Relative merits of different types of observations to constrain galaxy physical parameters

Camilla Pacifici

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ Paris VI

Spécialité:
Astronomie et Astrophysique

Présentée par

Camilla PACIFICI

pour obtenir le titre de:

DOCTEUR DE L'UNIVERSITÉ PIERRE ET MARIE CURIE

Mérites relatifs de différents types d'observations pour contraindre les paramètres physiques de galaxies

Thèse dirigée par Stéphane CHARLOT

Jury composé de :

Patrick BOISSÉ	Président du jury
Stéphane CHARLOT	Directeur de thèse
Bruno GUIDERDONI	Rapporteur
Roberto MAIOLINO	Rapporteur
Jérémy BLAIZOT	Examinateur
Emanuele DADDI	Examinateur

to those people I can call Friends

"We are stardust",
wrote Joni Mitchell,
and this is literally true
at the atomic level.

Whittet (2003) - Dust in the galactic environment

Contents

Abstract **ix**

Résumé **xi**

List of Figures **xxv**

List of Tables **xxviii**

1 Introduction **1**
1.1 The Universe . . . 1
1.2 Galaxies . . . 2
1.2.1 Galaxy morphologies . . . 3
1.2.2 Galaxy spectral energy distributions . . . 3
1.3 Outline . . . 7

2 Modeling galaxy spectral energy distributions **11**
2.1 The star formation history of galaxies . . . 11
2.1.1 The star formation history of the Universe . . . 12
2.1.2 The star formation history of individual galaxies . . . 12
2.2 Emission from stellar populations . . . 22
2.2.1 Stellar initial mass function . . . 22
2.2.2 Stellar evolutionary tracks . . . 24
2.2.3 Library of stellar spectra . . . 29
2.2.4 Modeling the spectral energy distribution of stellar populations 30
2.3 Nebular emission . . . 31
2.3.1 The interstellar medium . . . 33
2.3.2 Spectral features of the nebular emission . . . 36
2.3.3 Nebular emission models . . . 39
2.4 Attenuation by dust . . . 41
2.4.1 General properties of interstellar dust . . . 42
2.4.2 Modeling attenuation by dust in galaxies . . . 42
2.5 Absorption by the Intergalactic Medium . . . 49
2.6 Summary . . . 50

3 Relative merits of different types of rest-frame optical observations to constrain galaxy physical parameters **53**
3.1 Introduction . . . 53
3.2 Modeling approach . . . 55
3.2.1 Library of star formation and chemical enrichment histories . 55
3.2.2 Galaxy spectral modeling . . . 58
3.2.3 Library of galaxy spectral energy distributions . . . 63

3.2.4 Retrievability of galaxy physical parameters 65
3.3 Constraints on galaxy physical parameters from different types of observations 71
3.3.1 Constraints from multi-band photometry 71
3.3.2 Spectroscopic constraints 82
3.4 Application to a real sample 91
3.4.1 Physical parameters of SDSS galaxies 91
3.4.2 Influence of the prior distributions of physical parameters . . 95
3.5 Summary and conclusion 100

4 Constraining the physical properties of 3D-HST galaxies through the combination of photometric and spectroscopic data 103
4.1 Introduction 103
4.2 The data 104
4.3 Fits of the spectral energy distribution 107
4.3.1 Photometric approach 107
4.3.2 Spectroscopic approach 108
4.4 Results from photometric versus spectroscopic fits 111
4.4.1 Redshift 113
4.4.2 Mass and specific star formation rate 114
4.5 Possible causes of discrepancy between photometric and spectroscopic estimates 114
4.6 Summary and next steps 119

5 ACS and NICMOS photometry in the Hubble Ultra Deep Field 121
5.1 Introduction 121
5.2 The data 122
5.3 Modeling 122
5.3.1 Library of spectral energy distributions 124
5.3.2 The ultraviolet spectral slope 127
5.4 Fitting procedure 128
5.4.1 Estimates of the physical parameters 128
5.4.2 Preliminary pseudo-observed scene 130
5.5 Discussion 131
5.5.1 Comparison of redshift estimates 133
5.5.2 Correlation between ultraviolet spectral slope and optical depth of the dust 133
5.6 Summary and next steps 136

6 Conclusions 139

A Intrinsic correlations between spectral pixels 143

B Number of model galaxies in the library 147

C UDF data **149**

Bibliography **153**

Abstract

The light emitted by galaxies encloses important information about their formation and evolution. The spectral energy distribution of this light across the electromagnetic spectrum contains a myriad of details about the stellar, nebular and dust components of galaxies. To interpret these features in terms of constraints on physical parameters, we require sophisticated models of the spectral emission from galaxies.

In this thesis, we present a new approach to assess the relative merits of different types of observations to constrain galaxy physical parameters, such as stellar mass, star formation rate, metal enrichment of the gas and optical depth of the dust. To this goal, we build a comprehensive library of galaxy spectral energy distributions by combining the semi-analytic post-treatment of a large cosmological simulation with state-of-the-art models of the stellar and nebular emission and attenuation by dust. Using this library, we build a set of galaxy *pseudo-observations* by convolving spectral energy distributions of models with known parameters with different instrument response functions and then adding artificial noise. This approach allows us to quantify the accuracy and uncertainty to which physical parameters can be extracted from different types of observations. A main novelty of our approach is the ability to interpret simultaneously the stellar and nebular emission from galaxies, even at low spectral resolution.

We present three applications of our approach to the interpretation of spectroscopic and photometric observations of galaxies at different redshifts. We first analyze the medium-resolution, rest-frame optical spectra of a sample of $\sim 13,000$ nearby star-forming galaxies extracted from the Sloan Digital Sky Survey (SDSS). We show that the analysis of the combined stellar and nebular emission of these galaxies using our approach provides likelihood distributions of stellar mass, gas-phase oxygen abundance, total effective optical depth of the dust and specific star formation rate similar to those obtained in previous separate analyses of the stellar and nebular emission at the original (twice higher) SDSS spectral resolution.

Then, we apply our approach to the analysis of combined photometric and spectroscopic observations of a sample of galaxies at redshifts between 1 and 3. We explore the relative advantages and the complementarity of ultraviolet-optical-infrared photometry and rest-frame ultraviolet and optical spectroscopy to constrain the physical parameters of these galaxies.

Finally, we use our approach to simulate observations of primeval galaxies with the NIRSpec instrument onboard the future James Webb Space Telescope (JWST). We achieve this by feeding model spectra, calibrated using the deepest available photometric observations, into a preliminary version of the NIRSpec instrument performance simulator. We explore with particular attention the ability to interpret the rest-frame ultraviolet properties of primeval galaxies using our approach. This work will be valuable to quantify the improvements that low- and medium-resolution spectroscopy with JWST/NIRSpec will bring compared to currently available pho-

tometric datasets when exploring the physical parameters of the very first galaxies that formed in the early Universe.

The approach developed in this thesis can be used to extract valuable information from any kind of galaxy observation across the wavelength range covered by spectral evolution models as well as to plan for future galaxy observations.

Résumé

La lumière émise par les galaxies contient des informations importantes sur leur formation et leur évolution. La distribution spectrale d'énergie de cette lumière à travers le spectre électromagnétique contient elle même une myriade de détails sur les composants stellaire, nébulaire et de la poussière des galaxies. Pour interpréter ces caractéristiques en termes de contraintes sur certains paramètres physiques, nous avons besoin de modèles sophistiqués de l'émission spectrale des galaxies.

Dans cette thèse, nous présentons une nouvelle approche pour évaluer les mérites relatifs des différents types d'observations pour contraindre certains paramètres physiques des galaxies tels que la masse stellaire, le taux de formation stellaire, l'enrichissement en métaux du gaz et la profondeur optique de la poussière. Pour cet objectif, nous construisons une bibliothèque complète de distributions spectrales d'énergie des galaxies en combinant un traitement semi-analytique d'une grande simulation cosmologique avec des modèles de pointe de l'émission stellaire et nébulaire et de l'atténuation par la poussière. En utilisant cette bibliothèque, nous construisons un ensemble de *pseudo-observations* de galaxies par convolution des distributions spectrales d'énergie des modèles, dont nous connaissons les paramètres physiques, avec des fonctions de réponse de différents instruments, puis en ajoutant du bruit artificiel. Cette approche nous permet de quantifier la précision et l'incertitude à la quelle les paramètres physiques peuvent être extraites de différents types d'observations. Une principale nouveauté de notre approche est la capacité d'interpréter simultanément l'émission stellaire et nébulaire des galaxies, même à faible résolution spectrale.

Nous présentons trois applications de notre approche pour l'interprétation d'observations spectroscopiques et photométriques de galaxies à différents *redshifts*. Nous analysons d'abord des spectres optiques (*rest-frame*) en résolution moyenne d'un échantillon de $\sim$ 13,000 galaxies à proximité extraites du Sloan Digital Sky Survey (SDSS). Nous montrons que l'analyse de la composante stellaire combinée à la composante nébulaire de ces galaxies en utilisant notre approche fournit des distributions de probabilité de la masse stellaire, de l'abondance de l'oxygène en phase gazeuse, de la profondeur optique de la poussière et du taux spécifique de formation stellaire similaires à ceux obtenus dans des analyses séparées des composantes stellaire et nébulaire à la résolution spectrale originale du SDSS, qui est deux fois plus élevée.

Ensuite, nous appliquons notre approche à l'analyse combiné des observations photométriques et spectroscopiques d'un échantillon de galaxies à redshift entre 1 et 3. Nous explorons les avantages relatifs et la complémentarité de la photométrie ultraviolet, optique et infrarouge et de la spectroscopie ultraviolet et optique pour contraindre les paramètres physiques de ces galaxies.

Enfin, nous utilisons notre approche pour simuler des observations de galaxies primitives avec l'instrument NIRSpec à bord du future télescope spatial James

Webb (JWST). Nous réaliserons cela en utilisant des spectres, extraits de notre bibliothèque et calibré à l'aide des plus profondes observations photométriques disponibles, dans une version préliminaire du simulateur des performances de l'instrument NIRSpec. Nous explorons en particulier la capacité à interpréter les propriétés des longueurs d'onde ultraviolets (*reste-frame*) de galaxies primitives en utilisant notre approche. Ce travail sera très utile pour quantifier les améliorations que la spectroscopie de faible et moyenne résolution avec JWST / NIRSpec apporteront par rapport à des données photométriques actuellement disponibles en explorant les paramètres physiques des toutes premières galaxies qui se sont formées dans l'Univers primordial.

L'approche développée dans cette thèse peut être utilisée pour extraire des informations précieuses à partir de n'importe quel type d'observation de galaxies sur toute la gamme de longueurs d'onde couvertes par les modèles d'évolution spectrale ainsi que pour planifier de futures observations de galaxies.

List of Figures

1.1 A broad view of the most important telescopes and the wavelength ranges they cover. The atmospheric absorption coefficient is also shown. 3

1.2 Photometric images of the galaxy NGC1512 in different bands. Clockwise from left to right: far ultraviolet ($\sim$ 2000Å), near ultraviolet ($\sim$ 3000Å), optical ($\sim$ 5000Å), Hα emission (6563Å), near infrared ($\sim 1\mu$m), mid infrared ($\sim 10\mu$m) and far infrared ($\sim 100\mu$m). Credit: NASA. 4

1.3 Panel (a): spectral energy distribution of a galaxy at optical wavelengths (black line) overlaid are the five filter response functions of the SDSS. Panel (b): same spectral energy distribution as in panel (a) (black line), with the strongest emission lines labeled; the red line shows the stellar continuum underneath the lines. Panel (c): the same spectral energy distribution as in panel (a), computed at a 10-times lower resolution; red points mark the binning. Note that the ordinate scale in panel (c) is enlarged. 6

1.4 Flowchart of the general idea behind this thesis. 9

2.1 Fig. 9 from Madau et al. [1996], showing on the left the Universal metal ejection density, $\dot{\rho}_Z$, and on the right the total star formation rate density, $\dot{\rho}_*$, as a function of redshift. Triangle: Gallego et al. [1995]. Filled dots: Lilly et al. [1996]. Filled squares: lower limits from Hubble Deep Field images. 13

2.2 Fig. 5 from Cucciati et al. [2011], showing the total dust-corrected UV-derived star formation rate density as a function of redshift from the VVDS sample (red filled circles). The black dashed line is the star formation rate density as a function of z implied from the stellar mass density in Ilbert et al. [2010]. Other data sets are overplotted and labeled on the figure. All data have been converted into a star formation rate density with the scaling relation from Madau et al. [1998]. 13

2.3 Idealized star formation histories (panels a, b, c) and the relative spectral energy distributions (panels d, e, f). The model spectral energy distributions are computed using the latest version of Bruzual and Charlot [2003] at fixed solar metallicity, with a Chabrier initial mass function [Chabrier, 2003]. Emission lines are computed consistently following the prescription by Charlot and Longhetti [2001]. The attenuation by dust is neglected in these examples. 15

