f_{\rm b}$, c and λ_h. The confrontation of the predicted structural-dynamical correlations with those observed, including the scatters, allows us to probe the underlaying cosmological model and constrain the relevant astrophysical processes of disk galaxy formation.

10

INFLUENCE OF THE BARYON MASS FRACTION ON THE DISK GALAXY STRUCTURAL/DYNAMICAL CORRELATIONS. II. A POPULATION STATIC MODEL AND EVOLUTIONARY CONSEQUENCES

ABSTRACT

Using models of disk galaxies seeded inside ΛCold Dark Matter (ΛCDM) halos, we study the structural and dynamical statistical properties of a whole population of disk galaxies at $z = 0$. We are able to map the halo (cosmological) and astrophysical initial conditions into the main correlations among these properties. These conditions (inputs) are the distributions of the halo concentration and spin parameter, λ_h, as a function of halo mass $M_{\rm h}$, and of the baryonic mass growth efficiency given by the $M_{\rm b}/M_{\rm h}$ ratio, as a function of $M_{\rm h}$. The former come from ΛCDM cosmological simulations, and the latter from the new semi-empirical inferences for blue central galaxies presented in Part V. Our models show that the population of disk galaxies formed inside CDM halos are in good agreement with the observed shape, normalization and scatter of both the stellar and baryonic TFRs; in particular, there is not a zero point problem. Our disk galaxy population is also in good agreement with the observed shape, normalization and scatter of both the stellar and baryonic mass-radius relations. However, this consistency is attained only when the spin parameter of the disks is lower than that of their host halos, $\lambda_d < \lambda_h$, in particular for the massive ones, $M_{\rm h} > 10^{12}{\rm M}_{\odot}$. Thus, we constrain semi-empirically the dependence of λ_d on $M_{\rm h}$. We find a weak anti-correlation between the residuals of the TFR and those of the mass-radius relation. Thus, disk galaxies have a significant contribution of dark matter inside their optical radii. Moreover, our analysis based on the the disk-to-total maximum circular velocity ratio (which characterize the contribution of the baryonic to the total matter in the inner parts of the disk), shows that disk galaxies are dark matter dominated, i.e., disk galaxies are in the sub-maximum regime, as some direct pieces of evidence suggest. Future studies in large spectroscopical galaxy surveys (e.g., MaNGA) will confirm or not this result.

10.1 INTRODUCTION

Disk galaxies are the most common type of galaxies in the local Universe. In the context of the popular Λ−Cold Dark Matter (ΛCDM) cosmology, disks are generic structures formed inside CDM halos from the gas trapped within them. Therefore, it is expected that the structural and dynamical properties of the galactic disks are tightly related (1) to the "cosmological" initial conditions imprinted in their halos (halo virial mass, radius, concentration and spin parameter, $M_{\rm h}$, $r_{\rm vir}$, c, and λ_h, respectively), and (2) to the "astrophysical" initial or boundary conditions, as for example, the galaxy baryon mass fraction, $f_{\rm b}$=$M_{\rm b}/M_{\rm h}$, where $M_{\rm b}$ is the disk baryonic mass. Note that $M_{\rm h}$ refers to the total (baryonic + dark) mass inside the virial radius.

A large effort has been done in understanding the formation and evolution of disk galaxies in the context of the ΛCDM cosmology by means of evolutionary models/simulations and static population models (see for references Chapter 9). The former follow the assembly and evolution of individual disk galaxies inside the growing CDM halos; using this approach it is difficult to obtain a whole population of galaxies. The latter approach is not evolutionary, but it allows to "seed" disk galaxies in a whole population of CDM halos at a given epoch, and explore in this way the correlations between the main structural and dynamical properties of galaxies.

The main question posed in this Chapter is whether the recently inferred $f_{\rm b}$–$M_{\rm h}$ relation and its scatter is consistent with the structural/dynamical correlations of local disk galaxies. As mentioned above, these properties are tightly related to $f_{\rm b}$. This relation has been recently constrained for both blue and red central galaxies by using the abundance matching technique in Rodríguez-Puebla et al. [177, Chapter 6, see also Chapter 7]. Here we use this $f_{\rm b}$ relation and we will assume that blue galaxies are mostly disk-dominated galaxies. The $f_{\rm b}$–$M_{\rm h}$ relation for blue central galaxies is strongly dependent on $M_{\rm h}$, bell shaped, with a maximum value at $M_{\rm h} \sim 10^{12}$ $M_\odot$ not exceeding $f_{\rm b} = 0.04 << f_{\rm U}$, and with a relatively small scatter (0.17 dex). The fact that $f_{\rm b}$ is much smaller than the universal baryon mass fraction, $f_{\rm U} = \Omega_b/\Omega_m$, implies that the efficiency of central galaxy formation is low, most likely due to the negative feedback effects of the same galaxy over the gas inside and around it (see the introduction of Chapter 9 for a short discussion).

In order to explore the question introduced above, we will use static models developed on Mo, Mao & White [144] and Dutton et al. [80]. In the previous Chapter, we showed (by means of simplistic analytical expressions based on this model) how different initial conditions affect the structural/dynamical correlations of the baryonic disks seeded in CDM halos. Here, the full iterative model will be applied, including the corresponding distributions (scatters) of the initial condition parameters. In this way, the scatters around the predicted mean relations will be obtained. The scatters around relations such as the Tully-Fisher and the radius-mass contain also valuable information for understanding galaxy formation. Our model, under some assumptions, allows to obtain also the stellar mass, radius, and surface density of the disks, as well as their gas fractions. The scaling relations and other correlations of disk galaxies are better studied observationally for stellar quantities rather than baryonic ones.

We will also discuss some results obtained with full evolutionary semi-numerical models Firmani & Avila-Reese [82], Firmani, Avila-Reese & Rodríguez-Puebla [85]. These models follow self-consistently the formation and evolution of disks in isolated growing CDM halos. The baryon and stellar mass fractions in these models, rather than an input parameter, are obtained as the result of the gas infall, star formation (SF) process, and the SF–driven feedback (mainly due to the Supernovae, SNe). The physics of the latter process can be formulated in several ways. We explore in which cases the SN–driven outflows are able to shape the $M_{\rm s}$–$M_{\rm h}$ and $M_{\rm b}$–$M_{\rm h}$ relations semi-empirically inferred at $z \sim 0$ and whether other properties, as the SF rate vs stellar mass $M_{\rm s}$ at $z \sim 0$ and other epochs, can be predicted at the same time in agreement with observations.

The essence of the static model is presented in Section 2; the procedure to generate a full grid of models and the input distributions to be used are described in §§2.3. The model results regarding the scaling relations are presented and compared with

observations on §§3.1, while in §§3.2, the model results related to the internal luminous-to-dynamical mass compositions of disk galaxies as a function mass, surface density and rotation velocity are given and discussed in the light of observational inferences. In Section 4 we briefly discuss the results of evolutionary models constrained to reproduce the local $M_{\rm s}$–$M_{\rm h}$ relation. Our conclusions are given in Section 5.

10.2 SEEDING DISK GALAXIES IN CDM HALOS: THE STATIC MODEL

In the static model of galaxy disks Mo, Mao & White [144], an exponential disk of mass $M_{\rm b}$=$f_{\rm b}M_{\rm h}$ is seeded inside a CDM halo with a NFW profile (Navarro, Frenk & White [151]) characterized by its virial mass $M_{\rm h}$, concentration parameter c, and spin parameter λ_h. The scale radius $R_{\rm d}$ of the exponential (baryonic) disk is calculated under the assumption that the gas falls within the gravitational potential of the system until it reaches centrifugal equilibrium, having the gas the same angular momentum of the halo (detailed angular momentum conservation, $\lambda_d = \lambda_h$). The gravitational potential used to calculate the centrifugal equilibrium is the composition of the potentials due to the halo and the disk. However, the formation of the disk in the centre of the halo pulls gravitationally the dark matter. As a result the dark matter halo contracts; this contraction is calculated under the assumption of adiabaticity. The new gravitational potential is used to calculate a new $R_{\rm d}$ and so on; an iterative procedure is implemented until $R_{\rm d}$ and the gravitational potential do not change (see details in Mo, Mao & White [144] and Dutton et al. [80]).

The main aspects, assumptions, and equations describing the physics of the model were presented in Section 2 of the previous Chapter. The initial conditions for the model are $M_{\rm h}$, c, λ_h, and $f_{\rm b}$. The first three ones are related to the cosmological CDM given by N-body simulations of large volumes. The concentration weakly correlates with $M_{\rm h}$ and it presents a wide lognormal scatter (e.g., Muñoz-Cuartas et al. [150]. The spin parameter does not correlate with $M_{\rm h}$ and it presents also a lognormal scatter (e.g., Bett et al. [29]). Regarding $f_{\rm b}$, this is a parameter related to the astrophysical processes of galaxy formation (see Section 2 in Chapter 9). We will use the semi-empirical determination of $f_{\rm b}$ as a function of $M_{\rm h}$ (including the scatter) for blue galaxies obtained in Chapter 6.

Following, we detail the new features with respect to Mo, Mao & White [144] in our disk galaxy model.

10.2.1 *Adiabatic Contraction*

In order to calculate the gravitational drag of baryon matter over the dark matter, the key assumption is that the process of disk formation happens slowly in such a way that the halo potential changes in timescales larger than the typical halo dynamical timescale. The classical way to calculate the adiabatic contraction of CDM halos was introduced by Blumenthal et al. [33], and it is what we used in Section 2.1 of the previous Chapter (see for details therein). The calculation assumes spherical symmetry, homologous contraction (no shell crossing), circular particle orbits, and conservation of the radial angular momentum, which implies that $M_h(r)r$ is an invariant, where $M_h(r)$ is the mass

contained inside the radius r. However, the particle orbits in the hierarchically growing CDM halos in cosmological simulations are actually eccentric. Using high-resolution hydrodynamics simulations, Gnedin et al. [95] found that the classical adiabatic contraction model systematically over-predicts the amounts of dark matter density in the inner 5% of the virial radius. In order to improve the classical model over this shortcoming, they have proposed a simple modification, namely to introduce a generalized adiabatic invariant:

$$M(\bar{r})r = \text{cte}, \tag{128}$$

where $M(\bar{r})$ is the mass enclosed within the orbit-averaged radius given by

$$\frac{\bar{r}}{r_{\text{vir}}} = (0.85 \pm 0.05)\left(\frac{r}{r_{\text{vir}}}\right)^{0.8\pm0.02}. \tag{129}$$

This modification approximately accounts for orbital eccentricities of particles and reproduces simulation profiles to within 10% - 20%. This new adiabatic invariant roughly captures the complex interaction between the baryons and the dark matter halo.

10.2.2 *Star Formation and stellar quantities*

The static model is actually for baryonic disks, no matter if they are gaseous or stellar, and it is assumed that the disks all the time have an exponential surface density distribution. Since the model is non-evolutionary, it is not possible to follow the whole process of gas transformation into stars as well as secular dynamical processes or mergers that could change the exponential distribution of the gas and stellar disk components. However, based on the fact that the most general conditions for star formation are related to the local disk surface density and dynamics, we may roughly estimate the disk regions where stars are expected to be formed. Thus, a rough estimate of the stellar mass M_{s} and stellar half-mass radius, $R_{1/2}$, can be obtained for our disks. To that end, we implement a model close to the one used in Dutton et al. [80].

We assume that the inner regions of the exponential disks that obey the local Toomre [212] criterion for gravitational instability in the gas disk are mostly transformed into stars, while those that do not obey this criterion remain as gaseous. The gas Toomre criterion is expressed in terms of the surface density, the instability being applied when $\Sigma_{crit}(r) < \Sigma_d(r)$, where

$$\Sigma_{crit}(r) = \frac{\sigma_{gas}(r)\kappa(r)}{3.36QG}, \tag{130}$$

and σ_{gas} is the gas velocity dispersion, $\kappa(r)$ is the epicyclic frequency at the radius r, and Q is the Toomre instability parameter. Interstellar medium observations in our Galaxy as well as in a few near spiral galaxies show that σ_{gas} is low, with measured values of 5-10 km/s typically (see e.g., Tamburro et al. [204]), and without a significant change with radius and among different galaxies. Therefore, we assume here a constant value of $\sigma_{gas} = 6$ km/s. Regarding the Q parameter, in the original Toomre [212] paper, by definition it is equal to 1. However, by studying the instabilities in galaxy disks modeled

in numerical simulations, where the disks have some vertical structure, values of $Q \approx 1.5 - 2.5$ are found. Here, we assume $Q = 1.5$.

In the regions where the local criterion for instability is fulfilled, we compute the stellar surface mass density as

$$\Sigma_{\rm s}(r) = \Sigma_{\rm b}(r) - \Sigma_{crit}(r). \tag{131}$$

By integrating this last equation, we obtain the stellar mass enclosed at a radius r,

$$M_{\rm s}(r) = 2\pi \int_0^r [\Sigma_{\rm b}(r) - \Sigma_{crit}(r)] r' dr'. \tag{132}$$

The total stellar mass, $M_{\rm s}$, is computed up to the radius where $\Sigma_{\rm b}(r) = \Sigma_{crit}(r)$. From eq. (132) we may calculate also the radius where half of the stellar mass is attained, $R_{1/2}$. Note that the stellar disk is now a component of the baryonic disk, i.e., $\Sigma_{\rm b}(r) = \Sigma_{\rm s}(r) + \Sigma_{\rm g}(r)$, and since $\Sigma_{\rm b}(r)$ has an exponential distribution, $\Sigma_{\rm s}(r)$ will likely deviate from an exponential distribution.

