

UNIVERSIDAD DE GRANADA

Tesis doctoral

Mayo 2011

Galaxy Evolution with Large-Scale Galaxy Redshift Surveys

[PhD THESIS]

Antonio David Montero Dorta

Instituto de Astrofísica de Andalucía (CSIC)

Memoria de Tesis

presentada en la Universidad de Granada

para optar al grado de Doctor en Física

Director de tesis:

Francisco Prada Martínez

Editor: Editorial de la Universidad de Granada
Autor: Antonio David Montero Dorta
D.L.: GR 1232-2012
ISBN: 978-84-694-6116-7

Resumen

El modelo cosmológico Λ-CDM, aceptado actualmente, describe un Universo en expansión acelerada constituido en un 70% de energía oscura, en un 25% de materia oscura y en sólo un 5% de materia bariónica ordinaria. Dentro de este marco, se ha llegado a comprender la formación de estructuras a gran escala, a través de un proceso jerárquico de fusión de halos de materia oscura. Desafortunadamente, poco se sabe todavía acerca de la física que rige el enfriamiento, colapso y condensación de bariones, así como de los procesos de formación estelar dentro de los pozos de potencial de los halos de materia oscura. Sólo recientemente, con el desarrollo de los cartografiados de galaxias a gran escala y con la ayuda de simulaciones numéricas de N-cuerpos, hemos comenzado a identificar algunos de los procesos que gobiernan la evolución de las galaxias y su interacción con el entorno, entre una maraña de correlaciones entre propiedades galácticas. En los últimos años, la formación y evolución de galaxias se ha convertido en uno de los campos más activos y de mayor desarrollo dentro de la Astrofísica.

Durante mi doctorado, he tenido la oportunidad de abordar el estudio de la formación y evolución de galaxias desde ángulos diferentes aunque complementarios. Mi motivación fundamental ha sido, no obstante, la de proporcionar restricciones observacionales más precisas a los modelos teóricos actuales. Para mayor claridad, he dividido esta tesis, que es por lo tanto un compendio de distintos trabajos, en tres partes. En la primera parte se presenta un estudio sobre la función de luminosidad de galaxias en el universo cercano. Esta propiedad estadística de la población de galaxias ha sido medida con una precisión sin precedentes en el SDSS. La segunda parte está orientada al estudio de la evolución de galaxias desde $z \sim 1$. En concreto, se analizan propiedades estadísticas fundamentales de la población de galaxias a bajo corrimiento al rojo y se describe un método para obtener masas estelares con el cartografiado fotométrico de galaxias ALHAMBRA. Además, se proporciona una medición de la relación AGN - densidad del entorno a $z \sim 1$ en el DEEP2 GRS y se discute la evolución de la misma utilizando el SDSS. Finalmente, el avance en las teorías de formación y evolución de galaxias requiere de una mejora de las técnicas de cartografiado de galaxias. En la última parte de esta tesis trato de contribuir a este objetivo analizando posibles optimizaciones para espectrógrafos multi-objeto de fibras de nueva generación. Este trabajo se enmarca dentro del desarrollo del instrumento SIDE, que fue propuesto para el telescopio de 10 metros Gran Telescopio Canarias.

A bajo corrimiento al rojo, la caracterización de la población de galaxias ha sido posible gracias al desarrollo de cartografiados a gran escala de galaxias. En particular, el SDSS proporciona información espectroscópica acerca de aproximadamente 10^6 galaxias, permitiendo un estudio detallado de las

principales propiedades estadísticas de la población de galaxias. La función de luminosidad de galaxias es, sin duda, una de las más fundamentales de entre todas estas propiedades. En esta tesis se aprovecha el enorme incremento en la estadística de galaxias proporcionado por el SDSS Sixth Data Release, en comparación con previas estimaciones de la función de luminosidad, para presentar la estimación más detallada de la misma hasta la fecha, en todas las bandas fotométricas del SDSS. Se han encontrado, en este sentido, discrepancias importantes con las estimaciones previas de Blanton et al. (2003b). En concreto, se han obtenido pendientes más pronunciadas en el rango de galaxias débiles ($\alpha = -1.26$ en la banda ${}^{0.1}r$) y un exceso considerable de galaxias brillantes con respecto al ajuste de Schechter en las bandas más azules (de ~ 1.7 dex a magnitud $M_{{}^{0.1}u} \simeq -20.5$). Este exceso, que se ha denominado *bright-end bump* o bulto en la parte brillante, parece estar asociado con AGNs y galaxias con alta tasa de formación estelar. Además, se ha llevado a cabo una comparación preliminar con modelos semi-analíticos de formación y evolución de galaxias. Esto último se ha realizado en colaboración con el Carlton Baugh y su equipo del ICC y la Universidad de Durham.

Durante los últimos años, he estado involucrado en el desarrollo del proyecto Advance Large Homogeneous-Area Medium-Band Redshift Astronomical Survey está siendo liderado por el Instituto de Astrofísica de Andalucía. ALHAMBRA es un cartografiado fotométrico de galaxias dise
nado específicamente para el estudio de la evolución de galaxias, en el que se utilizan 20 filtros de banda estrecha que cubren completamente el rango óptico, además de las bandas JHK del infrarrojo cercano. ALHAMBRA nos permite detectar galaxias hasta magnitud $I_{AB} \sim 25$. En esta tesis se presenta un análisis preliminar de la población de galaxias de ALHAMBRA a bajo corrimiento al rojo, que, sin duda, facilitará notablemente la explotación científica del cartografiado. Las propiedades estadísticas fundamentales de esta población de galaxias, incluyendo número de cuentas de galaxias, distribuciones de color y de magnitud y funciones de luminosidad, han sido obtenidas y comparadas con otros cartografiados, especialmente con el SDSS. Además de esto, se ha llevado a cabo un análisis de síntesis de poblaciones estelares, utilizando el código Flexible Stellar Population Synthesis. Como resultado, se presenta la primera estimación de masas estelares fotométricas en ALHAMBRA. Nuestros resultados, en lo que a las propiedades estadísticas fundamentales de la población de galaxias se refieren, son consistentes con otros cartografiados. Además, demuestran el potencial de ALHAMBRA para aportar luz en el campo de la evolución de galaxias.

A alto corrimiento al rojo, la obtención de restricciones observacionales para los modelos de formación y evolución de galaxies se ve notablemente dificultada por los elevados tiempos de observación requeridos y por importantes efectos de selección. En estos modelos se asume que tanto los AGNs como el entorno juegan papeles determinantes. La actividad nuclear podría ser responsable, en este sentido, de la eliminación de los flujos de enfriamiento en galaxias masivas, y la consecuente paralización de la formación estelar. La influencia del entorno en las galaxias, por otro lado, está observacionalmente bien establecida (como ejemplo se puede citar la relación color-densidad o la relación morfología-densidad). En esta tesis se ha medido la dependencia de la fracción de AGNs con la densidad de su entorno local a $z \sim 1$, utilizando datos espectroscópicos tomados del DEEP2 Galaxy Redshift Survey, así como datos en rayos-X de Chandra del All-Wavelength Extended Groth Strip International Survey (AEGIS). Como resultado, se proporcionan evidencias que apuntan a que los AGNs de tipo LINER a alto corrimiento al rojo en el DEEP2 tienden a favorecer entornos relativamente densos, con respecto

al comportamiento global de la población de galaxias rojas. Por contra, los AGNs de tipo Seyfert o de rayos-X a $z \sim 1$ muestran poca (o nula) dependencia con la densidad del entorno, dentro de la misma muestra. Se han comparado estos resultados con una muestra de AGNs locales extraída del SDSS. A diferencia de lo que se observa a alto corrimiento al rojo, tanto LINERs como Seyferts en el SDSS tienden a encontrarse en entornos menos densos. Por otro lado, a $z \sim 1$ los Seyferts y LINERs rojos son aproximadamente igual de abundantes. A $z \sim 0$, sin embargo, se observa que la población de Seyferts rojos se ha visto reducida en un factor ~ 7 con respecto a la de LINERs. Este trabajo se ha llevado a cabo en colaboración con el Marc Davis, Darren Croton y el resto del equipo de DEEP2, durante una visita a la Universidad de California en Berkeley.

El avance en el campo de la formación y evolución de galaxias en las próximas décadas dependerá de nuestra capacidad de obtener muestras más amplias, más profundas y menos sesgadas de galaxias. Para entender las correlaciones entre las propiedades de las galaxias y para poder conectar distintas poblaciones a diferentes corrimientos al rojo se necesitan nuevos cartografiados espectroscópicos de galaxias. En la parte final de esta tesis se estudia la optimización de cartografiados para espectrógrafos de fibras y de gran campo de nueva generación. En concreto, se presenta un algoritmo optimizado para la asignación de fibras a galaxias previamente seleccionadas, que asegura que el máximo número de estas galaxias se observe, para un campo determinado, en las primeras exposiciones del espectrógrafo. Utilizando catálogos de galaxias distribuidas aleatoriamente así como catálogos artificiales de galaxias se ha estimado que la ganancia proporcionada por este algoritmo en comparación con un posicionamiento de fibras aleatorio llega a ser del 2% para los primeras exposiciones. Esto implicaría, para un cartografiado de nueva generación como BigBOSS, observar varios cientos de miles de objetos más o, incluso, reducir el area cubierta en ~ 350 grados cuadrados. Además, se han analizado posibles optimizaciones adicionales. En particular, se demuestra que la posibilidad de rotar el plano focal del instrumento mejora la eficiencia del proceso en un $\sim 3.5 - 4.5\%$, incluso si sólo pequeños ajustes son permitidos (de hasta 2 grados). Para instrumentos que permiten rotaciones mayores la ganancia esperada asciende a un $\sim 5 - 6\%$. Estos resultados, por lo tanto, apoyan fuertemente la conveniencia de permitir rotación del plano focal en futuros espectrógrafos, en lo que a la eficiencia del proceso de posicionamiento de fibras se refiere. Este trabajo fue concebido para el proyecto Super Ifu Deployable Experiment (SIDE), un espectrógrafo multi-objeto de fibras propuesto para el GTC.

Summary

The currently favored Λ-CDM model of cosmology describes an accelerated expanding Universe containing 70% of dark energy, 25% of collisionless cold dark matter and only 5% of ordinary baryonic matter. Within this framework, the formation of large-scale structure by the hierarchical merging of dark matter halos is relatively well understood. Unfortunately, the complex physics involved in the cooling, in-fall and condensation of baryons and the formation of stars within the potential wells of dark matter halos is still severely unconstrained. Only recently, with the advent of large-scale galaxy redshift surveys and with the help of N-body numerical simulations, have we commenced to identify some of the processes that drive the evolution of galaxies and their interaction with the environment, from an entangled web of observed correlations between galaxy properties. In the last years, the formation and evolution of galaxies has become one of the most active and rapidly expanding fields in Astrophysics.

During my PhD, I had the opportunity to approach the galaxy formation and evolution picture from different but complementary angles. In this sense, my thesis work was motivated by the challenge of providing more accurate observational constraints to current theories. For the sake of clarity, I have divided this dissertation, which is a compendium of different but related projects, into three parts. In the first part, I present and discuss in detail the luminosity function of galaxies in the nearby universe, which I have measured with unprecedented accuracy in the SDSS. The second part is devoted to galaxy evolution from $z \sim 1$. In particular, in this part I analyze some fundamental statistical properties of the low-redshift galaxy population and describe a method for obtaining stellar masses in the ALHAMBRA photometric survey. In addition, I provide a measurement of the the AGN - density relation at $z \sim 1$ in the DEEP2 Galaxy Redshift Survey, and discuss its evolution down to $z \sim 0$ using the SDSS. Finally, constraining galaxy formation and evolution theories in the near future requires the development of new survey spectrographs and the optimization of galaxy survey techniques. In the last part of this dissertation, I contribute to this goal by discussing survey optimizations for next-generation fiber-fed spectrographs such as the SIDE instrument, which was proposed for the 10-meter Gran Telescopio Canarias, or BigBOSS, for the 4-meter Mayall Telescope.

At low redshift, the characterization of the galaxy population has benefitted extensively from the development of large-scale galaxy redshift surveys. In particular, the SDSS provides photometric and spectroscopic information for $\sim 10^6$ galaxies, allowing for a detailed study of the main statistical properties of the galaxy population. The luminosity function of galaxies is one of the most fundamental of these properties. In this dissertation, I take advantage of the huge increase in galaxy statistics

provided by the Sixth Data Release of the SDSS to present the most accurate estimation of the low-redshift luminosity function of galaxies to date. Interestingly, I find several remarkable deviations as compared to the previous SDSS work of Blanton et al. (2003b). Namely, I obtain steeper faint end slopes in all bands ($\alpha = -1.26$ in the $^{0.1}r$ band) and a remarkable bright-end excess with respect to the Schechter fit in the bluer bands (of ~ 1.7 dex at $M_{^{0.1}u} \simeq -20.5$). Such an excess, that I call *bright-end bump*, seems to be associated with AGN and star-forming galaxies. A preliminary comparison with semi-analytic models of galaxy formation and evolution is also provided. This comparison has been performed in collaboration with Carlton Baugh and his team at the ICC and Durham University.

During the last years, I have been involved in the Advance Large Homogeneous-Area Medium-Band Redshift Astronomical Survey, which is lead by the Instituto de Astrofísica de Andalucía. ALHAMBRA is a photometric survey specifically designed for the study of galaxy evolution, that uses 20 narrow-band filters covering the entire optical range plus the near-infrared JHK bands, and reaches down to magnitude $I_{AB} \sim 25$. In this dissertation, I provide a preliminary low-redshift analysis that will facilitate the scientific exploitation of the survey. The main statistical properties of the low-z ALHAMBRA galaxy population, including galaxy number counts, color and magnitude distributions and luminosity functions have been obtained and compared with other surveys, especially with the SDSS. In addition, I have carried out a stellar population synthesis analysis, using the Flexible Stellar Population Synthesis code, which has allowed me to compute photometric stellar masses in ALHAMBRA. Our results on the main statistical properties of the ALHAMBRA low- redshift galaxy population are in good agreement with other surveys and confirm the potential of the survey to shed light into the evolution of galaxies.

At high redshift, providing observational constraints to galaxy formation and evolution models is remarkably hindered by telescope time requirements and strong selection effects. In these models, both active galactic nuclei (AGN) and environment are assumed to play key roles. Nuclear activity could be responsible for the suppression of cooling flows in massive galaxies and the consequent quenching of star formation. The influence of environment on galaxies is supported by undisputed observational evidence (i.e. color-density relation, morphology-density relation). I have measured the dependence of the AGN fraction on local environment at $z \sim 1$, using spectroscopic data taken from the DEEP2 Galaxy Redshift Survey, and Chandra X-ray data from the All-Wavelength Extended Groth Strip International Survey (AEGIS). I provide evidence that high redshift LINERs in DEEP2 tend to favour higher density environments relative to the red population from which they were drawn. In contrast, Seyferts and X-ray selected AGN at $z \sim 1$ show little (or no) environmental dependencies within the same underlying population. I compare these results with a sample of local AGN drawn from the SDSS. Contrary to the high redshift behaviour, I find that both LINERs and Seyferts in the SDSS show a slowly declining red sequence AGN fraction towards high density environments. Interestingly, at $z \sim 1$ red sequence Seyferts and LINERs are approximately equally abundant. By $z \sim 0$, however, the red Seyfert population has declined relative to the LINER population by over a factor of 7. This work was created during a research visit at the University of California at Berkeley, in collaboration with Marc Davis, Darren Croton and the rest of the DEEP2 team.

Progress in the galaxy formation and evolution field in the next decades will depend upon our capacity to collect deeper, larger and less biased galaxy samples. High-redshift spectroscopic surveys

are needed to disentangle correlations between galaxy properties and link different galaxy populations at different redshifts. In the last part of this dissertation, I discuss survey optimizations for next-generation wide-field fiber-fed spectrographs. In particular, I present an optimized algorithm for assigning fibers to targets that ensures that the maximum number of targets in a given target field is observed in the first few tiles (spectrograph exposures). Using randomly distributed targets and mock galaxy catalogs, I have estimated that the gain provided by our algorithm as compared to a random assignment can be as much as 2% for the first tiles. This would imply for a next-generation survey like BigBOSS saving for observation several hundred thousand objects or, alternatively, reducing the covered area in $\sim 350\ deg^2$. Additional optimizations of the fiber positioning process are also discussed. In particular, I show that allowing for rotation of the focal plane can improve the efficiency of the process in $\sim 3.5 - 4.5\%$ even if only small adjustments are permitted (up to 2 deg). For instruments that allow large rotations of the focal plane the expected gain increases to $\sim 5 - 6\%$. These results, therefore, strongly support focal plane rotation in future spectrographs, as far as the efficiency of the fiber positioning process is concerned. This work was primarily conceived for the Super Ifu Deployable Experiment (SIDE), a fiber-fed multi-object spectrograph proposed for the GTC.

Index

1

Introduction

1.1 The establishment of the standard cosmological model

By the spring of 1920, Harlow Shapley was a young and promising astronomer working at the Mount Wilson Observatory, where he had been hired by George Ellery Hale. At his mid-thirties, Shapley was well-known within the astronomical community for his studies on globular clusters, which were, by that time, very tightly-bound collections of stars that orbited the center of our galaxy. By determining the distances to these stellar systems using the period-luminosity relation for Cepheid variable stars, discovered by Henrietta Swan Leavitt, Shapley was the first to realize that the Milky Way was remarkably larger than previously believed and that the Sun was not, at least necessarily, at its center. He was also a passionate defender of the idea that the Universe extended no further than the Milky Way and that spiral nebulae were small gaseous objects laying within our own galaxy and not composed by stars at all.

Herber Doust Curtis was by that time an experienced astronomer of the Lick Observatory, where he had spent the last eighteen years and specialized in the study of spiral nebulae. He, in fact, continued the survey of nebulae initiated by James Edward Keeler with the Lick Crossley Telescope. He was also an enthusiastic convert to the *Island Universe* theory. According to this theory, spiral nebulae were located far beyond the ends of the Milky Way. They were, actually, independent galaxies in their own right. Such an idea was not really new, but it had been discarded by a considerable fraction of the astronomical community due to reported observational inconsistencies. In fact, the picture of visible nebulae as distant Milky Ways was first envisaged (at least in western cultures) by Thomas Wright in his *An original theory or new hypothesis of the universe* in 1750. A few years later, in 1755, Immanuel Kant took this idea to write in *Universal Natural History and Theory of the Heavens*:

"It is far more natural and conceivable to regard them as being not such enormous single stars but systems of many stars, whose distance presents them in such a narrow space that the light which is individually imperceptible from each of them, reaches us, on account of their immense multitude, in a uniform pale glimmer. Their analogy with the stellar system in which we find ourselves, their shape, which is just what it ought to be according to our theory, the feebleness of their light which demands a presupposed infinite distance: all this is in perfect harmony with the view that these elliptical figures

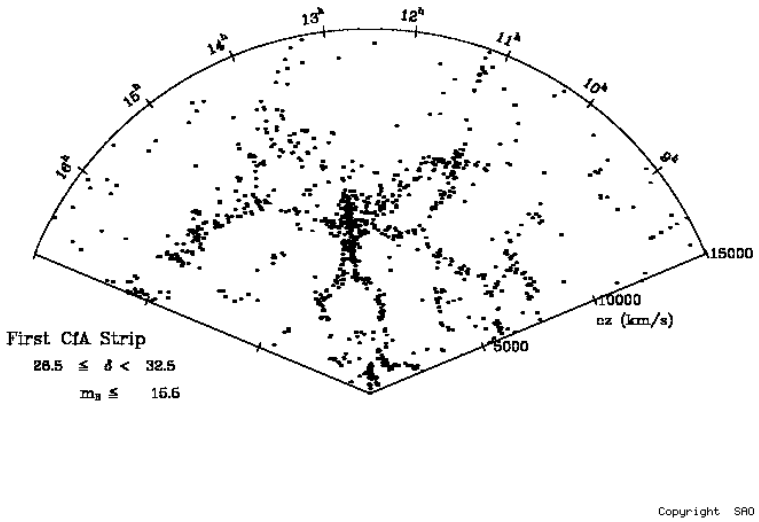

Figure 1.1 The distribution of galaxies in the first sample of the CfA Redshift Survey ($\sim 1,000$ galaxies). The CfA RS provided the first vision of large-scale structure in the Universe, with galaxies accumulating in thin filaments surrounding under-dense voids. Such a distribution is consistent with the cold dark matter paradigm.

are just island universes and, so to speak, Milky Ways, like those whose constitution we have just unfolded."

Shapley and Curtis represent the two ways of understanding the Universe as a whole at the beginning of the twentieth century. On 26 April 1920, they both joined the National Academy of Sciences Meeting, which took place at the Smithsonian Museum of Natural History, in Washington (USA), for a discussion entitled "The Scale of the Universe". In this session, that would pass into literature as the *Great Debate*, Curtis provided strong observational evidences backing up the idea that spiral nebulae were not simply gas accumulations somehow gravitationally bound to the Milky Way, but independent systems containing millions of stars, just like our own galaxy. In particular, Curtis noticed the huge recession velocities of nebulae as compared to typical star velocities observed, the similarities of their spectrum to that of stellar clusters and the insuperable difficulty to place them within a stellar evolution scheme. Moreover, he cited the high rate of novae within these objects and the presence of obscured regions, resembling those found at low galactic latitudes in the Milky Way. Herber Curtis has gone down in history as the winner of the Great Debate. More importantly, the famous discussion at the Smithsonian Museum of Natural History represents a change in the global perception of the Universe and a sort of initial point for a new discipline: the Extragalactic Astronomy. The final confirmation for the island universes paradigm came a few years later, when Edwin Hubble used the 2.5 - meter Hooker Telescope, precisely at the Mount Wilson Observatory, to resolve the outer regions of several spirals, including the Andromeda Nebula, discovering that they were populated by

millions of individual stars (Hubble, 1922). In particular, he found Cepheids stars and measuring their distance, he realized that spiral nebulae were too far away to be part of the Milky Way (Hubble, 1925).

The confirmation of the existence of other galaxies in the Universe was not the only outstanding discovery from Hubble. Using the Hooker Telescope, he also provided the first morphological classification of galaxies (at that time, still nebulae), establishing 3 main types (with several sub-types): ellipticals, spirals and lenticulars (Hubble, 1926). Hubble is mostly known, however, for his contribution to the observational confirmation of another extraordinary discovery: the expansion of the Universe. In 1916, Albert Einstein had presented his *general theory of relativity*, providing a unified description of gravity as a geometric property of *space-time* and giving rise to modern Cosmology. A few years later, in the early 1920's, Alexander Friedmann was able to solve Einsten's general relativity field equations. The classical solution, which is today dubbed Friedman-Lemaître-Robertson-Walker (FLRW) metric after the independent works of Friedmann, Georges Lemaître, Howard Robertson and Arthur Walker, describes a homogeneous and isotropic-expanding (or in principle, contracting) universe. Edwin Hubble and Milton Humason were able to corroborate such idea of an expanding universe when they found a velocity-distance relation for extragalactic nebulae, i.e. the Hubble's Law (Hubble & Humason, 1931). Combining their own measurements of galaxy distances (based on Swan Leavitt's period-luminosity relationship for Cepheids) with Vesto Slipher's measurements of the redshifts associated with the galaxies, Hubble and Humason discovered a rough proportionality of the distances of galaxies with their recession velocities with respect to the Earth.

The discovery that the Universe as a whole could actually be described by a set of mathematical equations motivated the development of both theoretical and observational Cosmology. In the early 1930's, George Lemaître gave a step further and suggested that the observed expansion of the Universe in forward time should require that the Universe contracted backwards in time. He speculated that the Universe might continue to do so until it could contract no further, bringing all the mass of the Universe into a single point, that he called the "primeval atom" (Lemaître, 1931). Within this model, which is consistent with the Friedman-Lemaître-Robertson-Walker (FLRW) metric, the Universe originated as an extremely hot and dense quantum containing both time and space. Some years later, Fred Hoyle would coin the name *Big Bang* to refer to Lemaître's idea. The Big Bang theory, however, was not totally accepted at the beginning, and several other so-called non-standard cosmologies arose in the 1930's. The Milne model, proposed by Edward Arthur Milne (Milne, 1932), rejected the general relativity formalism (curvature of space, expanding space) and obeyed the rules of the special relativity and the *Cosmological Principle*, which states that the properties of the Universe are the same for all observers. It provided, however, an unconvincing explanation for gravity. The oscillatory universe, advocated by Albert Einstein and Richard Tolman, or the tired light hypothesis, suggested by Zwicky in order to explain Hubble's Law are other examples of alternative cosmological models.

A much more reasonable alternative to the Big Bang theory was provided by Bondi & Gold (1948) and Hoyle (1948), who attempted to reconcile Einstein's general relativity, the cosmological principle and Hubble's Law. The so-called steady-state model guarantees that the Universe expands without violating (approximately) the cosmological principle, but requires a certain, undetectable amount of matter to be created continuously in order to keep the density of matter constant. The Universe is

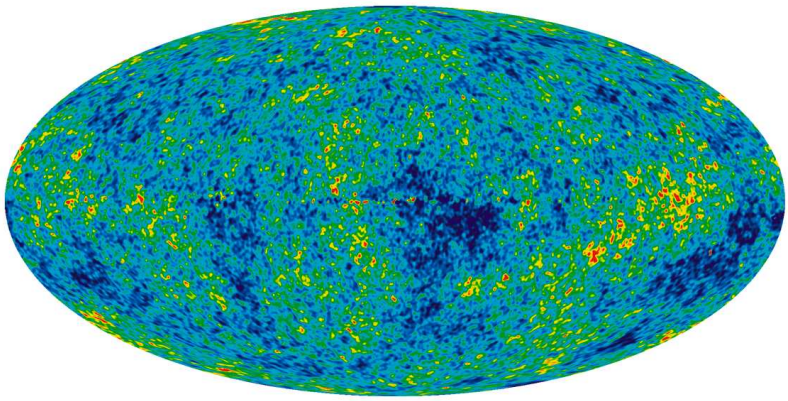

Figure 1.2 A detailed, all-sky picture of the early universe created from seven years of WMAP data. The image reveals 13.7 billion year old temperature fluctuations (shown as color differences) that correspond to the seeds that grew to become the galaxies. The signal from our Galaxy was subtracted using the multi-frequency data. This image shows a temperature range of ±200 microKelvin. Image taken from http://map.gsfc.nasa.gov/resources/imagetopics.html.

assumed to have no beginning and no end in both time and space. Even though the steady-state model received considerable support from the cosmology community, the validity of the Big Bang theory was progressively consolidated by both theoretical efforts and observational evidences. Much of this success is due to George Gamow, who advocated for the Big Bang theory and discovered that the observed abundances of hydrogen and helium in the Universe could be explained by *primordial nucleosynthesis* during Big Bang (Gamow, 1946). By 1960, it was difficult to accommodate several observational results within the scheme of an unchanged universe. In particular, the number counts of radio-emiting sources (quasars) or the apparent brightness and angular diameters of distant galaxies seemed to be inconsistent with such a model. The confirmation of the Big Bang theory as the most plausible cosmological model came in 1964 with the discovery of the cosmic microwave background (CMB) by Arno Penzias and Robert Wilson. The CMB, a thermal radiation filling the Universe almost uniformly and isotropically, had been theoretically predicted by Gamow and his collaborators Ralph Alpher and Robert Herman more than a decade before. The interpretation of the CMB as a *relic* radiation propagated from an epoch when the Universe was massively hotter ruled out the steady-state model once and for all.

In the 1970's, when the Big Bang theory was mostly accepted, the necessity to advocate for the existence of a considerable amount of non-baryonic *dark matter* in the Universe was well established within the community. Back in the 1930's, Zwicky had to claim for the need of such a mysterious matter in order to explain the circular velocities of galaxies in the Coma cluster, that must satisfy the virial theorem (Zwicky, 1937). The work of Vera Rubin and her associates in the decade of 1970 revitalized Zwicky's old ideas. Namely, she discovered that the velocity profiles of stars and

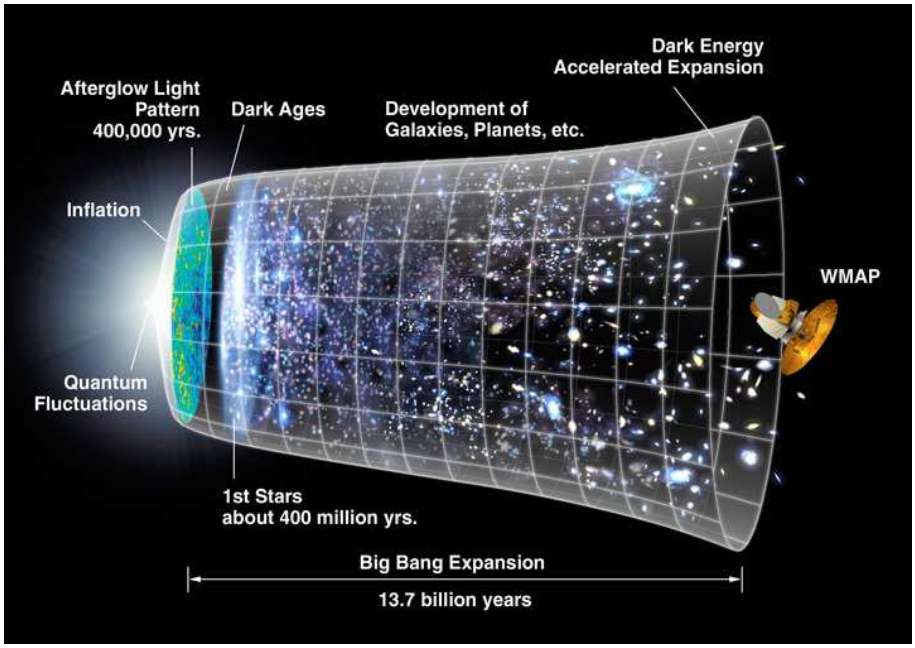

Figure 1.3 The timeline of the Universe, from the Big Bang to the present. Lines around the cylinder are about a billion years apart. The increasing diameter of the cylinder represents the expansion of the Universe. Current theory suggests a very rapid expansion at the beginning and observations show that the expansion has recently been accelerating.

HII regions in spiral galaxies remained roughly constant to great distances from the galactic center, instead of decreasing following the expected Keplerian dynamical behaviour (Rubin & Ford, 1970; Rubin et al., 1980). A number of observational evidences to support the existence of dark matter were provided in following years, including observations of X-ray emission produced by the hot gas in clusters of galaxies, the image magnifications produced by galaxy clusters acting as gravitational lenses or the morphology of the CMB. The majority of the astronomical community today believes that dark matter constitutes $\sim 75\%$ of the total matter in the Universe. Regarding its nature, there is a consensus that it must be formed by collisionless particles whose velocity dispersion in the early Universe is so small that fluctuations of galactic size or larger are not damped by free streaming (*cold dark matter*, CDM, Blumenthal et al. 1984). A CDM Universe would create structure in a hierarchical mode, with small objects collapsing first and merging in a continuous hierarchy to form more and more massive objects. Such a theoretical prediction was consistent with the distribution of galaxies in the Center for Astrophysics Redshift Survey (CfA RS, Huchra et al. 1983), which is considered the first proper galaxy redshift survey. In Figure 1.1, we show the first slice of data from the CfA RS, consisting of $\sim 1,000$ galaxies. Note that galaxies accumulate in thin filaments, which are surrounded by under-dense regions. The CDM paradigm would be further corroborated by subsequent large-scale galaxy surveys and by precise measurements of the CMB, obtained using satellites such as the Cosmic Background Explorer (COBE, see Fixsen et al. 1996) or the Wilkinson Microwave Anisotropy Probe (WMAP, see Komatsu et al. 2009). Find a detailed all-sky map of the CMB from WMAP in Figure 1.2.

In the early 1990's, one thing was fairly certain about the expansion of the Universe: gravity

must slow the expansion as time went on. However, in 1998, two independent groups made the latest extraordinary discovery that would change our perception of the Universe. Riess et al. (1998) and Perlmutter et al. (1999), used both ground-base telescopes and the Hubble Space Telescope (HST) to detect Type Ia supernovae (SNe Ia) in very distant objects, findind that 5 Gyrs ago, the Universe was expanding more slowly than it is today. Unavoidably, a form of energy driving such an accelerated expansion must be introduced. This still rather mysterious form of energy was called *dark energy*. Dark energy was contained in Eintein's equations in the form of a cosmological constant, λ (today written Λ), representing a uniform density and pressure, associated with the space-time itself (vacuum energy) and tending to oppose gravity. Other possibilities, which are commonly called *quintessence*, have been suggested to describe the nature of dark energy (see Peebles & Ratra 2003). A good review on the present status of cosmological discoveries, including dark matter and dark energy, can be found in Martínez & Trimble (2009).

After a century of immense theoretical and observational effort which has changed completely our vision of the Universe, we have come to the standard cosmological model, i.e. the Λ - CDM paradigm, which combines the Big Bang theory with a universe dominated by dark energy and cold dark matter. In particular, the Λ-CDM model describes a flat universe containing $\sim 5\%$ of baryonic matter, $\sim 23\%$ of dark matter and $\sim 72\%$ of dark energy. In Figure 1.3, we summarize in a cartoon the history of the Universe as described by the the standard model.

1.2 Galaxy formation in a hierarchical Universe

The establishment of the current cosmological model took place in parallel with the development of a theoretical framework for the formation and evolution of galaxies, which has recently become one of the most active and rapidly expanding fields of Astronomy. In this sense, the formation and evolution of galaxies, as we will briefly outline here, must not be separated from its cosmological context, i.e. the distribution of cold dark matter in an expanding Universe.

In a purely homogenous and isotropic Universe that perfectly satisfies the cosmological principle, the generation of structures such as galaxies, clusters or large filaments would simply not take place. In order to explain the observed large-scale structure of the Universe, we need to introduce some *initial perturbations*, from which structures can grow through *gravitational instability*. This idea originates in the work of Jeans (1902), who showed that, under appropriate conditions, a cloud would become unstable and begin to collapse when it lacked sufficient gaseous pressure support to balance the force of gravity. Unfortunately, general relativity alone is expectedly not capable of providing a complete description of the *Early Universe*, where initial perturbations should originate, as the prevailing temperature (energy) is so high that quantum effects are expected to be relevant. Even in the absence of a coherent theory for the generation of primordial density fluctuations, much effort was devoted from the 1940's to the 1970's to computing the linear evolution of these perturbation (e.g. Lifshitz 1946; Silk 1968; Peebles & Yu 1970). In the early 1970's, Harrison (1970) and Zel'Dovich (1970) independently predicted that only a certain type of fluctuations, with the same dimensionless amplitude on every scale, was consistent with the observed large-scale structure. A rather plausible (albeit not free from problems) explanation for the origin of initial perturbations came with the

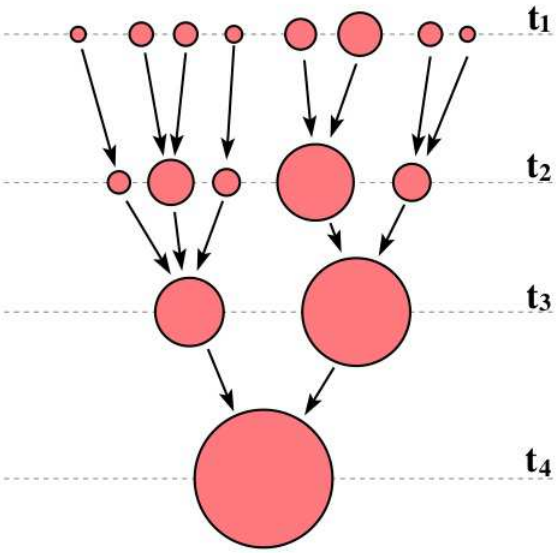

Figure 1.4 A scheme of a merger tree illustrating the merger history of a dark matter halo (where time goes downwards). The size of each circle represents the mass of the halo. Within the framework of the cold dark matter paradigm, dark matter haloes are formed by the merging of smaller progenitors. This figure was taken from Mo et al. (2010).

inflationary theory in the 1980's, which proposes that the Universe underwent a phase of exponential expansion driven by the vacuum energy of one or more quantum fields (see e.g. Guth 1981). Some of the predictions of this theory regarding the nature of these perturbations (adiabatic, Gaussian) along with the validity of the *Harrison-Zel'dovich initial perturbation spectrum* were corroborated in the 1990's, and subsequently, by the CMB experiments.

As mentioned before, in an expanding Universe dominated by non-relativistic matter, perturbations grow in time. Due to gravitational instability, a region with an initial density slightly higher than the mean will pull matter more strongly than average, thus becoming more over-dense. Conversely, under-dense regions will become progressively emptier as matter flows away from them. This happens in the so-called *linear regime*, where the density of the perturbation as compared to the mean is still small. In such conditions, the physical size of over-dense regions grows as well, due to the expansion of the Universe. However, in order to link the initial perturbations to the non-linear structures we see today, we need to understand the non-linear evolution of perturbations. In essence, when the density of the perturbation is large enough, the perturbation breaks away from expansion and starts to collapse ("turn-around"). Remarkable advances in the analytical computation of the collapse of perturbations and the formation of structures were made in the 1970's, such as the *Zel'Dovich approximation* for the generic collapse of a perturbation (Zel'Dovich, 1970) or the spherical collapse approximation Gunn & Gott (1972). Taking this last assumption and a gaussian field of perturbations, Press & Schechter (1974) were able to derive a formalism that can be used to compute the *cosmic mass function* of collapsed objects, which is in relatively good agreement with modern N-body cosmological simulations (with certain remarks, see e.g. Sheth et al. 2001; Sheth & Tormen 2002; Betancort-Rijo

& Montero-Dorta 2006).

Within the framework of the favored λ-CDM model, each initial perturbation contains both baryonic and non-baryonic dark matter in roughly their universal proportions. In this picture, collisionless cold dark matter would collapse first by a process called *violent relaxation*, that will lead the collapsed object to a quasi-equilibrium state. Such structures, which are called *dark matter halos*, are supposed to play a fundamental role in the galaxy formation process. The current theory of galaxy formation within dark matter halos is based on the theoretical research carried out in the late 1970's and the early 1980's. In particular, White & Rees (1978) proposed a two-stage theory for galaxy formation, where dark matter halos form first and galaxies form as a result of the cooling and fall of gas into the potential well of halos. Importantly, the process of halo formation is accepted today to be highly hierarchical, as it is schematically illustrated in Figure 1.4, which is taken from Mo et al. (2010). The idea, which is a theoretical implication of the cold dark matter assumption, is that dark matter haloes are formed by the merging of smaller progenitors.

The net effect of the cooling processes is that baryonic matter separates from dark matter and progressively accumulates in a *protogalaxy* at the center of the halo. The formation of centrifugally supported galaxy disks is assumed in current models to be caused by the conservation of the small angular momentum of the material as it flows inwards. Depending on the efficiency of the cooling, the gas cloud may fragment into small, high-density cores that may eventually form stars and give rise to a visible galaxy. Unfortunately, not much is known about the physics of *star formation process* itself. In particular, fundamental unknowns are the time-scale and the efficiency of the process, in terms of the fraction of the mass of the cloud that turns into stars. Regarding the distribution of mass with which stars form, i.e. *the initial mass function* (IMF), the consensus is that it follows a power law with the logarithm of mass, with an exponent hovering around ~ 1.3 in the mass range $10 \lesssim M/M_{\odot} \lesssim 1$ (Salpeter, 1955; Kroupa, 2001; Chabrier, 2003; van Dokkum, 2008). However, uncertainties remain regarding the fundamental low-mass end ($M/M_{\odot} \lesssim 1$) of the IMF, which can contribute remarkably to the mass of the galaxy. Another important uncertainty derives from the common assumption that the IMF is similar for any galaxy, independently of the environment (ubiquity) or the evolutionary state of the Universe (universality).

The two-stage theory for galaxy formation of White & Rees (1978), from which most current models derive, includes *feedback processes* in order to regulate star formation. This was motivated by observations indicating that the star formation process was highly inefficient. In particular, White & Rees (1978) invoked the ideas of Larson and collaborators, who carried out the first numerical simulations of galaxy formation (see Larson 1974 and subsequently), arguing that feedback from supernova explosions might explain the low surface brightness and low metallicity (fraction of metals) in dwarf galaxies. In current models of galaxy formation, feedback processes are assumed to play a fundamental role, although their nature is still severely unconstraint. In this sense, many of the basic elements of galaxy formation in a CDM Universe were already in place in the early 1980's (see Efstathiou & Silk 1983; Blumenthal et al. 1984). The development of N-body cosmological simulations in the last decades has allowed for a more detailed study of both the properties of CDM halos and the process of galaxy formation itself. Regarding the latter, it is necessary to mention the semi-analytic models of galaxy formation and evolution (SAMs), which combine the power of N-body simulations

to trace the evolution of dark matter halos with an analytical treatment of all the physical processes involved in the formation of galaxies within the potential wells of halos (see e.g. White & Frenk 1991; Somerville & Primack 1999; Cole et al. 2000). In very recent SAMs, an important feedback source for galaxy formation is provided by active galactic nuclei (AGN), which are those nuclei in galaxies that emit radiation powered by accretion onto a supermassive black-hole. Such an emission can re-heat the gas and suppress the cooling flows, thus causing the quenching of star formation in massive galaxies (Croton et al., 2006; Bower et al., 2006).

1.3 Characterizing the galaxy population: Large-scale galaxy redshift surveys

Since the beginning of the twentieth century, much effort has been devoted to mapping the contents of the Universe using both photometry and spectroscopy and the redshift as a distance estimator. It is worth noticing the pioneering works of Humason (1956) and Sandage (1978), who measured redshifts of bright galaxies from the Shapley-Ames photometric catalog (Shapley & Ames, 1932). The CfA Redshift Survey, as mentioned before, is considered the first proper redshift survey, specifically designed for the study of the 3-D distribution of galaxies. Importantly, the CfA RS took advantage of the improvements in the spectroscopic techniques and the observation of the 21-cm line of hydrogen, which marked the birth of spectral line radio astronomy. The first version of the catalogs contained a sample of about $2,500$ galaxies brighter than $M_B = 14.5$ extracted from the Zwicky photometric catalogs (see Zwicky & Humason 1961 and following papers). The CfA RS provided the community with a strong observational evidence for the existence of large scale structure in 3-D space in the Universe (Davis et al., 1985), an idea that had been envisaged long ago (mainly by analyzing the distribution of nabulae on the celestial sphere) by Shapley, Hubble, Zwicky, Gerárd de Vaucouleurs and others. Ever since the CfA RS we are positive that galaxies in the Universe do not distribute randomly, but accumulate in thin filaments which are separated by under-densed regions or voids (see Figure 1.1). Such a distribution is an observational evidence for the CDM paradigm.

This first vision of cosmic complexity encouraged the development of new imaging and spectrometric technology and, consequently, gave rise to a number of other redshift surveys that followed different approaches and strategies. In the last 3 decades, astronomers have collected progressively larger and deeper galaxy samples, not only to map the distribution of galaxies but also to investigate their formation and evolution and the expansion of the Universe. A good example is the Southern Sky Redshift Survey (SSRS, da Costa et al. 1988), which helped characterize the galaxy population in the late 1980's and early 1990's. The SSRS had to overcome the problem that there was no equivalent to the Zwicky catalogs in the southern hemisphere. Even so, it managed to obtain redshifts of about $5,400$ galaxies with $m_B \leq 15.5$ (in the last version; da Costa et al. 1994). Also worth mentioning are the Perseus-Pisces sample (Giovanelli & Haynes, 1991), where redshifts were estimated using the 21-cm line ($5,183$ galaxies) or the catalogs based on data from the Infrared Astronomical Satellite (IRAS; see Rowan-Robinson 1996), which commenced the systematic exploration of the infrared universe. A huge impulse to the survey field came with the multi-fiber spectrographs in the mid 1990's, which can measure the spectra of hundreds of objects simultaneously (before, spectra were taken object-to-object using slits). The advent of this technology caused an increase in galaxy statistics of a factor ~ 100

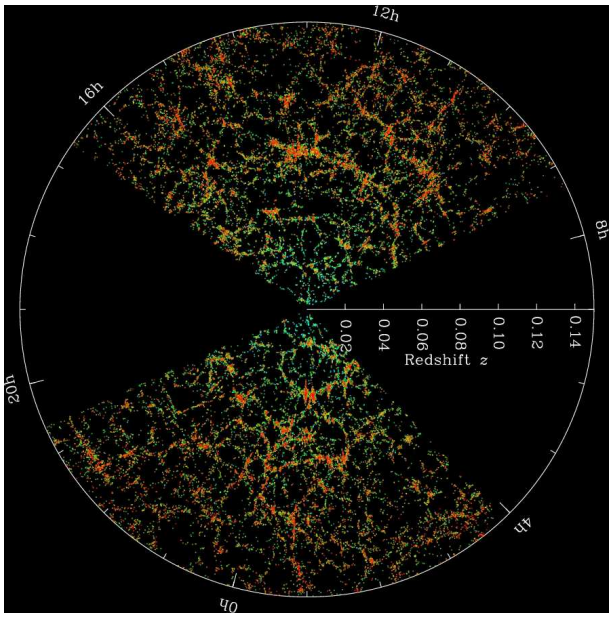

Figure 1.5 The SDSS is the largest galaxy (and stellar) survey ever compiled. The Eighth Data Realease contains extensive photometric and spectroscopic information for more than $1,000,000$ galaxies and quasars, which spread over 9200 deg^2 on the sky. We show here a slice through the SDSS 3-dimensional map of the distribution of galaxies up to $z \sim 0.15$. Galaxies are colored according to the ages of their stars, with the redder, more strongly clustered points showing galaxies with an older stellar population.

and also allowed for deeper studies. One of the first large-scale galaxy redshift surveys using multi-fiber spectrometric technology was the Las Campanas Redshift Survey (LCRS, Shectman et al. 1996), containing more than $25,000$ galaxies at an average redshift of $z \sim 0.1$. Around $250,000$ redshifts were collected by the the 2 degree Field Galaxy Redshift Survey (2dFGRS, Colless et al. 2001), which mapped an area of 1500 deg^2 on the sky. Finally, the Sloan Digital Sky Survey (SDSS, York et al. 2000) is the largest photometric and spectroscopic survey ever compiled, and provides an accurate map of the nearby universe at $z \lesssim 0.3$. The final version of the catalogs, the SDSS Eighth Data Release (Aihara et al., 2011), contains spectroscopic information for more than $1,000,000$ galaxies and quasars which spread over 9200 deg^2 on the sky (see Figure 1.5).

The statistical study of the galaxy population has not been restricted to the low-redshift universe ($z \lesssim 0.3$) or the relatively bright galaxies ($r \lesssim 20$). Recently, several deep galaxy surveys have been undertaken in order to explore the nature of the fainter and more distant objects in the Universe. Important observational constraints have been provided by the DEEP2 Galaxy Redshift Survey (Davis et al., 2003), the VIMOS-VLT Deep Survey (VVDS, Le Fevre et al. 2003), the Munich Near-Infrared Cluster Survey (MUNICS, Drory et al. 2001), the Cosmic Evolution Survey (COSMOS, Scoville et al. 2007) or the COMBO-17 Survey (Meisenheimer & Wolf, 2002). New constraints are expected from current surveys like the ALHAMBRA-Survey (Moles et al., 2008). These surveys can only reach down to magnitudes $m \gtrsim 22$ at the cost of reducing remarkably the mapped area ($\lesssim 5$ deg^2) as compared to large-scale low-z surveys like the SDSS or de 2dFGRS. The expensive requirements in terms of telescope time are, however, alleviated in some of them by using photometry as a rough proxy of the

spectral energy distribution (SED) of galaxies. Photometry can also be used for obtaining *photometric redshifts*. These type of galaxy surveys, which are dubbed *photometric* or *photo-z* surveys, lack the SED resolution provided by *spectroscopic* surveys (such as the SDSS, the 2dFGRS or the DEEP2 GRS), but can get considerably deeper, allowing for a statistical study of the fainter galaxies in the Universe. Examples of photo-z surveys mentioned above are the COMBO-17, COSMOS and the recent ALHAMBRA-Survey. ALHAMBRA, in particular, will provide photometry in 20 optical filters and the 3 near-infrarred JHK bands for $\sim 500,000$ galaxies at $z \lesssim 1$.

In the next decades, astronomers will be challenged to improve constraints on galaxy formation and evolution, the acceleration of the Universe and the nature of dark matter and dark energy. To this end, a number of large-scale photometric surveys are underway or being designed at the time of writing this dissertation. The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey (CANDELS, http://candels.ucolick.org/About.html) will use 902 assigned orbits of observing time with the Hubble Space Telescope (between 2011 and 2013) to provide imaging of the very distance Universe. The Javalambre - Physics of the Accelerating Universe Astrophysical Survey (J-PAS, http://www.iaa.es/∼benitez/jpas/survey.html) will use as many as 42 narrow-band filters a la ALHAMBRA to map a portion of 8000 deg^2 on the sky. Two other large-scale photometric surveys are expected to revolutionize the study of the Universe in the next decade: the Large Synoptic Survey Telescope (LSST, http://www.lsst.org/lsst) and the Panoramic Survey Telescope & Rapid Response System (Pan-STARRS, http://pan-starrs.ifa.hawaii.edu/public/). These imaging facilities will provide a massive statistical database ($\sim 10^{10}$ galaxies, from the Local Group to $z > 6$) of high-quality multi-band photometry.

There is a growing awareness in the community, however, that answering the most pressing questions in cosmology and in the field of galaxy formation and evolution also requires the accurate measurements of redshift and other fundamental galaxy properties that only spectroscopic surveys can provide (see a discussion supporting this claim in Bell et al. 2009). This approach is followed by the large and promising Baryon Oscillation Spectroscopic Survey (BOSS, Schlegel et al. 2009a). BOSS, whose scientific exploitation has already commenced, is a dark energy experiment designed to map 1.5 million galaxies up to $z \sim 0.75$ and $\sim 150,000$ quasars at $z = 2.5 - 3$, over a region of $10,000$ deg^2 on the sky. Its successor, the Big Baryon Oscillation Spectroscopic Survey (BigBOSS, Schlegel et al. 2009b), which is still in the design phase, would extend this coverage to $14,000$ deg^2, to map some 40 million galaxies. Spectroscopic facilities such as BigBOSS or the galaxy surveys conceived for the Large Sky Area Multi-Object Fibre Spectroscopic Telescope (LAMOST, Wang et al. 2009) are expected to lead the way in the next decade. To this end, much effort is being devoted to optimizing spectroscopic survey techniques and developing next-generation spectrometric technology (see Azzaro et al. 2010 and Morales et al. 2011).

1.4 Statistical properties of the galaxy population

The progress in the survey field in the last decades have made it necessary to develop data reduction pipelines and analysis tools to process and understand increasingly larger data sets. The motivation is to characterize in detail the distribution and main properties of the galaxy population in three-

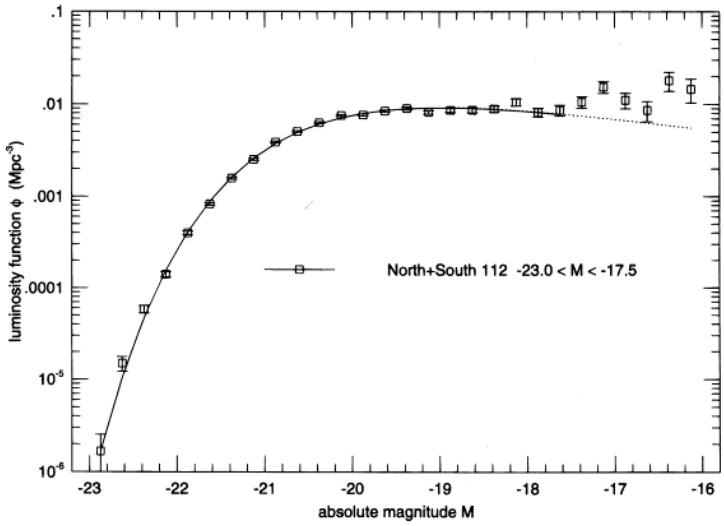

Figure 1.6 The luminosity function is one of the most fundamental statistical properties of the galaxy population. Here, we show the luminosity function of Lin et al. (1996), measured with the Las Campanas Redshift Survey. A Schechter fit to the data is shown in a dotted line.

dimensional space, providing observational constraints to both cosmological and galaxy formation models. The very first statistic to measure in a galaxy survey is the distribution of fluxes (apparent magnitudes) of galaxies, commonly dubbed *galaxy number counts* (GNC), which are expressed per unit magnitude and area on the sky. The general consensus today is that for relatively bright galaxies ($m_B \lesssim 21$) GNC are roughly consistent with the so-called *Euclidean model*, which assumes that the luminosity of galaxies remain constant in time and that the number of galaxies per unit volume is conserved after they are formed (see Montero-Dorta & Prada 2009 or Yasuda et al. 2001 for the SDSS GNC and Norberg et al. 2002 for the 2dFGRS GNC). For fainter objects ($m_B \gtrsim 22 - 23$), a non-evolving model is clearly not representative of the real Universe (see Campos 1996 for a good review on this subject).

A much better description of the average distribution of galaxy light in the Universe is provided by the luminosity function of galaxies (LF), which is the number density of galaxies per unit absolute magnitude. In Figure 1.6, we show an estimate of the LF in the r band in a sample of the Las Campanas Redshift Survey (19,000 galaxies at $< z > \sim 0.1$), taken from Lin et al. (1996). The shape of the LF, which typically obeys the parametric Schechter function (Schechter 1976; see an example of a Schechter fit represented by a dotted line in Figure 1.6), contains many valuable clues to the process of galaxy formation. The most prominent characteristic of the LF is the sharp cut-off for bright galaxies at $M_r \sim -21$, which has been associated with several feedback processes shutting down star formation in high-mass galaxies. At the faint end, observations indicate a relatively flat slope of the LF, which is in contrast to the very steep slope of the cosmic mass function. Uncertainties remain regarding the exact the exact value of this slope, though. Montero-Dorta & Prada (2009) or Norberg et al. (2002) advocate for a relatively steep slope (with the Schechter parameter $\alpha \sim -1.20$), whereas Blanton et al. (2003b) obtain a completely flat slope ($\alpha \sim -1$). Deviations from the Schechter

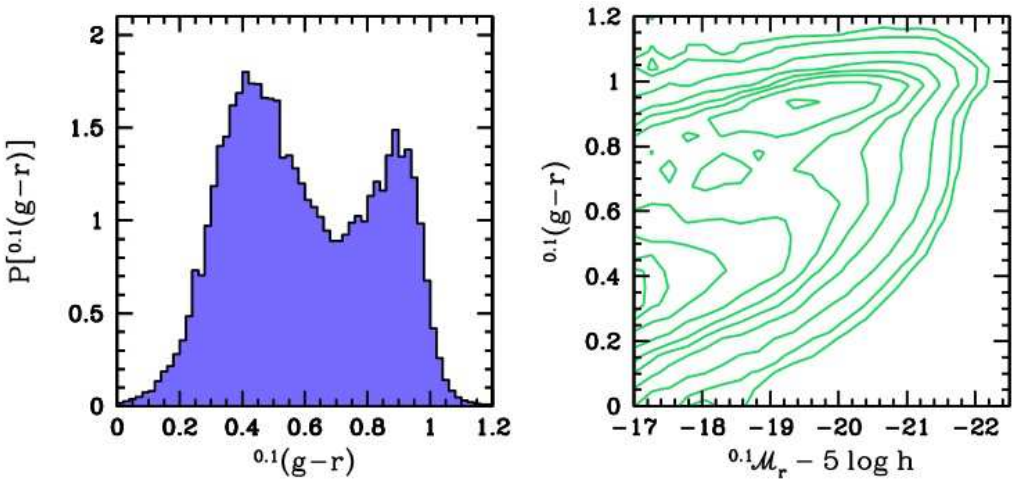

Figure 1.7 The distribution of galaxy colors in the low-z Universe is remarkably bimodal, with a prominent *red sequence* and a much more extended *blue cloud*. In the left-hand plot, we show the probability density of the $^{0.1}(g-r)$ color (where the superindex indicates that the color is K-corrected to z=0.1) in a sample comprised by 365,000 SDSS galaxies. On the right-hand side, a color-magnitude diagram, $^{0.1}(g-r) vs. ^{0.1}M_r$, for the same sample is provided. This figure was taken from Mo et al. (2010).

function have been found in both the field LF in blue bands (i.e. *bright-end bump*, Montero-Dorta & Prada 2009) and in the clusters LF (a faint-end upturn found by several groups, e.g. Popesso et al. 2005; Mercurio et al. 2006).

A number of other statistical properties of the galaxy population have been studied in recent years. Among the galaxy properties, *color* is one of the most important ones, as it contains important information about the *star formation history* of the galaxy. Massive stars, which are in general shorter-lived, emit a larger fraction of their total light at short wavelengths than lower-mass stars. This implies that bluer galaxies have younger stellar populations than redder galaxies. The color of a galaxy, however, depends on other properties of the stellar population, such as the metallicity (chemical composition) or the dust content. Ever since the work of Strateva et al. (2001), who used the SDSS, we know that the distribution of galaxy colors in the local Universe is highly bimodal, with a relatively narrow peak at the red end, i.e. the *red sequence*, and a much wider distribution of blue galaxies, i.e. the *blue cloud*. These features are associated (roughly) with two different galaxy morphologies: old early-type galaxies and younger late-type galaxies. In Figure 1.7, we illustrate *color bimodality* by showing the probability density of the $(g-r)$ color (left) and a color-magnitude diagram, $(g-r)$ vs. M_r (right), in a sample comprised by 365,000 SDSS galaxies. This figure was taken from Mo et al. (2010).

Galaxy number counts, luminosity functions and color distributions are the first-order statistics that characterize the galaxy population, as they are all basically derived from the apparent magnitude and the redshift of individual galaxies. In the last decades, much effort have been devoted to describing the global behavior of the galaxy population as far as other more fundamental galaxy properties are concerned. Progress in this sense has been possible thanks to the development of the *stellar population*

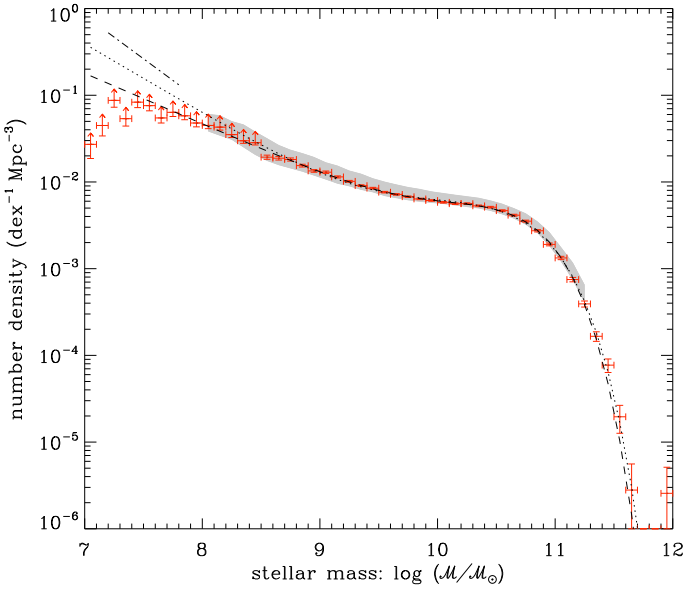

Figure 1.8 The observed stellar mass function at low redshift presents a remarkable upturn at the low-mass end, which implies that a single Schechter function is not capable of fitting the data. In this plot, the estimate of Baldry et al. (2008) for the local Universe ($z < 0.05$), which was obtained with a sample of 50,000 SDSS galaxies taken from the NYU VAGC, is shown in red symbols (with Poisson error bars). The different lines represent double-Schechter functions with varying faint-end slopes fitting the data. The shaded region shows the variation of the stellar mass function for different stellar mass estimates and redshift ranges. For more information see Baldry et al. (2008).

synthesis, which aims to interpret the integrated light that we observe from galaxies. Among these galaxy properties, stellar mass stands up as the most fundamental one. The distribution of stellar mass in the Universe is described by the stellar mass function (SMF), which is the number density of galaxies per unit stellar mass bin. It is rather accepted today that the shape of the stellar mass function differs from that of the luminosity function, even though stellar mass and luminosity are intrinsically connected. In fact, a remarkable faint-end upturn at $M_* \simeq 10^{10}\ M_\odot$ has been found up to, at least, $z \sim 0.5$, making it impossible to fit the stellar mass function with a single Schechter function (e.g. Baldry et al. 2008; Drory et al. 2009; Pozzetti et al. 2010). This is illustrated in Figure 1.8, where we show the stellar mass function of Baldry et al. (2008), obtained with a sample of 50,000 SDSS galaxies with $z < 0.05$ taken from the NYU VAGC (Blanton et al., 2005c). Note the importance of the shape of the SMF, where the *mass assembly history* of the Universe is imprinted.

Using stellar population synthesis techniques, astronomers have investigated a number of correlations between galaxy properties in order to disentangle the *galaxy formation and evolution picture*. A good example of this is the so-called *mass-metallicity relation*, which was first envisaged in Lequeux et al. (1979) , where luminosity was used as a surrogate for stellar mass. While stellar mass reflects the amount of gas locked up into stars, metallicity, another fundamental galaxy property, reflects the amount of gas reprocessed by stars and the exchange of gas between the galaxy and its environment. The mass-metallicity relation establishes that the larger the stellar mass content, the higher the gas-phase metallicity in star-forming late-type galaxies (see Tremonti et al. 2004).

The characterization of the galaxy population would be severely incomplete if the effect of environment were not taken into account. The fact that the distribution of different types of galaxies in the Universe depends strongly on the density of the surrounding environment was realized as early as the 1930's. Hubble & Humason (1931) found that denser environments (like clusters) were typically inhabited by early-type galaxies. This *morphology-density relation* was quantified more accurately by Dressler (1980), who analyzed the distribution of galaxy morphologies as a function of density in a sample of 55 clusters. From these and other subsequent works we know that spiral galaxies tend to inhabit low-density environments while elliptical galaxies show a strong preference for dense regions. The influence of environment is, of course, not only noticeable when galaxies are separated by morphology. Tightly connected with the morphology-density relation is the *color-density relation*, which is a fundamental aspect of color bimodality. By studying the (u-r) color distribution of 25,000 galaxies in the SDSS, Balogh et al. (2004) found that red galaxies tended to live in denser regions, in contrast to blue galaxies which populated any environment independently of its density.

The development of the large-scale galaxy survey field and the stellar population synthesis techniques have paved the way for more detailed studies on the environmental dependence of other fundamental galaxy properties. Today we know that the average galaxy that populates a dense environment is not only red and late-type, but also typically more massive, more concentrated, less gas-rich and has lower specific star formation rate (star formation per unit stellar mass) than the average galaxy inhabiting a less-dense environment. Figure 1.9, which was taken from Kauffmann et al. (2004), illustrate the environmental dependences of several galaxy properties in the local Universe, obtained from a sample of 122,000 SDSS DR1 galaxies (see also e.g. Hogg et al. 2003; Blanton et al. 2005a; Baldry et al. 2006; Weinmann et al. 2006). Much effort has been devoted to study the AGN - density relation

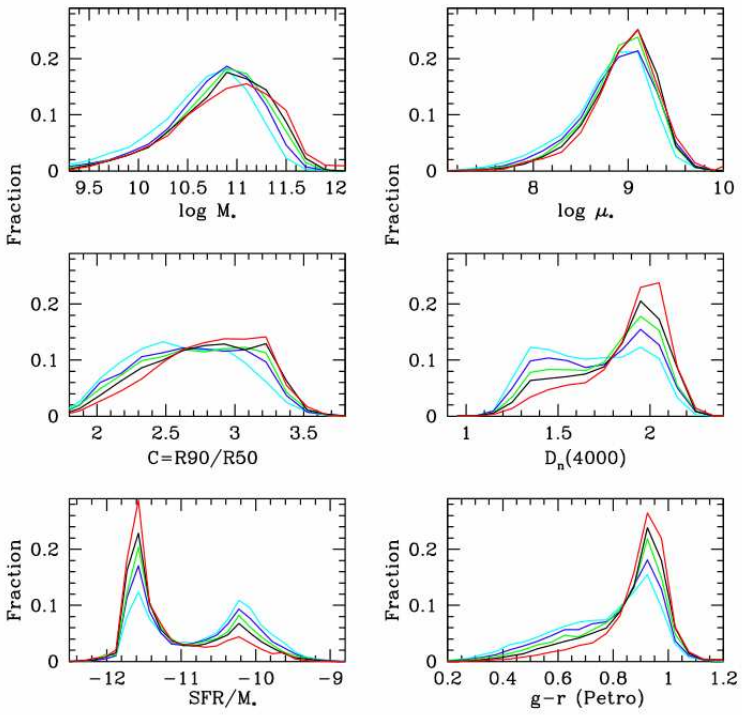

Figure 1.9 The fraction of galaxies with different physical properties depends on the density of the environment that they inhabit. In this plot, which was taken from Kauffmann et al. (2004), the fraction of the total stellar mass in the local universe contained in galaxies is shown as a function of the stellar mass ($\log M_*$), the mean stellar age (for which the λ4000 break strength, Dn(4000), is used as a surrogate), the specific star formation rate (SFR/M_*), the concentration index (R90/R50), the specific stellar mass surface density ($\log \mu_*$) and the (g - r) colour (K-corrected to $z = 0.1$). The different colour lines represent the different density bins as follows: cyan, 0 or 1 neighbour; blue, 2 - 3 neighbours; green, 4 - 6 neighbours; black, 7 - 11 neighbours; red, more than 12 neighbours. Kauffmann et al. (2004) used a sample of 122,000 galaxies taken from the SDSS DR1.

as well, motivated by the importance that nuclear activity seems to have in shaping the evolution of galaxies (see e.g. Miller et al. 2003; Kauffmann et al. 2004; Sorrentino et al. 2006; Georgakakis et al. 2007; Montero-Dorta et al. 2009). Note that the main difficulty here resides in the fact that each of these galaxy properties correlates with each other at fixed density in a rather complicated way. Much effort, both theoretical and observational, is still needed in order to disentangle this scenario and *identify the main processes that drive the formation and evolution of galaxies*. In particular, we are still yet to discover if is the interaction with the environment what shapes the evolution of galaxies (the nurture hypothesis) or if evolution is determined by the initial conditions at their birth (the nature hypothesis). It might also be a combination of both. For more information on the statistical properties of the galaxy distribution see Martínez & Saar (2002).

1.5 Challenges in the field of galaxy formation and evolution

The standard model of cosmology provides a satisfactory explanation for the hierarchical formation of large-scale structure in the Universe, even though the nature of dark matter and dark energy remains undisclosed. Moreover, the distribution and properties of dark matter halos have been relatively well characterized by N-body numerical simulations. Theorists will be challenged, however, to tune the Λ-CDM framework in order to accommodate several observational inconsistencies, basically involving small-scale structures. Examples of such problems are the observed emptiness of voids (Tikhonov & Klypin, 2009), the *missing satellites problem* (Klypin et al., 1999), which may be alleviated by the discovery of ultra-faint dwarfs in the Milky Way (see Kravtsov 2010), or the observationally inferred dark matter profiles of galaxy and cluster halos (see e.g. Prada et al. 2003; Broadhurst et al. 2008; Gentile et al. 2007). Current and future experiments such as BOSS, LSST, DES, PanSTARRS or BigBOSS are expected to shed light into the field, by exploring the nature of dark matter and dark energy and the perhaps inflationary processes that lead to the formation of large-scale structure in the Universe.