2.4 Fig. 10 from Kauffmann et al. [2003a]. Observed $g-r$ versus $r-i$ colors of a representative sample of SDSS galaxies (black points) for 2 particular redshifts (0.11 and 0.03). Blue points represent the model grid, computed applying Bruzual and Charlot [2003] models to exponentially declining star formation histories, with random superimposed burst of star formation. In the left panels, data are not corrected for dust attenuation, while in the right panels, data are corrected with an attenuation law of the form $\tau_\lambda \propto \lambda^{-0.7}$. The red arrow shows the predicted reddening vector. 16
2.5 Fig. 2 from De Lucia and Blaizot [2007]. The merger tree of a dark matter halo of $\sim 9 \times 10^{14}\ M_\odot$. The area of the symbols scales with the mass. Circles represent haloes that are part of the friends-of-friends group, while triangles show haloes that have not joined yet the friends-of-friends group. The halo which contains the main branch is marked in green. The trees on the right-hand side, which are not linked to the main halo, correspond to the other substructures identified in the friends-of-friends group at $z = 0$. 19
2.6 Stellar initial mass faction normalized to solar units. Salpeter [1955], blue line. Miller and Scalo [1979], magenta line. Kroupa [2001], green line. Chabrier [2003], dashed line. 23
2.7 The Hertzsprung-Russell diagram (H-R) shows the evolutionary stages of stars in terms of their luminosity and temperature. Different evolutionary paths for stars of different initial masses are plotted in different colors. The zero-age main sequence is shown as a black solid line. The labels indicate the main stages of the evolution. . . . 25
2.8 Fig. 2 and 3 from Heger et al. [2003]. Stellar mass at the time of final explosion or remnant formation (blue line), remnant mass (red line) and metals released in the interstellar medium (green fill and hatching) as a function of initial mass of the stars for solar metallicity (top panel) and for zero metallicity (bottom panel). 28
2.9 Top panel: isochrones from Bertelli et al. for different ages at solar metallicity, solid lines; isochrones from other models in the Padova database at solar metallicity, dashed lines. Bottom panel: zoom of the yellow region in the top panel, showing isochrones (ages between 0.06 and 16 Gyr) form Marigo et al. [2008] (Fig. 1) for pre-TP-AGB phase in green, and O-rich and C-rich configurations of the TP-AGB phase respectively in blue and red. 32
2.10 Representation of the different phases of the interstellar medium in terms of temperature and density. 34
2.11 A schematic representation of an H II region. The ionizing source is shown in the center. The region occupied by the ionized hydrogen (Stromgren sphere) is shown in blue. The region filled by the neutral hydrogen is shown in pink. (Figure from http://www.astro.cornell.edu/academics/) 36

2.12 Frequency variation of the free-bound continuous-emission coefficient $\gamma_\nu(T)$ for hydrogen (black solid line) and for helium (black thin solid line). Frequency variation of the two-photon-process coefficient (grey solid line). All lines correspond to a low-density regime at $T = 10,000$ K. Figure from Osterbrock and Ferland [2006]. 37

2.13 Typical spectrum of an H II region with the strongest lines labeled. (htp://frigg.physastro.mnsu.edu) 38

2.14 Extinction curves (A_λ/A_V) as a function of wavelength in the ultraviolet and optical rest-frame ranges. The red line shows the average extinction curve of the Milky Way as in Cardelli et al. [1989] ($R_V = 3.1$). The short-dashed line shows the extinction curve of the Small Magellanic Cloud bar, while the long-dashed line shows the average extinction curve of the Large Magellanic Cloud, both as in Gordon et al. [2003]. The blue line represents the extinction curve by Calzetti et al. [2000] and the green line shows a curve of the form $\lambda^{-0.7}$ normalized at 5500Å. 45

2.15 Trasmission coefficient in the intergalactic medium at three different redshifts: $z = 3.0$ (red line), $z = 4.5$ (green line), $z = 6.0$ (blue line). Dotted lines represent the Lyman limit (912Å) at the three different redshifts. 50

3.1 Example of galaxy star formation and chemical enrichment histories inferred from the semi-analytic post-treatment of the Millennium cosmological simulation: (a) star formation rate, ψ, and (b) interstellar metallicity, Z, plotted as a function of look-back time, for a galaxy with a present-day stellar mass of $1.5 \times 10^{10} M_\odot$. The vertical dashed line marks the evolutionary stage at which the galaxy is looked at in Fig. 3.3 (corresponding to a galaxy age of 8.4 Gyr). 56

3.2 Emission-line luminosities of galaxies computed using the models described in Section 3.2.2.2 (lines), compared with high-quality observations of a sample of 28,075 star-forming galaxies from the SDSS DR7 (black dots). The models assume constant star formation over the past 10 Myr. The data are corrected for attenuation by dust as described in Brinchmann et al. [2004]. (*a*) $L([\mathrm{O\,III}]\lambda5007)/L(\mathrm{H}\beta)$ versus $L([\mathrm{N\,II}]\lambda6584)/L(\mathrm{H}\alpha)$. (*b*) $L([\mathrm{O\,III}]\lambda5007)/L(\mathrm{H}\beta)$ versus $L([\mathrm{S\,II}]\lambda\lambda6716,6731)/L(\mathrm{H}\alpha)$. (*c*) $L([\mathrm{O\,III}]\lambda5007)/L(\mathrm{H}\beta)$ versus $L([\mathrm{N\,II}]\lambda6584)/L([\mathrm{O\,II}]\lambda3727)$. (*d*) $L([\mathrm{O\,III}]\lambda5007)/L(\mathrm{H}\beta)$ versus $L([\mathrm{O\,II}]\lambda3727)/L([\mathrm{O\,III}]\lambda5007)$. In each panel, lines of different colours refer to models with different zero-age ionization parameter $\log U_0$ (cyan: -1.5; red: -2.0; green: -2.5; magenta: -3.0; blue: -3.5). At fixed $\log U_0$, the lower dashed line corresponds to models with dust-to-metal ratio $\xi_\mathrm{d} = 0.1$, the solid line to models with $\xi_\mathrm{d} = 0.3$ and the upper dashed line to models with $\xi_\mathrm{d} = 0.5$. Along each line, dots mark the positions of models with different gas metallicity Z (0.10, 0.20, 0.50, 0.75, 1.00, 1.50, 2.00 and 3.00 times $Z_\odot$, from left to right; only low-metallicity models are considered for $\log U_0 = -1.5$). In (*a*), the gray long-short-dashed line shows the Kauffmann et al. [2003b] criterion to separate star-forming galaxies from AGNs. 61

3.3 Spectral energy distribution of the same model galaxy as in Fig. 3.1 at the age of 8.4 Gyr, computed using the models described in Sections 3.2.1 and 3.2.2. (*a*) At a spectral resolution of 5 Å FWHM ($R = 1000$ at 5000 Å). The black portion of the spectrum (λ=3600–7400 Å) is that used to retrieve galaxy physical parameters at this resolution. (*b*) Same as in (*a*), but at a spectral resolution of 50 Å FWHM ($R = 100$ at 5000 Å). (*c*) Same as in (*a*), but convolved with the SDSS *ugriz* filter response functions (shown at the bottom). The model galaxy has $f_\mathrm{SFH} = 0.3$ and current parameters (resampled using the distributions in Table 3.1) $\psi_\mathrm{S} = 0.08\,\mathrm{Gyr}^{-1}$, $Z = 0.4Z_\odot$, $\log U_0 = -2.8$, $\xi_\mathrm{d} = 0.3$, $\hat{\tau}_V = 1.0$, $\mu = 0.6$ and $n = 0.7$. In (*a*) and (*b*), the adopted pixel size is half the size of a resolution element. . . 66

3.4 Prior distributions of selected physical parameters of the 5 million galaxies in the spectral library generated in Section 3.2.3: (*a*) observer-frame absolute r-band stellar mass-to-light ratio, M_*/L_r; (*b*) fraction of current stellar mass formed during the last 2.5 Gyr, $f_{\rm SFH}$; (*c*) specific star formation rate, $\psi_{\rm S}$; (*d*) gas-phase oxygen abundance, $12 + \log({\rm O/H})$; (*e*) total effective V-band absorption optical depth of the dust, $\hat{\tau}_V$; (*f*) fraction of $\hat{\tau}_V$ arising from dust in the ambient ISM, μ. In each panel, the shaded histogram shows the distribution for all galaxies, while the solid histogram shows the contribution by star-forming galaxies alone. Non-star-forming galaxies are off scale (at $\log \psi_{\rm S} = -\infty$) in panel (*c*) and do not contribute to the distributions of interstellar parameters in panels (*d*)–(*f*). 69

3.5 Probability density functions of the same physical parameters as in Fig. 3.4 retrieved, using the Bayesian approach described in Section 3.2.4, from the pseudo-galaxy spectrum shown at the top. The spectrum was obtained by adding artificial noise with $\overline{\rm S/N} = 20$ to the $R = 100$ model spectrum of Fig. 3.3b. In each panel, the black triangle indicates the true value of the parameter of the pseudo-galaxy, the solid line the best estimate (50th percentile of the retrieved likelihood distribution) and the dashed lines the associated 68-percent confidence interval (16th and 84th percentiles). 70

3.6 Average probability density functions of the same 6 physical parameters as in Fig. 3.4 retrieved, using the Bayesian approach described in Section 3.2.4, from 5-band *ugriz* photometry with ${\rm S/N} = 30$ of a sample of 10,000 pseudo-galaxies (standard case). Each panel corresponds to a different parameter [from left to right: M_*/L_r, $f_{\rm SFH}$, $\psi_{\rm S}$, $12 + \log({\rm O/H})$, $\hat{\tau}_V$ and μ]. In each case, average probability density functions in 50 narrow bins of true parameter value were obtained by coadding and then renormalizing the probability density functions retrieved for the 10,000 pseudo-galaxies (see Section 3.3.1.1 for detail). Grey levels locate the 2.5th, 12th, 16th, 22nd, 30th, 40th, 60th, 70th, 78th, 84th, 88th and 97.5th percentiles of the average likelihood distribution in each bin. The solid line locates the associated median (best estimate) and the 2 dashed lines the 16th and 84th percentiles (68-percent confidence interval). For reference, 3 dotted lines indicate the identity relation and deviations by a factor of 2 (± 0.3 dex) between the retrieved and true parameter values. 72

3.7 (*a*) $g - i$ colour plotted against $u - g$ colour for a subset of 50,000 models from the galaxy spectral library assembled in Section 3.2.3. The models are colour-coded according to specific star formation rate, as indicated. (*b*) Same as (*a*), but without including the contribution by nebular emission to broadband fluxes. 74

3.8 Same as Fig. 3.6, but for 3 distinct alternatives to the standard case: (*top row*) adopting a signal-to-noise ratio S/N = 10 instead of 30; (*middle row*) using constraints from only the *ugr* photometric bands instead of *ugriz*; and (*bottom row*) not including nebular emission in the model library used to analyze the sample of 10,000 pseudo-galaxies. In all cases, the improvement factor I_σ and the gain in accuracy Δ (equations 3.3.13–3.3.14) are shown as a function of true parameter value at the bottom of each panel to quantify differences in the retrieved likelihood distributions relative to the standard case of Fig. 3.6. 77

3.9 Prior distributions of the same physical parameters as in Fig. 3.4 for the 5 million galaxies at redshifts between 0 and 1 in the spectral library assembled in Section 3.3.1.2. In each panel, the shaded histogram shows the distribution for all galaxies, while the solid histogram shows the contribution by star-forming galaxies alone. Non-star-forming galaxies are off scale (at $\log \psi_S = -\infty$) in panel (*c*) and do not contribute to the distributions of interstellar parameters in panels (*d*)–(*f*). 78

3.10 (*a*) Distribution of the difference between the retrieved best estimate (i.e. median) and true value of z_{obs}, in units of $1 + z_{obs}$, for a sample of 10,000 pseudo-galaxies at redshifts between 0 and 1 observed in the *ugriz* photometric bands. The shaded and solid histograms show the distributions obtained for S/N = 30 and S/N = 10, respectively. For reference, a dashed line indicates the $\sigma = 0.03$ Gaussian accuracy quoted by Ilbert et al. [2006], who analyzed CFHTLS ($S/N \gtrsim 30$) $u^*g'r'i'z'$ photometry of about 3000 galaxies with spectroscopic redshifts between 0.2 and 1.5. (*b*) Detail of the average retrieved probability density function plotted against z_{obs}, for S/N = 30. The lines and shading have the same meaning as in Fig. 3.6. (*c*) Same as (*b*), but for S/N = 10. 80

3.11 Average probability density functions of the same 6 physical parameters as in Fig. 3.9 retrieved, using the Bayesian approach described in Section 3.2.4, from 5-band *ugriz* photometry with S/N = 30 of samples of 10,000 pseudo-galaxies at random redshifts: (*top row*) drawn in the redshift range $0.2 < z_{obs} < 0.4$; and (*bottom row*) drawn in the redshift range $0.6 < z_{obs} < 0.8$. In each case, the redshift z_{obs} of a galaxy is assumed not to be known a priori, and the probability density functions are computed using models at all redshifts between 0 and 1 in the spectral library assembled in Section 3.3.1.2. In all panels, the lines and shading have the same meaning as in Fig. 3.6. . 81

3.12 Average probability density functions of the specific star formation rate, $\psi_{\rm S}$, gas-phase oxygen abundance, $12 + \log({\rm O/H})$, and dust attenuation optical depth in stellar birth clouds, $(1-\mu)\hat{\tau}_V$, retrieved, using the Bayesian approach described in Section 3.2.4, from the equivalent widths of optical emission lines in a sample of 10,000 star-forming pseudo-galaxies. The sample was extracted from the library of galaxy spectral energy distributions assembled in Section 3.2.3, assuming a median signal-to-noise ratio per pixel $\overline{\rm S/N} = 20$ and requiring 3σ measurements of all line equivalent widths. (*Top row*) using the equivalent widths of [O II]λ3727; Hβ; [O III]$\lambda\lambda$4959, 5007; [N II]λ6548+Hα+[N II]λ6584; and [S II]$\lambda\lambda$6716, 6731 at a spectral resolution of 50 Å FWHM ($R = 100$ at $\lambda = 5000$ Å). (*Bottom row*) using the equivalent widths of [O II]λ3727; Hβ; [O III]λ4959; [O III]λ5007; [N II]λ6548; Hα; [N II]λ6584; [S II]λ6716 and [S II]λ6731 at a spectral resolution of 5 Å FWHM ($R = 1000$ at $\lambda = 5000$ Å). In all panels, the lines and shading have the same meaning as in Fig. 3.6. 84