10.2.3 *The procedure*

In order to generate a catalog of mock disk galaxies and explore correlations and their scatters among the properties of the galaxies, we proceed as follows:

- To each halo of mass $M_{\rm h}$, a central galaxy with baryonic mass $M_{\rm b}$ is assigned randomly from a lognormal probability distribution $\mathrm{P}(M_{\rm b} \mid M_{\rm h})$. This distribution is defined by the mean $M_{\rm b} - M_{\rm h}$ relation and the width of the lognormal distribution, σ_b. We use the mean $M_{\rm b} - M_{\rm h}$ (or $f_{\rm b}$–$M_{\rm h}$) relation semi-empirically determined in Chapter 7 for central blue galaxies, with the scatter width $\sigma_b = 0.173$ dex.

- To each halo of mass $M_{\rm h}$, a concentration parameter is assigned randomly from a lognormal probability distribution $\mathrm{P}(c \mid M_{\rm h})$. This distribution is defined by the $c - M_{\rm h}$ relation and the width of the lognormal distribution σ_c. For this relation and the width of its scatter we use the fit to a large N-body cosmological simulation reported in Muñoz-Cuartas et al. [150].

- To each halo, a spin parameter is assigned randomly from a lognormal probability distribution $P(\lambda_h)$. The mean and width of the distribution are taken from the fits to the "Millenium" N-body cosmological simulation as reported in Bett et al. [29].

In the previous Chapter it was found that the radius–mass relation predicted by the static model steepens strongly at high halo masses, in potential conflict with observations. Since the radius is directly proportional to the initial spin parameter, λ_h, this suggests that for massive galaxies, some angular momentum is lost (transported) in such a way the final spin parameter, λ_d, is smaller and the radius is smaller. Interesting enough, by means of simple approaches for estimating from observations the galaxy spin parameter λ_d and for assigning a halo mass to the galaxies, some authors have found that λ_d decreases with $M_{\rm h}$ for massive galaxies (Jimenez, Verde & Oh [111], Cervantes-Sodi

et al. [47], Berta et al. [28]). This implies that the assumption of detailed angular momentum conservation for massive galaxies should be relaxed: in the formation process of these galaxies, the specific angular momentum of the gas is partially lost (transported). By means of N-body/hydrodynamic simulations indeed it was found that when the early assembly of a galaxy happens dominated by mergers, the gas trapped into the pre-galactic fragments looses efficiently its specific angular momentum because these fragments transfer their orbital angular momentum to the outer halo by tidal effects (Zavala, Okamoto & Frenk [256]). On the other hand, from the N-body cosmological simulations, it is well known that the merger rate increases as more massive are the halos (e.g., Zavala et al. [254]). Therefore, it is expected that for massive halos, the final spin parameter λ_d is systematically lower than the initial one in the halo, λ_h.

In view of the above discussion, for halos more massive than $M_{\rm h} = 10^{12}$ M$_\odot$, following the functionality found in Berta et al. [28], we assign a spin parameter taken from the same lognormal scatter given by the "Milleniumm" simulation but with a mean spin parameter given by $\log\lambda_d = a\log M_{\rm h} + b$. The values of a and b are fixed so that the observational stellar radius–mass relation is obtained, see below.

10.3 RESULTS

A grid of thousands of models has been generated according to the procedure described in §§10.2.3. We now, present the results regarding the stellar/baryonic scaling relations (§§10.3.1), and regarding the disk-to-total circular velocity ratio as a function of different disk galaxy properties.

10.3.1 *Scaling relations*

The upper panels of Fig. 45 show the z=0 model stellar and baryonic TF relations ($V_{\rm t,max}$ vs $M_{\rm s}$ and $M_{\rm b}$, respectively; gray shaded area, containing the 68% scatter at a given mass). Note that contrary to the previous Chapter, here (1) $V_{\rm t,max}$ is the actual maximum of the total (disk+ contracted halo) circular velocity profile, rather than this velocity at $2.2R_{\rm d}$; and (2) the masses are for both the stellar component and the total baryonic one. The black circles with error bars correspond to a homogenized sample of local high and low surface brightness (HSB and LSB) disk galaxies presented in Avila-Reese et al. [12]. The blue crosses are for data compiled in McGaugh [137].

Both the stellar and baryonic TFRs are well described by a power law that slightly starts to bend at lower masses. The slopes of the power laws ($\approx$ 0.28 and 0.29, respectively) are in excellent agreement with those fitted to observations in Avila-Reese et al. [12]. The data from McGaugh [137] imply shallower slopes. Note that McGaugh does not use $V_{\rm t,max}$ or a proxy to it, but the velocity where the rotation curves flattens. The larger the mass, the more concentrated the disks in such a way that their rotation curves are more peaked, and hence their maximum velocities are systematically larger than the velocities where the curve flattens (this is likely why the velocities of the McGaugh galaxies are smaller on average as the mass is larger with respect to those in the Avila-Reese et al. sample). McGaugh sample is also smaller and less complete in types and surface

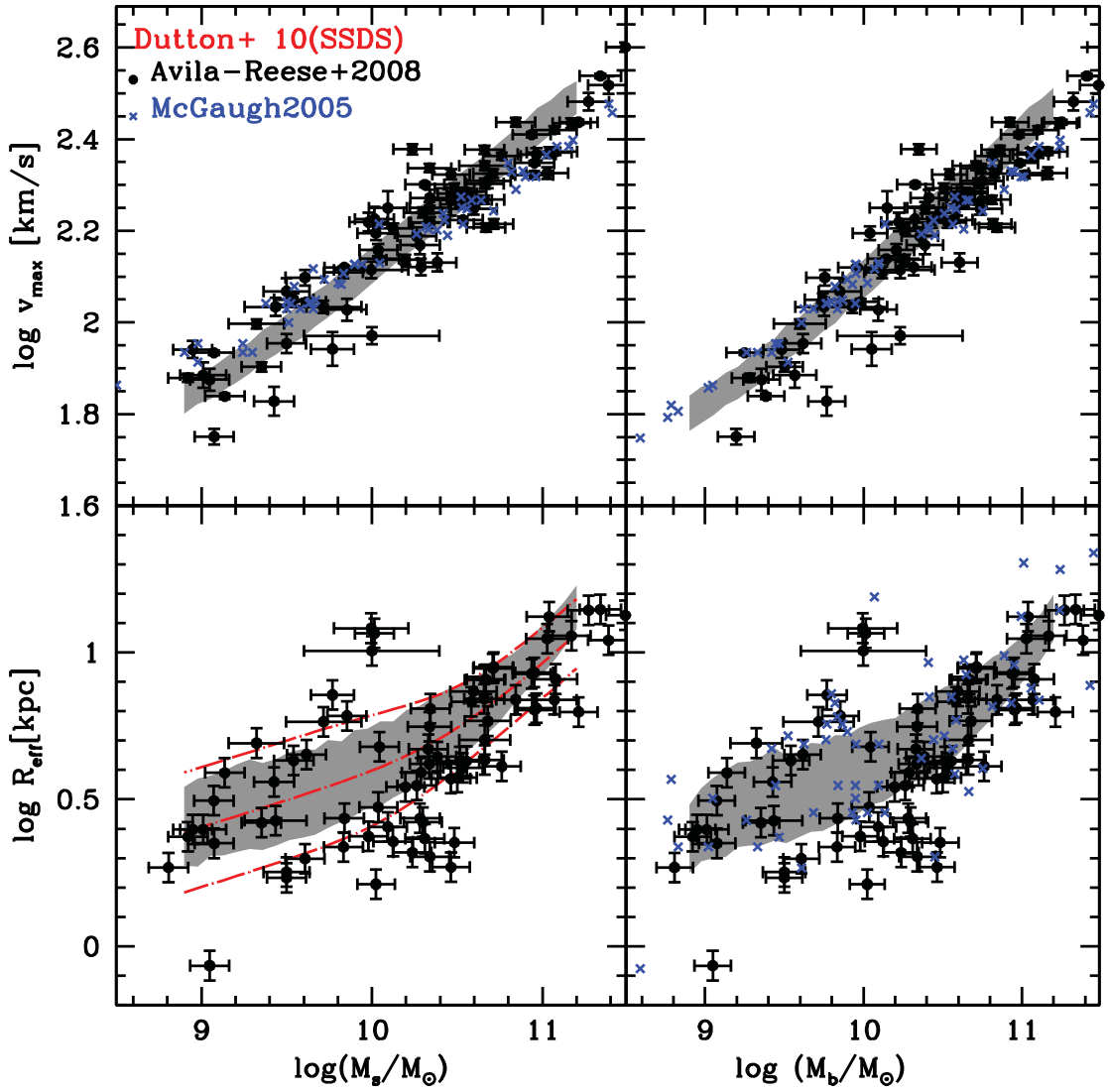

Figure 45: Predicted dependences of $V_{\rm m}$ and $R_{\rm eff}$ as a function of stellar mass (left panels) and baryon mass (right panels) by using the $f_{\rm b}(M_{\rm h})$ function inferred in Chapter 9 for blue central galaxies. The 1σ of the model population distributions is plotted with the dark-gray shaded regions. Observational data from Avila-Reese et al. [12] are plotted with black filled circles, inference reported in McGaugh [137] are plotted with the blue skeletal symbols while the 1σ of the distribution $R_{\rm eff}$ inferences obtained in Dutton et al. [78] from the SDSS DR are plotted with the red dotted long-dashed lines.

brightness values than the one compiled in Avila-Reese et al. In any case, for both samples, the zero points of the stellar and baryonic TFRs are similar and in good agreement with our model galaxies. Several previous semi-analytical and numerical works used to find that the zero-point of the predicted TFR is in conflict with observations in the sense that the ΛCDM-based predictions are shifted to larger $V_{t,max}$ than observations. We do not find this problem.

The slight bend of the TFRs at low masses is something that is also reported in observations. Unfortunately, the sample by Avila-Reese et al. is poor at masses below $M_s \sim 3 \times 10^9$ $M_\odot$ in such a way that this bend can not be seen, but works focused on dwarf galaxies show that their $V_{t,max}$ for a given M_s or luminosity are on average larger than those extrapolated from the TFR of massive (luminous) galaxies (e.g., McGaugh [137], De Rijcke et al. [70], Blanton, Geha & West [31], McGaugh et al. [139]).

Regarding the intrinsic scatter of the observational TFR (after subtracting the observational scatter), it is actually very small (see Avila-Reese et al. [12] and more references therein). Most of previous models and simulations used to predict scatters in the stellar TFR higher than in the observations (Eisenstein & Loeb [81], Firmani & Avila-Reese [82], Dutton et al. [80], Gnedin et al. [96], Avila-Reese et al. [12]). The intrinsic scatter of the stellar TFR obtained here is in better agreement with the observational one reported in Avila-Reese et al., thought yet it is slightly larger. The reason of the reduction of the scatter in our models with respect to previous works is mainly due to the low values of f_b and its dependence on M_h used as input here. Recall that the f_b–M_h relation and its scatter used here was inferred semi-empirically. It is known that the intrinsic scatter around the predicted TFR becomes smaller as lower is f_b (e.g., Firmani & Avila-Reese [82], Gnedin et al. [96]).

The lower panels of Fig. 45 show the z=0 model stellar and baryonic radius–mass relations (gray shaded area, containing the 68% scatter at a given mass). The radius here is the effective one (where half of the stellar or baryonic mass is contained). In the baryonic case, where the disk is by construction exponential, $R_{eff} = 1.68 \times R_d$. However, in the stellar case, the disk is not anymore exactly exponential and the effective radius in this case can be close to or smaller than the corresponding to an exact exponential disk; according to our disk instability algorithm of star formation, the stellar disk tends to be more concentrated than the total baryonic one. The solid circles with error bars are observations from the homogenized sample of Avila-Reese et al. [12, in this case, R_{eff} has been calculated as 1.68 times the R_d compiled by these authors]. The red dashed lines in the R_{eff}–M_s plane show the 16, 50 and 84% of the data from a large SDSS sample of blue galaxies reported in the Dutton et al. [78]. The blue crosses in the R_{eff}–M_b plane are the data from the McGaugh [137] compilation.

As concluded in the previous Chapter, the radius–mass relation suffers a curvature due to the strong curvature of the input f_b–M_h relation. The large observational sample of Dutton et al. [78] shows also such a curvature. However, we have shown in the previous Chapter that if $\lambda_d = \lambda_h$ (angular momentum conservation), then the radius–mass relation at the massive end is too steep, with radii too large as compared with observations. Therefore, as described in §§10.2.3, we lower systematically λ_d as more massive is the halo, following the semi-empirical results of Berta et al. [28]. This is why our pre-

dicted mass-radius relation agrees with observations towards higher masses. As a result of such an agreement, we find that:

$$log\lambda_d = -0.16 \times log(M_{\rm h}/10^{12}{\rm M}_\odot) + 0.3. \quad (133)$$

The normalization and scatter of the stellar and baryonic mass-radius relation are in good agreement with observations. The scatter in these relations is large, contrary to the case of the TFRs. As discussed in the previous Chapter, the scatter is mainly due the scatter in λ_h and hence it correlates strongly with the disk surface density (brightness); this is also seen in the observations (e.g., Zavala [253], Avila-Reese et al. [12]).

We also find a weak anti-correlation between the residuals of the TFR and those of the mass-radius relation, in particular for the baryonic case. The slopes of these anti-correlations are close to the ones of the observations [63, 253, 62, 12], which implies that real disk galaxies have a significant contribution of dark matter inside their optical radii.