The establishment of a coherent cosmological framework has paved the way for the study of the formation and evolution of galaxies. In fact, thanks to the seminal work of White & Rees (1978) and to subsequent studies, we have a general theory for the formation of galaxies within dark matter haloes. Unfortunately, the processes that regulate the cooling and condensation of baryons and the star formation at the center of dark matter halos are extremely complex. The further development of numerical simulations that solve the hydro-dynamical and gravitational forces between gas and dark matter is crucial to the understanding of the physics of galaxy formation and evolution.

As mentioned before, galaxies can be essentially of two kinds: disk galaxies (with the particular case of spiral galaxies) and elliptical galaxies. The consensus today, motivated by their somehow well-ordered shape, is that the evolution of disk galaxies is not determined by any major events (such as mergers), but by other *secular* or internal processes. The formation of these types of systems from the gravitational collapse of a rotating protogalactic cloud of gas within the gravitational potential well of a dark matter halo can be relatively well modeled by numerical simulations. These simulations must assume a halo density profile (which is relatively well characterized) and an initial angular momentum

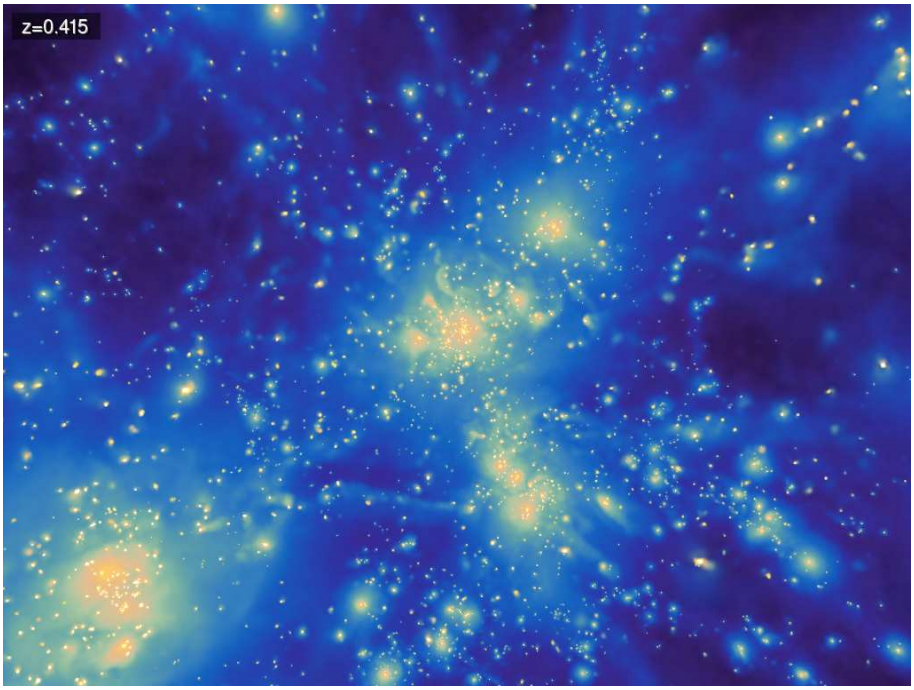

Figure 1.10 Numerical simulations have established as a fundamental tool for investigating the formation and evolution of galaxies. In this figure, we show a snapshot at $z = 0.45$ of a simulation of the evolution of the gas density in a galaxy cluster, performed within the MultiDark project. More information at http://www.multidark.org/MultiDark/.

for the in-falling gas, that must be conserved during the process. Understanding how the conservation of the angular momentum takes places is one of the most important challenges of galaxy formation. Find a good review on this topic in Mayer et al. (2008).

As for elliptical galaxies, numerical simulations have demonstrated that their chaotic shape can result from several major and minor mergers with other galaxies (this hypothesis was first formulated by Toomre 1977). These mergers are known to be capable of inducing star formation, which would take place in several phases. Such a scenario is easy to accommodate within the CDM hierarchical framework and have clear advantages with respect to the monolithic collapse scenario for ellipticals, where galaxy formation (and star formation) happens violently on a short time scale and evolution of the stellar population proceeds passively (see Larson 1975 and subsequent works). However, the merger scenario is far from being observationally consolidated and we still need to address how much of the evolution in elliptical galaxies is driven by mergers and how much by other secular processes. In this sense, recent works have tried to measure the evolution in the rate of merger events from $z \sim 1$ (see an example in Robaina et al. 2009), even in the absence of adequate high-redshift data.

Feedback processes regulating star formation are important ingredients in the galaxy formation and evolution picture. In order to understand the nature of the sharp cut-off at the bright-end of the observed luminosity function or its relatively shallow faint-end slope as compared to the halo mass function, advances in the understanding of AGN and supernova feedback or the effect of stellar wind-driven outflows are required. Semi-analytic models have proved to be a necessary tool for exploring

these processes.

Numerical simulations, semi-analytic models and other theoretical efforts must be complemented with improved observational constraints at high redshift in order to unveil the assembly of galaxies and their evolution. The idea is to link different galaxy populations at different redshifts and this requires a better statistical characterization of the Universe at $z \sim 1$, giving continuity to valuable efforts from spectroscopic surveys like the VVDS or DEEP2 and photo-z surveys like COSMOS or COMBO-17. More accurate measurements of even the most fundamental statistical properties such as the luminosity or the stellar mass function at $z \sim 1$ are needed. Aimed to provide this update is the current ALHAMBRA-Survey. Note, however, that typical L_* galaxies are believed to have assembled at $z = 2 - 3$, where star formation and black hole accretion activity peaked. Exploring the Universe at these redshifts has been possible thanks to photometric Lyman-break techniques (mainly due to Steidel and collaborators, see Steidel et al. 1999 as an example) and to deep surveys from the *Great Observatories*: HST, Chandra and Spitzer. These studies have been able to identify a number of galaxy populations with special features at different wavelength ranges. Characterizing these high-redshift galaxies and establishing connections with low-redshift galaxy populations is an ambitious and fundamental challenge for modern Astronomy. Next-generation surveys like the LSST, Pan-STARRS, LAMOST or BigBOSS will surely contribute to this goal. However, the prospection of new spectroscopic survey facilities, the development of spectrometric technology and the optimization of survey techniques are, as mentioned before, key to advance in this field. In this sense, the systematic study of the Universe at the critical redshift range $2 \lesssim z \lesssim 3$ will only be possible with large-scale spectroscopic surveys in the near-infrarred (Bell et al. 2009).

In this dissertation, I provide state-of-the-art observational constraints to several important aspects of the galaxy formation and evolution process, using large-scale galaxy redshift surveys. For the sake of clarity, this dissertation have been divided into 3 parts. In the first part, entitled *The luminosity function of galaxies in the nearby Universe*, I present the most accurate measurement of the luminosity function of galaxies at low redshift, obtained using the SDSS DR6 in all photometric ugriz bands. In addition, the observed luminosity function is compared with the predictions from semi-analytic models of galaxy formation and evolution. The second part, which is entitled *Galaxy evolution from* $z \sim 1$, focuses on two different galaxy evolution projects. Firstly, I analyze the main statistical properties of the low-redshift galaxy population and discuss a method for obtaining stellar masses in the ALHAMBRA photometric survey, using a preliminary sample of the data. This is intended to lay the foundations for future galaxy evolution studies at $z \sim 1$ with ALHAMBRA. Secondly, I measure de evolution of the AGN-density relation between $z \sim 1$ and $z \sim 0$ using the DEEP2 GRS and the SDSS DR4, respectively. In the third part, which is called *Future spectroscopic surveys*, I discuss survey optimizations for next-generation wide-field fiber-fed spectrographs. In particular, I present an optimized algorithm for fiber positioning that maximizes the number of target galaxies observed in the first spectrograph exposures. Finally, in the last part of this dissertation, I summarize the main conclusions of this work and outline the future work.

Part I

The luminosity function of galaxies in the nearby Universe

2

The SDSS DR6 Luminosity Functions of Galaxies

2.1 Introduction

The galaxy survey field has experienced a huge development in the last decades, thanks to the emergence of multi-fiber spectrographs, which can simultaneously measure hundreds of redshifts. Progressively larger and deeper galaxy samples have been obtained, thus allowing the statistical characterization of the galaxy population, especially at low redshift. In particular, the Sloan Digital Sky Survey is the largest photometric and spectroscopic survey ever compiled, providing an accurate map of the nearby universe, at $z \lesssim 0.3$. The SDSS Sixth Data Release (Adelman-McCarthy et al., 2008), that we use in this chapter, contains spectroscopic information for more than 1,000,000 galaxies and quasars which spread over 7425 deg^2 on the sky.

The luminosity function of galaxies (LF), i.e. the distribution of the number density of galaxies per unit absolute magnitude, is one of the most fundamental statistical properties of the galaxy population. The shape of the LF, which typically obeys the parametric Schechter function (Schechter 1976), contains many valuable clues to the process of galaxy formation. At the faint end ($M_r \gtrsim -21$), the logarithm of the LF follows a decreasing power law towards brighter absolute magnitudes. The slope of this power law is described by the Schechter parameter α, so that $\alpha = -1$ corresponds to a flat slope (the more negative this parameter the steeper the faint-end slope). At the bright end, this behavior breaks sharply, and the number density of galaxies decreases exponentially. This bright-end cut-off has been associated with several feedback processes shutting down star formation in high-mass galaxies. As an example, one of the first mechanisms invoked to explain the truncation of star formation in giant elliptical galaxies is the effect of supernova explosions in the inter-stellar medium. The energy released in these events can reheat to disruption the cold gas clouds or trigger strong winds that might eventually sweep away the cold gas supply of the galaxy. More recently, the outstanding discovery that most nearby luminous galaxies harbor a super-massive black hole (SMBH) with mass tightly correlated with the mass of the bulge has given rise to new theories. Namely, it has been suggested that feedback from an accreting SMBH might blow the gas content out of a galaxy, consequently shutting

down star formation and preventing any further growth (see Benson et al. 2003; Croton et al. 2006; Bower et al. 2006).

An accurate measurement of the luminosity function of galaxies at low redshift is crucial to constraining theories of galaxy formation and evolution. For this reason, the LF has been measured with a number of large-scale galaxy redshift surveys. With the LCRS, Lin et al. (1996) found that the LF that could be fitted by a Schechter function with $\alpha = -0.7$ down to $M_r \sim -17.5$. At fainter magnitudes, the authors discovered a considerable excess as compared to the Schechter fit, which was associated with a population of blue dwarf galaxies. These results are in good agreement with the previous findings of Marzke et al. (1994), who used a sample of about $9,000$ galaxies from the CfA RS (with $m_z \leq 15.5$). Loveday (1997), with the Stromlo-APM survey, obtained a completely flat faint-end slope ($\alpha = -1$) and similar deviations in the very faint end. The first estimations of the LF using the SDSS (in particular, the SDSS Commissioning Data) are presented in Blanton et al. (2001), who used a sample of $\sim 11,000$ galaxies brighter than $r = 17.6$ on a stripe of 140 deg^2. In this sample, the r-band LF is well fitted by a Schechter function with a steep faint-end slope ($\alpha = -1.20$) all the way down to $M_r \sim -16$ (also steep faint-end slopes are found in the u, g, i and z bands). Discrepancies with previous works are claimed to be caused by different photometry (the SDSS uses mainly Petrosian magnitudes, which measure a constant fraction of the galaxy's total light, regardless of the amplitude of its surface brightness profile) and by the fact that earlier works used remarkably shallower samples. The LF of Blanton et al. (2001) is in excellent agreement with the slightly subsequent paper of Norberg et al. (2002). In this work, a sample of more than $100,000$ galaxies taken from the 2dFGRS is used to estimate the b_J-band (similar to the SDSS g band) LF in the absolute magnitude interval $-16.5 > M_{bJ} - 5\log_{10} h > -22$. In this paper, the authors introduce a detailed treatment of survey incompleteness and a correction to account for the change in the luminosity of galaxies due to evolution. Observationally, the faint-end slope of the LF is still today a rather controversial issue. Note that providing an accurate observational constraint at this magnitude range is fundamental for theories trying to model the feedback processes regulating the formation of dwarf galaxies (see Benson et al. 2003 for a discussion on this issue). In Blanton et al. (2003b), the SDSS LF in all SDSS bands is re-calculated in a slightly larger sample than that of Norberg et al. (2002), corresponding to the SDSS DR2 ($\sim 150,000$ objects). The evolution correction is claimed to be responsible for the flattening of the LF, which is now well fitted (in the r-band) by a Schechter function with $\alpha \backsimeq -1$.

In this chapter, which is based on Montero-Dorta & Prada (2009), we take advantage of the large increase in galaxy statistics provided by the SDSS DR6 to calculate the luminosity functions of galaxies in the optical SDSS bands in the nearby Universe. In addition, we present luminosity densities derived from our LFs and galaxy number counts for the SDSS DR6 spectroscopic catalog. We intend to shed light both at the faint end of the LF, where most discrepancies come from, and at the bright-end, where statistics have always been poor and errors, consequently large. In section 2.2 we briefly describe the SDSS DR6, discuss our sample selection and comment on redshift completeness. In section 2.3 we present our results on the number counts, the luminosity funstions and the luminosity densities of galaxies in each one of the SDSS photometric bands. Finally, in Section 2.4 we discuss our results and in Section 2.5 we present a summary of our work. Throughout this chapter, unless otherwise

stated, we assume a standard ΛCDM concordance cosmology, with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $h = 1$. In addition, we use AB magnitudes.

2.2 The Sloan Digital Sky Survey Data Release 6. Data samples selection and redshift completeness

In this chapter we use the SDSS Sixth Data Release (Adelman-McCarthy et al., 2008) to measure important statistical properties of the low-z galaxy population. This data set almost completes the North Galactic Cap, containing photometric information for $\sim$ 290 million objects over 9583 deg^2. Around 1.27 million objects were selected for spectroscopy, covering an area of 7425 deg^2 on the sky. Important for this work, spectroscopy is available for $\sim 900,000$ galaxies down to magnitude $r \sim 17.77$ (York et al., 2000; Stoughton et al., 2002). Detailed information about the SDSS DR6 can be found in Adelman-McCarthy et al. (2008).

The SDSS DR6 was the largest spectroscopic survey of the nearby universe publicly available at the time of creating this work. Even today, only the last release of the SDSS, the SDSS DR8, is lightly larger, although not enough to alter our results remarkably. The SDSS collaboration have successively extended their catalogs since the times of the SDSS Early Data Release, improving enormously our capability of mapping the universe up to redshift $z \sim 0.3$. A few years ago, Blanton et al. (2003b) used the SDSS DR2 to estimate the luminosity function of galaxies. In this work, and thanks to the SDSS DR6, the size of our samples is a factor ~ 7 larger in the very blue u band and a factor between ~ 3 and ~ 5 in the other SDSS bands (g, r, i and z), as compared to Blanton et al. (2003b). This huge enhancement in the galaxy statistics will be especially useful in the bright end of the LF, where the number density is remarkably lower. In addition, we expect to reach deeper magnitudes in the faint end with respect to previous works. Galaxy number counts can also be estimated with significantly more accuracy. It is therefore well justified to update the current knowledge on the number counts, luminosity functions and luminosity densities of galaxies in the closeby Universe.

2.2.1 Data samples selection

We have drawn our samples from the NYU Value Added Galaxy Catalog DR6 (Blanton et al., 2005c). The NYU-VAGC is a compilation of galaxy catalogs cross-matched to the SDSS data products that includes a number of useful quantities derived from the photometric and the spectroscopic catalogs (such as K-corrections or absolute magnitudes). It also incorporates a precise and user-friendly description of the geometry of the survey.

Our Parent Sample (PS), which is the galaxy catalog from which all galaxy samples used in this chapter are extracted, is built from the Large Scale Structure sample (LSS) of the NYU-VAGC and contains all galaxies that satisfy the criteria of the SDSS Main Galaxy Sample (Strauss et al., 2002) - except for those which are close to a saturated bright star. With these first restrictions the PS is comprised by 947,053 galaxies that spread over 7280.35 deg^2 on the sky. From this catalog, each sample is drawn by applying the following cuts to the redshift and the apparent magnitude:

- $m_{min}(\mathrm{j}) < \mathrm{m(j)} < m_{max}(\mathrm{j})$

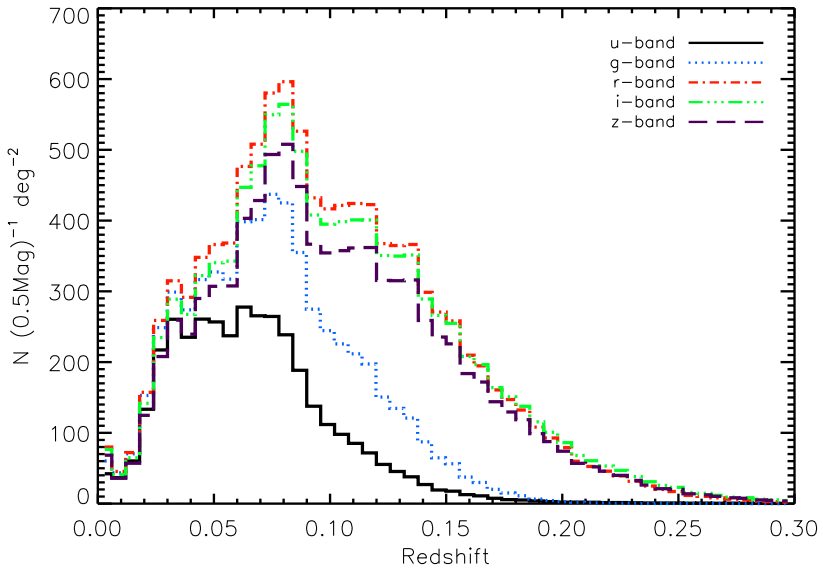

Figure 2.1 Redshift distributions in all SDSS photometric bands.

Band	Number	m_{min}	m_{max}	z_{min}	z_{max}	slope
u	153,411	16.45	18.81	0.020	0.170	0.615 ± 0.013
g	247,463	14.55	17.91	0.020	0.160	0.586 ± 0.009
r	437,565	13.93	17.6	0.020	0.220	0.591 ± 0.006
i	423,369	13.55	17.20	0.020	0.235	0.597 ± 0.007
z	378,833	13.40	16.87	0.020	0.24	0.604 ± 0.007

Table 2.1 Number of galaxies and limits in apparent magnitude and redshift in each SDSS sample. Motivation for each cut is discussed in section 2.2.1. The slope in units of mag^{-1} of galaxy number counts (see Section 2.3.1) within the apparent magnitude and redshift limits of each sample is also provided. Within these ranges, number counts are consistent with an Euclidean, non-evolving Universe.

- $z_{min}(\mathrm{j}) < \mathrm{z(j)} < z_{max}(\mathrm{j})$

where j=u, g, r, i, z. We use extinction-corrected Petrosian magnitudes and heliocentric redshifts. In Table 2.1, we show lower and upper limits of these quantities along with the number of galaxies for each sample.

At this point, it is necessary to clarify the motivation for each cut. The apparent magnitude limits of Table 2.1 are set to ensure that the effect of redshift incompleteness is small within our SDSS galaxy samples. In a spectroscopic survey and strictly speaking, redshift completeness is defined as the fraction of galaxies with a reliable redshift estimation of all galaxies lying within the survey covered area and brighter than the survey apparent magnitude limit. In practice, it is common to consider completeness relative to the target galaxy population, i.e. to those galaxies that have been selected for spectroscopy from the photometric catalog. Unless otherwise stated, this is the definition we will use throughout this chapter. In the SDSS, brightness-dependent redshift incompleteness starts to be

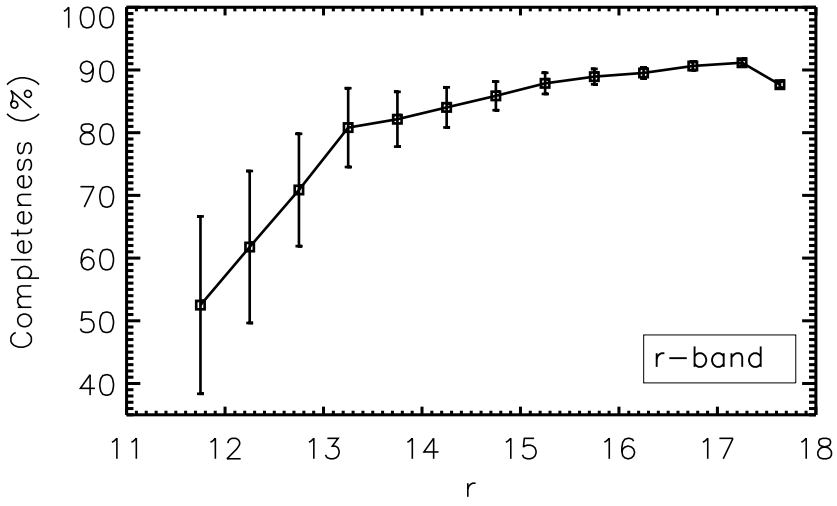

Figure 2.2 Redshift completeness vs. apparent magnitude in the r band. At $r \lesssim 14$ completeness decreases sharply. Errors have been estimated by propagating the poissonian uncertainties to redshift completeness.

important at $r \lesssim 15$. At the faint end, redshift incompleteness in all SDSS bands is of course dominated by the intrinsic faint limit of the Main Galaxy Sample (Strauss et al., 2002), i.e. $r = 17.77$. In order to set our limits, we have made use of the SDSS galaxy number counts, that will be properly discussed in Section 2.3.1. For each galaxy sample, we have taken the magnitude range where the number of galaxies rises at constant rate in each SDSS band, with a 0.1-dex deviation allowance (see Figure 4, where galaxy number counts have been scaled by an Euclidean, non-evolving model). In the r-band, we use a faint limit of $r = 17.6$, instead of $r = 17.77$. By restricting ourselves to this more conservative flux limit we can define a complete sample, avoiding some possible inconsistencies in the photometric calibration at $r > 17.6$ (see Blanton et al., 2001, for a detailed explanation). For consistency, in the rest of the SDSS bands we have ensured that the number of objects with $r > 17.6$ is less than 2%. Within the apparent magnitude ranges of Table 2.1, we estimate that redshift completeness is $\sim 85\%$ in all SDSS bands. In Section 2.2.2, we will discuss on redshift incompleteness issues in the SDSS in more detail.

In Figure 2.1, we show the redshift distribution for the Parent Sample in each SDSS photometric band. This figure illustrates the motivation for the redshift limits given in Table 2.1. The lower redshift limit is set to $z = 0.02$ to avoid the redshift incompleteness that affects the very bright and nearby galaxies. The upper redshift limit corresponds to the redshift at which 98% of objects are selected in each sample and is set for consistency.

2.2.2 Redshift Completeness

It is well known that the brightest galaxies in the SDSS are affected by severe redshift incompleteness. This effect is especially important for nearby galaxies. Apparently large and complex objects represent

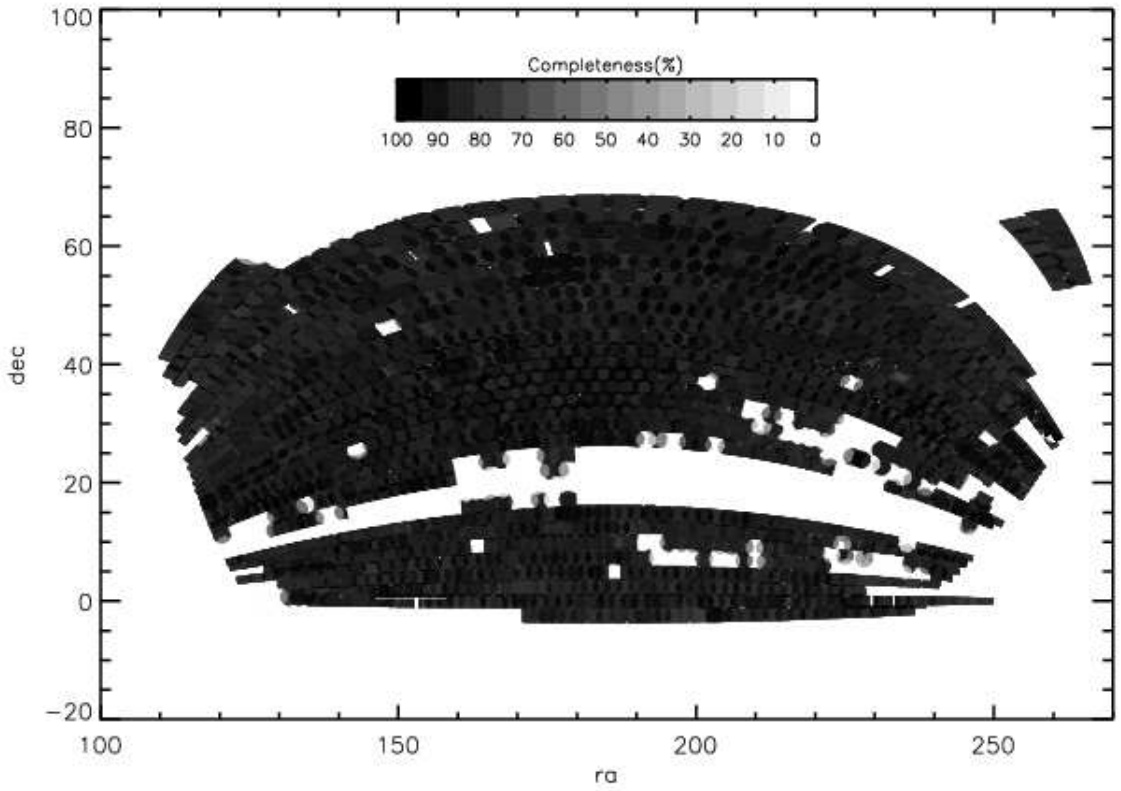

Figure 2.3 Angular redshift completeness in the SDSS DR6 spectroscopic catalog in the right ascension range $110^\circ < RA < 270^\circ$, that encompasses $\sim 99\%$ of the survey. Each polygon in the plot is an area of constant completeness. Polygons are also color-coded in white to black tones, representing from 0% to 100% completeness.

a major task for the photometric SDSS pipelines (Strauss et al., 2002). In order to illustrate this, we have cross-matched the SDSS DR5 and the Third Reference Catalog of Bright Galaxies (RC3, Corwin et al. 1994), which is supposed to be reasonably complete for very bright and nearby galaxies ($r \lesssim 15$ and $z \lesssim 0.05$). We found that around 66.5% of RC3 galaxies lying within the DR5 sky coverage, and therefore apparently imaged by the SDSS, have a corresponding entry in the photometric catalog. Moreover, only half of these photometric objects have a spectroscopic counterpart. Therefore, we estimate that the SDSS successfully gets redshift of $\sim$ 1/3 of the very bright and large galaxies.

In addition, and affecting the entire magnitude range, is the so-called fiber collisions problem. This source of incompleteness, which is due to the fact that fibers cannot be placed closer than $55''$, is responsible for most incompleteness in the SDSS data. Strauss et al. (2002) estimated that affects $\sim 6\%$ of all target galaxies.

In Figure 2.2, we show how redshift completeness varies with apparent magnitude in the r band. Redshift completeness falls from $\sim 85\%$ at $r \simeq 14$ to $\sim 50\%$ at $r \simeq 12$. At the faint end, however, it reaches a plateau at 90% (down to $r \simeq 17.8$), which is in agreement with Strauss et al. (2002). In the rest of the SDSS bands, redshift completeness not only decreases in the bright end, but also at the faint end. This is due to the intrinsic r-band faint limit of the Main Galaxy Sample at $r = 17.77$. Interestingly, the apparent magnitude, $m_{max}(j)$, at which this decrease occurs varies between bands as a result of the dispersion in galaxy colors. We have checked that, by imposing $m_{max}(j) > m(j) > m_{min}(j)$, we ensure that each galaxy sample is approximately 85% complete in any magnitude bin.

Another aspect of redshift completeness that should be carefully taken into account is its angular variation. In this chapter we deal with the calculation of number counts and luminosity functions in the entire SDSS DR6 spectroscopic catalog. These tasks require knowing the effective area covered by the galaxy sample on the sky. The survey masks, available at the NYU-VAGC website, provide us with this information. Redshift completeness is far from being uniform across the sky. In Figure 2.3, we show the angular sky redshift completeness of the SDSS DR6 spectroscopic catalog for a major part of the survey ($\sim 99\%$), just excluding objects outside the range $110^\circ < RA < 270^\circ$, for the sake of simplicity. The plot has been pixelized so that each pixel - or polygon - is an area of constant completeness. Polygons are also color-coded in white to black tones, the latter meaning 100% completeness. Approximately 10% of the total area is covered by polygons with less than 80% completeness. In principle, by taking these regions into account we would be overestimating the area covered on the sky by our samples. However, since the main source of incompleteness in the SDSS, as discussed above, comes from the fiber spacing constraint, one would expect the majority of the low-completeness polygons to lie in highly-dense regions on the sky. In order not to under-represent these regions, we have not excluded incomplete polygons in our analysis. We have checked, however, that the exclusion of these polygons would not alter our results in any significant way.

Finally, we cannot discard the possibility that the SDSS spectroscopic catalog is incomplete for very low surface brightness objects, i.e. $\mu_{r,50} \gtrsim 24$ (see Strauss et al. 2002). The presence of this selection effect in the data could in principle affect our results. However, we have evidence that the surface brightness of most galaxies in the SDSS, and consequently in our galaxy samples, is far greater than $\mu_{r,50} \simeq 24$ (Blanton et al., 2003b, 2005b).

2.3 Results

2.3.1 Number counts

In Figure 2.4 we plot with different symbols the logarithm of the number of galaxies per unit area and apparent magnitude (actually, half magnitude), scaled by an Euclidean model for all SDSS bands. We choose, arbitrarily, the following model for Euclidean counts:

$$N_{Euclidean} = 10^{0.6(x-18)} \tag{2.1}$$

where x represents the apparent magnitude in each SDSS band. Note that the Euclidean model assumes that the luminosity of galaxies remain constant in time and that the number of galaxies per unit volume is conserved after they are formed. It can be demonstrated that in such a model logarithmic GNC increase with slope 0.6. The zero-point of the model in Equation 2.1 gives the normalization of the GNC.

In the r band, galaxy number counts increase by a factor of about 10 from $r \sim 12$ to $r \sim 13.5$. This magnitude range is strongly affected by redshift incompleteness, as discussed above (see also Figure 2.2). From $r \sim 14$ to $r \sim 18$, counts rise at approximately the same rate as that of the model (Equation 2.1). At $r \sim 18$, where the SDSS spectroscopic faint-end limit is set ($r = 17.77$), galaxy number counts fall sharply.

In the rest of the bands, the behavior is very similar at the bright end and is also due to redshift incompleteness. In the z band, counts start to follow the Euclidean model at $z \sim 13.5$. In the i band, this happens at $i \sim 13.5$; in the g band, at $g \sim 14.5$ and, in the u-band, at $u \sim 16.5$. at the faint end, galaxy number counts fall roughly at $u \sim 19$, $g \sim 18$, $i \sim 17$ and $z \sim 17$ (see Table 2.1 for more accurate values). This decrease is obviously less pronounced in this band than it is in the r band, due to the different colors of galaxies. In addition, the slope of this faint-end decrease in the galaxy number counts is considerably steeper in the red bands than it is in the bluer bands. This is due to the fact that the dispersion in the $(u - r)$ colors in notably larger than that of the $(r - z)$ colors (see e.g. Blanton et al. 2003a).

Number counts are consistent with an Euclidean, non-evolving Universe in all SDSS bands within the magnitude and redshift ranges given in Table 2.1. Within these ranges, the slopes in units of mag^{-1} of the SDSS DR6 galaxy number counts are also listed in this table. Note that the small deviations that we find with respect to the Euclidean model are probably due to the presence of the large-scale structure in the SDSS.

Our results for the nearby universe are in agreement with a number of previous works (see Feulner et al. 2007 for a review). However, only Yasuda et al. (2001) presented galaxy number counts obtained using the SDSS . The authors used imaging data taken during the commissioning phase. They found an Euclidean-like behavior up to magnitude $m_{min} \sim 12$ - except for the u band, where $m_{min} \sim 14$ - (see their Figure 8). The lack of galaxies in the very bright end of Figure 2.4 with respect to photometric number counts from Yasuda et al. (2001) is partially due to the strong redshift incompleteness that affects the SDSS spectroscopic catalog (see Section 2.2.2). However, our results and those from Yasuda et al. (2001) are not directly comparable. Firstly, they used a very limited sample, in terms of sky

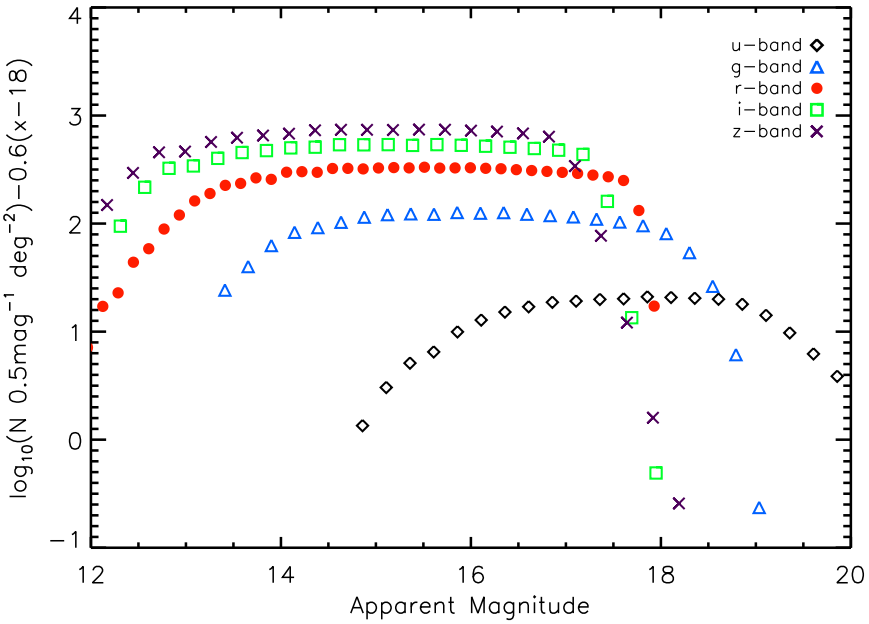

Figure 2.4 Galaxy number counts in all SDSS bands scaled by an Euclidean model in bins of half a magnitude. Poissonian errors are of similar size as symbols, so they are not shown.

coverage (~ 230 deg^2). Secondly, comparing spectroscopic and photometric results is always tricky.

Many galaxy formation and evolution models use observational constraints obtained with the SDSS spectroscopic catalog. Providing GNC for this catalog, for the first time, is therefore extremely useful.

2.3.2 Luminosity Functions

In order to estimate the luminosity function of galaxies in each SDSS photometric band, we take absolute magnitudes and K-corrections from the NYU VAGC DR6 LSS catalog. Following Blanton et al. (2003b), absolute magnitudes are calculated with the SDSS photometric bands shifted to $z = 0.1$. With this convention, the absolute magnitude of a galaxy in a given band shifted to $z = 0.1$, $M_{0.1j}$, can be constructed from its apparent magnitude at $z = 0$, m_j, and its redshift z as follows:

$$M_{0.1j} = m_j - 5log_{10}h - DM(z) - K_{0.1j}(z) \tag{2.2}$$

where $DM(z)$ is the distance modulus (which depends also on the cosmological parameters) and $K_{0.1j}(z)$, the K-correction for the galaxy in the shifted band $^{0.1}j$. In Chapter 5 and in the Appendix, we will discuss K-corrections in more detail. Blanton et al. (2003b) included another correction in Expression 2.2 to account for the evolution of the luminosity of the galaxy from redshift z to $z = 0.1$, the so-called evolution correction. Several authors have tried to implement a similar correction using different approaches (i.e. Norberg et al. 2002, for the 2dFGRS). It seems tempting to think, however, that this correction should be very small within the redshift ranges we use here. As this refinement is still subject to large uncertainties, we have opted to present uncorrected luminosity functions. Instead, we will discuss in detail the effect of evolution on our LFs in Section 2.3.3.

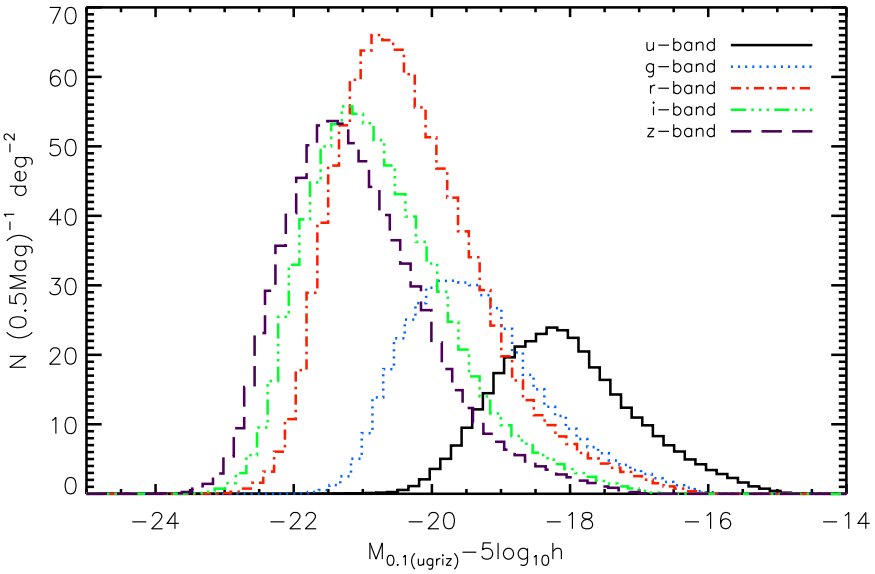

Figure 2.5 Absolute magnitude distribution in each SDSS galaxy sample. The shape of the distributions is very similar, but mean values move towards brighter magnitude bins from the $^{0.1}u$ band to the $^{0.1}z$ band, as redder objects are, on average, brighter than bluer objects.

In Figure 2.5, we show the distribution of K-corrected absolute magnitudes in each galaxy sample. The shape of the absolute magnitude distribution is very similar in all SDSS bands: a gaussian-like distribution slightly skewed to fainter magnitudes. However, mean values move towards brighter bins from the $^{0.1}u$ band ($M_{0.1_u} \sim -18$) to the $^{0.1}z$ band ($M_{0.1_z} \sim -21.5$), which is consistent with the fact that red objects are, on average, brighter than blue objects. In Figure 2.6, we also show the bimodal $^{0.1}u - ^{0.1}r$ color distribution of galaxies in our $^{0.1}r$-band sample. With a dashed line we represent the demarcation commonly used to separate red and blue objects (see Strateva et al. 2001).

We use the Stepwise Maximum Likelihood method (SWML; Efstathiou et al. 1988) to estimate the LF of galaxies, that is commonly expressed in the literature as $\Phi(L)$. This technique, which does not rely on any assumption about the shape of $\Phi(L)$, is based on the estimation of the probability that each galaxy in our sample is located at redshift z and has absolute magnitude M, while satisfying the survey limits. We refer the reader to Efstathiou et al. (1988) and Norberg et al. (2002) for more details about this method. The SWML method requires an independent estimation of the normalization constant, n. We use the following prescription proposed by Davis & Huchra (1982) based on the selection function of each galaxy, $\phi(z_i)$, and the maximum volume encompassed by the sample, V_{max}:

$$n = \frac{1}{V_{max}} \sum_i \frac{1}{\phi(z_i)} \tag{2.3}$$

In Figure 2.7, we show the SWML estimate of the SDSS DR6 LF in the $^{0.1}r$ band. In addition, we over-plot in a dashed line the corresponding best-fitting Schechter function, which can be expressed in the following way:

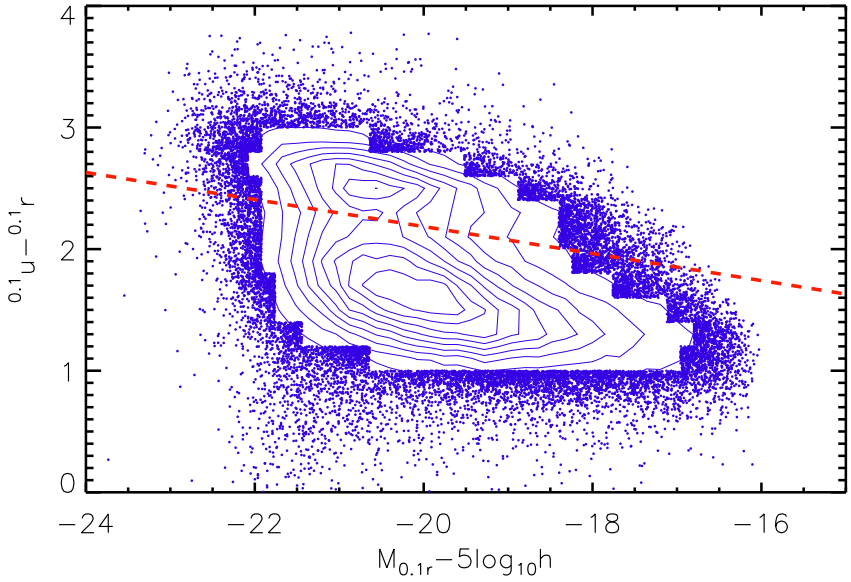

Figure 2.6 The $(^{0.1}u - ^{0.1}r)$ vs. $M_{0.1_r}$ color-magnitude diagram in the $^{0.1}r$ band. The dashed line represents the demarcation commonly used to separate red and blue galaxies.

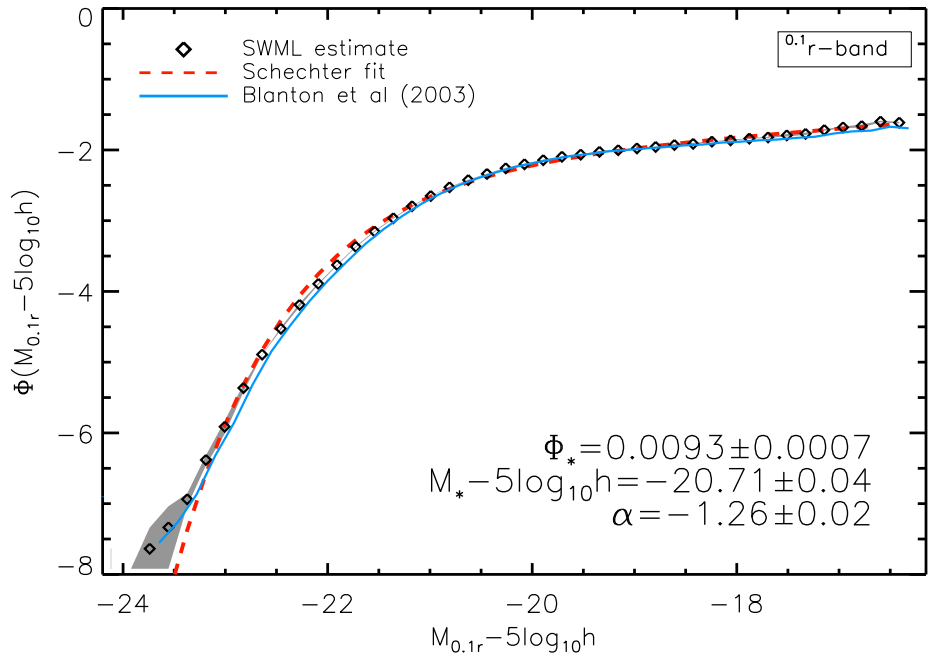

Figure 2.7 The $^{0.1}r$-band SDSS DR6 Luminosity Function. The SWML LF estimate is shown in diamonds. The dashed line represents the best-fitting Schechter function and the solid line, the $^{0.1}r$-band LF from Blanton et al. (2003b). best-fitting values of Schechter parameters α, M_* and Φ_* are also shown. Shaded regions represent the 1σ uncertainty calculated using a bootstrapping technique.

$$\Phi(M) = 0.4\log(10)\Phi_* 10^{-0.4(M-M_*)(\alpha+1)} \exp\left[-10^{-0.4(M-M_*)}\right] \quad (4)$$

where α, M_* and Φ_* are the three parameters to fit. Values of these parameters for the best-fitting Schechter function are provided in Table 2.2. For comparison, we also show the LF of Blanton et al. (2003b) in a solid line. This comparison will be addressed in the Discussion section. To calculate errors in the SWML estimates of the LF we perform a bootstrapping analysis using 1,000 random sub-samples of 1/3 the number of objects in each sample. In Figure 2.7, shaded regions represent the 1σ uncertainty obtained from this method.

Because of the big number statistics that we have, with about 450,000 galaxies in the ${}^{0.1}r$ band, errors are only significant at the very bright end of the LF. At the faint end, we can go down to $M_{{}^{0.1}r} \sim -16.5$, which means that we can build the LF with unprecedented precision within a very large range of magnitudes. As we will see below, the above statements hold for all SDSS bands. Our ${}^{0.1}r$-band LF is reasonably well fitted by a Schechter LF with a faint-end slope $\alpha = -1.26$. It is only at the very bright end where this best-fitting Schechter LF starts to underestimate our LF. At $M_{{}^{0.1}r} \lesssim -23.5$, statistics are poor and errors become increasingly large.

In Figure 2.8 we present, in the same way as in Figure 2.7, SWML estimates of the LF in bands ${}^{0.1}u$, ${}^{0.1}g$, ${}^{0.1}i$ and ${}^{0.1}z$, as well as their corresponding best-fitting Schechter LF. Values of best-fitting Schechter parameters are also given in Table 2.2. As in the ${}^{0.1}r$ band, errors are only significant at the very bright end of the ${}^{0.1}u$, ${}^{0.1}g$, ${}^{0.1}i$ and ${}^{0.1}z$ band LFs. In addition, we can go down to very faint magnitudes without losing precision.

In the very blue ${}^{0.1}u$ band, the shape of our SDSS DR6 LF is consistent with a Schechter LF with a slightly positive faint-end slope (corresponding to $\alpha = -1.05$). However, at the bright end, we find a remarkable excess with respect to the best-fitting Schechter LF. This excess, of ~ 1.7 dex at $M_{{}^{0.1}u} \simeq -20.5$, is very significant within the magnitude range $-20.5 < M_{{}^{0.1}u} \lesssim -22$. In the ${}^{0.1}g$ band, this *bright-end bump* (BEB) weakens considerably, but it is probably still significant, even though errors are large according to our bootstrapping analysis. In this band, our SDSS LF is very well fitted by a Schechter LF with a positive faint-end slope, corresponding to $\alpha = -1.10$. Only at the very bright-end, where the excess is still noticeable, do we find some discrepancy. Below, we provide a preliminary analysis and discussion on the nature of this bump at the bright end of the ${}^{0.1}u$-band LF, that may have important implications in terms of galaxy formation and evolution.

In the redder bands we find a positive faint-end slope, corresponding to $\alpha = -1.14$ in the ${}^{0.1}i$ band and $\alpha = -1.26$ in the ${}^{0.1}z$ band. The BEB has diminished but is still clearly significant in the ${}^{0.1}i$ band and only disappears completely in the very red ${}^{0.1}z$ band. It is interesting to note that, from the ${}^{0.1}u$ band to the ${}^{0.1}z$ band, the shape of the SWML estimate of the SDSS LF changes following a clear pattern. The faint-end slope increases towards the redder bands (see Table 2.2), being almost flat in the ${}^{0.1}u$ band and remarkably steep in the ${}^{0.1}z$-band. In this sense, the ${}^{0.1}r$ band SDSS LF seems to slightly deviate from this trend. In the faint-end, we find a slope that is a bit larger than we could expect ($\alpha = -1.26$), but this could be just a consequence of the fact that the entire SDSS

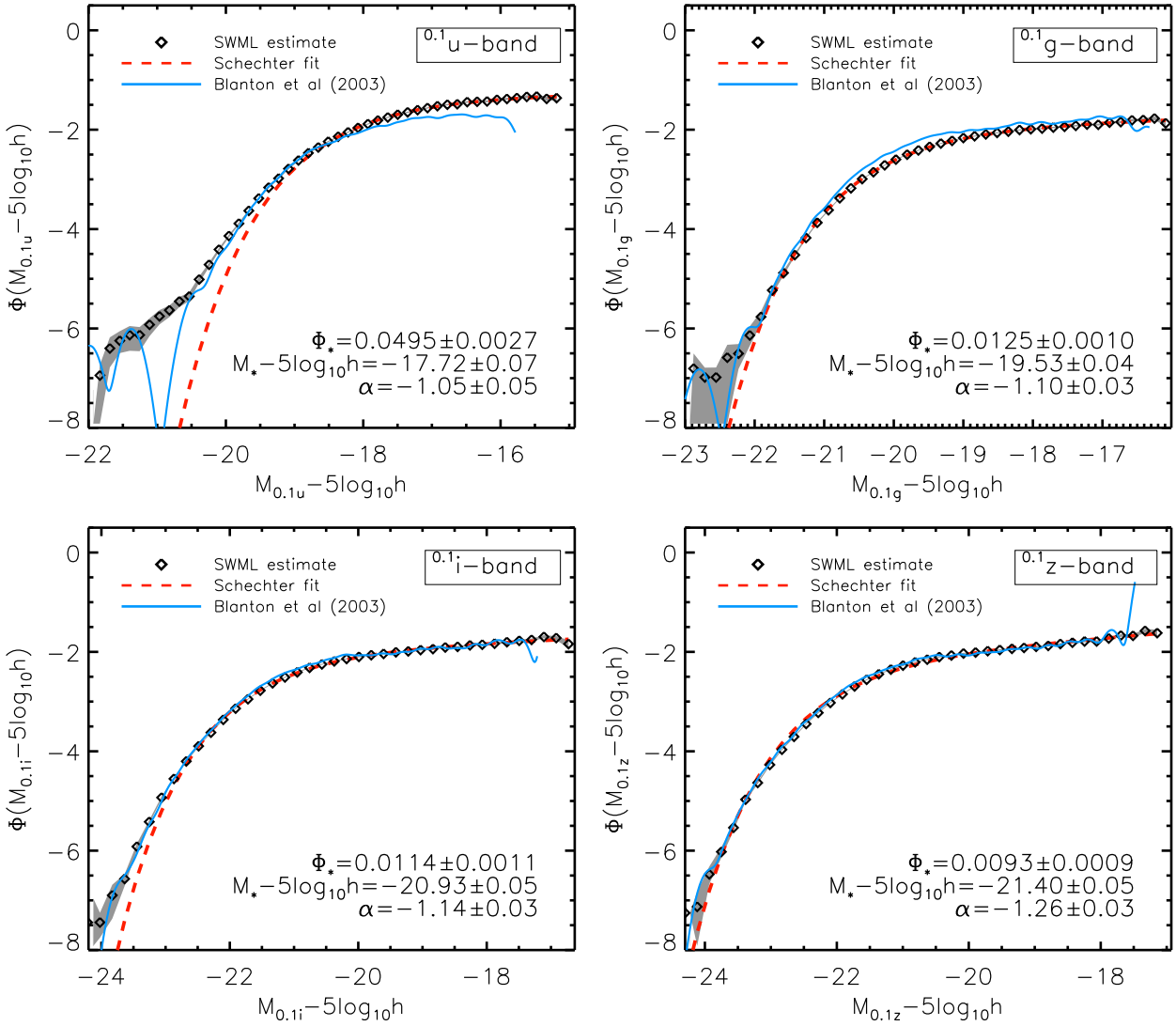

Figure 2.8 The SDSS DR6 luminosity functions in bands $^{0.1}u$, $^{0.1}g$, $^{0.1}i$ and $^{0.1}z$. SWML LF estimates are shown in diamonds and best-fitting Schechter functions are represented by dashed lines. In addition, we over-plot in each panel the LF from Blanton et al. (2003b). best-fitting values of Schechter parameters α, M_* and Φ_* are also shown. Shaded regions represent the 1σ uncertainty calculated using a bootstrapping technique.

This work			
Band	$\Phi_*(10^{-2}h^3$ Mpc$^{-3})$	$M_* - 5log_{10}h$	α
$^{0.1}u$	4.95 ± 0.27	-17.72 ± 0.07	-1.05 ± 0.05
$^{0.1}g$	1.25 ± 0.10	-19.53 ± 0.04	-1.10 ± 0.03
$^{0.1}r$	0.93 ± 0.07	-20.71 ± 0.04	-1.26 ± 0.02
$^{0.1}i$	1.14 ± 0.11	-20.93 ± 0.05	-1.14 ± 0.03
$^{0.1}z$	0.93 ± 0.09	-21.40 ± 0.05	-1.26 ± 0.03
Blanton et al. (2003b)			
Band	$\Phi_*(10^{-2}h^3$ Mpc$^{-3})$	$M_* - 5log_{10}h$	α
$^{0.1}u$	3.05 ± 0.33	-17.93 ± 0.03	-0.92 ± 0.07
$^{0.1}g$	2.18 ± 0.08	-19.39 ± 0.02	-0.89 ± 0.03
$^{0.1}r$	1.49 ± 0.04	-20.44 ± 0.01	-1.05 ± 0.01
$^{0.1}i$	1.47 ± 0.04	-20.82 ± 0.02	-1.00 ± 0.02
$^{0.1}z$	1.35 ± 0.04	-21.18 ± 0.02	-1.08 ± 0.02

Table 2.2 Values of Schechter parameters α, M_* and Φ_* of the best-fitting Schechter function in all SDSS bands for this work and for Blanton et al. (2003b). Note that it is not convenient to compare LFs by just looking at their best-fitting Schechter parameters. Also bear in mind that both estimates are not strictly comparable, due to differences in the sample selection, the treatment of galaxy evolution and, especially, the size of the samples. In Section 2.3.3 and Section 2.4 we discuss on the importance of these inconsistencies.

spectroscopic sample was selected in this band.

The BEB that shows up clearly in the $^{0.1}u$-band LF, and partially in the $^{0.1}g$-band and $^{0.1}i$-band LFs, is an interesting discovery that may have implications for our understanding of galaxy formation and evolution. In order to investigate the nature of the objects that populate it, we have selected all galaxies brighter than -20.5 in the $^{0.1}u$-band sample. We will hereafter refer to this population as BEB galaxies and to their corresponding sample, which is comprised of 252 objects, as the BEB sample. It is convenient to remind that, although here we focus on the $^{0.1}u$ band, significant bright-end excesses have been found in both the $^{0.1}g$ band and the $^{0.1}i$ band. It is in the very blue $^{0.1}u$ band, however, were this population stands out more prominently.

We have visually inspected the spectra of all galaxies in the BEB sample. In addition, we have taken the spectral classification of each individual galaxy, based on emission line ratios, from the NYU VAGC. According to this analysis, about 60% of objects have a typical QSO or Seyfert I spectrum, $\sim 8\%$ of sources are classified as starburst (SB) galaxies, $\sim 12\%$ as star-forming (SF) galaxies and $\sim 20\%$ of objects are normal galaxies that show no significant emission lines in their spectra (typically elliptical galaxies). We hereafter consider three types of BEB galaxies: QSOs/Seyfert I's ($\sim 60\%$), SB/SF galaxies ($\sim 20\%$) and normal galaxies ($\sim 20\%$). In Figure 2.9 we plot in the $(^{0.1}u - ^{0.1}r)$ vs. $M_{^{0.1}u}$ colour-magnitude diagram (CMD) of the $^{0.1}u$-band sample the three types of BEB galaxies discussed above. In general, these galaxies form a relatively tight sequence in the bright-end of the CMD, showing a considerable color dispersion. Note that, in contrast to Figure 2.6, density contours are now log-spaced and hence, these objects occupy an extremely underpopulated region in the CMD. In the left-hand plot, QSOs/Seyferts I's show the smallest color dispersion and are, on average, the bluest: $\langle(^{0.1}u - ^{0.1}r)\rangle = 0.68$. Both SF/SB galaxies (middle plot) and normal galaxies (right-hand plot) show

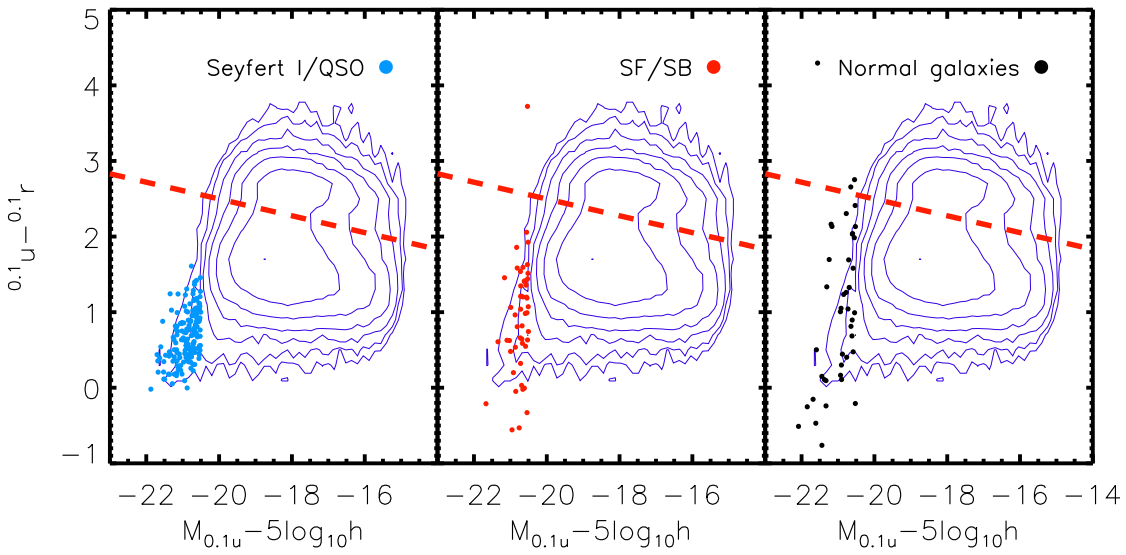

Figure 2.9 The $(^{0.1}u - ^{0.1}r)$ vs. $M_{0.1u}$ color-magnitude diagram for the three types of bright-end bump galaxies considered: QSOs/Seyferts I's (left-hand plot), SF/SB galaxies (middle plot) and normal galaxies (right-hand plot). The underlying CMD of the entire $^{0.1}u$-band sample from which BEB galaxies are selected is shown in log-spaced contours.

much larger color dispersion and are, on average, considerably redder, with $\langle(^{0.1}u - ^{0.1}r)\rangle = 0.89$ and $\langle(^{0.1}u - ^{0.1}r)\rangle = 0.99$, respectively. In the $^{0.1}u$ band, QSOs/Seyfert I's and normal galaxies are, on average, the brightest among the BEB galaxies: $\langle M_{0.1u}\rangle = (-20.91^{QSO/SI}, -20.75^{SF/SB}, -20.99^{normal})$.

BEB galaxies are typically at high redshift relative to the average redshift of the $^{0.1}u$-band sample from which they are drawn. The mean redshift in the BEB sample is $\langle z\rangle \sim 0.155$ while in the entire $^{0.1}u$-band sample is $\langle z\rangle \sim 0.080$. Note that the redshift limit of $z < 0.17$ that we have imposed for consistency in the $^{0.1}u$-band sample turns out to be slightly restrictive for BEB galaxies, according to their redshift distribution. However, the above results remain unchanged as far as the mean properties of the BEB galaxies are concerned, when we extend this limit to $z \sim 0.2$. The size of the BEB sample would increase to $\sim 1,000$ objects, though.

At this point, it is necessary to remark that this is a preliminary analysis. We have performed a rough classification of the BEB galaxies, and this is still subject to some uncertainty. However, our aim is to provide a first approach to the nature of this population of galaxies. In Section 2.4 we speculate on possible implications of these results.

Finally, in Figure 2.10, we present, in the same way as in Figure 2.7 and Figure 2.8, SWML estimates of the SDSS $^{0.1}r$-band LF for blue and red galaxies separately, along with their corresponding best-fitting Schechter LF. Values of best-fitting Schechter parameters are also shown in the figure. The demarcation that we use to separate blue and red objects is represented by a dashed line in Figure 2.6. The SDSS $^{0.1}r$-band LF of blue galaxies is well fitted by a Schechter LF with $\alpha = -1.41$. In contrast, the SDSS $^{0.1}r$-band LF of red galaxies has a negative faint-end slope, corresponding with $\alpha = -0.81$. At the bright end ($M_{0.1r} \lesssim -21.5$), the blue LF falls remarkably below the red LF.

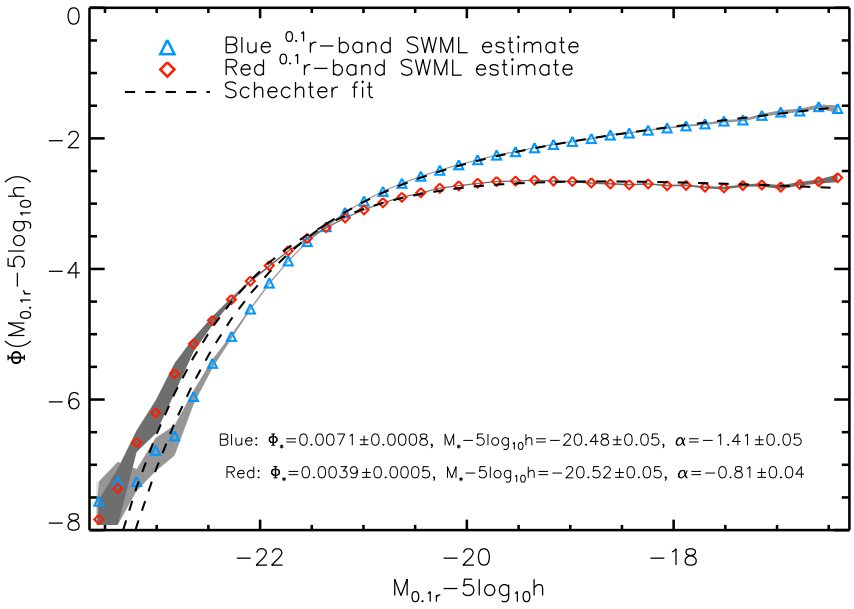

Figure 2.10 The $^{0.1}r$-band SDSS DR6 luminosity function for blue and red galaxies separately. The SWML LF estimates are shown in diamonds. The dashed lines represents the best-fitting Schechter function. best-fitting values of Schechter parameters α, M_* and Φ_* for both blue and red galaxies are also shown in the figure. Shaded regions represent the 1σ uncertainty calculated using a bootstrapping technique.

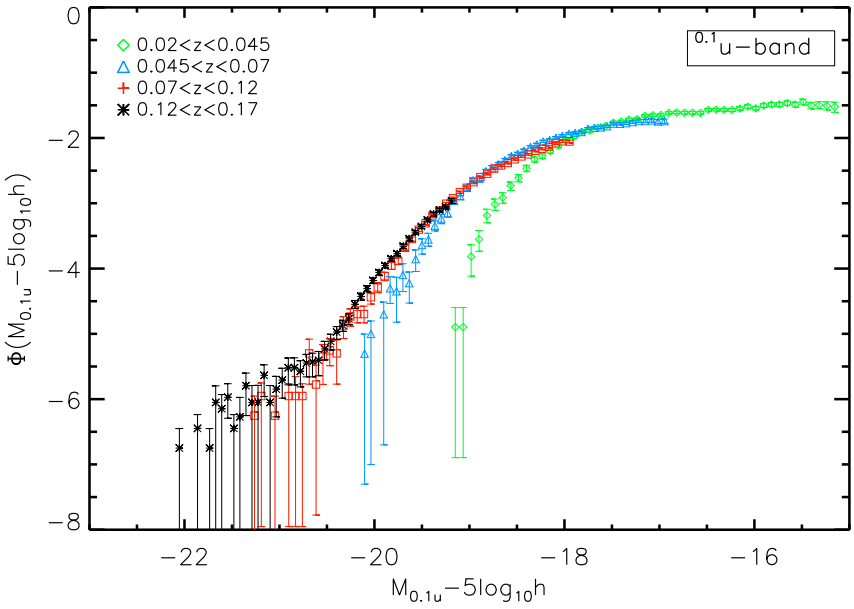

Figure 2.11 The evolution of the $^{0.1}u$-band LF in four redshift slices. Error bars represent the 1σ uncertainty calculated using a bootstrapping technique. The number of objects in each redshift bin is listed in Table 2.3.

2.3.3 The effect of galaxy evolution

In previous sections, we assumed that the evolution in the luminosity of galaxies within the redshift intervals considered had a negligible effect on the SDSS LFs. This idea is based on the little redshift depth of our samples ($z \lesssim 0.2$). In this section, we elaborate on the validity of such an assumption. Accounting for galaxy evolution is a rather difficult task. In principle, one should expect that different types of galaxies evolve with time in different ways. The evolutionary path of an average galaxy of a given type not only depends on intrinsic properties such as mass but also on the environment, with a number of possible physical processes being proposed (i.e. minor and major mergers, harrassment, tidal disruption). Unfortunately, our knowledge of these and other fundamental processes involved in a galaxy's evolution seems insufficient to treat this problem in an accurate way. Some previous works, however, have tried to account for evolution in a simple way, using different approaches. Blanton et al. (2003b) implemented a correction based on the change in the absolute magnitude distribution with redshift (Blanton, private communication). When correcting to $z = 0.1$, this evolution correction ranges from ~ -0.15 mag for objects at $z \sim 0$ to $+0.2$ mag for galaxies at $z \sim 0.2$ and is responsible, according to this study, for the flattening of the faint-end slopes of their LFs. On the other hand, Norberg et al. (2002) used four different evolution corrections depending on the spectral type of the galaxy to calculate the b_j-band LF for the 2dFGRS. Even after applying this correction, they find a LF in the b_j-band (which is roughly comparable to our SDSS $^{0.1}g$-band) with a remarkably steep faint-end slope, corresponding to $\alpha \sim -1.2$. Although efforts like these are extremely valuable and contribute to the understanding of the effect of evolution on the LF of galaxies, the approximations they use are still subject to major uncertainties. These uncertainties, which derive from our still limited knowledge of the process of galaxy evolution itself, become especially significant when dealing with relatively narrow redshift ranges, like those we use in this work.

In this work, we use a different approach and we investigate the evolution of the LF by binning our galaxy samples by redshift. The aim of this section is to show that accounting for evolution is not likely to alter, in any significant way, our main results, i.e. the prominent bright-end excess in the $^{0.1}u$-band LF and the relatively steep faint-end slopes in all SDSS bands. For the sake of simplicity, we will address this question by studying the evolution of the $^{0.1}u$-band and the $^{0.1}r$-band LFs.

In Figure 2.11, we plot the $^{0.1}u$-band LF in 4 different redshift bins. The slicing has been designed to ensure that the number of objects is similar in the first 3 redshift bins and enough to guarantee a reliable estimation of the LF in the more distant one. The number of galaxies in each slice is listed in Table 2.3. Interestingly, we find that the $^{0.1}u$-band LF seems to move towards brighter magnitudes with redshift. The BEB appears only in the more distant slices, at $z > 0.10$. The bright-end incompleteness of the SDSS makes it difficult to know if this is just a selection effect (as the sampling volume is smaller at lower redshift) or a natural property of this population of galaxies. In the redshift interval $0.12 < z < 0.17$, the BEB stands up clearly. It is important to stress that by narrowing our redshift ranges we minimize the effect of evolution. Figure 2.11 proves the presence of a prominent bright-end excess in the $^{0.1}u$-band LF at $z \sim 0.15$. We could expect a standard evolution correction (as those cited above) to produce mainly a small horizontal shift of the BEB towards fainter magnitudes in the $z = 0.1$ $^{0.1}u$-band LF of Figure 2.8, but not of course the suppression of the excess.