3.13 Average probability density functions of the same 6 physical parameters as in Fig. 3.4 retrieved, using the Bayesian approach described in Section 3.2.4, from low-resolution optical spectra of a sample of 10,000 pseudo-galaxies. The spectra cover the wavelength range from $\lambda = 3600$ to 7400 Å at the resolution of 50 Å FWHM ($R = 100$ at $\lambda = 5000$ Å) with median signal-to-noise ratio per pixel $\overline{\rm S/N} = 20$. (*Top row*) for a sample of 10,000 pseudo-galaxies extracted randomly from the spectral library assembled in Section 3.2.3. The improvement factor I_σ and the gain in accuracy Δ (equations 3.3.13–3.3.14) are shown as a function of true parameter value at the bottom of each panel to quantify differences in the retrieved likelihood distributions relative to the standard case of Fig. 3.6. (*Bottom row*) for a sample of 10,000 star-forming pseudo-galaxies extracted from the same spectral library, with the requirement that the net Hα+[N II] emission equivalent width be greater than 5 Å. In all panels, the lines and shading have the same meaning as in Fig. 3.6. 86

3.14 Same as Fig. 3.13, but: (*top row*) adopting a median signal-to-noise ratio per pixel $\overline{\rm S/N} = 5$ instead of 20; (*bottom row*) not including nebular emission in the model library used to analyze the sample of 10,000 pseudo-galaxies. In both cases, the improvement factor I_σ and the gain in accuracy Δ (equations 3.3.13–3.3.14) are shown as a function of true parameter value at the bottom of each panel to quantify differences in the retrieved likelihood distributions relative to the standard case of Fig. 3.6. 87

3.15 Average probability density functions of the same 6 physical parameters as in Fig. 3.4 retrieved, using the Bayesian approach described in Section 3.2.4, from medium-resolution optical spectra of a sample of 10,000 pseudo-galaxies. The spectra cover the wavelength range from $\lambda = 3600$ to 7400 Å at the resolution of 5 Å FWHM ($R = 1000$ at $\lambda = 5000$ Å) with median signal-to-noise ratio per pixel $\overline{\mathrm{S/N}} = 20$. (*Top row*) for a sample of 10,000 pseudo-galaxies extracted randomly from the spectral library assembled in Section 3.2.3. The improvement factor I_σ and the gain in accuracy Δ (equations 3.3.13–3.3.14) are shown as a function of true parameter value at the bottom of each panel to quantify differences in the retrieved likelihood distributions relative to the standard case of Fig. 3.6. (*Bottom row*) for a sample of 10,000 star-forming pseudo-galaxies extracted from the same spectral library, with the requirement that the net Hα emission equivalent width be greater than 5 Å. In all panels, the lines and shading have the same meaning as in Fig. 3.6. 90

3.16 Estimates of physical parameters retrieved from the spectra of 12,660 SDSS star-forming galaxies (degraded to a resolution of 5 Å FWHM) using the Bayesian approach described in Section 3.2.4 plotted against estimates of the same parameters from different sources, as indicated. (*a*) Stellar mass, $M_\star$. (*b*) Gas-phase oxygen abundance, $12 + \log(\mathrm{O/H})$. (*c*) Total V-band attenuation optical depth of the dust, $\hat{\tau}_V$. (*d*) Specific star formation rate, ψ_S. In each panel, the likelihood distributions from previous studies were reconstructed from the published 16th, 50th and 84th percentiles and combined with the likelihood distributions obtained in this work to generate 2D probability density functions. The contours depict the normalized co-added 2D probability density function including all 12,660 galaxies in the sample (on a linear scale 25 levels, the outer edge of the lightest grey level corresponding to the 96th percentile of the 2D distribution). The solid line is the identity relation. 93

3.17 Two-dimensional probability density functions of the fraction of the current stellar mass formed during the last 2.5 Gyr, $f_{\rm SFH}$, versus stellar mass, $M_\star$, specific star formation rate, $\psi_{\rm S}$, gas-phase oxygen abundance, $12 + \log({\rm O/H})$, total V-band attenuation optical depth of the dust, $\hat{\tau}_V$, and fraction of $\hat{\tau}_V$ arising from dust in the ambient ISM, μ. All parameters were retrieved from the spectra of the same 12,660 SDSS star-forming galaxies as in Fig. 3.16 (degraded to a resolution of 5 Å FWHM) using the Bayesian approach described in Section 3.2.4. In each panel, the contours depict the normalized co-added 2D probability density function including all 12,660 galaxies in the sample (on a linear scale of 25 levels, the outer edge of the lightest grey level corresponding to the 96th percentile of the 2D distribution). All physical quantities pertain to the regions of galaxies probed by the 3-arcsec-diameter SDSS fibers. 95

3.18 Prior distributions of the same physical parameters as in Fig. 3.4 for the 5 million galaxies in the spectral library generated in Section 3.4.2. These distributions are those arising, at the mean redshift $z = 0.07$ of the SDSS sample of Section 3.4.1, from the original semi-analytic post-treatment of the Millennium cosmological simulation of Springel et al. [2005] by De Lucia and Blaizot [2007, see text for detail]. In each panel, the shaded and thin solid histograms have the same meaning as in Fig. 3.4. In panels (*a*), (*c*), (*d*) and (*e*), the thick solid histograms show the distributions of the best estimates of M_*/L_r, $\psi_{\rm S}$, $12 + \log({\rm O/H})$ and $\hat{\tau}_V$ from Gallazzi et al. [2005], Brinchmann et al. [2004] and Tremonti et al. [2004] for the same sample of 12,660 SDSS star-forming galaxies as in Fig. 3.16. 97

3.19 Same as Fig. 3.16, but using the prior distributions in Fig 3.18 instead of those in Fig. 3.4 to retrieve the probability density functions of $M_\star$, $12 + \log({\rm O/H})$, $\hat{\tau}_V$ and $\psi_{\rm S}$ for the 12,660 SDSS star-forming galaxies in the sample. 98

4.1 Galaxy #1705. The G141 spectrum (red dotted line), the part of the spectrum used in the analysis ($1.1 \leq \lambda \leq 1.6\mu$m, blue solid line), J-band effective wavelength (magenta dashed line) and square band of 0.1 μm width (magenta solid line). 106

4.2 Top panel: MAGPHYS fit ($z_{\rm grism} = 2.21$ is the input redshift) to the FIREWORKS photometry (red dots) of galaxy #1799; the black spectrum is the best-fit model and the blue spectrum is the pure stellar spectrum (i.e. unattenuated by dust) corresponding to this best-fit model. Middle panel: zoom of the observed-frame wavelength range from the ultraviolet to the near-infrared; FIREWORKS photometry (red dots), MAGPHYS best-fit model spectrum (black line), 3D-HST spectrum (green line) with error bars (grey). Bottom panels: histograms showing the likelihood distributions of the fraction of total dust luminosity contributed by the diffuse ISM, f_μ, the total optical depth seen by young stars in birth clouds, τ_V, the optical depth seen by stars in the diffuse ISM, $\mu\tau_V$, the star formation rate, ψ, the specific star formation rate, ψ_S, the stellar mass, M_*, and total dust luminosity, $L_{\rm d}^{\rm tot}$. 109

4.3 Prior distributions of selected physical parameters of the 2,500,000 galaxies in the spectral library: (a) rest-frame absolute V-band stellar mass-to-light ratio, M_*/L_{55}; (b) fraction of the current galaxy stellar mass formed during the last Gyr, $f_{\rm SFH}$; (c) specific star formation rate, $\psi_{\rm S}$; (d) gas-phase oxygen abundance, $12 + \log({\rm O/H})$; (e) total effective V-band absorption optical depth of the dust, $\hat{\tau}_V$; (f) fraction of $\hat{\tau}_V$ arising from dust in the ambient ISM, μ. In each panel, the shaded histogram shows the distribution for all galaxies, while the solid histogram shows the contribution by star-forming galaxies alone. Non-star-forming galaxies are off scale (at $\log(\psi_{\rm S}) = \infty$) in panel (c) and do not contribute to the distributions of interstellar parameters in panels (d)–(f). 110

4.4 Spectral fit of galaxy #1799, using this thesis approach, with no prior information on redshift. Top panel: the red line is the observed 3D-HST spectrum, and the blue line is the best-fit model spectrum. Medium panel: residuals between the observed and the best-fit model spetra. Bottom panels: histograms showing the likelihood distributions of stellar mass, fraction of stellar mass formed in the last Gyr, specific star formation rate, gas-phase oxygen abundance, total effective optical depth of the dust and redshift. 112

4.5 Comparison between the redshift estimates obtained with our spectroscopic approach, with the redshift provided in the catalogue ($z_{\rm grism}$). Green dots represent the median estimates when redshift is treated as a free parameter. Blue crosses represent the median estimates obtained fitting each galaxy only with models in the range $z_{\rm grism} \pm 0.3$. Error bars are the 16 and 84 percentiles in the likelihood distributions. 113

4.6 Same as Fig. 4.4. Spectral fit of galaxy #5435, using the approach developed in this thesis. No constraint is applied on the redshift. . . 115

4.7 Same as Fig. 4.4. Spectral fit of galaxy #5434, using the approach developed in this thesis. The redshift is constrained to be within 0.3 of $z_{\rm grism}$. 116

4.8 Left panel: comparison between stellar mass estimates obtained with our spectroscopic approach and the ones obtained with MAGPHYS. Right panel: comparison between specific star formation rate estimates obtained with our spectroscopic approach and the ones obtained with MAGPHYS. In both panels, green dots represent the median estimates when redshift is treated as a free parameter. Blue squares represent the median estimates obtained when the redshift is constrained to be within 0.3 of $z_{\rm grism}$. Error bars represents the 16th and 84th percentiles of the likelihood distributions. 117

4.9 Top panel: MAGPHYS fit ($z_{\rm grism} = 1.02$ is the input redshift) to the FIREWORKS photometry (red dots) of galaxy #5209; the black spectrum is the best-fit model and the blue spectrum is the pure stellar spectrum (i.e. unattenuated by dust) corresponding to this best-fit model. Middle panel: zoom of the observed-frame wavelength range from the ultraviolet to the near-infrared; FIREWORKS photometry (red dots), MAGPHYS best-fit model spectrum (black line), 3D-HST spectrum (green line) with error bars (grey). Bottom panels: histograms showing the likelihood distributions of the fraction of total dust luminosity contributed by the diffuse ISM, f_μ, the total optical depth seen by young stars in birth clouds, τ_V, the optical depth seen by stars in the diffuse ISM, $\mu\tau_V$, the star formation rate, ψ, the specific star formation rate, ψ_S, the stellar mass, M_*, and total dust luminosity, $L_{\rm d}^{\rm tot}$. 118

5.1 ACS and NICMOS filters: B in blue, V in magenta, i' in cyan, z' in green, J in orange and H in red. In gray, five model galaxies at redshifts 2.15, 3.55, 4.40, 5.78 and 6.72. Flux is expressed in AB magnitudes in the observed-range between 0.3 and 2 μm. The increasing effect of the IGM absorption with redshift is visible blueward of the Lyman-α line. 123

5.2 Panel *a*: $V-i'$ versus $i'-z'$ colors. Panel *b*: $i'-z'$ versus $z'-J$ colors. In both plots, dots represent model galaxies in different ranges of redshift: $z < 3.5$ black, $3.5 < z < 4.5$ red, $4.5 < z < 5.5$ yellow, $5.5 < z < 6.5$ green, $6.5 < z < 7.5$ magenta. Grey squares with error bars represent the selected HUDF galaxies. Photometric redshifts from the catalogue by Coe et al. [2006] label a few high-redshift galaxies to identify the outliers in panel (a) and check their position in panel (b). 124

5.3 Prior distributions of selected physical parameters of the 500,000 galaxies in the spectral library: (a) observer-frame absolute H-band stellar mass-to-light ratio, M_*/L_H; (b) redshift, z; (c) specific star formation rate, $\psi_{\rm S}$; (d) gas-phase oxygen abundance, $12 + \log({\rm O/H})$; (e) total effective V-band absorption optical depth of the dust, $\hat{\tau}_V$; (f) fraction of $\hat{\tau}_V$ arising from dust in the ambient ISM, μ. In each panel, the shaded histogram shows the distribution for all galaxies, while the solid histogram shows the contribution by star-forming galaxies alone. Non-star-forming galaxies are off scale (at $\log(\psi_{\rm S}) = \infty$) in panel (c) and do not contribute to the distributions of interstellar parameters in panels (d)–(f). 126

5.4 Left panel: rest-frame ultraviolet spectral energy distribution of a galaxy ($z = 3.41$) extracted form the model library (black) and the associated power-law best fit (green). The wavelength windows used to perform the fit are highlighted in cyan. Right panel: distribution of the rest-frame ultraviolet slope in the model library for all galaxies (shaded histogram) and for galaxies with negligible 2175Å bump only (solid histogram). 127

5.5 ID 5299. Top panel: ACS observed magnitudes (magenta dots), NICMOS observed magnitudes (red dots), ACS and NICMOS magnitudes of the best-fit model galaxy (blue squares), best-fit model spectrum when fitting both ACS and NICMOS magnitudes (black line), best-fit model spectrum when fitting ACS data alone (green line). Middle panel: residuals between best-fit model magnitudes and data (crosses). Bottom panels: probability density functions retrieved for stellar mass, specific star formation rate, gas-phase oxygen abundance, total effective optical depth of the dust, rest-frame ultraviolet slope, redshift. In each panel, the shaded histogram was derived by fitting both ACS and NICMOS data, while the solid histogram was derived using ACS data alone. In the redshift panel, the vertical red line shows the photometric redshift listed in the Coe et al. [2006] catalogue. 129