10.3.2 *The baryonic-to-dark matter ratio in disk galaxies*

The $M_{\rm s}$–$M_{\rm h}$ or $M_{\rm b}$–$M_{\rm h}$ relations connect the masses of galaxies to those of the CDM halos as cosmological objects, which are much more extended than galaxies. From an observational point of view, the dynamical mass distribution (sum of baryonic and dark matter) is easier to obtain within the optical radii of galaxies. The static models presented here, by construction, offer the possibility to link both aspects of the galaxy/halo systems, i.e., their global galaxy-to-virial halo mass ratio and the galaxy-to-dynamical mass ratios at radii within the optical galaxy.

The dynamical mass, $M_{\rm dyn}$, is the sum of the disk and dark matter components at a given internal radius. In our models, where centrifugal equilibrium and circular motions are assumed, $M_{\rm dyn}(r) = V_t^2(r)r/G$. For a thin exponential disk, the ratio of the disk (baryonic) to total mass at a given radius is proportional to the square of the ratio of disk to total circular velocity at that radius, $M_{\rm b}/M_{\rm dyn} \propto (V_d/V_t)^2$. The disk circular velocity for a thin exponential disk can be calculated from the observed disk surface brightness (density) profile by using the Freeman [88] solution. Thus, the V_d/V_t ratio is in principle an observational quantity that traces the ratio of luminous (baryonic) to total (including dark matter) mass inside disk galaxies (see Zavala [253] for more explanations).

The V_d/V_t ratio can be defined for example at $2.2R_{\rm d}$ as in the previous Chapter. Since detailed rotation curves are typically not available, but proxies to their maxima are more frequently reported, we can use from our static full model $V_{\rm t,max}$, the total circular velocity at its maximum, which for disk dominated galaxies is attained at $\approx 2.2R_{\rm d}$, but for more dark matter dominated galaxies, is attained at larger radii. Instead, the maximum of the disk circular velocity $V_{\rm d,max}$, is by definition at $2.2R_{\rm d}$ and it can be calculated easily from the observable mass and scale radius $R_{\rm d}$ of the exponential disk Freeman [88]. Therefore, in order to compare with observations, we will report the $V_{\rm d,max}/V_{\rm t,max}$ ratio for our models.

Figure 46 shows the $V_{\rm d,max}/V_{\rm t,max}$ ratio vs the central stellar and baryonic surface densities (left upper and lower panels) and the stellar and baryonic masses (right upper and lower panels) for our model results (gray shaded area display the 68% scatter). The squares and solid circles are the observational inferences from the Zavala et al. [255]

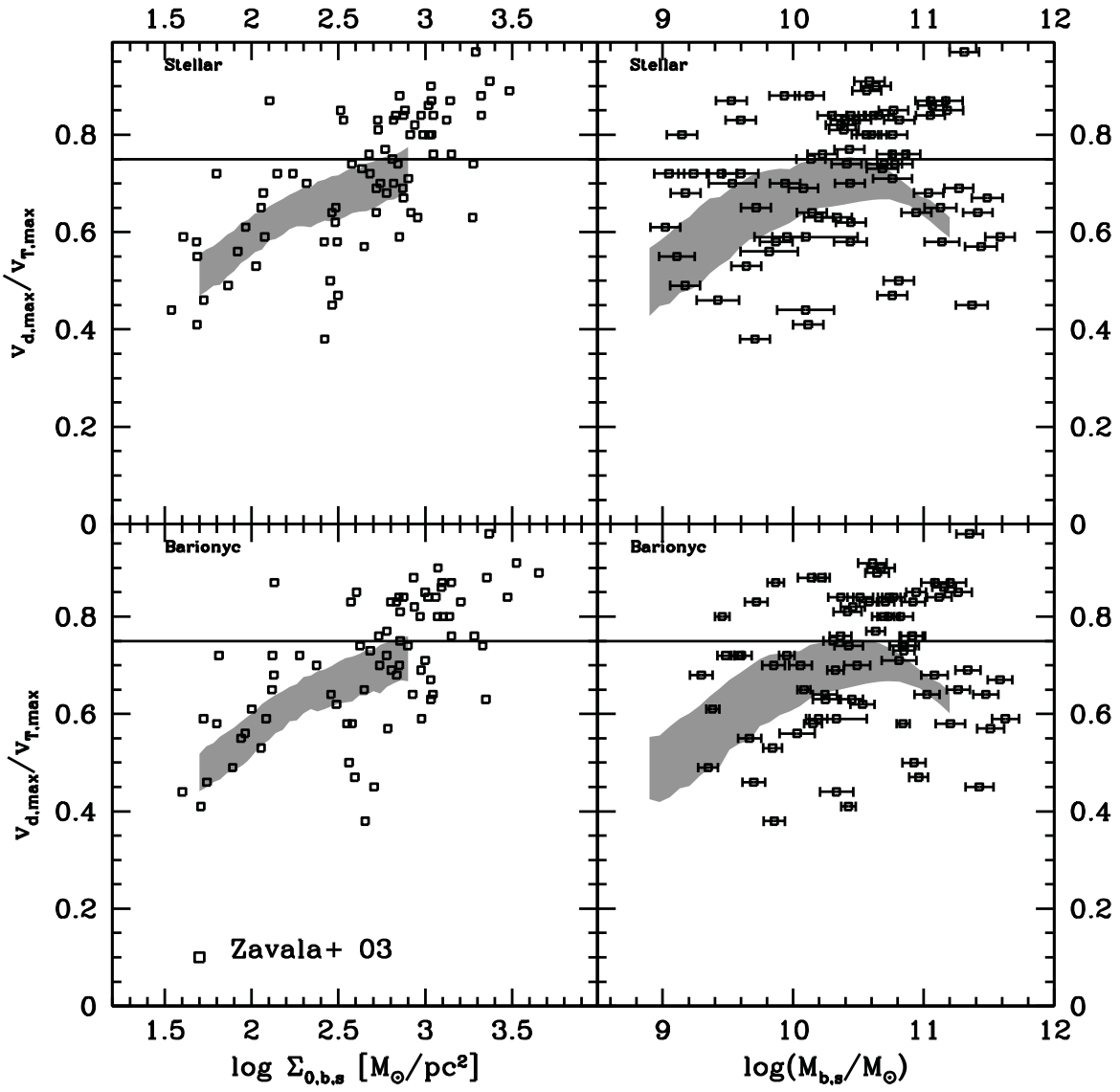

Figure 46: Predicted dependences of V_d-to-$V_{t,max}$ at the $2.2R_d$ R_{eff} as a function of stellar/baryon surface mass density (left panels) and stellar/baryon mass (right panels), by using the $f_b(M_h)$ function inferred in Chapter 9 for blue central galaxies. The 1σ of the model population distributions is plotted with the dark-gray shaded regions. Observational data from Avila-Reese et al. [12] are plotted with black filled circles, inference reported in Zavala et al. [255] are plotted with the open squares symbols.

and Avila-Reese et al. [12] sample. The horizontal solid line indicates the minimal value (0.75) above which a disk is maximum in the rotation curve decomposition Sackett [179], i.e., luminous matter dominates inside the radius at which the rotation curve has its maximum (e.g, see the recent review by van der Kruit & Freeman [221]).

As anticipated in the previous Chapter, the low baryon fractions used as input in the models imply that disk galaxies are not maximum disks, that is, at radii around the one where $V_{t,max}$ is attained, dark matter already dominates. The $V_{d,max}/V_{t,max}$ ratio decreases towards lower surface densities and masses, although with a large scatter in the latter case. Observations show qualitatively similar trends as shown in Fig. 46 (see also Herrmann & Ciardullo [106]). Both observations and our ΛCDM-based models show that the $V_{d,max}/V_{t,max}$ ratio is more dependent on the surface density of the disks than on their masses (see also Zavala et al. [255], who used f_b=0.05=const). The weak dependence on mass according to our models is such that for massive galaxies the $V_{d,max}/V_{t,max}$ ratio instead of continue increasing, decreases with mass. This is a clear footprint of the shape of the used f_b–M_h relation. Unfortunately, the level of precision of the current observational inferences and the low numbers do not allow yet to test this interesting feature.

The observational inferences plotted in Fig. 46 show that there are some galaxies with $V_{d,max}/V_{t,max} > 0.75$, something that does not happen for the ΛCDM-based models with the semi-empirical f_b–M_h relation; the latter are dark matter dominated, even for the highest surface density disks. If the adiabatic halo contraction implemented in our models would be weaker or even, instead of contraction, the inner halo regions would expand due to galaxy formation (feedback effects; Dutton et al. [80], then the $V_{d,max}/V_{t,max}$ could be higher than we obtained. The question whether disk galaxies can be maximal or not is highly debated from the observational point of view. As mentioned in the previous Chapter, some authors find that all galaxies are in the sub-maximum regime Bottema [35], Kregel, van der Kruit & de Grijs [116], Kregel, van der Kruit & Freeman [117], but see Sackett [179].

More recently, by means of integral field stellar and gas kinematics, Bershady et al. [27] find that the disk-to-total rotation velocity ratio increases with $V_{t,max}$ but it does not overcome ~ 0.7, i.e. galaxies are sub-maximal. The method used by these authors breaks the disk-halo degeneracy by obtaining independent measures of the total dynamical mass and dynamical disk-mass surface density. In Fig. 47, the $V_{d,max}/V_{t,max}$ ratio vs $V_{t,max}$ is plotted for our models and compared to the fit that Bershady et al. [27] provide to their 30 disk galaxies. The observational inferences imply that disk galaxies could be even more dark matter dominated than we are finding for the ΛCDM-based models and including adiabatic halo contraction.

10.4 CONCLUSIONS

A disk+halo static model based on Mo, Mao & White [144] and Dutton et al. [80] was used to probe whether the recently inferred f_b–M_h relation for blue galaxies and its scatter (see Chapter 9) is consistent with the structural/dynamical correlations of local disk galaxies. Contrary to the previous Chapter, where we used simplistic analytical expressions based on these models, here the full iterative model has been applied, including

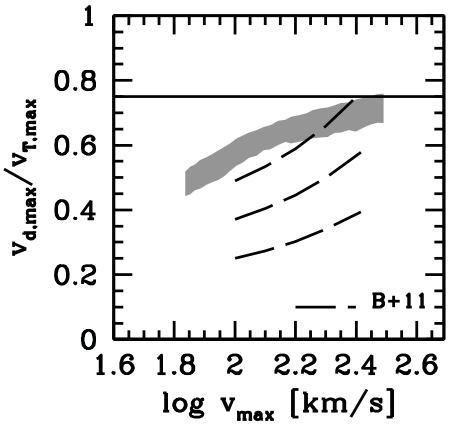

Figure 47: $V_{\rm d,max}/V_{\rm t,max}$ ratio as a function of $V_{\rm t,max}$. The resulting 1σ distribution predicted by the models is plotted by the shaded region, while the fitted relation (along with its uncertainty) obtained in Bershady et al. [27] is plotted with the long dashed lines.

the corresponding distributions (scatters) of the initial condition parameters. This has allowed us to study not only the mean of the structural/dynamical correlations but also their scatters, and compare directly with observations. for both stellar and baryonic quantities.

Our main findings are summarized as follows:

- The zero point and the slope of both the stellar and baryonic TFRs are in excellent agreement with observational inferences [e.g., 12]. This is in contrast to some previous ΛCDM-based semi-analytical and numerical works that find that the zero-point of the predicted TFR is in conflict with observations. While both TFRs are well described by a power law, in detail we find that it slightly starts to bend at lower masses, as well as at the largest masses. We also find that the scatter of the stellar TFR obtained here is in reasonable agreement with the observations.

- The shape, the normalization and the scatter of the stellar and baryonic mass-radius relation are in good agreement with observations [e.g., 78, 12] To attain this agreement between models and observations we have relaxed the assumption of detailed angular momentum conservation by lowering systematically the value of λ_d with respect to λ_h towards higher masses. Our results lead to a new semi-empirical inference of the λ_d dependence on mass. There is a weak anti-correlation between the residuals of the TFR and those of the mass-radius relation, which implies that disk galaxies have a significant contribution of dark matter inside their optical radii.

- Based on the disk-to-total maximum circular velocity ratio, we observe that models galaxies are not maximum disks (at the 1σ level), even for the most massive or highest density disks. This is because the $f_{\rm b}(M_{\rm h})$ relation used as an input in

the models has very low values, even at the peak of this relation. Therefore, on average, model disk galaxies are dark matter dominated at optical radii. This is consistent with recent observational reports that find that all galaxies are in the sub-maximum regime.

- The disk-to-total maximum circular velocity ratio correlates better with the disk surface density than with M_s. This ratio as a function of $V_{t,max}$ tends to increase with the surface density and $V_{t,max}$ but it does not overcome ~ 0.7, i.e. galaxies are sub-maximal. A comparison with observational inferences in Bershady et al. [27] suggest that disk galaxies could be even more dark matter dominated than what is found with our models.

The population of $z = 0$ disk galaxies generated with our model by using the semi-empirical f_b–M_h relation and its scatter for blue central galaxies constrained in Chapter 9, is in good agreement with the structural/dynamical correlations of disk galaxies. In the particular case of the radius-mass relation, we had to relax the assumption of detailed angular momentum conservation, specially for large-mass disks. This implies that during the formation and secular evolution of galaxies, some processes of angular momentum transport play an important role in massive galaxies. These processes are likely also behind the formation of a bulge and a supermassive black hole in its center. We will report results by including bulges in our disk galaxies elsewhere.