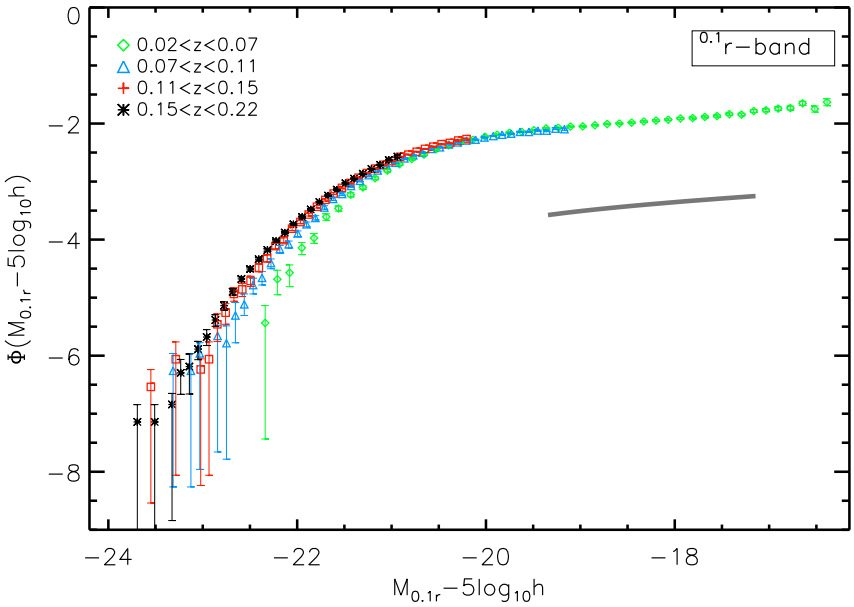

Figure 2.12 The evolution of the $^{0.1}r$-band LF in four redshift slices. Errors bars represent the 1σ uncertainty calculated using a bootstrapping technique. The number of objects in each redshift bin is listed in Table 2.3. The solid line represents a faint-end slope corresponding to $\alpha = -1.26$.

In fact, the existence of the BEB can be somehow inferred from previous works. Although noisy, clear signs of a bright-end excess can be found in the $^{0.1}u$-band and the $^{0.1}g$-band LFs of Blanton et al. (2003b) (see Figure 2.8). Similarly, a statistically significant bright-end deviation from the best-fitting Schechter LF is obtained, although not reported, in the b_j-band LF of Norberg et al. (2002).

In Figure 2.12 we present, in the same format as in Figure 2.11, the evolution of the $^{0.1}r$-band LF. As Figure 2.12 clearly shows, the faint-end of the SDSS LF is almost exclusively determined by low-redshift galaxies ($z \lesssim 0.07$). At higher redshifts the SDSS is not deep enough to capture the faint objects. The faint-end slope that we obtain in the first redshift bin ($0.02 < z < 0.07$) is absolutely consistent with that of the $z = 0.1$ $^{0.1}r$-band LF of Figure 2.8. To illustrate this, a solid line with a slope corresponding to $\alpha = -1.26$ has been plotted. A similar argument as that invoked to show that accounting for evolution could not produce the suppression of the BEB can be used here. As the faint-end slope of the LF is almost entirely determined by a narrow redshift range, it is unlikely that a realistic evolution correction could change this slope significantly. This assumption is in agreement with the theoretical work of Khochfar et al. (2007), where the evolution of the faint-end slope of the LF is investigated using semi-analytic models of galaxy formation. The authors report a very small variation of α at low redshift, being less than 1% between $z = 0.02$ and z 0.10.

To summarize, although allowing for galaxy evolution in our galaxy samples in a proper way could produce slight adjustments to our LFs (that could even translate into small variations in the very sensitive best-fitting Schechter parameters, mainly in M_*), it seems unlikely, as we have shown above, that this correction could produce a qualitative alteration of our main results. In any case, it still convenient not to forget the issue of evolution when comparing our best-fitting Schechter parameters

$^{0.1}u$-band		$^{0.1}r$-band	
Redshift bin	Number	Redfshift bin	Number
$0.02 < z < 0.045$	40,738	$0.02 < z < 0.07$	124,608
$0.045 < z < 0.07$	46,310	$0.07 < z < 0.11$	142,183
$0.07 < z < 0.12$	54,843	$0.11 < z < 0.15$	102,871
$0.12 < z < 0.17$	9,737	$0.15 < z < 0.22$	66,070

Table 2.3 Number of galaxies in each redshift slice of the $^{0.1}u$-band and the $^{0.1}r$-band sample, corresponding to the evolution plots of Figure 2.11 and Figure 2.12, respectively.

with those from Blanton et al. (2003b). In Section 2.4, we will discuss other inconsistencies between both works that are more likely to be responsible for the discrepancies we find.

2.3.4 Petrosian magnitudes vs. model magnitudes

Petrosian magnitudes have been widely used in SDSS works. These magnitudes, which are based on a modified form of the Petrosian (1976) system, measure galaxy fluxes within a circular aperture whose radius is defined by the shape of the azimuthally averaged light profile. The motivation is to measure a constant fraction of the total light of the galaxy, independent of the position and distance of the object. The SDSS, however, provides other measures of fluxes and magnitudes, each of them being appropriate for a different type of science. In particular, model magnitudes, which are calculated by fitting an exponential or a de Vaucouleurs light profile, are also regarded as a reliable estimate of the flux of a galaxy. These magnitudes have the advantage of measuring almost unbiased galaxy colors. In Figure 2.13 we show the scatter between these two different flux estimates in the SDSS DR6. In the right-hand plot, the r-band scatter is presented as a function of Petrosian apparent magnitude. Note how the scatter increases towards fainter magnitudes in this band. Throughout the entire range, Petrosian magnitudes measure less flux on average than model magnitudes, which is consistent with the way both measures are defined. In the fainter end, this loss can translate into as much as 0.5 mag. The cumulative distributions of the scatter in bands u,r and z are shown in the left-hand plot of Figure 2.13. We have excluded the distributions in bands g and i for the sake of clarity. Below, we address the question of how the choice of model magnitudes over the more conventional Petrosian magnitudes would affect our LF estimates.

We have obtained SDSS DR6 LFs using model magnitudes. In order to be consistent in our comparison, we have followed a similar selection scheme to that discussed in Section 2.2.1. In addition, absolute magnitudes have been calculated in the same way as in Equation 2.2 and K-corrections for model magnitudes have been taken from the NYU-VAGC. In Figure 2.14, we present, in the format used throughout the chapter, the SWML estimate of the SDSS DR6 LF in the $^{0.1}r$ band, obtained using model magnitudes. Instead of showing the LF of Blanton et al. (2003b) and in order to facilitate the comparison, we have over-plotted our SDSS DR6 LF obtained using Petrosian magnitudes. In the same way we show, in Figure 2.15 , SWML estimates of the SDSS DR6 LF for model magnitudes in bands $^{0.1}u$, $^{0.1}g$, $^{0.1}i$ and $^{0.1}z$. Selection limits for the new samples and values of best-fitting Schechter parameters for the model LFs in all bands are listed in Table 2.4 .

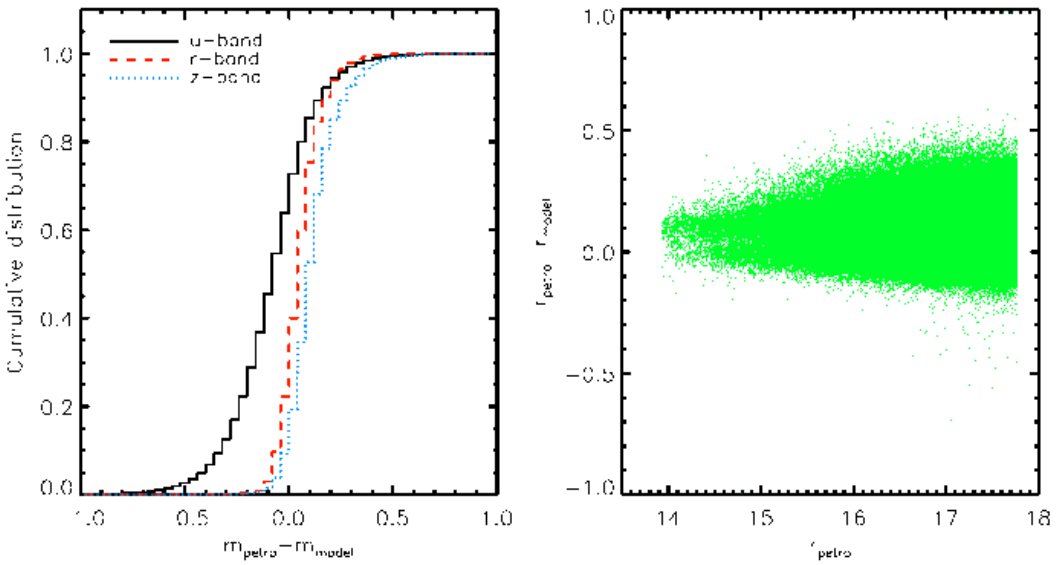

Figure 2.13 The scatter between Petrosian and model magnitudes. In the left-hand plot, the cumulative distribution of the scatter is shown in bands u,r and z. The distributions in bands g and i have been excluded for the sake of clarity. In the right-hand plot, we show the r-band scatter as a function of Petrosian apparent magnitude.

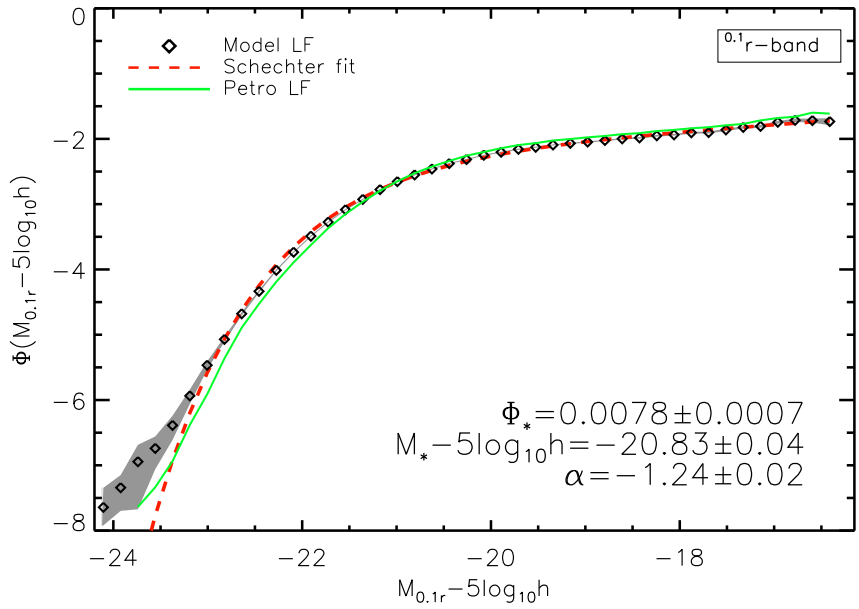

Figure 2.14 The $^{0.1}r$-band SDSS DR6 LF estimated using model magnitudes. The SWML LF estimate is shown in diamonds. The dashed line represents the best-fitting Schechter function and the solid line, the $^{0.1}r$-band SDSS DR6 LF estimated using Petrosian magnitudes. best-fitting values of Schechter parameters α, M_* and Φ_* are also shown. Shaded regions represent the 1σ uncertainty calculated using a bootstrapping technique.

Band	Number	m_{min}	m_{max}	z_{min}	z_{max}	$\Phi_*(10^{-2}h^3$ Mpc$^{-3})$	$M_* - 5log_{10}h$	α
u	159,018	16.45	18.90	0.020	0.170	3.07 ± 0.23	-17.70 ± 0.05	-1.00 ± 0.04
g	256,952	14.55	17.91	0.020	0.160	1.19 ± 0.08	-19.59 ± 0.04	-1.10 ± 0.03
r	466,280	13.93	17.6	0.020	0.220	0.78 ± 0.07	-20.83 ± 0.04	-1.24 ± 0.02
i	461,928	13.55	17.20	0.020	0.235	0.81 ± 0.11	-21.13 ± 0.03	-1.17 ± 0.03
z	422,643	13.40	16.85	0.020	0.240	0.77 ± 0.08	-21.52 ± 0.03	-1.20 ± 0.03

Table 2.4 Values of Schechter parameters α, M_* and Φ_* of the best-fitting Schechter function in all SDSS bands for the SDSS DR6 LF estimated using model magnitudes.

As we can see in Figure 2.14 and Figure 2.15, the scatter between Petrosian and model magnitudes has a noticeable impact on the LF. In the $^{0.1}u$ band, the LF is shifted downwards at the faint end, where the slope also flattens slightly: $\alpha = -1.00$ for the model LF and $\alpha = -1.05$ for the Petrosian LF. In contrast, in the rest of the photometric bands, the main variation with respect to the Petrosian LF involves M_*, which is brighter in the model LF. This variation tends to increase towards the redder bands, reaching a value of ~ 0.2 mag in the $^{0.1}i$-band and ~ 0.12 in the very red $^{0.1}z$-band. This pattern is clearly consistent with the scatter in each band.

Note that some of the differences discussed above might be due, not only to the scatter in magnitude but also to slight differences in the sample selection with respect to the Petrosian LFs. This is, at least partially, behind the flattening of the faint-end slope in the $^{0.1}u$-band LF, where we have extended the apparent magnitude faint limit in almost 0.1 mag. As we have stressed several times throughout this chapter, the faint end of the bluer bands (especially the $^{0.1}u$ band) is extremely sensitive to variations in the faint limit of the sample, as a consequence of the large dispersion of the $^{0.1}u - ^{0.1}r$ color. Finally, and important for this work, the results shown in Figure 2.15 seem to rule out the possibility that one of our most prominent findings, the so-called bright-end bump, is an artefact caused by inconsistencies in the photometry.

2.3.5 Luminosity Densities

In order to estimate the luminosity density, ρ, in all SDSS bands, we have integrated within the absolute magnitude ranges defined by the limits shown in Table 2.5. In this table, we list ρ in all bands obtained using the SWML estimates of the luminosity functions shown in Figure 2.7 and Figure 2.8. Luminosity densities are expressed in *AB mags* $(\mathrm{Mpc}^3h^{-3})^{-1}$. We also give in Table 2.5 the effective wavelength, λ_{eff}, corresponding to each SDSS photometric band. Errors have been calculated using a similar bootstrapping technique as that discussed in Section 2.3.2. We have used 1000 random subsamples of 1/3 the total number of objects in each sample.

As expected, the Universe at $z \sim 0.1$ is immensely more luminous in the redder SDSS bands than it is in the bluer ones (about 2.5 *mag* between the $^{0.1}u$ band and the $^{0.1}z$ band). Our luminosity densities, obtained using the SWML estimate and shown in Table 2.5, are in agreement with those from Blanton et al. (2003b). Only in the $^{0.1}g$ band do discrepancies translate into a remarkable difference in the luminosity density. In this band, the nearby universe is ~ 0.8 mag brighter per unit volume, according to this work. In the rest of the bands, both estimates differ in ~ 0.1 mag $(\mathrm{Mpc}^3h^{-1})^{-1}$, i.e. less than 1%.

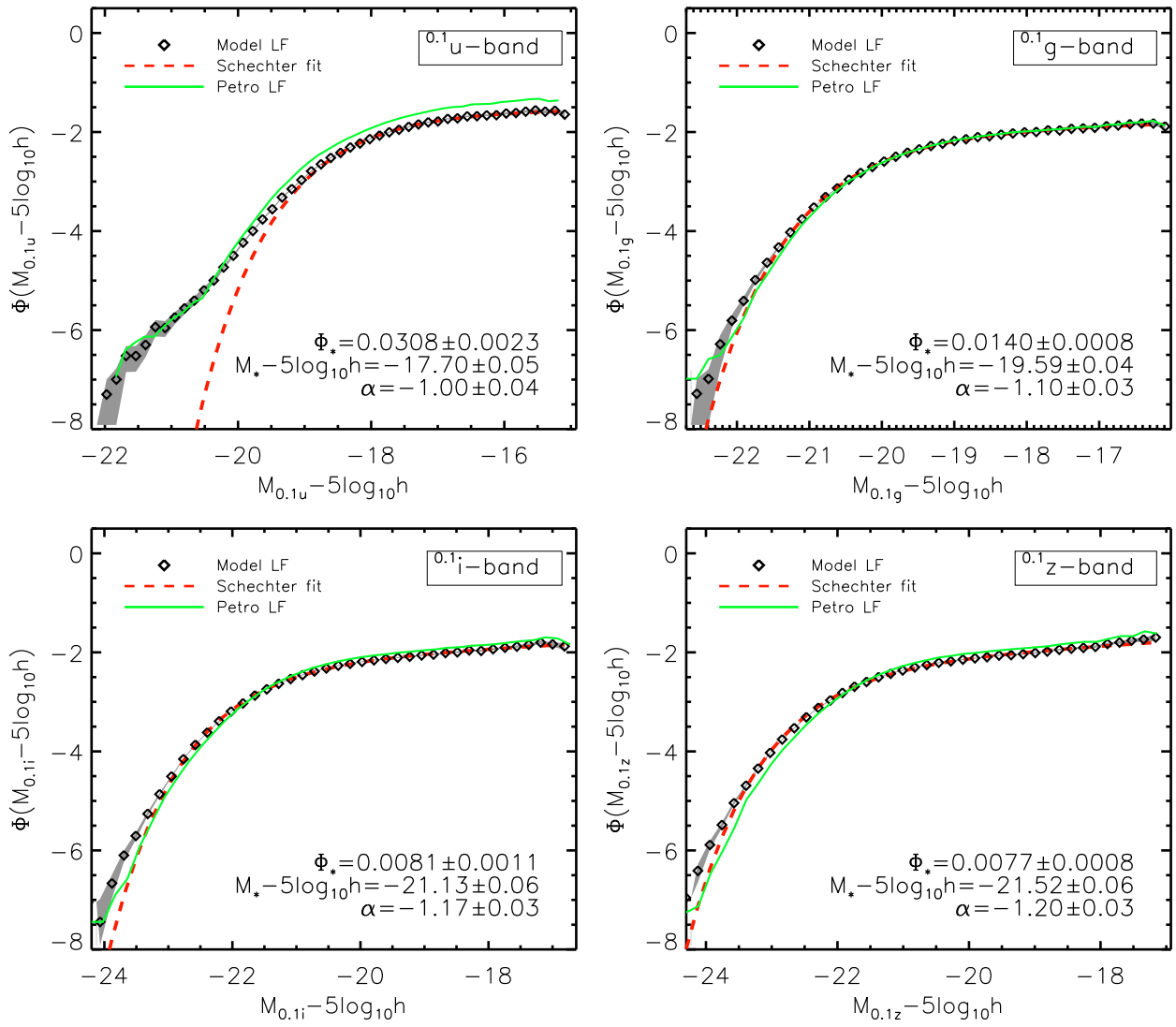

Figure 2.15 The SDSS DR6 LFs in bands $^{0.1}u$, $^{0.1}g$, $^{0.1}i$ and $^{0.1}z$ estimated using model magnitudes. SWML LF estimates are shown in diamonds and best-fitting Schechter functions are represented by dashed lines. In addition, we over-plot in each panel the SDSS DR6 LF estimated using Petrosian magnitudes. best-fitting values of Schechter parameters α, M_* and Φ_* are also shown. Shaded regions represent the 1σ uncertainty calculated using a bootstrapping technique.

Band	λ_{eff}(Å)	Absolute Magnitude Range	$\rho_{SWML} + 2.5log_{10}h$	$\rho_{fit} + 2.5log_{10}h$	$\rho_{B03} + 2.5log_{10}h$
$^{0.1}u$	3216	$-22.34 < M_{^{0.1}u} < -15.11$	-14.009 ± 0.014	-14.377	-14.10 ± 0.15
$^{0.1}g$	4240	$-24.09 < M_{^{0.1}g} < -16.01$	-14.386 ± 0.021	-14.783	-15.18 ± 0.03
$^{0.1}r$	5595	$-25.48 < M_{^{0.1}r} < -16.32$	-15.814 ± 0.015	-15.803	-15.90 ± 0.03
$^{0.1}i$	6792	$-26.02 < M_{^{0.1}i} < -16.72$	-16.355 ± 0.016	-16.138	-16.24 ± 0.03
$^{0.1}z$	8111	$-26.23 < M_{^{0.1}z} < -17.05$	-16.661 ± 0.018	-16.491	-16.56 ± 0.02

Table 2.5 Luminosity densities in AB mag $(\mathrm{Mpc}^3h^{-3})^{-1}$ for all SDSS photometric bands calculated using the SWML estimates of the luminosity functions of Figure 2.7 and Figure 2.8, ρ_{SWML}, and the best-fitting Schechter LF of Table 2.2, ρ_{fit}). Absolute magnitude ranges for the integration and the effective wavelength corresponding to each band, λ_{eff}, are also provided. Errors in ρ_{SWML} have been calculated using a bootstrapping technique. In addition, we give for comparison the luminosity densities from Blanton et al. (2003b), ρ_{B03} .

In Table 2.5, we also provide luminosity densities in each band obtained using the Schechter fits to the SWML estimates of the LFs presented in Section 3.2. As expected, the differences between these luminosity densities and those obtained with the SWML estimates are greater in the blue bands, where the Schechter LF provides a worse fit to our data points.

2.4 Discussion

The main results presented in this chapter are the SDSS DR6 luminosity functions of galaxies in the nearby universe ($z \sim 0.1$). A few years ago, Blanton et al. (2003b) used an early version of the SDSS (namely, the DR2) to calculate their widely-used SDSS LFs. The latest releases of the SDSS have provided, however, a huge increase in galaxy statistics. In particular, with the SDSS DR6, the size of our galaxy samples has increased by a factor ~ 7 in the very blue $^{0.1}u$ band and by a factor between ~ 3 and ~ 5 in the rest of the SDSS photometric bands, as compared to Blanton et al. (2003b) . In addition, we have achieved a high redshift completeness in our galaxy samples. Firstly, we have guaranteed that the effect of brightness-dependent redshift incompleteness is small within the magnitude ranges considered. Secondly, we have ensured high completeness and avoided contamination from galaxies with $r > 17.6$ by means of carefully chosen apparent magnitude faint limits. These advances make our SDSS DR6 LFs substantially more precise than those from Blanton et al. (2003b) at both the bright and the faint end. The LFs of Blanton et al. (2003b) seem to be compatible, however, with our results; although notable differences, which are surely physically significant, exist. Namely, at the bright end of the LF in the bluer bands (especially in the $^{0.1}u$ band) we find a remarkable excess, which was very noisy in Blanton et al. (2003b) due to their lack of statistics. At the faint end, we obtain steeper slopes in all SDSS bands, especially in the $^{0.1}u$-band (where the DR6 statistics allow us to go about $0.5 - 1$ mag deeper as compared to Blanton et al. 2003b) and, in less extent, in the $^{0.1}g$ band.

With regard to the faint end of the LF, a possible explanation for these discrepancies (apart from the consequence of the huge improvement in galaxy statistics) could come from differences in the sample selections. It is well known that the faint end of the LF in the blue bands is very sensitive to variations in the flux limit of the sample, due to the large dispersion in the the $(g-r)$ and, especially,

the $(u-r)$ color. In Blanton et al. (2003b), the flux limit is set at $u = 18.36$ and $g = 17.69$ in these blue bands, which means ~ 0.5 and ~ 0.2 mag brighter than in this work, respectively. This inconsistency could perfectly explain the discrepancies found in Figure 2.8, as moving this limit towards brighter magnitudes produce a flattening of the faint-end slope. Even slight variations in the calculation of the LF itself could cause noticeable differences in these bands. Similarly, these inconsistencies could be responsible for the much smaller deviations that we find in the redder bands, where the LF is proved to be much more robust. In this sense, it is worth mentioning that relatively small variations in the shape of the LF may translate into considerable changes in the values of the best-fitting Schechter parameters. It is not convenient, therefore, to make comparisons between different LFs by just looking at these best-fitting Schechter parameters. Finally, we cannot neglect completely the effect of evolution in our galaxy samples, but as we have shown in Section 2.3.3, we expect it to be small enough not to change our faint-end slope significantly, given the uncertainties we are dealing with.

We have also seen that the bright-end bump that we discovered in the bluer bands (and also in the ${}^{0.1}i$ band) is statistically very significant, according to our standard bootstrapping error analysis. Moreover, as we have mentioned above, we also find clear but noisy evidence of its existence in Blanton et al. (2003b). We have also checked that this excess is not a consequence of any of the limits that we have imposed to define our samples. It is neither an artefact originated by photometric errors in Petrosian magnitudes nor a consequence of neglecting galaxy evolution within our redshift intervals. This is, therefore, a remarkable result that may have strong implications for galaxy evolution.

From a preliminary analysis on the nature of the BEB in the ${}^{0.1}u$-band LF, we have discovered that it is mostly populated by star-forming (including starbursts) and active galaxies ($\sim 80\%$). The spectra of these galaxies is consistent with what we expect from QSOs/Seyferts I ($\sim 60\%$) and SF/SB galaxies ($\sim 20\%$). It seems, therefore, that an important fraction of the light that we receive from the brightest galaxies in the ${}^{0.1}u$-band would come from nuclear activity. Only about 20% of galaxies in the bright-end bump seem to be normal galaxies, showing no remarkable emission lines in their spectra.

Evidence of the existence of a similar bright-end bump as the one discussed above has been found in the semi-analytic models of galaxy formation by Croton et al. (2006). They report a bump at the bright end of the b_j-band LF of blue galaxies in their models (see their Figure 11). This excess is almost exclusively caused in the model by galaxies with very high SFRs (note that AGN are not implemented in these models) . The authors conclude that these "bump" galaxies are analogous to ultra-luminous infra-red galaxies (ULIRGs) and attribute the excess to an inadequate dust modeling within these objects. A more detailed study is needed not only to answer this question but to fully understand the origin and the nature of the BEB galaxy population.

We note that we have not applied any correction for the light attenuation produced by dust inside each galaxy. Interstellar dust is known to have a strong impact on the flux of an individual galaxy, especially in the bluer bands, depending on the galaxy's inclination, bulge-to-total ratio and wavelength of observation (see Driver et al. 2008 and Driver et al. 2007 for a realistic treatment of this issue). Correcting for this effect would therefore be necessary in any modeling that aims to constrain the stellar densities or the star formation histories of the universe.

In the next chapter, we present first results on a comparison between our SDSS DR6 luminosity

functions and the predictions from the semi-analytic models of galaxy formation and evolution of Bower et al. (2006). As mentioned before, SAMs are an excellent tool for investigating the physical processes and observational effects that shape the luminosity function of galaxies. These preliminary results will provide new insights on the main results presented in this chapter.

2.5 Chapter conclusions

In this chapter, we make use of the SDSS Sixth Data Release to estimate the number counts, luminosity functions and luminosity densities of galaxies in all SDSS photometric bands. The SDSS DR6 contains redshifts for $\sim 1,000,000$ galaxies down to magnitude $r \sim 17.77$, covering ~ 7400 deg^2 on the sky. The huge increase in the galaxy statistics with respect to the previous work of Blanton et al. (2003b) and the adequate treatment of brightness-dependent redshift incompleteness in our samples have allowed us to estimate the galaxy LFs of the nearby universe with unprecedented accuracy. In addition, we have calculated, for the first time, the galaxy number counts in all photometric bands using spectroscopic SDSS data. Luminosity densities in all SDSS bands have also been computed.

The main results of this work can be summarized as follows:

- The SDSS DR6 galaxy number counts in all SDSS photometric bands are consistent with an Euclidean, non-evolving Universe within a magnitude range that is limited by redshift incompleteness in the bright end and by the intrinsic apparent magnitude r-band limit of the survey at the faint end.

- The SDSS DR6 LFs of galaxies are compatible with those of Blanton et al. (2003b). However, notable differences exist. At the bright end of the $^{0.1}u$-band LF we find a remarkable excess, of ~ 1.7 dex at $M_{^{0.1}u} \simeq -20.5$ with respect to the best-fitting Schechter LF. This bright-end bump is very strong in the $^{0.1}u$ band and weakens in the $^{0.1}g$ band, fading away towards the very red $^{0.1}z$ band. At the faint end, we obtain steeper slopes in all bands, especially in the $^{0.1}u$ band and, in less extent, in the $^{0.1}g$ band. In addition, we can reach $0.5-1$ magnitude deeper without losing precision.

- We believe the discrepancies we find with respect to Blanton et al. (2003b) are primary due to the great increase in the sizes of the galaxy samples, of a factor ~ 7 in the $^{0.1}u$ band and between ~ 3 and ~ 5 in the rest of the SDSS bands, and to differences in the sample selection (especially in the lower limit in flux). We have also investigated the effect that accounting for galaxy evolution within our samples would have on our LFs. We have shown that, although it could produce slight adjustments to our LFs, it is unlikely to modify the main results presented in this work substantially.

- A preliminary analysis of the origin of the BEB seen in the $^{0.1}u$-band SDSS DR6 LF reveals that it is comprised of QSO and Seyferts I galaxies ($\sim 60\%$), star-forming galaxies and starbursts ($\sim 20\%$) and normal galaxies ($\sim 20\%$). It seems, therefore, that an important fraction of this exceeding luminosity might come from nuclear activity.

- We have estimated the SDSS DR6 LF using model magnitudes and found some clear differences with respect to the Petrosian LF. In the $^{0.1}u$ band, the faint-end slope of the LF flattens remarkably and, in the rest of the SDSS bands, the LF moves slightly to the bright end. These alterations are consistent with the scatter between model and Petrosian magnitudes in each band.

- The $^{0.1}r$-band SDSS DR6 LF of blue galaxies is consistent with a Schechter LF with a remarkably steep faint-end slope, corresponding to $\alpha = -1.41$. The $^{0.1}r$-band SDSS DR6 LF of red galaxies has, however, a slightly decreasing faint-end slope, corresponding to $\alpha = -0.81$.

- The SDSS DR6 luminosity densities of galaxies are in very good agreement with Blanton et al. (2003b) in bands $^{0.1}u$, $^{0.1}r$, $^{0.1}i$ and $^{0.1}z$. In the $^{0.1}g$ band, however, we find that the Universe in the redshift range considered is ~ 0.8 magnitude brighter per unit volume.

The state-of-the-art results presented in this chapter may be used to constrain a variety of aspects regarding star formation histories or feed-back processes in galaxy formation models. However, much effort is still needed in the survey field to fully understand the mechanisms that drive the evolution of galaxies in the Universe. This is especially necessary at high-z, where galaxy statistics are still very poor. In the next chapter, we will explore the implications of our results using semi-analytic models of galaxy formation and evolution.

3

Predictions from semi-analytic models of galaxy formation and evolution: luminosity function

In this chapter, we briefly introduce an ongoing follow-up project that aims to shed some light into the main processes governing the formation and evolution of galaxies, by comparing the theoretical predictions from semi-analytic models of galaxy formation (SAMs) with the observational constraints provided by the SDSS at low-redshift. To this purpose, we have worked in collaboration with Carlton Baugh and his group at the Institute for Computational Cosmology at Durham University (UK), who have contributed remarkably to the development of SAMs in the last decade. The first step in this study is to look into the average statistical properties of the low-z galaxy population. This chapter is organized as follows. In Section 3.1, we will briefly describe the simulations and the SAMs. In Section 3.2 we will present some preliminary results on the comparison between our SDSS estimates of the luminosity functions and the predictions from SAMs. Throughout this chapter, we assume a λ-CDM concordance model with $\Omega_m = 0.7$, $\Omega_\lambda = 0.3$ and $h = 1$.

3.1 The mock galaxy catalogs

The mock galaxy catalogs that we use in this study are based on the semi-analytic models described in Bower et al. (2006), which are implemented on the Millennium cosmological simulation (Springel et al., 2005). The Millennium Simulation, carried out by the Virgo consortium, followed the evolution of $\sim 10^{10}$ dark matter particles from redshift $z = 127$ in a cubic periodic volume of $500h^{-1}$ Mpc on a side. Note that the collapse of density fluctuations and the subsequent hierarchical formation of large-scale structure in the Universe is a highly non-linear process which can only be approached by means of N-body numerical simulations. In the Millennium Simulation, approximately 20 million halos with more than 20 particles of mass $8.6 \times 10^8 h^{-2} M_\odot$ were formed. For further details see (Springel et al., 2005).

Bower et al. (2006) use the GALFORM semi-analytic models (Cole et al., 2000; Benson et al., 2003), which are based on the pioneering work of White & Frenk (1991). GALFORM employs a Monte-Carlo algorithm to follow the evolution of dark matter particles and generate merger trees in

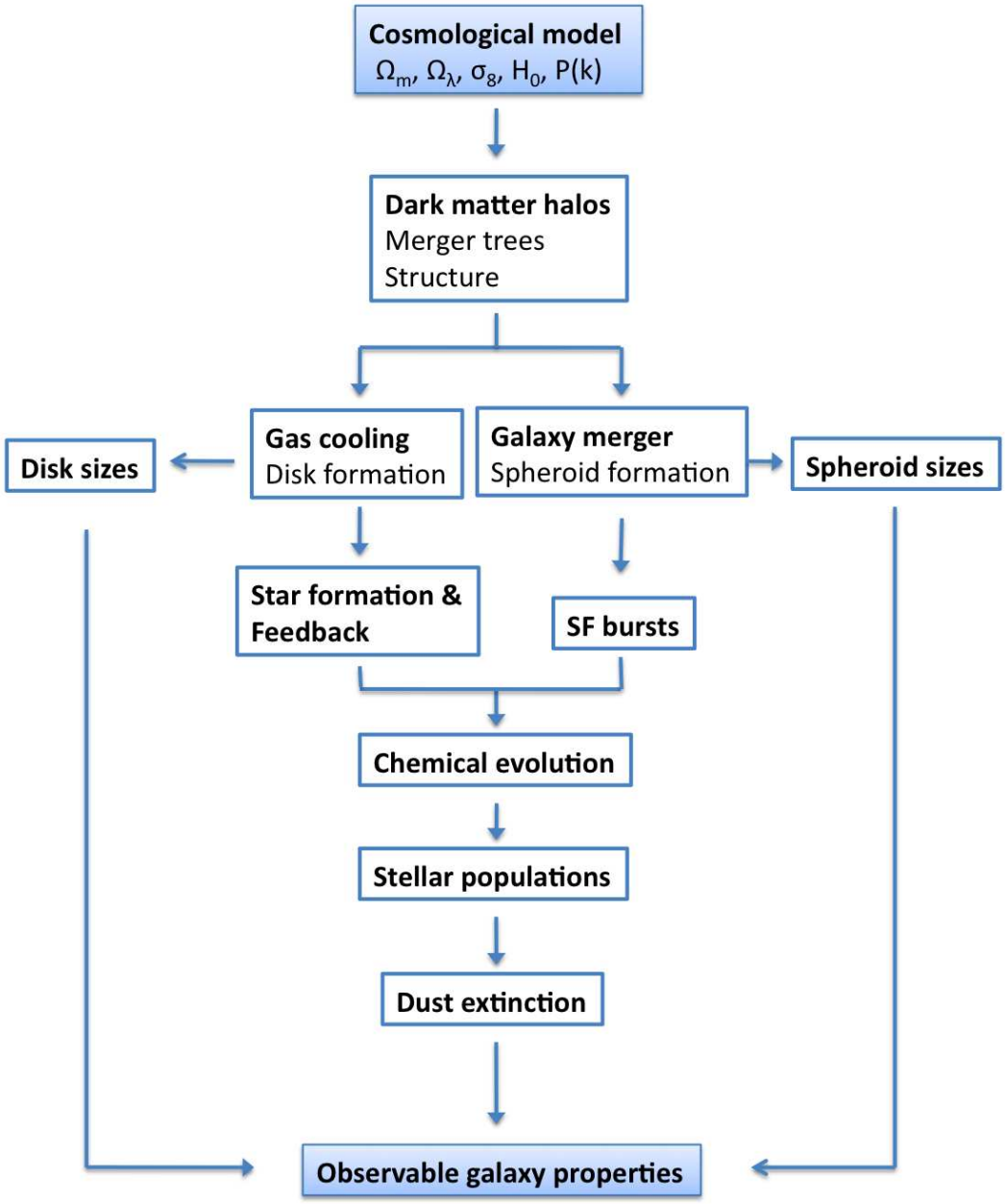

Figure 3.1 A scheme taken from Cole et al. (2000) showing the main processes modeled by GALFORM. The starting point is a λ-CDM cosmogony, which is characterized by a set of cosmological parameters. Here, Ω_m is the matter density, Ω_λ is the dark energy density, σ_8 is the amplitude of the density fluctuations at 8 $Mpc\ h^{-1}$, H_0 is the Hubble constant and $P(k)$ is the power spectrum of the density fluctuations. The different modules combine to produce predictions for observable galaxy properties.

the Millennium Simulation, which describe the hierarchical formation of dark matter halos. Within this scheme, galaxy formation is assumed to occur gravitationally in a process where hot baryons cool and condensate to form stars (White & Rees, 1978). GALFORM uses a variety of analytical recipes and techniques to model complex processes such as the radiative cooling of gas to form rotationally supported galactic disks, the star formation in galactic disks, the supernova feedback, the chemical enrichment of the ISM, the formation of galactic spheroids through mergers, the merger-induced star formation bursts, the spectroscopic evolution of stellar populations or the effect of dust on galaxy luminosities and colors. In Figure 3.1, we reproduce a scheme from Cole et al. (2000) showing the main modules into which GALFORM is divided. Here, the starting point is a λ-CDM cosmogony, which is characterized by a set of cosmological parameters. The different prescriptions then combine to produce predictions for observable galaxy properties. In the SAMs of Bower et al. (2006) the GALFORM models are extended mainly to track the formation and evolution of AGN. Supermassive black holes at the centers of galaxies are assumed to produce the quenching of star formation by reheating the cold gas and, consequently, suppressing the cooling flows. This process, that was first implemented in SAMs by Croton et al. (2006) as an AGN "radio mode", is thought to be responsible for, among other observational evidences, the sharp cut-off at the bright end of the luminosity function, or the fact that very massive galaxies in the Universe are red and passive-evolving instead of young, blue and star-forming. For a complete description of the SAMs used in this chapter see (Cole et al., 2000) and Bower et al. (2006).

3.2 The Luminosity function of galaxies

We report here on a preliminary comparison between our SDSS DR6 LFs and the theoretical predictions from the Bower et al. (2006) models. The theoretical LFs have been calculated without imposing any observational limits, other than a redshift range. This means that we have not mimicked the SDSS apparent magnitude cut ($r < 17.77$), so we take into account all objects above the resolution limit of the simulation. In Figure 3.2, we show the ${}^{0.1}u$-band luminosity function of galaxies obtained with the mock catalogs in 4 different redshift slices, at $z = 0$, $z = 0.05$, $z = 0.10$ and $z = 0.20$ (this figure was provided by Carlton Baugh and his team at Durham). In addition, our SDSS DR6 LF (quoted as Montero-Dorta & Prada 2009 in the legend) is plotted in dots and the Blanton et al. (2003b) LF is represented by a solid line. The following conclusions can be drawn from this first comparison:

- At the faint end, we find a remarkable discrepancy between our observational constraints and the predictions from the SAMs, which are in better agreement with Blanton et al. (2003b). The LF predicted by the models falls below the LF of both Blanton et al. (2003b) and this work at ${M_{0.1}}_u > -18$. In addition, the mock LF seems to present a steeper faint-end slope than that of Blanton et al. (2003b).

- The Bower et al. (2006) models predict a extremely slight evolution (if any) of the LF within the redshift limits considered. This seems to support the argument made in Chapter 2, that an evolution correction should be negligible within our SDSS sample.

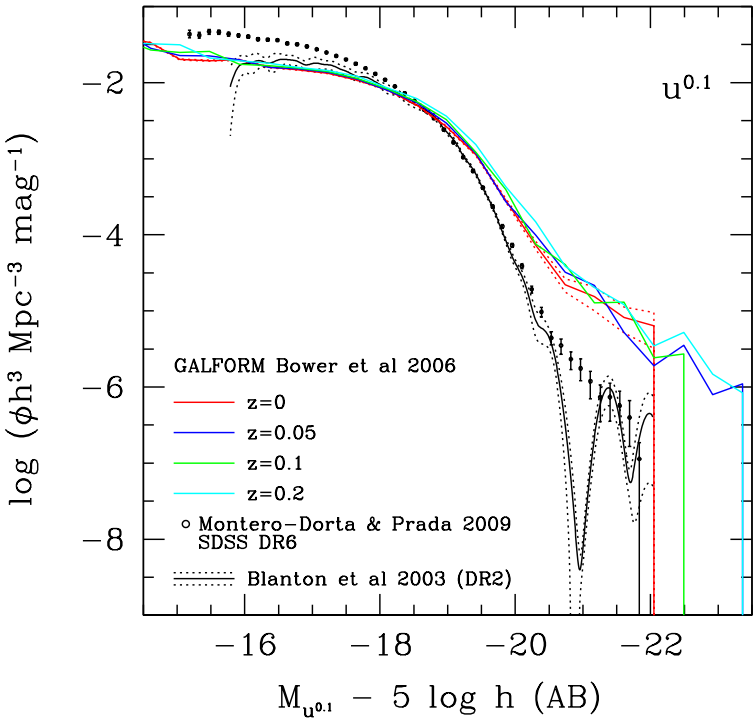

Figure 3.2 The $^{0.1}u$-band luminosity function of galaxies obtained with the Bower et al. (2006) semi-analytic models in four redshift slices at $z = 0$ (red), $z = 0.05$ (blue), $z = 0.10$ (green) and $z = 0.20$ (cyan). For comparison, the SDSS LF presented in Chapter 2 (Montero-Dorta & Prada, 2009) and that from Blanton et al. (2003b) are shown in dots and a solid line, respectively. The 1σ error region for the latter is represented by a dotted line.

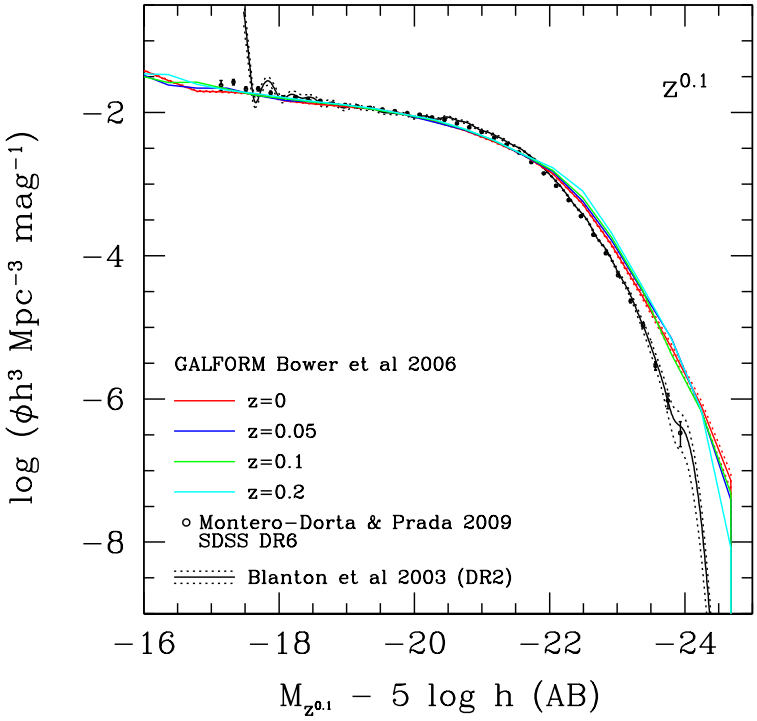

Figure 3.3 The same as in Figure 3.2 for the $^{0.1}z$ band.

- The AGN feedback implemented in the models is not enough to reproduce the exponential cut-off at the bright end of the LF. However, we find an encouraging result. The shape of the mock LF at the bright end resembles that of our SDSS DR6 LF, presenting a similar bump, although remarkably more pronounced than the BEB.

Similar results are found when comparing the SDSS and the mock LFs in the $^{0.1}g$ band. Namely, a discrepancy in the normalization between our estimation and that from Blanton et al. (2003b) and the models, a steeper faint-end slope in the models that is consistent with our findings and a remarkable bump at the bright end, which has however decayed considerably as compared to the $^{0.1}u$-band LF (this is also in agreement with our work). In the redder bands, the agreement between observations and theoretical predictions is better, as expected, although remarkable discrepancies remain. As an example, in Figure 3.3 we show, in the same format as in Figure 3.2, the results from this comparison in the $^{0.1}z$ band. No evolution is found in the mock LF in this band. In addition, the models predict a steeper faint-end slope, which is in slightly better agreement with our work than with Blanton et al. (2003b).

The results presented here seem to support the argument that an evolution correction is unlikely to produce a significant flattening of the faint-end slopes of our SDSS LFs. In order to investigate this in detail, the next step is to reproduce the observational effects in the mocks, as so far we are considering the entire box. This, in fact, poses an interesting question regarding the real underlying

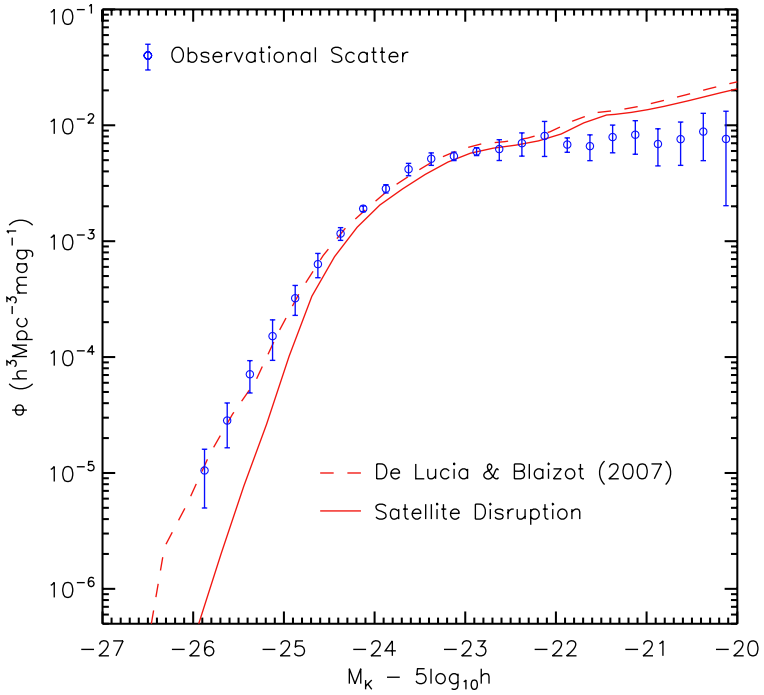

Figure 3.4 The tidal stripping of stars from satellite galaxies during mergers is one of the physical processes that might explain the bright-end excess reported in the model LF as compared with the observations. This figure, which was taken from Henriques & Thomas (2010), shows the effect of tidal disruption on the NIR K-band LF at $z = 0$. The dashed line represents the LF computed in the semi-analytic models of De Lucia & Blaizot (2007) and the solid line shows the LF obtained with the same models with the addition of stellar stripping. Finally, the data points represent the mean of a combined set of observations from Cole et al. (2001), Bell et al. (2003) and Jones et al. (2006).

LF in the Universe, as Figure 3.2 and Figure 3.3 suggest that its shape (mainly the faint-end slope) might remain constant if galaxies of any luminosity would be taken into account. Several groups have found evidence, however, of remarkably steeper faint-end slopes in the field LF obtained in deeper galaxy samples (see an example in Blanton et al. 2005b with the SDSS and Wolf et al. 2003 with the COMBO-17 survey). We will return to this topic in Chapter 4, where we present first results on the LF in the very deep ALHAMBRA photometric survey.

Especial attention deserves the discrepancy found in the bluest SDSS bands at the faint end. As we discussed in the previous chapter, the shape of the LF is very sensitive to the apparent magnitude limits, especially in the faint end, due to the large color dispersion and the r-band limit of the survey ($r \leq 17.77$). We have also found that the shape of the LF is even sensitive to the method used in the calculation. In order to explore the nature of these deviations, we plan to extract comparable samples from the mock catalogs, trying to reproduce all observational effects.

As for the semi-analytic models, the biggest challenge is still to discover the mechanisms that cause the sharp cut-off in the luminosity function at the bright end and prevent galaxies more massive than $\sim 10^{12}\ M_{\odot}h^{-2}$ to grow. In the Bower et al. (2006) SAMs that we use here, AGN feedback is set to suppress cooling flows completely, preventing hot gas to cool in the host halos of massive galaxies.

This feedback can certainly produce a cut-off in the luminosity function, but is still insufficient to fully reproduce the observations, as Figure 3.2 and Figure 3.3 show.

Several physical processes must be explored in more detail in order to unveil the nature of this bright-end discrepancy. In particular, the fact that cooling is switched off in massive galaxies in the models might suggest that the excess is due to an inaccurate treatment of galaxy mergers. In current SAMs, red passively-evolving elliptical galaxies are assumed to form through a series of mergers. In most of these events, dynamical friction plays a crucial role. Essentially, during the merger of two dark matter halos, galaxies in the less massive halo become the satellite galaxies of the more massive one. These satellite galaxies progressively lose their energy and angular momentum, which are transmitted to dark matter particles. Satellite galaxies are decelerated and eventually sink to the center of the more massive dark matter halo, if they are not disrupted by tidal forces. The Bower et al. (2006) models used a simple prescription for the merger timescale of satellite galaxies, i.e. the so-called Chandrasekhar's formula, which is based on the early work of Chandrasekhar (1943), who discussed dynamical friction for an idealized case where a rigid object moves through a field of collisionless matter particles (see Binney, J. and Tremaine, S. 1987; Lacey & Cole 1993 for a generalization with dark matter particles). However, other more accurate prescriptions for the merger timescales have been proposed, which might have strong implications in the formation and evolution of massive elliptical galaxies.

Another physical process that could explain the excess of bright galaxies in the SAMs is the tidal stripping of stellar material from satellite galaxies during mergers. In essence, tidal forces can sweep stars and cool gas from in-falling satellite galaxies during mergers, and inject this material into the inter-cluster medium. This process would not only alter the timescale for mergers but also reduce the amount of mass which is accreted by the central galaxy. The effect of satellite disruption on the luminosity function of galaxies is illustrated in Figure 3.4, which was taken from Henriques & Thomas (2010) . In this figure, the near-infrarred K-band LF at $z = 0$ computed in the semi-analytic models of De Lucia & Blaizot (2007) is shown in a dashed line. The solid line represents the LF obtained with the same models with the addition of stellar stripping. Finally, the data points represent the mean of a combined set of observations from Cole et al. (2001), Bell et al. (2003) and Jones et al. (2006). Note that the switch-on of stellar stripping makes a strong impact on the bright-end of the local luminosity function.

Part II

Galaxy evolution from $z \sim 1$

4

A Low-Redshift Calibration of the ALHAMBRA Survey

4.1 Introduction

In order to identify the main physical processes that drive the evolution of galaxies, an accurate and unbiased characterization of the galaxy population at different redshifts is required. At $z \sim 0$, large-scale galaxy redshift surveys like the SDSS or the 2dFGRS have made such a statistical characterization possible. We are, in fact, at a stage where we can fine-tune our measurements on the luminosity function, the stellar mass function, the galaxy clustering, the correlations between galaxy properties or the influence of environment. Unfortunately, the Universe is remarkably fainter at $z \sim 1$, and there is no equivalent to such huge surveys available yet. Galaxy redshift surveys like the DEEP2 GRS, VVDS, COMBO-17 or COSMOS have, however, allowed us to explore these epochs and provide observational constraints to galaxy formation and evolution models.

The ALHAMBRA-Survey is an ongoing project lead at the Instituto de Astrofísica de Andalucía that aims to alleviate the lack of quality data at high redshift. ALHAMBRA is specifically designed for the study of the evolution of the contents of the Universe from $z \sim 1$. To this purpose, it utilizes the 3.5-meter Telescope at Calar Alto Observatory to cover a region of 4 deg^2 on the sky, providing photometric data for $\sim 500,000$ galaxies down to magnitude $I(AB) \sim 25$. In order to account for cosmic variance, the ALHAMBRA covered area is divided into 8 different fields, located at different pointings. ALHAMBRA uses 20 contiguous, equal-width, medium-band filters covering from 3500 $\mathring{A}$ to 9700 $\mathring{A}$, plus the standard JHK near-infrared bands. Photometric redshifts are estimated using the Bayesian Photo-Z method (Benítez, 2000) and a precision of at least dz/z=0.015 for $I(AB) < 24$ is expected. At the present time, most of the observations have been carried out and data reduction and photo-z estimation have reached to a state where we can commence the scientific exploitation of the survey. Detailed information about the survey design can be found in Moles et al. (2008).

Galaxy redshift surveys can be classified into two groups, according to the techniques involved in measuring redshifts. Spectroscopic surveys (e.g. SDSS, 2dFGRS, DEEP2 GRS), which use spectroscopy, provide precise SED determination but have expensive requirements for telescope time.

This usually imposes important restrictions in terms of photometric depth. Photometric surveys (e.g. COMBO-17, COSMOS or ALHAMBRA), by contrast, use photometry to obtain a less accurate but faster (in terms of telescope time) estimation of redshift and SED. While they cannot achieve the same resolution to that provided by spectroscopic surveys, they can get considerably deeper, allowing the study of the fainter galaxies. The main potential of the ALHAMBRA-Survey resides in the combination of a very good SED determination for a photometric survey (which is guaranteed by its photometric system) and a very deep photometry. This combination makes ALHAMBRA an excellent tool for shedding light into the galaxy evolution picture and exploring the high-redshift galaxy population. In particular, ALHAMBRA will provide the most accurate observational constraints to date on fundamental properties of the high-z galaxy population, such as the luminosity function and the stellar mass function. In addition, the relatively large area covered by ALHAMBRA will allow for the study of the evolution of the large-scale structure and, in particular, the number and content of galaxy clusters at different redshifts. The unbiased nature of the survey will also help characterize many different kinds of objects, ranging from emission-line galaxies to the diverse types of AGNs and stars in our own Galaxy.

This chapter is devoted to analyzing some fundamental statistical properties of the ALHAMBRA population of galaxies. To this end, the ALHAMBRA data will be compared with other surveys, such as the SDSS, COMBO-17 or COSMOS. Throughout this chapter, we will restrict ourselves to a low redshift sample ($0.1 < z < 0.3$). The idea is to perform a sort of calibration of the survey at low redshift, where our knowledge on the galaxy population is remarkably more consolidated. It is important to stress that, due to the early state of the ALHAMBRA data, this study must be regarded as a preliminary analysis, aimed to lay the foundations for the study of galaxy evolution with ALHAMBRA. The chapter is organized as follows. In Section 4.2, we discuss on the photo-z estimation in ALHAMBRA using the BPZ method. In Section 4.3, we describe our sample selection, including the method used for separating galaxies from stars and artificial detections. In Section 4.4, we present the ALHAMBRA galaxy number counts. Section 4.5 is devoted to describing the calculation of K-corrected ALHAMBRA absolute magnitudes. In Section 4.6, we discuss the ALHAMBRA color distributions. In Section 4.7, the ALHAMBRA low-redshift luminosity functions of galaxies is presented and compared with previous works from other surveys. Finally, in Section 4.8, we summarize the main results of this chapter.

4.2 Photometric redshifts

We start by showing in Figure 4.1 the ALHAMBRA filter system, which is composed of 20 contiguous, top-hat filters with full width half maximum $FWHM = 310$ $\mathring{A}$, covering the entire optical range from 3500 to 9700$\mathring{A}$. The choice of a filter system is an important aspect in the design of a photometric survey as it basically determines not only the photometric depth but also the photometric redshift accuracy. For a given wavelength coverage, the use of broader filters provide more photometric depth but photometric redshift accuracy decreases due to color-redshift degeneracies. These degeneracies can be alleviated by using narrow filters at the cost of losing photometric depth. The ALHAMBRA filter system is by design an optimal compromise between photometric depth and photometric redshift

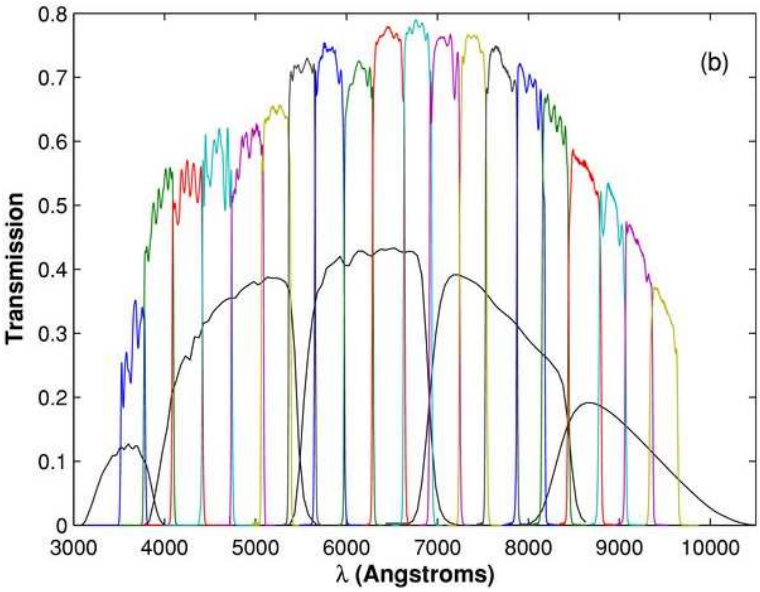

Figure 4.1 The ALHAMBRA system in the optical range is comprised by 20 contiguous, equal-width filters extending from 3500 to 9700 $\mathring{A}$. This figure was directly taken from Aparicio Villegas et al. (2010). Transmissions have been estimated by taking into account atmospheric extintion, mirror reflectivity and CCD quantum efficiency. Overplotted are the SDSS *ugriz* bands. Note that the u and, in less extent, the z bands are only partially covered by the ALHAMBRA photometric system.

accuracy, both being necessary requirements for the science case persued. A complete justification of the convenience of this system can be found in Benítez et al. (2009). In this work, the authors conclude that a system with 20 constant-width, non-overlapping filters like that of ALHAMBRA can reach only 0.1 mag shallower than a system with 4-5 broader filters. Redshift accuracy, however, increases by as much as a factor 3. We refer the reader to Benítez et al. (2009) for further information.

The set of filters shown in Figure 4.1 provides, as explained above, many advantages when it comes to estimating photometric redshifts. The photo-z's used in this work were obtained using the BPZ method. Let us briefly outline how this method, which is based on Bayesian inference, works. Within the framework of Bayesian probability, the problem of photometric redshift estimation can be posed as finding the probability of a galaxy having redshift z given the data, and the so-called prior information, which includes any knowledge relevant to the hypothesis under consideration ideally not already contained in the data. The basic priors on which the BPZ method relies are the redshift distributions per galaxy type and the galaxy clustering in the redshift range covered by the survey. In Benítez (2000), the author explains how these priors could even be extracted from the data sample in which the photo-z's are obtained. An important advantage of Bayesian statistics is that the accuracy of the redshift estimation can be characterized in a way that has no equivalent in other statistical approaches, the so-called $ODDS$ parameter. In a few words, the $ODDS$ parameter is a measure of the width of the photo-z probability distribution for each galaxy. Narrower distributions mean more certain photo-z determinations and consequently higher values for the $ODDS$ parameter (which ranges from 0 to 1). The $ODDS$ parameter allow us to select galaxy samples with reliable photo-z's.

In this way, the properties of individual galaxies can be determined more accurately and, at the same time, we can estimate the statistical properties of an entire galaxy sample in an optimal fashion. For more information on the BPZ technique, see Benítez (2000).

The photo-z accuracy that we expect for ALHAMBRA is that of a state-of-the-art photometric survey. This is not only guaranteed by the optimal optical coverage, but also by the additional help provided by the NIR bands. The importance of the NIR bands in the photo-z estimation resides in the fact that they help breaking the degeneracy between low-redshift and high-redshift galaxies that comes from misidentifying the Balmer and the Lyman breaks. From simulations, Moles et al. (2008) anticipated a typical photometric redshift accuracy of $\delta z/(1+z) \leq 0.015$ for objects with $I(AB) \leq 24$. Even for very faint objects, down to $I(AB) \sim 25$, few objects are expected to have photo-z errors greater than 3%, according to Moles et al. (2008).

In this work, we use a preliminary version of the ALHAMBRA photo-z catalogs. According to the collaborators in charge of the photo-z estimation, the average photo-z uncertainty in these catalogs is $\delta z/(1+z) \leq 0.015$. However, photo-z errors in some magnitude ranges (especially for very faint objects, i.e. $I_{AB} > 24$), are expectedly larger than the above predictions. In addition, the photo-z accuracy might reportedly be lower at very low redshifts, due to the fact that the BPZ has been calibrated mainly with higher-redshift priors. We are aware of these uncertainties, so we will be very cautious when discussing the results presented in this chapter. Our results, however, seem to be consistent with our knowledge of the nearby Universe, as we will show below. This first low-redshift calibration of the ALHAMBRA data is a necessary step that confirms the potential of the survey to shed light into the evolution of galaxies from $z \sim 1$.

4.3 Sample selection

ALHAMBRA comprises 8 rectangular fields, each of them covering an area of $\sim 1 \times 0.25$ deg^2 on the sky. That makes a total surveyed area of ~ 4 deg^2. Fields are located at different pointings in order to account for cosmic variance and overlap with surveys like the SDSS or the DEEP2 RS to facilitate the photo-z calibration. Because of the characteristics of the detectors, each field is divided into 4 sub-fields. Approximately 25% of the total area on the sky surveyed by ALHAMBRA is used in this work. That makes a total of ~ 1 deg^2 (1.06 deg^2, 17 sub-fields), which is a representative fraction of the data.

The ALHAMBRA photometric catalogs provide, for each object, fluxes in the 20 optical filters and the J,H and K bands in the NIR. This massive amount of photometric information offers new ways of statistically investigating the evolution of the galaxy population, by means of a number of color-magnitude combinations. The color information provided by ALHAMBRA will surely help constrain current models of galaxy formation and evolution, especially at high redshift. However, and before we can take advantage of all ALHAMBRA's potential, we need to find a way to make the data comparable to other well-known surveys. To this end, we have constructed synthetic SDSS magnitudes from the ALHAMBRA narrow-filter optical magnitudes. Note that both the SDSS u and z bands are not well covered by the ALHAMBRA system (see Figure 4.1). For this reason, we will only make use of the synthetic g, r, i bands; hereafter quoted as g^A, r^A and i^A.

Arguably the most problematic step in the selection of an adequate galaxy sample in ALHAMBRA is the identification of galaxies among non-galactic sources, the so-called star/galaxy separation (S/G separation). In the next section, we describe the method that we followed to complete this critical task.

4.3.1 Star/Galaxy separation

Object detection in each of the ALHAMBRA fields is performed in a deep synthetic image that we call DEEP, which is a combination of a certain set of filter images with the highest signal-to-noise. From this DEEP image all type of sources above a certain brightness threshold are extracted using SExtractor. Consequently, raw catalogs include not only galaxies, our desired targets, but also stars, saturated pixels and other artifacts that we need to remove. Naively, cleaning these catalogs might seem straightforward, but in practice, it turns out to be a rather tricky task. Most problems come with very faint detections. In some cases, literally, there is no real way to distinguish an ultra faint galaxy from a star or any spurious object. This must be carefully treated in ALHAMBRA, where we expect to reach as deep as $i^A \sim 25$. At the bright end, as we will show below, the process of substracting stars is less uncertain. At these ranges (especially at the very bright end) the S/G separation has to be treated with the same care, as missidentified objects can have a strong impact in ranges where statistics are poor.

In Cristóbal-Hornillos et al. (2009), a succesful S/G separation is performed in the ALHAMBRA NIR bands, based on the work of Huang et al. (1997). The authors used a $g^A - r^A$ versus $J - K_s$ diagram to identify stars, instead of the $B - I$ versus $I - K$ diagram proposed by Huang et al. (1997). Unfortunately, our interests in this chapter are focused on the optical range, making the above method unfeasible. Using the same S/G separation criteria as those from Cristóbal-Hornillos et al. (2009) would introduce a strong bias in our sample, as we can reach considerably deeper in the optical range than we do in the NIR. A modification of the $g^A - r^A$ vs. $J - K_s$ diagram to include only optical magnitudes (e.g. $g^A - r^A$ vs. $r^A - z^A$) would also prove useless, as the power of this method precisely resides in the IR color, tracing the presence of dust and AGB stars in galaxies.

Motivated by Leauthaud et al. (2007), we will rely on a purely morphological method that uses only SExtractor parameters. SExtractor provides a stellar identification index called CLASS_STAR, that ranges from 0 for an extended source to 1 for a point source. The general consensus, however, is that CLASS_STAR alone is not enough to perform an efficient S/G separation (Cristóbal-Hornillos et al., 2009; Gardner, 1995; Huang et al., 1997; Leauthaud et al., 2007). As expected, the ambiguity of this parameter stands out at the faint end. Figure 4.2 shows three possible methods to clean our catalog using SExtractor parameters, including CLASS_STAR. In all three panels, crosses (in any color) represent detections in a randomly selected ALHAMBRA field (Field 8, CCD 21). Note that we use the DEEP image to perform the S/G separation, and consequently a DEEP magnitude ($m_{\rm DEEP}$) that has not been properly calibrated (here we have added a constant 30 to each magnitude arbitrarily). This implies that magnitudes are not fully comparable between different fields and, consequently, fields have to be treated independently. Just to provide a reference, in Figure 4.3 we plot $m_{\rm DEEP}$ versus the synthetic i^A band.

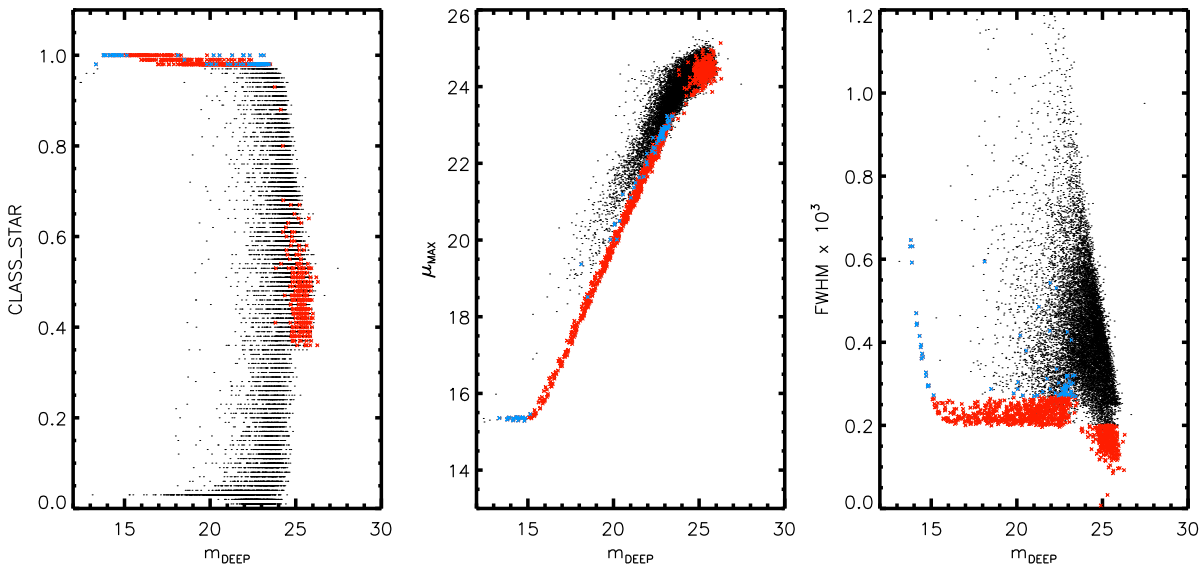

Figure 4.2 Three SExtractor parameters are shown as a function of magnitude in the synthetic detection image DEEP for a given ALHAMBRA field: CLAS_STAR (left), μ_{MAX} (middle) and FWHM (right). Red points represent non-galactic sources (stars and artifacts) and blue points are uncertain classifications which have been discarded in order not to contaminate our samples. The CLASS_STAR parameter alone is not enough to perform an efficient S/G separation. The μ_{MAX} vs m_{DEEP} diagram is a much accurate method to classify objects, although it becomes ambiguous at faint magnitudes, where point sources are indistinguishable from very small galaxies. We propose a combination of the CLASS_STAR parameter and a $FWHM$ vs m_{DEEP} diagram as described in Section 4.3.1.