5.6 Example of the multi-object spectrograph (MOS) scene obtained by placing different point-like galaxies 'behind' different micro-shutters. Light is dispersed using NIRSpec low-resolution prism ($R = 100$). . . 131

5.7 Zoom of the scene presented in Figure 5.6. In correspondence of every galaxy, 3 vertically-adjacent micro-shutters are open. In each case, the radiation of the Zodiacal light dominates the background in the upper and lower micro-shutters. The continuum emission from the galaxies and some emission lines (red isolated pixels) can be detected in the central micro-shutter. 132

5.8 Differences in redshift estimates between our fits and the values reported in Coe et al. [2006] catalogue. Gray squares represent the fits using ACS and NICMOS magnitudes. Black dots represent the fits when ACS magnitudes alone are available. For each measurement, the error bar indicates the sum of the 16th–84th percentile range of the likelihood distribution retrieved from our analysis and the 1-sigma uncertainty in the Coe et al. estimate (≈ 0.3). The red dotted line marks the location of perfect agreement between the two redshift estimates. 134

5.9 Estimates of rest-frame ultraviolet slope, β, versus total effective optical depth of the dust, $\hat{\tau}_V$. Grey dots: 50,000 random galaxies from the model library. Red ($z < 4$) and magenta ($z > 4$) squares: median estimates fitting both ACS and NICMOS data. Green dots: median estimates fitting ACS data alone. The black dashed line shows the relation derived by Meurer et al. [1999] for nearby, ultraviolet-selected starburst galaxies (see text for details). 135

A.1 Representation of the mean pseudo-observational error correlation matrix $Corr_{i,i'}$ (equation A.0.7) across the wavelength range from 3600 to 7400 Å, derived as described in the text from the residuals of the fits of 3000 pseudo-observed spectral energy distributions with the library of 5 million models assembled in Chapter 3. The color scale is indicated at the top. 145

B.1 Uncertainty and accuracy (defined as in the caption of Table 3.2) of the stellar mass-to-light ratio M_*/L_r (black curves) and specific star formation rate ψ_S (blue curves) retrieved from 2 types of spectral fits, plotted against number of spectral energy distributions in the model library. (*a*) & (*b*) As retrieved from medium-resolution spectroscopy. (*c*) & (*d*) As retrieved form multi-band photometry (note that the vertical scales are different from those in Panels *a* and *b*). 148

List of Tables

2.1 Basic characteristics of the most recent N-body cosmological simulations. We cite Gottlöber et al. [2001], Hatton et al. [2003], the Millennium Simulation by Springel et al. [2005], the Millennium-II simulation by Boylan-Kolchin et al. [2009], the Horizon simulation by Teyssier et al. [2009], the Horizon Run by Kim et al. [2009] and the most recent Bolshoi simulation by Klypin et al. [2011]. For these simulations we report: the size of the box in h^{-1} Mpc, the number of dark matter particles, the mass of the single particle in h^{-1} solar masses, the redshift where the initial conditions are created, the number of timesteps for which data are saved, and the cosmological parameters (omega matter, $\Omega_{\rm m}$, omega lambda, Ω_Λ, the Hubble constant at redshift 0 in units of 100 km s^{-1} Mpc^{-1}, h, and the root mean squared amplitude of linear mass fluctuations in $8h^{-1}$ Mpc spheres at $z = 0$, σ_8). 18

2.2 Examples of theoretical and empirical stellar libraries. We report name, type, wavelength range, FWHM around 5500 Å and reference. 30

2.3 Average temperatures and densities of the gas in the interstellar medium in different phases. 33

3.1 Prior distributions of the current (i.e. averaged over the last 10 Myr) physical properties of star-forming galaxies in the library of star formation and chemical enrichment histories assembled in Sections 3.2.1 and 3.2.2. 64

3.2 Summary of the constraints on galaxy physical parameters retrieved from different types of optical observations using the method developed in Section 3.2. For each parameter and each type of observation, the quoted uncertainty (defined as half the 16th–84th percentile range of the retrieved likelihood distribution) and accuracy (defined as the absolute difference between the retrieved best estimate and true parameter value) are median quantities determined from the analysis of 10,000 pseudo-observations in Section 3.3. 99

4.1 ID and spectroscopic redshifts of the galaxies in the sample described in Section 4.2. 105

4.2 Scaling factors of the galaxies in the sample defined as the ratio between the median flux in a square band of 0.1μm width, centered on the J-band effective wavelength, and the FIREWORKS mean flux in the J-band. 106

4.3 Differences in the prescriptions adopted in MAGPHYS and in the spectroscopic approach described in this thesis. 117

Chapter 1

Introduction

"A Nebula is a celestial object, often of irregular form and brightness, appearing like a mass of luminous fog. ... Among elliptic nebulae, the signal object is the great nebula in Andromedae.", New Astronomy, Todd [1906].

1.1 The Universe

At the beginning of the 20th century, the measured distances towards objects in the sky were quite uncertain and the Milky Way was believed to be the only system in the Universe. In 1925, Edwin Hubble discovered that certain objects, called *nebulae*, were too far away to be part of the Milky Way. He suggested that the Universe extended much beyond our Galaxy and that those *nebulae* had intrinsic sizes similar to the Milky Way. The evidence that changed the vision of the Universe was that the light reaching us from distant galaxies is shifted to the red compared to what is seen in laboratories. This *redshift* is defined as

$$z = \frac{\lambda_{\rm obs} - \lambda_{\rm em}}{\lambda_{\rm em}}\,, \tag{1.1.1}$$

where $\lambda_{\rm em}$ is the wavelength at which light is emitted by a distant galaxy and $\lambda_{\rm obs}$ is the wavelength at which this light is detected by the observer. Hubble observed a correlation between the redshift and the distance of these galaxies, known as the *Hubble law*

$$cz = H_0\, d\,, \tag{1.1.2}$$

where cz is the recession velocity[1] of the galaxy, d is the distance to the galaxy and H_0 is the Hubble constant. Redshift can therefore be used as a measure of distance to a galaxy. Hubble concluded also that extragalactic objects are moving away from the Milky Way [Hubble, 1929] and that the Universe is expanding.

The expansion of the Universe can be understood in the framework of Friedmann's solution to Einstein field equations of general relativity[2]. In the standard model of Big Bang cosmology, the Universe expanded and cooled from a hot and dense initial state about 13.7 Gyr ago. It cooled sufficiently to allow energy to be converted into various subatomic particles which then combined to form atoms.

[1] This is based on the Doppler effect, $\Delta\lambda/\lambda = v/c$ and is valid for non-relativistic objects ($z \ll 1$). This relation can be extended in special and general relativity, and be applied to objects with $z > 1$.

[2] Einstein field equations describe gravitational interactions as a result of spacetime being curved by matter and energy.

Recent observations of the cosmic microwave background radiation and of distant supernovae suggest that this expansion is accelerating (see for example, Spergel et al. 2003, Lamarre et al. 2003, Riess et al. 1998) under the influence of a mysterious dark energy. This component can be formalized with the *cosmological constant* Λ in Einstein equations[3].

The other main component that appears to be driving the evolution of the Universe is dark matter. Dark matter neither emits nor scatter light. It interacts only gravitationally and thus cannot be observed directly. The dark mass component revealed by the rotational curve of galaxies and by gravitational lensing looking through galaxy clusters are evidences of the presence of this dark matter.

In the standard model, *cold* (i.e. non-relativistic) dark matter accounts for ~ 23 percent, while dark energy constitutes ~ 72 percent of the mass-energy density of the Universe. Only the remaining 5 percent is in the form of visible baryonic matter (stars, gas and dust). The general idea is that dark matter structures grow from weak density fluctuations present in the homogeneous and rapidly expanding early universe. These fluctuations are amplified by gravity, turning into the structures that we observe today.

Although many cosmological simulations have been developed to describe the evolution of the baryons in these dark-matter structures, the complexity of the physical processes involved (gas cooling mechanisms, feedback, the process of star formation) has prevented the full understanding of the formation and evolution of the galaxies we see today.

Progress in this area has come mainly from the advent of deep sky observations at the end of the 20th century. In fact, because of the finite speed of light, the light reaching us from distant galaxies provides images from the past. Therefore, observations of large samples of galaxies at various redshifts should help us better understand the formation and evolution of galaxies in a hierarchical Universe.

1.2 Galaxies

Over the past two decades, the advent of large ground-based and space-based observatories has largely improved our knowledge of galaxies. These observatories have allowed us to probe the light emitted by galaxies at different energies or wavelengths, from γ-rays to radio. Figure 1.1 provides an overview of the current and future instruments enabling the collection of galaxy lights across the electromagnetic spectrum. In this thesis we focus on the observation of galaxies at ultraviolet, optical and near-infrared wavelengths. We now recall a few basic concepts about the interpretation of galaxy observations.

[3]The cosmological constant was first proposed by Albert Einstein as a modification of his original theory of general relativity to achieve a stationary universe. Einstein abandoned the concept after Hubble discovered the expansion of the Universe. The discovery of a cosmic acceleration has recently renewed interest in Λ.

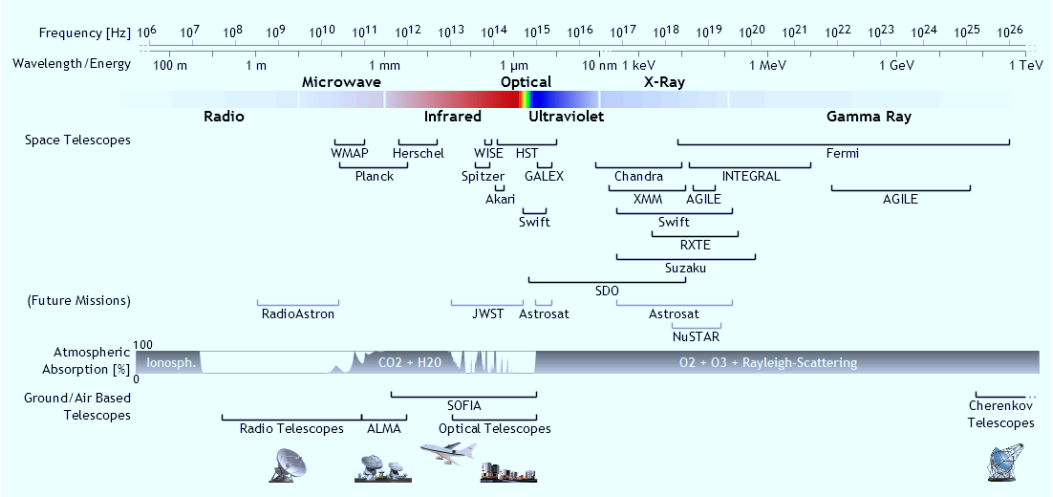

Figure 1.1: A broad view of the most important telescopes and the wavelength ranges they cover. The atmospheric absorption coefficient is also shown.

1.2.1 Galaxy morphologies

A wide variety of galaxy shapes and colors can be observed in the Universe. Hubble compiled a classification of galaxies based on morphology [Hubble, 1926]. In this scheme, galaxies are divided into *ellipticals* (E), *lenticulars* (S0), *spirals* (Sa, Sb, Sc, ...), *barred spirals* (SBa, SBb, SBc, ...) and *irregulars* (Irr). Ellipticals and lenticulars are also referred to as *early-type* galaxies, while spirals can be of either *early type* (Sa) or *late type* (Sc, ...). The morphology of galaxies tends to correlate with the *color*[4]. Early-type galaxies generally look red, contain little gas and host negligible star-formation. Late-type galaxies contain more gas and dust and host a larger amount of young blue stars. Strateva et al. [2001] noted that local galaxies exhibit a bimodal distribution in color space: a red peak is populated mainly by non-star-forming galaxies of morphological type earlier than Sa; a blue peak is populated primarly by star-forming galaxies of morphological type later than Sb. At high redshift, morphological studies are more difficult because galaxies have smaller apparent sizes.

1.2.2 Galaxy spectral energy distributions

Stars and the interstellar medium in galaxies emit radiation across the full electromagnetic spectrum. As an example, Figure 1.2 shows images of the galaxy NGC1512 observed in different bands from the ultraviolet to the infrared. The ultraviolet light reflects the emission from the youngest stars and is preferentially absorbed by in-

[4]The color of a galaxy is defined as the ratio of the flux observed in a long-wavelength band and the flux observed in a short-wavelength band.