The fact that the baryon fraction, used as an input in our models, is very low, inevitably implies that on average disk galaxies are dark dominated on their inner parts, i.e., galaxies are not maximum disks. While observationally there is an ample debate on whether or not disk galaxies are maximum the evidences points that the majority of the disk are in the sub-maximum regime Bottema [35], Kregel, van der Kruit & de Grijs [116], Kregel, van der Kruit & Freeman [117], Bershady et al. [27]. Future analysis with high resolution spectroscopy galaxy surveys, (e.g., CALIFA, MaNGA), would shed light on whether galaxies are sub-maximum disks as our model predicts.

Finally, we should say that when evolutionary models of disk galaxies are used, in order to reproduce the local M_s–M_h relation that we have constrained along this Thesis, strong feedback effects should be introduced. For galaxies in halos smaller than $\sim 10^{12}$ $M_\odot$, the main feedback mechanism invoked in the literature is the one of SN-driven outflows. There are several types of SN-driven feedback, e.g., by energy conservation or by momentum conservation. In Firmani, Avila-Reese & Rodríguez-Puebla [85], we have used a semi-numerical model of disk galaxy evolution and introduced different types of SN-driven outflows in order to reproduce the obtained M_s–M_h relation for blue galaxies (Part V) in halos smaller than $\sim 10^{12}$ $M_\odot$. In that work, we have found that the peculiar shape and normalization of the low-mass M_s–M_h relation can be reproduced when strong momentum conservation outflows are used, taking into account that some gas should be re-accreted (without re-accretion, the energy conservation strong outflows are better suited to reproduce the M_s–M_h relation).

However, in Firmani, Avila-Reese & Rodríguez-Puebla [85] we have found that in spite that the local M_s–M_h relation of blue galaxies can be reproduced by introducing very strong feedback-driven outflows, it is impossible to reproduce the observed specific star formation rate (SSFR) vs M_s relation. Besides, the model results show a evolution

of the M_s/M_h ratio that seems to be opposite to what is inferred from semi-empirical inferences for low-mass galaxies (see Chapter 5). Therefore, we conclude that the strong feedback-driven outflows usually invoked in the literature, are not able to explain crucial semi-empirical inferences for late-type galaxies as is the M_s–M_h relation and its evolution as well as the observed SSFR-M_s relation (downsizing). Our semi-numerical evolutionary model results were lately shown to be similar to the more sophisticated N-body/hydrodynamics cosmological simulations of low-mass galaxies (Colín et al. [53], Avila-Reese et al. [6]). Therefore, we conclude that some astrophysical ingredients are missing in our current understanding of low-mass galaxy evolution. These ingredients can be related to the process of conversion of neutral gas into molecular one or to the process of gas conversion into stars, and/or to feedback effects that rather than ejecting the gas, keep it no forming stars.

Part VII

CONCLUSIONS

11

CONCLUSIONS

En general, en la presente Tesis se ha tratado la conexión entre las galaxias y los (sub)halos de materia obscura mediante enfoques estadísiticos. En la Parte II, la cual engloba los Capítulos 2 y 3, se han estudiado las relaciones de masa estelar galaxia central-masa halo, M_*-$M_{\rm h}$, y masa estelar galaxia satélite-masa subhalo, m_*-$m_{\rm sub}$, por separado. Como resultado, se desarrollo un modelo estadístico el cual es capaz de constreñir por separado las relaciones locales M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}$, donde $m_{\rm sub}$ puede referirse a la masa de los subhalos al tiempo de observación, $m_{\rm sub}^{\rm obs}$, y al momento de su acreción, $m_{\rm sub}^{\rm acc}$. Este modelo representa una extensión a la técnica de la correlación de abundancias. La virtud de este modelo radica en relacionar de manera auto-consistente la función de masa estelar de galaxias (FMEG) centrales y satétiles, la función de masa de halos/subhalos de materia oscura fría, las funciones condicionales de masa estelar (FCME) de satélites y la función de correlación de dos puntos (FC2P). Esto nos ha permitido no solo constreñir las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}$ por separado si no sondear su robustez, y ultimadamente discernir que tan viable es la suposición, comúnmente encontrada en la literatura, de que las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}$ son idénticas.

En la Parte III, la cual es conformada por los Capítulos 4 y 5, se exploró las implicaciones de las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}$ en la ocupación de galaxias en halos de materia oscura. Esto se ha hecho utilizando el modelo estadístico desarrollado en la Parte II de esta Tesis. En particular, nos hemos enfocado en estudiar la brecha en masa estelar existente entre el satélite más masivo y la galaxia central de un grupo/cumulo en función de $M_{\rm h}$, así como también diferentes distribuciones de probabilidades para la población de satélites tan masivos como Fornax en galaxias tipo Vía Lćatea (VL).

En el Capítulo 6 correspondiente Parte IV, hemos extendido nuestro modelo estadístico, para estudiar por separado la evolución con el corrimiento al rojo de las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}$, así como también la evolución de la fracción de galaxias satélites utilizando los datos del catastro COSMOS.

En la Parte IV, la cual engloba los Capítulos 7 y 8, hemos estudiado la relación local M_*-$M_{\rm h}$ de galaxias centrales no solo por su masa sino por colores. Como resultado, en el Capítulo 7 se ha extendido la técnica de la correspondencia de las abundancias. Para esto, fue necesario suponer la función de masa de halos que albergan galaxias azules centrales a partir de motivaciones teorico-observacionales. En ese Capítulo, también hemos estudiado la relación masa bariónica galaxia central-masa halo, M_b-$M_{\rm h}$. En el Capítulo 8, presentamos un enfoque más general, el cual esta basado en el modelo semi-empirico desarrollado en la Parte II. Este nuevo enfoque nos permite incluir información sobre la función de correlación de galaxias rojas y azules. La virtud de este modelo, es que no solo permite constreñir de manera robusta las relaciónes M_*-$M_{\rm h}$ de galaxias rojas y azules sino también constreñir de manera empirica la función de masa de halos que albergan galaxias azules centrales.

Por último, en la Parte V, que engloba los Capítulos 9 y 10, hemos desarrollado un modelo estático de galaxia de disco sumergido dentro de un halo de materia oscura fría. En este modelo las relaciones estructurales y dinámicas de las galaxias de disco quedan bien caracterizadas en términos de los parámetros cosmológicos de entrada; la masa del halo $M_{\rm h}$, el parametro de concentración c y el parametro de giro λ_h, y del parametro astrofísico $f_{\rm b} = M_b/M_{\rm h}$, la fracción bariónica. Esto nos ha permitido sondear la $f_{\rm b}$ de galaxias azules obtenida previamente en la Parte IV a la luz de las de las relaciones estructurares y dinámicas de las galaxias de disco.

A continuación presentamos las conclusiones más generales obtenidas en esta Tesis. Las conclusiones más especificas son presentadas al final de cada Capítulo.

- Bajo la suposición de que la relación de masa entre galaxias satélites y subhalos es idéntica a la relación entre galaxias centrales y halos, suposición comúnmente empleada en la técnica de la correspondencia de las abundancias, lleva a predecir una función de masa de galaxias satélites, funciónes de correlación de dos puntos y funciones condicionales de masa estelar de satélites en desacuerdo con lo que se infiere mediante los grandes catastros de galaxias. Esta conclusión fue obtenida definiendo la masa de los subhalos tanto al tiempo de observación, $m_{\rm sub}^{\rm obs}$, como al momento de su acreción, $m_{\rm sub}^{\rm acc}$.

- Mediante el modelo desarrollado en la Parte II, que es una extensión de la correspondencia de las abundancias, el modelo ocupaciónal de halos y la función condicional de masa de galaxias satélites, hemos probado la robustez de las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}$, al utilizar diferentes combinaciones de datos observacionales, tales como las funciones de masa estelar total, de galaxias centrales y satélites, y la función de correlación de dos puntos. En todos los casos estudiados, las incertidumbres asociadas a las relaciones de masa resultaron mucho más pequeñas que las incertidumbres intrínsecas de las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}$(~ 0.17 dex). La robustez de estas relaciones es valida definiendo la masa de los subhalos tanto al tiempo de observación, $m_{\rm sub}^{\rm obs}$, como al momento de su acreción, $m_{\rm sub}^{\rm acc}$.

- En general, la relación m_*-$m_{\rm sub}^{\rm obs}$ es muy diferente a la relación m_*-$m_{\rm sub}^{\rm acc}$, y a su vez ambas son diferentes a la relación M_*-$M_{\rm h}$ de galaxias centrales. A una masa estelar fija, la masa de los subhalos, tanto para $m_{\rm sub}^{\rm obs}$ como para $m_{\rm sub}^{\rm acc}$, es más pequeña que la de los halos de galaxias centrales $M_{\rm h}$. Esta diferencia incrementa a bajas masas y son más fuertes cuando se utiliza $m_{\rm sub}^{\rm obs}$.

- Encontramos que el ancho de la disperción intrínseca asociada a las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}^{\rm obs}$ son estrechas; $\sigma_c = 0.168 \pm 0.051$ dex y $\sigma_s = 0.172 \pm 0.057$ dex, respectivamente

- A una masa de halo fija, la distribución de masa del satélite más masivo se encuentra desplazada a bajas masas con respecto a la distribución de masa de las galaxias centrales, esto es, las galaxias centrales son galaxias estadísticamente especiales en el halo (grupo). Sin embargo, esta diferencia se hace menor con la masa del halo, al punto que en halos suficientemente masivos ambas distribuciones se traslapan, aproximadamente $\sim$ 15%. En este caso, las galaxias centrales podrían

ser una realización del la cola masiva de la función condicional de masa estelar de satélites.

- Con respecto a la población de satélites masivos en galaxias tipo VL, obtuvimos que la probabilidad de encontrar 2 satélites en el rango de masa de las Nube de Magallanes (NM) es de $\sim$ 6.6% (o $\sim$ 0.08% si ademas se pide la condición que no haya satétlites más masivos que la Nube Mayor de Magallanes, NMaM). La probabilidad de tener un satétlite con $m_* \geq m_{\rm LMC}$, dos satétlites con masa mayores a la Nube Menor de Magallanes, NMeM, $m_* \geq m_{\rm SMC}$, o tres con masas mayores o iguales a la de Sagitario, Sgr, $m_* \geq m_{\rm Sgr}$, son respectivamente el 26.1%, 14.5% y el 14.3%. La distribución del numero ocupacional de galaxias satélites en galaxias tipo VL es bastante amplia. Sin embargo, el hecho que la VL tenga dos satélites tan masivos como las NMs la hace menos común, pero no un caso atípico.

- No es posible constreñir la masa del halo de la VL mediante la abundancia de sus satélites masivos. No obstante, nuestros resultados muestran que a un nivel de 1σ, la masa de galaxias tipo VL no es menor a $1.38 \times 10^{12} M_\odot$.

- Nuestros resultados muestran que en la cosmología ΛCDM no existe un problema de satélites perdidos, al menos en el caso de los más masivos. Sin embargo, confirmamos un potencial problema de dinámica interna: la velocidad circular máxima de los satélites enanos, por abajo de $\sim 10^8 M_\odot$, parece ser sobreestimada por un factor $\sim 1.3 - 2$ de lo que típicamente es inferido en base a observaciones de los satélites enanos de la VL. Es plausible que este problema se refiera solo a galaxias satélites y no a galaxias centrales.

- Hemos calculado la función de correlación angular de dos puntos, $\omega(\theta)$, utilizando el catastro COSMOS en cuatro intervalos de corrimiento al rojo, $z = [0.2, 0.4]$, $[0.4, 0.6]$, $[0.6, 0.8]$ y $z = [0.8, 1]$ y para diferentes umbrales de masa estelar. Fijando un intervalo de corrimiento al rojo, la amplitud de la función de correlación se incrementa con la masa estelar. Fijando un umbral de masa estelar, galaxias a bajos corrimientos al rojo se encuentran más aglomeradas.

- Utilizando $\omega(\theta)$ y la FMEG total del COSMOS, hemos constreñido la evolución con el corrimiento al rojo de las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}^{\rm acc}$, por separado. Nuestros resultados muestran que la evolución con el corrimiento al rojo de las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}^{\rm acc}$ es débil . No obstante, la masa de halo donde se da el máximo del cociente $f_{*,\rm peak} = M_*/M_{\rm h}^{\rm peak}$ ha aumentado en casi un factor 2 desde $z \sim 1$, mientras que su amplitud se ha mantenido relativamente constante desde $z \sim 1$, con un valor de $f_{*,\rm peak} \sim 0.03$. En el caso de las galaxias satélites, la masa $m_{\rm acc}^{\rm peak}$, la masa del subhalo donde se da el máximo del cociente $f_{*,\rm peak} = m_*/m_{\rm acc}^{\rm peak}$, es constante para todos los corrimientos al rojo. Su amplitud decrece conforme aumenta el corrimientos al rojo, hasta parecerse a la de galaxias centrales.

- El hecho que las relaciones M_*-$M_{\rm h}$ y m_*-$m_{\rm sub}^{\rm acc}$ no difieren significativamente entre sí y que tampoco evolucionen significativamente con el corrimiento al rojo, cumple

con los requerimientos que supone la técnica de la correspondencia de las abundancias. Es plausible que este hecho brinde más pistas sobre porque es tan exitosa esta técnica.