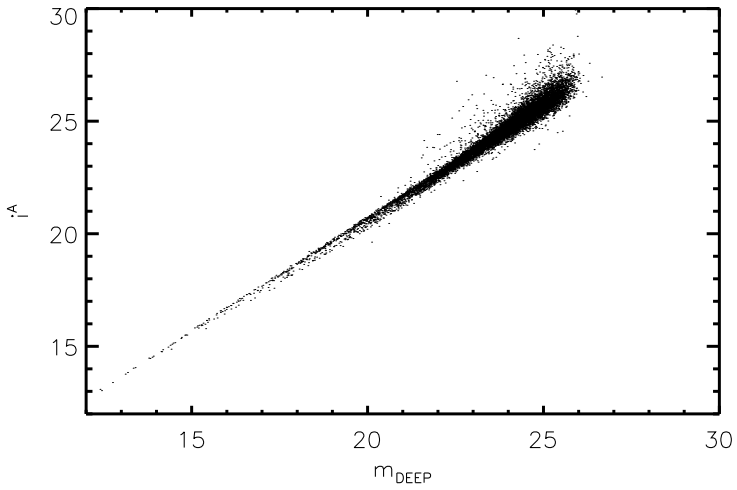

Figure 4.3 The synthetic ALHAMBRA i^A band plotted as a function of magnitude in the synthetic detection image DEEP, for a given ALHAMBRA field.

In the left-hand plot of Figure 4.2, the CLASS_STAR parameter is shown as a function of $m_{\rm DEEP}$. Stars, as point sources, are expected to occupy the very top part of this diagram. However, as we move to fainter magnitudes, setting a threshold turns out to be rather arbitrary. In the middle plot, we illustrate the method proposed by Leauthaud et al.(2007), based on a μ_{MAX} vs. m_{DEEP} diagram. The μ_{MAX} parameter represents the peak surface brightness above the background level, which is proportional to magnitude for point sources. This is definitely a more efficient method, as stars form a well-defined line in the diagram. Again, problems arise at faint magnitudes, where point sources become indistinguishable from very small galaxies. Even at brighter magnitudes, as Figure 4.2 shows, it turns out to be hard to perform the classification in the region where the line of stars enters the spot where the majority of galaxies are located.

Instead of the Leauthaud et al. (2007) diagram we propose a FWHM vs. $m_{\rm DEEP}$ diagram, as that shown in the right-hand panel of Figure 4.2, combined with the CLASS_STAR parameter. The motivation behind this method is simply to take advantage of the fact that the FWHM of stars is typically independent of magnitude and smaller than that of galaxies. Consequently, in our FWHM vs. m_{DEEP} plot, stars lie in a well-defined horizontal band while galaxies basically spread along a vertical, slightly tilted, branch. In order to separate stars (or non-galactic sources) from galaxies we proceed as follows:

- We set a threshold in CLASS_STAR so that the entire band of stars is covered, up to the gap between this band and the galaxy cloud (see the right-hand panel of Figure 4.2). In this particular field this natural demarcation occurs at $m_{\rm DEEP} \sim 23.5$. Note that $m_{\rm DEEP} \sim 23.5$ corresponds to $i^A \sim 23.5/24$ (see Figure 4.3).

- At fainter magnitudes ($m_{\rm DEEP} > 24$) we cannot rely on CLASS_STAR to identify non-galactic sources. We expect, however, the large majority of detections to be galaxies. We exclude objects with a extremely small FWHM, even smaller than that of stars, under the assumption that these are spurious, unreal detections.

In Figure 4.2 both red and blue objects are identified as stars using the above criteria. Note that all these sources below $m_{\rm DEEP} \sim 23.5$ sit on the line of stars in the μ_{MAX} vs. $m_{\rm DEEP}$ diagram and have high values of CLASS_STAR. Here we have highlighted in blue those objects which satisfy the above criteria (and actually lie in the line of stars in the middle plot) but have a relatively large FWHM, comparable to those of galaxies. At very bright magnitudes these objects are clearly artifacts (Leauthaud et al., 2007), but at faint magnitudes their classification is dubious. These objects make up $\sim 1.5\%$ of all objects around $m_{DEEP} \sim 23.5$ in our example and, in the worst-case scenario, this fraction can rise up to $\sim 6\%$ in other fields. This is still a small fraction of the entire galaxy sample so we have decided to discard them in order not to contaminate our catalogs. In future works, we will examine this type of sources in detail in order to confirm this preliminary classification.

At fainter magnitudes ($m_{\rm DEEP} > 24$), objects classified as non-galactic sources have a very low $FWHM$. These objects are located in the regions where galaxies lie in the μ_{MAX} vs. $m_{\rm DEEP}$ diagram and are considered ambiguous by CLASS_STAR. There is finally a group of very faint objects with a similar FWHM as that of stars. In these ranges of magnitude, ultra-faint galaxies can perfectly have

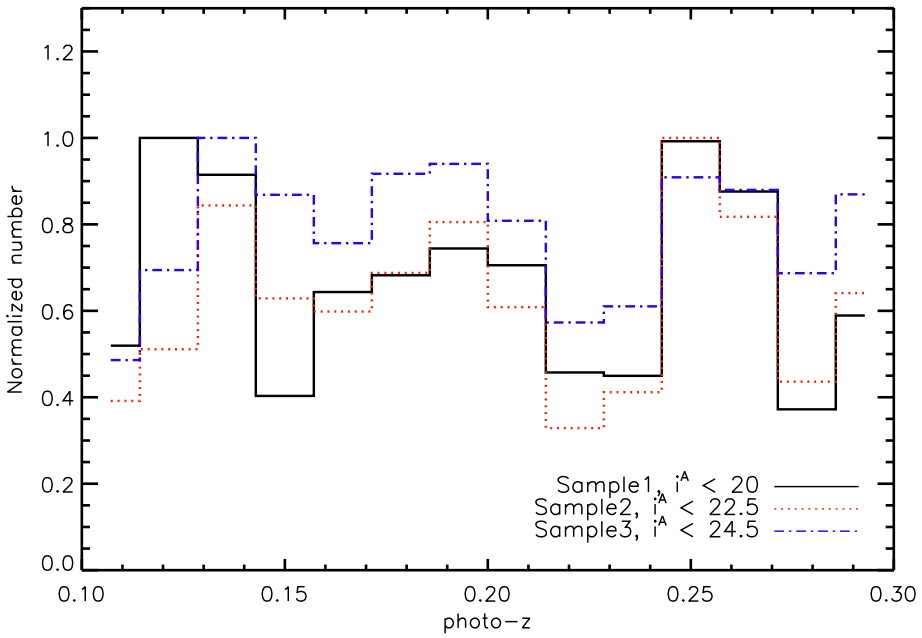

Figure 4.4 Normalized redshift distribution in Sample1 ($i^A < 20$), Sample2 ($i^A < 22.5$) and Sample3 ($i^A < 24.5$).

the FWHM of a point-like source. For this reason, we have classified these objects, otherwise very rare, as galaxies. We would like to stress that we are moving here in a very uncertain region, where a precise classification is hardly possible. Also note that, in any case, these objects have magnitudes $i^A > 25$ and, therefore, lie outside the range where the photo-z estimation is reliable.

4.3.2 Sample limits

After applying the S/G separation method described in the previous section, only some minor adjustments are needed in order to create our main galaxy catalog (hereafter, MC), from which all our galaxy samples are drawn. These adjustments, include selecting objects with a good photo-z estimation (imposing $ODDS > 0.99$) and excluding areas in each field where signal-to-noise is very low. Finally, the MC is comprised by 220,772 galaxies. Note that we expect that the fraction of non-galactic sources in the MC is smaller than $\sim 1\%$. Once we have made sure that the fraction of stars and spurious artifacts is negligible in the MC, we proceed with the selection of some adequate low-redshift galaxy samples, in order to characterize de ALHAMBRA galaxy population.

Firstly, we select objects in the redshift range $0.10 < z < 0.30$. The main motivation for this selection is to allow for a relatively fair comparison with the SDSS, which basically covers a redshift range $0.02 \lesssim z \lesssim 0.25$, with the redshift distribution peaking somewhere around $z = 0.1$. Due to the small volume sampled by ALHAMBRA at low redshift, we had to extend our redshift limits to 0.1 and 0.3 in order to alleviate the effect of cosmic variance. After making these cuts, we extract three different galaxy samples with different photometric depths. For the sake of simplicity, we have opted to select these samples using a single band, somehow imitating the SDSS selection. We have chosen

the i^A band to this purpose. This choice is motivated mainly by the fact that the SED variation in the redder bands (typically tracing the older stars) is smaller than in the bluer bands, which are also more affected by dust extinction. The first sample, Sample1, is selected so that $i^A \leq 20$ and this is the closer, in terms of depth, that we can get to the SDSS (wich is limited at $r \sim 17.77$). Again, ALHAMBRA is severely incomplete for very bright objects at these redshifts due to the small volume surveyed. At some point, it will be convenient to extent our analysis to fainter magnitudes. We will then use Sample2, where $i^A \leq 22.5$, and Sample3, where $i^A \leq 24.5$. Sample1 comprises $1,290$ objects, Sample2, $4,700$ objects and Sample3, 13,098 objects. Finally, in Figure 4.4, we show the redshift distribution in each of the samples. Note that, mainly due to the faint apparent limits that we have imposed, the distribution of objects is pretty uniform across our redshift range.

4.4 Galaxy Number Counts

Counting the number of galaxies per unit area and apparent magnitude is a basic way to photometrically characterize a survey. Cristóbal-Hornillos et al. (2009) presented a detailed analysis on the GNC in the NIR JHK bands in a 0.5 square-degree portion of the ALHAMBRA survey. We have studied counts in the optical range in the MC, that, as we have mentioned in the previous section, comprises $220,772$ objects - once all non-galactic sources have been removed. In the upper-left panel of Figure 4.5 we present GNC in each of the ALHAMBRA narrow filters. Curves in this plot are rainbow-colored from black to red, representing from the bluest to the reddest filter. The depth of ALHAMBRA is apparent here, with the redder filters peaking around $23.5-24$ and the bluest reaching as deep as $m \sim 25$. Note that the survey is photometrically shallower towards the redder bands. This is a consequence of the fact that detection in ALHAMBRA is performed in a single synthetic image, the so-called DEEP image, which is a weighted combination of several filters.

In order to put ALHAMBRA in context with other surveys we will use synthetic SDSS bands. As expressed in previous sections, ALHAMBRA provides vast photometric information that first needs to be simplified in order to understand the data. In the remaining three plots of Figure 4.5 we show GNC in bands g^A, r^A and i^A (crosses). Note that, as can be seen in Figure 4.1, the SDSS u and z bands are not completely covered by the ALHAMBRA filter system, so we cannot use them in our comparisons. According to the GNC shown in Figure 4.5, ALHAMBRA seems to be photometrically complete up to an apparent magnitude of 25 in the g^A band. In the redder bands the survey is slightly shallower, especially in the i^A band, where GNC start to decline somewhere around $i^A \sim 24$. At the bright end, ALHAMBRA is clearly incomplete below $g^A \lesssim 17$, $r^A \lesssim 16.5$ and $i^A \lesssim 16$.

ALHAMBRA GNC have been compared with three photometric galaxy surveys. First of all, we have used the SDSS photometric catalog Yasuda et al. (2001), mainly to get an idea of completeness in the bright end (blue line). It is important to remind the reader at this point that ALHAMBRA is a pencil beam survey that maps a considerably smaller volume than the SDSS, especially at low redshift. This obviously affects ALHAMBRA at the bright end, where objects are remarkably scarcer, as a comparison with the SDSS reflects. In the mid range, the agreement is good in the g and the r band. Only in the redder i band do we find some discrepancies as compared with the SDSS photometric catalog, with ALHAMBRA capturing more obects per apparent magnitude bin and sq deg.

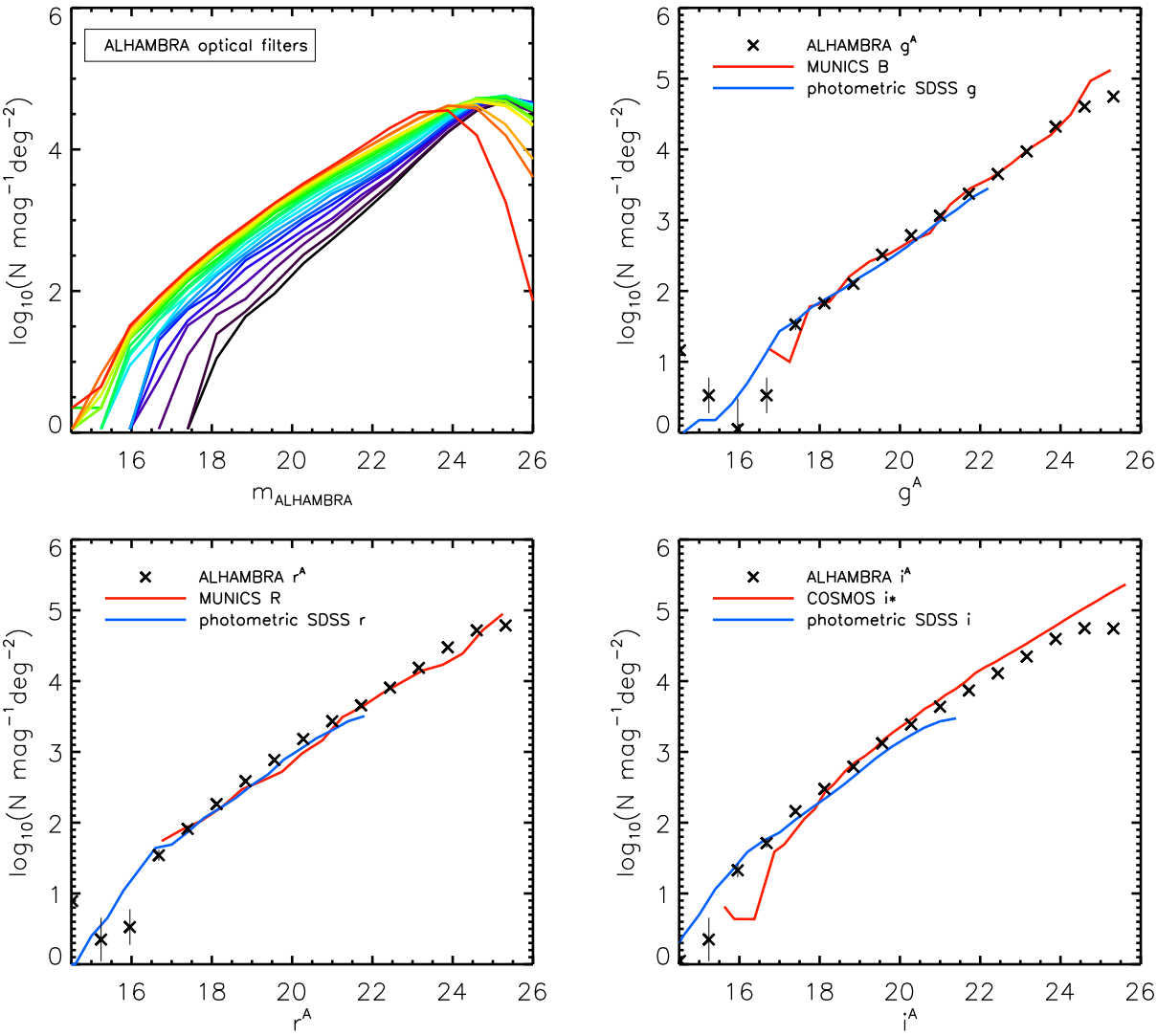

Figure 4.5 ALHAMBRA GNC in each of the optical filters (top left panel) and in the synthetic g^A, r^A and i^A bands (crosses, rest of the panels). In each of the panels showing synthetic-band GNC, the SDSS GNC from Yasuda et al. (2001) are shown for comparison in the corresponding band (blue line). In addition, ALHAMBRA GNC in the g^A and r^A bands are compared with counts from the MUNICS survey in the B and R bands, respectively (red lines). The ALHAMBRA i^A-band GNC are compared with COSMOS (i^* band). Finally, an Euclidean model is shown in a dashed line as a reference.

In the g and r bands, we have compared ALHAMBRA with the Munich Near-Infrared Cluster Survey (MUNICS, Drory et al. 2001), a wide-field medium-deep imaging survey in the NIR and optical ranges. We find an excellent agreement with GNC in the MUNICS B and R bands (red lines), which are comparable with the SDSS g and r bands. MUNICS GNC have been taken from Feulner et al. (2007), where the authors used a sample of 0.26 deg^2 from the total area of 1 deg^2 mapped by the survey.

Finally, ALHAMBRA GNC in the i^A band have been compared with GNC from the Cosmic Evolution Survey (COSMOS, Scoville et al. 2007). COSMOS is an extremely deep photometric survey which uses 15 bands covering from the X-ray to the radio to map an area of 2 deg^2 on the sky. The data, which corresponds to the COSMOS i^+ band (see Scoville et al. 2007 for more information) is represented in a red line in Figure 4.5 and was taken from Capak et al. (2007). It appears clear from this comparison that ALHAMBRA is shallower than COSMOS in the i band by at least 1 magnitude. Note that COSMOS, which was actually selected in the i^+ band, is claimed to be 91% complete at $i^+ = 25$ and even an outstanding 50% complete at $i^+ = 27.4$, according to Capak et al. (2007). In Figure 12 from Capak et al. (2007) we can see how ALHAMBRA is consistent with other deep catalogs such as the Hawaii-Hubble Deep Field (Capak et al., 2004) or the Canada France Deep Fields (McCracken et al., 2003).

4.5 K-corrections and absolute magnitudes

We have obtained absolute magnitudes in the three samples defined in Section 4.3.2: Sample1 ($i^A < 20$), Sample2 ($i^A < 22.5$) and Sample3 ($i^A < 24.5$). The absolute magnitude of an object (M), as we have shown in Chapter 2, can be estimated from the apparent magnitude (m) and the redshift (z) in the following way:

$$M = m - 5log_{10}h - DM(z) - K(z) - E(z) \tag{4.1}$$

where $DM(z)$ is the distance modulus and $K(z)$ and $E(z)$ are the so-called K-correction and evolution correction for the galaxy, respectively. The K-correction (Oke & Sandage, 1968; Hogg et al., 2002), allows us to transform the flux (or, equivalently, the absolute magnitude) that we measure from an object at redshift z at a certain passband to the rest-frame flux at the desired passband. The evolution correction, on the other hand, is included to account for the evolution in the luminosity of a galaxy from redshift z to the rest-frame redshift (commonly $z = 0$). We have not computed a detailed evolution correction so far to our galaxy samples. However, as will show in Chapter 5 and in the Appendix, our K-correction includes an estimation of the evolution correction as well. For higher redshift ranges, especially when studying the evolution of the luminosity function or the stellar mass function, the evolution correction must be taken under great consideration.

In Chapter 2, we mentioned that SDSS K-corrections were taken from the NYU-VAGC DR6. In ALHAMBRA, unfortunately, no K-corrections are available so we need to find a way to correct our magnitudes by ourselves. There are several well-known codes for estimating K-corrections publicly available, such as *kcorrect* (Blanton & Roweis, 2007), which is the code used in the NYU-VAGC. Some

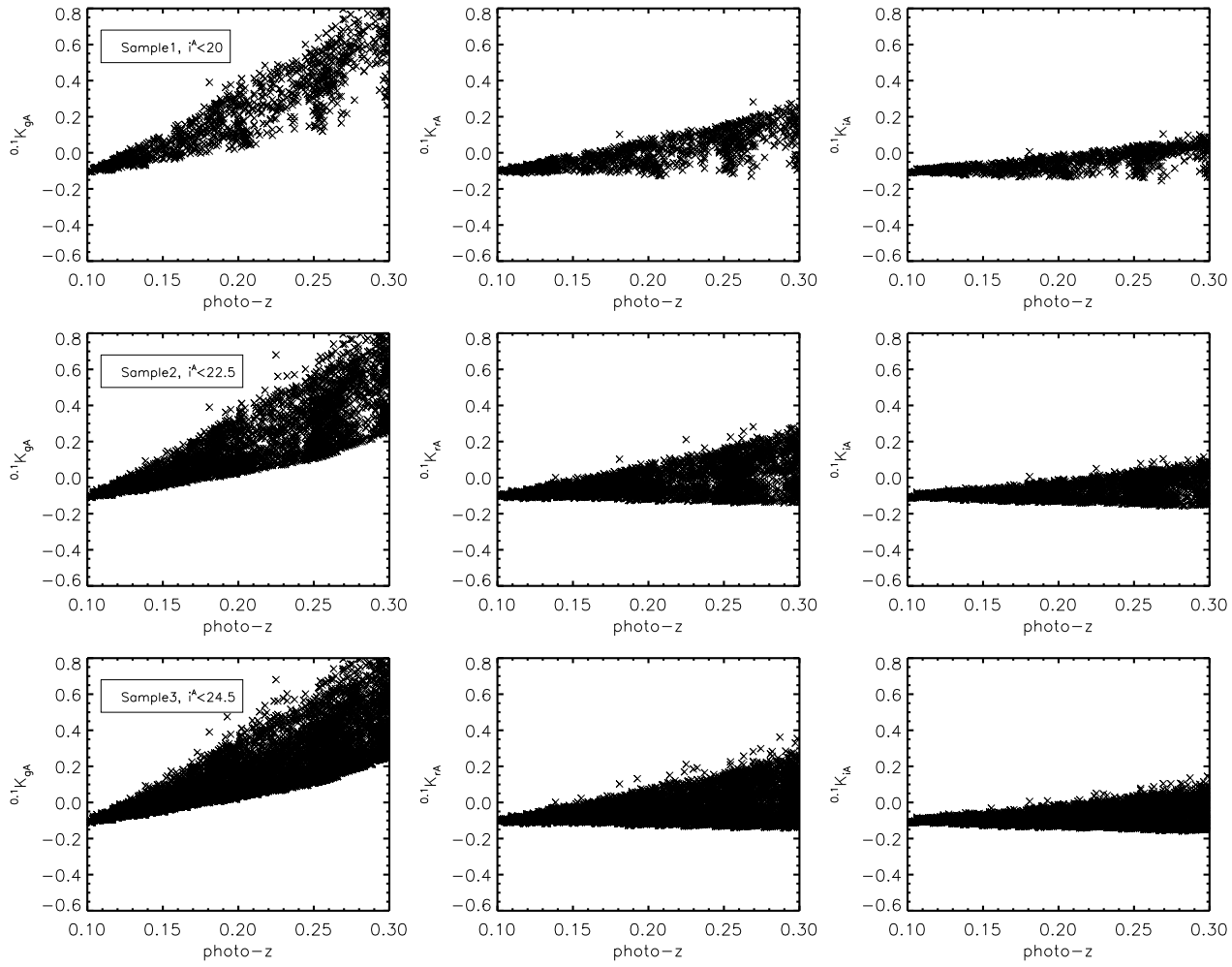

Figure 4.6 K-corrections to $z = 0.1$ as a function of photometric redshift for the three ALHAMBRA samples considered and for bands g^A, r^A and i^A. ALHAMBRA K-corrections have been estimated using the Flexible Stellar Population Synthesis code (Conroy et al., 2009) and are consistent, generally speaking, with the SDSS K-corrections of Blanton & Roweis (2007).

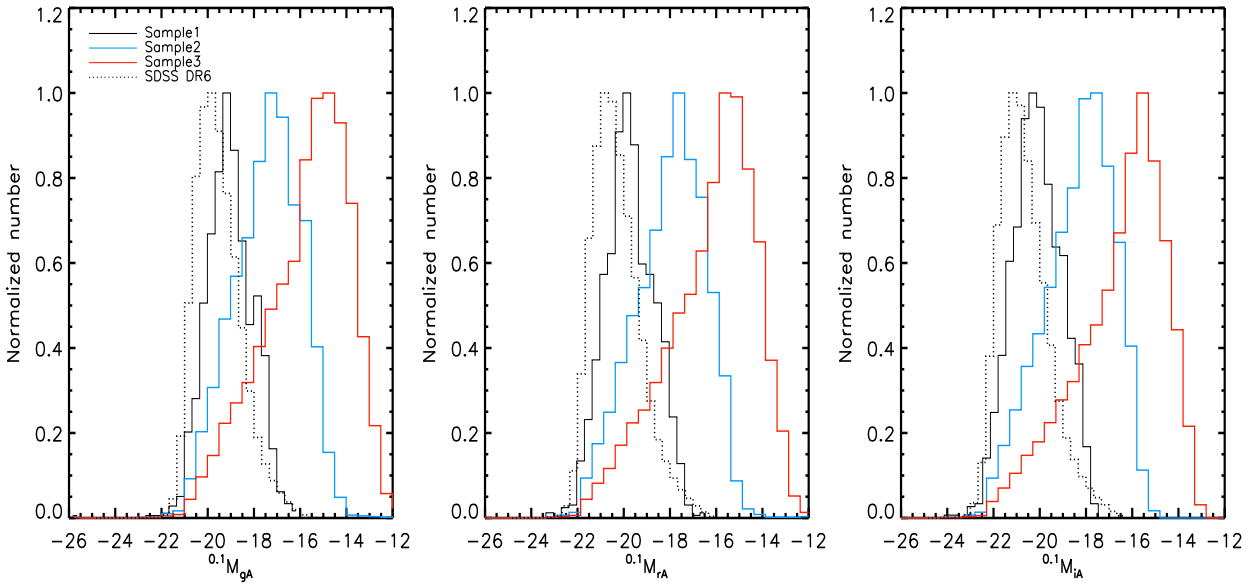

Figure 4.7 Distribution of K-corrected ALHAMBRA absolute magnitudes in bands g^A, r^A and i^A for Sample1 (black solid line), Sample2 (blue line) and Sample3 (red line). K-corrections have been estimated to $z = 0.1$ to facilitate comparison with the SDSS DR6 (dashed lines).

authors have also provided analytical expressions for K-corrections in several widely-used filters as a function of redshift, normally for different galaxy types (e.g. Chilingarian et al. 2010; Westra et al. 2010). Instead of using existing software or relying on analytical approximations, we have developed our own code for estimating K-corrections in ALHAMBRA, using the Flexible Stellar Population Synthesis code (FSPS, Conroy et al. 2009). In the next chapter, we will describe in detail FSPS, which have been basically used for estimating photometric stellar masses. In addition, the method for obtaining K-corrections, based on the formalism presented in detail and thoroughly discussed in Hogg et al. (2002) and Blanton & Roweis (2007), is described in the Appendix. It is convenient to anticipate, however, that the key aspect of this estimation is the SED fitting. Once we have obtained a model for the SED of a given galaxy, it is straightforward to estimate K-corrections. We point the reader to Chapter 5 and the Appendix for further details about this calculation.

In Figure 4.6, we show ALHAMBRA K-corrections in the synthetic g^A, r^A and i^A bands as a function of redshift for the 3 samples considered: Sample1 in the upper panel, Sample2 in the the middle and Sample3 in the lower panel. Note that in this figure, the rest-frame redshift is set at $z = 0.1$, instead of $z = 0$. In some cases, it is convenient to blue-shift the rest frame so that K-corrections, which represent a non-negligible source of uncertainty in the estimation of absolute magnitudes, are smaller. Blueshifting K-corrections and consequently absolute magnitudes to $z = 0.1$ is a common procedure in a number of SDSS works, which is motivated by the fact that the SDSS galaxy distribution peaks around this redshift. From now on in this chapter, and unless otherwise stated, we follow this prescription in order to facilitate comparison with the SDSS. The K-corrections presented in Figure 4.6 are in very good agreement with those from Blanton & Roweis (2007), as Figure 16 from that work reveals. Note that K-corrections are typically smaller towards the redder bands,

in any of the samples, reflecting the fact that galaxy spectra are on average flatter towards longer wavelengths. Another interesting feature which is worth highlighting from Figure 4.6 is the increase in the dispersion towards higher redshifts. This increase is, in general, not due to uncertainties in the photometry or in the SED fitting, which would tend to be larger in Sample2 and Sample3. This is a natural dispersion that reflects the presence of different types of galaxies, with very different spectral shapes, in our samples. The reader can find a good illustration of this in Figure 5 from Westra et al. (2010), where predictions for K-corrections as a function of redshift are shown, in the g and r bands, for several galaxy models with varying age.

In Figure 4.7 we present the distribution of K-corrected (to $z = 0.1$) absolute magnitudes in bands g^A, r^A and i^A for the 3 samples considered. It is important to remind at this point that this work must be regarded as a preliminary study that is intended to lay the foundations for future galaxy evolution studies with ALHAMBRA. The absolute magnitudes shown in Figure 4.7 are expected to be affected by the two main sources of uncertainty that we have to deal with here: errors in the photometry and errors in the estimation of the photo-z's, which by the time this work was created were still not in their final version. Errors in the photometry are of course expected to be larger towards fainter apparent magnitudes. This implies a more ambiguous photo-z estimation, which turns out to be harder also for very nearby objects. Sample2 and Sample3 will therefore give us a valuable idea of how the survey behaves at faint magnitude ranges but the reader must be especially careful not to over-interpret our results from these samples. We have over-plotted in each panel of Figure 4.7 the distribution of absolute magnitudes in the SDSS DR6 (spectroscopic catalog), in the photometric g, r and i bands. The ALHAMBRA distributions cover a wide range of magnitudes in any of the bands, peaking around an absolute magnitude of ~ -19.5 for Sample1 and ~ -15 for the very deep Sample3. Note that even in Sample1 the ALHAMBRA distributions peak approximately 1 magnitude fainter than the SDSS distributions. The SDSS spectroscopic catalog is restricted to $r^{0.1} < 17.77$, which roughly corresponds to $i^{0.1} < 17.2$, still almost 3 magnitudes fainter than the ALHAMBRA Sample1.

4.6 Color-magnitude distributions

Galaxy colors are a simple but useful tool for characterizing a galaxy population in a statistical way. A remarkable statistical feature of the low-redshift galaxy population is the color bimodality, which was first found by Strateva et al. (2001) by studying the optical colors of $147,920$ galaxies brighter than $g = 21$ in the SDSS. They found that the distribution of galaxy colors in the SDSS is clearly bimodal, with two peaks that correspond to late- and early-type galaxies, respectively. Also, Balogh et al. (2004) analyzed the u - r color distribution of $24,346$ galaxies in the SDSS and found that red galaxies and blue galaxies have completely different behaviors with respect to the density of the environment. Namely, at fixed luminosity there is a clear preference from red galaxies to inhabit denser regions, whereas blue galaxies seem to populate any environment independently of the density. In Figure 4.8 we show the ${}^{0.1}(g-r)$ color distribution in ALHAMBRA (Sample1 and Sample2) and in the SDSS (the ${}^{0.1}(u-r)$ color is not reliable in ALHAMBRA). Note that the SDSS galaxy population (represented by a dashed line in Figure 4.8) presents a remarkable overabundance of red objects in a tight range of colors, the so-called red sequence, and an extended distribution of bluer objects, forming a blue cloud.

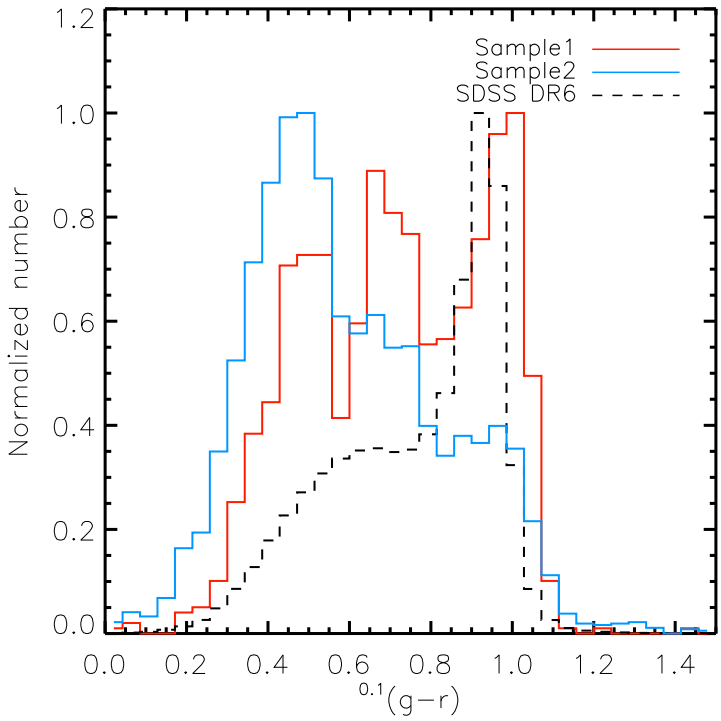

Figure 4.8 The $^{0.1}(g-r)$ color distribution in the ALHAMBRA Sample1 (red line) and Sample2 (blue line) and in the SDSS DR6 (dashed line). The ALHAMBRA distributions of $^{0.1}(g-r)$ colors deviate considerably from that of the SDSS. In Sample1, the red sequence is still predominant, although blue objects are considerably more numerous than in the SDSS. Sample2 is clearly dominated by a blue population.

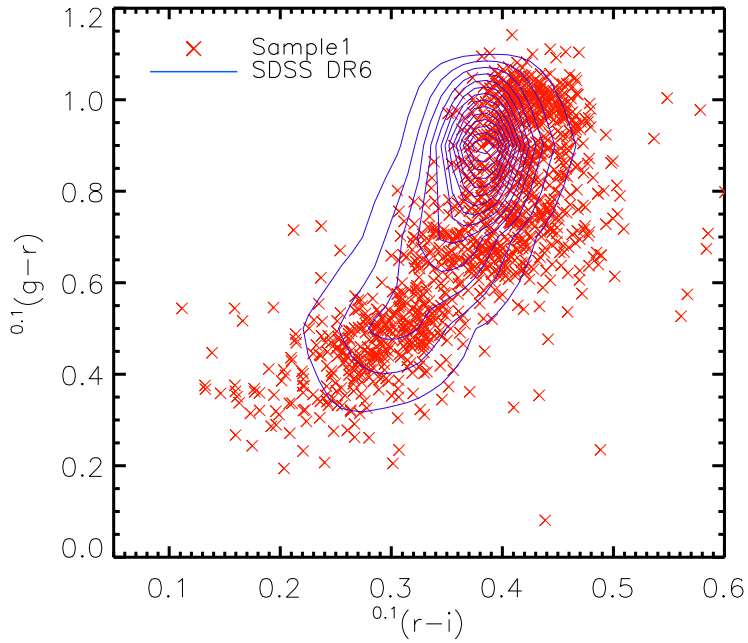

Figure 4.9 The $^{0.1}(g-r)$ vs. $^{0.1}(r-i)$ color-color diagram for the ALHAMBRA Sample1 (red crosses) and for the SDSS DR6 (blue contours).

This configuration is believed to be the product of the evolution of blue star-forming galaxies into red galaxies through several processes of star formation quenching (including galaxy mergers) across cosmic time (see Faber et al. 2007 for further details). In Figure 4.8 we see how the ALHAMBRA distributions of $^{0.1}(g-r)$ colors deviate considerably from that of the SDSS. In Sample1 (red line), the red sequence is still predominant, although blue objects are considerably more numerous than in the SDSS. Sample2 is clearly dominated by a blue population. This discrepancy is probably due to the average luminosity of the samples. The simplest explanation is that, as we go deep into i^A magnitudes and we explore progressively fainter absolute magnitude ranges, the fraction of blue objects increases. This can be easily seen in the color-magnitude diagram of Figure 2.6 in Chapter 2, where the red sequence is located in a relatively narrow magnitude range as compared to the blue cloud that extends into deeper magnitudes (see also Figure 1 from Balogh et al. 2004).

In Figure 4.9 we compare the ALHAMBRA Sample1 and the SDSS DR6 in a $^{0.1}(g-r)$ vs $^{0.1}(r-i)$ color-color diagram. These are the two main colors that we can use for comparisons. The ALHAMBRA distribution, which is shown in red crosses in this plot, is roughly consistent with the SDSS, represented by blue contours. However, the ALHAMBRA $^{0.1}(r-i)$ color distribution is slightly redder than that of the SDSS, mainly for large $^{0.1}(g-r)$ colors (by ~ 0.03 magnitudes). Note also that contours in this plot show the prominence of the red sequence in the SDSS. In the ALHAMBRA Sample1, however, the blue population is almost as numerous as the red population.

Finally, in Figure 4.10 we present the ALHAMBRA $^{0.1}(g-r)$ vs $^{0.1}M_r$ color-magnitude diagram (CMD) for Sample1, which gives us an idea of the distribution of colors as a function of luminosity in ALHAMBRA. Again, the agreement is good with the SDSS DR6, represented by blue contours.

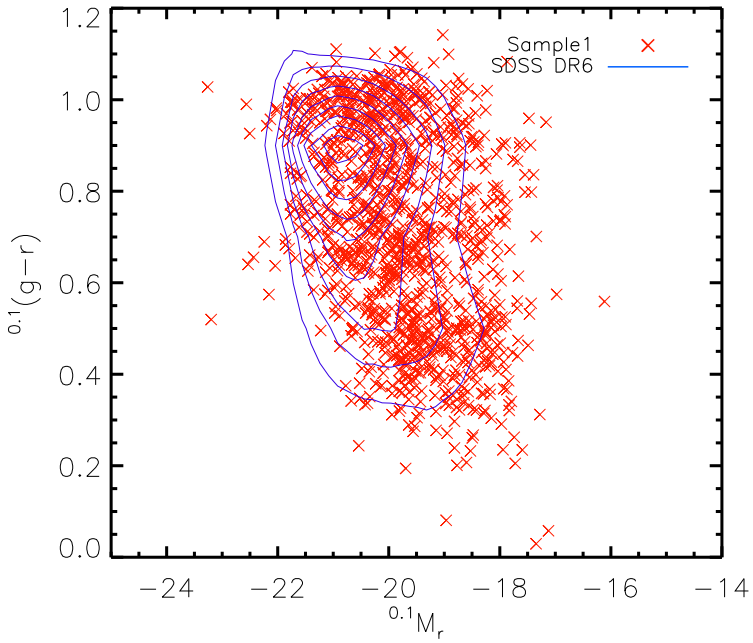

Figure 4.10 The ALHAMBRA $^{0.1}(g-r)$ vs. $^{0.1}M_r$ color-magnitude diagram for Sample1 (red crosses) and for the SDSS DR6 (blue contours).

Also, a deviation of the $^{0.1}M_r$ distribution towards fainter magnitudes is found, which is consistent with Figure 4.7.

4.7 The ALHAMBRA low-redshift luminosity functions

4.7.1 Methodology and sample definition

In previous sections we provide a first approach to the ALHAMBRA data, restricting ourselves to a pseudo low-redshift sample ($0.1 < z < 0.3$). We have shown that galaxy number counts and the distribution of absolute magnitudes and colors seem reasonable as compared to other well-known surveys. The depth of our samples (or even cosmic variance) appear to be the reason, however, that we are mapping a considerably bluer population than the SDSS. Taking this into account, in this section we provide the first estimation of ALHAMBRA luminosity functions in both the synthetic i^A and r^A bands.

In order to calculate LFs we will use two methods: the Stepwise Maximum Likelihood Method (SWML, Efstathiou et al. 1988) and the $1/V_{max}$ method (Schmidt, 1968). We will focus here on describing the second of these methods in the form proposed by Davis & Huchra (1982). For more information about the SWML method see Chapter 2 and Efstathiou et al. (1988). In addition, the reader can find a complete discussion on statistical methods for estimating galaxy luminosity functions in Takeuchi et al. (2000).

The $1/V_{max}$ method is a simple non-parametric method for calculating LFs which is based on the estimation of the maximum volume where a galaxy in our sample could be found while satisfying

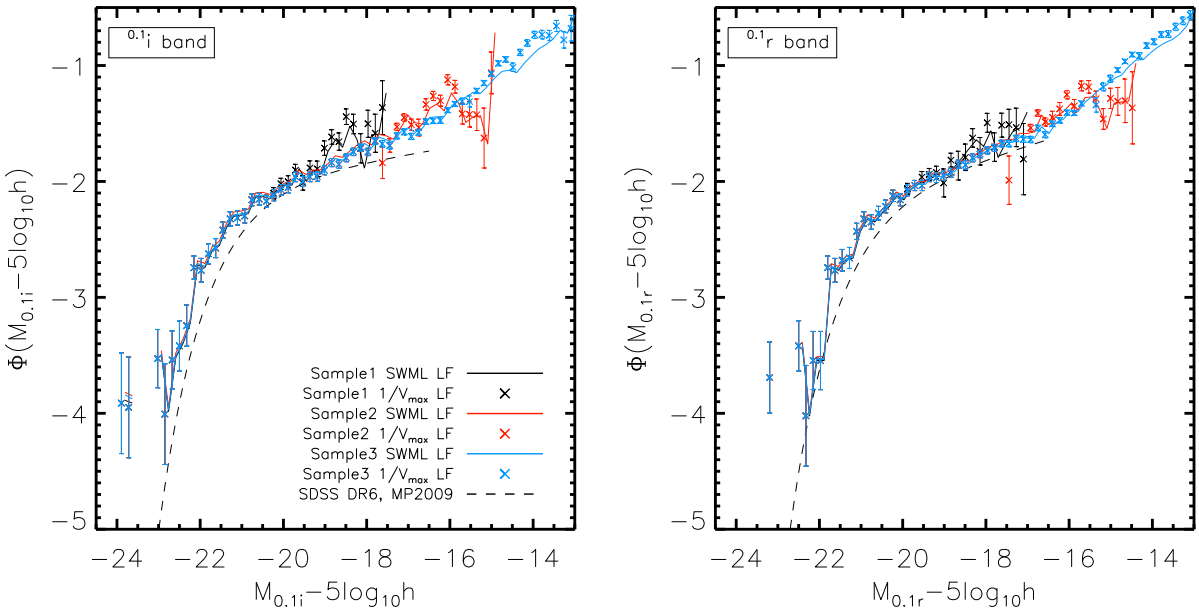

Figure 4.11 The ALHAMBRA luminosity functions in both the i^A (left) and r^A (right) bands. In each panel, the three samples selected in each band are plotted in black (Sample1), red (Sample2) and blue (Sample3), where crosses represent the LFs obtained with the $1/V_{max}$ method and lines represent those estimated with the SWML method. Errors are only shown for the $1/V_{max}$ estimates for the sake of clarity. We have over-plotted the Schechter fits to the SDSS DR6 LFs discussed in Chapter 2 (and also presented in Montero-Dorta & Prada 2009)

the survey constraints. Let us assume that we have the following lower and upper limits in apparent magnitude and redshift in our sample:

- $m_l < \mathrm{m} < m_u$
- $z_l < \mathrm{z} < z_u$

with a survey solid angle Ω. The differential luminosity function, $\Phi(M)$, at a given absolute magnitude bin, ΔM, is just given by the sum of the density contributions of each individual galaxy in the considered luminosity bin:

$$\Phi(M)\Delta M = \sum_i \frac{1}{V_{max,i}(M_i, z_i, SED_i)} \tag{4.2}$$

where M_i is the absolute magnitude, z_i is the redshift, SED_i is the spectral energy distribution that represent the shape of the spectrum and $V_{max,i}$ is the maximum volume for each galaxy. $V_{max,i}$ can be expressed as

$$V_{max,i} = \int_\Omega \int_{z_{min,i}}^{z_{max,i}} \frac{d^2V}{d\Omega dz} dz d\Omega \tag{4.3}$$

The critical step in this calculation is the estimation of $V_{max,i}$ and, in particular, the main difficulty resides in the estimation of the corresponding redshift range, $z_{min,i}$ and $z_{max,i}$. In practice, we need to solve the following equations iteratively for each galaxy:

$$m_u = M_i + 5log_{10}h + DM(z^*_{min,i}) + K(z^*_{min,i})$$
$$m_l = M_i + 5log_{10}h + DM(z^*_{max,i}) + K(z^*_{max,i})$$

and, finally,

$$z_{min,i} = max[z_l, z^*_{min,i}]$$
$$z_{max,i} = min[z_u, z^*_{max,i}]$$

In order to correctly implement the $1/V_{max}$ method, knowning the K-correction for each galaxy is not enough. We need to be able to estimate the K-correction that a galaxy with a given M_i and SED_i would hypothetically have if placed at any redshift within the limits of our sample. Again, we can simply use analytical approximations for K-corrections as a function of redshift and color (or galaxy type). A much better solution, however, can be obtained by using our accurate method based on the FSPS code mentioned above.

In previous sections, we have defined three samples selected from our MC in the synthetic i^A band. We will use these samples to estimate the ALHAMBRA LF in the i^A band. In order to facilitate comparisons with other surveys, we will also present the ALHAMBRA LF in the r^A band

Band	Sample	Number	m_l	m_u	z_l	z_u
r	1	1,206	16.5	20.5	0.1	0.3
	2	4,294	16.5	23	0.1	0.3
	3	10,614	16.5	25	0.1	0.3
i	1	1,290	16	20	0.1	0.3
	2	4,700	16	22.5	0.1	0.3
	3	13,098	16	24.5	0.1	0.3

Table 4.1 Number of galaxies and apparent magnitude and redshift limits for each of the samples selected in order to estimate ALHAMBRA LFs in the r^A and i^A bands.

and to this purpose we will define in a similar way three other samples selected from the MC in this band. In Table 4.1 we list apparent magnitude and redshift limits along with the number of objects for each sample. Note that the apparent magnitude limits have been set 0.5 mag fainter in the r^A band than in the i^A band, which is roughly consistent with the ALHAMBRA GNC from Figure 4.5 and also from the SDSS (spectroscopic) GNC from Figure 1.4, Chapter 2.

4.7.2 Results

In Figure 4.11, we present the ALHAMBRA LFs in both the i^A band (left) and the r^A band (right). In each panel, the three samples selected in each band are plotted in black (Sample1), red (Sample2) and blue (Sample3); where crosses represent the LFs obtained with the $1/V_{max}$ method and lines represent the LFs estimated with the SWML method. Errors are only shown for the $1/V_{max}$ estimates for the sake of clarity. We have over-plotted the Schechter fits to the SDSS DR6 LFs discussed in Chapter 2 (and also presented in Montero-Dorta & Prada 2009). The ALHAMBRA LFs shown in Figure 4.11 present some noticeable characterics, as compared to the SDSS DR6 LFs. First of all, we find a remarkable excess of bright galaxies with respect to the SDSS LF, that might also be understood as a brighter M^*, considering the smoothness of the function. Secondly, we obtain a very steep slope at the faint end, below $M^A \sim -20$. Not so evident but still noticeable is the slight discrepancy in the normalization, with the SDSS DR6 LFs falling behind the ALHAMBRA LFs.

Considering the state of the ALHAMBRA data used in this work, it is hard to provide a full interpretation of the results shown in Figure 4.11. In particular, we cannot give a definite answer to whether the shape of the LF shown in Figure 4.11 responds to real statistical properties of the low-redshift galaxy population or if it is otherwise affected by dominant systematics. Let us first briefly discuss on the systematics we might suffer from. Naively, we would expect that uncertainties in the photometry were not particularly relevant at the bright end, where the flux estimation is more reliable. However, it is known that objects are considerably scarce at these ranges even in the SDSS. Cosmic variance might therefore be an issue, especially at low redshifts (note that we are sampling a volume that is $\sim 1,000$ times smaller than that of the SDSS). At the faint end, we have to deal with extremely low-surface brightness galaxies and, consequently, with a considerably more ambiguous flux estimation, that might affect the slope of the LF.

Besides the above considerations, which are common in most LF works, we must not forget the fact

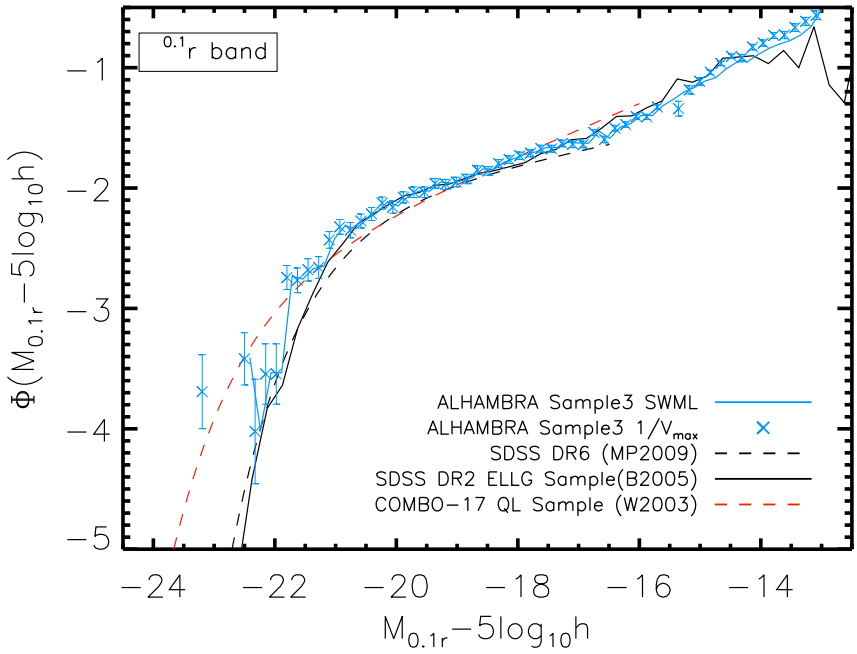

Figure 4.12 The ALHAMBRA r^A-band LF in the very deep Sample3 ($r^A < 25$) is compared with previous estimates obtained with deep samples. The ALHAMBRA LF is shown in the same format as in Figure 4.11. The black solid line represents the LF of Blanton et al. (2005b), obtained with an extremely low lumonisity galaxy (ELLG) sample taken from the SDSS DR2. The red dashed line shows the Schechter fit to the COMBO-17 LF of Wolf et al. (2003), obtained in a quasi-local (QL) sample. Finally, in a black dashed line, we show the Schechter fit to the SDSS DR6 LF discussed in Chapter 2 (and also presented in Montero-Dorta & Prada 2009).

that we are using photometric redshifts. In essence, each photo-z must be considered as a probability distribution, which differs from the "classical" concept of spectroscopic redshift. In simulations undertaken with ALHAMBRA galaxy mock catalogs (provided by R. Wechsler and her group at Stanford University) we have checked that a Poissonian distribution of photo-z errors with FWHM equal to the average uncertainty in the ALHAMBRA photo-z's ($\delta z/(1+z) \leq 0.015$) would not alter the shape of the LF but would add a slight uncertainty to the normalization of the LF (note that normalization might also be affected by cosmic variance). However, we do expect photo-z accuracy to change (on average) across the entire range of magnitudes, in a non-trivial way. At the faint end, uncertainties in the photometry should correlate with uncertainties in the photo-z estimation. At the very bright end, we have learnt within the collaboration that this estimation might be noisy for very large and luminous objects. The next step in order to advance in the scientific exploitation of ALHAMBRA is therefore to characterize its photo-z and detection incompleteness, a purpose several collaborators are devoting considerable effort to.

Bearing the above remarks in mind, we can speculate on the implications of the results shown in

Figure 4.11. Note that evidences can be found in the literature to support the argument that the deviations from the SDSS DR6 shown in Figure 4.11 are due to real intrinsic statistical properties of the galaxy population. In particular, a steep slope in the very faint end of the LF has been reported in several works, that typically use very low redshift samples. A first example of this can be found in Loveday (1997), who reports a considerable excess of very low luminosity galaxies with respect to a best-fitting Scechter function in the Stromlo-APM survey. A similar result is found by Lin et al. (1996) in the Las Campanas Redshift Survey (LCRS). In Figure 4.12 we compare the ALHAMBRA r^A-band LF (only Sample3 is shown, in blue) whith other well-known low-redshift works. A Schechter fit to the r-band COMBO-17 LF in a quasi-local sample (QL Sample, Wolf et al. 2003) is represented by a red dashed line. This sample covers a redshift range $0.2 < z < 0.4$. Note that COMBO-17 is analoguos to ALHAMBRA in terms of covered area (also reaching down to $R \lesssim 24$). Interestingly, the LF from Wolf et al. (2003) presents a remarkable bright-end excess of galaxies as compared to the SDSS DR6 LF (shown in a black dashed line) and a steeper slope at the faint end (reportedly $\alpha \sim -1.5$), similarly to ALHAMBRA. The authors are not able to give conclusive explanations but speculate on differences in the photometry and the evolution of dwarf galaxies (from $z \sim 0.35$ to $z \sim 0.05$) as possible reasons for the above discrepancies, respectively. Blanton et al. (2005b) analyzed the properties and luminosity function of extremely low lumonisity galaxies (ELLG) in a low-redshift sample with the SDSS DR2. Correcting for the strong low surface brightness selection effects, the authors predict a LF with $\alpha \sim -1.5$ (or even steeper) reaching down to $M_r \sim -12.5$ in the r band. This LF, shown in Figure 4.12 in a solid black line, is in surprising agreement with the ALHAMBRA LF in almost the entire absolute magnitude range, except for the very bright end. Albeit still preliminary, these results would suggest that the dwarf galaxy population in the nearby Universe is considerably larger than what would be inferred from the SDSS (or the 2dF RS) LF (see Chapter 2).

4.7.3 A note on the faint-end upturn

In Figure 4.11 and Figure 4.12, an upturn is noticeable in the ALHAMBRA LF at $M^A \sim -17$. This implies that the LF is not well fitted by a single Schechter function, and a double Schechter is needed. This is an intriguing result that must be investigated in more detail in the future. A similar upturn have been found by several groups in the low-mass end (below 10^{10} $M_\odot$) of the field stellar mass function (SMF; Baldry et al. 2008 and before, Pozzetti et al. 2010, Drory et al. 2009) and also the cluster SMF (Popesso et al., 2006; Jenkins et al., 2007). Regarding the LF, which is intrinsically connected with the SMF, the presence of the faint-end upturn is only accepted (at least to a certain extent) in the clusters LF. Popesso et al. (2005) find two clearly distinctable components with an observed upturn of faint galaxies somewhere below $M_r \sim -18$. A similar result is found by Mercurio et al. (2006) with the Shapley Optical Survey. This would suggest that not only the faint-end upturn but also the steep faint-end slope found in the ALHAMBRA LF might be due to cosmic variance. We might be mapping an over-dense region of the Universe with the presence of a large filament. We use, however, different fields in this work, which are located at different pointings precisely to overcome cosmic variance. A simple way to investigate this issue is to undertake a proper field-to-field analysis of the ALHAMBRA LF (even though the statistics for each field may be small).

In the field, the most accurate estimations of the LF come from large-scale surveys such as the SDSS or the 2dF RS (both at $z \sim 0.1$). The faint-end limit in apparent magnitude makes it hard to study the dwarf galaxy population (see Montero-Dorta & Prada 2009 or Blanton et al. 2003b for the SDSS and Norberg et al. 2002 for the 2dF RS). Interestingly, faint-end upturns have been found with both the SDSS and the 2dF RS when restricting the selection to very low redshift samples (in order to reach into deeper absolute magnitudes) or splitting the galaxy sample into different spectral types, respectively. In Figure 4.12, a faint-end upturn is noticeable in the SDSS LF from Blanton et al. (2005b) at $M_r \sim -19$, i.e., at considerably brighter magnitudes than that apparently found in ALHAMBRA. Madgwick et al. (2002) find that the Schechter function provides an inadequate fit of the 2dF RS LF calculated over the magnitude range $-22 \leq M_{\mathrm{b}_J} - 5\ \log(h) \leq -13$, especially for the most passive and star-forming galaxies. Finally, the LFs calculated in the deeper galaxy samples provided by the Stromlo-APM survey (Loveday, 1997) or the LCRS (Lin et al., 1996) present and excess of dwarf galaxies which makes them incompatible with a single Schechter fit.

Understanding the faint-end of the ALHAMBRA LF and, in particular, checking that both the steep faint-end slope and the faint-end upturn are not due to cosmic variance or errors in the photo-z estimation is crucial as it is in these magnitude ranges where ALHAMBRA is expected to shed light into the galaxy formation and evolution picture.

4.8 Chapter Conclusions

ALHAMBRA is an ongoing photo-z survey primarily designed for the study of the evolution of the contents of the Universe from high redshift ($z \sim 1$). To this purpose, it uses 20 contiguous, equal-width filters covering the optical range, along with the J, H and K bands of the NIR, to map an area of 4 deg^2 on the sky down to a magnitude $I(AB) \sim 25$. At the present time, most of the observations have been carried out and data reduction and photo-z estimation have reached to a state where we can commence the scientific exploitation of the survey. In this chapter, we use a portion of ~ 1 deg^2 and we restrict ourselves to a pseudo low-redshift sample ($0.1 < z < 0.3$) to investigate some important statistical properties of the ALHAMBRA galaxy population, including magnitude and color distributions, galaxy number counts and galaxy luminosity functions. The photo-z estimation in ALHAMBRA is, however, still preliminary so we must take our results with much care.

The main results of this study can be summarized as follows:

- The ALHAMBRA galaxy number counts in the synthetic g^A, r^A and i^A bands are in good agreement with the SDSS photometric catalog and with other deep surveys such as MUNICS (g^A and r^A bands) and COSMOS (i^A band). Importantly, GNC start to decline around $g^A \sim 25$, $r^A \sim 24.5$ and $r^A \sim 24$, confirming the extraordinary photometric depth of ALHAMBRA.

- We have estimated ALHAMBRA K-corrections using the Flexible Stellar Population Synthesis (FSPS, Conroy et al. 2009) code. Our ALHAMBRA K-corrections are qualitatively consistent with the SDSS K-corrections from Blanton & Roweis (2007). With these K-corrections we have obtained ALHAMBRA absolute magnitudes.

- Due to the depth of the survey, the ALHAMBRA blue population is remarkably predominant, as compared to the red population. This implies that ALHAMBRA maps a considerably bluer galaxy population than the SDSS. This said, the ALHAMBRA $(g-r)^{0.1}$ versus ${}^{0.1}(r-i)$ and the ${}^{0.1}(g-r)$ vs ${}^{0.1}M_r$ diagrams are consistent with the SDSS.

- We have estimated ALHAMBRA luminosity functions of galaxies in the r^A band and the i^A band, using both the $1/V_{max}$ and the SWML method. ALHAMBRA LFs have been calculated in three different samples selected in each band with progressively fainter apparent magnitude limits. As compared to the SDSS, the ALHAMBRA LFs present three noticeable characteristics: an excess of bright galaxies, a slightly higher normalization and a remarkable faint-end slope ($\alpha \sim -1.5$).

- At the bright end, cosmic variance, due to the relatively small volume sampled by ALHAMBRA at $z < 0.3$, and/or uncertainties in the photo-z estimation might be causing these discrepancies (this could also explain the normalization issue). A similar excess of bright galaxies is found in a similar sample extracted from the photo-z COMBO-17 survey (Wolf et al., 2003).

- At the faint end, photometry is less reliable and, consequently, the photo-z estimation is more ambiguous. However, clear evidence of similar steep slopes at the faint end can be found in literature, when comparing with more adequate galaxy samples. Wolf et al. (2003) found a very pronounced faint-end slope in a pseudo local sample extracted from COMBO-17. Remarkable excesses were reported also in old works with deep samples taken from the Stromlo-APM survey and the LCRS (Loveday 1997, Lin et al. 1996, respectively). Finally, we find an excellent agreement in almost the entire absolute magnitude range (at least for $M_r > -21$) with Blanton et al. (2005b), who used a local SDSS galaxy sample to explore the very faint-end of the LF.

- We report a faint-end upturn in the ALHAMBRA LFs at $M^A \sim -17$. This upturn implies that the ALHAMBRA LFs cannot be fitted by a single Schechter function. A few works have claimed the existence of faint-end upturns in the field LF before, mainly when analyzing very low-redshift samples or when splitting the samples into different galaxy types. A similar excess is relatively accepted in the cluster LF. This is a very remarkable preliminary result that must be confirmed in future works.

5

Estimating Photometric Stellar Masses in ALHAMBRA

5.1 Introduction

In Chapter 4, we presented a preliminary study on some fundamental statistical properties of the ALHAMBRA survey at low redshift, including galaxy number counts, color and magnitude distributions and luminosity functions. These are the first-order statistics that characterize a galaxy population, as they are all derived from the luminosity and the redshift of individual galaxies. ALHAMBRA was designed to shed light into the physical processes involved in shaping the evolution of galaxies in the last $\sim$ 7 Gyr of cosmic time. Looking into these fundamental processes requires information about other more fundamental galaxy properties, such as stellar mass, age or metallicity. This chapter is dedicated to describing a methodology for determining photometric stellar masses (and other stellar population properties) in ALHAMBRA.

It is today accepted that the visible mass constitutes a small fraction of the total mass of the Universe, which is dominated by dark matter. Zwicky (1937) was the first to claim the need for an invisible matter in order to explain the circular velocities of galaxies in the Coma cluster using the virial theorem. Several other proofs of the existence of dark matter came up subsequently, including rotation curves of spiral galaxies (Rubin et al., 1980) or the anisotropies in the Cosmic Microwave Background (Bennett et al. 1996 with the Cosmic Background Explorer, COBE). Today, the majority of the Astrophysics and the Cosmology community believes that galaxies are embedded in dark matter halos, which extend well beyond the luminous baryonic cloud and contains at least an order of magnitude more mass than the galaxy itself. The formation of galaxies is, of course, tightly connected to the formation of dark matter halos. We expect baryons to cool and collapse into the potential well of the halo much faster than non-interacting dark matter particles. The relation between halos and galaxies in the context of large-scale structure is therefore a fundamental aspect in the study of the formation and evolution of galaxies. In order to investigate this halo-galaxy connection, reliable methods for characterizing both the mass of the halo and the baryonic mass of the galaxy are required.

The properties of halos have been probed with several indirect methods, all of them presenting

important limitations, especially at very large scales (of a few times the radius of the visible component). These methods include weak and strong lensing techniques, rotation velocity of galaxies, X-ray temperature in clusters or satellite dynamics of isolated galaxies. For the luminous matter, which is the main subject of this chapter, these mainly kinematic/dynamical measures are not effective (precisely due to the presence of dark matter). In fact, before the missing mass problem was raised, the use of these methods lead to incorrect mass-to-light ratios (M/L, i.e. the amount of mass per unit luminosity), of the order of a few hundreds. We have to rely, therefore, on galaxy luminosity as a surrogate of stellar mass. Galaxy luminosities have important limitations as well, as they typically trace the mass of particular stellar populations, mainly the younger more luminous stars, instead of that of the entire system. In addition, and especially in the bluer optical wavelengths, they are strongly affected by the presence of interstellar dust, which can bias the stellar mass estimation remarkably. Measuring the stellar mass requires, therefore, understanding the physics of stars, gas and dust along with the initial conditions at the birth of a galaxy. All these elements determine the shape of a galaxy SED, which is the observable that we can measure. Modeling all the processes involved in shaping the SED of a galaxy, the so-called stellar population synthesis (SPS), is, of course, one of the most fundamental and ambitious challenges in Astrophysics.

The aim of the SPS is to connect the physical properties of galaxies with the observed SEDs. The first attempts to interpret the integrated light (colors, line indices, and mass-to-light ratios) that we observe from galaxies appear in the 1970's and early 1980's. In these pioneering works, trial and error analyses and other similar techniques were used to reproduce the SED of galaxies by means of a linear combination of individual stellar spectra from different type of stars in the H-R diagram. A very early example of this approach can be found in Spinrad & Taylor (1971), where a library of stellar spectra from main-sequence, subgiants and giant stars in the solar neighborhood is used to synthesize model SEDs for the nearby galaxies M31, M32 and M81. However, these first models had too many free parameters and failed to provide a convincing physical interpretation of the galaxy SED. In the last decades, much effort have been devoted to improving the modeling by characterizing each of the elements that determine the shape of a galaxy SED. Ever since the works of Tinsley (1978) or Bruzual (1983), to name same of the first, the SPS includes a model for the initial distribution of stellar masses, i.e. the Initial Mass Function (IMF), evolutionary tracks for stars with different mass, a prescription for dust attenuation and chemical enrichment and a model for the star formation history (SFH) of the galaxy. Unfortunately, and despite the major advances made in the last decades, the SPS is still severely limited. Some of the largest uncertainties, in this sense, are related to determination of the IMF or the characterization of relatively unknown phases in the evolution of stars that can contribute extensively to the galaxy luminosity at certain wavelengths (i.e. horizontal branch, post-asymptotic giant branch phase etcetera). A detailed discussion on the most dominant uncertainties in the SPS can be found in Conroy et al. (2009).

In this chapter, we use the Flexible Stellar Population Synthesis code (FSPS, Conroy et al. 2009) to provide a preliminary SPS analysis in ALHAMBRA using the entire set of 20+3 filters and focusing on the determination of M/Ls and photometric stellar masses. This work is intended to lay the foundations for a complete SPS analysis that includes the determination of several other important galaxy properties such age, metallicity or SFH in ALHAMBRA. The chapter is organized as follows.

We will start, in Section 5.2, by providing a basic review on the main aspects of the SPS, including the modeling of single stellar populations (SSPs) and composite stellar populations (CSPs), and describing the FSPS code. In Section 5.3, we will present a preliminary grid of FSPS galaxy models, discuss on model-data agreement and describe the SPS fitting methodology in ALHAMBRA. In Section 6.3, we presents our results on SPS fitting, M/Ls and stellar masses. Finally, we will summarize the main results of this chapter in Section 5.8.

5.2 A basic review on Stellar Population Synthesis. The FSPS code

In order to estimate ALHAMBRA photometric stellar masses we use the Flexible Stellar Population Synthesis code, created by C. Conroyy. FSPS is a flexible SPS package that allows the user to compute simple stellar populations for a range of IMFs and metallicities, and with a variety of assumptions regarding the average properties of the stellar population. From these SSPs we can generate composite stellar populations with varying star formation histories and dust attenuation prescriptions. The code is written in FORTRAN and is publicly available at https://www.cfa.harvard.edu/∼cconroy/FSPS.html.

In this section, we briefly outline the steps involved in the SPS as implemented in FSPS and describe the main parts of the code itself. This section is based on Conroy et al. (2009).

5.2.1 SSP modeling

The motivation behind the SPS modeling is to interpret the integrated light (colors, line indices, mass-to-light ratios) that we observe from galaxies. Ideally, we want to determine what mix of stars gives rise to the observations. However, the physical processes involved in shaping the energy output, or the SED, of a galaxy are numerous and complex, so several assumptions are needed. The first step in the SPS is the modeling of single stellar populations, which are the building blocks of the SPS. A SSP is defined as a set of stars with the same age and metallicity. The spectrum, $S_{SSP}(t, Z)$, from one of this coeval set of stars of metallicity Z at time t after birth can be described as

$$S_{SSP}(t, Z) = \int_{M_i^l}^{M_i^u(t)} \Phi(M_i)\Lambda[L_{\rm Bol}(M_i, Z, t), T_{\rm eff}(M_i, Z, t), Z]\, dM_i, \tag{5.1}$$

where M_i is the initial mass of the star, M_i^l and M_i^u are lower and upper mass cutoffs, Φ is the IMF, Λ is a spectrum from the stellar library, and $L_{\rm Bol}$ and $T_{\rm eff}$ are the bolometric luminosity and effective temperature of a star of mass M_i and metallicity Z. In essence, what we do in Equation 5.1 is simply to add the contribution of each individual star in order to construct the spectrum of the entire single-age single-metallicity stellar population. Let us discuss in more detail the elements of Equation 5.1.