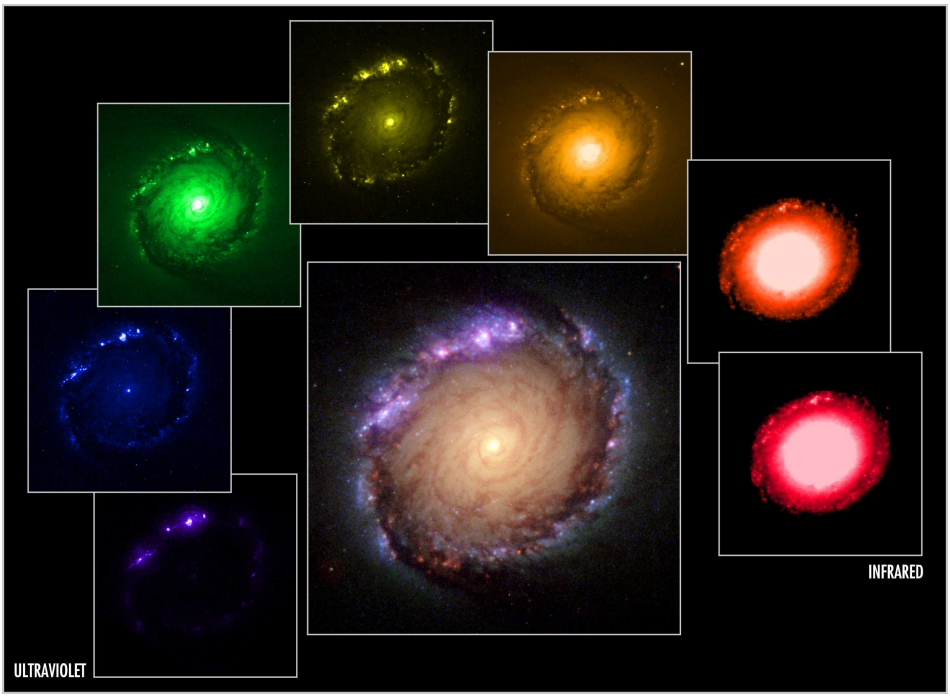

Figure 1.2: Photometric images of the galaxy NGC1512 in different bands. Clockwise from left to right: far ultraviolet ($\sim$ 2000Å), near ultraviolet ($\sim$ 3000Å), optical ($\sim$ 5000Å), Hα emission (6563Å), near infrared ($\sim$ 1μm), mid infrared ($\sim$ 10μm) and far infrared ($\sim$ 100μm). Credit: NASA.

terstellar dust. Older stars radiate mainly at longer wavelengths (near-infrared). The emission by gas is characterized by narrow emission lines (such as, Hα line at 6563Å). Dust absorbs and scatters the ultraviolet and optical light emitted by stars and gas, and reradiates this at mid- an far-infrared wavelengths (note that in Fig. 1.2 optical and infrared images are almost complementary).

The light emitted by galaxies at different wavelengths therefore contains valuable information about the properties of the stars, gas and dust within them. This light can be probed by means of different types of observations: integrating the flux through broad- and narrow-band filters (photometry), or dispersing the light in wavelength at different spectral resolutions (spectroscopy).

Figure 1.3a shows an example of spectral energy distribution of a galaxy at optical wavelengths. Overlaid are the five filter response functions, *u g r i z*, of the Sloan Digital Sky Survey (SDSS, York et al. 2000). The flux gathered through filters of this type is generally expressed in *magnitudes*. In this thesis, we always compute magnitudes in the AB system:

$$mag = -2.5 \log_{10} f_\nu - 48.6 \,, \tag{1.2.3}$$

where

$$f_\nu = \frac{\int \frac{d\nu}{h\nu} S_\nu F_\nu}{\int \frac{d\nu}{h\nu} S_\nu} \tag{1.2.4}$$

is the mean photon weighted flux through the filter, expressed in erg per second per square centimeter per hertz, and S_ν is the filter response function. Broad-band photometric observations are usually characterized by good signal-to-noise ratio. They can trace the main features of galaxy spectral energy distributions.

More refined information about galaxy spectral energy distributions, such as strong emission- and absorption-line features, can be obtained using narrow-band photometry ($\sim$ 100Å band-width). For example, the equivalent widths of strong emission lines can be estimated by combining narrow-band and broad-band observations. The equivalent width of a line is defined as[5]

$$EW = \int_{\lambda_1}^{\lambda_2} d\lambda \, \frac{F_\lambda - C_\lambda}{F_\lambda} \,, \tag{1.2.5}$$

where λ_1 and λ_2 define the wavelength range sampled by the line, F_λ is the observed flux per unit wavelength (that can be probed by narrow-band imaging) and C_λ is the flux per unit wavelength of the continuum under the line (whose mean value can be estimated from broad-band imaging). An accurate flux calibration is not crucial when measuring equivalent widths, since they involve only flux ratios.

More detailed information about the emission and absorption features in galaxy spectral energy distributions can be gathered by appealing to spectroscopy (Fig. 1.3b). A disadvantage is that spectroscopy requires far more telescope time

[5] We take the convention of positive equivalent widths for emission lines.

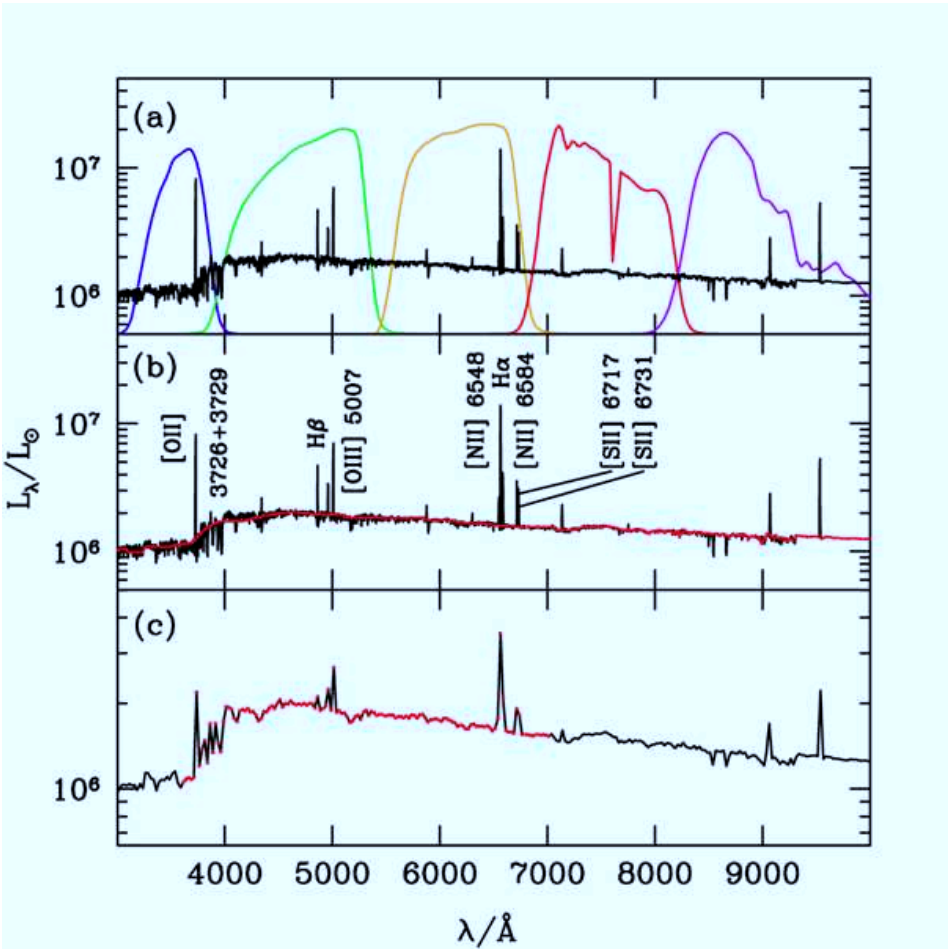

Figure 1.3: Panel (a): spectral energy distribution of a galaxy at optical wavelengths (black line) overlaid are the five filter response functions of the SDSS. Panel (b): same spectral energy distribution as in panel (a) (black line), with the strongest emission lines labeled; the red line shows the stellar continuum underneath the lines. Panel (c): the same spectral energy distribution as in panel (a), computed at a 10-times lower resolution; red points mark the binning. Note that the ordinate scale in panel (c) is enlarged.

than photometry to reach similar signal-to-noise ratio. In this contest, low-resolution spectroscopy represents an interesting trade-off between broad-band photometry and high-resolution spectroscopy. As shown in Figure 1.3c, emission and absorption features in this case are still detectable, although disentangling these features from the stellar continuum becomes problematic. A major contribution of this thesis is to develop a new approach to extract valuable information from low-resolution spectroscopic observations of galaxies.

1.3 Outline

In this thesis we investigate the relative merits of different types of observations to constrain the stars, gas and dust content of galaxies. Our main motivation is to extract the best constraints from the large amount of high-quality data that are being gathered on galaxies at various redshifts using modern telescopes, as well as to help plan for future galaxy surveys. To this aim, we develop an original approach to characterize physical properties of galaxies, e.g. stellar mass, star formation history, gas-phase metallicity, based on the combined interpretation of the stellar and nebular emission. A main feature of our approach is the high level of sophistication of the prescriptions used to build a comprehensive library of galaxy spectral energy distributions.

In Chapter 2, we introduce the different tools involved in the modeling of spectral energy distributions of galaxies. We show how physically motivated star formation and chemical enrichment histories can be derived from the semi-analytic post-treatment of a large-scale cosmological simulation. Also, we show how state-of-the-art models of stellar spectral synthesis, nebular emission and attenuation by dust can be used to describe the emission from stars and the interstellar medium in galaxies.

In Chapter 3, we present a new approach to constrain galaxy physical parameters from the combined interpretation of the stellar and nebular emission, using a comprehensive library of model spectral energy distributions. We appeal to *pseudo-observations* to assess the relative merits of photometric and spectroscopic observations, to constrain galaxy physical parameters. Then, we apply our approach to the interpretation of a sample of $\sim$13,000 high quality SDSS galaxies.

In Chapter 4, we further apply our approach to the analysis of a sample of 12 spectra extracted from an ongoing survey of high-redshift galaxies with the new wide field camera on board the Hubble Space Telescope (3D-HST, PI Dr. Pieter van Dokkum). The sample consists of star-forming galaxies at $z \sim 2$, observed in the optical rest-frame at low resolution and low signal-to-noise ratio. We compare our estimates of stellar mass and specific star formation rate with estimates derived using a photometric approach. This analysis is part of a project started in November 2011 with Dr. Elisabete da Cunha and Prof. Hans-Walter Rix, and it is to be concluded in summer 2012.

In Chapter 5, we apply our approach to the analysis of a sample of galaxies at

even higher redshifts selected among the deepest photometric observations in the Hubble Ultra Deep Field. The sample consists of 55 galaxies in the redshift range $2 < z < 8$ observed at optical and near-infrared wavelengths. We focus in particular on the correlation between the shape of the rest-frame ultraviolet spectral energy distribution and the optical depth of the dust.

In Chapter 6, we summarize this work and present our conclusions.

More detail about some technical points can be found in Appendices A and B. The photometric data used in Chapter 5 are listed in Appendix C.

Figure 1.4 shows a schematic representation of our approach.

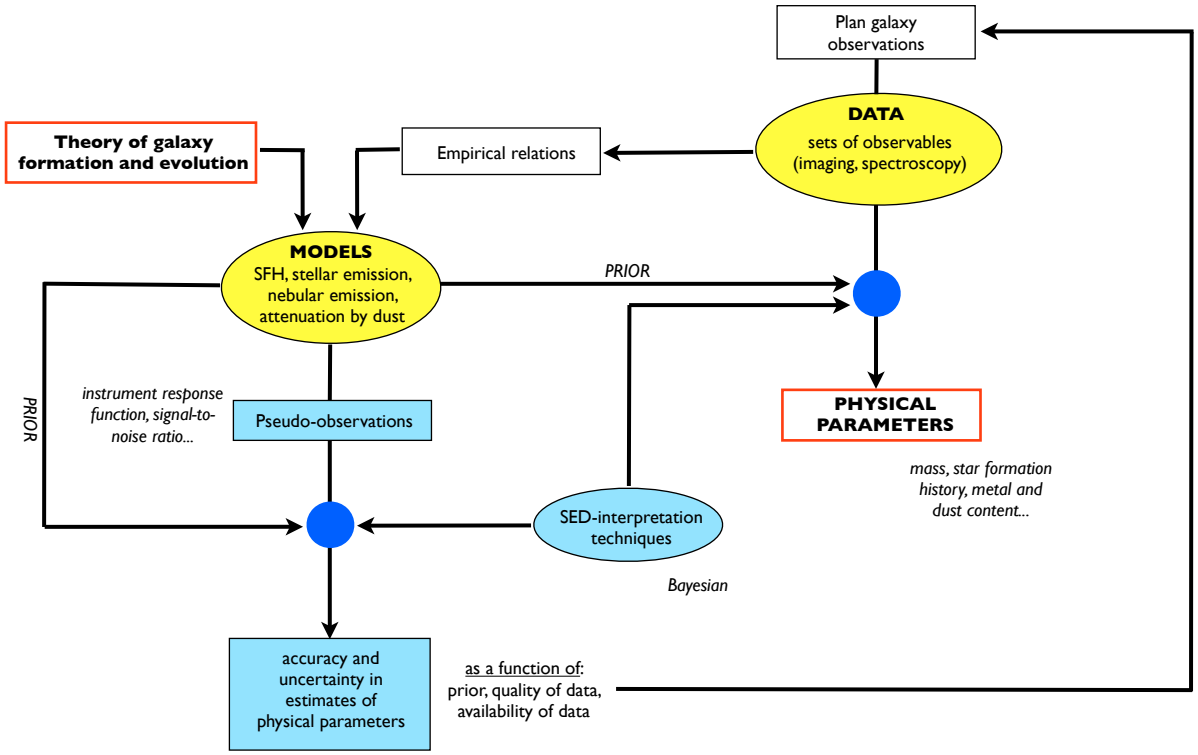

Figure 1.4: Flowchart of the general idea behind this thesis.

Chapter 2

Modeling galaxy spectral energy distributions

The modeling of the light emitted by the different constituents of galaxies (stars, gas and dust) is required to interpret multi-wavelength observations in terms of constraints on physical parameters, such as star formation rate, metallicity and dust content. In this Chapter, we introduce the different types of techniques that have been developed to model the spectral energy distributions of galaxies, with the goal to constrain the histories of star formation and chemical enrichment. In Section 2.1, we introduce different approaches to study the star formation history of galaxies. In Section 2.2, we describe models of the emission from stellar populations, and in Section 2.3, models of the emission from the gas heated by young stars. The effects of dust on the light emitted by stars and gas are addressed in Section 2.4. At the end (Section 2.5), we also mention the necessity to account for absorption by the intergalactic medium when modeling the emission from distant galaxies.