- Las relaciones m_*-M_h y M_b-M_h (donde M_b es la masa en estrellas más la masa en gas, i.e., la masa bariónica) de galaxias centrales azules y rojas no difieren significativamente de la relación de galaxias centrales, y por lo tanto tampoco entre sí. Estas diferencias se encuentran dentro del 1σ del la incertidumbre asociada a las relaciones de masa. Tampoco hemos encontrado una clara indicación de segregación de las relaciones M_*-M_h por color, no obstante, la M_*-M_h de galaxias centrales rojas parece encontrarse por arriba de la M_*-M_h de galaxias azules centrales.

- Las suposiciones hechas en el Capítulo 7 sobre los halos que albergan galaxias azules centrales, aquellos que no han sufrido una fusión mayor desde $z \sim 0.8$ y que no viven en gurpos con más de tres miembros, son consistentes con la función de correlación de dos puntos de rojas y azules observada.

- La fracción de masa de halos que albergan galaxias azules centrales es una función que depende fuertemente de la masa del halo. Esta muestra que la fracción máxima de halos azules es del $\sim$ 80% y se da en masas de $M_\mathrm{h} \sim 10^{11.4} M_\odot$. Después de estas masas la fracción cae, alcanzando un $\sim$ 50% en halos de masa $M_\mathrm{h} \sim 10^{12} M_\odot$ y un $\sim$ 4% en halos tan masivos como $M_\mathrm{h} \sim 10^{15} M_\odot$.

- Utilizando un modelo estático de galaxia de disco, hemos sondeado como afecta el parametro de entrada la f_b a las relaciones estructurales y dinámicas. Nuestros resultados muestran que mientras las relaciones Tully-Fisher estelar (TFE) y bariónica (TFB) son bastantes robustas, las relaciones radio efectivo-masa bariónica y radio efectivo-masa estelar son afectadas significativamente por la forma de f_b. Similarmente, el cociente velocidad circular del disco a velocidad circular total al máximo también se ve afectada significativamente por la forma de f_b. No obstante, cuando utilizamos la f_b constreñida para las galaxias azules en el Capitulo 9, todas las relaciones estructurales y dinámicas presentan un ligero doblez, esto es más evidente en las relaciones radio efectivo-masa bariónica y radio efectivo-masa estelar, así como también en el cociente velocidad circular del disco a velocidad circular total al máximo.

- En general, utilizando como dato de entrada en nuestro modelo de galaxias de disco la f_b constreñida para las galaxias azules, encontramos que la normalización así como las dispersiones intrínsecas en relaciones estructurales y dinámicas son consistentes con las observaciones. Esto implica que la relación TFE y TFB no presentan el problema del punto cero reportado en trabajos previos.

- Suponer conservación detallada del momento angular, $\lambda_d = \lambda_h$, lleva a encontrar galaxias masivas con radios muy extendidos, es decir, galaxias de muy bajo brillo superficial. Con el fin de reproducir las observaciones, encontramos que λ_d debe adquirir sistemáticamente valores menores a λ_h con la masa M_h, i.e., la suposición de conservación de momento angular detallado debe ser relajada.

- Los valores tan bajos encontrado de galaxias azules para f_b, inevitablemente lleva a encontrar que al radio donde la curva de rotación total del sistema (halo+disco) alcanza su máximo las galaxias son dominadas por materia oscura. En las observaciones, ligeras indicaciones existen de que este puede ser el caso. Sin embargo, más inferencias observacionales con datos de mucha más alta resolución son necesarios para dilucidar cuando esto es cierto o no.

Part VIII

APPENDIX

A

THE SPATIAL CLUSTERING OF GALAXIES IN THE HOD MODEL

Here we review the main ideas used to infer the spatial clustering of galaxies based on a HOD model. We assume that the most massive galaxy in terms of stellar mass within a halo of mass $M_{\rm h}$ is its central galaxy. Consequently the remaining galaxies are all satellites. We let them follow the mass density profile of the host halo. We denote the cumulative number of central and satellite galaxies with stellar masses greater than m_* as N_c and N_s, respectively.

The two point correlation function is decomposed into two terms,

$$1+\xi_{\rm gg}(r) = [1+\xi^{\rm 1h}_{\rm gg}(r)] + [1+\xi^{\rm 2h}_{\rm gg}(r)], \tag{134}$$

with $1+\xi^{\rm 1h}_{\rm gg}(r)$ describing galaxy pairs within the same halo (the one-halo term), and $1+\xi^{\rm 2h}_{\rm gg}(r)$ describing the correlation between galaxies occupying different halos (the two-halo term).

We compute the one-halo term as

$$1+\xi^{\rm 1h}_{\rm gg}(r) = \frac{1}{2\pi r^2 n_g^2}\int_0^\infty \frac{\langle N(N-1)\rangle}{2}\lambda(r)\phi_{\rm h}(M_{\rm h})dM_{\rm h}, \tag{135}$$

for pairs separated by a distance $r \pm dr/2$. Here $\langle N(N-1)\rangle/2$ is the total mean number of galaxy pairs within halos $M_{\rm h}$ following a pair distribution $\lambda(r)dr$, and a mean number density $n_g(m_*)$. The contribution to the total mean number of galaxy pairs from central-satellite pairs and satellite-satellite pairs is

$$\begin{aligned}\frac{\langle N(N-1)\rangle}{2}\lambda(r)dr = \langle N_c\rangle\langle N_s\rangle\lambda_{c,s}(r)dr \\ + \frac{\langle N_s(N_s-1)\rangle}{2}\lambda_{s,s}(r)dr.\end{aligned} \tag{136}$$

As commonly assumed in HOD models, the number of central-satellite pairs follow the normalized mass density halo profile, taken to be of Navarro, Frenk & White [151] shape. The number of satellite-satellite pairs is then related to the normalized density profile convolved with itself.

Halo profiles are defined in terms of the total halo mass and the concentration parameter. We use the relation between concentration parameter $c_{\rm NFW}$ and halo mass obtained by Muñoz-Cuartas et al. [150] from fits to N-body simulations.

Based on results of high-resolution N-body [115] and hydrodynamic simulations of galaxy formation [262], we model the second moment of satellite galaxies, $\langle N_s(N_s-1)\rangle$, as a Poisson distribution with mean $\langle N_s\rangle^2 = \langle N_s(N_s-1)\rangle$.

We compute the two-halo term as

$$\xi^{\rm 2h}_{\rm gg}(r) = b_g^2\zeta^2(r)\xi_m(r), \tag{137}$$

where $\xi_m(r)$ is the non-linear matter correlation function [196], $\zeta(r)$ is the scale dependence of dark matter halo bias [209], and

$$b_g = \frac{1}{n_g}\int_0^\infty b(M_{\rm h})\langle N(>m_*|M_{\rm h})\rangle\phi_{\rm h}(M_{\rm h})dM_{\rm h}, \tag{138}$$

is the galaxy bias with $b(M_{\rm h})$ the halo bias function [191].

Once we have calculated $\xi_{\rm gg}(r)$, we relate it to the projected correlation function, $w_{\rm p}(r_{\rm p})$, integration over the line of sight,

$$w_{\rm p}(r_{\rm p}) = 2\int_0^\infty \xi_{\rm gg}(\sqrt{r_{\rm p}^2+x^2})dx. \tag{139}$$

B

THE MARKOV CHAIN MONTE CARLO METHOD

The basic idea behind the Markov Chain Monte Carlo Method (MCMC) is to provide a simple method for sampling random numbers from some generic probability distribution $P(x)$. This is motived by the fact that in many cases one knows how to write the equation for the probability distribution $P(x)$, but is not trivial to generate random numbers from $P(x)$. The benignity of the MCMC is that you do not even need to know how to write the probability distribution $P(x)$. This is the situation which one faces when try to sample the a set of parameters, $\mathbf{a} = (a_1, ..., a_n)$ that maximize the likelihood function $\mathcal{L}(\mathbf{a}) \propto \exp(-\chi^2(\mathbf{a})/2)$.

One of the most popular examples of a MCMC is the Metropolis-Hasting algorithm (MHa, see e.g., Metropolis [141]). The idea in the MHa is to generate a sequence of states $(x_0, x_1, ..., x_n, ...)$ such a that for a sufficient large n (ideally, $n \to \infty$) we can guarantee that $x_n \sim P(x)$, i.e., we have sampled completely the distribution function $P(x)$. In fact, the above idea is known as a Markov chain, (Mc). Below, we describe how can one generate samples from the distribution function $P(x)$.

B.1 THE METROPOLIS-HASTING ALGORITHM

Without loss of generality, suppose that the current state of the Mc is x_n and we want to generate the next state x_{n+1}. The generation of x_{n+1} in the Mc requires of two steps:

1. On the first step, one generates a candidate, denoted by x_c, from a proposal distribution function, denoted by $Q(x_c|x_n)$. Here the proposal distribution depends on the current state of the Mc, x_n. The importance of this step is to use a distribution function $Q(x_c|x_n)$ easy to handy, i.e., that we know how to generate random numbers from $Q(x_c|x_n)$. Typically in the literature a normal distribution function centered on the current state of x_n and standard deviation of σ is what is used, $Q(x_c|x_n) \sim \text{Normal}(x_n, \sigma)$. Therefore, our candidate x_c is a random realization from $\text{Normal}(x_n, \sigma)$. In the following, we will denote the normal distribution function by $Q(x_c|x_n, \sigma)$. Observe that σ is a free parameter that should be adjusted.

2. On the second step, we will ask whether x_c is a good candidate for x_{n+1}. To than end, we need to compute the acceptance probability of x_c given by:

$$A(x_n \to x_c) = \min\left(1, \frac{P(x_c)}{P(x_n)} \times \frac{Q(x_n|x_c, \sigma)}{Q(x_c|x_n, \sigma)}\right). \tag{140}$$

   To decide whether the candidate x_c is either accepted or rejected, a random number between 0 and 1 from a uniform distribution should be generated, then;

$$x_{n+1} = \begin{cases} x_c & \text{if} \quad u \leq A(x_n \to x_c) \\ x_n & \text{if} \quad u > A(x_n \to x_c) \end{cases} \tag{141}$$

we decide to accept the candidate x_c when $x_{n+1} = x_c$, otherwise we decide to reject the candidate x_c and $x_{n+1} = x_n$.

Essentially, the above two-stage process is the entirely MHa. Nevertheless, few comments are needed at this point. Firstly, observe that the ratio $P(x_c)/P(x_n)$ is independent on the normalizing constant of the distribution $P(x)$. This will be very convenient for our porpuse because obtaining the normalizing constant from the likelihood function is not trivial, hence, we are allowed to write: $\mathcal{L}(\mathbf{a}) \propto \exp(-\chi^2(\mathbf{a})/2)$. Second, because our proposal distribution is symmetric, i.e., $Q(x_n|x_c, \sigma) = Q(x_c|x_n, \sigma)$ for all x_c and x_n, then the ratio in Eq. 140 is simply $Q(x_n|x_c, \sigma)/Q(x_c|x_n, \sigma) = 1$.

B.2 THE MCMC METHOD IN MULTIDIMENSIONS

So far the above explanation focused in the one dimensional case. The generalization for sampling random vectors from some distribution $P(\mathbf{x})$[1] is immediate. In essence, the idea is the same since one should define a Mc to generate a very large sequence of states $(\mathbf{x_0}, \mathbf{x_1}, ..., x_n, ...)$ to sample the distribution function $P(\mathbf{x})$. Here each $\mathbf{x_i}$ is a N-dimensional vector. For example N can be the number of model parameter for fitting some data and the vector $\mathbf{x_i}$ represents the ith-array of the best fit model parameters.

To extent the above two-stage process there are few technical modifications that one needs to include in step one. The value of each component, $x_{c,i}$, of the vector x_c is generated from a Normal distribution function centered on the current state of the ith-component and standard deviation of σ_i, $Q_i(x_{c,i}|x_n, \sigma_i)$. In general, for a candidate x_c there is no need to assume that all its components have the same parameters x_n and σ_i, rather they will depend on the problem to be solve. Finally, note that in the most basic conception of MCMC method in multidimensions the normal distributions Q_i are independent between each other, but in general they can be generated, see Chapter 6.

1 Note that $P(\mathbf{x})$ is a real number

C

THE FITTING PROCEDURE

From the fit to the data, we constrain the ten free parameters of model by maximizing the likelihood function $\mathcal{L} \propto \exp(-\chi^2/2)$. Table 1 lists the different combination of observational constrains used in this paper.

For each *GSMF*, the χ^2's are defined as:

$$\chi^2(\phi^{\text{author}}_{\text{tot,cen,sat}}) = \frac{1}{N_{\text{bin}}} \sum_{i=1}^{N_{\text{bin}}} \left(\frac{\phi_{g^i,\text{model}} - \phi_{g^i,\text{obs}}}{\sigma^i_{\text{obs}}} \right)^2, \tag{142}$$

where N_{bin} is the number of bins in the total/central/satellite *GSMF* reported for each author with an ith value of $\phi_{g^i,\text{obs}}$ and an error of σ^i_{obs}. The ith value of the total/central/satellite *GSMF* computed in the model is denoted as $\phi_{g^i,\text{model}}$.