- **The IMF**

  The IMF, represented by Φ in Equation 5.1, is the distribution of stars per stellar mass bin along the zero-age Main Sequence (MS). The MS is a band on the Hertzsprung-Russell (HR) diagram along which $\sim 90\%$ of visible stars lie, while producing energy by the conversion of hydrogen to

helium by nuclear fusion in their cores. The zero-age MS is the time when a star first joins the main sequence on the HR diagram. Consequently, the IMF is the number of stars per volume (n) that were originally created per mass bin in a galaxy. Namely:

$$\Phi(logM) = \frac{dn}{dlogM}. \tag{5.2}$$

This is the way the IMF was originally defined by Salpeter (1955). Ever since this pioneering work, the IMF has been derived from the LF of stars in our Galaxy (or in the Large Magellanic Cloud and the Small Magellanic Cloud at most). In Salpeter (1955), the IMF is found to be reasonably well fitted by a power law with an exponent $\alpha \sim 2.35$, where $\Phi(logM) \propto (M/M_\odot)^{-\alpha}$, in the mass range $0.5 \lesssim M/M_\odot \lesssim 10$, where $M_\odot$ is the solar mass. Much effort has been devoted to estimating the IMF after this first study (Scalo, 1986, 1998; Kroupa, 2001; Chabrier, 2003; van Dokkum, 2008). Today, there is some consensus that the IMF follows a power law for masses $M \gtrsim 1M_\odot$, with an exponent that seems not to deviate too much from the Salpeter value (Scalo 1986 claims that $\alpha = 2.7$). In sub-solar mass ranges ($M \lesssim 1M_\odot$) a departure from this behavior, as firstly proposed by Miller & Scalo (1979), is accepted, although here the IMF is remarkably under-constrained. This is a rather important issue, as even though the contribution from these stars to the total luminosity of a galaxy is small, their contribution to the total mass might be large (note that these are also the longest lasting stars). General forms for the IMF have been proposed by several authors, where major discrepancies come from the low-mass end. As an example, the following expression was proposed by van Dokkum (2008):

$$\Phi \equiv \frac{dn}{d\ln M} = \begin{cases} A_l(0.5n_c m_c)^{-x} \exp\left[\frac{-(\log M - \log m_c)^2}{2\sigma^2}\right] & (M \leq n_c m_c) \\ A_h M^{-x} & (M > n_c m_c) \end{cases}, \tag{5.3}$$

with $A_l = 0.140$, $n_c = 25$, $x = 1.3$, $\sigma = 0.69$, and $A_h = 0.158$. Here, the characteristic mass m_c is the parameter to be varied. This expression with $m_c = 0.08$ is similar to other IMFs, such as the Chabrier (2003) IMF and the Kroupa (2001) IMF.

The IMF plays a crucial role in SPS modeling. In fact, the determination of the IMF is key to understanding galaxy evolution as the distribution of stellar mass in a galaxy determines the luminosity, the surface brightness, the chemical enrichment, etcetera. Note, for instance, that the evolution of passively evolving systems (those which are not undergoing any star formation event) is determined by the evolution of old stellar populations. In these populations, the turnoff point (at which stars abandon the MS to enter the giant branch) occurs approximately at $M = 1M_\odot$, which is within the range of masses at which the IMF starts to be uncertain. The logarithmic slope of the IMF at these values is therefore critical for the evolution of red passive galaxies. This is one of the main sources of uncertainty in SPS modeling. Another fundamental uncertainty of the SPS modeling derives from the assumption that the IMF is similar for any galaxy, independently of the environment (ubiquity) or the evolutionary state of the Universe (universality).

Throughout this chapter, we will use two different IMFs: the Salpeter (1955) IMF (hereafter, simply, Salpeter IMF) and the Kroupa (2001) IMF (Kroupa IMF). As mentioned above, the Salpeter IMF is simply a power law with $\alpha = 2.35$. The Kroupa IMF, is a step-wise power law with 3 different exponents: $\alpha = 2.30$ at $M/M_{\odot} > 0.5$, $\alpha = 1.30$ at $0.08 < M/M_{\odot} < 0.5$ and $\alpha = 0.3$ at $M/M_{\odot} < 0.08$. Mass-to-light ratios and stellar masses obtained with these IMFs are known to differ in ~ 0.15 dex.

- **Stellar Evolution tracks**

Once we have adopted an IMF, i.e., a model for the distribution of stellar masses at the zero-age MS, we need to assume an evolution for stars at different mass bins. This is what we call stellar evolution tracks. A number of stellar evolution libraries from different groups are publicly available (e.g., Bertelli et al. 1994; Girardi et al. 2000; Cioni et al. 2006; Schaller et al. 1992; Cassisi et al. 2000; Yi et al. 2001; Dotter et al. 2007). From these stellar evolution tracks, it is possible to build isochrones, which give us the position in the HR diagram (bolometric luminosity and the effective temperature) of a star of mass M and metallicity Z at certain time t after it entered the zero-age MS. Essentially, functions $L(M_i, Z, t)$ and $T(M_i, Z, t)$ in Equation 5.1 are directly derived from these isochrones.

The FSPS package allows the user to choose between two different sets of models: the latest library from the Padova group (Marigo & Girardi, 2007; Marigo et al., 2008) and the BaSTI library (Pietrinferni et al., 2004). Throughout this chapter, we use the former, which is set by default in the code. This library covers a range of initial stellar masses $0.15 \leq M/M_{\odot} \leq 100$ with ages between $10^{6.6}$ yr and $10^{10.2}$ yr. Stellar evolution tracks are available for equally-spaced metallicities in the range $10^{-4} < Z < 0.030$ ($\Delta log(Z) = 0.1$). The Padova models are supplemented in the very low mass range ($0.10 < M/M_{\odot} < 0.1$) with the non-evolving stellar models of Baraffe et al. (1998). As mentioned above, these stars do not contribute to the total luminosity of the galaxy, but they do to the total stellar mass.

- **Spectral libraries**

Spectral libraries provide the last ingredient needed to generate SSP models: the stellar spectra ($\Lambda[L_{\rm Bol}, T_{\rm eff}, Z]$ in Equation 5.1). Given a certain range of metallicities, we need a well-sampled set of models with varying bolometric luminosity and effective temperature which is representative of the stellar population in the Universe. The FSPS package uses the semi-empirical BaSeL3.1 library (Lejeune et al., 1997, 1998; Westera et al., 2002). This library covers a wavelength range $91\mathring{A} < \lambda < 160\mu m$ with a resolving power $\lambda/\Delta\lambda \approx 200 - 500$. The spectra of very hot stars ($T_{eff} > 50000K$), which are not included in the library, are approximated as pure blackbodies in FSPS. Also added to the BaSeL3.1 library are the spectra of thermally pulsating asymptotic giant branch (TP-AGB) stars, which dominate the radiation of intermediate-age stellar populations. These spectra are taken from the compilation of Lançon & Mouhcine (2002). For a detailed discussion on uncertainties related to the BaSeL3.1 library see Conroy et al. (2009).

Note, finally, that the mass of a coeval population of stars (SSP) can be obtained from the IMF, Φ, in the following way:

$$M(t) = \int_{M_i^l}^{M_i^u(t)} \Phi(M_i)\, M_{\rm evol}(M_i) dM_i \; + \; M_{\rm rem,} \tag{5.4}$$

where $M_{\rm evol}(M_i)$ is the evolved mass of a star of initial mass M_i and $M_{\rm rem}$ accounts for all the mass contained in remnants such as white dwarfs, neutron stars and black holes. $M_{\rm rem}$ is estimated in FSPS following the prescription of Renzini & Ciotti (1993). Namely:

$$\begin{array}{lcl} \text{If } M_i \geq 40 M_o & \longrightarrow & 0.5 M_i \text{ black hole} \\ \text{If } 8.5 M_i \leq M_i < 40 M_o & \longrightarrow & 1.4 M_i \text{ neutron star} \\ \text{If } M_i < 8.5 M_o & \longrightarrow & 0.077 M_i + 0.48 \text{ white dwarf} \end{array} \tag{5.5}$$

5.2.2 CSP Modeling

We have seen that SSPs can be derived from three main ingredients: an IMF, a set of isochrones and a spectral library. Equation 5.1 describes the spectral evolution of a coeval population of stars with metallicity Z. This description turns out to be rather incomplete for galaxies, where star formation can take place during long periods of cosmic time, or even occur in several outbursts. The time-dependent spectrum of a galaxy is therefore made of the contribution of stars of different ages. The star formation history is the main ingredient when generating CSPs. Following Conroy et al. (2009), the star formation rate of a galaxy (SFR, Ψ) can be expressed as follows:

$$\Psi(t) = \frac{(1-C)}{\tau} \frac{e^{-t/\tau}}{e^{-T_{\rm start}/\tau} - e^{-T_{\rm univ}/\tau}} + \frac{C}{T_{\rm univ} - T_{\rm start}}, \qquad \text{with } T_{\rm start} \leq t \leq T_{\rm univ}, \tag{5.6}$$

where Ψ is normalized so that one solar mass of stars is formed over the age of the Universe, $T_{\rm univ}$. The star formation is allowed to start at a time $T_{\rm start}$. Note that this expression permits the star formation to be described by two separate components. The first one decreases exponentially from $T_{\rm start}$ with a timescale given by τ, which is the time at which this SF component has dropped by a factor e. The second one is a constant SF mode, which is characterized by C. With Expression 5.6 and the SSP of Equation 5.1 we have everything in hand to generate a CSP, by simply adding the contribution of different SSPs with varying ages weighted by the SFH. Following Conroy et al. (2009), the spectrum of a composite stellar population with metallicity Z after a time t can be expressed as follows:

$$S_{\rm CSP}(t) = \int_0^t \Psi(t-t')\, S_{\rm SSP}(t', Z)\, e^{-\hat{\tau}_\lambda(t')}\, dt'. \tag{5.7}$$

Note that we have included an exponential factor $e^{-\hat{\tau}_\lambda(t)}$, characterized by a time-scale $\hat{\tau}_\lambda$, in order to describe the attenuation of starlight by dust. The transfer of radiation through the interstellar medium (ISM) is an important process that must be treated with the same accuracy that the production of stellar radiation itself. Throughout this work, we will use the dust model of Charlot & Fall

(2000), which is set by default in FSPS. In this model, the time-scale $\hat{\tau}_\lambda$ can be represented by two components in the following way:

$$\hat{\tau}_\lambda(t) \equiv \begin{cases} \hat{\tau}_1(\lambda/5500\ \mathring{A})^{-0.7} & t \leq 10^7\ \text{yr} \\ \hat{\tau}_2(\lambda/5500\ \mathring{A})^{-0.7} & t > 10^7\ \text{yr} \end{cases} \tag{5.8}$$

where $\hat{\tau}_1$ and $\hat{\tau}_2$ are the time scales for the attenuation of young and old populations, respectively. Conroy et al. (2009) find that the values of $\hat{\tau}_1 = 1$ and $\hat{\tau}_2 = 0.3$ are in good agreement with a wide range of observations.

5.2.3 FSPS Outline

We devote this section to providing a brief description of the FSPS package in order to facilitate the reader's comprehension of this chapter. The FSPS package contains several FORTRAN routines that allow the user to compute both SSPs and CSPs, along with some IDL routines that are intended to help the user manipulate the output. Important for this work, one of the main advantages of the FSPS package is that it is structured in a way that facilitates its integration in larger FORTRAN codes. In addition, the user can produce a large number of SPS models relatively fast. The main FSPS routines/modules are the following:

- **sps_vars.f90**: This module must be called at the beginning of every program, before any other routine is used. It sets up two types of variables: common variables and parameters. Parameters are set up once in this module and cannot be changed, while common variables can be modified in other routines. Below we list the most relevant common variables and parameters that we use in this work.

- **SSP_GEN**: This routine computes the SSP specified by a set of parameters and common variables. It outputs the time-dependent mass, bolometric luminosity and spectrum of an SSP model defined in sps_vars.f90.

- **COMPSP**: This routine takes the outputs of SSP_GEN (mass, bolometric luminosity and spectrum) as inputs and computes the CSP defined by the parameters set. It outputs spectra and/or magnitudes in all the bands implemented in the code.

- **GETMAGS**: This routine takes the redshift and the spectrum as inputs to produce magnitudes in any of the bands contained in the code.

As we mentioned above, parameters and common variables, which must be set up in module sps_vars.f90, define the SSP and CSP, as they describe the properties of the stellar populations. Below, we define some of the most relevant ones. The complete set of parameters and common variables is well documented in https://www.cfa.harvard.edu/∼cconroy/FSPS.html. Let us start with the common variables:

- **tuniv**: It sets the age of the Universe.

- **imf_type**: This common variable allows the user to choose between different IMF types. In particular, 4 different IMFs are available: Salpeter1955, Kroupa2001, Chabrier2003 and van Dokkum2008.

- **isoc_type**: This common variable defines the isochrone type. Isochrones are constructed from two different stellar evolution libraries: the Padova library (Marigo & Girardi, 2007; Marigo et al., 2008) and the BaSTI library (Pietrinferni et al., 2004).

- **dust_type**: Common variable defining the extinction curve for dust around old stars. The user can choose between 4 different prescriptions: a power law, the Milky Way extinction law parameterized by Cardelli et al. (1989), the Calzetti et al. (2000) attenuation curve and the attenuation curve models from Witt (2000).

The flexibility provided by FSPS when it comes to computing SPS models is reflected in the long list of parameters available to describe the stellar population (some of them associated with the above common variables). Some of the most important ones are:

- **zred**: This is the redshift at which the spectrum must be placed when computing the magnitudes. If $zred = 0$, magnitudes are rest-framed. If we set $zred$ at the redshift of the galaxy, observed-frame magnitudes will be output.

- **zmet**: Metallicity of the stellar population. The user can choose between 22 different values for the Padova isochrones, ranging from 0.0002 to 0.03, and 10 different values for the BaSTI isochrones, where $0.0003 < zmet < 0.04$. Note that the solar metallicity is assumed to be $Z_o \backsimeq 0.0190$.

- **sfh**: This parameter defines the star formation history of the system, normalized so that one solar mass of stars is formed over the age of the Universe. The *sfh* parameter can be set to zero for SSPs. The user can also provide a tabulated SFH or chose the five-parameter SFH that is implemented in the code. These SFH parameters are documented below.

- τ **(SFH)**: This SFH parameter is the time in Gyr that it takes for the SFH to decrease by a factor e. Permitted values cover the range $10^{-2} < \tau < 10^2$.

- **C (SFH)**: The fraction of mass formed in a constant mode of star formation ($0 \leq C \leq 1$).

- **tage (SFH)**: This parameter defines the age of the system, not weighted by the SFH. In other words, the SF starts at $t = 0$ and ends at $t = tage$.

- **fburst, tburst (SFH)**: These SFH parameters define the fraction of mass created by a single burst of SF and the time (or age of the Universe) at which this event occurs, respectively.

- **imf1, imf2, imf3**: These parameters are related with the common variable imf_type, described above, and must be set up if the piece-wise Kroupa IMF is selected. They represent the logarithmic slopes of the IMF over the ranges $0.08 < M/M_\odot < 0.5$, $0.5 < M/M_\odot < 1$ and $1 < M/M_\odot < 100$, respectively.

- **dust1, dust2**: Dust parameters describing the attenuation of young and old stellar light, respectively. These parameters are equivalent to $\hat{\tau}_1$ and $\hat{\tau}_2$ in the Charlot & Fall (2000) prescription of Equation 5.8. For the sake of simplicity, we will hereafter name these parameters $d1$ and $d2$.

- **dell, delt**: Logarithmic shift in bolometric luminosity end effective temperature, respectively, of the isochrone of TP-AGB stars.

- **sbss**: Number of Blue Straggler stars (BS) per unit Horizontal Branch (HB) stars. According to Conroy et al. (2009), plausible values are restricted to $sbss \lesssim 10$.

- **fbhb**: Fraction of HB stars that are blue. According to Conroy et al. (2009), a plausible range is $0 < fbhb \lesssim 0.5$.

For a more detailed discussion on each parameter, see Conroy et al. (2009).

5.3 SPS fitting in ALHAMBRA

This Chapter is devoted to developing a framework for obtaining fundamental physical properties from ALHAMBRA spectra. In particular, we will focus on the determination of ALHAMBRA photometric stellar masses using the FSPS code, which was briefly described in the previous section. At this point, it is important to make the following remark. The ALHAMBRA survey will extensively benefit from the contribution of different works that pursue similar scientific goals with different aproaches. In this sense, the SPS analysis that we propose here using FSPS can complement any similar analysis that uses BPZ (which is the main photo-z code in the collaboration) or even other approaches. It is precisely this synergy between different approaches that can lead to an optimal estimation of galaxy properties in ALHAMBRA, and, more generally, to an optimal scientific exploitation of the survey.

We will restrict ourselves to the same i^A-band selected low-z samples that we defined in Chapter 4. We believe this is the natural step to follow up the statistical analysis that we presented previously. The method that we propose here for SPS fitting in ALHAMBRA can be, however, trivially extended to higher redshift bins. As shown in Table 4.1 from Chapter 4, Sample1 was limited to $i^A = 20$ ($1,290$ objects), Sample2 to $i^A = 22.5$ ($4,700$ objects) and Sample3 to $i^A = 24.5$ ($10,098$ objects).

5.3.1 A FSPS grid of models

One of the main advantages of FSPS is that it allows the user to easily compute a large number of galaxy models. The goal is to build a library or a grid of models representative of the ALHAMBRA galaxy population. Once we have created the grid, we will fit the ALHAMBRA SEDs to find the best-fitting model in the grid for each galaxy. Each ALHAMBRA SED consists of a set of 23 magnitudes. FSPS can output magnitudes in a number of bands, including the SDSS bands. The ALHAMBRA filter system, however, is not included in the package, so the first thing we had to do was to add each ALHAMBRA filter/band (wavelength and transmission) to file allfilters.dat. In Chapter 4, we used synthetic SDSS bands to check that the main statistical properties of the ALHAMBRA distribution of galaxies are consistent with previous surveys such as the SDSS. Now, it makes all sense to use the

Common variable/ Parameter	Description	ALHAMBRA Grid
imf_type (CV)	Type of IMF	Salpeter, Kroupa
dust_type (CV)	Dust Prescription	Power law (Equation 5.8)
isoc_type (CV)	Stellar evolution library	Padova
$z_{\rm met}(P)$	Metallicity	[0.0006, 0.0020, 0.0061, 0.019]
$z_{\rm red}(P)$	Redshift	100 lineally-spaced values [0.1,0.3]
sfh (P)	SFH prescription	5-parameter SFH (Equation 5.6)
τ (P)	SFH e-folding time (Gyrs)	10 log-spaced values [0,10]
C (P)	SFH Constant mode	[0,0.25,0.5,0.75,1]
$t_{\rm age}$ (P)	Age of the system	$T_{\rm univ}$ at z_i
t_{burst}, f_{burst} (P)	Single SF burst	No bursts allowed
d2 (P)	Old stellar light attenuation	[0, 0.25,0.5,0.75,1]
d1 (P)	Young stellar light attenuation	$d2 * 3$
delt (P)	$\Delta log(T_{\rm eff})$ in TP-AGB	0
dell (P)	$\Delta log(L_{\rm bol})$ in TP-AGB	0
sbss (P)	Specific frequency of BS stars	0
fbhb (P)	Fraction of HB stars	0

Table 5.1 Summary of the main FSPS common variables (CV) and parameters (P) that define the ALHAMBRA grid. The grid comprises 1,000 models per redshift bin, i.e. a total of 1000 × 100 = 100,000 models.

complete set of narrow filters and take advantage of the excellent spectral coverage of ALHAMBRA to derive photometric stellar masses.

Building a grid of models with FSPS is equivalent to selecting the most relevant parameters and common variables and choosing a range of variation and a convenient sampling for each of them. Several grids have been generated during the creation of this work. In Table 5.1, we define the FSPS model grid used in this chapter. The range of variation chosen for each parameter is consistent with Conroy et al. (2009). Note that, throughout this work, we will use both a Salpeter IMF and a Kroupa IMF, so we have generated 2 different grids. The parameters shown in Table 5.1 define a set of 1,000 models for each redshift bin. An important assumption that we make is that $t_{\rm age} = T_{\rm univ}(z_i)$, which means that the age of each system equals that of the Universe at the corresponding redshift bin. We have divided our redshift range ($0.1 < z < 0.3$) into 100 bins, making a total of 100,000 models. Important for this work, the uncertainty in redshift in the SPS fitting due to this binning is considerably smaller than the expected photo-z uncertainty in ALHAMBRA ($\Delta z \approx 0.015 * z$).

At this point, it is convenient to take a look at the models in the grid and see how varying the main parameters shown in Table 5.1 modifies the shape of the resulting SED. For the SFH, we use the 5-parameter prescription of Expression 5.6, assuming that $t_{\rm age}$ is fixed at $t_{\rm age} = T_{\rm univ}(z_{\rm i})$ and not allowing for SF bursts (t_{burst}=0 and f_{burst}=0). The SFH is therefore defined by the e-folding time τ and a constant SF mode, which is characterized by C. In Figure 5.1 we show the effect of varying the first of these parameters, τ, in the SED of ALHAMBRA rest-frame ($z_i = 0$) models, for two different metallicities: $Z = 0.0006$ (sub-solar, left) and $Z = 0.019$ (solar, right). In order to isolate this effect, we assume that no constant SF mode or dust is present, so the SFH is defined by an unattenuated decaying SF. The flux shown in Figure 5.1, F_λ, is expressed in units $ergs/cm^2/s/\mathring{A}$ normalized at $\lambda = 10000\ \mathring{A}$. Note that by increasing τ, we can slow down the decay in the SF, so we can generate

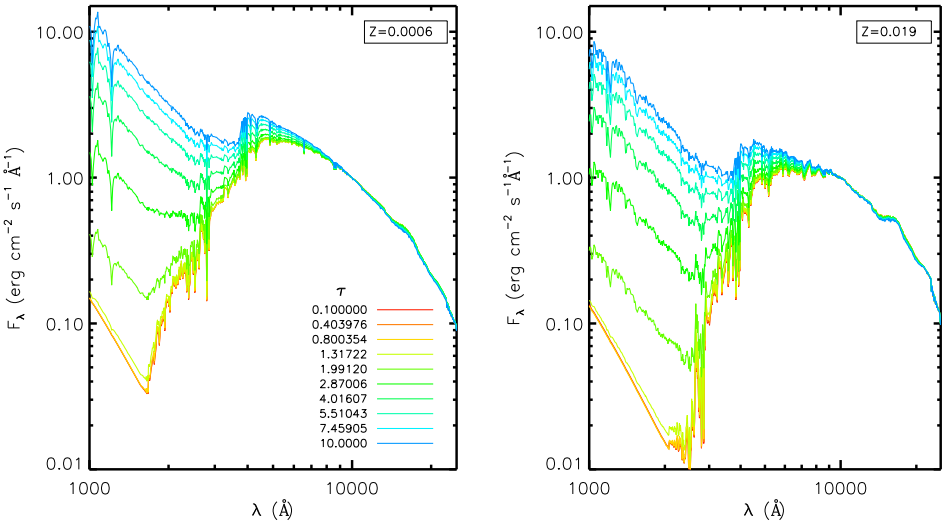

Figure 5.1 Spectra of ALHAMBRA models with varying SFH e-folding time, τ, for two different metallicities: sub-solar ($Z = 0.0006$) and solar ($Z = 0.019$). No dust ($d1 = 0$ and $d2 = 0$) and a single decaying mode for the SF are assumed ($C = 0$). The flux, F_λ, is normalized at $\lambda = 10000$ $\mathring{A}$. Note that by increasing τ we can generate increasingly younger stellar populations. The SEDs of young stellar population present a characteristic strong emission at short wavelengths.

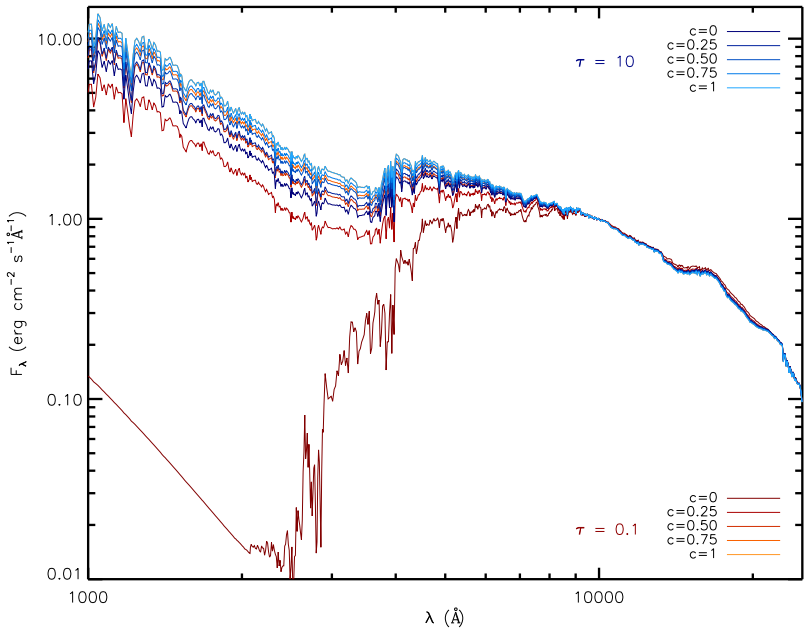

Figure 5.2 The effect of adding a constant SF mode in two sets of ALHAMBRA models, where no dust attenuation is assumed. In blue and red colors, a young stellar population ($\tau = 10$) and an old stellar population ($\tau = 0.1$) are shown, respectively.

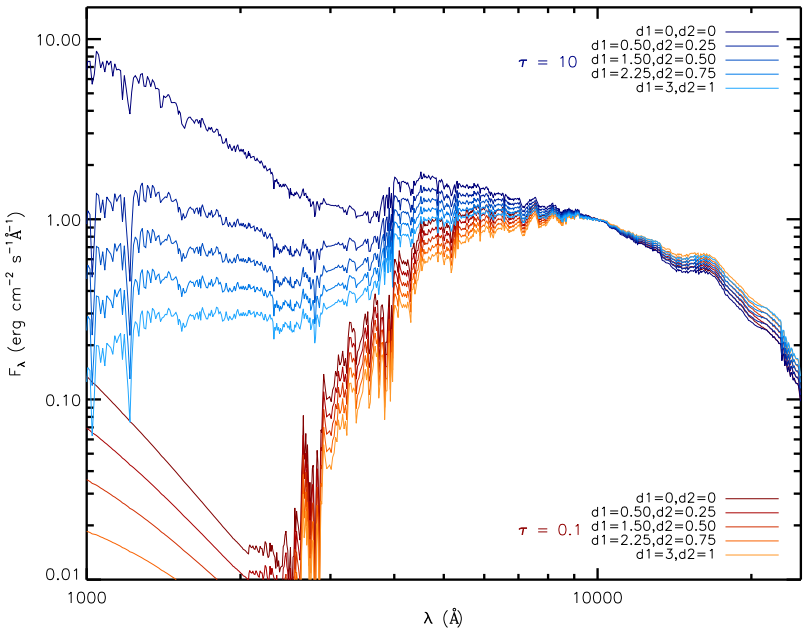

Figure 5.3 The effect of dust in two sets of ALHAMBRA models, where a single decaying mode for the SF is assumed ($C = 0$). In blue and red colors, a young stellar population ($\tau = 10$) and an old stellar population ($\tau = 0.1$) are shown, respectively, for varying dust attenuation. Parameters $d1$ and $d2$ represent the dust extinction around young and old stars, respectively. The flux, F_λ, is normalized at $\lambda = 10000$ $\mathring{A}$. Dust extinction has a stronger effect in the younger system.

increasingly younger stellar populations. This translates into a progressively larger energy output towards the bluer range of the spectrum, spanning two orders of magnitude in flux. In the near-IR the variation is less pronounced. Important for this work, our choice of parameters implies that we are so far not covering extremely blue systems such as young starburst and very young stellar populations (< 25 Myr) . The energy output of this type of systems relative to the red side of the spectrum could exceed in at least one order of magnitude that of the bluest model in Figure 5.1. It might be necessary in following versions of the ALHAMBRA grid to modify $t_{\rm age}$ and especially allow for SF bursts to represent these galaxy populations.

The other parameter that defines the SFH in the ALHAMBRA grid is the parameter C, which characterizes a constant SF mode. In Figure 5.2 we show the effect of adding such a mode to the SFH in two set of models, where again no dust extinction is considered and a solar metallicity is assumed. Represented by blue lines is a set of young stellar population models with $\tau = 10$ and varying C (which ranges from 0 to 1). As expected, the effect of adding a constant mode to such a young population is moderate. A much stronger impact is achieved when we apply this mode to a very old population ($\tau = 0.1$, red lines in Figure 5.2). Note that a SF mode characterized by $C = 0.25$ (25% of stellar mass form in this way in the galaxy) is enough to transform a very old and red passive-evolving galaxy model into a model that could represent a young system, such an irregular galaxy or a starburst. Figure 5.2 also illustrates the unavoidable degeneracy in the grid. This degeneracy also appears when considering other parameters such as dust or metallicity and must be carefully taken into account in

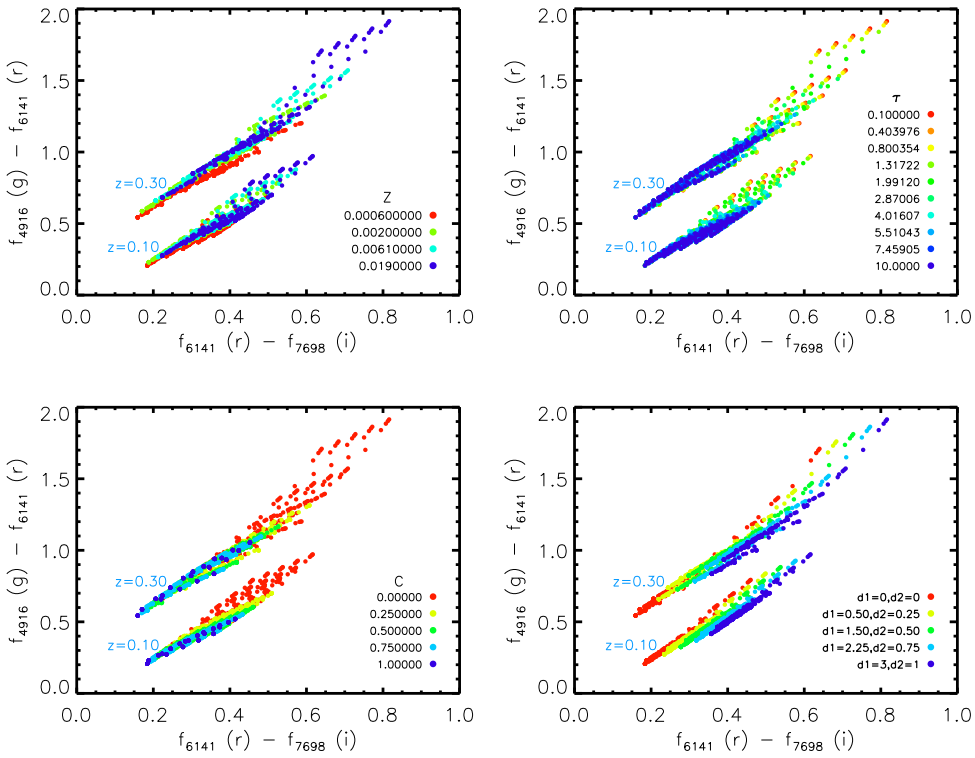

Figure 5.4 The effect of the variation of the FSPS grid parameters on the model colors shown in a f_{4916} (g) - f_{6141} (r) vs. f_{6141} (r) - f_{7698} (i) color-color diagram at two different redshift slices: $z = 0.10$ and $z = 0.30$. In each panel, models with different values of one of the following FSPS parameters are shown in different colors: Z, τ, C and the dust parameters ($d1$ and $d2$).

a rigorous SPS fitting.

In Figure 5.3 we can visually evaluate the effect of dust attenuation in a galaxy SED, in the same format as in Figure 5.1 and Figure 5.2. Again, in order to isolate the effect, a simply decaying SF and a solar metallicity are assumed. Also, we consider a young stellar population ($\tau = 10$, blue colors) and an old stellar population ($\tau = 0.1$, red colors). Note that following Conroy et al. (2009) we assume that the attenuation around young stars, $d1$, is proportional to the attenuation around old stars, $d2$, so that $d1 = 3d2$. The presence of dust in the inter-stellar medium has a stronger effect towards progressively shorter wavelengths, so the visual consequence is a flattening of the blue side of the spectrum. Part of the energy absorbed by dust in the blue and UV range is re-radiated in the IR, which is noticeable in Figure 5.3. The fact that the bluer, younger system in Figure 5.3 suffers more from dust extinction is a direct consequence of the fact this attenuation is stronger in the surroundings of young stars ($d1 = 3d2$).

In the next section, we will evaluate how well the FSPS grid presented above represent the ALHAMBRA galaxy population. A simple but effective way to do this is by using galaxy colors, which give us a rough idea of the shape of the SED. Before we proceed in this direction, it is convenient to see how the different parameters used in the ALHAMBRA grid distribute in a typical color-color

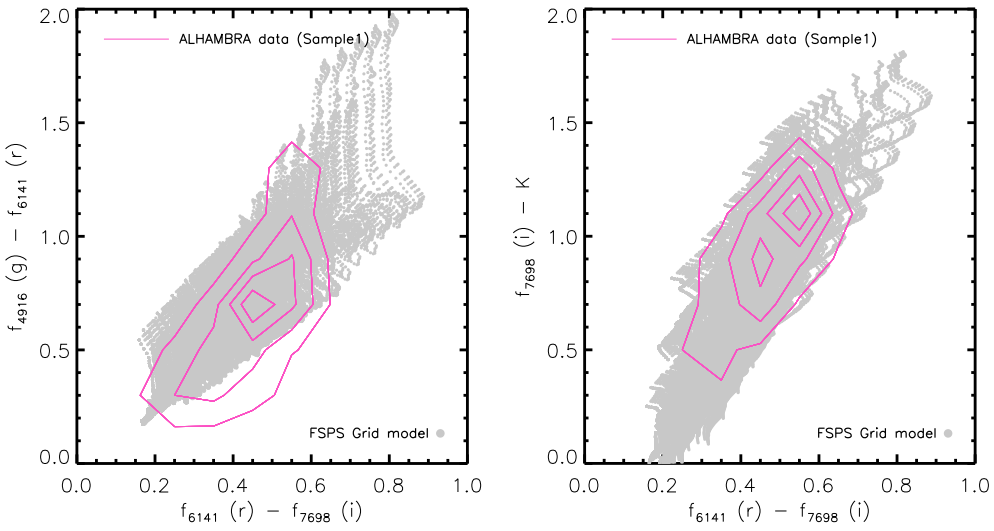

Figure 5.5 The entire FSPS set of models (grey dots) and the ALHAMBRA Sample1 data (magenta contours) shown in two different color-color diagrams, one representing the Optical range (left) and the other one the NIR (right). In the Optical range, where we use a f_{4916} (g) - f_{6141} (r) vs. f_{6141} (r) - f_{7698} (i) color-color diagram, the models cover reasonably well the ALHAMBRA data, except for a very blue region. In the NIR, we opt for a f_{6141} (r) - f_{7698} (i) vs. f_{7698} (i) - K diagram and find an excellent agreement between the data and the models.

diagram. This is illustrated in Figure 5.4 using three ALHAMBRA filters: f_{4916}, f_{6141} and f_{7698}. Note that the effective wavelength of these filters ($\lambda = 4916\ \mathring{A}, 6141\ \mathring{A}, 7698\ \mathring{A}$) are similar to those of the SDSS g, r and i bands so the diagrams shown in Figure 5.4 roughly reproduce a typical (r-i) vs. (g-r) color-color diagram. In each panel of Figure 5.4, two sets of models, one at $z = 0.1$ and the other one at $z = 0.3$, are represented by dots. Each of these panels focuses on one of the following parameters: Z, τ, C and the dust parameters ($d1$ and $d2$), so that colors indicate the variation of the corresponding parameter. Regarding the metallicity, Z, it seems clear that systems with sub-solar metallicities (especially $Z = 0.0006$) tend to have bluer r-i and g-r colors, according to Figure 5.4. A much clearer trend is found when looking at the e-folding SFH time, τ. Models with progressively shorter τ or, equivalently, with older stellar population, are increasingly redder, especially in g-r colors. Trivially, a similar behavior (even more pronounced) is obtained when increasing the fraction of stellar mass formed in a constant SF mode, C. As expected from Figure 5.2, the inclusion of this mode makes the system rapidly become red, which suggests that a finer sampling might be needed for this parameter. Finally, the presence of dust produces a reddening that is stronger in (r-i) colors.

5.3.2 FSPS models vs. ALHAMBRA data

One of the key steps in the SPS fitting is the generation of a family of models that is representative of the galaxy population. In Figure 5.5 we compare our FSPS grid with real ALHAMBRA data (Sample1) in both the optical range and the NIR. Note that it is not necessary to use K-corrected

magnitudes for either the models nor the real data, as FSPS is set up to produce magnitudes in the observed frame. This must be taken into account in Figure 5.5 where the reddening is amplified by the effect of redshift. In the left-hand panel of Figure 5.5 we show a f_{6141} (r) - f_{7698} (i) vs. f_{4916} (g) - f_{6141} (r) color-color diagram, as those shown in Figure 5.4, where contours represent real ALHAMBRA data and grey points symbolize grid models. The distribution of models is, as expected, more extended than that of the data. It is especially remarkable the population of extremely red models, which (as can be inferred from Figure 5.4) correspond to very old passive evolving objects in the highest redshift bins, which are not present in ALHAMBRA. Important for this work, the ALHAMBRA data appears to be relatively well covered by this preliminary grid, except for the very blue side of the diagram, in f_{4916} (g) - f_{6141} (r) colors. This discrepancy might be in principle due to errors in the photometry or in the photo-z estimation in ALHAMBRA, as we use a preliminary version of the catalogs in this work. We might also be lacking very young stellar populations in our grid (this is noticeable in Figure 5.1). In the future, we plan to allow for variations in the starburst parameters ($t_{\rm burst}$ and $f_{\rm burst}$) and investigate whether we can improve the performance of our grid.

In the right-hand panel of Figure 5.5, we modify the above color-color diagram in order to assess the performance of the grid in the NIR. In particular, we plot real data and models in a f_{6141} (r) - f_{7698} (i) vs. f_{7698} (i) - K color-color diagram. Again, the models cover a larger region of the color-color diagram, as compared to the data. In this case, however, the data seems to be very well covered by the models. Note that this was somehow expected, as the variability of typical galaxy SEDs is remarkably smaller in these wavelength ranges.

The color-color diagrams of Figure 5.5 are very helpful in the sense that they give us an idea of the performance of our model grid. However, they only provide information on a few filters, where photometric errors are not taken into account. In the next sections, we will be able to assess the performance of our model grid in a more accurate way, when we discuss the SPS fitting in ALHAMBRA.

5.3.3 SPS fitting: methodology

The method for finding the best-fitting model in the grid to a given ALHAMBRA galaxy is extremely simple. We first select the redshift slices (in the grid) which are compatible with the measured photo-z of the object, given the estimated precision of the photo-z's ($\Delta z \approx 0.015 * z$). This reduces the problem considerably, so the solution can be obtained faster. Then we simply perform a χ^2 minimization method over the ALHAMBRA 20+3 magnitudes. An important point here is that the absolute magnitudes output by the code are normalized so that 1 $M_\odot$ of stars is formed over the age of the Universe. The normalization of the model spectrum must be considered, therefore, as a free parameter that we marginalize. The χ^2 of each galaxy i when fitted with model j can be expressed as follows:

$$\chi^2_{i,j} = \sum_k \frac{(M_{\rm obs,i,k} - (M_{\rm model,j,k} + N_{\rm i,j}))^2}{\sigma^2_{\rm i,k}} \tag{5.9}$$

where $M_{\rm obs,i,k}$ is the ALHAMBRA absolute magnitude of galaxy i in filter k, $M_{\rm model,j,k}$ is the absolute magnitude of model j in the same filter as output by FSPS, $N_{\rm i,j}$ is the normalization of

galaxy i with respect to model j and $\sigma_{\rm i,k}$ is the uncertainty in the estimation of the ALHAMBRA absolute magnitude in filter k. In order to marginalize over $N_{\rm i,j}$ we can rearrange Equation 5.9 to obtain:

$$N_{\rm i,j} = \frac{\sum_k \frac{M_{\rm obs,j,k} - M_{\rm model,j,k}}{\sigma^2_{\rm i,k}}}{\sum_k \sigma^{-2}_{\rm i,k}} \tag{5.10}$$

This expression is used for calculating $N_{\rm i,j}$. Now $N_{\rm i,j}$ can be inserted into Equation 5.9 to obtain the following expression for χ^2:

$$\chi^2_{i,j} = \sum_k \frac{(M_{\rm obs,i,k} - M_{\rm model,j,k})^2}{\sigma^2_{\rm i,k}} - N^2_{\rm i,j} \sum_k \sigma^{-2}_{\rm i,k} \tag{5.11}$$

which is the expression that we have to minimize to find the best-fitting model of each galaxy within the grid. The goodness of a fit is approximately determined by the value of χ^2 for the best-fitting model, which is of course the model with the lowest χ^2 among all models in the grid compatible with the redshift of the object. This value is usually divided by the number of degrees of freedom of the object, which is simply the number of data points to fit (the number of filters available, 23 ad maximum) minus the number of free parameters (5 in our case: Z, τ, C, $d1$ or $d2$ and the normalization, $N_{\rm i,j}$). Note that not always do we have all the ALHAMBRA filters available for each object, as of course we suffer from some photometric incompleteness. All χ^2 values quoted in this chapter have been divided by the number of degrees of freedom.

5.4 Results

5.4.1 SPS fitting

In this section, we present our results on the SPS fitting in the ALHAMBRA Sample1, which is comprised by ~ 1000 objects brighter than $i^A = 20$. In following sections, we will extend our analysis to fainter magnitudes, where the fitting is of course more uncertain. A remarkable characteristic that we noticed in the ALHAMBRA data (in the version of the catalogs that we use in this chapter) is the extremely small photometric errors quoted in the catalogs for bright objects. In Figure 5.6, we illustrate this by showing photometric errors (σ) as a function of magnitude in the ALHAMBRA filter f_{7698} (i) for the entire sample (Sample3, $i^A < 24.5$), where errors are expressed in magnitudes. Note that at $i^A \lesssim 20$, photometric errors fall well below 0.01 magnitudes ($< 1\%$ error in flux), which is smaller than a reasonable calibration error. Errors seem to be underestimated towards brighter magnitudes, where this calibration error dominates. In Figure 5.7 we plot χ^2 as a function of the synthetic i^A band (used for selecting the samples) in the ALHAMBRA Sample3. The unnaturally small photometric errors translate into very high χ^2 values at the brighter end. Until we can use a more updated version of the ALHAMBRA data (an almost definitive version is available now), we will have to live with these caveats, which are, in any case, not critical for the purposes of this work, as we will show below.

In order to investigate the performance of our grid, we have carried out a visual inspection of

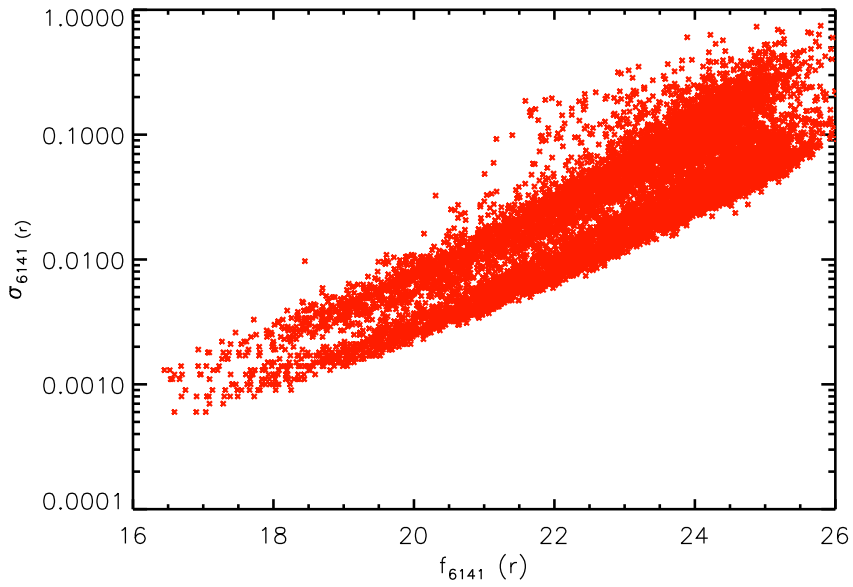

Figure 5.6 Photometric errors as a function of magnitude in the ALHAMBRA filter f_{7698} (i) for the Sample3 ($i^A < 24.5$). Errors at the bright end seem to be unnaturally small.

the fits in Sample1. From this visual inspection, it is fair to say that our preliminary grid performs more than decently, providing a reasonably good fit for most objects. It is still necessary to find a way to quantify the goodness of our fits in order to be able to select different subsamples or identify problems. A procedure commonly used when photometric errors are not totally reliable consists of a modification of χ^2. Namely, we can just replace the photometric error of each object by its magnitude (data). In the left-hand panel of Figure 5.7, we show the distribution of this modified goodness of fit parameter (χ^2_{mod}), which peaks somewhere in the vicinity of $\log_{10}(\chi^2_{mod}) \sim -5$. In the right-hand panel of Figure 5.8, we show the distribution of the maximum difference in magnitudes between the model and the data (in any filter) for Sample1. This distribution peaks around ~ 0.15 magnitudes, which we believe is a good result for this preliminary FSPS grid.

To alleviate the error problem without discarding all the error information we can opt to increase the photometric errors up to a reasonable value. In order to investigate this, we have proceeded by adding, for every object, an additional error equals to $\sigma = 0.03$ mag in any filter where $\sigma < 0.03$ mag. By doing this we are just somehow adding an artificial calibration error at the bright end, where it should actually dominate. Note that such a threshold might be conservative, in the sense that the resulting errors are maybe a bit larger than we might expect form ALHAMBRA. We have visually checked that varying this threshold within a reasonable range does not change our fits qualitatively. This is a common procedure in SPS fitting when some of the bands are problematic. An example of this can be found in Domínguez et al. (2011), where the extragalactic background light (EBL) is studied using different passbands. The authors follows the same method to increase the photometric errors in the DEEP2 bands in order to give some flexibility to the fitting. In what follows, we adopt this method in order to improve the SED fitting. We want to stress, however, the importance of measuring well-estimated photometric errors in the process of extracting galaxy properties such as

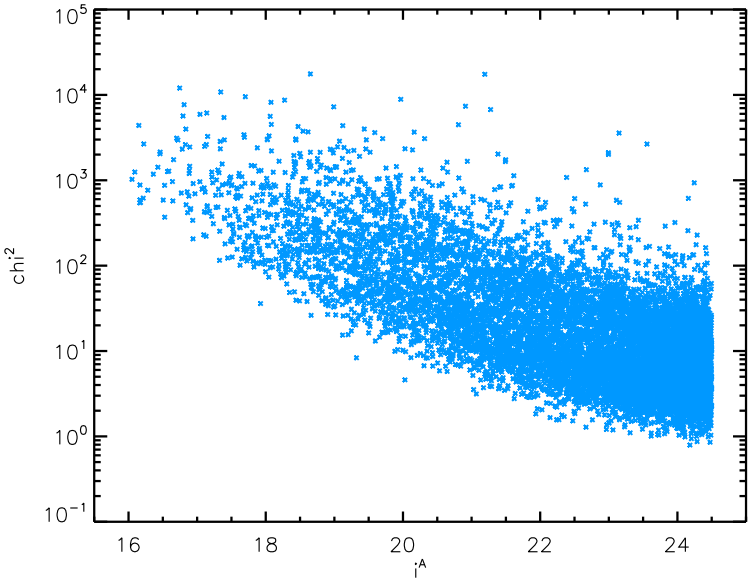

Figure 5.7 The χ^2 for the best-fitting model as a function of magnitude in the ALHAMBRA filter f_{7698} (i) for the ALHAMBRA Sample3 ($i^A < 24.5$). The unusually small photometric errors that we have, especially at the bright end, translate into high χ^2 values.

SFH, ages or metallicities from observed spectra. In fact, this is a critical aspect in the estimation of the error budget associated with a SPS analysis.

In the left-hand panel of Figure 5.8, we show the distribution of χ^2 in the ALHAMBRA Sample1, where photometric errors have been artificially increased in the way previously described. The χ^2 distribution now peaks somewhere around ~ 3 and is remarkably skewed to the right. On the right-hand side of Figure 5.8, we show the accumulated distribution of the χ^2. From our visual inspection, we have checked that fits are reasonably good up to (at least) $\chi^2 \sim 8 - 10$. Note that approximately 90% of objects in the sample have $\chi^2 < 8$. A threshold for good fits could be set at $\chi^2 < 6$, which roughly corresponds to a 1σ dispersion from the median value. This cut might turn out to be slightly restrictive, however, according to our visual analysis.

We will now take a look at some examples of real ALHAMBRA spectra and their corresponding best-fitting models within our preliminary FSPS grid. This will illustrate the meaning of the χ^2 values shown in Figure 5.9. In Figure 5.10, we present some examples of good fits. Each row in this figure corresponds to a different galaxy, with increasing goodness of fit, i.e. $\chi^2 \sim 0.6$, 2 and 2.3, respectively. In each example, the ALHAMBRA data is represented by crosses and the best-fitting model is shown in a red line. The fitting is shown in both magnitudes (not K-corrected), in the left-hand panel and in fluxes, in units of $10^{-17} f_\lambda$ in the right-hand panel. We have chosen galaxies of different spectral types using the BPZ preliminary classification: an elliptical galaxy (top), a spiral galaxy (middle) and a SF galaxy (bottom). Despite some slight discrepancies, the performance of the grid is very good for the examples shown in Figure 5.10, which are representative of objects below the peak of the χ^2 distribution in Figure 5.9. The χ^2, the redshift and the FSPS parameters of the corresponding best-fitting models are also shown for each galaxy.

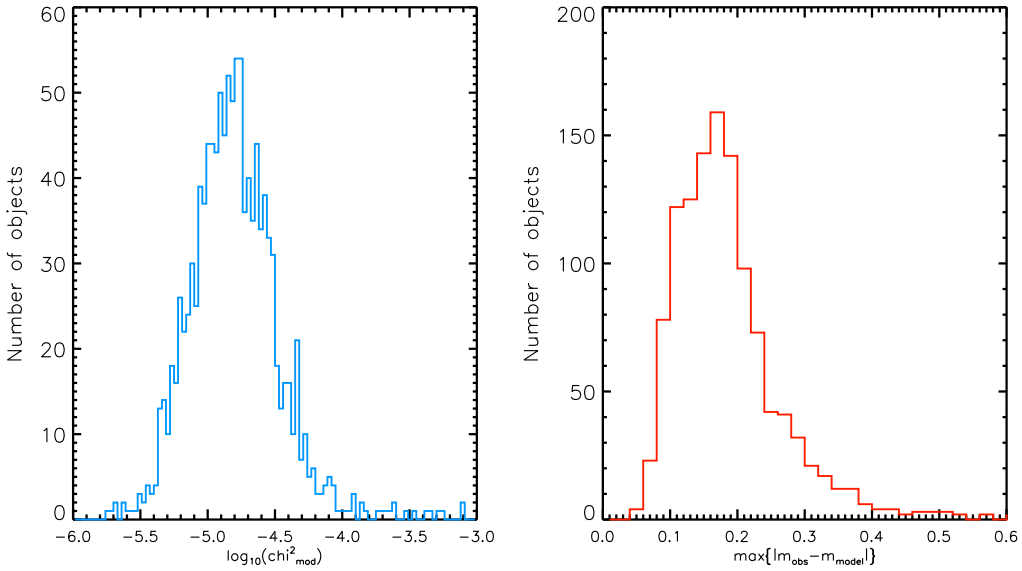

Figure 5.8 Left: The distribution of a modified χ^2 goodness of fit parameter (χ^2_{mod}) in the ALHAMBRA Sample1. In order to find the best-fitting model in the FSPS grid we have replaced the photometric error in Equation 5.11 by the estimated absolute magnitude of the object. Right: The distribution of the maximum difference in magnitudes between the best-fitting model and the data.

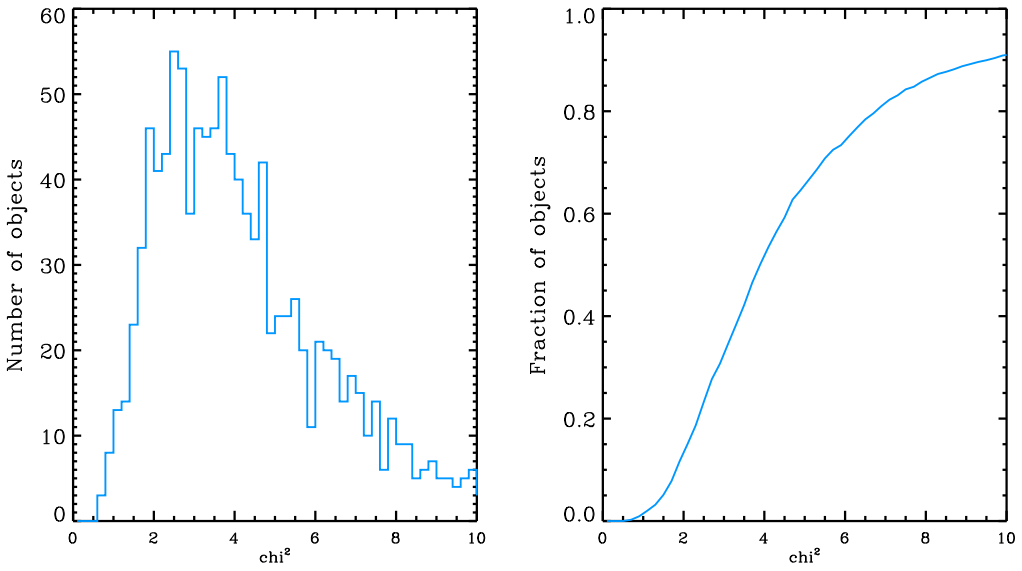

Figure 5.9 Left: The distribution of the χ^2 in the ALHAMBRA Sample1, after adding a constant 0.03 mag to the photometric error, σ, in any filter where $\sigma < 0.03$ mag. Right: The cumulative fraction of the χ^2 in the same galaxy sample.

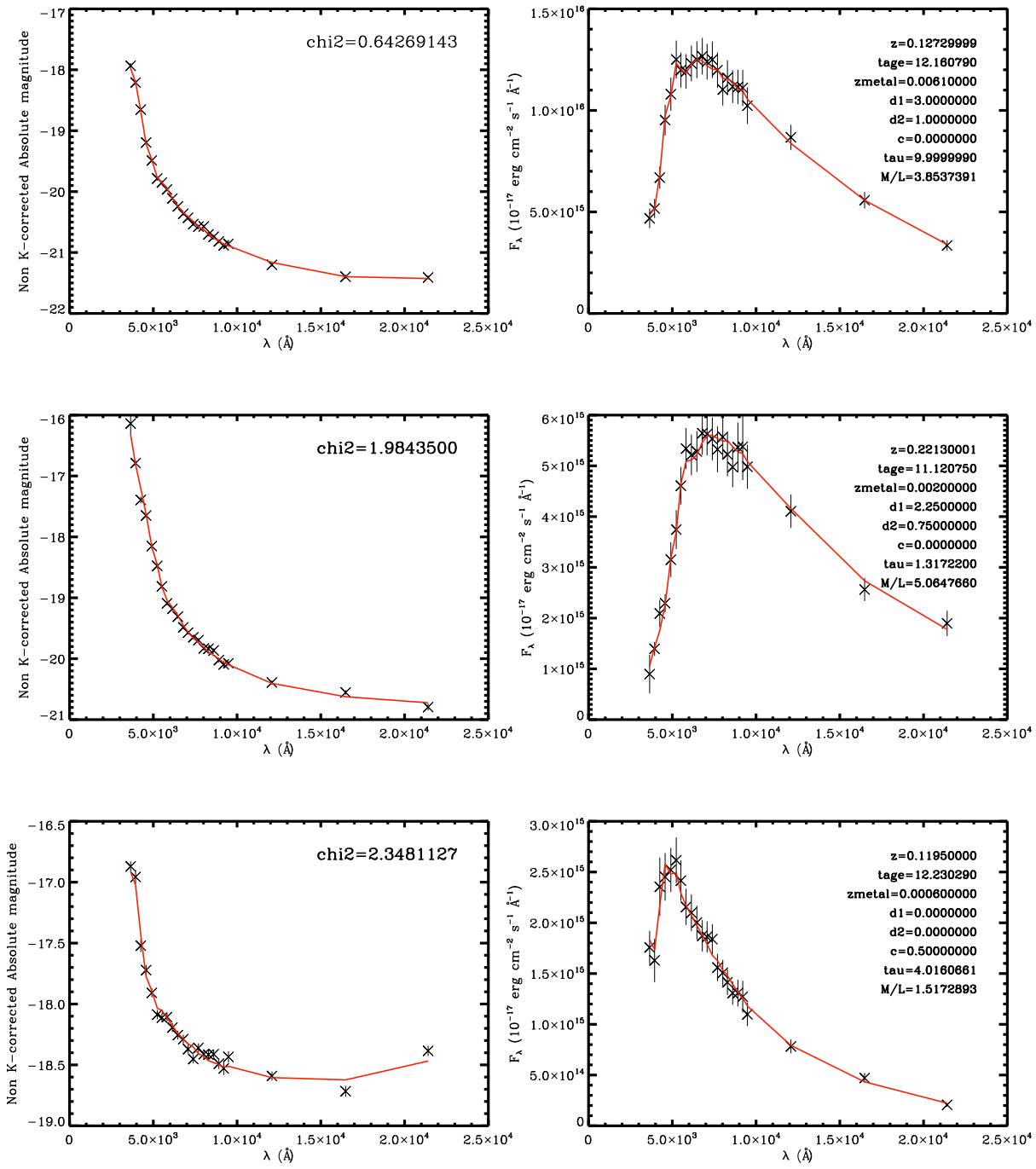

Figure 5.10 Examples of well-fitted real ALHAMBRA spectra (crosses) and their corresponding best-fitting models (red lines). In each row, the data and the model of a single object are shown in magnitudes (not K-corrected, left) and fluxes (in units of 10^{-17} f_λ, right). The examples shown correspond to an elliptical galaxy, a spiral galaxy and a SF galaxy, respectively, according to a preliminary BPZ classification. For each galaxy, the redshift, χ^2 and best-fitting FSPS parameters are also listed.

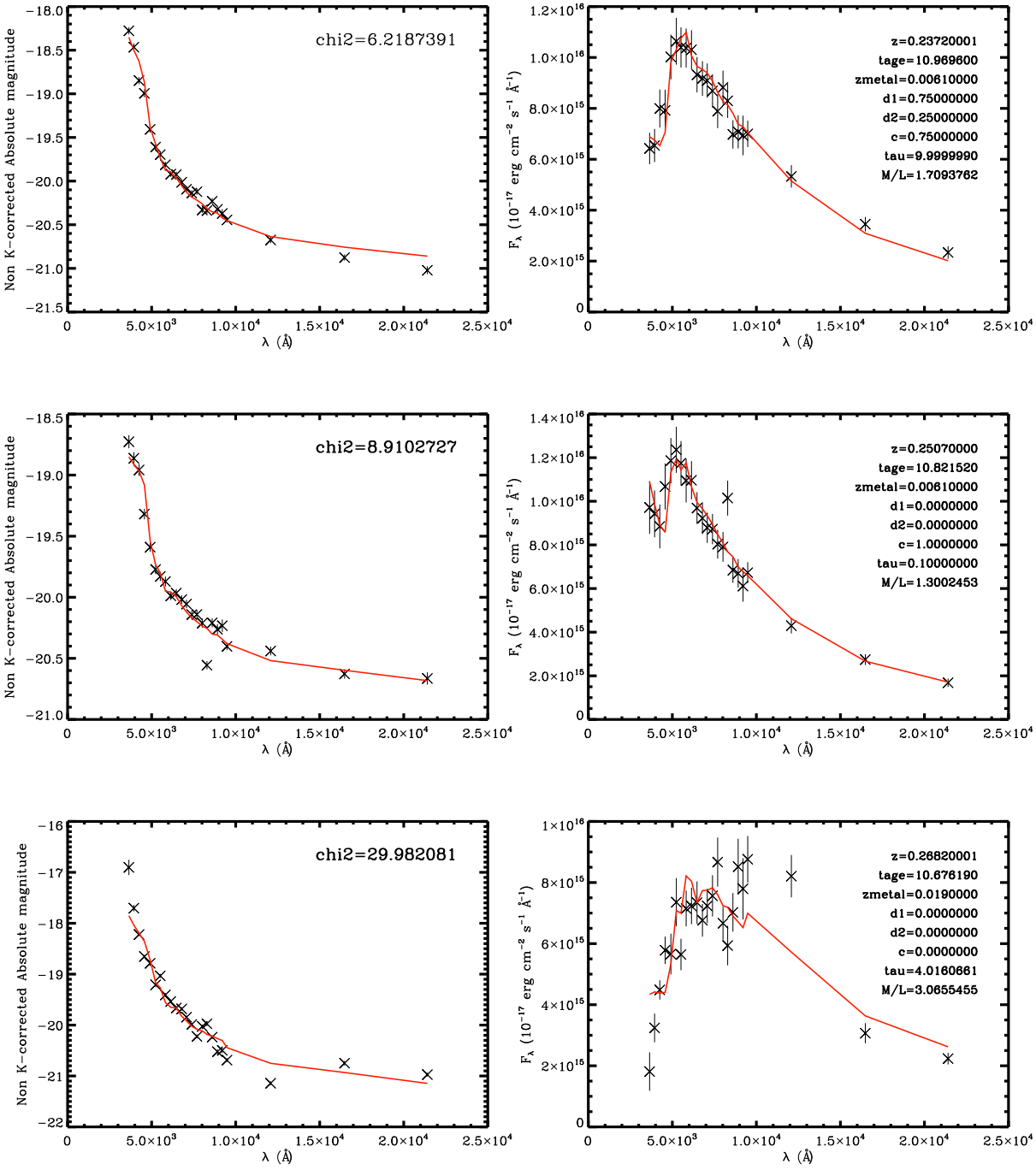

Figure 5.11 The same as in Figure 5.10 but showing examples of fits with larger χ^2 values. The first 2 rows correspond to SF galaxies with $\chi^2 = 6.21$ and $\chi^2 = 8.91$, respectively. Note that the fits are acceptable, despite the large χ^2. In the last row, we show an example of a catastrophic fit ($\chi^2 \sim 30$) to an elliptical galaxy.

In Figure 5.11, we show, in the same format as in Figure 5.10, some examples of ALHAMBRA galaxies for which our FSPS grid provides a worse fit. The first example corresponds to a SF galaxy with $\chi^2 = 6.22$. Even though we find some discrepancies in the NIR bands, the fit can still be considered reasonably good. In the second example, we show a galaxy with an even younger stellar population and a best-fitting model with $\chi^2 = 8.91$. Note that, again, the fit is not bad for such a large χ^2, especially considering the intrinsic uncertainties in the estimation of stellar masses. This example also illustrates a remarkable advantage of the ALHAMBRA filter system: the capacity to capture emission lines. Finally, we provide an example of a catastrophic fit with $\chi^2 \sim 30$, corresponding to an elliptical galaxy with, apparently, several emission lines.

5.4.2 Grid caveats

We have shown that the performance of the FSPS grid described in Section 5.3.1 is good for the majority of the galaxies in the ALHAMBRA Sample1 ($i^A < 20$). This is illustrated in Figure 5.10 and Figure 5.11. Let us now briefly discuss on the main aspects that we need to investigate in order to improve the fitting. Figure 5.12 is intended to give a rough idea of the ALHAMBRA filters where most discrepancies come form. In particular, we show the distribution of the maximum deviations across the ALHAMBRA spectrum, comprised by 20 optical filters and the JHK bands in the NIR. This figure shows that most discrepancies come from the very blue end of the ALHAMBRA spectrum. Frequently throughout the sample, as we have visually checked, the best-fitting model overestimates the flux in the bluest filters, by typically $\sim 0.1 - 0.2$ mag. At this point, we cannot give a definite explanation for this effect. As we are not really including very young populations and starbursts we might expect our models not to be blue enough in some cases. An inadequate variation range of other FSPS parameters, such as the dust parameters or metallicity, might also be causing these problems. We know, however, that the very blue filters are problematic in ALHAMBRA, so it is important to firstly check that the photometry is consistent in these filters. At the other end of the ALHAMBRA spectrum, we also encounter some problems in the reddest filter of the optical range, whereas in the NIR, the grid displays a good performance, as expected from the color-color diagram of Figure 5.5.

The photo-z catalogs used in this work, that where obtained using the BPZ method, provide a preliminary spectral classification of the ALHAMBRA galaxies. We have used this classification to analyze how the FSPS grid performs for different types of objects. In Figure 5.13 we show the χ^2 distribution for 4 different galaxy types in the ALHAMBRA Sample1: elliptical galaxies (black), spiral galaxies (red), irregular galaxies (green) and SF galaxies (blue). As we have checked through our visual inspection, elliptical and spiral galaxies are, in general, considerably better fitted than younger systems such as irregular galaxies or SF galaxies.

Several updates can be conceived in order to improve this preliminary FSPS grid. Namely, we need to increase the sampling rate in τ and the metallicity (also allowing for super-solar values). We should definitely include younger stellar populations at each redshift slice and explore the burst parameters. The best strategy in order to evaluate the effect of these updates in our grid is to take an individual galaxy representative of each of the spectral types discussed above and try to fit them using a broader parameter set. We hope that this type of fine tuning can help us improve the fitting and, especially,

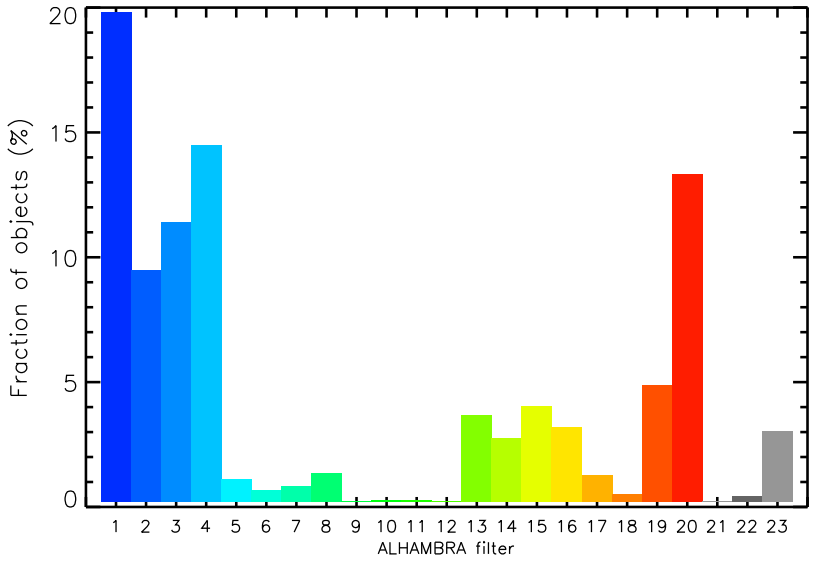

Figure 5.12 Distribution of the largest deviations across the ALHAMBRA spectrum. Most discrepancies come from the very blue filters and, in less extend, the redder filters in the optical range.