2.1 The star formation history of galaxies

The star formation history is the evolution of the rate of star formation as a function of galaxy age. The light arising from a galaxy should reflect the past episodes of star formation it underwent, through the spectral signatures of different stellar generations and the implied enrichment of the interstellar medium. Dissecting the integrated light of an individual galaxy into components of different ages is a complicated process. If one considers the galaxy population as a whole, a more direct way to gain at least some clues on the global star formation history of the Universe is to explore the star formation rate density at different cosmic epochs. This can be achieved because measurements of the current rate of star formation at any cosmic epoch through the emission from bright young stars is less challenging than interpreting the spectral signatures of old stars in the spectra of today's galaxies. To interpret, instead, the evolution of individual galaxies, requires more complicated models of the history of star formation. In this Section, we first briefly recall the conclusions that can be drawn from analyses of the global star formation rate density of the Universe at various cosmic epochs and the different sources of uncertainty affecting these analyses. Then, we describe in more detail the models that have been developed to follow the star formation and chemical enrichment histories of individual galaxies, in the framework of a hierarchical universe, and the successes and limitations of such models.

2.1.1 The star formation history of the Universe

The study of the star formation history of galaxies through the assessment of the star formation rate density at different epochs has been pioneered by Lilly et al. [1996] and Madau et al. [1996], combining observations from the local Universe to $z \sim 3-4$. Hopkins [2004] and Hopkins and Beacom [2006] have improved the statistics out to $z \sim 6$ and have analyzed the various sources of uncertainties that can contribute to the normalization of the star formation rate density. More recently, Cucciati et al. [2011] have investigated the star formation rate density from $z \sim 0.05$ to $z \sim 4.5$ using data from a single galaxy redshift survey (VIMOS-VLT Deep Survey, VVDS, Garilli et al. 2008, Le Fèvre et al. 2004, Le Fèvre et al. 2005), to avoid merging different datasets. In their work, they assess also the various sources of uncertainties, and in particular, the treatment of dust attenuation.

In Figures 2.1 and 2.2, respectively, we present the original and the most recent determinations of the evolution of the star formation rate density of the Universe as a function of redshift. Despite the inhomogeneity of the galaxy samples used in these studies and the differences in the adopted star-formation-rate indicators (far ultraviolet, far infrared, Hα emission line luminosity, radio observations), the Universe appears to have undergone a phase of most active star formation around $z \sim 1-2$. Other uncertainties affect this measurement, leading to large error bars. For example, the correction for attenuation by dust at ultraviolet and optical wavelengths and the interpretation of dust emission at infrared wavelengths can alter estimates of the star formation rate. Moreover, the lack of statistics beyond redshift of 2–3 precludes the interpretation of the decline of star-formation at cosmic epochs earlier than $z \sim 1$.

Although the evolution of the star formation rate density gives us important clues about the global history of star formation of the Universe, the interpretation of the results of Figures 2.1 and 2.2 is very limited. To gain more insight into the evolution of different types of galaxies, we thus need to investigate in more detail the star formation history of individual galaxies.

2.1.2 The star formation history of individual galaxies

Constraining the star formation history of individual galaxies requires models that can describe the spectral evolution implied by different scenarios of star formation and chemical enrichment. By comparing such models with observations, one can infer the most likely scenario for the evolution of individual galaxies. In practice, even idealized representations of the star formation history can reproduce reasonably well the colors of nearby galaxies. Here, we first briefly describe different idealized representations of the star formation history of individual galaxies. We then describe more sophisticated approaches based on detailed cosmological simulations.

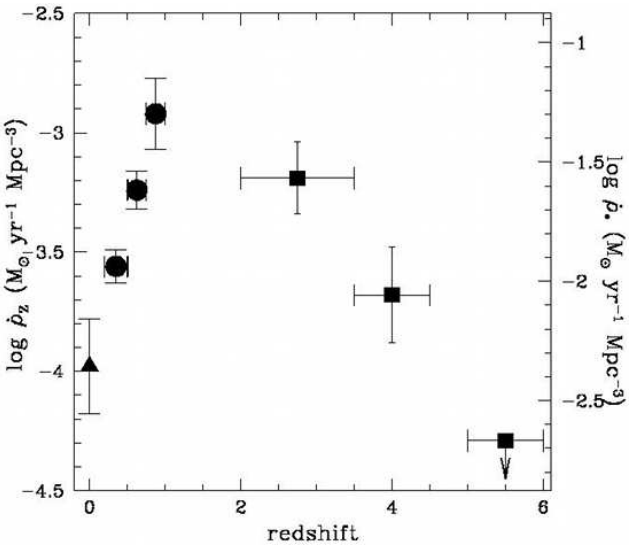

Figure 2.1: Fig. 9 from Madau et al. [1996], showing on the left the Universal metal ejection density, $\dot{\rho}_Z$, and on the right the total star formation rate density, $\dot{\rho}_*$, as a function of redshift. Triangle: Gallego et al. [1995]. Filled dots: Lilly et al. [1996]. Filled squares: lower limits from Hubble Deep Field images.

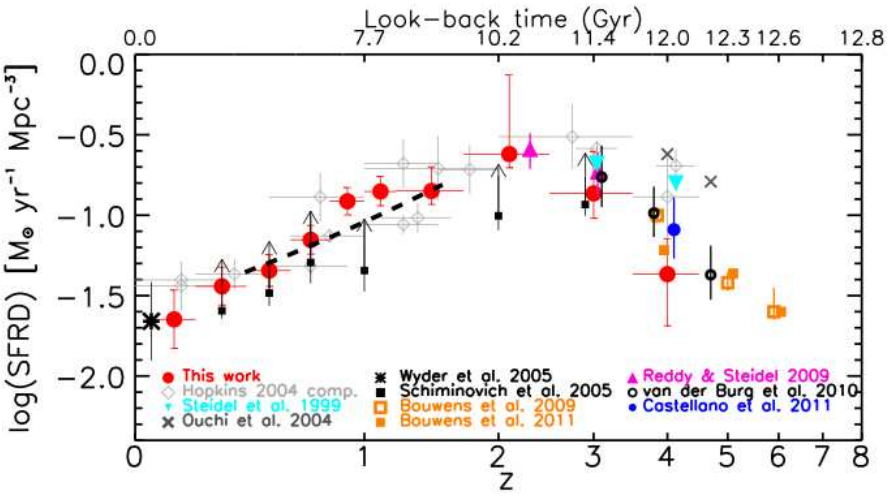

Figure 2.2: Fig. 5 from Cucciati et al. [2011], showing the total dust-corrected UV-derived star formation rate density as a function of redshift from the VVDS sample (red filled circles). The black dashed line is the star formation rate density as a function of z implied from the stellar mass density in Ilbert et al. [2010]. Other data sets are overplotted and labeled on the figure. All data have been converted into a star formation rate density with the scaling relation from Madau et al. [1998].

2.1.2.1 Simple idealized models of star formation history

Galaxy star formation histories have been traditionally approximated by simple analytic functions: single bursts of star formation, constant star formation rate, exponentially declining star formation rate or a combination of these three. Figure 2.3 shows 3 templates of star formation histories and the corresponding spectral energy distributions, computed using the latest version of Bruzual and Charlot [2003] for fixed solar metallicity, with a Chabrier initial mass function [Chabrier, 2003]. Emission lines are computed consistently following the prescription by Charlot and Longhetti [2001]. Attenuation by dust is neglected in this example. A constant star formation history (panel a) gives a relatively blue spectral energy distribution (panel e). A single burst of star formation at early times (panel b) does not include the emission from any young stars, resulting in a fairly red spectral energy distribution (panel e). In panel (c), we show an exponentially declining star formation history, with e-folding time 2.5 Gyr, and the corresponding spectral energy distribution is shown in panel (f).

Previous studies (e.g. Kauffmann et al. 2003a, Brinchmann et al. 2004, da Cunha et al. 2008) have shown that these simplistic formulations of star formation histories are able to reproduce the colors of low-redshift galaxies. For example, Kauffmann et al. [2003a] show that exponentially declining star formation histories, with random superimposed bursts of star formation, reproduce reasonably well the colors of a representative sample of galaxies from the SDSS at redshift 0.03 and 0.11, once corrected for dust attenuation (see Figure 2.4). We note that these authors also correct the SDSS colors for the contamination by emission lines. We will return in Chapter 3 (Figure 3.7) on the non-negligible influence of nebular emission lines when interpreting the observed colors of star-forming galaxies.

Wuyts et al. [2009] specifically investigate the limitations arising from the interpretation of complex star formation histories of galaxies using idealized representations of the type shown in Figure 2.3. Their approach consists of 2 steps: 1 – building 'realistic' galaxy spectral energy distributions based on star formation histories derived from a sophisticated simulation of galaxy formation (they adopt a simulation based on a smoothed-particle-hydrodynamic code, see Section 2.1.2.2); 2 – comparing the colors of these galaxies with colors of model galaxies computed using simplistic star formation histories. They consider the spectral evolution over a period of only 2 Gyr, as they focus on studies of high-redshift galaxies. Wuyts et al. [2009] show that the spectral energy distributions of galaxies derived from a cosmological simulation are better represented by exponentially declining star formation histories with e-folding time of 300 Myr than by single-burst and constant star formation histories. They also find that, in the case of starburst galaxies, age and mass are systematically underestimated when adopting simple exponentially declining star formation histories. They show that models allowing for secondary bursts of star formation on top of an exponentially declining star formation history allow for larger total stellar masses, providing a better match for blue objects than without secondary bursts (see also Papovich et al. 2006, Erb et al. 2006, Wuyts

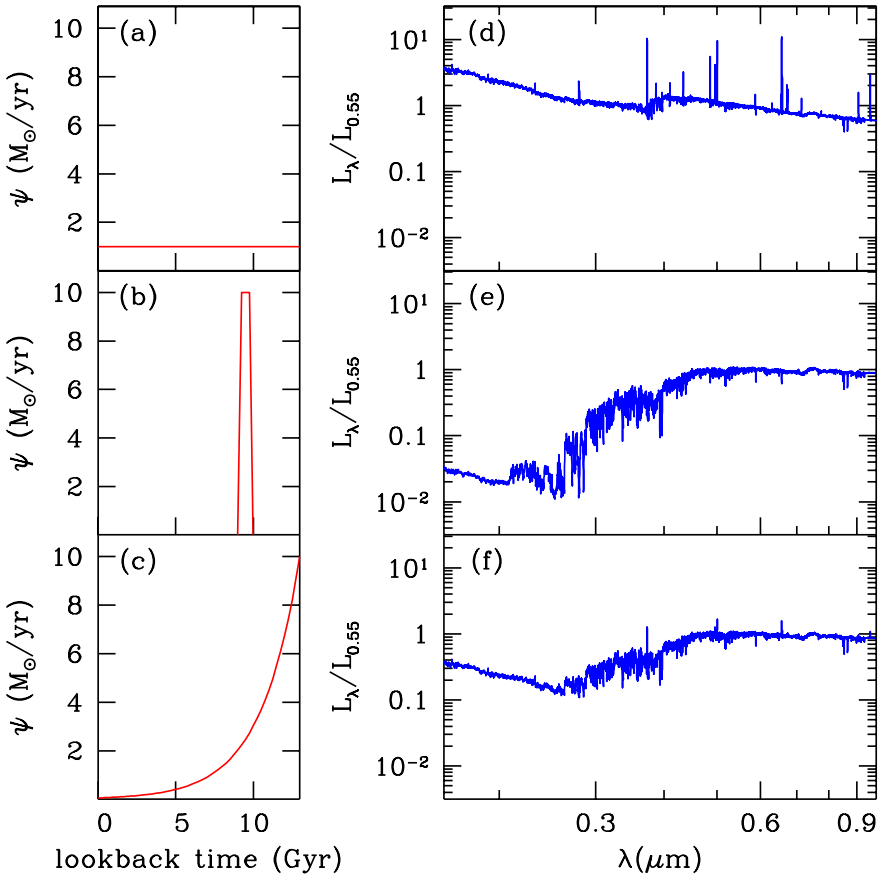

Figure 2.3: Idealized star formation histories (panels a, b, c) and the relative spectral energy distributions (panels d, e, f). The model spectral energy distributions are computed using the latest version of Bruzual and Charlot [2003] at fixed solar metallicity, with a Chabrier initial mass function [Chabrier, 2003]. Emission lines are computed consistently following the prescription by Charlot and Longhetti [2001]. The attenuation by dust is neglected in these examples.