For the Yang et al. [250] projected 2PCFs, the χ^2 is defined as:

$$\chi^2(w^{\text{Y11}}_{\text{p,bin}}) = \frac{1}{N_{\text{s,bin}} N_{\text{r,bin}}} \sum_{i=1}^{N_{\text{s,bin}}} \sum_{j=1}^{N_{\text{r,bin}}} \left(\frac{w_{\text{p},\text{model}}^{i,j} - w_{\text{p},\text{obs}}^{i,j}}{\sigma^{i,j}_{\text{obs}}} \right)^2, \tag{143}$$

where $N_{\text{s,bin}}$ is the number of stellar mass bins, $N_{\text{r,bin}}$ denotes the number of bins in the 2PCF, $w_{\text{p},\text{obs}}^{i,j}(w_{\text{p},\text{model}}^{i,j})$ is the amplitude of the observed (modeled) 2PCF in the jth projected distance bin of the ith stellar mass bin.

First, we find the set of parameters, $\mathbf{a} = (a_1, ..., a_n)$, that minimizes χ^2 using the Powell's directions set method in multi-dimensions, Press et al. [167]. Then, the resulting set of parameters is used as the starting point to sample the parameter space with the MCMC method. In most of our cases $n = 10$. We also need to specify for each parameter a proposed distribution, which generates the candidate for sampling the parameter space. We assume that each proposed distribution is Gaussian distributed. The standard deviation for each parameter, $\sigma(a_i)$, is calculated from the covariance matrix. The covariance matrix or error matrix of $\mathbf{a}$ is defined as the inverse of the $n \times n$ matrix $\alpha = \epsilon^{-1}$, computed according to

$$\alpha_{kl} = \frac{1}{2} \frac{\partial^2 \chi^2(\mathbf{a})}{\partial a_k \partial a_l}. \tag{144}$$

Therefore, the standard deviations in the parameters correspond to the square roots of the terms in its diagonal, i.e., $\sigma(a_i) = \sqrt{\epsilon_{ii}}$. We consider these numbers as our best initial guess for the diagonal covariance matrix of the model parameters. Then the covariance matrix for the proposed matrix is computed according to the formula given in [77]; $\epsilon^{\text{p}}_{ii} = \frac{2.4}{n}^2 \epsilon_{ii}$, with n the number of parameters to be fitted.

Using these results, we sample a first chain with 100,000 models, from which we compute the diagonal of the covariance matrix, ϵ^{c}_{ii}. If the ratio of each prior, $\sqrt{\epsilon_{ii}}$, to each element of the resulting diagonal covariance matrix, $\sqrt{\epsilon^{\text{c}}_{ii}}$, lies in the range

$0.8 \leq \sqrt{\epsilon_{ii}/\epsilon_{ii}^{c}} \leq 1.2$, then we initialize a second chain with 1.5×10^6 elements for the model analysis; else, we repeat the procedure $j-$times until the ratio of the covariances of the previous chain with the last one reachs the condition $0.8 \leq \sqrt{\epsilon_{ii}^{j-1}/\epsilon_{ii}^{j}} \leq 1.2$, that is to say, until there is not a sufficiently significant improvement in the standard deviations of the model parameters. The $j-$covariance matrix for the proposed distribution is given by $\epsilon_{ii}^{p,j} = \frac{2.4}{n}^2 \epsilon_{ii}^{j}$. Then, we run a last chain with 1.5×10^6 elements for the model analysis. This procedure usually takes one or two iterations. For all the chains, we find a convergence ratio in each parameter lower than 0.01 [77].

D

JUSTIFICATION TO EQUATION 76

Suppose that there exist a probability distribution function $f(x)$ which is negligible outside the bin $(\eta - \epsilon, \eta + \epsilon)$ and we would like to estimate the following integral:

$$\int_{\infty}^{\infty} g(x)f(x)dx, \tag{145}$$

where $g(x)$ is a continuos function. Here, we will assume that η is the mean of the distribution $f(x)$. Expanding $g(x)$ in Taylor series around η

$$g(x) = g(\eta) + g'(\eta)(x-\eta) + ... + g^{(n)}(\eta)\frac{(x-\eta)^n}{n!}, \tag{146}$$

where $g^{(n)}$ denotes the n-derivate with respect to x. Inserting this into Eq. 145, we obtain,

$$\int_{\infty}^{\infty} g(x)f(x)dx = g(\eta) + g''(\eta)\frac{\sigma^2}{2} + ... + g^{(n)}(\eta)\frac{\mu_n}{n!}. \tag{147}$$

Here, σ is the standard deviation of $f(x)$. More generally μ_n is the nth moment of the distribution function $f(x)$, defined as:

$$\mu_n = \int_{\infty}^{\infty} (x-\eta)^n f(x)dx. \tag{148}$$

If $g(x)$ is approximated by a parabola, then;

$$\int_{\infty}^{\infty} g(x)f(x)dx = g(\eta) + g''(\eta)\frac{\sigma^2}{2}, \tag{149}$$

which is the solution to Eq. 145 at a second order. In Eq. 73 we have that $f(x) \rightarrow P_{\rm B,c}(M_*|M_{\rm h})$ and $g(x) \rightarrow \mathcal{C}_{gr,\rm B}(M_*)$, in principle what we only need to compute is the following: $d^2\mathcal{C}_{gr,\rm B}(M_*)/d\log {M_*}^2$. Because observationally $\mathcal{C}_{gr,\rm B}(M_*)$ is well characterize by a straight line then $d^2\mathcal{C}_{gr,\rm B}(M_*)/d\log {M_*}^2 \sim 0$. Therefore $\mathcal{C}_{gr,\rm B}(M_{\rm h}) \approx \mathcal{C}_{gr,\rm B}(M_{*,\rm B}(M_{\rm h}))$ provides an excellent approximation. Similarly, for the total and the red mean relations. This implies that the fraction of blue central galaxies as a function of halo mass is well approximated by,

$$f_{\rm B}(M_{\rm h}) \approx \frac{\mathcal{C}_{gr}(M_*(M_{\rm h})) - \mathcal{C}_{gr,\rm R}(M_{*,\rm R}(M_{\rm h}))}{\mathcal{C}_{gr,\rm B}(M_{*,\rm B}(M_{\rm h})) - \mathcal{C}_{gr,\rm R}(M_{*,\rm R}(M_{\rm h}))}. \tag{150}$$

E

THE MEAN COLOR-TO-STELLAR MASS RELATION OF BLUE AND RED CENTRAL GALAXIES

One important input of the model described in Chapter 8 is the mean color-to-stellar mass relation of blue and red galaxies. In this Appendix, we obtain these relations from the galaxy color distribution of central galaxies by using the Yang et al. [249] group catalog.

First at all, we divide the catalog for centrals into two large subsamples namely that of blue and that of red centrals. Here, we define blue and red galaxies by using the same color cut in the color-magnitude diagram as in YMB09. Once we have obtained these blue and red subsamples, we divide each of them into 32 stellar masses bins of width 0.1 dex over the stellar mass range $[10^{8.5}, 10^{11.7}]$ $M_{\odot}$. Then, each stellar mass bin is divided into 22 color bins of width 0.05 each. Figure 48 shows these color distributions at nine different stellar mass bins. Observe that each color distribution in the figure has been renormalized by its maximum value. The blue circles in the figure indicates the observed color distributions for the blue sample of central galaxies, while the red circles presents the same but for red centrals. The solid lines represent the fit to each blue and red color distribution by using Gaussian distribution models. In what follows, we will denote these color distributions by $\mathcal{B}(C_{\rm gr}|M_*)$ and $\mathcal{R}(C_{\rm gr}|M_*)$ for the blue and the red subsamples, respectively.

The relation between the total color distribution, $P(C_{\rm gr}|M_*)$, and the distributions $\mathcal{B}(C_{\rm gr}|M_*)$ and $\mathcal{R}(C_{\rm gr}|M_*)$ is

$$P(C_{\rm gr}|M_*) = \mathcal{B}(C_{\rm gr}|M_*) f_{*,B}(M_*) + \mathcal{R}(C_{\rm gr}|M_*) f_{*,R}(M_*). \tag{151}$$

Here $f_{*,B}(M_*)$ and $f_{*,R}(M_*)$ denote the fraction of blue and red central galaxies with stellar masses M_*. Notice that by definition, the observed mean color-to-stellar mass relation is

$$C_{\rm gr}(M_*) = C^{\rm B}_{\rm gr}(M_*) f_{*,B}(M_*) + C^{\rm R}_{\rm gr}(M_*) f_{*,R}(M_*) \tag{152}$$

where $C^{\rm B}_{\rm gr}(M_*)$ and $C^{\rm R}_{\rm gr}(M_*)$ denote the observed mean color-to-stellar mass relation for blue and red central galaxies. The quantities, $C^{\rm B}_{\rm gr}(M_*)$ and $C^{\rm R}_{\rm gr}(M_*)$ are the average relations obtained after fitting the observed distributions $\mathcal{B}(C_{\rm gr}|M_*)$ and $\mathcal{R}(C_{\rm gr}|M_*)$.

Because of galaxies at a fixed luminosity, and at a fixed stellar mass, are observed only within a certain redshift range within the sample, it is then expected that the color distribution suffers of incompleteness effects. This issue can be accounted by assigning to each galaxy a weight $1/V_{\rm max}$ where $V_{\rm max}$ is the maximum volume over which the galaxy will be observed within the sample of volume V_S. However, this volume corrections depend mainly on the luminosity (stellar mass) and in less degree on the color such a that correction is not important for the color distributions $\mathcal{B}(C_{\rm gr}|M_*)$ and $\mathcal{R}(C_{\rm gr}|M_*)$. Nevertheless, the total color distribution $P(C_{\rm gr}|M_*)$ depends on the fraction of blue

and red centrals at each stellar mass bin, see Eq. 151, for which volume corrections are important. We account for this correction by using the fact that the fraction $f_{\rm s}(M_*)$ is the ratio $\phi_{\rm B,c}(M_*)/\phi_{\rm g,c}(M_*)$, where $\phi_{\rm B,c}(M_*)$ and $\phi_{\rm g,c}(M_*)$ are the blue and the total observed central GSMFs, respectively. In addition that $f_{\rm R}(M_*) = 1 - f_{\rm B}(M_*)$.

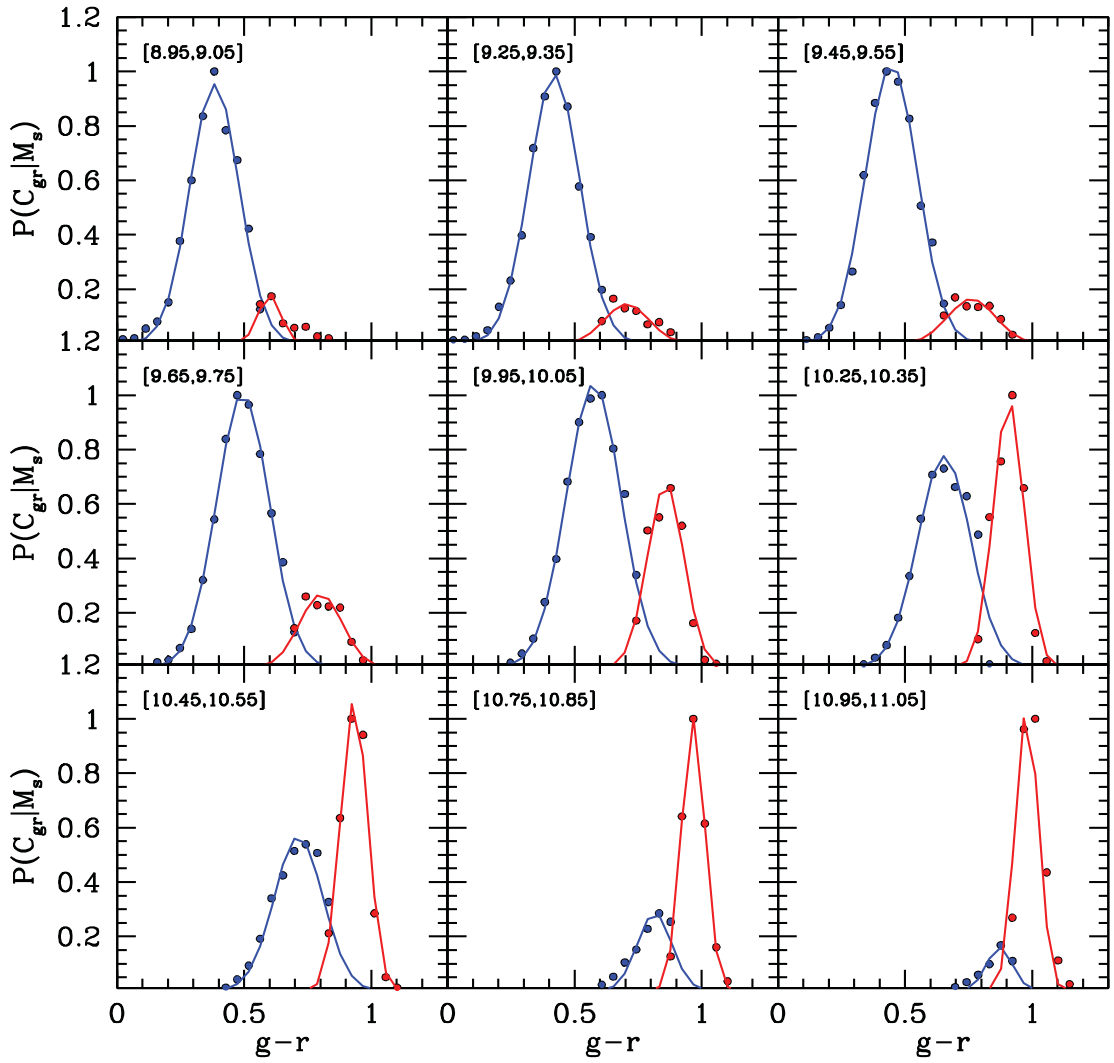

Figure 48: Color distribution for blue and red central galaxies in various stellar mass bins. The blue circles indicate the observed color distributions for the blue sample of central galaxies, while the red circles presents the same information but for red. The solid blue and red lines show the results of the Gaussian fitting models to the data for the blue and red distributions. Each color distribution in the figure has been renormalized by its maximum value.