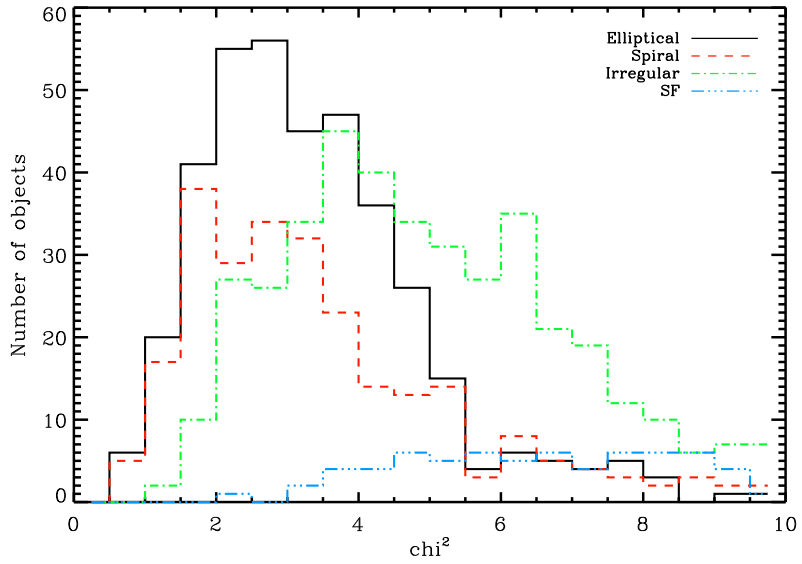

Figure 5.13 The χ^2 distribution for different spectral types as classified by BPZ: elliptical galaxies (black), spiral galaxies (red), irregular galaxies (green) and SF galaxies (blue). The FSPS grid displays a considerably better performance for older stellar populations.

solve the blue-end issue. Note however, that we have demonstrated that a very simple FSPS grid as that described in Section 5.3.1 can provide a reasonably good fitting in ALHAMBRA, at least for the brighter objects.

5.4.3 Mass-to-light ratios

Once we have found the best-fitting model to an ALHAMBRA galaxy in the FSPS grid, it is straightforward to extract the M/L of the model, and, subsequently, an estimation of the stellar mass of the object. Note that we have set FSPS to produce observed-frame magnitudes, so we have not K-corrected our magnitudes before hand to perform the fitting. In order to obtain K-corrected M/Ls we proceed as follows. We first fit the ALHAMBRA uncorrected magnitudes in the grid. The FSPS code provides the stellar mass of the model (the *smass* parameter, normalized to 1 $M_\odot$) and the absolute magnitudes in any of the ALHAMBRA filters. As these magnitudes are observed-frame, what we do is to take the magnitudes of the corresponding model at $z = 0$ in the grid. If we transform theses magnitudes into luminosities we can compute the M/L through any given filter. This is the method used by Bell et al. (2003) in order to obtain M/Ls in the SDSS by means of an SPS analysis. Note that by proceeding in this way, we are not only K-correcting our magnitudes but also including an estimation of the evolution correction, as the stellar population of the model at $z = 0$ has evolved from the corresponding redshift of the galaxy. It is convenient to remind that the spectrum output by FSPS is normalized so that 1 $M_\odot$ of stars formed during the lifetime of the system. However, the mass of the model, *smass*, that we take from FSPS, is smaller than 1 $M_\odot$ (typically $\sim 0.7\ M_\odot$). FSPS accounts for the fact that a remarkable fraction of the mass formed in stars is lost via stellar winds or supernovae and injected back into the interstellar medium. Finally, the mass-to-light ratio at a given filter k, M/L_k can be obtained from the following expression:

$$M/L_{\rm k} = \frac{smass}{10^{-0.4({\rm M_k}+48.60)}\left(\frac{c}{\lambda^2_{\rm eff,k}}\right)\ L^{-1}_{\odot,{\rm k}}} \tag{5.12}$$

where *smass* and M_k represent the stellar mass and the absolute magnitude of the best-fitting model at $z = 0$, respectively, $\lambda_{\rm eff,k}$ is the effective wavelength of filter k, $L_{\odot,{\rm k}}$ is the solar luminosity through filter k and c is the light speed. Unless otherwise stated, the results shown from now on were obtained with a Salpeter IMF.

In Figure 5.14, we show the distribution of the logarithm of the K-corrected M/L in Sample1, in 3 different ALHAMBRA filters: f_{4916} (g), f_{6141} (r) and f_{7698} (i). We have excluded unreliable fits by imposing the rather restrictive condition $\chi^2 < 6$, which includes 80% of the sample. We have compared our results with the color-dependent parametrizations of Bell et al. (2003), that we will hereafter name Bell03. Bell et al. (2003) obtained the following fits for the mass-to-light ratios in the SDSS bands as a function of the $z = 0$ K-corrected (g-r) color:

$$\begin{aligned}
\log_{10}(M/L)_{\rm g} &= -0.306 + 1.097(g-r)\\
\log_{10}(M/L)_{\rm r} &= -0.499 + 1.519(g-r)\\
\log_{10}(M/L)_{\rm i} &= -0.222 + 0.864(g-r)
\end{aligned} \tag{5.13}$$

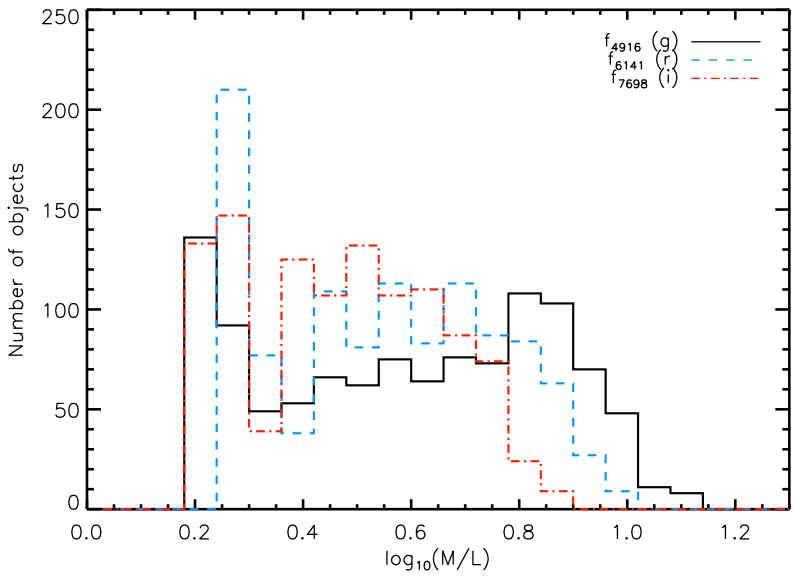

Figure 5.14 The distribution of K-corrected M/Ls in the ALHAMBRA Sample1 in filters f_{4916} (g), f_{6141} (r) and f_{7698} (i).

where a diet-Salpeter IMF was used. We have modified the above expressions by adding a constant 0.15 *dex* in order to make them consistent with a Salpeter IMF (Bell et al., 2003). In Figure 5.15 we show the logarithm of the K-corrected M/L as a function of the f_{4916} (g) - f_{6141} (r) color in the same filters used in Figure 5.14. Mass-to-light ratios have been obtained using both the FSPS code (contours, dots) and the prescriptions of Bell et al. (2003) (red line). We find a tight $\log_{10}(M/L)$-color relation in all filters, with a slope that is in good agreement with the SDSS prescription of Bell et al. (2003). However, there seems to be a shift in the zero point (of ~ 0.05 dex). Note that the comparison is obviously not totally fair, as the prescriptions were obtained using the SDSS broad bands and we use narrow-band filters. In addition, we find that the M/Ls tend to a plateau for very red colors, instead of increasing linearly (in logarithmic scale) with color. A similar behavior but less pronounced was found by Kauffmann et al. (2003b) in a sample of 122,808 SDSS galaxies.

Mass-to-light ratios in the synthetic g^A, r^A and i^A bands have also been computed in the ALHAMBRA Sample1. In order to do so, we have taken the magnitudes in these bands from the best-fitting model, which was fitted using the entire set of ALHAMBRA filters. This means that we have not properly fitted the synthetic bands, which might result in uncertainties in the estimation of the M/L. Using the $(g^{\rm A} - r^{\rm A})$ color, we have also obtained M/Ls following the prescriptions of Equation 5.13. Note that the agreement is now excellent, in both the slope and the zero-point. This is shown in Figure 5.16, in the same format as in Figure 5.15.

5.4.4 Photometric stellar masses

Once we have obtained mass-to-light ratios in a given filter, it is trivial to estimate the stellar mass (SM) of the galaxy. We just need to multiply our K-corrected M/L, obtained from the best-fitting

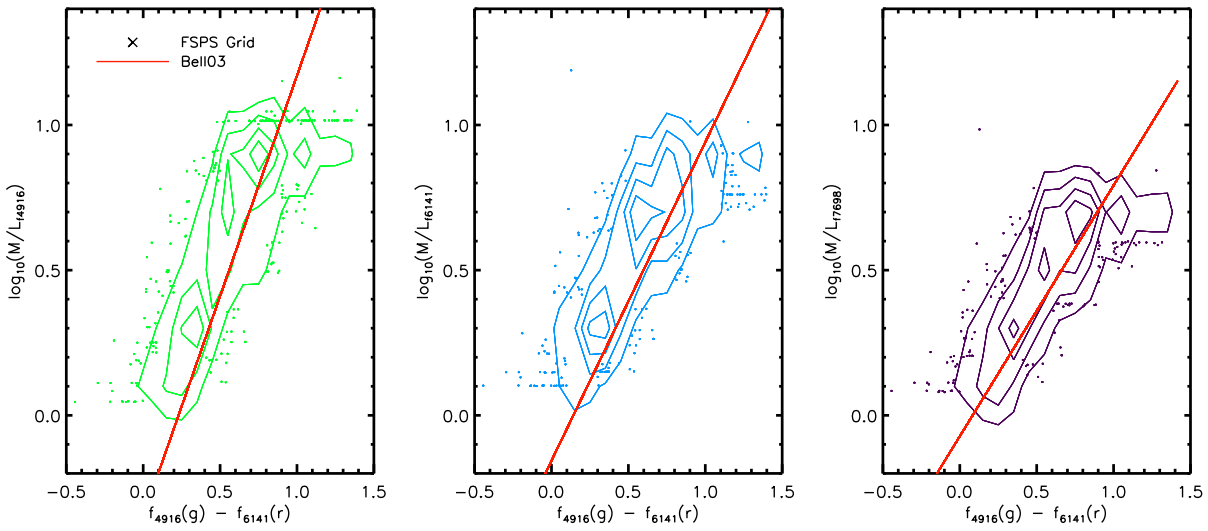

Figure 5.15 Mass-to-light ratios in the ALHAMBRA f_{4916} (g), f_{6141} (r) and f_{7698} (i) filters as function of the f_{4916} (g) - f_{6141} (r) color in Sample1. M/Ls obtained with the FSPS code are shown in crosses and the M/L-color relation of Bell et al. (2003) is represented by a red line.

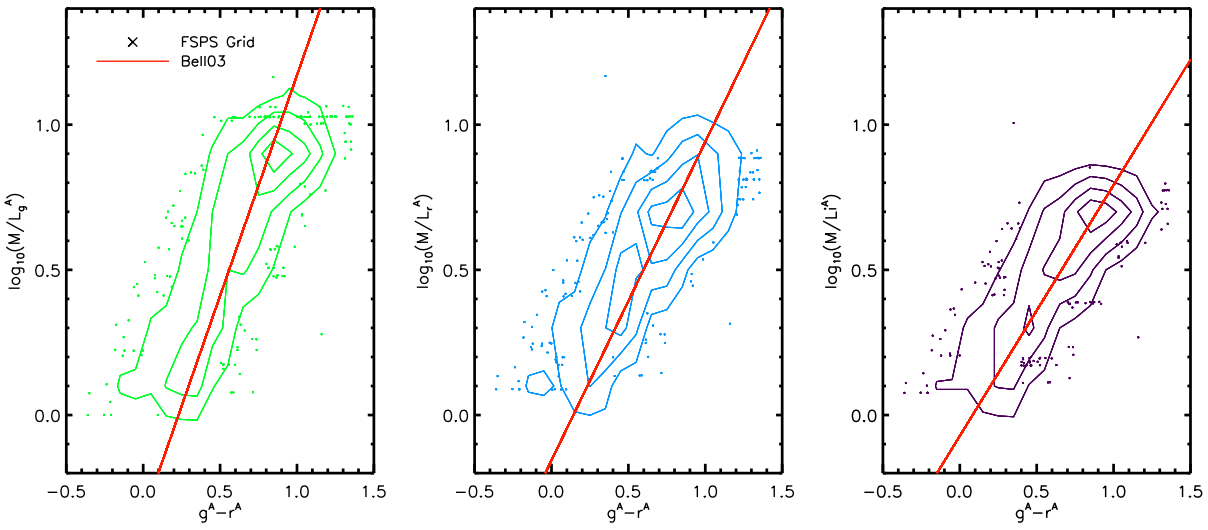

Figure 5.16 The same as in Figure 5.16 but using the ALHAMBRA synthetic $g^{\rm A}$, $r^{\rm A}$ and $i^{\rm A}$ bands.

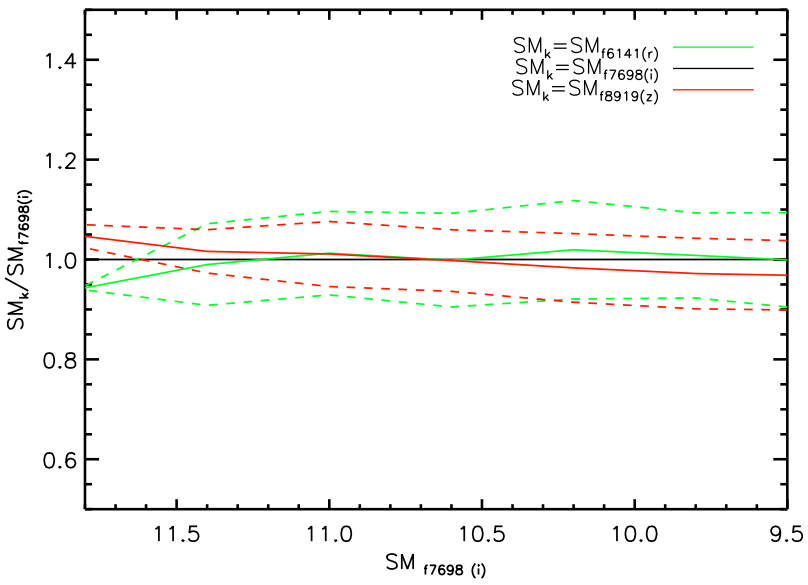

Figure 5.17 Discrepancy in the stellar mass estimated using filters f_{6141} (r), f_{7698} (i) and f_{8919} (z) as a function of the stellar mass, $SM_{\mathrm{f7698(i)}}$, in the ALHAMBRA Sample1.

model, by the real K-corrected luminosity of the galaxy in the corresponding filter. In principle, as we fit the entire spectrum from the optical range to the NIR we might expect to get an estimation of the stellar mass that is independent of the filter. In practice, however, errors in the fitting will make an impact in our stellar mass estimation. In this sense, it is clear from Figure 5.12 that taking a blue filter is not a good strategy. Note that a difference in ~ 0.15 mag between the model and the data in a given filter translates into a 40% uncertainty in the mass estimation. In Figure 5.17, we illustrate the discrepancy in the stellar mass estimation through filters f_{6141} (r), f_{7698} (i) and f_{8919} (z) - we have excluded the blue filters for the reason discussed above. In particular, we show in solid lines the ratio of the stellar mass obtained using each of the above filters, SM_{k}, to the stellar mass obtained through filter f_{7698} (i), $SM_{\mathrm{f7698(i)}}$, as a function of the stellar mass $SM_{\mathrm{f7698(i)}}$. The dashed lines represent the 1σ dispersion on both sides of the distributions. Figure 5.17 shows that the 3 estimates are in good agreement.

The main motivation of this work is to provide a method for estimating photometric stellar masses in ALHAMBRA, taking advantage of its spectral coverage and resolution, with 20 filters in the optical range plus the 3 bands in the NIR. In addition, it is very interesting to investigate how the estimates obtained with narrow filters differ from those obtained with broad band filters. In Figure 5.18 we provide a first approach to this study by showing, in the same format as in Figure 5.17, the discrepancy in the stellar mass obtained with the narrow f_{7698} (i) filter and the synthetic i^A band. This result must be taken with special caution because, as explained before, we have not properly fitted the synthetic bands. The mean of these distributions differs in $\sim 5\%$. This is an important aspect of this work that must be looked into in the future.

We have compared the distribution of our stellar masses with that from the nearby Universe. In

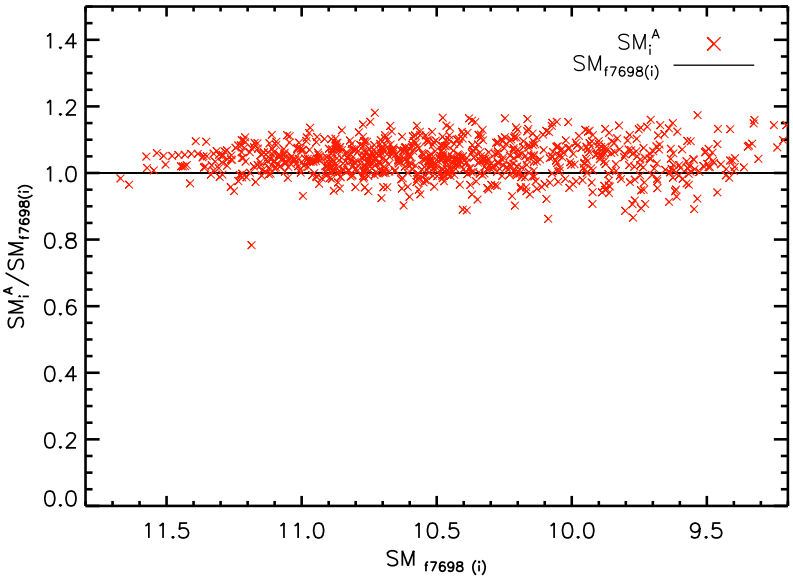

Figure 5.18 Discrepancy in the stellar mass estimated using the narrow f_{7698} (i) filter and the synthetic i^A band.

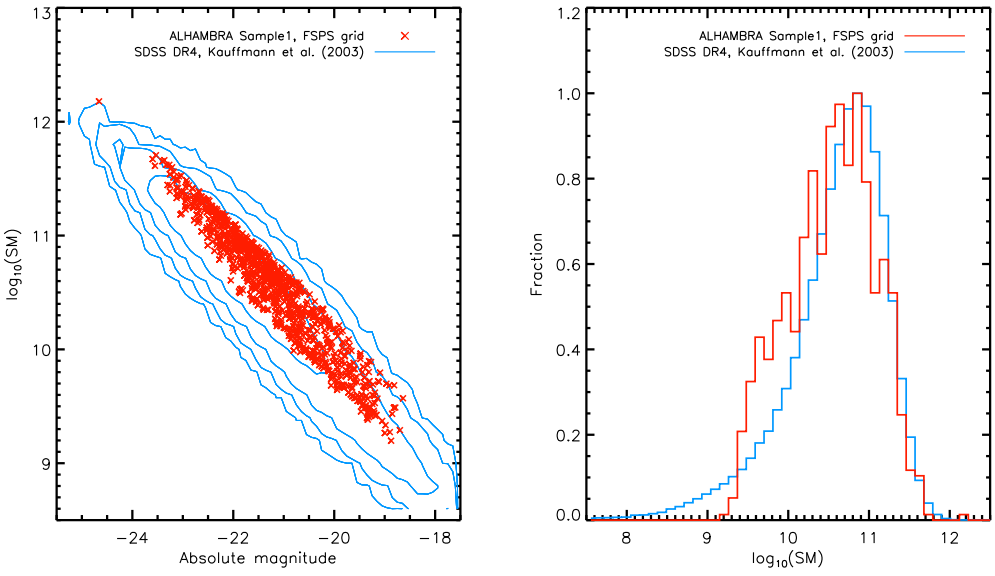

Figure 5.19 Left: Logarithm of the stellar mass as a function of K-corrected (to $z = 0$) absolute magnitude for both the ALHAMBRA Sample1 (crosses) and the SDSS DR4 (contours). The SDSS estimates were taken from Kauffmann et al. (2003b). The absolute magnitudes are in the z band for the SDSS and in the f_{8919} (z) filter for ALHAMBRA. Right: The distribution of stellar mass in the same ALHAMBRA (red) and SDSS (blue) samples. We have assumed $h = 0.7$ and a Kroupa IMF.

particular, we have used the stellar masses obtained by Kauffmann et al. (2003a) with the SDSS. To this purpose, we have downloaded an extension of the catalog used in Kauffmann et al. (2003a), which includes the DR4 (the original work used the DR1), from the webpage *http* : *//www.mpa* − *garching.mpg.de/SDSS/index_old.html*. The updated sample contains $\sim 400,000$ galaxies with $14.5 < r < 17.77$. In order to make both estimations comparable, we have set $h = 0.7$, K-corrected our magnitudes (and M/Ls) to $z = 0.1$ and used a Kroupa IMF. Note that the masses obtained with a Kroupa IMF are approximately a factor 2 larger than those obtained with a Salpeter IMF (Kauffmann et al., 2003b; Bell et al., 2003). In Figure 5.19, we show the result of this ALHAMBRA-SDSS comparison. In the left-hand panel, we plot the logarithm of the stellar mass as a function of absolute magnitude for both samples (where crosses represent ALHAMBRA galaxies and contours show the distribution of SDSS galaxies) . For the absolute magnitude, we use the SDSS z band for the SDSS sample (as this is the only one available in the catalog) and the f_{8919} (z) filter for ALHAMBRA. For consistency, stellar masses are also estimated using this filter in ALHAMBRA, which seems to be a very good choice, according to our fits. We find a tight relation between stellar mass and absolute magnitude in ALHAMBRA, which is in good agreement with the SDSS. We notice, however, that the slope is slightly flatter in ALHAMBRA. In the right-hand panel, we show the distribution of stellar mass for both the SDSS (blue line) and ALHAMBRA (red line). Both distributions peak somewhere close to 10^{11} $M_{\odot}$, with the ALHAMBRA distribution peaking at slightly smaller stellar masses. Note that the ALHAMBRA sample is limited at $i^A = 20$, so we are clearly mapping a much fainter, lower-mass population. However, we do not get galaxies with masses lower than $\sim 10^9$ $M_{\odot}$ in ALHAMBRA, whereas the SDSS does. This might be due to cosmic variance, as the volume mapped by the SDSS is huge as compared to that of ALHAMBRA.

5.5 Extension to deeper samples

In the previous sections, we focused on the estimation of photometric stellar masses in the ALHAMBRA Sample1, which is comprised by $\sim 1,000$ galaxies with $16 < i^A < 20$. In this section, we will briefly discuss on the extension of this analysis to Sample3, which contains $\sim 10,000$ galaxies with $16 < i^A < 24.5$. This sample represents the real potential of the ALHAMBRA survey. Note that, for the sake of simplicity, we will not explicitly consider Sample2 ($16 < i^A < 22.5$), which is, in any case, contained in Sample3.

When we go deep into faint magnitudes, the uncertainty in the photometry increases considerably, as Figure 5.6 reveals. We might expect that well-estimated photometric errors compensate the larger model-data discrepancy, so a similar χ^2 distribution as that shown in Figure 5.9 would be obtained. In practice, we find that the χ^2 distribution for Sample3 peaks close to that of Sample1, but the distribution is less pronounced, with a longer tail of high χ^2 values. This is illustrated in Figure 5.20 where we show the χ^2 distribution in Sample3, in the same format as in Figure 5.9, but extending the x axis up to $ch^2 = 20$. The χ^2 value that we set previously in order to exclude uncertain fits, i.e. $\chi^2 = 6$, seems a bit too conservative for this sample. Such a value would leave out approximately half of the sample, as the left-hand panel of Figure 5.20 demonstrates. Note that from a visual inspection of Sample1, we had checked that χ^2 values of up to $8 - 10$ yielded a reasonably good fit. The next

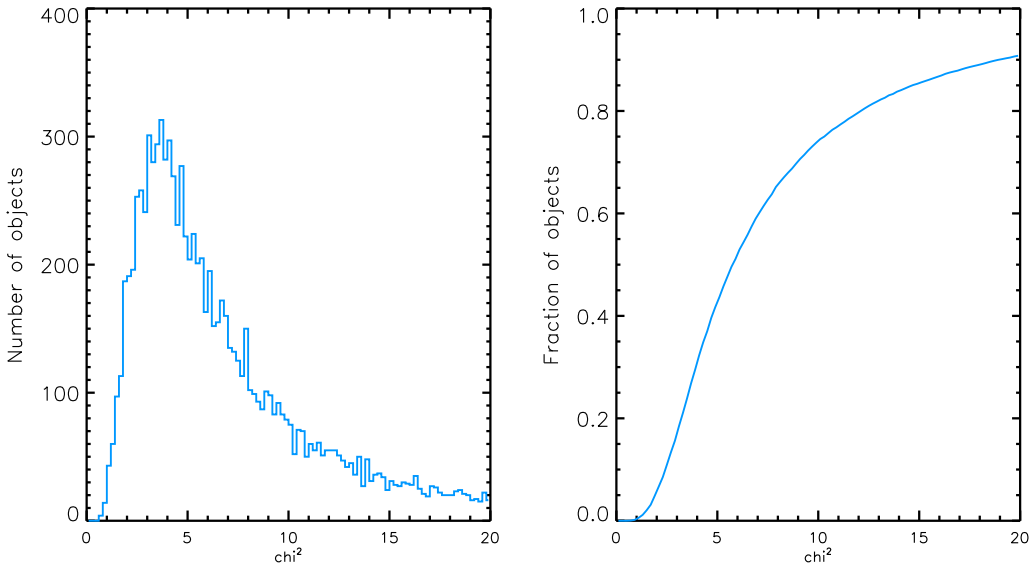

Figure 5.20 The same as in Figure 5.9 but for the ALHAMBRA Sampl3 ($16 < i^A < 24.5$).

step, as a follow up, is to update our FSPS and see wether we can improve our fitting and reduce the χ^2.

Until we can improve the performance of our grid, we will stick to our conservative χ^2 cut to provide a first look at the stellar mass distribution in the ALHAMBRA Sample3. In Figure 5.21, we compare the distribution of stellar masses obtained using the f_{8919} (z) filter with the SDSS DR4 sample of Kauffmann et al. (2003b), in the same format as in Figure 5.19. In the left-hand plot, the flatter slope in the ALHAMBRA narrow-filter $log_{10}(SM)$ - absolute magnitude relation as compared with the SDSS is very noticeable in Sample3. The right-hand plot of Figure 5.21 illustrates the potential of ALHAMBRA to shed light into the very faint, dwarf galaxy population, with a stellar mass distribution peaking at $SM \sim 10^{8.5}\ M_{\odot}$.

5.6 Uncertainties in the determination of stellar population properties

We have shown that the distributions of both M/Ls and photometric stellar masses preliminary derived with our FSPS grid in the ALHAMBRA low-redshift samples are consistent with previous results from the SDSS. The fact that photometric and, especially, photo-z errors are, however, not well estimated in our sample has prevented us from deriving errors for these galaxy properties. As it has been widely pointed out, the determination of uncertainties in SPS is as important as the determination of galaxy properties itself.

In many SPS works (see an example in Kauffmann et al. 2003a) the determination of galaxy properties is performed by means of a Bayesian approach. In this approach, the likelihood that a galaxy has a given value of a parameter can be computed by weighting each model in the grid by the probability function $exp(-\chi^2/2)$ and then simply binning the probabilities as a function of the parameter value. By doing so, we can obtain a probability distribution for a given galaxy property.

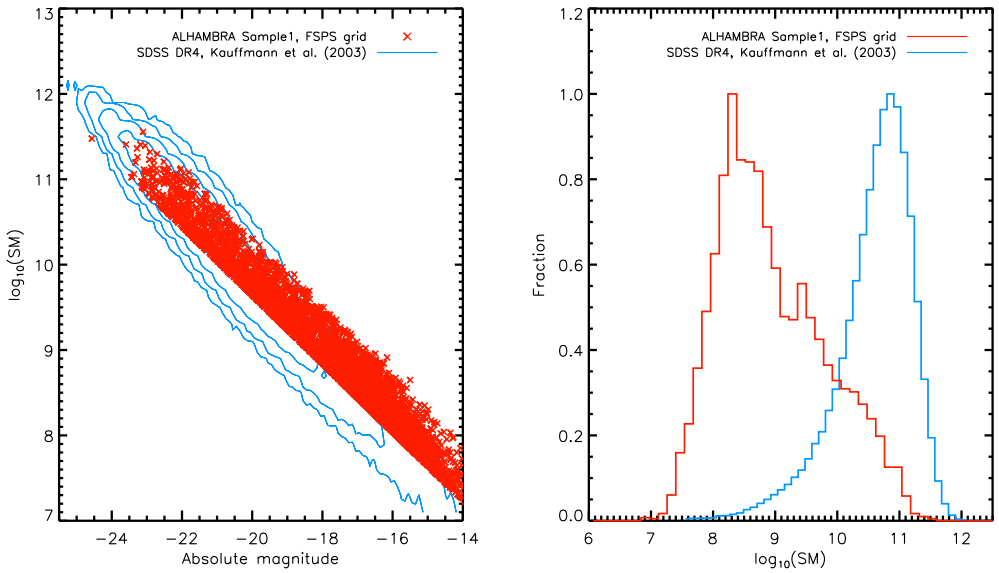

Figure 5.21 The same as in Figure 5.9 but for the ALHAMBRA Sampl3 ($16 < i^A < 24.5$).

The most probable value of the galaxy property corresponds to the peak value of this distribution and the most typical value is its median. By doing so instead of simply taking the parameter value of the best-fitting model, we can obtain an expectedly more accurate parameter value, which takes into account the probability for each model. More importantly, this method provides a simple way to derive confidence intervals for each parameter, by simply taken the values encompassing a given fraction of the distribution.

The error budget for SPS parameters in ALHAMBRA must take into account the uncertainty in the estimation of photo-z's. As mentioned before, photo-z's should be seen as probability distributions, instead of values with associated errors. In this sense, a simple way to include the uncertainties in the photo-z's in the SPS fitting is to marginalize on the distribution of possible redshifts, by weighting the χ^2 by an additional factor that penalizes models in different redshift slices. By proceeding in this way, we can obtain a probability distribution for the parameter values that takes into account the uncertainty in the estimation of the photo-z's.

Note, finally, that interpreting the light emitted by galaxies is extremely complex. The galaxy properties derived from stellar populations synthesis are inevitably affected by systematic uncertainties in the models. Despite the immense progress made in the last decades, there are still important unknowns on the IMF, on several phases of the stellar evolution and on the effect of dust, to name but a few. Comprehensive analyses on uncertainties in stellar population synthesis can be found in Marchesini et al. (2009) and in a series of works by Conroy et al. (Conroy et al. 2009 and subsequently).

5.7 The ALHAMBRA stellar mass function

The stellar mass function (SMF), which is the distribution of the number density of galaxies per stellar mass interval, is arguably the best tracer of galaxy evolution among all the statistical properties of

the galaxy population. The fact that it takes into account the maximum volume, V_{max}, for each galaxy within the sample allows us to compare our stellar mass distribution with previous estimates from other surveys. We have calculated the SMF in the ALHAMBRA Sample3, which, as we have mentioned before, is limited at $i^A = 24.5$. The SMF, that we represent by Ψ, can be estimated in a similar way to the luminosity function, using a simple $1/V_{max}$ method, where the absolute magnitude is replaced by the stellar mass, SM. Namely:

$$\Psi(SM)\Delta SM = \sum_i \frac{1}{V_{max,i}(M_i, z_i, SED_i)} \tag{5.14}$$

where $V_{max,i}(M_i, z_i, SED_i)$ is the maximum volume where galaxy i can be found while satisfying the apparent magnitude limits of the survey. These quantities are exactly the same that we used for estimating the ALHAMBRA LF in Chapter 4.

In a magnitude-limited sample, the minimum stellar mass for which the survey is complete depends on both the redshift and the stellar mass-to-light ratio. In fact, determining the selection function in stellar mass for this type of surveys in an exact way is hardly possible, as it requires knowledge of the intrinsic distribution of M/Ls in the galaxy population above and below the apparent magnitude limit. Several methods have been used before to estimate the stellar mass completeness limit, SM^{lim}, in flux-limited galaxy samples. A common approach is to take the M/L of a maximally old and red stellar population at the apparent mass limit of the survey (e.g., Drory et al. 2005; Fontana et al. 2006; Borch et al. 2006; Bundy et al. 2006; Pérez-González et al. 2008). Instead of this rather conservative approach, we will follow the method used by Pozzetti et al. (2010). For each galaxy in the sample, we calculate the minimum stellar mass, SM^{min}, that the galaxy with stellar mass SM could have while satisfying the survey limits. In practice, SM^{min} can be calculated in the following way:

$$\log_{10} SM^{min} = \log_{10} SM + 0.4(m - m_{lim}) \tag{5.15}$$

where m and m_{lim} are the apparent magnitude of the galaxy and the apparent magnitude limit of the survey, respectively. We will use filter f_{7698} (i) for estimating stellar masses, i.e. $m = f_{7698}(i)$ in Equation 5.15. Note that such a choice produces some dispersion in SM^{min}, as our sample is selected in the synthetic i^A band. We have checked, however, that such a dispersion in SM^{min} has little effect in the estimation of SM^{lim}. In order to obtain a reasonable value for SM^{lim}, we take the 20% fainter galaxies in our sample, following Pozzetti et al. (2010). We then define SM^{lim} as the stellar mass threshold that encompasses 95% of the distribution of SM^{min} in this sub-sample. The 20% threshold is consistent with the color-magnitude relation and is set to include only galaxies with a typical M/L close to the magnitude limit. Other methods for estimating SM^{min} that will be considered in the future include using galaxy mock catalogs (Meneux et al., 2009) and progressively deeper galaxy samples (Marchesini et al., 2009).

In Figure 5.22, a preliminary estimate of the SMF in the ALHAMBRA Sample3 is shown in crosses with Poisson error bars. Note that these purely statistical errors do not take into account the uncertainty in the photo-z estimation and the SPS fitting. They must be regarded, therefore, as lower limits to the real uncertainty, that will be estimated through a complete error budget analysis in the

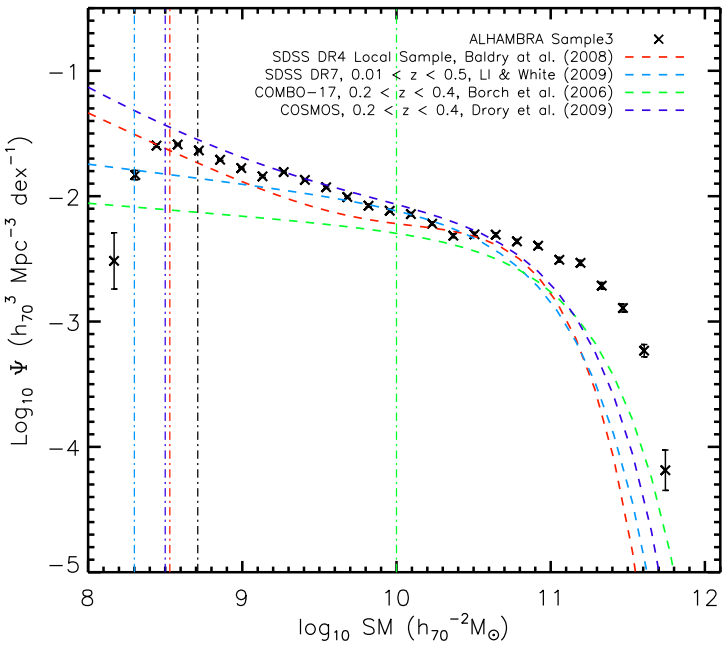

Figure 5.22 The Stellar Mass Function in the ALHAMBRA Sample3 ($i^A < 24.5$), represented by crosses, is compared with previous estimates from other surveys. The ALHAMBRA SMF was obtained with a Kroupa IMF. Borch et al. (2006) used a Kroupa et al. (1993) IMF, which is similar to the Chabrier2003 and the Kroupa IMFs. The local SMF of Baldry et al. (2008) was obtained with a combination of these two IMFs. Both Li & White (2009) and Drory et al. (2009) adopted a Chabrier2003 IMF. Vertical lines show the stellar mass limit for each estimate. Finally, Poison error bars have been added to the ALHAMBRA SMF and $H_0 = 70 km/s/Mpc$ ($h_{70} = H_0/70$) is assumed in all estimates.

future (see Section 5.6). We have adopted a Kroupa IMF and assumed $H_0 = 70 km/s/Mpc$ (in fact SM and Ψ are expressed as a function of h_{70} in Figure 5.22, so that $h_{70} = H_0/70$). The ALHAMBRA SMF is compared with estimates from different surveys and at different redshift ranges. Namely, we show, in a red line, the local SMF of Baldry et al. (2008), obtained with the SDSS DR4. We also show, in cyan, the estimate of Li & White (2009), where a sample of $\sim 500,000$ SDSS galaxies with $0.01 < z < 0.5$ taken from the NYU VAGC DR7 was used. In addition, we compare our ALHAMBRA SMF with two analogous deep photometric surveys: COMBO-17 (Borch et al., 2006) and COSMOS (Drory et al., 2009). These estimates were obtained at the redshift range $0.2 < z < 0.4$, which is roughly similar to that considered in this work, and are represented by a green and a blue line, respectively. Note that all these SMFs use either a Kroupa et al. (1993), a Kroupa2001 or a Chabrier2003 IMF, which are all comparable, and have been corrected to $h_{70} = 1$ when necessary. Finally, vertical lines show, in the same color code as that used for the SMFs, the stellar mass limit claimed for each estimate.

The most remarkable characteristic of the preliminary ALHAMBRA SMF, as compared with earlier works, is the prominent excess that we find at the high-mass end. Namely, at $M_* \gtrsim 10^{11}\ h_{70}^{-2} M_\odot$, we find ~ 200 galaxies, where according to previous estimates, we should barely find a few for the entire volume mapped by ALHAMBRA. Such a discrepancy seems to be connected with the bright-end excess found in the ALHAMBRA LFs presented in Chapter 4 (Figure 4.11 and Figure 4.12). Similarly, it will be hard to give a definite explanation for the high-mass excess in the SMF, until we can use the latest ALHAMBRA catalogs, where photo-z's are expectedly more reliable and errors are better characterized. From what we have learnt in this chapter, however, our method for estimating stellar masses seems to be unlikely to cause these deviation. In particular, no unexpected trend is found in the mass-magnitude relation of Figure 5.21. Moreover, the shift found in the M/L - color relation as compared to Bell et al. (2003), that is visible in Figure 5.15, is not large enough, of course, to be responsible for the high-mass excess (we have checked that a similar excess is found when synthetic bands are used for estimating stellar masses). The discrepancies are not due to catastrophic failures in the SPS fitting either, as $\chi \sim 3.4$ (with $\sigma \sim 1.5$) at $SM \gtrsim 10^{11}\ h_{70}^{-2} M_\odot$.

Cosmic variance and photo-z errors are, at this point, the strongest candidates to explain the high-mass excess found in the ALHAMBRA SMF. Regarding the first, preliminary analysis on the distribution of cosmological distances to galaxies in our sample would not indicate the presence of prominent large-scale structures within the volume mapped by ALHAMBRA. However, we cannot rule out this possibility yet. Uncertainties in the photo-z estimation, on the other hand, are known to be capable of generating strong deviations at the high-mass end of the SMF (and the bright end of the LF), as compared to estimates that use spectroscopic redshifts. On the exponential part of the SMF, uncertainties in redshift will tend to scatter objects preferentially from lower stellar mass bins, where objects are exponentially more abundant, to higher stellar mass bins. This effect can be strong even in the absence of catastrophic failures as it is shown in Figure 4 from Drory et al. (2009), who used the COSMOS photo-z survey. The expected overall precision of our photo-z estimates should guarantee, however, a small systematic effect at the high-mass end, by no means as large as that found in Figure 5.22. Therefore, everything seems to point towards catastrophic photo-z failures as the cause of the discrepancy found at the high-mass end of the SMF.

At the low-mass end, the effect of photo-z uncertainties is expected to be considerably smaller

(see Drory et al. 2009). At these ranges, the slope of the ALHAMBRA SMF seems to be relatively consistent with that of the local SMF of Baldry et al. (2008) and the COSMOS SMF of Drory et al. (2009), which was obtained in the redshift range $0.2 < z < 0.4$. Note that for the SMF of Li & White (2009) we took the single-Schechter fit, which does not represent the faint-end upturn conveniently. This relative agreement at the low-mass end is a promising sign before applying our SPS method to the latest ALHAMBRA catalogs. In this sense, a priority with the final catalogs will be to find out the nature of the high-mass excess in the SMF (and the bright-end excess in the LF). As expressed before, the BPZ, at least in the catalogs used in this work, was calibrated to the median redshift of ALHAMBRA ($z \sim 0.7$), so it might happen that these discrepancies disappear when we extend our analysis to higher redshift bins.

5.8 Chapter conclusions

In this chapter, we provide a method for estimating photometric stellar masses in the ALHAMBRA survey, using stellar population synthesis. In particular, we propose Flexible Stellar Population Synthesis, which is a user-friendly and extremely fast code for generating single and composite stellar population models, for a variety of assumptions regarding the average properties of the stellar population. We have built a simple library of galaxy models with varying SFHs, metallicities, dust attenuation and redshift. With this grid, we have performed a SPS fitting in the same low-z sample used in Chapter 4, with $0.1 < z < 0.3$. Both mass-to-light ratios and photometric stellar masses have been estimated in ALHAMBRA, for the first time. The main conclusions from this study can be summarized as follows:

- The colors of the ALHAMBRA galaxy population in the narrow filters are well reproduced by our preliminary FSPS grid in both the optical range and the NIR. In the optical range, however, we detect that ALHAMBRA galaxies with very blue f_{4916} (g) - f_{6141} (r) are not covered by the models, which might be due to photometric errors. This might also be a consequence of our FSPS parameter choice, as we have not included starburst and very young SF galaxies.

- From a visual inspection of the fits in the ALHAMBRA Sample1 ($16 < i^A < 20$), we conclude that the performance of our FSPS grid is good, as it provides a reasonably good fit for the majority of the galaxies in the sample.

- We have detected that the photometric errors are underestimated for relatively bright objects ($i^A < 20$) in the catalogs used in this work, which hinders the SPS fitting and, especially, the derivation of a reasonable χ^2 goodness-of-fit parameter. In order to perform the fitting without losing all photometric error information, we have added a constant error of 0.03 mag in any band showing this problem.

- A conservative cut for the χ^2 of good fits can be set at $\chi^2 = 6$, which includes $\sim 75\%$ of the ALHAMBRA Sample1. From a visual inspection of the fits in this sample, we conclude that the FSPS grid performs reasonably well up to $\chi^2 \sim 10$, i.e. for 90% of the galaxies in the sample.

- Most of the model-data discrepancies come from the very blue ALHAMBRA filters. In addition, our FSPS grid preforms considerably better for older stellar populations than for younger stellar populations.

- Mass-to-light ratios in narrow filters in ALHAMBRA (with $\chi^2 = 6$) are in good agreement with the Bell et al. (2003) parametrizations. However, we notice a shift in the zero point of ~ 0.05 mag. The M/Ls obtained with synthetic SDSS bands are in excellent agreement with these prescriptions.

- We have presented ALHAMBRA photometric stellar masses for both the Sample1 and the entire Sample3 ($16 < i^A < 24.5$). The distribution of the stellar mass is in good agreement with the SDSS DR4 stellar masses of Kauffmann et al. (2003b). However, we find a slightly flatter slope in the stellar mass - absolute magnitude relation as compared with the SDSS.

- The ALHAMBRA distribution of stellar masses derived for the entire Sample3 peaks at $\sim 10^{8.5}\ M_\odot$, which reflects the enormous potential of the survey to shed light into the very faint, low-mass galaxy population.

- We have presented a preliminary estimate of the ALHAMBRA stellar mass function in Sample3, which is comprised by $\sim 10,000$ galaxies with $i^A < 24.5$. At the high-mass end, the ALHAMBRA SMF shows a clearly artificial excess as compared to previous estimates, which is probably due to catastrophic photo-z errors. At the low-mass end, where the effect of photo-z uncertainties is expectedly less important, the preliminary ALHAMBRA SMF is in relatively good agreement with previous estimates from different surveys.

6

The DEEP2 Galaxy Redshift Survey: The Red Sequence AGN Fraction and its Environment and Redshift Dependence

6.1 Introduction

Both active galactic nuclei (AGN) and local environment play key roles in shaping galaxy evolution. It is now understood that AGN are those nuclei in galaxies that emit radiation powered by accretion onto a supermassive black hole. Although this realisation has proved useful for explaining many observed characteristics of these active objects, there are still many unsolved problems, especially related to the physics of the accretion process itself. In the recent years much effort has been invested in studying the global properties of AGN as a unique population in the context of galaxy formation. In this chapter, we focus on a fundamental question: the dependence of the fraction of galaxies that have AGN on the density of the local environment at $z \sim 1$, and the evolution of this dependence to $z \sim 0$.

At low redshift, many authors have investigated various correlations between galaxy properties and environment. It is now well established that there exists a relationship between morphology and density (Oemler 1974 and Dressler 1980), in that star-forming disk-dominated galaxies tend to inhabit less dense regions of the Universe than "quiescent" or inactive elliptical galaxies. Moreover, additional (and related) dependencies with environment have been found, such as with stellar mass, luminosity, colour, recent and past star formation, star formation quenching, surface brightness, and concentration (to name but a few) (e.g. Kauffmann et al., 2004; Balogh et al., 2004; Hogg et al., 2004; Blanton et al., 2005a; Bundy et al., 2006).

In this scenario of entangled correlations it is useful to investigate the dependence of AGN properties on the local environment, especially since AGN are believed to play an important part in shaping galaxy evolution. This has sometimes been a rather controversial issue. In the local Universe, Miller et al. (2003) found no dependence on environment of the fraction of spectroscopically selected AGN, using the SDSS Early Data Release. This result is in good agreement with Sorrentino et al. (2006) who

used the much larger SDSS DR4. However, many other authors have claimed the existence of a strong link between nuclear activity and environment, at least for specific AGN types. Kauffmann et al. (2004) found that intermediate luminosity optically selected AGN (Seyfert IIs) favoured underdense environments, while low-luminosity optically selected AGN (Low-Ionization Nuclear Emission-line Regions; hereafter, LINERs) showed no density dependence, within the SDSS DR1. Similarly, lower-luminosity AGN were found to have a higher clustering amplitude than high-luminosity AGN by Wake et al. (2004) and Constantin & Vogeley (2006). Radio-loud AGN have been noted to reside preferentially in mid-to-high density regions and tend to avoid underdense environments (Zirbel, 1997; Best, 2004).

At high redshift the study of both galaxies and AGN, and their relation to the environment, has been restricted by the lack of adequate data. Only in recent years, with the emergence of quality large-scale probes of the high redshift galaxy population, such as the DEEP2 Galaxy Redshift Survey (Davis et al., 2003) or the VIMOS-VLT Deep Survey (VVDS, Le Fevre et al., 2003), have we reached the stage where we can begin to measure the statistics of galaxy evolution in some detail. Using DEEP2, Cooper et al. (2006) found that many of the low redshift galaxy correlations with environment are already in place at $z \sim 1$. However important differences exist. The colour-density relation, for instance, tends to weaken towards higher redshifts (Cooper et al., 2007a; Cucciati et al., 2006). Also, bright blue galaxies are found, on average, in much denser regions than at low redshift. Such a population inverts the local star formation-density relation in overdense environments (Cooper et al., 2007b; Elbaz et al., 2007). This inversion may be an early phase in a galaxy's transition onto the red sequence through the process of star formation quenching. The truncation of star formation in massive galaxies is believed to be tightly connected with nuclear activity (see e.g. Croton et al., 2006; Bower et al., 2006, for more information). Further investigation reveals that post-starburst (aka. K+A or E+A) galaxies (e.g. Dressler & Gunn, 1983) are galaxies "caught in the act" of quenching and are in transit to the red sequence. These predominantly "green valley" objects reside in similar environments to regular star forming galaxies (Hogg et al. 2006; Nolan et al. 2007; Yan et al. 2007 in prep.) supporting the picture that star formation precedes AGN-triggered quenching, which precedes retirement onto the red sequence.

Georgakakis et al. (2007) were one of the first to study the environments of X-ray selected AGN at $z \sim 1$ using a sample of 58 sources drawn from the All-Wavelength Extended Groth Strip International Survey (AEGIS, Davis et al. 2007). The authors found that these galaxies avoided underdense regions with a high level of confidence. Nandra et al. (2007) show that the same AGN reside in host galaxies that populate from the top of the blue cloud to the red sequence in colour-magnitude space. They speculate that such AGN may be the mechanism through which a galaxy stays red. Similar ideas have become a popular feature of many galaxy formation models that implement lower luminosity (i.e. non-quasar) AGN to suppress the supply of cooling gas to a galaxy, hence quenching star formation through a process of "starvation" (e.g. Croton et al., 2006; Bower et al., 2006).

In this chapter we study the environmental dependence of nuclear activity in red sequence galaxies within a carefully chosen sample of both X-ray and optically selected AGN, drawn from the AEGIS Chandra catalogue and the DEEP2 Galaxy Redshift Survey, respectively. The chapter is organised as follows. In Section 6.2 we describe our AGN selection. In Section 6.3 we present our main result: the AGN fraction of red sequence galaxies at $z \sim 1$ as a function of environment for three types of AGN

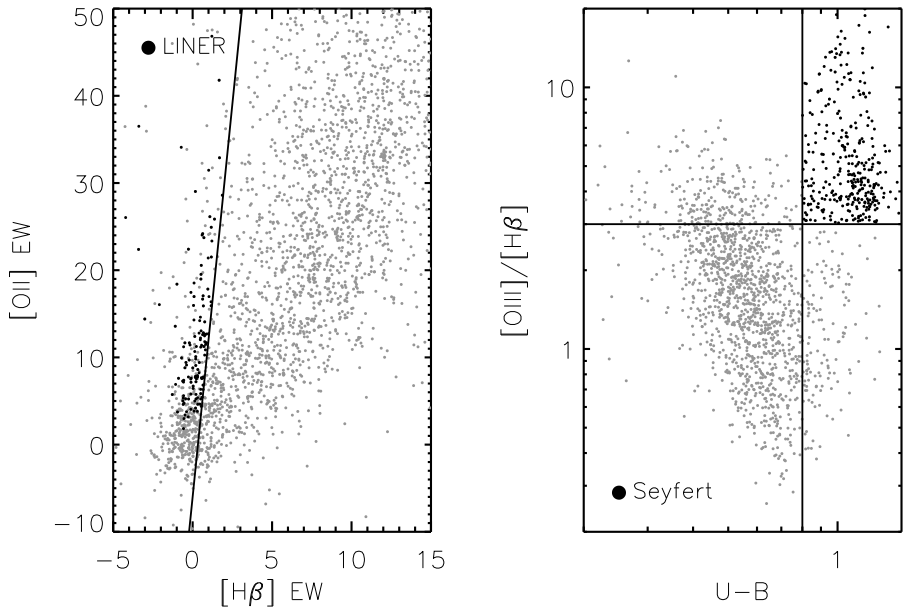

Figure 6.1 Two panels that show our AGN selection of Seyferts and LINERS within the DEEP2. The left panel plots [OII] EW versus Hβ EW for objects with accurate redshifts ($Q \geq 3$), δ_3 environment measures, and covered [OII], [OIII] and Hβ (grey points). LINERs (black points) are selected using the empirical demarcation of Equation 6.1 along with the colour cut defined by Equation 6.2. The right panel shows the line ratio [OIII]/Hβ plotted against $(U - B)$ rest-frame colour for the same DEEP2 sample (grey points). Seyferts (black points) are selected to have [OIII]/H$\beta \geq 3$ and rest-frame colour $(U-B) > 0.8$, as denoted by the horizontal and vertical lines. See Section 6.2.1 for further details.

(LINERs, Seyferts and X-ray selected). We undertake a comparison between our high-z results and those derived from a low-z sample drawn from the SDSS in Section 6.4. Finally, in Sections 6.5 and 6.6 we provide a discussion and brief summary. Throughout, unless otherwise stated, we assume a standard ΛCDM concordance cosmology, with $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $w = -1$, and $h = 1$. In addition, we use AB magnitudes unless otherwise stated.

6.2 Galaxy and AGN Selection

Our primary galaxy and AGN samples are drawn from the DEEP2 Galaxy Redshift Survey (Davis et al., 2003, 2005), a project designed to study galaxy evolution and the underlying large-scale structure out to redshifts of $z \sim 1.4$. The survey utilized the DEIMOS spectrograph (Faber et al., 2003) on the 10-m Keck II telescope and targeted $\gtrsim 50\,000$ galaxies covering ~ 3 deg^2 of sky over four widely separated fields. In each field, targeted galaxies are observed down to an apparent magnitude limit of $R_{\rm AB} < 24.1$. Important for this work, the spectral resolution of the DEIMOS spectrograph is quite high, R $\approx$ 5000, spanning an observed wavelength range of $6500 < \lambda < 9200$Å. This allows us to confidently identify AGN candidates through emission-line ratios down to low equivalent widths. Such objects form a core part of the data analysed in this chapter. More details on the DEEP2 survey

design and galaxy detection can be found in Davis et al. (2003, 2005, 2007) and Coil et al. (2007).

To study the dependence of the AGN fraction on local environment, for each galaxy we use the pre-calculated projected third-nearest-neighbour distance, $D_{p,3}$, and surface density, $\Sigma_3 = 3/(\pi D_{p,3}^2)$ (taken from Cooper et al., 2005). This density measure is then normalised by dividing by the mean projected surface density at the redshift of the galaxy in question, yielding a quantity denoted by $1+\delta_3$. Tests using mock galaxy catalogues show that δ_3 is a robust environment measure that minimises the role of redshift-space distortions and edge effects. See Cooper et al. (2005) for further details and comparisons with other commonly used density estimators.

To complement our optical catalogue we employ Chandra X-ray data from the All-Wavelength Extended Groth Strip International Survey (AEGIS, Davis et al. 2007). The AEGIS catalogue provides a panchromatic measure of the properties of galaxies in the Extended Groth Strip (EGS) covering X-ray to radio wavelengths. The EGS is part of DEEP2, constituting approximately one sixth of its total area. This allows us to cross-correlate each X-ray detection with the optical catalogue to identify each galaxy counterpart. In this way environments can be determined for the X-ray AGN sources.

Selecting objects from both the DEEP2 spectroscopic and Chandra (AEGIS) X-ray catalogues provides two different AGN populations that are embedded in the same underlying large-scale structure. To differentiate the two in the remainder of the chapter, we hereafter refer to the first as the optically selected sample (OSS) and the second as the X-ray selected sample (XSS). In the following sections we describe the OSS and XSS populations in more detail.

6.2.1 The optically selected AGN sample (OSS)

Optically (or spectroscopically) selected AGN in the DEEP2 survey can be divided into two main classes, LINERs and Seyferts, distinguished primarily through the spectral lines present and their strength. Although the physical processes that differentiate one class from the other are still not well understood, the identification of each class is never-the-less well defined. We restrict our analysis to the redshift range $0.72 < z < 0.85$ to ensure that all chosen AGN spectral indicators are visible within the covered wavelength range and that the environment measure is sufficiently reliable. This will be the redshift interval from which all our OSS results are taken. Furthermore, to facilitate a fair comparison between both AGN types, only objects on the red sequence or in the green valley are included (defined below). This will also allow us to compare with a low redshift sample (see Section 6.4). For a complete discussion of the spectroscopic detection of AGN in the DEEP2 survey see Yan et al. (2011). Below we will briefly outline our LINER and Seyfert selection in turn.

LINERs

As discussed in Yan et al. (2006), LINERs are a population of emission-line galaxies with high equivalent width (EW) ratio [OII]/Hα (or [OII]/Hβ). Specifically, we select a complete sample of LINERs using the division in [OII]/Hβ EW space given in Yan et al. (2006):

$$\mathrm{EW}\big([\mathrm{O_{II}}]\big) > 18\,\mathrm{EW}\big(\mathrm{H}\beta\big) - 6 \tag{6.1}$$

The left panel of Figure 6.1 illustrates this selection by plotting [OII] EW against Hβ EW for the entire DEEP2 sample with accurate redshifts ($Q \geq 3$), δ_3 environment measures, covered [OII], [OIII] and Hβ (for consistency with Seyfert selection – see below), and redshift window $0.72 < z < 0.85$ (grey points). The solid line indicates the empirical demarcation of Equation 6.1. Since quiescent galaxies with no line emission also satisfy this criteria, the inequality relation alone is not sufficient. Thus, we further require LINERs to have significant detection (2σ) of [OII]. As Hβ emission is expected to be weak in LINERs (Yan et al., 2006), we do not require significant detection on Hβ.

The error on Hβ EW emission is large due to the difficulty in measuring it after subtracting the stellar absorption. The above LINER selection has contamination from star-forming galaxies whose Hβ EW is underestimated. From a study of SDSS galaxies, Yan et al. (2006) concluded that LINERs are almost exclusively found in red sequence galaxies. Therefore, we adopt an additional colour cut to remove this contamination, which is the same used by Willmer et al. (2006):

$$(U-B) > -0.032 M_B + 0.322 \tag{6.2}$$

Our final LINER sub-sample with all of the above constraints is comprised of 116 objects and is overplotted in the left panel of Figure 6.1 with black points. Note that within the SDSS a strong vertical branch can be seen (see Figure 2 of Yan et al. 2006, where they use Hα instead of Hβ). This branch is significantly weaker at $z \sim 1$ in the DEEP2 data. This is due in part to the greater errors on Hβ in the DEEP2 data, and in part to the domination of red galaxies in the SDSS sample (due to the SDSS selection criteria).

Seyferts

Seyferts require different selection techniques than LINERs. Following the method of Yan et al (in prep.), we identify Seyferts in DEEP2 using a modified Baldwin-Phillips-Terlevich (BPT) diagram (Baldwin et al., 1981). Historically, the BPT diagram has been a reliable tool for determining the source of line emission from a galaxy. By plotting the line ratios [OIII] λ5007/Hβ against [NII] λ6583/Hα one can visually differentiate Seyferts, LINERs and star-forming galaxies. However, Hα is not available in the DEEP2 spectra at $z \gtrsim 0.4$ as it is redshifted into the infrared. For this reason, we use a modified BPT diagram which replaces the line ratio [NII] λ6583/Hα with the rest-frame $(U-B)$ colour. This is possible because both are rough proxies for metallicity. Tests done on SDSS samples demonstrate that such a substitution is able to produce a clean and complete selection criterion for Seyferts (Yan et al., 2011).

In the right panel of Figure 6.1 we illustrate our Seyfert selection by showing the modified BPT diagram for the same underlying sample used to select LINERs (grey points). This figure shows that the modified BPT diagram has a similar two branching structure to the original BPT diagram. Seyferts are selected to have [OIII]/H$\beta \geq 3$ and rest-frame colour $(U-B) > 0.8$ (horizontal and vertical lines, respectively). For cases in which Hβ is not positively detected, we use a 2σ lower limit on [OIII]/Hβ. With such criteria we obtain 131 Seyferts in the range $0.72 < z < 0.85$ where all spectral signatures for both Seyferts and LINERs are normally available (black points). Selecting only red sequence (or green valley) objects facilitates a fair comparison with LINERs and is consistent with their typical

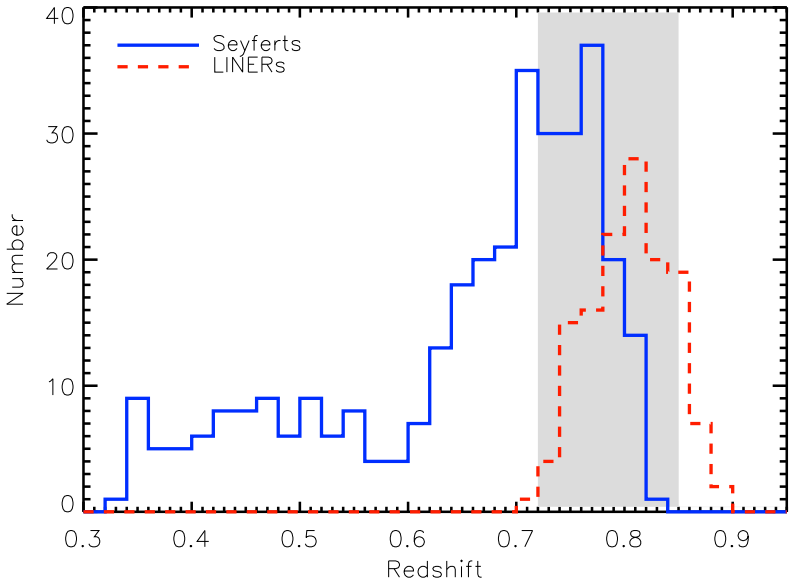

Figure 6.2 The redshift distribution for our optically selected AGN sample (OSS, Section 6.2.1). The distribution of Seyferts is given by the blue solid line, while the distribution of LINERs is given by the red dashed line. The shaded region denotes the redshift window $0.72 < z < 0.85$ from which our final OSS sample is drawn. Within this window both populations are cleanly identified spectroscopically and the effect of selection is small in both sub-samples.

position in the colour-magnitude diagram (Yan et al., 2006).

Redshift distributions

In Figure 6.2.1 we show the redshift distribution of both LINERs (dashed line) and Seyferts (dotted line) drawn from the selection given in each panel of Figure 6.1. The DEEP2 Seyfert population extends from $z \approx 0.35$ to $z \approx 0.85$, peaking at around 0.75. For LINERs the distribution is much more concentrated, extending from 0.72 to 0.9 and peaking at around 0.8. Note that the peak for both is dominated by the DEEP2 survey galaxy selection and not an intrinsic peak in the AGN distribution. As discussed previously, the redshift window where both populations can be cleanly identified spectroscopically is $0.72 < z < 0.85$, denoted by the shaded region. This range maximises AGN coverage while insuring that selection effects are minimised in both samples.

6.2.2 The X-Ray selected AGN sample (XSS)

AEGIS Chandra X-ray sources within the EGS field are optically and spectroscopically identified by cross-correlating with the DEEP2 photometric and redshift catalogues, following the prescriptions presented by Georgakakis et al. (2007). They cover X-ray luminosities of $10^{41} \lesssim \mathrm{L_X(erg/s)} \lesssim 10^{44}$ in host galaxies of luminosity $-19 \lesssim M_B - 5\log h \lesssim -22$. The base X-ray sample comprises a total of 113 reliably matched objects.

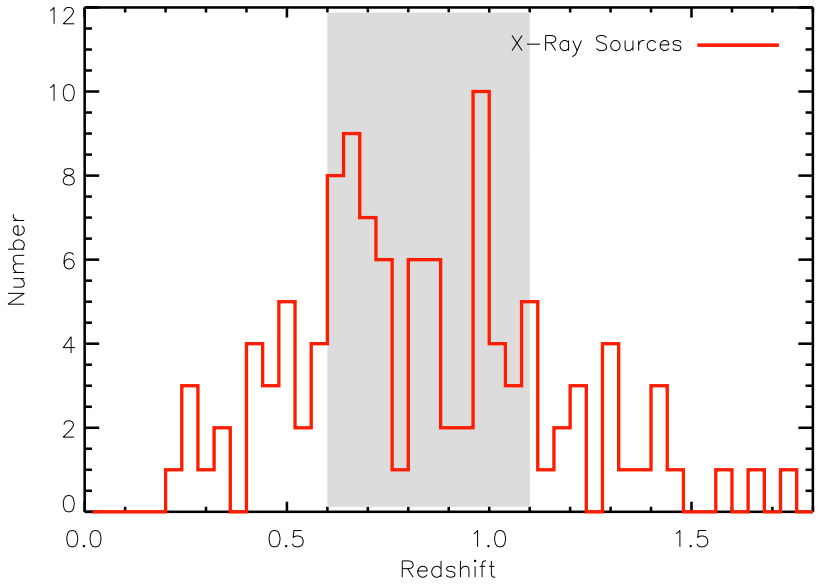

Figure 6.3 The redshift distribution for our X-ray selected AGN sample (XSS, Section 6.2.2), shown by the red solid line. The grey shaded region denotes the redshift range $0.6 < z < 1.1$, from which our final XSS is drawn. This range was chosen to be approximately comparable to the OSS (Figure 6.2.1) while simultaneously maximising the number and completeness in the sample.

In Figure 6.2.1 we show the redshift histogram of our X-ray catalogue. To extract a sample that is as closely comparable to the OSS as possible while simultaneously maximising AGN number and completeness, we restrict the X-ray sources to the redshift range $0.6 < z < 1.1$. This is wider than the OSS redshift window; however, both samples (OSS and XSS) have similar redshift means, and we assume that the evolution effects for sources outside the OSS redshift window do not dominate our results (or at least does not differ significantly from evolution in the red sequence population itself). In this redshift range the number of reliable X-ray AGN drops to 68, including 52 red-ward of $(U-B) > 0.8$ (i.e. a green valley cut), and 36 red-ward of Equation 6.2 (i.e. a red sequence cut).

6.2.3 AGN in colour-magnitude space

In Figure 6.4 we show the colour-magnitude diagram for LINERs (left panel), Seyferts (middle panel) and X-ray AGN (right panel). The demarcations given by the solid and dashed lines represent the conventions adopted to separate the blue cloud from the green valley, and the latter from red sequence objects (Equation 6.2), respectively (Willmer et al. 2006; Yan et al. in prep.). Here, LINERs are red sequence galaxies by definition. As explain above, this restriction is supported by the fact that local LINERs are almost exclusively red (Yan et al., 2006). For Seyferts, $\sim 80\%$ lie on the red side of the CMD, with the remainder residing in the green valley. Finally, for the XSS AGN, $\sim 50\%$ of the sources are red, $\sim 25\%$ are green, and the remaining $\sim 25\%$ blue. The grey contours in each panel show the underlying DEEP2 color-magnitude diagram within the same redshift range.