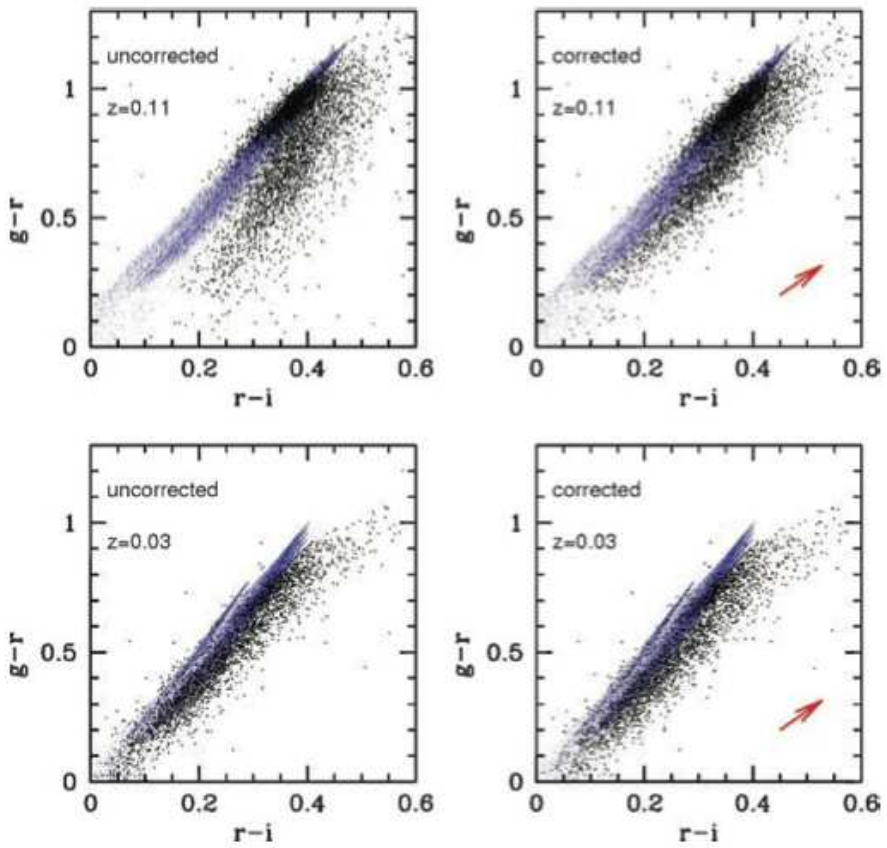

Figure 2.4: Fig. 10 from Kauffmann et al. [2003a]. Observed $g-r$ versus $r-i$ colors of a representative sample of SDSS galaxies (black points) for 2 particular redshifts (0.11 and 0.03). Blue points represent the model grid, computed applying Bruzual and Charlot [2003] models to exponentially declining star formation histories, with random superimposed burst of star formation. In the left panels, data are not corrected for dust attenuation, while in the right panels, data are corrected with an attenuation law of the form $\tau_\lambda \propto \lambda^{-0.7}$. The red arrow shows the predicted reddening vector.

et al. 2007). In any case, the results of Wuyts et al. [2009] suggest that sophisticated simulations of galaxy formation can help us better constrain the history of star formation in galaxies than simple idealized models.

2.1.2.2 History of star formation in a hierarchical universe

In recent years, sophisticated star formation and metal enrichment histories of galaxies have been modeled by appealing to the combination of cosmological N-body simulations (to trace the evolution of dark matter haloes) with semi-analytical recipes or gas-dynamics codes (to simulate the behavior of baryons in the dark-matter structures). In this Section, we first present a few simulations of the evolution of dark-matter particles. Then, we describe possible treatments for the baryonic component. For simplicity, we analyze these two components separately, although often single codes follow their evolution simultaneously.

Cosmological N-body simulations rely on the standard cosmological scenario that structures grow from weak density fluctuations present in the otherwise homogeneous and rapidly expanding early Universe. These fluctuations are amplified by gravity, turning into the structures we observe today.[1] The evolution of these fluctuations is a highly non-linear process, which can be assessed most readily through numerical simulations. In this framework, cold dark matter is assumed to be made of elementary particles which interact only gravitationally. The representation of this fluid as an N-body system is a good approximation, which improves as the size of the simulation increases (in number of particles and size of the simulated box).

The first N-body simulation used 300 particles [Peebles, 1970]. Cosmological simulations have since improved in size and resolution. Cosmological simulations built to study the large-scale structure of the Universe tend to favor large size and low resolution. In contrast, simulations designed to study galaxy formation tend to favor resolution over size. One of the most widely used simulations today is the Millennium Simulation [Springel et al., 2005]. It is one of the largest existing simulations, with sufficient mass resolution to obtain good statistical samples of rare objects, such as massive cluster haloes. The possibility to sample large volumes is required to build mock catalogues for future galaxy surveys. Kim et al. [2009] hold so far the record of the largest simulation (the Horizon Run) with 69.9 billion particles and a box of 6.6 Gpc on a side, but with low mass resolution compared to the Millennium. Smaller volumes with higher mass resolution (much more computationally demanding) can improve the understanding of galaxy formation tracing the evolution of low-mass galaxies. An example which fulfills this requirement is the Millennium II simulation [Boylan-Kolchin et al., 2009]. Among the other N-body cosmological simulations that have been run more recently, we can cite the Horizon simulation by Teyssier et al. [2009] and the most recent Bolshoi simulation by

[1] The validity of this model is supported by the observation of the cosmic microwave background with WMAP (Bennett et al. 2003, Spergel et al. 2003) and PLANCK [Lamarre et al., 2003], along with the observation of the apparent acceleration of the cosmic expansion inferred from studies of distant supernovae (Riess et al. 1998, Perlmutter et al. 1999).

Simulation and reference	Size (h^{-1} Mpc)	N of part.	Part M (h^{-1} $M_\odot$)	z	N steps	Cosmological parameters (Ω_m	Ω_Λ	h	σ_8)
[Gottlöber et al., 2001]	60	1.7×10^7	1.1×10^9	10	25	0.30	0.70	0.70	1.00
[Hatton et al., 2003]	150	1.7×10^7	5.5×10^8	35.59	100	0.33	0.67	0.67	0.88
Millennium Simulation [Springel et al., 2005]	500	$\sim 10^{10}$	8.6×10^8	127	64	0.25	0.75	0.73	0.90
Millennium-II [Boylan-Kolchin et al., 2009]	100	$\sim 10^{10}$	6.9×10^6	127	68	0.25	0.75	0.73	0.90
Horizon Simulation [Teyssier et al., 2009]	2000	6.87×10^{10}	7.7×10^9	40	737	0.24	0.76	0.73	0.77
Horizon Run [Kim et al., 2009]	6592	6.99×10^{10}	2.96×10^{11}	23	6	0.26	0.74	0.72	0.77
Bolshoi [Klypin et al., 2011]	250	8.0×10^9	1.35×10^8	80	180	0.27	0.73	0.70	0.82

Table 2.1: Basic characteristics of the most recent N-body cosmological simulations. We cite Gottlöber et al. [2001], Hatton et al. [2003], the Millennium Simulation by Springel et al. [2005], the Millennium-II simulation by Boylan-Kolchin et al. [2009], the Horizon simulation by Teyssier et al. [2009], the Horizon Run by Kim et al. [2009] and the most recent Bolshoi simulation by Klypin et al. [2011]. For these simulations we report: the size of the box in h^{-1} Mpc, the number of dark matter particles, the mass of the single particle in h^{-1} solar masses, the redshift where the initial conditions are created, the number of timesteps for which data are saved, and the cosmological parameters (omega matter, Ω_m, omega lambda, Ω_Λ, the Hubble constant at redshift 0 in units of 100 km s^{-1} Mpc^{-1}, h, and the root mean squared amplitude of linear mass fluctuations in $8h^{-1}$ Mpc spheres at $z = 0$, σ_8).

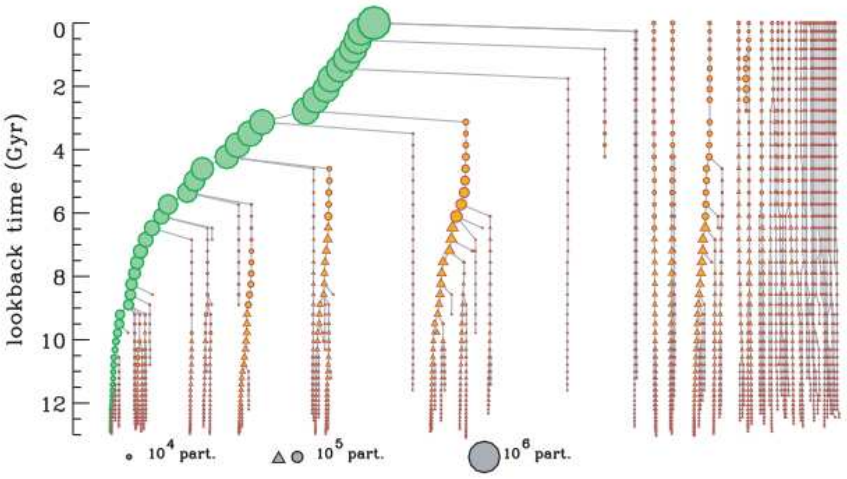

Figure 2.5: Fig. 2 from De Lucia and Blaizot [2007]. The merger tree of a dark matter halo of $\sim 9 \times 10^{14}\ M_{\odot}$. The area of the symbols scales with the mass. Circles represent haloes that are part of the friends-of-friends group, while triangles show haloes that have not joined yet the friends-of-friends group. The halo which contains the main branch is marked in green. The trees on the right-hand side, which are not linked to the main halo, correspond to the other substructures identified in the friends-of-friends group at $z = 0$.

Klypin et al. [2011]. These simulations were all run in the framework of a ΛCDM Universe, with slightly different parameter variants. Their main characteristics are given in Table 2.1.

The formation and evolution of galaxies in the framework of dark-matter N-body simulations is dictated by the merger history of dark-matter haloes, which baryons follow. We show in Figure 2.5 an example of merger tree of a dark-matter halo computed with the Millennium Simulation (from Fig. 2 of De Lucia and Blaizot [2007]). In this example, the halo has a mass of about $9 \times 10^{14}\ M_{\odot}$ at $z = 0$ and lies on top of the plot, with all the progenitors plotted downward as a function of lookback time. The green branch represents the *main branch*, which follows the evolution of the *main progenitors* of the final halo. In this framework, the evolution of the halo can be drawn as a series of accretion events. The other progenitors are shown in orange: circles indicate haloes which belong to the same *friends-of-friends*[2] group, while triangles mark haloes which have not yet joined the friends-of-friends group. The trees on the right-hand side, which are not connected to the main branch, are substructures which have not merged into the main halo.

[2] *friends-of-friends* is a simple algorithm [Davis et al., 1985], which links two particles in the same group if their distance is less than a certain linking length. This distance can be also expressed as a density and in general friends-of-friends algorithms pick up structures with density ≈ 200 times higher than the mean density of the system.

Within the hierarchy of haloes produced by dark-matter N-body simulations, the evolution of the baryons can be followed using different approaches: semi-analytic models (SAM), smoothed-particle-hydrodynamic techniques (SPH), adaptive mesh refinement methods (AMR) and the latest moving-mesh code AREPO. We briefly summarize below the main characteristics of these 4 approaches (in Chapter 3 in this thesis we focus on the first one).

Semi-analytic models (see for example Bower et al. 2006 and references therein) are a simple and powerful approach to describe the evolution of baryons in a cosmological context, without computing the hydrodynamic evolution of the gas. These models can be performed through a post-treatment of dark-matter N-body simulations, thus they are computationally efficient. The design of semi-analytic models requires the specification of a few key parameters, defined using empirical relations (feedback from supernovae, AGN, stellar winds). It is worth pausing here to describe in slightly more detail the characteristics of the Croton et al. [2006] model, which we will be using in Chapter 3.

Croton et al. [2006] (see also De Lucia and Blaizot 2007) propose an approach to study the formation and evolution of galaxies in the framework of the Millennium N-body Simulation [Springel et al., 2005, see Table 2.1]. The full dataset from the dark-matter simulation is stored in 60 time steps between $z = 20$ and $z = 0$ and four more steps at $z = 30, 50, 80$ and 127. The merger trees do not have to be re-run to explore the parameter space. At $z \sim 0$ the time resolution is of approximately 300 Myr. The code produces a friends-of-friends group catalogue, saving only haloes with at least 20 particles ($\sim 1.7 \times 10^{10}$ $M_\odot$). At $z = 0$ this procedure identifies 17.7×10^6 friends-of-friends groups, down from a maximum of 19.8×10^6 at $z = 1.4$, when groups are more abundant but of lower mass on average. Once the resulting merging history is stored, it is possible to perform the semi-analytic treatment. Croton et al. [2006] follow the standard paradigm by White and Frenk [1991], as adapted for implementation into high-resolution N-body simulations by Springel et al. [2001] and De Lucia et al. [2004]. This assumes that when a dark-matter halo collapses, the baryons in it collapse as well. Baryons are initially in the form of diffuse gas of primordial composition. They transform into stars and heavy elements during the evolution of the halo. When the gas collapses on to a central object (which is assumed to be a cold gas disk), episodes of star formation, either quiescent or in a burst, are ignited. The prescription for quiescent star formation is based on Kennicutt [1998], while starbursts arise in merger events, between the central galaxy of a halo and a satellite galaxy. When the mass of the satellite is small compared to the central galaxy, the event is called *minor merger* and the stars of the satellite are added to the bulge of the central galaxy, causing a minor starburst. When the two masses are comparable, the starburst is more significant, with the merger destroying the discs of both galaxies to form a spheroid in which all stars are placed. Croton et al. [2006] find that galaxies are better represented when accretion and cooling are inefficient. Active galactic nuclei and supernova feedback have an important role in suppressing cooling flows. In particular, supernova events inject gas, metals and energy into the surrounding medium, reheating cold-disc gas and possibly ejecting

gas from the surrounding halo. This process prevents the gas from cooling down and also contributes to the metal enrichment of the gas.

In contrast to the semi-analytic model described above, the other approaches (smoothed particle hydrodynamics, adaptive mesh refinement and moving mesh) run in parallel to N-body dark-matter simulations. Smoothed particle hydrodynamics methods (see for example Katz et al. 1996, Springel et al. 2001) are called *Lagrangian* methods. They discretize mass, using a set of fluid particles to model the flow. They are particularly suited to follow the gravitational growth of structures and resolution is automatically increased in the central regions of galactic haloes, with no discontinuous jumps. The main disadvantage is that they appear to suppress fluid instabilities and thus they have to rely on artificial viscosity.