BIBLIOGRAPHY

[1] Abbas U., Sheth R. K., 2006, MNRAS, 372, 1749

[2] Amorín R., Alfonso J., Aguerri J. A. L., Muñoz-Tuñón C., Cairós L. M., 2009, A&A, 501, 75

[3] Angulo R. E., Lacey C. G., Baugh C. M., Frenk C. S., 2009, MNRAS, 399, 983

[4] Arraki K. S., Klypin A., More S., Trujillo-Gomez S., 2012, ArXiv e-prints

[5] Avila-Reese V., 2006, ArXiv Astrophysics e-prints

[6] Avila-Reese V., Colín P., González-Samaniego A., Valenzuela O., Firmani C., Velázquez H., Ceverino D., 2011, ApJ, 736, 134

[7] Avila-Reese V., Colín P., Gottlöber S., Firmani C., Maulbetsch C., 2005, ApJ, 634, 51

[8] Avila-Reese V., Firmani C., 2000, RMxAA, 36, 23

[9] Avila-Reese V., Firmani C., 2011, in Revista Mexicana de Astronomia y Astrofisica Conference Series, Vol. 40, Revista Mexicana de Astronomia y Astrofisica Conference Series, pp. 27–35

[10] Avila-Reese V., Firmani C., Hernández X., 1998, ApJ, 505, 37

[11] Avila-Reese V., Firmani C., Klypin A., Kravtsov A. V., 1999, MNRAS, 310, 527

[12] Avila-Reese V., Zavala J., Firmani C., Hernández-Toledo H. M., 2008, AJ, 136, 1340

[13] Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469

[14] Baldry I. K. et al., 2012, MNRAS, 421, 621

[15] Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay A. S., 2004, ApJ, 600, 681

[16] Baldry I. K., Glazebrook K., Driver S. P., 2008, MNRAS, 388, 945

[17] Bauer A. E. et al., 2013, MNRAS

[18] Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101

[19] Behroozi P. S., Conroy C., Wechsler R. H., 2010, ApJ, 717, 379

[20] Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57

[21] Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJS, 149, 289

[22] Bellazzini M., Correnti M., Ferraro F. R., Monaco L., Montegriffo P., 2006, A&A, 446, L1

[23] Benson A. J., 2010, PhR, 495, 33

[24] Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38

[25] Berlind A. A., Weinberg D. H., 2002, ApJ, 575, 587

[26] Bernardi M., Shankar F., Hyde J. B., Mei S., Marulli F., Sheth R. K., 2010, MNRAS, 404, 2087

[27] Bershady M. A., Martinsson T. P. K., Verheijen M. A. W., Westfall K. B., Andersen D. R., Swaters R. A., 2011, ApJ, 739, L47

[28] Berta Z. K., Jimenez R., Heavens A. F., Panter B., 2008, MNRAS, 391, 197

[29] Bett P., Eke V., Frenk C. S., Jenkins A., Helly J., Navarro J., 2007, MNRAS, 376, 215

[30] Blanton M. R., Berlind A. A., 2007, ApJ, 664, 791

[31] Blanton M. R., Geha M., West A. A., 2008, ApJ, 682, 861

[32] Blanton M. R., Moustakas J., 2009, ARA, 47, 159

[33] Blumenthal G. R., Faber S. M., Flores R., Primack J. R., 1986, ApJ, 301, 27

[34] Bolzonella M. et al., 2010, A&A, 524, A76

[35] Bottema R., 1993, A&A, 275, 16

[36] Boylan-Kolchin M., Besla G., Hernquist L., 2011, MNRAS, 414, 1560

[37] Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2011, MNRAS, 415, L40

[38] Boylan-Kolchin M., Bullock J. S., Kaplinghat M., 2012, MNRAS, 422, 1203

[39] Boylan-Kolchin M., Bullock J. S., Sohn S. T., Besla G., van der Marel R. P., 2013, ApJ, 768, 140

[40] Boylan-Kolchin M., Ma C.-P., Quataert E., 2008, MNRAS, 383, 93

[41] Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., 2010, MNRAS, 406, 896

[42] Brooks A. M., Zolotov A., 2012, ArXiv e-prints

[43] Bryan G. L., Norman M. L., 1998, ApJ, 495, 80

[44] Busha M. T., Wechsler R. H., Behroozi P. S., Gerke B. F., Klypin A. A., Primack J. R., 2011, ApJ, 743, 117

[45] Cacciato M., van den Bosch F. C., More S., Li R., Mo H. J., Yang X., 2009, MNRAS, 394, 929

[46] Capak P. et al., 2007, ApJS, 172, 99

[47] Cervantes-Sodi B., Hernandez X., Park C., Kim J., 2008, MNRAS, 388, 863

[48] Chabrier G., 2003, PASP, 115, 763

[49] Chae K.-H., 2010, MNRAS, 402, 2031

[50] Chou M.-Y. et al., 2007, ApJ, 670, 346

[51] Cole A. A., Tolstoy E., Gallagher, III J. S., Smecker-Hane T. A., 2005, AJ, 129, 1465

[52] Cole S. et al., 2001, MNRAS, 326, 255

[53] Colín P., Avila-Reese V., Vázquez-Semadeni E., Valenzuela O., Ceverino D., 2010, ApJ, 713, 535

[54] Collister A. A., Lahav O., 2005, MNRAS, 361, 415

[55] Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486

[56] Conroy C. et al., 2007, ApJ, 654, 153

[57] Conroy C., Wechsler R. H., 2009, ApJ, 696, 620

[58] Conroy C., Wechsler R. H., Kravtsov A. V., 2006, ApJ, 647, 201

[59] Cooray A., 2006, MNRAS, 365, 842

[60] Cooray A., Sheth R., 2002, PhR, 372, 1

[61] Coupon J. et al., 2012, A&A, 542, A5

[62] Courteau S., Dutton A. A., van den Bosch F. C., MacArthur L. A., Dekel A., McIntosh D. H., Dale D. A., 2007, ApJ, 671, 203

[63] Courteau S., Rix H.-W., 1999, ApJ, 513, 561

[64] Cuesta A. J., Prada F., Klypin A., Moles M., 2008, MNRAS, 389, 385

[65] Davis M. et al., 2007, ApJ, 660, L1

[66] de Boer T. J. L. et al., 2012, A&A, 544, A73

[67] de la Torre S. et al., 2010, MNRAS, 409, 867

[68] de Lapparent V., Slezak E., 2007, A&A, 472, 29

[69] De Lucia G., Kauffmann G., Springel V., White S. D. M., Lanzoni B., Stoehr F., Tormen G., Yoshida N., 2004, MNRAS, 348, 333

[70] De Rijcke S., Zeilinger W. W., Hau G. K. T., Prugniel P., Dejonghe H., 2007, ApJ, 659, 1172

[71] De Rossi M. E., Avila-Reese V., Tissera P., Gonzalez-Samaniego A., Pedroza S., 2013, submitted

[72] de Rossi M. E., Tissera P. B., Pedrosa S. E., 2010, A&A, 519, A89

[73] di Cintio A., Knebe A., Libeskind N. I., Yepes G., Gottlöber S., Hoffman Y., 2011, MNRAS, 417, L74

[74] Driver S. P. et al., 2011, MNRAS, 413, 971

[75] Drory N. et al., 2009, ApJ, 707, 1595

[76] Drory N., Salvato M., Gabasch A., Bender R., Hopp U., Feulner G., Pannella M., 2005, ApJ, 619, L131

[77] Dunkley J., Bucher M., Ferreira P. G., Moodley K., Skordis C., 2005, MNRAS, 356, 925

[78] Dutton A. A., Conroy C., van den Bosch F. C., Prada F., More S., 2010, MNRAS, 407, 2

[79] Dutton A. A., van den Bosch F. C., 2009, MNRAS, 396, 141

[80] Dutton A. A., van den Bosch F. C., Dekel A., Courteau S., 2007, ApJ, 654, 27

[81] Eisenstein D. J., Loeb A., 1996, ApJ, 459, 432

[82] Firmani C., Avila-Reese V., 2000, MNRAS, 315, 457

[83] Firmani C., Avila-Reese V., 2009, MNRAS, 396, 1675

[84] Firmani C., Avila-Reese V., 2010, ApJ, 723, 755

[85] Firmani C., Avila-Reese V., Rodríguez-Puebla A., 2010, MNRAS, 404, 1100

[86] Flynn C., Holmberg J., Portinari L., Fuchs B., Jahreiß H., 2006, MNRAS, 372, 1149

[87] Fontanot F., De Lucia G., Monaco P., Somerville R. S., Santini P., 2009, MNRAS, 397, 1776

[88] Freeman K. C., 1970, ApJ, 160, 811

[89] Gallazzi A., Bell E. F., 2009, ApJS, 185, 253

[90] Gao L., Frenk C. S., Boylan-Kolchin M., Jenkins A., Springel V., White S. D. M., 2011, MNRAS, 410, 2309

[91] Gao L., White S. D. M., Jenkins A., Stoehr F., Springel V., 2004, MNRAS, 355, 819

[92] Geha M., Blanton M. R., Masjedi M., West A. A., 2006, ApJ, 653, 240

[93] Giocoli C., Tormen G., Sheth R. K., van den Bosch F. C., 2010, MNRAS, 404, 502

[94] Giocoli C., Tormen G., van den Bosch F. C., 2008, MNRAS, 386, 2135

[95] Gnedin O. Y., Kravtsov A. V., Klypin A. A., Nagai D., 2004, ApJ, 616, 16

[96] Gnedin O. Y., Weinberg D. H., Pizagno J., Prada F., Rix H.-W., 2007, ApJ, 671, 1115

[97] Governato F., Mayer L., Brook C., 2008, in Astronomical Society of the Pacific Conference Series, Vol. 396, Formation and Evolution of Galaxy Disks, Funes J. G., Corsini E. M., eds., p. 453

[98] Governato F. et al., 2012, MNRAS, 422, 1231

[99] Gritschneder M., Lin D. N. C., 2013, ApJ, 765, 38

[100] Grogin N. A. et al., 2011, ApJS, 197, 35

[101] Guo Q., Cole S., Eke V., Frenk C., 2011, MNRAS, 417, 370

[102] Guo Q., White S., Li C., Boylan-Kolchin M., 2010, MNRAS, 404, 1111

[103] Hansen S. M., Sheldon E. S., Wechsler R. H., Koester B. P., 2009, ApJ, 699, 1333

[104] Hayashi K., Chiba M., 2012, ApJ, 755, 145

[105] Heinämäki P., Einasto J., Einasto M., Saar E., Tucker D. L., Müller V., 2003, A&A, 397, 63

[106] Herrmann K. A., Ciardullo R., 2009, ApJ, 705, 1686

[107] Hoeft M., Yepes G., Gottlöber S., Springel V., 2006, MNRAS, 371, 401

[108] Ilbert O. et al., 2009, ApJ, 690, 1236

[109] James P. A., Ivory C. F., 2011, MNRAS, 411, 495

[110] Jiang C. Y., Jing Y. P., Li C., 2012, ApJ, 760, 16

[111] Jimenez R., Verde L., Oh S. P., 2003, MNRAS, 339, 243

[112] Kajisawa M. et al., 2009, ApJ, 702, 1393

[113] Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102

[114] Kovač K. et al., 2010, ApJ, 708, 505

[115] Kravtsov A. V., Berlind A. A., Wechsler R. H., Klypin A. A., Gottlöber S., Allgood B., Primack J. R., 2004, ApJ, 609, 35

[116] Kregel M., van der Kruit P. C., de Grijs R., 2002, MNRAS, 334, 646

[117] Kregel M., van der Kruit P. C., Freeman K. C., 2005, MNRAS, 358, 503

[118] Lacerna I., Rodríguez-Pueba A., Avila-Reese A., Hernandez-Toledo H., 2013, submitted

[119] Landy S. D., Szalay A. S., 1993, ApJ, 412, 64

[120] Lares M., Lambas D. G., Domínguez M. J., 2011, AJ, 142, 13

[121] Leauthaud A., Tinker J., Behroozi P. S., Busha M. T., Wechsler R. H., 2011, ApJ, 738, 45

[122] Leauthaud A. et al., 2012, ApJ, 744, 159

[123] Li C., Kauffmann G., Jing Y. P., White S. D. M., Börner G., Cheng F. Z., 2006, MNRAS, 368, 21

[124] Li C., White S. D. M., 2009, MNRAS, 398, 2177

[125] Li Y.-S., White S. D. M., 2008, MNRAS, 384, 1459

[126] Lin Y.-T., Mohr J. J., 2004, ApJ, 617, 879

[127] Liu L., Gerke B. F., Wechsler R. H., Behroozi P. S., Busha M. T., 2011, ApJ, 733, 62

[128] Longair M. S., 2008, Galaxy Formation

[129] Lovell M. R. et al., 2012, MNRAS, 420, 2318

[130] Macciò A. V., Stinson G., Brook C. B., Wadsley J., Couchman H. M. P., Shen S., Gibson B. K., Quinn T., 2012, ApJ, 744, L9

[131] Magliocchetti M., Porciani C., 2003, MNRAS, 346, 186

[132] Mandelbaum R., Seljak U., Hirata C. M., 2008, JCAP, 8, 6

[133] Mandelbaum R., Seljak U., Kauffmann G., Hirata C. M., Brinkmann J., 2006, MNRAS, 368, 715