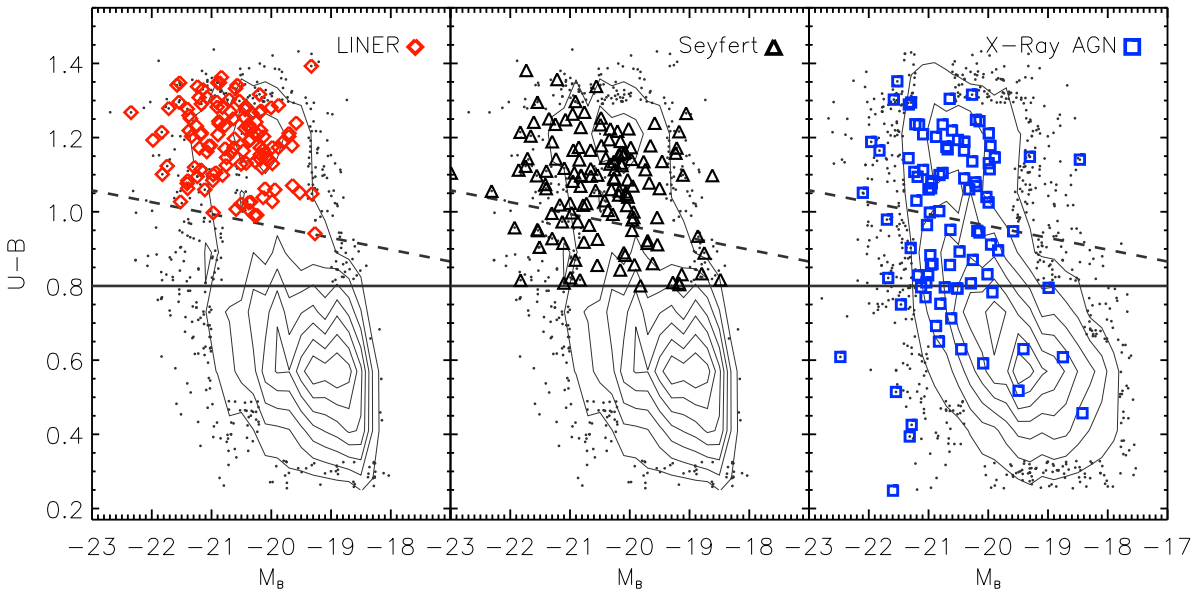

Figure 6.4 The colour-magnitude diagram for LINERs (left panel, red diamonds), Seyferts (middle panel, black triangles) and X-ray AGN (right panel, blue squares). The demarcations given by the solid and dashed lines represent the conventions adopted to separate the blue cloud from the green valley, and the latter from red sequence objects (Equation 6.2), respectively (Willmer et al., 2006). The LINER sub-sample is composed of 116 objects, all of them lying on the red sequence by definition. The Seyferts sub-sample is composed of 131 objects, with 97 of them on the red sequence and 34 in the green valley. Finally, from our X-ray sample of 68 objects, 36 sources are red, 16 are green, and the remaining 16 blue. The underlying CMD of the population from which all AGN are drawn is shown in each panel with grey contours and black points. This parent population, in the left-hand and middle panels, is comprised of objects with accurate redshifts ($Q \geq 3$), δ_3 environment measures, covered [OII], [OIII] and $H\beta$; and redshift between 0.72 and 0.85. In the right-hand panel, the grey contours and black points represent all objects in the EGS field with accurate redshifts ($Q \geq 3$) and δ_3 environment measures; and redshift between 0.6 and 1.1.

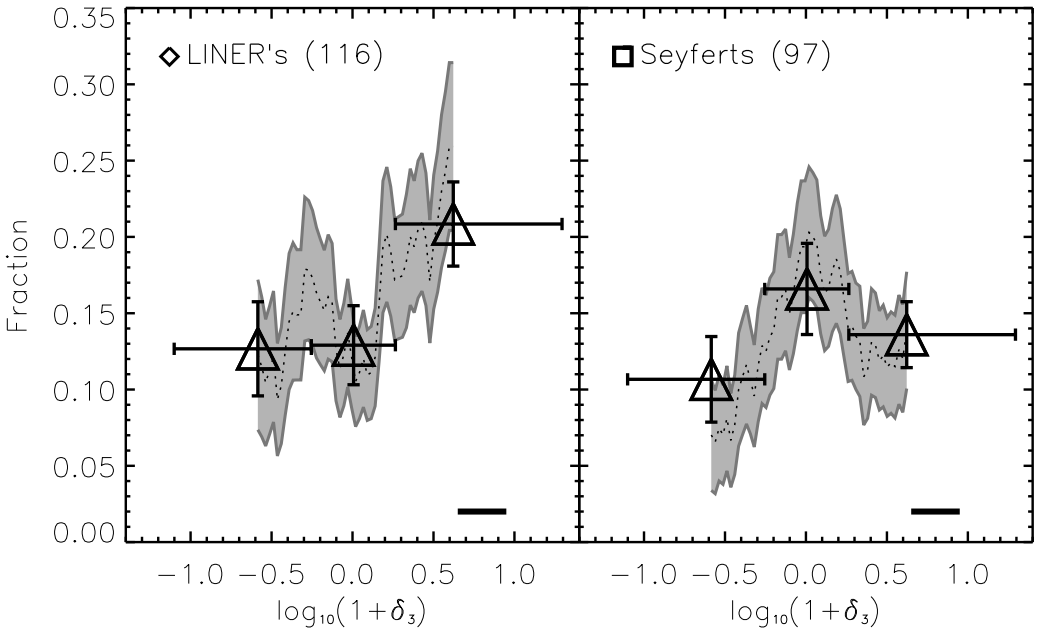

Figure 6.5 The AGN fraction in red sequence galaxies versus local galaxy over-density, δ_3, for LINERs (left panel) and Seyferts (right panel). For each, the respective symbols (diamonds for LINERs and squares for Seyferts) show the median measure in bins of low, mean and high density. Vertical error-bars represent the Poissonian uncertainty, while horizontal error-bars show the size of each density range. We also show how the AGN fraction varies smoothly with environment using a sliding box of width 0.3 dex shifted from low to high density in increments of 0.025 dex (dotted lines with shaded area showing the 1σ uncertainty in the sliding fraction). The overall fraction of LINERs and Seyferts is plotted with horizontal dashed lines. This figure shows some evidence that LINERs tend to favour high density environments relative to the underlying red sequence, whereas Seyferts have little (or no) environment dependencies.

6.2.4 Errors and completeness

Our greatest source of error is that of noise from small number statistics, given the low number of AGN we have available in the DEEP2 and AEGIS surveys in any particular environment bin. Such current-generation high-redshift catalogues are thus still limited in the extent to which the statistical nature of the AGN population can be examined. All errors calculated in this paper were determined by propagating the Poissonian uncertainties on the number of objects. Due to the small number statistics, the errors obtained with this method will dominate any cosmic variance effects in the observed fields (Newman & Davis, 2002).

It should be noted that the DEEP2 survey is by design incomplete. At $z \sim 1$, approximately 60% of the actual objects are observed by the telescope. Moreover, redshifts are successfully obtained for around 70% of the target parent population (based on tests with blue spectroscopy, most failures are objects at $z > 1.4$ (Steidel, private communication). This should be carefully considered in any statistic that counts absolute numbers of objects. In our work, however, we deal with *relative* numbers of objects, i.e. the AGN fraction. We assume, to first order (and to the level of uncertainty given by the Poisson error), that any variation in redshift success or targeting rate between the AGN sample and the red sequence parent population is the same in low density regions as it is in high density regions. In principle, one may expect an easier detection (or even an easier redshift estimation) of an object identified as an AGN than the one for a "regular" red sequence object (due to the presence of remarkable features in the spectrum). However both Cooper et al. (2005) and Gerke et al. (2005) found that DEEP2 selection rates are essentially independent of local density.

Finally, because of the different Seyfert and LINER selection we find some inevitable (but small) overlap between the two populations, 7% of the total in our case, where a single object has been classified as both AGN types. We have re-calculated all our results excluding these dual class objects and find only trivial differences. For the sake of maximising statistics we have not removed such objects from the OSS, however note that they may constitute an interesting sub-population whose physical implications warrant further investigation.

6.3 Results

In this section we present our primary result: the dependence of the AGN fraction in the red sequence on local environment density. We will also extend the analysis to include green valley objects.

Figure 6.5 presents the density dependence of the fraction of $z \sim 1$ red sequence AGN, for LINERs (left panel) and Seyferts (right panel) separately. In each panel, the respective symbols show the median measure in bins of low, mean and high density environments (each of them encompassing one third of the OSS), where the horizontal error-bars indicate the width of each bin, and the vertical error-bars show the Poisson uncertainty in the measured fraction, as described in Section 6.2.4. We also show how the AGN fraction varies smoothly with environment using a sliding box of width 0.3 dex, shifted from low to high density in intervals of 0.025 dex (dotted line). The accompanying grey-shaded regions correspond to the sliding 1σ uncertainties in the sliding fraction.

Evidence for a trend in the behaviour of the LINERs is quite apparent in Figure 6.5, suggesting the possibility that these objects tend to favour high density environments, and in a way stronger

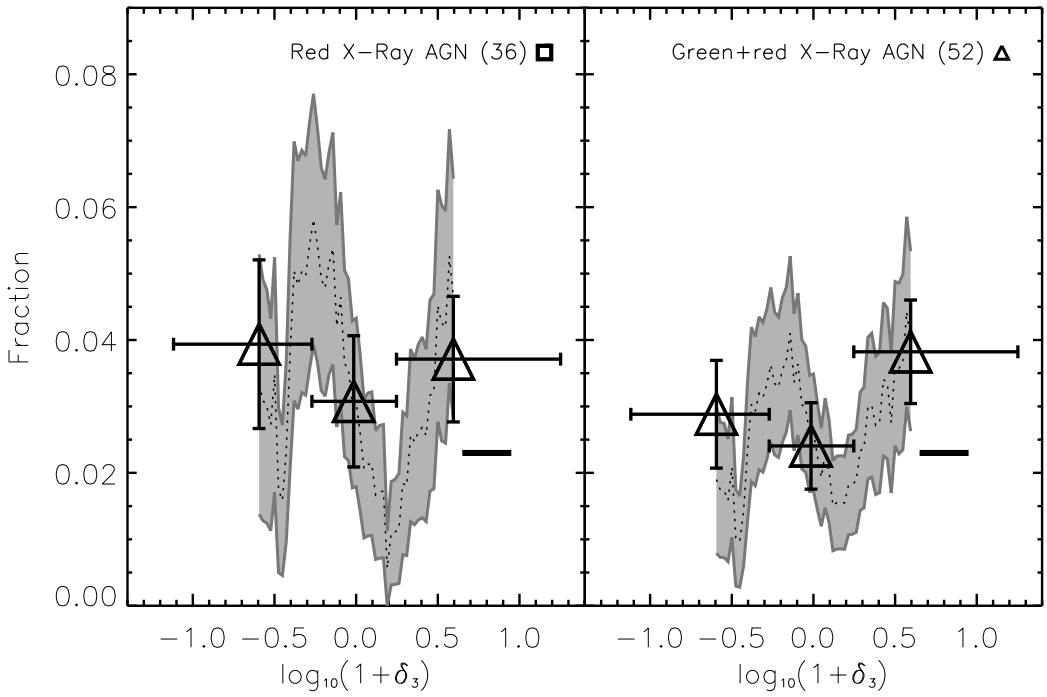

Figure 6.6 The AGN fraction versus environment for our X-ray selected sample (XSS), using the same format as in Figure 6.5. In the left panel, squares show the fraction of red sequence X-ray AGN in the three density bins considered. In the right panel we extend this analysis to include green valley AGN. Note that this extension does not change the results in any significant way. XSS AGN behave more like Seyferts than LINERs (see Figure 6.5), with the fraction showing only a weak (or no) environmental dependence to within the errors.

than the majority of red sequence galaxies. This is in contrast to the behaviour of red Seyferts, which show little (or no) environment dependence relative to the red sequence. This is a key result that will be discussed in more detail in the following sections.

We now consider the X-ray catalogue drawn from the AEGIS Chandra imaging. Figure 6.6 presents the X-ray selected AGN fraction versus local environment (note that the same format used in Figure 6.5 has been applied here). In the left panel we show the fraction of red X-Ray AGN and in the right panel we extended the analysis to include green valley X-ray AGN and galaxies. Including green valley objects does not significantly alter our results. Red sequence X-ray selected AGN appear to behave similarly to optically selected Seyferts in terms of their lack of environmental preference, and differently from the LINER population in high density environments. This is in agreement with the results of Georgakakis et al. (2007), also using the AEGIS data. We note that improved statistics could show trends at levels equal to or smaller than that which can be measured here given our errors.

We have tested the significance of the results in Figures 6.5 and 6.6 in a number of ways. Since LINERs show the most interesting environment dependencies we will focus our tests on this population. We randomly draw 1000 sub-samples from the red sequence and replace the LINER sample with each of these random populations. After repeating our analysis for each we find that only 2% of the random sub-samples show similar density dependencies to the LINER population (i.e. results at least as pronounced as the one in Figure 6.5) . The LINER environment dependence seen in the left panel of Figure 6.5 deviates by at least 2σ (actually almost 2.5σ) from a random selection of red sequence galaxies. Additionally, we can confirm that the trend in Figure 6.5 is not due to an implicit dependence of colour or magnitude on environment within the red sequence. This was checked by repeatedly replacing the LINER sample with randomly drawn objects with the same colour or colour and magnitude distributions, and comparing their density distributions with that of the real LINERs. The mean density of the LINER population is $1 + \delta_3 = 0.37 \pm 0.06$, almost double that for randomly colour selected samples which have $1 + \delta_3 = 0.20 \pm 0.01$, and randomly colour and magnitude selected samples with $1 + \delta_3 = 0.19 \pm 0.01$. Similar tests were performed on Seyferts and X-Ray AGN.

6.4 A Comparison with Local AGN in the SDSS

Our results thus far suggest that the $z \sim 1$ red sequence LINER fraction depends on environment in a way that is different from Seyferts. This dependence takes the form of an increase in the relative abundance of LINERs in higher density environments. In this section we address the question of AGN fraction evolution. Specifically, do local red sequence LINERs also favour dense environments and Seyferts show little environment dependence?

Our low redshift AGN sample is drawn from the Sloan Digital Sky Survey (SDSS, York et al. 2000) spectroscopic DR4 catalogue (Adelman-McCarthy et al., 2006). The SDSS DR4 covers almost 5000 square degrees of the sky in five filters (*ugriz*) to an apparent magnitude limit of $r = 17.7$. The redshift depth is approximately $z \sim 0.3$, with a median redshift of $z = 0.1$. DR4 consists of $\sim 400\,000$ galaxies. The same environment measure is applied for consistency with the DEEP2 analysis above (see Cooper et al. 2007a for full details).

To measure the low redshift AGN fraction we follow a similar procedure to that used for the high-z

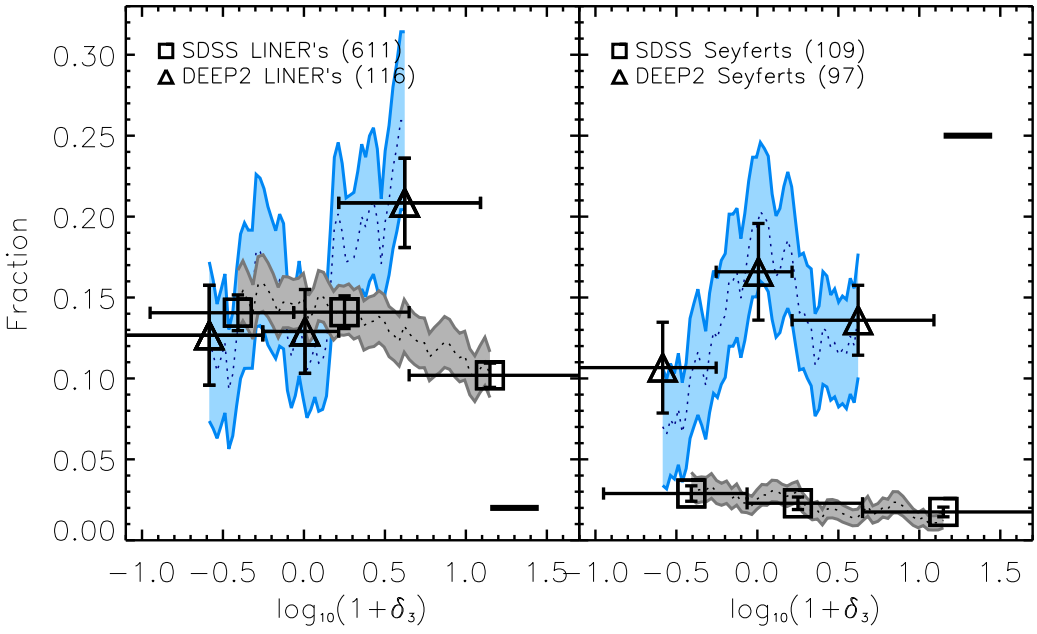

Figure 6.7 The SDSS red sequence AGN fractions versus environment, plotted using the same format from Figures 6.5 and 6.6. The left panel shows the fraction of low-z SDSS LINERs (squares) and, for comparison, the high-z DEEP2 LINER fraction (triangles) reproduced from Figure 6.5. The right panel shows the fraction of SDSS Seyferts, with the equivalent DEEP2 result again reproduced from Figure 6.5 for comparison. SDSS LINERs and Seyferts both show a decreasing AGN fraction towards high density environments, unlike that seen in DEEP2. At $z \sim 1$, LINERs and Seyferts are approximately equally abundant, whereas by $z \sim 0$ the relative abundance of Seyferts to LINERs has dropped by approximately a factor of 7.

results. This procedure isolates a well defined red sequence population and identifies AGN within it. The base red sequence population is constructed by selecting SDSS galaxies within the redshift interval $0.05 < z < 0.15$ and applying the rest-frame colour cut $(U - B) > -0.032M_B + 0.483$ (Cooper et al. 2007a, in agreement with the previous analysis by Blanton 2006). For consistency with our DEEP2 sample we take a faint absolute magnitude limit of $M_{\rm B} - 5\log h = -20$ (representing the approximate faint-end of the red sequence at $z \sim 0.8$ within the DEEP2 data) and evolve it 0.88 magnitudes to mimic the evolution in the galaxy luminosity function between DEEP2 and SDSS (assuming evolution of 1.3 magnitudes per unit redshift from Willmer et al. 2005 and mean DEEP2 and SDSS redshifts of 0.78 and 0.1, respectively). With these constraints the underlying low-z red galaxy sample is composed of 5335 objects.

To select AGN from the SDSS red sequence sample we use the same set of criteria described in Section 6.2.1 with the following modifications. Since SDSS spectra have a much higher signal-to-noise than DEEP2 spectra the emission line detections are easier. This will result in differences between the two AGN samples, as SDSS AGN will include weaker optical AGN than the DEEP2 can detect. Therefore, we determine a different line detection criterion by comparing the errors in the emission line EW measurements. Typical line measurement errors in DEEP2 are almost exactly twice as large as those in SDSS. We thus change all 2σ line detection criteria to 4σ for selecting AGN in the SDSS. The final low redshift AGN sample is comprised of 720 objects, 611 being LINERs and 109 Seyferts. This should be contrasted with the high-z sample which has 213 objects, of which 116 are LINERs and 97 Seyferts.

In Figure 6.7 we present the SDSS red sequence AGN fractions versus environment (this figure follows the same format used in Figures 6.5 and 6.6). The left panel shows the fraction of local red sequence LINERs (squares) and, for comparison, the high redshift LINER fraction (triangles) reproduced from Figure 6.5. The right panel shows the Seyfert fraction in the SDSS red population (squares) and similarly for DEEP2 (triangles, from Figure 6.5).

The left panel of Figure 6.7 reveals a different LINER trend with environment at $z \sim 0$ than that seen at $z \sim 1$. LINERs in the SDSS show no indication of favouring high density regions relative to other environments. In fact, it is very statistically significant that SDSS LINERs tend to reside more in mean-to-low density environments and clearly disfavour those of high density. In the right panel of Figure 6.7 Seyferts also follow a clear trend of decreasing AGN fraction towards denser SDSS environments. This is in contrast to the weak (or no) environmental trend in the high redshift DEEP2 Seyfert population.

It is important to note that comparing the overall amplitudes of the LINER and Seyfert fractions between DEEP2 and SDSS is dangerous, as subtleties in the selection of the underlying red sequence can shift the absolute values around somewhat. The relative trends across environment *within* a population are much more robust however, and these can be compared between high and low redshift (for example, we do not expect any selection effects to have a significant density dependence). Additionally, the difference in the abundance *between* Seyferts and LINERs at a given redshift can be contrasted. Seyferts and LINERs are approximately equally abundant at $z \sim 1$. By $z \sim 0$ however, the Seyfert population has diminished relative to the LINER population by over a factor of 7. This decline in the relative number of Seyfert AGN by redshift zero will be discussed in the next section.

6.5 Discussion

6.5.1 Previous measures of AGN and environment

It is difficult to make direct comparisons of our results with previously published works. This is because past environment studies have tended to focus on the AGN fraction of *all* colours of host galaxies, and also to mix both LINER and Seyfert classes into a combined AGN population. Our selection is restricted to the red sequence only (and also green valley), which allows us to compare high and low redshift populations and also study the AGN–environment connection without the colour–environment correlation.

Locally, a number of SDSS measures of AGN and environment have been made. Using the SDSS early data release, Miller et al. (2003) found no dependence on environment for the spectroscopically selected AGN fraction in a sample of 4921 objects. Specifically, the authors report no statistically significant decrease in the AGN fraction in the densest regions, although their densest points visually suggest such a trend. This result is broadly consistent with the results of both LINERs and Seyferts in Figure 6.7, even though we only consider red sequence objects.

Kauffmann et al. (2003a) also found little environment dependence of the overall fraction of detected AGN in a sample drawn from the SDSS DR1. However, they do report different behaviour when the sample is broken into strong AGN ($\log L[\mathrm{OIII}] > 7$, "Seyferts") and weak AGN ($\log L[\mathrm{OIII}] < 7$, "LINERs"). For Seyfert they find a significant preference for low-density environments, especially when hosted by more massive galaxies. This is consistent with our SDSS findings in Figure 6.7 and different to what we find at $z \sim 1$ in the DEEP2 fields. For LINERs, Kauffmann et al. measure little environment dependence, whereas we find a significant decline in the SDSS LINER fraction in our overdense bins. The explanation for this difference may come from our removal of possible contaminating star forming galaxies by restricting our analysis to the red sequence. Also, we impose a higher line detection threshold on the SDSS data to provide a fair comparison with DEEP2 (see Section 6.4). Finally, Kauffmann et al. (2003a) required that all lines for the BPT diagnosis were detected and this implies biasing the LINERs sample towards the strongest objects.

Between redshifts $z = 0.4$ and $z = 1.35$, Cooper et al. (2007a) show that *red galaxies* within the DEEP2 survey favour overdense environments, although the blue fraction in clusters does become larger as one moves to higher redshift (see also Gerke et al., 2007; Coil et al., 2007). At all redshifts there exists a non-negligible red fraction in underdense environments, which evolves only weakly if at all. Nandra et al. (2007) show that the host DEEP2 galaxies of X-ray selected AGN within the EGS field ($\sim 1/6$ of the DEEP2 survey volume) occupy a unique region of colour-magnitude space. These objects typically live at the top of the blue cloud, within the green valley, or on the red sequence. Georgakakis et al. (2007) measure the mean environment of this population and confirm that, on average, they live in density regions above that of the mean of the survey. They find this to be true for all host galaxy magnitudes studied ($M_B \lesssim -21$) and colours ($U - B \gtrsim 0.8$) (note the DEEP2 red sequence begins at $U - B \sim 1$). However, given limited sample sizes, they were not able to establish whether the environment distribution of the X-ray AGN differed from that of the red population, rather than the DEEP2 population as a whole.

6.5.2 Understanding the sequence of events

From our results alone a comprehensive understanding of the different environment trends within the AGN population from high to low redshift is not possible. However, some speculation and interpretation can be made by drawing on our broader knowledge of these active objects from the literature.

One possible scenario posits that LINERs and Seyferts occur in different types of galaxies. In this picture, LINERs are often associated with young red sequence galaxies (Graves et al., 2007) and are especially common among post-starburst (K+A) galaxies (Yan et al., 2006). These galaxies would already be into the quenched phase of their evolution but still relatively young. Merger triggered starbursts and subsequent quasar winds are a possible mechanisms to produce rapid star formation shut down in such objects (Hopkins et al., 2006). The gas rich merging events required in this scenario are common in overdense environments at $z \sim 1$ as clusters and massive groups assemble. By $z \sim 0$, however, the activity in these environments has mostly ended. Hence, if this picture is correct, one may expect an over-abundance of red sequence LINERs in dense environments at high redshift (since both star formation and rapid quenching is common) that is not seen locally. This may be consistent with the trends found in the left panel of Figure 6.7.

Seyfert galaxies, on the other hand, could be objects in transition from the blue cloud to red sequence (Groves et al., 2006), whose AGN are thought to be initiated by internal processes (and not mergers), inferred from their often found spiral structure (e.g. M77) (mergers act to destroy such structure). From this, one may expect our red sequence Seyfert population to represent the tail of the colour distribution of transitioning objects whose dependence on environment is determined by secular mechanisms and who would evolve accordingly. At high redshift, disk galaxies are commonly found in all environments, including the most dense. In contrast, overdense regions in the local Universe are dominated by passive ellipticals and show an absence of spirals. This would be broadly consistent with our findings in the right panel of Figure 6.7, where the most significant evolution in the red Seyfert fraction arises from a depletion in overdense regions relative to other environments, from high redshift to low.

Alternatively, some authors claim that LINERs and Seyferts form a continuous sequence, with the Eddington rate the primary distinguishing factor (Kewley et al., 2006). In this scenario, Seyferts are young objects with actively accreting black holes. As the star formation begins to decay so does the accretion rate, and the galaxy enters a transition phase. Eventually, a LINER-like object emerges, with an old stellar population and very low supermassive BH accretion rate. This picture is supported by recent studies in voids from Constantin et al. (2007). At high redshift, our results show that red Seyferts and LINERs are approximately equally abundant. By $z \sim 0$ however, the Seyfert population has declined relative to the LINER population by over a factor of 7. This may be interpreted as the natural transformation of Seyferts into LINERs with time, within a galaxy population which is smoothly reddening from $z \sim 1$ to $z \sim 0$. Moreover, the fact that high-z LINERs reside preferentially in high-density environments may imply that this Seyfert-LINER transition is more efficient in dense regions of the Universe.

6.6 Chapter conclusions

In this chapter we measure the dependence of the AGN fraction of red galaxies on environment in the $z \sim 1$ DEEP2 Galaxy Redshift Survey and local $z \sim 0.1$ SDSS. We restrict our analysis to the red sequence to maintain a clean and consistent selection of AGN at high and low redshift, and this also reduces the additional effects of environment associated with galaxy colour.

Our results can be summarised as follows:

(i) High redshift LINERs at $z \sim 1$ in DEEP2 appear to favour higher density environments relative to the red sequence from which they are drawn. In contrast, Seyferts and X-ray selected AGN at $z \sim 1$ show much weaker (or no) environmental dependencies within the same underlying population. Extending our analysis to include green valley objects has little effect on the results.

(ii) Low redshift LINER and Seyfert AGN in the SDSS both show a slowly declining red sequence AGN fraction towards high density environments. This is in contrast to the high redshift result.

(iii) At $z \sim 1$, Seyferts and LINERs are approximately equally abundant. By $z \sim 0$ however, the Seyfert population has declined relative to the LINER population by over a factor of 7.

It is important to remember that such measures are difficult to make with current data, and hence we remain limited by statistics to the extent to which we can physically interpret our results. Regardless, a robust outcome of our analysis is the differences between LINER and Seyfert AGN populations in high density regions, and between high and low redshift in all environments. Our results indicate that a greater understanding of both AGN and galaxy evolution may be possible if future analyses simultaneously focus on the detailed subdivision of different AGN classes, host galaxy properties, and their environment.

Part III

Future spectroscopic surveys

7

Fiber Assignment in Next-generation Wide-field Spectrographs

7.1 Introduction

In the next decade astronomers will be challenged to constrain the nature of dark matter, dark energy and the perhaps inflationary processes which generated structure in the Universe. Much effort will also be devoted to shedding light into the physics of galaxy formation and evolution. Observationally, large-scale galaxy surveys have consolidated as the most efficient way to approach these and other fundamental studies. Currently, the complexity and the scale of the phenomena involved in these studies demand progressively larger and deeper galaxy samples. In addition, more accurate measurements of fundamental physical properties from astronomical objects (such as kinematics, temperature, gravity/mass, chemical abundances or age) are needed. These requirements can only be satisfied through spectroscopy, which has become critical to further astrophysical understanding. There is a growing awareness in the community that answering many of the pressing astrophysical and cosmological questions of the coming decades will require undertaking vast deep spectroscopic surveys, mainly on large aperture telescopes (see Peacock & Schneider 2006; Bell et al. 2009). The development and optimization of spectroscopic survey techniques and strategies represents therefore a necessary step in order to increase the efficiency of our experiments and provide theorists with more accurate observational constraints.

Multi-object spectroscopy can be performed with traditional slits or modern optical fibers. Both approaches present advantages and disadvantages, but the latter provides much more versatility in terms of object collection and spectral coverage, which makes it more suitable for large-scale spectroscopic facilities. A number of fiber-fed spectrographs intended for large-scale galaxy surveys have been proposed recently. A good example is the Wide-Field Multi-Object Spectrograph (WFMOS, Bassett et al. 2005), which was proposed for the 8.2-m Subaru Telescope and aimed for a detailed investigation into the nature of dark energy and into galaxy formation and evolution. Similar scientific motivations supported the proposal of a wide-field fiber-fed spectrograph, the Super Ifu Deployable Experiment (SIDE, Prada et al. 2008), for the 10-m Gran Telescopio Canarias. None of these instruments were

finally built but they are excellent examples of state-of-the-art survey spectrographs aiming to fulfill next-generation scientific requirements. Several spectrographs have otherwise been accepted for construction, such as the Fibre Multi-Object Spectrograph (FMOS, Kimura et al. 2010), a near-infrared instrument which is already mounted on the Subaru Telescope. Already in the last stages of commissioning is also the set of 16 multi-fiber spectrographs for the Large Sky Area Multi-Object Fibre Spectroscopic Telescope (LAMOST, Wang et al. 2009). LAMOST was conceived to carry out several wide-field spectroscopic surveys focusing on both the structure of the Milky Way and the large-scale structure of the Universe. Finally, as an example of next-generation spectroscopic facilities, we will mention the Big Baryonic Oscillation Spectroscopic Survey (BigBOSS, Schlegel et al. 2009b). BigBOSS is a proposed ground-base dark energy experiment to study baryon acoustic oscillation with an all-sky galaxy redshift survey, making use of a multi-object fiber-fed spectrograph on the Mayall 4-m telescope at KPNO.

For all fiber-fed spectrographs (as those mentioned above) a primordial and common technical problem is the positioning of fiber ends, which must match the projected position of objects in the focal plane. In order to solve this difficulty, the concept used in most recently-proposed fiber-fed spectrographs (SIDE, LAMOST, BigBOSS, etcetera) consists of an array of fiber positioners covering the entire focal plane which is able to position all fiber heads simultaneously. This solution reduces drastically the reconfiguration time of the system as compared to the most common alternative based on a pick-and-place device. Important for this work, in a real survey normally a single configuration of the array of positioners is not enough to observe all (or even a given required fraction) of the objects in a target field. Several configurations (or tiles) are needed to reach a given completeness, depending on the typical number of objects per fiber positioner. The purpose of this chapter is to present an optimized way to assign fibers to objects so that the maximum number of objects is assigned in the first tiles. We also discuss on some additional ways of optimizing the fiber positioning process, depending on the capabilities of the instrument.

This chapter is organized as follows. In Section 7.2, we define some important concepts and provide the nomenclature that we use throughout this work. In Section 7.3, we briefly describe a general system consisting of an array of positioners covering the intrument focal plane. In Section 7.4, we present our optimized fiber positioning algorithm, that we call draining algorithm, and assess its performance as compared to a random assignment in catalogs of randomly distributed objects. Section 7.5 is dedicated to evaluating further optimizations such as rotation of the focal plane. In Section 7.6, we test the efficiency of our optimizations in mock galaxy catalogs drawn from cosmological simulations. Finally, in Section 7.7, we summarize our main results and discuss on their implications.

7.2 Nomenclature and definitions

In order to facilitate the reader's comprehension we will first introduce some basic concepts that will be frequently quoted in this chapter. Namely:

- **Target:** Each of the objects that we plan to observe.
- **Target field:** An area on the sky that we plan to observe.

- **Target density:** Total number of targets per deg^2 in a given target field.
- **Position angle (PA):** The angle that describes the rotation of the focal plane around an arbitrary fixed axis perpendicular to it.
- **Tile:** A spectrograph exposure at a given telescope pointing and position angle.
- **Tile density:** Number of tiles per deg^2 in a given target field.
- **Completeness:** The minimum fraction of targets that we need to observe in a given target field in order to satisfy the scientific requirements of the survey.
- **Tiling:** The complete process of minimizing the number of tiles needed to observe a given target field with the requested completeness. This process not only involves finding the optimal number of tiles but also the position of each tile in the target field (see Blanton et al. 2003c).
- **Positioner:** A mechanical device for positioning a fiber in the desired location within a certain region of the focal plane of the instrument. The region that each positioner can patrol is called *patrol disc*.
- **Fiber density:** Number of fibers (hence positioners) per deg^2 in a given target field.
- **Target-to-positioner ratio (η):** The number of targets per positioner or, equivalently, the target density divided by the fiber density.
- **Fiber collision:** A conflict that occurs when two or more targets are so closely located in the focal plane that they cannot be observed simultaneously, due to the physical size of the positioners. The minimum separation that prevents a situation like this from happening depends on the geometry of the positioner.

Note that the main intention of this work is to present a complete and optimized method for fiber assignment. We will therefore, and unless otherwise stated, consider a target field of approximately the size of the focal plane of the spectrograph. Only rotation of the focal plane will be allowed as an additional optimization. The process of tiling itself, that involves positioning tiles on a large region of the sky (as compared to the size of a tile), will only be discussed in a qualitative way.

7.3 Focal plane description

The concept for the focal plane that we outline here follows a standard design that can be extrapolated to most future fiber-fed multi-object spectrographs. In these instruments, the focal plane is populated with an array of fiber positioners, distributed in a hexagonal pattern, so that a single device is used to position each fiber head in the desired location within its patrol disc. Each positioner is therefore devoted to observing a single target. In a few words, the fiber positioning robot is a collection of positioners, all identical, distributed over an array which covers the entire focal plane so that all possible positions can be reached by at least one positioner (see Fig. 7.3). In order to cover the whole focal plane, a certain degree of overlap between patrol discs is needed so some regions of the

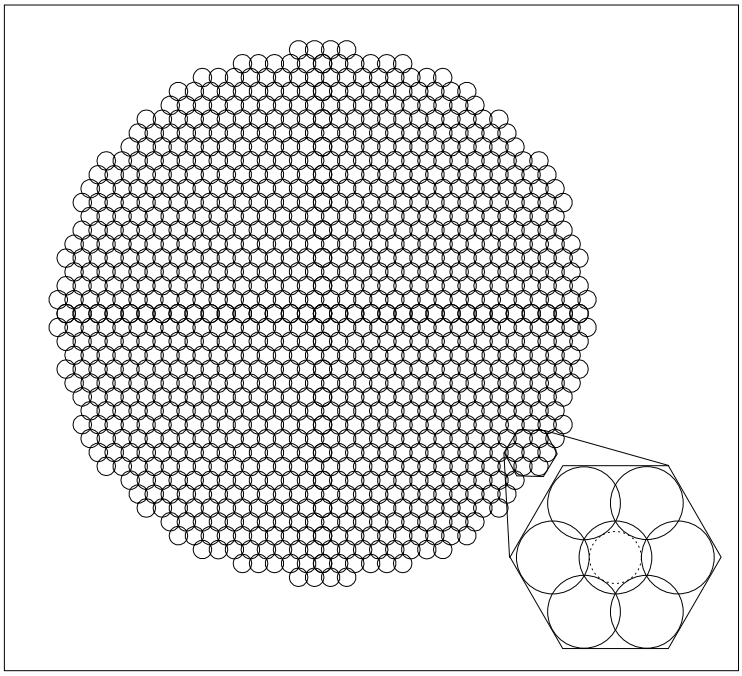

Figure 7.1 Sketch of a fiber positioning robot represented by an array of patrol discs covering the entire focal plane. Each positioner of the robot is dedicated to one of these patrol discs. This example features 1003 patrol discs for the 992-mm diameter field of view at the GTC. The detail shows a group of 7 patrol discs which overlap in a hexagonal pattern.

focal plane can be reached by more than one positioner (two or maximum three), hence rising the possibility of fiber collisions. In Fig. 7.2 we show a view of a real subset of a fiber positioner robot, with 19 positioners in hexagonal pattern (Azzaro et al., 2010).

This concept offers a number of advantages as compared to others. It is robust, scalable and easy to service and maintain (failure of one positioner causes the loss of one target only). Operationally, it provides extremely short reconfiguration times and allows an efficient real-time correction of differential atmospheric dispersion. It is, however, specifically conceived for wide-field surveys. The reason is that positioners cannot be densely packed onto a small portion of the focal plane. Consequently, the system is efficient for rather uniform distributions of targets or, alternatively, for large areas where any possible bias is smoothed. Find a complete discussion on this fiber positioning concept in Azzaro et al. (2010).

The results presented in this chapter on the optimization of the fiber assignment process are based on a complete simulation of the focal plane of the SIDE spectrograph, which adopted a fiber positioning robot as that described above. SIDE is an example of a state-of-the-art fiber-fed instrument capable of efficiently undergoing next-generation large-scale spectroscopic surveys. The focal plane of the SIDE spectrograph was designed as an array of 1003 positioners covering the 20-arcmin field of view at GTC. Important for this work, the results obtained with this simulation are valid for any focal plane

Figure 7.2 A view of a subset of a positioner array mounted on its holder. Here, 19 units are shown as an example.

Table 7.1 SIDE and BigBOSS focal plane parameters.

	FOV (deg^2)	Target density ($targets/deg^2$)	Number of positioners	Fiber density ($fibers/deg^2$)	η
SIDE	0.085	11500	1003	11500	1
BigBOSS	7	3500	5000	714	5

consisting in an array of positioners as that outlined in this section. The key parameter to describe the efficiency of the fiber assignment process, as we will see below, is the target-to-positioner ratio, η (instead of other parameters such as the size of the focal plane or the number of positioners). In order to better illustrate our results we will also discuss another real-life example: BigBOSS. In Table 7.1 we list some of the relevant parameters for both SIDE and BigBOSS.

7.4 Fiber assignment algorithm

In this section we briefly describe the philosophy behind our optimized fiber positioning algorithm: the draining algorithm. We also discuss its performance as compared to a simple random approach. Let us consider a generic instrument with a fiber positioning robot like that described in the previous section covering the entire focal plane with n positioners. A set of targets is used to populate the focal plane according to a given target-to-positioner ratio. We will first assume that targets are randomly distributed in the focal plane. The design of the system considered implies that each target is reachable

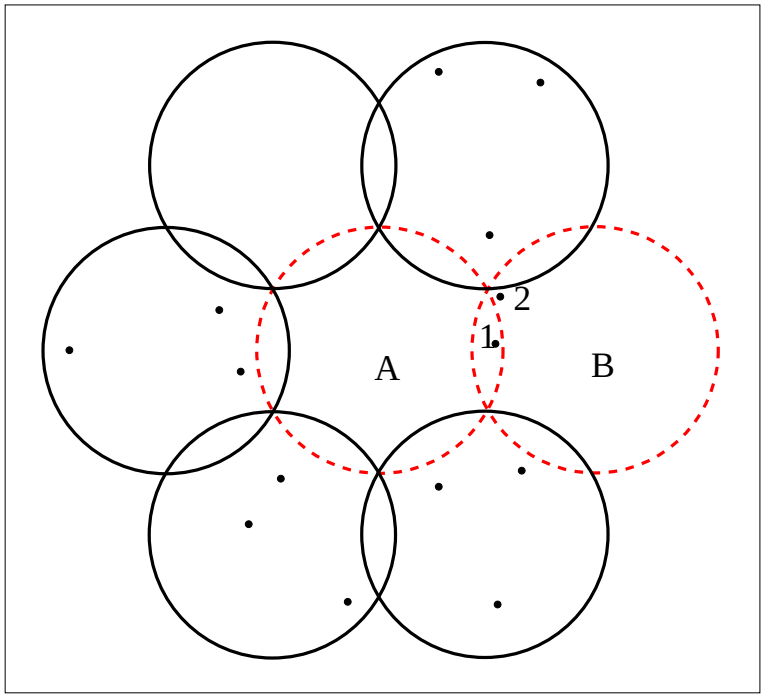

Figure 7.3 Sketch showing an array of 7 patrol discs which illustrates the advantages of the draining algorithm presented in this chapter. Note that an algorithm based on proximity would assign target 1, which lies in a common area shared by positioner A and B, to positioner B, consequently preventing operation of positioner A. This could also happen with a random assignment. The draining algorithm, however, would assign target 1 to positioner A and target 2 to positioner B, thus causing both targets to be observed in the current tile (consequently maximizing the number of assignments).

by one, two or maximum three positioners.

The motivation for optimizing the fiber assignment process derives from the fact that a single tile is normally not enough to observe all targets at a given pointing, or even a fraction of them. At this point and for the sake of understanding how the fiber assignment algorithm works, it is convenient to consider tiles as the different exposures that are needed to observe a given fraction of all targets at a given pointing. Tiles should be seen in this context, therefore, as iterations in the fiber assignment process. In practice, the number of tiles needed at a given pointing is basically set by η and the required completeness. As we will see below, only in cases where a very small η (~ 0.5) and a relatively low completeness ($\sim 80\%$) are required, will a single iteration (tile) be sufficient. These are of course very rare cases. The main goal here is to maximize the fraction of all targets observed after a given number of tiles.

Several strategies can be developed in order to assign targets to positioners. The simplest approach is to select randomly for each positioner any of the targets lying within the corresponding patrol disc. We could also impose a proximity criterion for the assignment, as that conceived for LAMOST. Instead of that, our motivation is to achieve an optimal solution, where optimal here means that the maximum number of targets is observed in the first tiles. The idea is to concentrate as many targets as possible in the first tile, and then repeat the process for the second and subsequent tiles until the required

completeness is reached. We want to move, whenever possible, targets towards the first tiles, and this is possible because some of the targets will fall into the common areas of the patrol discs (where they can be observed by two or, in a few cases, three positioners). In essence, the following steps must be implemented:

- For each target, identify the positioners which can access it.
- For each positioner, create a list of targets to be observed. In case a target can be observed by more than one positioner, it must be assigned to the positioner with the shortest list. The first target of this list would be observed in the first tile (iteration), the second target in the second tile and so forth.
- Optimize the assignments. The above configuration can be optimized by moving targets between different lists (objects can be moved between lists if they belong to more than one list, i.e. they are observable by more than one positioner). The idea is to shorten the long lists in favor of the shorter ones, so that all lists tend to the same length. In practice, we would search list by list for targets that could be moved to shorter lists. This process must be iterated until the optimal configuration is achieved.

To illustrate typical situations where the draining algorithm is clearly advantageous we show in Fig. 7.3 a simple sketch of an array of 7 positioners, represented by their corresponding patrol discs. Note that in a real focal plane each patrol disc overlaps with 6 other patrol discs (obviously excluding those in the border). In this example most of the targets (dots) sit inside exclusive areas, i.e. regions that are only accessible by a single positioner. Target 1, however, lies in the common region shared by positioner A and positioner B. An algorithm based on proximity would assign target 1 to positioner B, consequently preventing operation of positioner A. This could also happen with a random assignment. The draining algorithm proposed here, however, would assign target 1 to positioner A and target 2 to positioner B, thus forcing both targets to be observed in the current tile and consequently increasing the number of assignments. Our tests indicate that, over a large number of positioners, situations like this happen with relevant frequency so the draining algorithm achieves a sensibly better performance than other approaches, as we will show below. Following this idea, it is rather straightforward to identify the conflicting cases and formalize this method into an algorithm so that the optimal sequence of tiles is found.

In order to assess the performance of the draining algorithm, we have generated a set of 100 catalogs of randomly distributed targets for systems with four different target-to-positioner ratios, namely $\eta = 0.5, 1, 3, 5$. Note that, as shown in Table 7.1, the values of $\eta = 1$ and $\eta = 5$ correspond to the two examples considered in this paper, i.e. SIDE and BigBOSS, respectively. We have implemented both the draining algorithm and a simple approach where targets are randomly assigned to positioners. In Table 7.2, we show the average total fraction of all targets assigned to positioners after each of the first 10 tiles (cumulative fractions) using both these methods, for the target-to-positioner ratios considered. The tile at which 80% of the total number of targets is successfully assigned to positioners has been highlighted for each η. The number of tiles needed to reach this completeness (that we have selected arbitrarily) obviously increases with the target-to-positioner ratio, ranging from only

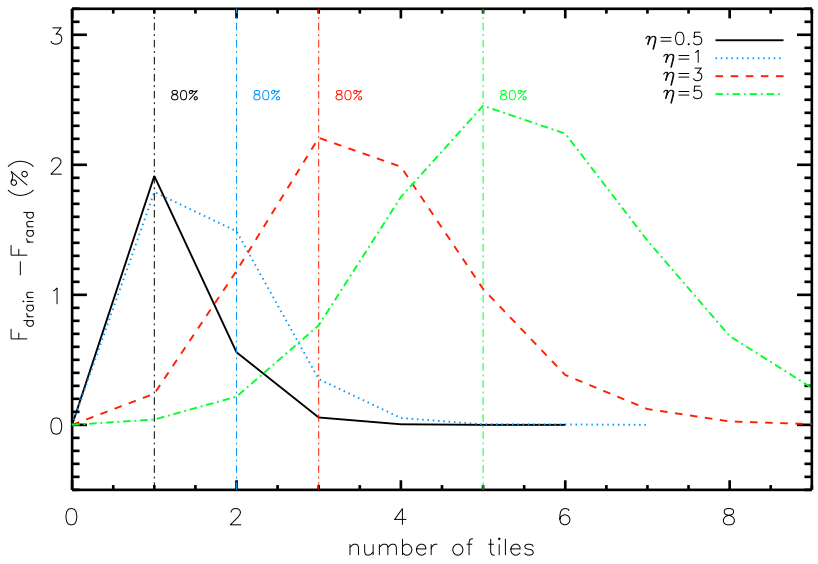

Figure 7.4 Difference in the cumulative fraction of targets assigned with the draining algorithm described in Section 7.4 ($F_{\rm drain}$) and with a random approach ($F_{\rm rand}$) shown as a function of the number of tiles for different target-to-positioner ratios, η. Vertical lines represent the number of tiles needed to reach the 80% completeness level for each η. Simulations have been performed with 100 catalogs where targets are randomly distributed across the focal plane.

Table 7.2 Detailed comparison between the average cumulative fractions of targets assigned with the draining algorithm and with a random approach as a function of the number of tiles for different target-to-positioner ratios, η. For each η, the tile at which the 80% completeness level is highlighted in bold fonts. Simulations have been performed with 100 catalogs where targets are randomly distributed across the focal plane. Standard deviations of the cumulative fractions are shown in brackets. We also show results obtained with our draining algorithm combined with rotation of the focal plane, i.e. draining + ROT.

Number of tiles	Method	$\eta = 0.5$	$\eta = 1$	$\eta = 3$	$\eta = 5$
1	Draining	**0.837 (0.027)**	0.687 (0.013)	0.336 (0.002)	0.209 (0.001)
	Random	**0.817 (0.025)**	0.669 (0.012)	0.334 (0.002)	0.208 (0.001)
	Draining + ROT	**0.866 (0.027)**	0.707 (0.011)	0.339 (0.001)	0.211 (0.000)
2	Draining	0.983 (0.031)	**0.937 (0.017)**	0.626 (0.004)	0.411 (0.001)
	Random	0.977 (0.030)	**0.922 (0.017)**	0.614 (0.004)	0.409 (0.001)
	Draining + ROT	0.998 (0.031)	**0.966 (0.016)**	0.644 (0.003)	0.417 (0.001)
3	Draining	0.998 (0.032)	0.989 (0.019)	**0.832 (0.007)**	0.599 (0.002)
	Random	0.998 (0.031)	0.986 (0.019)	**0.809 (0.007)**	0.592 (0.002)
	Draining + ROT	1 (0.031)	0.999 (0.017)	**0.866 (0.006)**	0.615 (0.001)
4	Draining	0.999 (0.032)	0.998 (0.019)	0.940 (0.009)	0.762 (0.003)
	Random	0.999 (0.031)	0.998 (0.019)	0.920 (0.008)	0.744 (0.003)
	Draining + ROT		1 (0.017)	0.971 (0.008)	0.788 (0.003)
5	Draining	0.999 (0.032)	0.999 (0.019)	0.982 (0.009)	**0.881 (0.005)**
	Random	0.999 (0.031)	0.999 (0.019)	0.972 (0.009)	**0.856 (0.006)**
	Draining + ROT			0.998 (0.009)	**0.913 (0.004)**
6	Draining	1 (0.032)	0.999 (0.019)	0.995 (0.010)	0.950 (0.006)
	Random	1 (0.031)	0.999 (0.019)	0.991 (0.009)	0.928 (0.006)
	Draining + ROT			1 (0.009)	0.978 (0.006)
7	Draining		1 (0.019)	0.999 (0.010)	0.981 (0.007)
	Random		1 (0.019)	0.997 (0.009)	0.967 (0.006)
	Draining + ROT				0.997 (0.006)
8	Draining			0.999 (0.019)	0.994 (0.007)
	Random			0.999 (0.009)	0.987 (0.006)
	Draining + ROT				0.999 (0.006)
9	Draining			0.999 (0.010)	0.998 (0.007)
	Random			0.999 (0.009)	0.995 (0.006)
	Draining + ROT				1 (0.006)
10	Draining			0.999 (0.010)	0.999 (0.007)
	Random			0.999 (0.009)	0.998 (0.006)
	Draining + ROT				

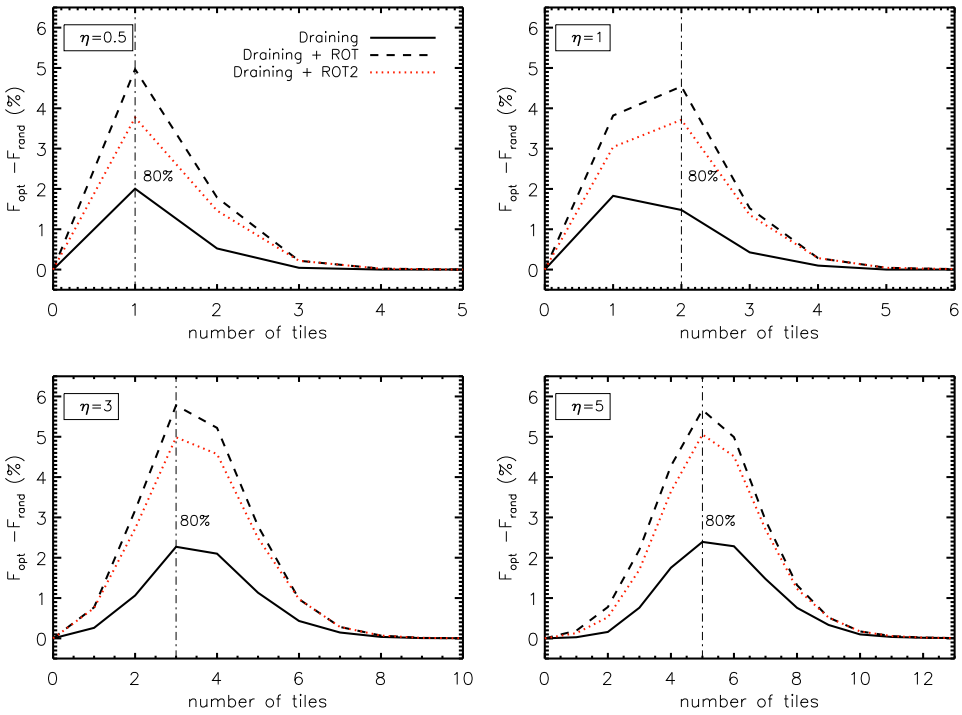

Figure 7.5 Cumulative percentage of targets gained with respect to a random assignment by using the draining algorithm described in Section 7.4 alone (solid line) and the same algorithm allowing rotation of the focal plane as discussed in Section 7.5 (including ROT and ROT2), as a function of the number of tiles for different target-to-positioner ratios, η. Vertical lines represent the number of tiles needed to reach the 80% completeness level for each η. Simulations have been performed with 100 catalogs of randomly distributed targets.

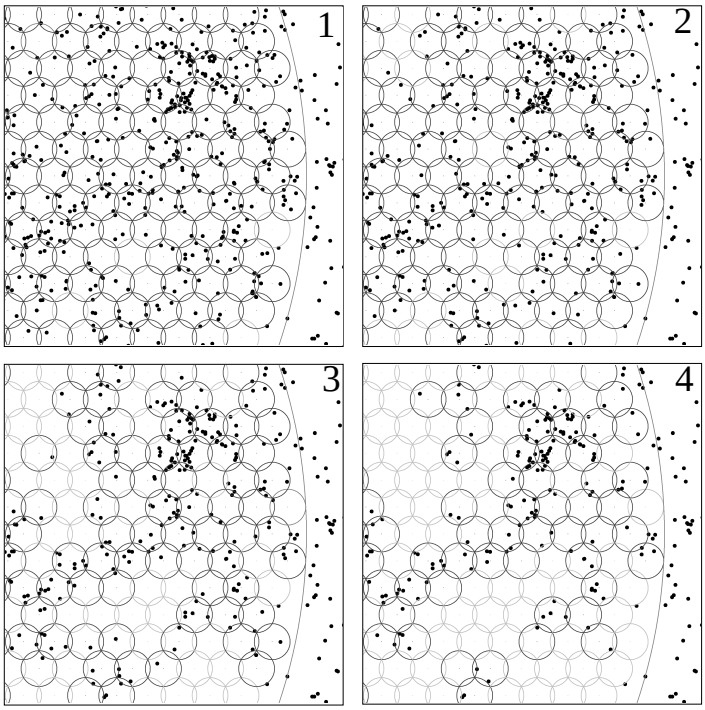

Figure 7.6 Portion of the focal plane of an instrument, represented by an array of patrol discs, in a sequence of the first tiles. The over-plotted dots symbolize the projected positions of targets belonging to a mock galaxy catalog with a target-to-positioner ratio $\eta \sim 3$. The first panel in this figure (upper-left panel) illustrates the initial situation where the entire sample of targets is to be observed. Patrol discs with at least one target inside have been highlighted. The draining algorithm maximizes the number of targets observed in each tile. In this example, $\sim 75\%$ of targets have been selected after 3 tiles.

1-2 tiles for $\eta \leq 1$ to up to 5 tiles for $\eta = 5$. At this point it is necessary to remind that the scope of this optimization is not to reduce drastically the average number of tiles needed to achieve a certain completeness, but to increase the fraction of targets observed in the same number of tiles. As we can see in Table 7.2, the draining algorithm presented here increases the assignments in approximately 2% as compare to a greedy procedure at the 80% completeness level.

Fig. 7.4 also illustrates the performance of the draining algorithm as compared to a random assignment. Here, the difference in the cumulative percentage of targets assigned with our optimized algorithm (F_{drain}) and with a random approach (F_{rand}) is shown as a function of the number of tiles. Fig. 7.4 shows that the gain provided by this optimized method increases slightly with η, ranging from $\sim 1.7\%$ for $\eta = 0.5$ to $\sim 2.5\%$ for $\eta = 5$. Note that, due to the nature of this optimization, the improvement occurs in the first few tiles. In the following sections we will show how an improvement like this can have a strong effect in the design of a survey.

An important issue related to the fiber assignment process in this type of fiber-fed spectrographs is the so-called fiber collision problem. This concept was used to describe the fact that in the SDSS (York et al., 2000) fibers could not be placed closer than $55''$ arcsec. Spectrographs that feature a fiber positioning robot as that described in Section 7.3 present a similar problem, as due to the

physical size of the positioners, frequently two or more targets cannot be observed simultaneously (i.e. in the same tile). The results shown in this work were obtained assuming that positioners have no physical size, hence neglecting the possibility of fiber collisions. The typical number of collisions as a function of the target-to-positioner ratio, η, will depend on the geometry of the system, i.e. on the number and the size of the positioners. The problem, however, can be solved naturally within the framework of the draining algorithm. Fiber collisions will occur exclusively in common areas and this method is specifically designed to deal with objects in these regions conveniently, and to ensure that the maximum number of targets is observed in the first tiles. We have checked that the draining algorithm alone can solve naturally almost all collisions, so the results presented in Fig. 7.4 and Table 7.2 are not affected by our assumption. Note that the number of these events in catalogs of randomly distributed targets is absolutely negligible. Even in real catalogs, where fiber collisions are more frequent, the draining algorithm provides an optimal solution, as we will show in Section 7.6.

7.5 Additional optimizations

So far we have presented an optimized method for fiber positioning, the draining algorithm, that can be easily implemented in any future fiber-fed spectrograph. In this section we explore additional ways to optimize the number of targets observed that require some pre-designed capabilities from our instrument. In particular, we will demonstrate that it is possible to achieve a much better performance by simply allowing the focal plane of the instrument to rotate.

In Section 7.2 we defined the position angle, PA, describing the rotation of our focal plane. Note that an array of positioners as that described in Section 7.3 is geometrically symmetric under PA rotations of 60^o (see an example in Fig. 7.3). Let us first assume that our instrument is capable of rotating the focal plane up to this angle and consequently all possible configurations are accessible. Rotating the focal plane implies that the number of targets sitting inside each patrol disc changes. More importantly, it changes the number of positioners with at least one target within the patrol disc. It seems convenient, therefore, to find the configuration that optimizes the efficiency of the assignment process. In this section, we combine this position angle optimization (ROT optimization) with the draining algorithm (draining + ROT).

The draining algorithm is a simple method which optimizes the fiber assignment process by ensuring that the maximum number of targets is assigned in the first tiles. Such a method provides an optimized solution for our observation plan that normally includes several tiles , depending on η and the completeness requested. The easiest way to insert the ROT optimization into this scheme is to simply allow for rotation at the beginning of the process and find the configuration that maximizes the number of positioners with at least one target in the patrol disc. However, it is obviously much more efficient to permit rotation not only at the beginning of the process but also between tiles (iterations of the process). This ensures that the maximum number of positioners is in use in each tile and requires a trivial modification of the above scheme. Now, the draining algorithm is applied "from scratch" in each iteration of the process on the targets left unobserved from the previous iteration and after the optimal PA is found. Obviously, only the first tile from the solution provided by our optimized algorithm will eventually be performed for each step. This does not jeopardize the proper optimization

Table 7.3 The same as in Table 2 but using mock catalogs extracted from the Bolshoi Simulation. For each target-to-positioner ratio, η, we used a set of 9 catalogs.

Number of tiles	Method	$\eta = 0.5$	$\eta = 1$	$\eta = 3$	$\eta = 5$
1	Draining	0.737	0.599	0.321	0.207
	Random	0.718	0.585	0.319	0.206
	Draining + ROT	0.766	0.622	0.327	0.208
2	Draining	**0.934**	**0.855**	0.576	0.398
	Random	**0.923**	**0.840**	0.566	0.394
	Draining + ROT	**0.969**	**0.897**	0.599	0.406
3	Draining	0.983	0.950	0.755	0.565
	Random	0.978	0.940	0.738	0.556
	Draining + ROT	0.999	0.980	0.791	0.584
4	Draining	0.996	0.984	**0.863**	0.701
	Random	0.995	0.979	**0.847**	0.687
	Draining + ROT	1	0.998	**0.905**	0.730
5	Draining	0.9999	0.994	0.926	**0.803**
	Random	0.9999	0.993	0.913	0.786
	Draining + ROT		1	0.962	**0.838**
6	Draining	1	0.998	0.960	0.874
	Random	1	0.997	0.951	**0.856**
	Draining + ROT			0.986	0.909
7	Draining		1	0.978	0.921
	Random		0.999	0.972	0.905
	Draining + ROT			0.995	0.952
8	Draining		0.999	0.988	0.950
	Random		0.999	0.984	0.938
	Draining + ROT			0.998	0.976
9	Draining		1	0.993	0.968
	Random		1	0.991	0.959
	Draining + ROT			0.999	0.989
10	Draining			0.996	0.980
	Random			0.995	0.972
	Draining + ROT			0.999	0.995

of the process by any means, as the algorithm is specifically designed so that the maximum number of assignments is achieved in the first tile.

We have assessed the relevance of the ROT optimization in the same 100 catalogs of randomly distributed targets that we used in Section 7.4. In order to isolate the effect of rotation on the efficiency of the process, we need to conveniently address what we call the "border scan effect". Usually the target field is a circle on the sky encompassing the most external patrol discs of the array of positioners. As we rotate the focal plane, some of the targets lying between the outermost perimeter of the patrolled area, which is not a circle, and the border of the target field can become accessible to some positioners. It is important not to overestimate the improvement by including targets which would be out of reach of the patrolled area if no rotation were applied. In order to avoid this "contamination", we have restricted ourselves to an area of the focal plane where the total number of targets remains constant under any rotation.

In Table 7.2 we presented the average cumulative fractions of targets assigned with the draining algorithm and with a random approach in the first 10 tiles for different values of η. We also show for comparison the same fractions obtained using the combined draining + ROT algorithm described in this section. Interestingly, allowing for rotation reduces drastically the total number of tiles needed, namely only 4 instead of 7 for $\eta = 1$, as an example. This is not critical as far as a real observation is concerned, unless a 100% completeness were required. More importantly, the increase in the efficiency of the assignment in the relevant tiles provided by the PA optimization is very significant as compared to the case were rotation is not permitted: $3-4\%$ even if we were to use the draining algorithm and $5-6\%$ if we adopted a greedy aproach. This improvement is shown in a clear way in Fig. 7.5, which is similar to Fig. 7.4 but including the cumulative percentage of targets gained with respect to a random approach when rotation of the focal plane is allowed.

There are instruments where only slight rotations of the focal plane are permitted. As an example, the BigBOSS spectrograph is expected to be capable of rotating its focal plane by no more than $\sim 2^o$. We have studied the effect that a limited ROT optimization would have in the efficiency of the fiber assignment process by allowing our simulated focal plane to rotate a maximum of $\pm 2^o$ (ROT2 optimization). In Fig. 7.5 we show in a dotted line the result of applying this test on our set of random catalogs. Interestingly, even a small rotation like this produces a remarkable optimization: $3.5-5\%$ at a completeness level of 80%. Also, an alternative to the ROT optimization for instruments where rotation of the focal plane is not permitted might be found in tweaking the telescope pointing position around the center of the target field. Again, this should lead to configurations where targets are more conveniently spread over the array of positioners, consequently increasing the efficiency of the assignment process. It is therefore reasonable to expect an optimization like this to produce similar results as those found with the ROT optimization.

7.6 Application to mock galaxy catalogs

In previous sections we discussed the performance of our optimized algorithm for fiber positioning (including rotation or, alternatively, small shifts of the focal plane of our instrument) when implemented in catalogs of randomly distributed targets. However, it is well known that galaxies in the

Universe are far from being randomly distributed. Galaxies are actually clustered, forming filaments in the 3-dimensional space that are separated from each other by under-dense regions. In a similar manner would targets be distributed in the 2-dimensional projection of this space onto the focal plane of our instrument. Consequently, in a realistic situation, targets would accumulate in certain regions of our focal plane, leaving others relatively underpopulated. We have simulated the performance of our optimizations in real-life situations by making use of galaxy mock catalogs extracted from the Bolshoi simulation (Klypin et al., 2010). Important for this work, the clustering properties of mock galaxies match those of real galaxies with good accuracy. The reason for using mock catalogs instead of real catalogs is that they provide us with more flexibility when it comes to selecting different samples with different number densities. We have performed a 2-D projection of a simulation box of 250 Mpc/h on a side situated at redshift 1 onto our focal plane and used the circular velocity of haloes as an empirical threshold for selecting different densities. In order to allow for a fair comparison with the results presented in the previous sections, we have selected catalogs with target-to-positioner ratios of ~ 0.5, 1, 3 and 5. Simulations were performed with 9 different realizations for each η, obtained by rotating the box conveniently.

In Fig. 7.6, we show a portion of the focal plane of an instrument, represented by an array of patrol discs, in a sequence of the first tiles. The over-plotted dots symbolize the projected positions of targets belonging to a mock galaxy catalog with $\eta \sim 3$. The first panel in this figure illustrates the initial situation where the entire sample of targets is to be observed. Note that, as explained above, targets on the focal plane of the spectrograph, rather than being randomly distributed, accumulate in filaments. The rest of the panels show the distribution of targets assigned to each tile by using the optimized algorithm described in Section 7.4. In this example, $\sim 75\%$ of targets have been selected after 3 tiles.

The performance of our optimizations with mock galaxy catalogs is presented in the same format as in previous sections in Table 7.3 and Fig. 7.7. A direct consequence of the presence of clustering in our catalogs is that the fraction of targets assigned to each tile decreases significantly, as a comparison between Table 7.2 and Table 7.3 demonstrates. Trivially, the probability that a single positioner has to deal with several targets is now higher and, consequently, it becomes harder to move targets towards the first tiles. As an example, with $\eta \sim 1$ and two tiles we could observe almost 94% of all targets in a random catalog, even whithout allowing for rotation of the focal plane. In a real catalog we could only assign 85% of targets in the same number of tiles. Similarly, in a real catalog with $\eta \sim 5$ we would need 5 tiles to barely reach a completeness of 80%, at least 8% below our expectations from random catalogs. Again, we refer the reader to Table 7.3 for the exact fractions of targets assigned with a random approach, with the draining algorithm alone and with the draining algorithm complemented with rotation of the focal plane. In order to analyze the performance of our optimizations as compared to a random approach we point the reader to Fig. 7.7. This figure shows that the gain provided by our optimizations as compare to a greedy approach when implemented in a real-life situation is consistent with what we could infer from random catalogs. A closer inspection, however, reveals that the gain provided by the draining algorithm alone is slightly smaller now, falling below 2%, whereas the gain achieved by combining this algorithm with a ROT optimization remains in the range of $5 - 6\%$. If only slight rotations were allowed (ROT2 optimization) we could still improve the efficiency of the

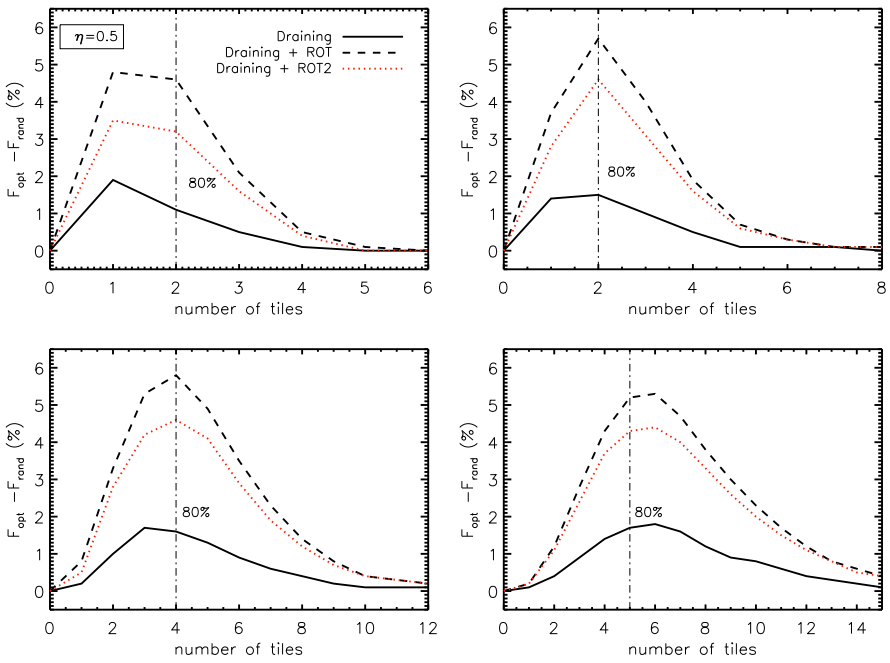

Figure 7.7 The same as in Fig. 7.5 but using galaxy mock catalogs extracted from the Bolshoi simulation. For each target-to-positioner ratio, η, 9 mock catalogs were used.

process in $3.5 - 4.5\%$.

Note that, as mentioned previously, the results shown in this chapter were obtained ignoring fiber collisions, as the effect of these depends on the geometry of the fiber positioning robot itself, and, therefore, cannot be extrapolated to any fiber-fed spectrograph of this kind. The typical number of these events obviously increases in mock galaxy catalogs (and hence in real catalogs) due to the fact that objects are more clustered (the fraction of collisions is almost negligible in random catalogs). Even in real catalogs the fraction of objects in conflict remains small for SIDE: $\lesssim 1\%$ even for $\eta = 5$. However, an important advantage of the draining algorithm is that collisions can be solved optimally, and additional methods are not needed. The results presented in Table 7.3 and Fig. 7.7 would basically not change if collisions were taken into account. Only very slight variations are expected in the very last tiles, which are not relevant in real-life observations.

The results presented in this section confirm that, despite the physical restrictions of this state-of-the-art fiber positioning robots, it is possible to optimize a survey strategy in a remarkable way just by assigning targets to tiles and positioners conveniently. In addition, these results represent a strong support for allowing the focal plane of future instruments to rotate. We have shown how even a slight rotation produces a remarkable optimization. In the next section we discuss on the implications of our results.

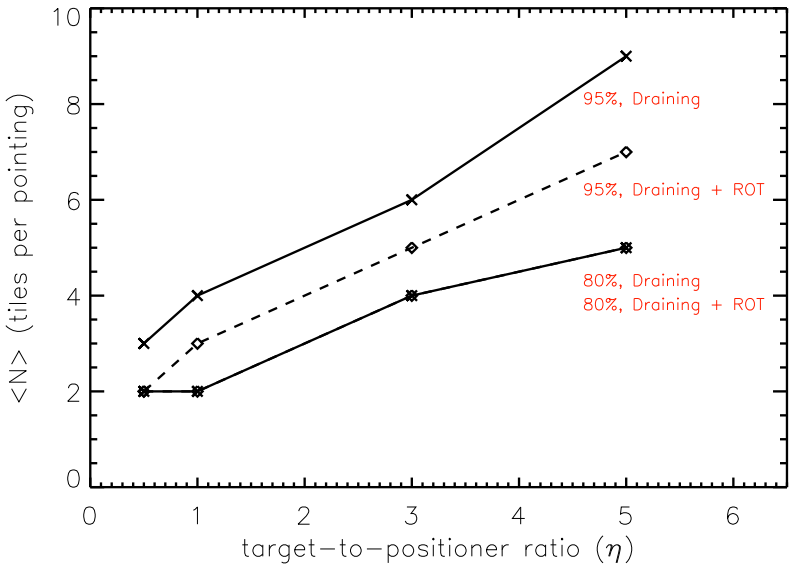

Figure 7.8 Average number of tiles per pointing, $< N >$, as a function of the target-to-positioner ratio, η, for 2 different completeness levels: 80% and 95%

7.7 Discussion and chapter conclusions

The results on the optimization of the fiber positioning process that we present in this chapter are valid for any focal plane consisting of an array of positioners as that described in Section 7.3. To first order, therefore, our results are not dependent on the size of the tile or the number of fibers per tile, as long as positioners are arrayed in a hexagonal pattern as that shown in Fig. 7.3. This is a standard concept in state-of-the-art fiber-fed spectrographs, such as some of the ones proposed for the GTC or others like BigBOSS, LAMOST, etcetera. Let us illustrate the relevance of our results with a simple example based on the future survey BigBOSS. BigBOSS is planned to map an area of approximately $14,000$ sq. deg. The expected target density reportedly hovers around $3,500$ $targets/deg^2$ with a required completeness of $\sim 80\%$. This would yield a huge survey, of a few tens of millions of objects. According to our estimations, by using the draining algorithm presented here instead of a random assignment, we could save for observation $\sim 2\%$ of all targets, that is, several hundred thousand objects (expectedly $> 700,000$). Note that a sample like this would by itself triple the size of the 2dF Galaxy Redshift Survey (Colless et al., 2001) and would be comparable to the SDSS. If we could take advantage of the fact that the BigBOSS instrument allows for slight rotations of the focal plane, the above gain would increase in at least a factor 2, which represents a remarkable improvement. The increase in the efficiency of the fiber positioning process that these optimizations provide could also reduce the costs of the survey, in terms of observing time. In fact, by simply implementing the draining algorithm and according to our estimations, we could collect the same number of objects in a remarkably smaller area: between ~ 350 deg^2 and $\sim$ 700 deg^2 smaller depending on the feasibility of rotation. Assuming that we need an average of 5 tiles per pointing, that would mean saving between ~ 50 and ~ 100 tiles.

The intention of this work is not only to present some remarkable fiber positioning optimizations

but also to provide some basic guidelines for an efficient instrument requirement definition for fiber-fed multi-object spectrographs. Fig. 7.8 is intended to summarize some of our main results in a way that may be useful for future surveys. In this figure we show the average number of tiles per pointing, $< N >$, as a function of the target-to-positioner ratio, η, for 2 different completeness levels: 80% and 95%. These values define a reasonable range for the completeness required in a real survey. Note that η and the completeness requested determine completely (at least to first order) the average number of tiles per pointing and, consequently, the (approximate) tile density of our survey. The parameter η is the ratio of the target density to the fiber density. The target density is a requirement that strongly depends on the science case. The fiber density is restricted by current technology as there is a limit in the size of the positioner. Ideally, we want to increase our fiber density as much as possible, consequently decreasing η, so that fewer tiles are needed to complete our survey. This obviously reduces the observation time and consequently the cost of the survey.

Finally, we summarize the main results of this chapter as follows:

- We have presented an optimized algorithm for assigning fibers to targets in next-generation fiber-fed multi-object spectrographs: the draining algorithm. Our method is very easy to implement and ensures that the maximum number of targets in a given target field is observed in the first few tiles. Using both catalogs of randomly distributed targets and mock galaxy catalogs drawn from cosmological simulations we have estimated that the gain provided by the draining algorithm as compared to a random assignment can be as much as 2% for the first tiles.

- The fiber collision problem can be solved easily and in an optimal way within the framework of the draining algorithm.

- We also discuss additional optimizations of the fiber positioning process. In particular, we have shown that allowing for rotation of the focal plane can improve the efficiency of the process in $\sim 3.5 - 4.5\%$ even if only small adjustments are permitted (up to 2^o, ROT2 optimization). For instruments that allow large rotations of the focal plane (ROT optimization) the expected gain increases to $\sim 5 - 6\%$. These results, therefore, strongly support focal plane rotation in future spectrographs, as far as the efficiency of the fiber positioning process is concerned

- An alternative to the ROT optimization for instruments where rotation of the focal plane is not permitted might be found in tweaking the telescope pointing position around the center of the target field. We expect an optimization like this to produce similar results as those found with the ROT optimization.

- As an example of a future large-scale galaxy survey, we have discussed the implications of our optimizations if applied to BigBOSS ($\sim 14,000\ deg^2$, $\sim 40M$ objects). We have estimated that by using the draining algorithm instead of a random assignment we could save for observation several hundred thousand objects (expectedly $> 700,000$, a sample comparable to the SDSS). Alternatively, we could collect the total number of objects expected for the entire survey in a remarkably smaller area: $\sim 350\ deg^2$ smaller (equivalent to ~ 50 tiles). This figures would typically double if we could perform focal plane rotations of at least 2^o.

- We provide the average tile density as a function of the target-to-positioner ratio, η, and completeness, which are the fundamental quantities that defines a survey or the instrument requirements. These results are applicable to most next-generation fiber-fed multi-object spectrographs.

Part IV

Conclusions and future work

Conclusion remarks

Almost a century after Harlow Shapley and Herber Curtis debated on the scale of the observed universe at the Smithsonian Museum of Natural History in Washington, USA, our global vision of the Universe has experienced a dramatic change, thanks to the combined efforts of observational and theoretical astronomers, cosmologists, particle physicists and computer scientists. In the last decades, we have even attempted to unveil the formation of the most complex structures in the Universe: galaxies.

In fact, we have come up with a reasonable theory for the formation of galaxies within hierarchically growing dark matter halos. In such an achievement, both photometric and spectroscopic large-scale galaxy redshift surveys have played an important role, allowing us to characterize the low-to-moderate redshift galaxy population in a statistical fashion ($z \lesssim 0.6$). At higher redshifts, poor statistics and selection effects have hindered our efforts remarkably. However, important constraints on the evolution of the properties of galaxies have been obtained using pencil beam surveys, studying specific galaxy populations at different wavelength ranges or analyzing very deep images from the Hubble Space Telescope. From a wide range of observations and thanks to the development of stellar population synthesis, we have been able to establish statistically-significant correlations between galaxy properties at different redshifts, and the dependence of several of these on the surrounding environment. However, the physical processes driving fundamental changes on galaxies are diverse and complex. The challenge today is to identify these fundamental processes, by linking different galaxy populations at different redshifts and tracing the average evolution of galaxy properties.

Disentangling the galaxy formation and evolution picture requires accurate observational constraints at both low and high redshift. This dissertation is a compendium of different works intended to alleviate this necessity. In this sense, the state-of-the-art results that we have presented will surely have strong implications to current models.

The luminosity function of galaxies is one of the most fundamental statistical properties of the galaxy population, as it contains important information on the formation and evolution of galaxies. In Chapter 2, I took advantage of the huge increase in galaxy statistics provided by the SDSS DR6 as compared to earlier estimations (e.g. Blanton et al. 2003b with the SDSS, Norberg et al. 2002 with the 2dFRS) to present the most accurate estimation of the luminosity function to date, in a sample of $\sim 500,000$ galaxies. Two main updates have been found, as compared to the broadly-used SDSS luminosity function of Blanton et al. (2003b). Namely, we obtain steeper faint-end slopes in all SDSS photometric bands and a prominent excess of bright galaxies in the bluer bands, i.e. the bright-end bump. This bump seems to be comprised mainly by star-forming galaxies and AGN. These important

deviations have evident implications on current theories. In fact, the steeper faint-end slopes would imply that less feedback is required in dwarf galaxies. The bright-end bump may indicate that AGN feedback in massive galaxies is not as efficient as we thought. Interestingly, both these features are relatively consistent with the luminosity function obtained with the semi-analytical models of Bower et al. (2006).

The ALHAMBRA photometric survey is one of the current projects aiming to alleviate the lack of quality galaxy data at high redshift ($z \sim 1$) with a combination of good photo-z accuracy, photometric depth and large statistics. In particular, the ALHAMBRA photo-z accuracy will exceed that of previous photo-z surveys (such as COSMOS or COMBO-17), thanks to its photometric system, with 20 narrow-band filters covering the optical range plus the JHK bands in the near-infrarred. In Chapter 4, I used a preliminary catalog to analyze the main statistical properties of the ALHAMBRA low-redshift galaxy population, in a sample of $\sim 10,000$ galaxies with synthetic SDSS magnitude $i^A < 24.5$. Even though preliminary data was used, these observed properties are consistent with previous results from other surveys. Only in the bright-galaxy range have I found remarkable deviations, that seem to be caused by catastrophic photo-z errors. An important outcome of this analysis is the ALHAMBRA luminosity function, that has been estimated down to synthetic absolute magnitude $M_i^{0.1} \sim -13.5$. Interestingly, at the faint-end, where photo-z errors are expectedly not so relevant, the preliminary ALHAMBRA luminosity function presents a very steep slope, which seems to be in agreement with the work of Blanton et al. (2005a), who used a very local sample of the SDSS. These results, that must be confirmed in the future, would add important constraints on the formation of dwarf galaxies.

The low-z calibration of the ALHAMBRA survey presented in Chapter 4 lays the foundations for future galaxy evolution studies with ALHAMBRA. The expectedly good photo-z accuracy of ALHAMBRA will allow us, not only to explore the evolution of the luminosity function, but also to obtain fundamental galaxy properties from individual galaxies. In Chapter 5, I have proposed and described a methodology for estimating photometric stellar masses in ALHAMBRA, using the flexible and user-friendly FSPS code (Conroy et al., 2009). This method has been applied to the same low-z sample of ∼10,000 galaxies used in Chapter 4. The distributions of mass-to-light ratios and stellar masses obtained are consistent with broadly-accepted results from the SDSS. The stellar population synthesis method developed for ALHAMBRA can be used to obtain other fundamental properties of the stellar population, such as age, metallicity or SFH.

A common way to investigate the evolution of galaxies involves measuring correlations between fundamental galaxy properties at different redshifts. In Chapter 6, I have looked into the dependence of different type of nuclear activity among red galaxies on the density of the environment (i.e. the AGN-density relation) at both $z \sim 1$, with the DEEP2 Galaxy Redshift Survey, and $z \sim 0$ with the SDSS. I have provided evidence that high redshift LINERs in DEEP2 tend to favour higher density environments relative to the red population from which they were drawn. In contrast, Seyferts and X-ray selected AGN at $z \sim 1$ show little (or no) environmental dependencies within the same underlying population. At low redshift, both LINERs and Seyferts in the SDSS show a slowly declining red sequence AGN fraction towards high density environments. Interestingly, at $z \sim 1$ red sequence Seyferts and LINERs are approximately equally abundant. By $z \sim 0$, however, the red Seyfert population has declined relative to the LINER population by over a factor of 7. This was the first

systematic study on the evolution of this important observational constraint.

Finally, Chapter 7 was devoted to exploring survey optimizations in next-generation spectroscopic facilities. Even though photo-z surveys have proven capable of providing key insights on the high-redshift universe, answering the most fundamental questions about the formation and evolution of galaxies will require a spectroscopic follow-up. The prospection of future large-scale spectroscopic surveys, which can provide more accurate observational constraints to current theories, is therefore of great importance. In Chapter 7, I have focused on the optimization of the number of targets for next-generation fiber-fed wide-field spectrographs such as SIDE and BigBOSS. In particular, I have presented an optimized algorithm for assigning fibers to targets, i.e. the draining algorithm, that ensures that the maximum number of targets in a given target field is observed in the first few tiles (spectrograph exposures). Additional optimizations have also been discussed. In particular, focal plane rotation provides a remarkable increase in the efficiency of the fiber positioning process. These optimizations will surely contribute to increasing galaxy statistics in future surveys.

Consideraciones finales

Casi un siglo después de que Harlow Shapley y Herber Curtis se citaran para debatir acerca del tamaño del Universo en el Museo Smithsonian de Historia Natural de Washington, Estados Undos, nuestra visión global del Universo ha experimentado un cambio dramático, gracias al esfuerzo conjunto de astrónomos observacionales y teóricos, cosmólogos, físicos de partículas y científicos computacionales. En las últimas décadas, incluso nos hemos aventurado a estudiar la formación de las estructuras más complejas del Universo: las galaxias.

De hecho, hemos sido capaces de desarrollar una teoría que explica la formación de galaxias en el seno de halos de materia oscura que crecen de manera jerárquica. Éxitos como este han sido posibles en gran parte gracias a los cartografiados a gran escala de galaxias, tanto fotométricos como espectrscópicos, que nos han permitido caracterizar estadísticamente la población de galaxias a bajo y medio corrimiento al rojo ($z \lesssim 0.6$). A más altos corrimientos al rojo, la pobre estadística y los efectos de selección han dificultado nuestros estudios considerablemente. Sin embargo, se han obtenido restricciones observacionales importantes para la evolución de las propiedades físicas de las galaxias utilizando cartografiados de tipo pencil beam (de campo pequeño pero profundos), estudiando poblaciones concretas de galaxias en distintos rangos de longitudes de onda o analizando imágenes muy profundas del Telescopio Espacial Hubble. A partir de un amplio espectro de observaciones y gracias al desarrollo de la síntesis de poblaciones estelares, hemos sido capaces de establecer correlaciones estadístcamente significativas entre propiedades de galaxias a diferentes corrimientos al rojo, además de la dependencia de éstas con el entorno. Sin embargo, los procesos físicos que rigen estos cambios fundamentales en las galaxias son diversos y complejos. El reto al que nos enfrentamos está en identificar estos procesos fundamentales, en conectar poblaciones de galaxias a diferentes corrimientos al rojo y en trazar la evolución promedio de las propiedaes de las galaxias.

Para avanzar en el campo de la formación y evolución de galaxias se requieren restricciones observacionlaes precisas tanto a bajo como a alto corrimiento al rojo. Esta tesis es un compendio de distintos trabajos cuyo objetivo común no es otro que aliviar esta necesidad. En este sentido, los resultados punteros que se han presentado pueden tener fuertes implicaciones en los models teóricos actuales.

La función de luminosidad es una de las propiedades estadísticas más fundamentales de la población de galaxias, ya que contiene importante información sobre la formación y evolución de galaxias. In el Capítulo 2, se aprovecha el enorme aumento en la estadística de galaxias proporcionado por el SDSS DR6 en comparación con estimaciones anteriores de la función de luminosidad (como las de

Blanton et al. 2003b con el SDSS o Norberg et al. 2002 con el 2dFRS) para presentar la estimación más precisa de la misma hasta la fecha, en una muestra de $\sim 500,000$ galaxias. Se han encontrado dos diferencias fundamentales con respecto a la ampliamente aceptada función de luminosidad de Blanton et al. (2003b). Concretamente, se obtienen pendientes más pronunciadas en la parte débil en todas las bandas fotométricas así como un prominente exceso de galaxias brillantes en las bandas más azules, que hemos denominado *bright-end bump* (o bulto en la parte brillante). Este bulto parece estar compuesto fundamentalmente por galaxias con formación estelar reciente y por AGNs. Estas importantes discrepancias tienen implicaciones evidentes en las teorías actuales. En particular, las pendientes más pronunciadas en la parte débil implicarían un *feedback* o retroalimentación en la formación estelar menos eficiente en galaxias enanas. Del mismo modo, el bulto observado en la parte brillante podría indicar que la retroalimentación producida por AGNs en galaxias masivas es igualmente menos eficiente de lo esperado. Es muy interesante el hecho de que características más o menos similares se observen en la función de luminosidad obtenida con los modelos semi-analíticos de Bower et al. (2006).

El cartografiado fotométrico ALHAMBRA es uno de los proyectos que en la actualidad tratan de aliviar la falta de datos de calidad a alto corrimiento al rojo ($z \sim 1$), con una combinación de buena precisión en los photo-z, profundidad fotométrica y amplia estadística. En particular, la precisión en los photo-z de ALHAMBRA se espera que mejore la de otros cartografiados fotométricos (como COSMOS o COMBO-17), gracias a su excelente sistema fotométrico. Este sistema se compone de 20 filtros de banda estrecha que cubren completamente el rango óptico y las bandas JHK del infrarrojo cercano. En el Capítulo 4 se utiliza una versión preliminar de los datos para analizar las propiedades estadísticas principales de la población de galaxias a bajo corrimiento al rojo en ALHAMBRA, con una muestra de $\sim 10,000$ galaxias con magnitud sintética del SDSS $i^A < 24.5$. Aunque se ha utilizado un catálogo preliminar, se ha comprobado que estas propiedades en ALHAMBRA son consistentes con resultados anteriores obtenidos con otros cartografiados. Únicamente en el rango de galaxias brillantes se han encontrado discrepancias significativas, que parecen ser causadas por errores catastróficos en los photo-z. Un resultado importante de este análisis es la función de luminosidad en ALHAMBRA, que ha sido estimada hasta una magnitud sintética absoluta de $M_i^{0.1} \sim -13.5$. En la parte débil de la función de luminosidad, donde cabe esperar que lo errores en los photo-z sean poco relevantes, se encuentra una pendiente muy pronunciada, que parece estar de acuerdo con las estimaciones de Blanton et al. (2005a), donde se utiliza una muestra de galaxias locales del SDSS. Este resultado, que debe ser confirmado en el futuro, añadiría restricciones importantes a los modelos de formación y evolución de galaxias enanas.

La calibración a bajo corrimiento al rojo del cartografiado ALHAMBRA que se ha presentado en el Capítulo 4 sienta las bases para futuros estudios de la evolución de las galaxias con ALHAMBRA. La alta precisión en la estimación de los photo-z que se espera en ALHAMBRA nos permitirá, no sólo explorar la evolución en la función de luminosidad, sino también obtener propiedades fundamentales de galaxias individuales. En el Capítulo 5, se propone y se describe una metodología para estimar masas estelares fotométricas en ALHAMBRA, utilizando el código de síntesis de poblaciones estelares FSPS (Conroy et al., 2009). Este método ha sido aplicado a la misma muestra de $10,000$ galaxias a bajo corrimiento al rojo utilizada en el Capítulo 4. Las distribuciones de las relaciones masa-luminosidad y

de la propia masa estelar que se han obtenido son consistentes con resultados ampliamente aceptados del SDSS. El método de síntesis de poblaciones estelares desarrollado para ALHAMBRA puede ser utilizado para obtener otras propiedades fundamentales de la población estelar, como la edad, la metalicidad o la historia de formación estelar.

Una buena manera de estudiar la evolución de las galaxias consiste en medir correlaciones entre sus propiedades físicas fundamentales de las galaxias a distintos corrimientos al rojo. En el Capítulo 6, se investiga la dependencia de diferentes tipos de actividad nuclear en galaxias rojas con la densidad del entorno (relación AGN-densidad) tanto a $z \sim 1$, con el DEEP2 GRS como a $z \sim 0$ con el SDSS. Como resultado, se proporcionan evidencias que indican que los AGNs de tipo LINER en el DEEP2 tienden a favorecer entornos más densos con respecto a la población roja de la que fueron extraidos. Por contra, los AGNs de tipo Seyfert o de rayos-X a $z \sim 1$ muestran poca (o nula) dependencia con la densidad del entorno, dentro de la misma muestra. A bajo corrimiento al rojo, tanto LINERs como Seyferts en el SDSS parecen preferir ligeramente entornos menos densos. Por otro lado, a $z \sim 1$ los Seyferts y LINERs rojos son aproximadamente igual de abundantes. A $z \sim 0$, sin embargo, se observa que la población de Seyferts rojos se ha visto reducida en un factor ~ 7 con respecto a la de LINERs. Este trabajo fue el primer estudio sistemático sobre la evolución de la relación AGN-densidad.

Finalmente, el Capítulo 7 de esta tesis se ha dedicado a la optimización de futuros cartografiados que utilicen espectrógrafos de nueva generación. Los cartografiados fotométricos han demostrado ser capaces de proporcionar importantes claves para entender el Universo a alto corrimiento al rojo. Sin embargo, para dar respuesta a las preguntas más acuciantes dentro del campo de la formación y evolución de galaxias es necesaria la ayuda de la espectroscopía. La prospección de futuros cartografiados espectroscópicos a gran escala de galaxias que puedan proveer de restricciones observacionales más precisas a los modelos teóricos se presenta en la actualidad como una cuestión de gran importancia. En el Capítulo 7 de esta tesis se trata la optimización del número de galaxias observadas en cartografiados que utilicen espectrógrafos de gran campo, alimentados con fibras y de nueva generación, como SIDE y BigBOSS. En concreto, se presenta un algoritmo optimizado para la asignación de fibras a galaxias previamente seleccionadas, el *draining algorithm* o algoritmo de drenado, que asegura que el máximo número de estas galaxias es observado en las primeras exposiciones del espectrógrafo. Además, se han discutido optimizaciones adicionales como la rotación del plano focal, que proporciona un aumento significativo de la eficiencia del proceso de posicionado de fibras. Estas optimizaciones sin duda contribuirán al aumento de la estadística de galaxias en futuros cartografiados.

Future work

During my PhD, I had the opportunity to investigate several important aspects of the very wide processes of galaxy formation and evolution, from different but complementary angles. My goal for the next years is to make use of this training to look into the most important questions that remains unanswered in this field. My approach is simply to confront theory and observation, taking advantage of the enormous amount of multi-frequency data that is available now (SDSS, ALHAMBRA, BOSS) and will become available in the near future (BigBOSS, Pan-Starrs, LSST, etcetera). Regarding the theoretical side, I am especially interested in the potential of semi-analytic models to shed light into the galaxy formation and evolution picture.

Galaxy evolution with the ALHAMBRA-Survey

Some of my scientific plans for the future involve continuing my research with the ALHAMBRA-Survey. The next step, in this sense, is to make use of the latest version of the photo-z catalogs, where most of the caveats that we have found are expectedly removed. With these catalogs, we plan to go through the main statistical properties of the ALHAMBRA galaxy population, focusing on those aspects where discrepancies with other surveys have been identified. The bright end and the faint end of the luminosity function deserve, in this sense, further investigation. In addition, a good characterization of both photometric and photo-z errors along with the completeness and selection function of the survey is certainly needed.

As for the determination of ALHAMBRA photometric stellar masses, the idea is to improve the SPS fitting by extending and improving the model grid, in order to solve the discrepancies found at the blue end of the ALHAMBRA spectrum. In addition, well-estimated photometric and photo-z errors will allow us to derive an error budget in the estimation of stellar masses, taking into account uncertainties in the stellar population synthesis. We will then be in a better position to calculate the stellar mass function and elucidate what causes the high-mass excess that we found, as compared to previous results from other surveys.

In parallel, we plan to extend our analysis to higher redshift bins (up to $z \sim 1$). As we have learnt from surveys like COSMOS or COMBO-17, statistics like the luminosity function and the stellar mass function can be estimated with enough accuracy from photo-z surveys. The depth of ALHAMBRA will be crucial to explore the faint-end of the luminosity function and the high-mass end of the stellar mass function at high redshift. Regarding the SMF, we will be able to investigate the high-mass upturn found by several groups at $z \lesssim 0.5$.

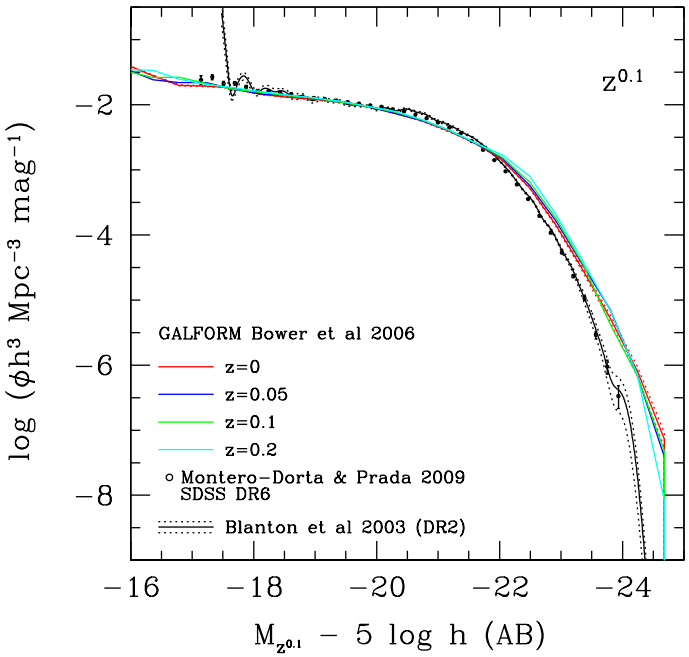

Figure 7.9 The $^{0.1}z$-band luminosity function of galaxies obtained with the Bower et al. (2006) semi-analytic models in four redshift slices at $z = 0$ (red), $z = 0.05$ (blue), $z = 0.10$ (green) and $z = 0.20$ (cyan). For comparison, the SDSS observational constraints presented in Chapter 2 (Montero-Dorta & Prada, 2009) and those from Blanton et al. (2003b) are shown in dots and a solid line, respectively. The 1-σ error region for the latter is represented by a dotted line.

The potential photo-z accuracy of ALHAMBRA, with its 20 narrow-band filters covering the optical range plus the 3 bands in the NIR, combined with our calibrated SPS machinery will allow us to obtain other fundamental galaxy properties (e.g, ages, metallicities, SFH) at different redshift bins. We will then be able to provide quality and statistically significant observational constraints on correlations between galaxy properties at high redshift.

Finally, I will continue the collaboration with Risa Wechsler from Stanford University that I started with a 3-month research visit in 2009 to produce an initial set of ALHAMBRA mock galaxy catalogs, based on the abundance halo matching technique. These catalogs will be of great help in the future, especially for the estimation of uncertainties in statistics like the luminosity function or the stellar mass function. The idea is to improve these mock catalogs by comparison with real data from the latest version of the ALHAMBRA catalogs.

Semi-analytic models of galaxy formation and evolution

Another one of my main objectives for the near future is to continue with the collaboration that I started with Carlton Baugh and his team at Durham University, aimed to compare observational constraints extracted from the SDSS with the predictions from semi-analytical models of galaxy formation and evolution. We have obtained first results on the luminosity function of galaxies, which are shown in Chapter 3. The idea is to explore the discrepancies between the models and the data (especially at the bright end) in order to shed light into the observational effects and the physical processes that determine the shape of the luminosity function. Regarding the latter, some of processes to investigate are the effect of supernova explosions in the inter-stellar medium or feedback from an accreting supermassive black hole, which are usually invoked to explain the observed sharp cut-off of the luminosity function at the bright end.

The analysis will not only be restricted to the luminosity function of galaxies. Our plan is to explore other fundamental statistical properties of the galaxy population, such as the stellar mass function or the correlation function and clustering properties.

Survey optimization for next-generation spectroscopic facilities

As mentioned before, in order to provide theorists with more accurate observational constraints we need larger and deeper galaxy samples but we also need to optimize survey techniques and strategies to increase the efficiency of our experiments. The prospection and design of future galaxy evolution surveys have represented an important part of my PhD work, within the framework of the SIDE instrument at the Instituto de Astrofísica de Andalucía. In the next years, I would like to continue with this line of research. At this moment, I am involved in BigBOSS, which is a proposed ground-based dark energy experiment to study baryon acoustic oscillations and the growth of structure with an all-sky galaxy redshift survey. BigBOSS would also provide unprecedented statistics of spectra in the critical range for massive galaxy assembly, where galaxy formation and evolution models are still severely unconstrained. The group lead at the IAA-CSIC by my thesis supervisor, Francisco Prada, is actively involved in this project, not only in the design of the focal plane of the future instrument but also in the optimization of the survey strategy. In particular, the plan for the future is to implement

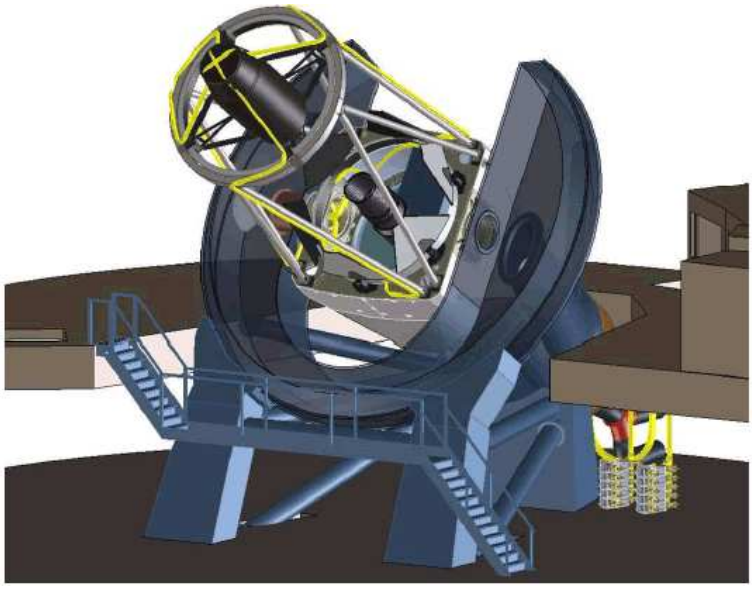

Figure 7.10 A cartoon illustrating the BigBOSS instrument, installed at the Mayall 4-m telescope. The yellow trace is a ber routing path from the focal plane to the spectrograph room incorporating ber spooling locations to accommodate the inclination and declinaton motions of the telescope. The two stack-of-five spectrograph arrays are adjacent to the telescope base at the end of the ber runs. BigBOSS is the most ambitious dark energy experiment ever conceived, intended to obtain redshifts of ~ 40 million objects. More information at http://bigboss.lbl.gov/.

the survey optimizations presented in Chapter 7 to BigBOSS and explore further optimizations.

Dark energy experiments

The formation and evolution of galaxies is tightly connected to the formation and evolution of dark matter haloes and the growth of structure. I am therefore very interested in Cosmology and, in particular, in dark energy experiments such as BOSS and BigBOSS. These surveys are intended to explore the nature of dark matter and dark energy, and, perhaps, even the signatures of early-Universe inflation. BOSS and BigBOSS will provide important constraints on the expansion of the Universe and the growth of structure. They are expected to shed light into the halo-galaxy connection and the formation of galaxies as well. At the moment, I am involved in a project using BOSS data, in collaboration with my thesis supervisor, Francisco Prada (IAA-CSIC), Ginevra Favole (IAA-CSIC), Rita Tojeiro (ICG, Portsmouth) and other BOSS members. The idea is to combine observational data from BOSS with analytical formalisms and cosmological numerical simulations to constrain cosmological parameters. In addition, we will be able to measure the signature of primordial non-Gaussianity in initial conditions, by analyzing the clustering of galaxies and the abundance of rare objects - such as voids - in the Universe. The project requires the characterization of the BOSS galaxy population, which is basically composed of luminous red galaxies. The first step is to understand the selection function of the survey and to estimate statistical properties such as the luminosity and the stellar mass function and the correlation function.

Related publications

- **The DEEP2 Galaxy Redshift Survey: the red sequence AGN fraction and its environment and redshift dependence**; Montero-Dorta, Antonio D.; Croton, Darren J.; Yan, Renbin; Cooper, Michael C.; Newman, Jeffery A.; Georgakakis, Antonis; Prada, Francisco; Davis, Marc; Nandra, Kirpal; Coil, Alison; 2009, MNRAS, 392, 125.

  We measure the dependence of the AGN fraction on local environment at $z \sim 1$, using spectroscopic data taken from the DEEP2 Galaxy Redshift Survey, and Chandra X-ray data from the All-Wavelength Extended Groth Strip International Survey (AEGIS). To provide a clean sample of AGN we restrict our analysis to the red sequence population; this also reduces additional colour–environment correlations. We find evidence that high redshift LINERs in DEEP2 tend to favour higher density environments relative to the red population from which they are drawn. In contrast, Seyferts and X-ray selected AGN at $z \sim 1$ show little (or no) environmental dependencies within the same underlying population. We compare these results with a sample of local AGN drawn from the SDSS. Contrary to the high redshift behaviour, we find that both LINERs and Seyferts in the SDSS show a slowly declining red sequence AGN fraction towards high density environments. Interestingly, at $z \sim 1$ red sequence Seyferts and LINERs are approximately equally abundant. By $z \sim 0$, however, the red Seyfert population has declined relative to the LINER population by over a factor of 7. We speculate on possible interpretations of our results.

- **The SDSS DR6 luminosity functions of galaxies**, Montero-Dorta, Antonio D.; Prada, Francisco; 2009, MNRAS, 399, 1106.

  We present number counts, luminosity functions (LFs) and luminosity densities of galaxies obtained using the Sloan Digital Sky Survey Sixth Data Release in all SDSS photometric bands. Thanks to the SDSS DR6, galaxy statistics have increased by a factor of ~ 7 in the u-band and by a factor between ~ 3 and ~ 5 in the rest of the SDSS bands with respect to the previous work of Blanton et al. (2003b). In addition, we have ensured a high redshift completeness in our galaxy samples, mainly by minimizing the effect of brightness-dependent incompleteness. With these advances, we have estimated very accurate SDSS DR6 LFs in both the bright and the faint end. In the $^{0.1}r$-band, our LF is well fitted by a Schechter LF with parameters $\Phi_* = 0.93 \pm 0.07$,

$M_* - 5log_{10}h = -20.71 \pm 0.04$ and $\alpha = -1.26 \pm 0.02$. As compared with previous results, we find some notable differences. In the bright end of the $^{0.1}u$-band LF we find a remarkable excess, of ~ 1.7 dex at $M_{^{0.1}u} \simeq -20.5$, with respect to the best-fit Schechter LF. This excess weakens in the $^{0.1}g$-band, fading away towards the very red $^{0.1}z$-band. A preliminary analysis on the nature of this bright-end bump reveals that it is comprised of QSOs/Seyferts I's ($\sim 60\%$), Starbursts and star-forming galaxies ($\sim 20\%$) and normal galaxies ($\sim 20\%$). It seems, therefore, that an important fraction of this excess luminosity may come from nuclear activity. In the faint end of the SDSS DR6 luminosity functions, where we can reach $0.5-1$ magnitude deeper than the previous SDSS LF estimations, we obtain a steeper slope, that increases from the $^{0.1}u$-band, with $\alpha = -1.05 \pm 0.05$, to the very red $^{0.1}z$-band, with $\alpha = -1.26 \pm 0.03$. We have also investigated the effect of galaxy evolution on our LFs. These state-of-the-art results may be used to constrain a variety of aspects of star formation histories and/or feed-back processes in galaxy formation models.

- **Fiber Assignment in Next-generation Wide-field Spectrographs**, Morales, Isaac; Montero-Dorta, Antonio D.; Azzaro, Marco; Prada, Francisco; Sanchez, Justo; Becerril, Santiago; 2011, MNRAS, in press.

We present an optimized algorithm for assigning fibers to targets in next-generation fiber-fed multi-object spectrographs. The method, that we named draining algorithm, ensures that the maximum number of targets in a given target field is observed in the first few tiles. Using randomly distributed targets and mock galaxy catalogs we have estimated that the gain provided by the draining algorithm as compared to a random assignment can be as much as 2% for the first tiles. This would imply for a survey like BigBOSS saving for observation several hundred thousand objects or, alternatively, reducing the covered area in ~ 350 deg^2. An important advantage of this method is that the fiber collision problem can be solved easily and in an optimal way. We also discuss additional optimizations of the fiber positioning process. In particular, we show that allowing for rotation of the focal plane can improve the efficiency of the process in $\sim 3.5-4.5\%$ even if only small adjustments are permitted (up to 2^o). For instruments that allow large rotations of the focal plane the expected gain increases to $\sim 5-6\%$. These results, therefore, strongly support focal plane rotation in future spectrographs, as far as the efficiency of the fiber positioning process is concerned. Finally, we discuss on the implications of our optimizations and provide some basic hints for an optimal survey strategy based on the number of targets per positioner.

Appendix

K-correction estimation with FSPS

In this Appendix, we describe in detail the method for estimating K-corrections in the ALHAMBRA survey using the Flexible Stellar Population Synthesis code (FSPS, Conroy et al. 2009). The nomenclature used throughout this Appendix is taken from Hogg et al. (2002). As expressed before, the K-correction accounts for the fact that, due to the expansion of the Universe, we measure the flux (or magnitude) of a source through a passband that is different from that at which it was emitted (Humason, 1956; Oke & Sandage, 1968). We can then start by defining two different passbands: an observed-frame passband and an emitted-frame or rest-frame passband. In essence, we want to convert our observed-frame magnitudes to emitted-frame magnitudes. The apparent magnitude that we measure from a source at an observed-frame passband R, $m_{\rm R}$, can be expressed as a function of the intrinsic rest-frame absolute magnitude at an emitted passband Q, $M_{\rm Q}$, as follows:

$$m_{\rm R} = M_{\rm Q} + DM(z) + K_{\rm QR} \tag{A.1}$$

where DM is the distance modulus, z is the redshift and $K_{\rm QR}$ is the K-correction between passbands R and Q . The distance modulus, $DM(z)$, accounts for the dimming in luminosity between z and a reference distance set at 10 pc and is defined by

$$DM = 5\log_{10}\left[\frac{D_{\rm L}}{10pc}\right] \tag{A.2}$$

where $D_{\rm L}$ is the luminous distance, which depends on the cosmological parameters. We can also define an observed-frame wavelength (λ_o) and an emitted-frame wavelength (λ_e), which are related to each other by the following equation:

$$\lambda_{\rm e} = \frac{\lambda_{\rm o}}{1+z} \tag{A.3}$$

The apparent magnitude of a source is related to its flux density per unit wavelength, $f_\lambda(\lambda_o)$ ($erg\ s^{-1}\ cm^{-2}\ \mathring{A}^{-1}$), in the following way:

$$m_{\rm R} = -2.5\log_{10}\left[\frac{\int d\lambda_o \lambda_o f_\lambda(\lambda_o) R(\lambda_o)}{\int d\lambda_o \lambda_o g^R_\lambda(\lambda_o) R(\lambda_o)}\right] \tag{A.4}$$

where $g^R_\lambda(\lambda_o)$ is the flux density per unit wavelength for the standard source and $R(\lambda_o)$ is the instrument response, i.e. the probability that a photon of wavelength λ_o is counted by the CCD detector. The ALHAMBRA magnitudes are in the AB system, which is defined so that a source with a flat spectrum with $g^{AB}_\nu(\nu) = 3631\ Jy$ ($1\ Jy = 10^{-23}\ erg\ s^{-1}\ cm^{-2}\ Hz^{-1}$) has magnitude $m = 0$. It is trivial to derive $g^{AB}_\lambda(\lambda)$ from $g^{AB}_\nu(\nu)$ using the relation $\lambda\nu = c$, where c is the light speed. Note that Equation A.5 uses observed-frame wavelengths, λ_o.

The absolute magnitude is defined as the apparent magnitude that a source would have if it were placed at a distance of 10 pc. Similarly to Equation A.5, the absolute magnitude of an object at an emitted-frame passband Q can be expressed as follows:

$$M_{\rm Q} = -2.5\log_{10}\left[\frac{\int d\lambda_e \lambda_e \frac{L_\lambda(\lambda_e)}{4\pi(10pc)^2} Q(\lambda_e)}{\int d\lambda_e \lambda_e g^R_\lambda(\lambda_e) Q(\lambda_e)}\right] \tag{A.5}$$

where now the integrals are over emitted-frame wavelengths. Note that $Q(\lambda_e)$ is the instrument response for passband Q. The flux density per unit wavelength at a given wavelength λ_e is now simply the spectral luminosity of the source, $L_\lambda(\lambda_e)$, divided by the area of a sphere of radius 10 pc. In a similar way the luminosity per unit wavelength of the source in emitted-frame wavelengths , $L_\lambda(\lambda_e)$, can be related to its observed flux:

$$L_\lambda(\lambda_e) = 4\pi D^2_{\rm L}(1+z) f_\lambda(\lambda_o) \tag{A.6}$$

where factor $(1+z)$ accounts for the redshift of wavelengths. Finally, inserting Equation A.6 into Equation A.5, making use of Equation A.3 and rearranging Equation A.1, we can obtain a theoretical expression for the K-corretion. Namely:

$$K_{\rm QR} = -2.5\log_{10}\left[\frac{1}{(1+z)}\frac{\int d\lambda_o \lambda_o f_\lambda(\lambda_o) R(\lambda_o) \int d\lambda_e \lambda_e g^Q_\lambda(\lambda_e) R(\lambda_e)}{\int d\lambda_o \lambda_o g^R_\lambda(\lambda_o) R(\lambda_o) \int d\lambda_e \lambda_e f_\lambda((1+z)\lambda_e) Q(\lambda_e)}\right] \tag{A.7}$$

In most cases, as in this work, the observed-frame passband and the emitted-frame passband are the same, but simply redshifted. This trivially implies that $g^R_\lambda(\lambda) = g^Q_\lambda(\lambda)$ so Equation A.7 can be simplified considerably:

$$K_{\rm QR} = -2.5\log_{10}\left[\frac{1}{(1+z)}\frac{\int d\lambda_o \lambda_o f_\lambda(\lambda_o) R(\lambda_o)}{\int d\lambda_e \lambda_e f_\lambda((1+z)\lambda_e) R(\lambda_e)}\right] \tag{A.8}$$

In practice, it is rather convenient to work with emitted-frame luminosities $L_\lambda(\lambda_e)$ instead of observed-frame fluxes, f_λ. Equation A.8 can be easily transformed using Equation A.6:

$$K_{\rm QR} = -2.5\log_{10}\left[\frac{1}{(1+z)}\frac{\int d\lambda_o \lambda_o L_\lambda(\frac{\lambda_o}{1+z}) R(\lambda_o)}{\int d\lambda_e \lambda_e L_\lambda(\lambda_e) R(\lambda_e)}\right] \tag{A.9}$$

This is the equation that we use for estimating K-corrections in ALHAMBRA. In practice, we just need two ingredients to calculate K-corrections: a model that fits the spectrum of a given galaxy reasonably well and the instrument response for each passband, $R(\lambda)$. $R(\lambda)$ must take into account

not only the filter transmission, but also the atmosphere extinction, the telescope (mirror) reflectivity and the CCD efficiency. The contribution from these effects is however small, especially because any shift in the filter transmission cancels off in Equation A.9.

Solving Equation A.9 for an ALHAMBRA galaxy using FSPS is extremely simple. We first fit the ALHAMBRA magnitudes, as discussed in Chapter 6, using our grid of galaxy models. For the corresponding best-fit model, we need to extract the entire spectrum from FSPS. Note that FSPS can output both magnitudes in a number of filters and spectra. From the spectrum output by FSPS, the luminosity, $L_\lambda(\lambda)$, can be easily obtained:

$$L_\lambda(\lambda) = 10^{-0.4n} S_\nu^{\mathrm{FSPS}} \left(\frac{c}{\lambda^2}\right) \; L_\odot \quad (erg\ s^{-1}\ \mathring{A}^{-1}) \tag{A.10}$$

where $L_\odot$ is the solar luminosity in erg/s ($L_\odot = 3.839\ 10^{33}\ erg/s$) and S_ν^{FSPS} is the best-fit spectrum from FSPS, in units of $L_\odot/Hz$. Finally, n is the normalization of the model (the FSPS models are normalized to 1 $M_\odot$). Note that the luminosity that we obtain from the fit corresponds to the observed-frame luminosity of the object, $L_\lambda(\lambda_o)$. In order to obtain $L_\lambda(\lambda_e)$, we can just blueshift the luminosity that we derive from the fitting to $z = 0$. Instead of doing this, it is much easier to derive an emitted-frame luminosity by taking the spectrum of the best-fit model at $z = 0$ from the grid. This has an important advantage. By selecting the best-fit model at $z = 0$ we are also accounting for the change in luminosity due to evolution. Note that, in the FSPS grid presented in Chapter 6, a model at $z = 0$ is not only blueshifted as compared to the "same" model at redshift z, but the shape of its SED has changed, to account for the aging of the stellar population. In this sense, we can estimate a K+E correction, which will be very useful when we extend our analysis to higher redshift slices.

Finally, it is usually convenient to blueshift the emitted frame (rest frame) from $z = 0$ to a given redshift z_0, so that a smaller K-correction is required. This is a common procedure in a number of SDSS works, where K-corrections are blueshifted to $z_0 = 0.1$, where the SDSS redshift distribution peaks.

Bibliography

Adelman-McCarthy J. K., Agüeros M. A., Allam S. S., et al., 2006, ApJS, 162, 38

Adelman-McCarthy J. K., Agüeros M. A., Allam S. S., et al., 2008, ApJS, 175, 297

Aihara H., Allende Prieto C., An D., et al., 2011, ApJS, 193, 29

Aparicio Villegas T., Alfaro E. J., Cabrera-Caño J., et al., 2010, AJ, 139, 1242

Azzaro M., Becerril S., Vilar C., et al., 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 7735 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference

Baldry I. K., Balogh M. L., Bower R. G., et al., 2006, MNRAS, 373, 469

Baldry I. K., Glazebrook K., Driver S. P., 2008, MNRAS, 388, 945

Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5

Balogh M. L., Baldry I. K., Nichol R., Miller C., Bower R., Glazebrook K., 2004, ApJ, 615, L101

Baraffe I., Chabrier G., Allard F., Hauschildt P. H., 1998, A&A, 337, 403

Bassett B. A., Nichol B., Eisenstein D. J., 2005, Astronomy and Geophysics, 46, 5, 050000

Bell E., Davis M., Dey A., et al., 2009, in astro2010: The Astronomy and Astrophysics Decadal Survey, vol. 2010 of ArXiv Astrophysics e-prints, arXiv:0903.3404

Bell E. F., McIntosh D. H., Katz N., Weinberg M. D., 2003, ApJS, 149, 289

Benítez N., 2000, ApJ, 536, 571

Benítez N., Moles M., Aguerri J. A. L., et al., 2009, ApJ, 692, L5

Bennett C. L., Banday A. J., Gorski K. M., et al., 1996, ApJ, 464, L1+

Benson A. J., Bower R. G., Frenk C. S., Lacey C. G., Baugh C. M., Cole S., 2003, ApJ, 599, 38

Bertelli G., Bressan A., Chiosi C., Fagotto F., Nasi E., 1994, A&AS, 106, 275

Best P. N., 2004, MNRAS, 351, 70

Betancort-Rijo J. E., Montero-Dorta A. D., 2006, ApJ, 650, L95

Binney, J. and Tremaine, S., 1987, Galactic dynamics

Blanton M. R., 2006, ApJ, 648, 268

Blanton M. R., Dalcanton J., Eisenstein D., et al., 2001, AJ, 121, 2358

Blanton M. R., Eisenstein D., Hogg D. W., Schlegel D. J., Brinkmann J., 2005a, ApJ, 629, 143

Blanton M. R., Hogg D. W., Bahcall N. A., et al., 2003a, ApJ, 594, 186

Blanton M. R., Hogg D. W., Bahcall N. A., et al., 2003b, ApJ, 592, 819

Blanton M. R., Lin H., Lupton R. H., et al., 2003c, AJ, 125, 2276

Blanton M. R., Lupton R. H., Schlegel D. J., et al., 2005b, ApJ, 631, 208

Blanton M. R., Roweis S., 2007, AJ, 133, 734

Blanton M. R., Schlegel D. J., Strauss M. A., et al., 2005c, AJ, 129, 2562

Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517

Bondi H., Gold T., 1948, MNRAS, 108, 252

Borch A., Meisenheimer K., Bell E. F., et al., 2006, A&A, 453, 869

Bower R. G., Benson A. J., Malbon R., et al., 2006, MNRAS, 370, 645

Broadhurst T., Umetsu K., Medezinski E., Oguri M., Rephaeli Y., 2008, ApJ, 685, L9

Bruzual A. G., 1983, ApJ, 273, 105

Bundy K., Ellis R. S., Conselice C. J., et al., 2006, ApJ, 651, 120

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Campos A., 1996, in Dark Matter in the Universe, edited by S. Bonometto, J. R. Primack, & A. Provenzale, 123–+

Capak P., Aussel H., Ajiki M., et al., 2007, ApJS, 172, 99

Capak P., Cowie L. L., Hu E. M., et al., 2004, AJ, 127, 180

Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245

Cassisi S., Castellani V., Ciarcelluti P., Piotto G., Zoccali M., 2000, MNRAS, 315, 679

Chabrier G., 2003, PASP, 115, 763

Chandrasekhar S., 1943, ApJ, 97, 255

Charlot S., Fall S. M., 2000, ApJ, 539, 718

Chilingarian I. V., Melchior A., Zolotukhin I. Y., 2010, MNRAS, 405, 1409

Cioni M., Girardi L., Marigo P., Habing H. J., 2006, A&A, 448, 77

Coil A. L., Newman J. A., Croton D., et al., 2007, ArXiv e-prints, 708

Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168

Cole S., Norberg P., Baugh C. M., et al., 2001, MNRAS, 326, 255

Colless M., Dalton G., Maddox S., et al., 2001, MNRAS, 328, 1039

Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486

Constantin A., Hoyle F., Vogeley M. S., 2007, ArXiv e-prints, 710

Constantin A., Vogeley M. S., 2006, ArXiv Astrophysics e-prints

Cooper M. C., Newman J. A., Coil A. L., et al., 2007a, MNRAS, 376, 1445

Cooper M. C., Newman J. A., Croton D. J., et al., 2006, MNRAS, 370, 198

Cooper M. C., Newman J. A., Madgwick D. S., Gerke B. F., Yan R., Davis M., 2005, ApJ, 634, 833

Cooper M. C., Newman J. A., Weiner B. J., et al., 2007b, ArXiv e-prints, 706

Corwin Jr. H. G., Buta R. J., de Vaucouleurs G., 1994, AJ, 108, 2128

Cristóbal-Hornillos D., Aguerri J. A. L., Moles M., et al., 2009, ApJ, 696, 1554

Croton D. J., Springel V., White S. D. M., et al., 2006, MNRAS, 365, 11

Cucciati O., Iovino A., Marinoni C., et al., 2006, A&A, 458, 39

da Costa L. N., Geller M. J., Pellegrini P. S., et al., 1994, ApJ, 424, L1

da Costa L. N., Pellegrini P. S., Sargent W. L. W., et al., 1988, ApJ, 327, 544

Davis M., DEEP Team, Extended Groth Strip Collaboration, 2005, in Bulletin of the American Astronomical Society, vol. 37 of Bulletin of the American Astronomical Society, 1299–+

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Davis M., Faber S. M., Newman J., et al., 2003, in Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II. Edited by Guhathakurta, Puragra. Proceedings of the SPIE, Volume 4834, pp. 161-172 (2003)., edited by P. Guhathakurta, vol. 4834 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 161–172

Davis M., Guhathakurta P., Konidaris N. P., et al., 2007, ApJ, 660, L1

Davis M., Huchra J., 1982, ApJ, 254, 437

De Lucia G., Blaizot J., 2007, MNRAS, 375, 2

Domínguez A., Primack J. R., Rosario D. J., et al., 2011, MNRAS, 410, 2556

Dotter A., Chaboyer B., Ferguson J. W., et al., 2007, ApJ, 666, 403

Dressler A., 1980, ApJ, 236, 351

Dressler A., Gunn J. E., 1983, ApJ, 270, 7

Driver S. P., Popescu C. C., Tuffs R. J., et al., 2007, MNRAS, 379, 1022

Driver S. P., Popescu C. C., Tuffs R. J., Graham A. W., Liske J., Baldry I., 2008, ApJ, 678, L101

Drory N., Bundy K., Leauthaud A., et al., 2009, ApJ, 707, 1595

Drory N., Feulner G., Bender R., et al., 2001, MNRAS, 325, 550

Drory N., Salvato M., Gabasch A., et al., 2005, ApJ, 619, L131

Efstathiou G., Ellis R. S., Peterson B. A., 1988, MNRAS, 232, 431

Efstathiou G., Silk J., 1983, Fund. Cosmic Phys., 9, 1

Elbaz D., Daddi E., Le Borgne D., et al., 2007, A&A, 468, 33

Faber S. M., Phillips A. C., Kibrick R. I., et al., 2003, in Instrument Design and Performance for Optical/Infrared Ground-based Telescopes. Edited by Iye, Masanori; Moorwood, Alan F. M. Proceedings of the SPIE, Volume 4841, pp. 1657-1669 (2003)., edited by M. Iye, A. F. M. Moorwood, vol. 4841 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 1657–1669

Faber S. M., Willmer C. N. A., Wolf C., et al., 2007, ApJ, 665, 265

Feulner G., Goranova Y., Hopp U., et al., 2007, MNRAS, 378, 429

Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A., Wright E. L., 1996, ApJ, 473, 576

Fontana A., Salimbeni S., Grazian A., et al., 2006, A&A, 459, 745

Gamow G., 1946, Physical Review, 70, 572

Gardner J. P., 1995, ApJS, 98, 441

Gentile G., Tonini C., Salucci P., 2007, A&A, 467, 925

Georgakakis A., Nandra K., Laird E. S., et al., 2007, ApJ, 660, L15

Gerke B. F., Newman J. A., Davis M., et al., 2005, ApJ, 625, 6

Gerke B. F., Newman J. A., Faber S. M., et al., 2007, MNRAS, 376, 1425

Giovanelli R., Haynes M. P., 1991, ARA&A, 29, 499

Girardi L., Bressan A., Bertelli G., Chiosi C., 2000, A&AS, 141, 371

Graves G. J., Faber S. M., Schiavon R. P., Yan R., 2007, ArXiv e-prints, 707

Groves B., Kewley L., Kauffmann G., Heckman T., 2006, New Astronomy Review, 50, 743

Gunn J. E., Gott III J. R., 1972, ApJ, 176, 1

Guth A. H., 1981, Phys. Rev. D, 23, 347

Harrison E. R., 1970, Phys. Rev. D, 1, 2726

Henriques B. M. B., Thomas P. A., 2010, MNRAS, 403, 768

Hogg D. W., Baldry I. K., Blanton M. R., Eisenstein D. J., 2002, ArXiv Astrophysics e-prints

Hogg D. W., Blanton M. R., Brinchmann J., et al., 2004, ApJ, 601, L29

Hogg D. W., Blanton M. R., Eisenstein D. J., et al., 2003, ApJ, 585, L5

Hogg D. W., Masjedi M., Berlind A. A., Blanton M. R., Quintero A. D., Brinkmann J., 2006, ApJ, 650, 763

Hopkins P. F., Somerville R. S., Hernquist L., Cox T. J., Robertson B., Li Y., 2006, ApJ, 652, 864

Hoyle F., 1948, MNRAS, 108, 372

Huang J., Cowie L. L., Gardner J. P., Hu E. M., Songaila A., Wainscoat R. J., 1997, ApJ, 476, 12

Hubble E., Humason M. L., 1931, ApJ, 74, 43

Hubble E. P., 1922, ApJ, 56, 400

Hubble E. P., 1925, ApJ, 62, 409

Hubble E. P., 1926, ApJ, 64, 321

Huchra J., Davis M., Latham D., Tonry J., 1983, ApJS, 52, 89

Humason M. L., 1956, Vistas in Astronomy, 2, 1620

Jeans J. H., 1902, Royal Society of London Philosophical Transactions Series A, 199, 1

Jenkins L. P., Hornschemeier A. E., Mobasher B., Alexander D. M., Bauer F. E., 2007, ApJ, 666, 846

Jones D. H., Peterson B. A., Colless M., Saunders W., 2006, MNRAS, 369, 25

Kauffmann G., Heckman T. M., Tremonti C., et al., 2003a, MNRAS, 346, 1055

Kauffmann G., Heckman T. M., White S. D. M., et al., 2003b, MNRAS, 341, 33

Kauffmann G., White S. D. M., Heckman T. M., et al., 2004, MNRAS, 353, 713

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006, MNRAS, 372, 961

Khochfar S., Silk J., Windhorst R. A., Ryan Jr. R. E., 2007, ApJ, 668, L115

Kimura M., Maihara T., Iwamuro F., et al., 2010, PASJ, 62, 1135

Klypin A., Kravtsov A. V., Valenzuela O., Prada F., 1999, ApJ, 522, 82

Klypin A., Trujillo-Gomez S., Primack J., 2010, ArXiv e-prints, arXiv:1002.3660

Komatsu E., Dunkley J., Nolta M. R., et al., 2009, ApJS, 180, 330

Kravtsov A., 2010, Advances in Astronomy, 2010

Kroupa P., 2001, MNRAS, 322, 231

Kroupa P., Tout C. A., Gilmore G., 1993, MNRAS, 262, 545

Lacey C., Cole S., 1993, MNRAS, 262, 627

Lançon A., Mouhcine M., 2002, A&A, 393, 167

Larson R. B., 1974, MNRAS, 169, 229

Larson R. B., 1975, MNRAS, 173, 671

Le Fevre O., Vettolani G., Maccagni D., et al., 2003, in Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II. Edited by Guhathakurta, Puragra. Proceedings of the SPIE, Volume 4834, pp. 173-182 (2003)., edited by P. Guhathakurta, vol. 4834 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference, 173–182

Leauthaud A., Massey R., Kneib J., et al., 2007, ApJS, 172, 219

Lejeune T., Cuisinier F., Buser R., 1997, A&AS, 125, 229

Lejeune T., Cuisinier F., Buser R., 1998, A&AS, 130, 65

Lemaître G., 1931, Nature, 127, 706

Lequeux J., Peimbert M., Rayo J. F., Serrano A., Torres-Peimbert S., 1979, A&A, 80, 155

Li C., White S. D. M., 2009, MNRAS, 398, 2177

Lin H., Kirshner R. P., Shectman S. A., et al., 1996, ApJ, 464, 60

Loveday J., 1997, ApJ, 489, 29

Madgwick D. S., Lahav O., Baldry I. K., et al., 2002, MNRAS, 333, 133

Marchesini D., van Dokkum P. G., Förster Schreiber N. M., Franx M., Labbé I., Wuyts S., 2009, ApJ, 701, 1765

Marigo P., Girardi L., 2007, A&A, 469, 239

Marigo P., Girardi L., Bressan A., Groenewegen M. A. T., Silva L., Granato G. L., 2008, A&A, 482, 883

Martínez V. J., Saar E., 2002, Statistics of the Galaxy Distribution, Chapman amp

Martínez V. J., Trimble V., 2009, in Cosmology Across Cultures, edited by J. Rubiño-Martín, J. Belmonte, F. Prada, & A. Alberdi , vol. 409 of Astronomical Society of the Pacific Conference Series, 47–+

Marzke R. O., Huchra J. P., Geller M. J., 1994, ApJ, 428, 43

Mayer L., Governato F., Kaufmann T., 2008, Advanced Science Letters, 1, 7

McCracken H. J., Radovich M., Bertin E., et al., 2003, A&A, 410, 17

Meisenheimer K., Wolf C., 2002, Astronomy and Geophysics, 43, 3, 030000

Meneux B., Guzzo L., de la Torre S., et al., 2009, A&A, 505, 463

Mercurio A., Merluzzi P., Haines C. P., et al., 2006, MNRAS, 368, 109

Miller C. J., Nichol R. C., Gómez P. L., Hopkins A. M., Bernardi M., 2003, ApJ, 597, 142

Miller G. E., Scalo J. M., 1979, ApJS, 41, 513

Milne E. A., 1932, Nature, 130, 9

Mo H., van den Bosch F. C., White S., 2010, Galaxy Formation and Evolution

Moles M., Benítez N., Aguerri J. A. L., et al., 2008, AJ, 136, 1325

Montero-Dorta A. D., Croton D. J., Yan R., et al., 2009, MNRAS, 392, 125

Montero-Dorta A. D., Prada F., 2009, MNRAS, 399, 1106

Morales I., Montero-Dorta A. D., Azzaro M., Prada F., Sanchez J., Becerril S., 2011, ArXiv e-prints

Nandra K., Georgakakis A., Willmer C. N. A., et al., 2007, ApJ, 660, L11

Newman J. A., Davis M., 2002, ApJ, 564, 567

Nolan L. A., Raychaudhury S., Kabán A., 2007, MNRAS, 375, 381

Norberg P., Cole S., Baugh C. M., et al., 2002, MNRAS, 336, 907

Oemler A. J., 1974, ApJ, 194, 1

Oke J. B., Sandage A., 1968, ApJ, 154, 21

Peacock J., Schneider P., 2006, The Messenger, 125, 48

Peebles P. J., Ratra B., 2003, Reviews of Modern Physics, 75, 559

Peebles P. J. E., Yu J. T., 1970, ApJ, 162, 815

Pérez-González P. G., Rieke G. H., Villar V., et al., 2008, ApJ, 675, 234

Perlmutter S., Aldering G., Goldhaber G., et al., 1999, ApJ, 517, 565

Petrosian V., 1976, ApJ, 210, L53+

Pietrinferni A., Cassisi S., Salaris M., Castelli F., 2004, ApJ, 612, 168

Popesso P., Biviano A., Böhringer H., Romaniello M., 2006, A&A, 445, 29

Popesso P., Böhringer H., Romaniello M., Voges W., 2005, A&A, 433, 415

Pozzetti L., Bolzonella M., Zucca E., et al., 2010, A&A, 523, A13+

Prada F., Azzaro M., Rabaza O., Sánchez J., Ubierna M., 2008, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 7014 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference

Prada F., Vitvitska M., Klypin A., et al., 2003, ApJ, 598, 260

Press W. H., Schechter P., 1974, ApJ, 187, 425

Renzini A., Ciotti L., 1993, ApJ, 416, L49+

Riess A. G., Filippenko A. V., Challis P., et al., 1998, AJ, 116, 1009

Robaina A. R., Bell E. F., Skelton R. E., et al., 2009, ApJ, 704, 324

Rowan-Robinson M., 1996, in Mapping, Measuring, and Modelling the Universe, edited by P. Coles, V. Martinez, & M.-J. Pons-Borderia, vol. 94 of Astronomical Society of the Pacific Conference Series, 171–+

Rubin V. C., Ford Jr. W. K., 1970, ApJ, 159, 379

Rubin V. C., Ford W. K. J., . Thonnard N., 1980, ApJ, 238, 471

Salpeter E. E., 1955, ApJ, 121, 161

Sandage A., 1978, AJ, 83, 904

Scalo J., 1998, ArXiv Astrophysics e-prints

Scalo J. M., 1986, Fund. Cosmic Phys., 11, 1

Schaller G., Schaerer D., Meynet G., Maeder A., 1992, A&AS, 96, 269

Schechter P., 1976, ApJ, 203, 297

Schlegel D., White M., Eisenstein D., 2009a, in astro2010: The Astronomy and Astrophysics Decadal Survey, vol. 2010 of ArXiv Astrophysics e-prints, 314–+

Schlegel D. J., Bebek C., Heetderks H., et al., 2009b, ArXiv e-prints, arXiv:0904.0468

Schmidt M., 1968, ApJ, 151, 393

Scoville N., Aussel H., Brusa M., et al., 2007, ApJS, 172, 1

Shapley H., Ames A., 1932, Annals of Harvard College Observatory, 88, 41

Shectman S. A., Landy S. D., Oemler A., et al., 1996, ApJ, 470, 172

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Sheth R. K., Tormen G., 2002, MNRAS, 329, 61

Silk J., 1968, ApJ, 151, 459

Somerville R. S., Primack J. R., 1999, MNRAS, 310, 1087

Sorrentino G., Radovich M., Rifatto A., 2006, A&A, 451, 809

Spinrad H., Taylor B. J., 1971, ApJS, 22, 445

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

Steidel C. C., Adelberger K. L., Giavalisco M., Dickinson M., Pettini M., 1999, ApJ, 519, 1

Stoughton C., Lupton R. H., Bernardi M., et al., 2002, AJ, 123, 485

Strateva I., Ivezić Ž., Knapp G. R., et al., 2001, AJ, 122, 1861

Strauss M. A., Weinberg D. H., Lupton R. H., et al., 2002, AJ, 124, 1810

Takeuchi T. T., Yoshikawa K., Ishii T. T., 2000, ApJS, 129, 1

Tikhonov A. V., Klypin A., 2009, MNRAS, 395, 1915

Tinsley B. M., 1978, ApJ, 222, 14

Toomre A., 1977, ARA&A, 15, 437

Tremonti C. A., Heckman T. M., Kauffmann G., et al., 2004, ApJ, 613, 898

van Dokkum P. G., 2008, ApJ, 674, 29

Wake D. A., Miller C. J., Di Matteo T., et al., 2004, ApJ, 610, L85

Wang X., Chen X., Zheng Z., Wu F., Zhang P., Zhao Y., 2009, MNRAS, 394, 1775

Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006, MNRAS, 366, 2

Westera P., Lejeune T., Buser R., Cuisinier F., Bruzual G., 2002, A&A, 381, 524

Westra E., Geller M. J., Kurtz M. J., Fabricant D. G., Dell'Antonio I., 2010, PASP, 122, 1258

White S. D. M., Frenk C. S., 1991, ApJ, 379, 52

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

Willmer C. N. A., Faber S. M., Koo D. C., et al., 2005, ArXiv Astrophysics e-prints

Willmer C. N. A., Faber S. M., Koo D. C., et al., 2006, ApJ, 647, 853

Witt A. N., 2000, J. Geophys. Res., 105, 10299

Wolf C., Meisenheimer K., Rix H., Borch A., Dye S., Kleinheinrich M., 2003, A&A, 401, 73

Yan R., Ho L. C., Newman J. A., et al., 2011, ApJ, 728, 38

Yan R., Newman J. A., Faber S. M., Konidaris N., Koo D., Davis M., 2006, ApJ, 648, 281

Yasuda N., Fukugita M., Narayanan V. K., et al., 2001, AJ, 122, 1104

Yi S., Demarque P., Kim Y., et al., 2001, ApJS, 136, 417

York D. G., Adelman J., Anderson Jr. J. E., et al., 2000, AJ, 120, 1579

Zel'Dovich Y. B., 1970, A&A, 5, 84

Zirbel E. L., 1997, ApJ, 476, 489

Zwicky F., 1937, ApJ, 86, 217

Zwicky F., Humason M. L., 1961, ApJ, 133, 794