[134] Martínez H. J., O'Mill A. L., Lambas D. G., 2006, MNRAS, 372, 253

[135] Martínez H. J., Zandivarez A., Merchán M. E., Domínguez M. J. L., 2002, MNRAS, 337, 1441

[136] Mashchenko S., Wadsley J., Couchman H. M. P., 2008, Science, 319, 174

[137] McGaugh S. S., 2005, ApJ, 632, 859

[138] McGaugh S. S., Schombert J. M., Bothun G. D., de Blok W. J. G., 2000, ApJ, 533, L99

[139] McGaugh S. S., Schombert J. M., de Blok W. J. G., Zagursky M. J., 2010, ApJ, 708, L14

[140] Meneux B. et al., 2009, A&A, 505, 463

[141] Metropolis N., 1985, in Lecture Notes in Physics, Berlin Springer Verlag, Vol. 240, Lecture Notes in Physics, Berlin Springer Verlag, Alcouffe R., Dautray R., Forster A., Ledonois G., Mercier B., eds., p. 62

[142] Milosavljević M., Miller C. J., Furlanetto S. R., Cooray A., 2006, ApJ, 637, L9

[143] Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution

[144] Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319

[145] More S., van den Bosch F. C., Cacciato M., Mo H. J., Yang X., Li R., 2009, MNRAS, 392, 801

[146] More S., van den Bosch F. C., Cacciato M., Skibba R., Mo H. J., Yang X., 2011, MNRAS, 410, 210

[147] Moster B. P., Naab T., White S. D. M., 2013, MNRAS, 428, 3121

[148] Moster B. P., Somerville R. S., Maulbetsch C., van den Bosch F. C., Macciò A. V., Naab T., Oser L., 2010, ApJ, 710, 903

[149] Moustakas J. et al., 2013, ApJ, 767, 50

[150] Muñoz-Cuartas J. C., Macciò A. V., Gottlöber S., Dutton A. A., 2011, MNRAS, 411, 584

[151] Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

[152] Neistein E., Li C., Khochfar S., Weinmann S. M., Shankar F., Boylan-Kolchin M., 2011, MNRAS, 416, 1486

[153] Niederste-Ostholt M., Belokurov V., Evans N. W., Peñarrubia J., 2010, ApJ, 712, 516

[154] Norberg P. et al., 2002, MNRAS, 332, 827

[155] Norberg P. et al., 2001, MNRAS, 328, 64

[156] Ogiya G., Mori M., 2012, ArXiv e-prints

[157] Olsen K. A. G., Zaritsky D., Blum R. D., Boyer M. L., Gordon K. D., 2011, ApJ, 737, 29

[158] Padilla N., Lambas D. G., González R., 2010, MNRAS, 409, 936

[159] Papastergis E., Cattaneo A., Huang S., Giovanelli R., Haynes M. P., 2012, ApJ, 759, 138

[160] Peñarrubia J., McConnachie A. W., Navarro J. F., 2008, ApJ, 672, 904

[161] Peacock J. A., Smith R. E., 2000, MNRAS, 318, 1144

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

[163] Pérez-González P. G. et al., 2008, ApJ, 675, 234

[164] Poggianti B. M. et al., 2006, ApJ, 642, 188

[165] Pozzetti L. et al., 2010, A&A, 523, A13

[166] Press W. H., Schechter P., 1974, ApJ, 187, 425

[167] Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in FORTRAN. The art of scientific computing

[168] Przybilla N., Tillich A., Heber U., Scholz R.-D., 2010, ApJ, 718, 37

[169] Purcell C. W., Zentner A. R., 2012, JCAP, 12, 7

[170] Reddick R. M., Wechsler R. H., Tinker J. L., Behroozi P. S., 2013, ApJ, 771, 30

[171] Reyes R., Mandelbaum R., Gunn J. E., Nakajima R., Seljak U., Hirata C. M., 2012, MNRAS, 425, 2610

[172] Reyes R., Mandelbaum R., Hirata C., Bahcall N., Seljak U., 2008, MNRAS, 390, 1157

[173] Robertson B., Yoshida N., Springel V., Hernquist L., 2004, ApJ, 606, 32

[174] Robotham A. S. G. et al., 2012, MNRAS, 424, 1448

[175] Roche N. D., Almaini O., Dunlop J., Ivison R. J., Willott C. J., 2002, MNRAS, 337, 1282

[176] Rodríguez-Puebla A., Avila-Reese V., Drory N., 2013, ApJ, 767, 92

[177] Rodríguez-Puebla A., Avila-Reese V., Firmani C., Colín P., 2011, RMxAA, 47, 235

[178] Rodríguez-Puebla A., Drory N., Avila-Reese V., 2012, ApJ, 756, 2

[179] Sackett P. D., 1997, PASA, 14, 11

[180] Sakamoto T., Chiba M., Beers T. C., 2003, A&A, 397, 899

[181] Sales L. V., Wang W., White S. D. M., Navarro J. F., 2013, MNRAS, 428, 573

[182] Santini P. et al., 2012, A&A, 538, A33

[183] Schechter P., 1976, ApJ, 203, 297

[184] Schulz A. E., Mandelbaum R., Padmanabhan N., 2010, MNRAS, 408, 1463

[185] Scoville N. et al., 2007, ApJS, 172, 38

[186] Seljak U., 2000, MNRAS, 318, 203

[187] Shankar F., Lapi A., Salucci P., De Zotti G., Danese L., 2006, ApJ, 643, 14

[188] Shao Z., Xiao Q., Shen S., Mo H. J., Xia X., Deng Z., 2007, ApJ, 659, 1159

[189] Shen S., Mo H. J., White S. D. M., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, MNRAS, 343, 978

[190] Sheth R. K., Hui L., Diaferio A., Scoccimarro R., 2001, MNRAS, 325, 1288

[191] Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

[192] Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

[193] Simha V., Weinberg D. H., Davé R., Fardal M., Katz N., Oppenheimer B. D., 2012, MNRAS, 423, 3458

[194] Skibba R. A., van den Bosch F. C., Yang X., More S., Mo H., Fontanot F., 2011, MNRAS, 410, 417

[195] Smith M. C. et al., 2007, MNRAS, 379, 755

[196] Smith R. E. et al., 2003, MNRAS, 341, 1311

[197] Springel V. et al., 2008, MNRAS, 391, 1685

[198] Stanimirović S., Staveley-Smith L., Jones P. A., 2004, ApJ, 604, 176

[199] Stewart K. R., Bullock J. S., Barton E. J., Wechsler R. H., 2009, ApJ, 702, 1005

[200] Strigari L. E., Bullock J. S., Kaplinghat M., Diemand J., Kuhlen M., Madau P., 2007, ApJ, 669, 676

[201] Strigari L. E., Frenk C. S., White S. D. M., 2010, MNRAS, 408, 2364

[202] Strigari L. E., Wechsler R. H., 2012, ApJ, 749, 75

[203] Stringer M. J., Benson A. J., 2007, MNRAS, 382, 641

[204] Tamburro D., Rix H.-W., Leroy A. K., Mac Low M.-M., Walter F., Kennicutt R. C., Brinks E., de Blok W. J. G., 2009, AJ, 137, 4424

[205] Tamm A., Tempel E., Tenjes P., Tihhonova O., Tuvikene T., 2012, A&A, 546, A4

[206] Tempel E., Saar E., Liivamägi L. J., Tamm A., Einasto J., Einasto M., Müller V., 2011, A&A, 529, A53

[207] Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

[208] Tinker J. L., Conroy C., Norberg P., Patiri S. G., Weinberg D. H., Warren M. S., 2008, ApJ, 686, 53

[209] Tinker J. L., Weinberg D. H., Zheng Z., Zehavi I., 2005, ApJ, 631, 41

[210] Tinker J. L., Wetzel A. R., 2010, ApJ, 719, 88

[211] Tollerud E. J., Boylan-Kolchin M., Barton E. J., Bullock J. S., Trinh C. Q., 2011, ApJ, 738, 102

[212] Toomre A., 1964, ApJ, 139, 1217

[213] Tremaine S. D., Richstone D. O., 1977, ApJ, 212, 311

[214] Trujillo-Gomez S., Klypin A., Primack J., Romanowsky A. J., 2011, ApJ, 742, 16

[215] Tucker D. L. et al., 2000, ApJS, 130, 237

[216] Vale A., Ostriker J. P., 2004, MNRAS, 353, 189

[217] Vale A., Ostriker J. P., 2008, MNRAS, 383, 355

[218] van den Bosch F. C., 2000, ApJ, 530, 177

[219] van den Bosch F. C., Aquino D., Yang X., Mo H. J., Pasquali A., McIntosh D. H., Weinmann S. M., Kang X., 2008, MNRAS, 387, 79

[220] van den Bosch F. C., Tormen G., Giocoli C., 2005, MNRAS, 359, 1029

[221] van der Kruit P. C., Freeman K. C., 2011, ARA, 49, 301

[222] Vogelsberger M., Zavala J., Loeb A., 2012, MNRAS, 423, 3740

[223] Wake D. A., Franx M., van Dokkum P. G., 2012, ArXiv e-prints

[224] Wang J., Frenk C. S., Navarro J. F., Gao L., Sawala T., 2012, MNRAS, 424, 2715

[225] Wang L., Jing Y. P., 2010, MNRAS, 402, 1796

[226] Wang W., White S. D. M., 2012, MNRAS, 424, 2574

[227] Wang Y., Yang X., Mo H. J., van den Bosch F. C., Katz N., Pasquali A., McIntosh D. H., Weinmann S. M., 2009, ApJ, 697, 247

[228] Watson D. F., Berlind A. A., McBride C. K., Hogg D. W., Jiang T., 2012, ApJ, 749, 83

[229] Watson D. F., Berlind A. A., Zentner A. R., 2011, ApJ, 738, 22

[230] Watson D. F., Berlind A. A., Zentner A. R., 2012, ApJ, 754, 90

[231] Watson D. F., Conroy C., 2013, ArXiv e-prints

[232] Wechsler R. H., Bullock J. S., Primack J. R., Kravtsov A. V., Dekel A., 2002, ApJ, 568, 52

[233] Wei L. H., Kannappan S. J., Vogel S. N., Baker A. J., 2010, ApJ, 708, 841

[234] Weinberg D. H., Colombi S., Davé R., Katz N., 2008, ApJ, 678, 6

[235] Weinberg M. D., 1999, in Astronomical Society of the Pacific Conference Series, Vol. 165, The Third Stromlo Symposium: The Galactic Halo, Gibson B. K., Axelrod R. S., Putman M. E., eds., p. 100

[236] Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006, MNRAS, 366, 2

[237] Wetzel A. R., Tinker J. L., Conroy C., 2012, MNRAS, 424, 232

[238] White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

[239] White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

[240] Wilkinson M. I., Evans N. W., 1999, MNRAS, 310, 645

[241] Wojtak R., Mamon G. A., 2013, MNRAS, 428, 2407

[242] Wolf J., Bullock J. S., 2012, ArXiv e-prints

[243] Xue X. X. et al., 2008, ApJ, 684, 1143

[244] Yang X., Mo H. J., Jing Y. P., van den Bosch F. C., Chu Y., 2004, MNRAS, 350, 1153

[245] Yang X., Mo H. J., van den Bosch F. C., 2003, MNRAS, 339, 1057

[246] Yang X., Mo H. J., van den Bosch F. C., 2008, ApJ, 676, 248

[247] Yang X., Mo H. J., van den Bosch F. C., 2009, ApJ, 695, 900

[248] Yang X., Mo H. J., van den Bosch F. C., 2009, ApJ, 693, 830

[249] Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, ApJ, 671, 153

[250] Yang X., Mo H. J., van den Bosch F. C., Zhang Y., Han J., 2012, ApJ, 752, 41

[251] Yniguez B., Garrison-Kimmel S., Boylan-Kolchin M., Bullock J. S., 2013, ArXiv e-prints

[252] Yoo J., Weinberg D. H., Tinker J. L., Zheng Z., Warren M. S., 2009, ApJ, 698, 967

[253] Zavala J., 2003, BSc Thesis, National Autonomous University of Mexico

[254] Zavala J., Avila-Reese V., Firmani C., Boylan-Kolchin M., 2012, MNRAS, 427, 1503

[255] Zavala J., Avila-Reese V., Hernández-Toledo H., Firmani C., 2003, A&A, 412, 633

[256] Zavala J., Okamoto T., Frenk C. S., 2008, MNRAS, 387, 364

[257] Zavala J., Vogelsberger M., Walker M. G., 2013, MNRAS, 431, L20

[258] Zehavi I. et al., 2011, ApJ, 736, 59

[259] Zehavi I. et al., 2005, ApJ, 630, 1

[260] Zentner A. R., Berlind A. A., Bullock J. S., Kravtsov A. V., Wechsler R. H., 2005, ApJ, 624, 505

[261] Zentner A. R., Bullock J. S., 2003, ApJ, 598, 49

[262] Zheng Z. et al., 2005, ApJ, 633, 791

[263] Zheng Z., Coil A. L., Zehavi I., 2007, ApJ, 667, 760

[264] Zolotov A. et al., 2012, ApJ, 761, 71

DECLARATION

This Thesis is a presentation of my original work. Where other source of information have been used they have been acknowledge accordingly.

México, Distrito Federal, August 2013

Aldo A. Rodríguez Puebla,
August 28, 2013