Adaptive-mesh-refinement codes follow an *Eulerian* approach (see for example Cen and Ostriker 1992, Teyssier 2002). In this approach, space is discretized and fluid variables are represented on a mesh. The building blocks of the mesh are rectangular patches of various sizes. The position and dimension of the patches depend on speed and memory constraints. AMR codes are faster than smoothed particle hydrodynamics codes and do not need artificial viscosity. Usually, individual galaxies are poorly resolved and only a very fine mesh can allow the code to reach the resolving power of an SPH code. Moreover, the results may change when a velocity field is applied to the system (no Galilean invariance).

The approach proposed by Springel [2010] combines the advantages of Lagrangian and Eulerian methods by letting the mesh itself move. This is the basic principle of the new code AREPO. The mesh is defined as a set of discrete mesh-generating points, which move freely. AREPO allows the continuous adjustment of resolution and it is Galilean invariant when the mesh is moved along with the flow. It inherits the geometric flexibility of smoothed-particle-hydrodynamic codes and also the absence of artificial viscosity of adaptive-mesh-refinement codes.

Scannapieco et al. [2011] show that when SPH, AMR or AREPO codes are run on the same initial conditions, the differences in the codes themselves and the different treatments for cooling, star formation and feedback cause large variations in the resulting stellar mass, size, morphology and gas content of the galaxies at the end of the evolution. For example, the stellar mass in AREPO and in adaptive mesh refinement codes is typically twice as high as that in smoothed particle hydrodynamics simulations. All three codes tend to produce too many massive galaxies compared to theoretical expectations and also to observations. The peak of star formation is predicted to arise at redshifts around $z \sim 4$ with basically insignificant star formation in recent times. This as well does not match with observations (see Section 2.1.1). The codes differ also in the gas fraction at $z = 0$: AREPO predicts a larger gas fraction compared to smoothed particle hydrodynamics methods, while adaptive-mesh-refinement codes tend to predict gas fractions intermediate between these two.

A general limitation of all the approaches mentioned above is that they tend to produce a too large fraction of quiescent galaxies, along with too small spirals (see for example Weinmann et al. 2006, Guo et al. 2010).

2.2 Emission from stellar populations

In this Section, we describe the techniques used to model the emission from the primary component of galaxies: stellar populations. Stellar population synthesis models have been developed for over 40 years to interpret the light from galaxies in terms of physical parameters. The first attempts to model and interpret spectra of galaxies relied on *trial and error* analyses ([Spinrad and Taylor, 1971, Faber, 1972, O'connell, 1976, Turnrose, 1976, Pritchet, 1977]). In this technique, one reproduces the spectral energy distribution of an observed galaxy with a linear combination of individual stellar spectra of various spectral types and luminosity classes taken from a comprehensive library. The solution is found by minimizing numerically the difference between models and data and by invoking additional constraints such as positive star numbers, increasing number of stars with decreasing mass on the main sequence and consistent numbers of evolved red giant stars and main-sequence progenitors. This technique was abandoned in the early 80's because the number of free parameters was too large to be constrained by typical galaxy spectra. More recent models are based on the *evolutionary population synthesis* technique (see Guiderdoni and Rocca-Volmerange 1987, Bruzual and Charlot 1993, Leitherer and Heckman 1995, Maraston 1998, and references therein). This technique is based on the property that stellar populations can be expanded in series of instantaneous starbursts, called *simple stellar populations* (SSPs). The main adjustable parameters are the stellar initial mass function, the star formation rate and the rate of chemical enrichment as a function of time.

We describe below the main ingredients required to perform this modeling: 1 – a prescription for the mass distribution of newly born stars; 2 – a theory for the evolution of stars of given initial mass and chemical composition; 3 – a library of stellar spectral energy distributions to describe the emission from any single star.

2.2.1 Stellar initial mass function

Stars of different initial mass evolve differently, thus the observed properties of a simple stellar population at any time depend on the stellar initial mass function. Salpeter [1955] proposes a simple parametrization of the initial mass function from counts of main-sequence stars in the solar neighborhood. This representation is a single power law which has long been considered universal [$\phi(m) = dN/dm \propto m^{-(1+x)}$, with $x \approx 1.35$]. Only later in the Seventies, a few studies started to reveal discrepancies from this simple form when observing different samples of stars (for example from globular clusters or the Magellanic Clouds) and improving the statistics for very low-mass stars. Scalo [2005] reviews different prescriptions for the initial mass function and suggests that the main problems in the computation of the initial mass function arise from the possible incompleteness of the data, the uncertainties in the evolution of massive stars and the corrections of star counts for extinction by dust.

The currently most popular analytic representations of the stellar initial mass

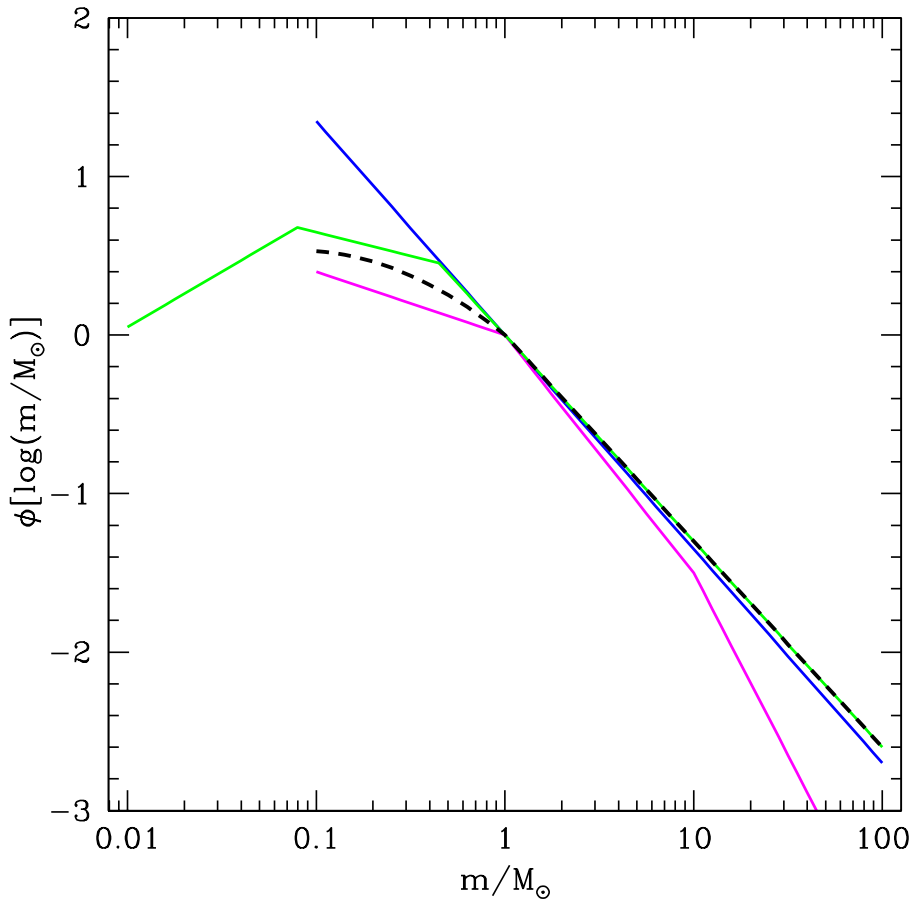

Figure 2.6: Stellar initial mass faction normalized to solar units. Salpeter [1955], blue line. Miller and Scalo [1979], magenta line. Kroupa [2001], green line. Chabrier [2003], dashed line.

function are:

- Salpeter [1955]
- Miller and Scalo [1979]
- Kroupa [2001]
- Chabrier [2003].

We plot the four functions in Figure 2.6, all normalized to the same value at 1 $M_\odot$.

The initial mass function by Salpeter [1955] (blue line) is expressed as

$$\phi[\log(m)] \propto m^{-1.35} \tag{2.2.1}$$

In Figure 2.6, the Miller and Scalo [1979] stellar initial mass function (magenta line) is defined as

$$\phi[\log(m)] \propto \begin{cases} m^{-0.4} & \text{for } 0.1 \leqslant m/M_{\odot} < 1.0 \\ m^{-1.5} & \text{for } 1.0 \leqslant m/M_{\odot} < 10 \\ m^{-2.3} & \text{for } 10 \leqslant m/M_{\odot}\,. \end{cases} \tag{2.2.2}$$

In the stellar initial mass functions favored by Kroupa [2001] and Chabrier [2003] the high-mass slope (stars more massive than 1 $M_{\odot}$) is similar to the one found by Salpeter [1955]. The improved observational constrains in low-mass stars allowed a better characterization of the stellar initial mass function for $m < 1M_{\odot}$. The Kroupa [2001] stellar initial mass function (green line in Figure 2.6) is defined by

$$\phi[\log(m)] \propto \begin{cases} m^{0.7} & \text{for } 0.01 \leqslant m/M_{\odot} < 0.08 \\ m^{-0.3} & \text{for } 0.08 \leqslant m/M_{\odot} < 0.50 \\ m^{-1.3} & \text{for } 0.50 \leqslant m/M_{\odot}\,, \end{cases} \tag{2.2.3}$$

while the Chabrier [2003] initial mass function (dashed line) is defined by

$$\phi[\log(m)] \propto \begin{cases} \exp\left[-\frac{(\log m - \log m_{\rm c})^2}{2\sigma^2}\right] & \text{for } 0.1 \leqslant m/M_{\odot} < 1.0 \\ m^{-1.3} & \text{for } 1.0 \leqslant m/M_{\odot}\,, \end{cases} \tag{2.2.4}$$

where $m_{\rm c} = 0.08$ and $\sigma = 0.69$.

The shape of the low-mass end of the stellar initial mass function is extremely important. Even if these stars are very faint, they contribute significantly to the mass of a galaxy. Figure 4 of Bruzual and Charlot [2003] shows the differences in the evolution of the colors and mass-to-light ratios of simple stellar populations (see Section 2.2.4) computed for different initial mass functions. Basically, a stellar initial mass function which predicts a higher number of stars at low mass [Salpeter, 1955] causes systematically higher mass-to-light ratio, compared to the other forms. In all applications in this thesis, we always adopt the Chabrier [2003] initial mass function. We must also define lower and upper mass limits. The lower mass limit usually corresponds to the minimum mass required to ignite hydrogen burning in the stellar core, i.e. about $0.1M_{\odot}$ (we note that the stellar initial mass function can be extended down to lower masses to include brown dwarfs; e.g. Kroupa [2001]). The upper limit is less well defined, with a conventional value around 100 $M_{\odot}$. Roughly, above this value, a star cannot radiate energy fast enough to remain stable.

2.2.2 Stellar evolutionary tracks

Stars of different initial masses burn their fuel at different rates, thus they evolve through different paths in the Hertzsprung-Russell (H-R) diagram. In this diagram, shown in Figure 2.7, luminosity is plotted as a function of effective temperature of a star. The position of a star in the H-R diagram depends on chemical composition, mass and age. A star is on the zero-age main sequence (black line) when thermonuclear reactions have begun in the core and it has reached hydrostatic equilibrium.

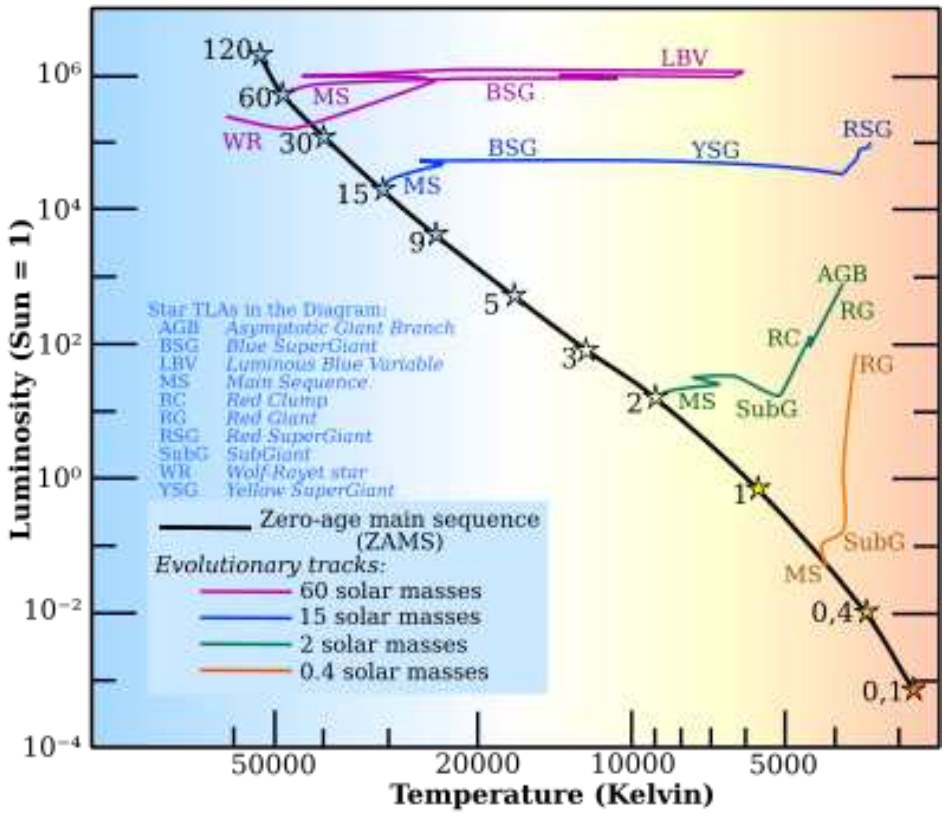

Figure 2.7: The Hertzsprung-Russell diagram (H-R) shows the evolutionary stages of stars in terms of their luminosity and temperature. Different evolutionary paths for stars of different initial masses are plotted in different colors. The zero-age main sequence is shown as a black solid line. The labels indicate the main stages of the evolution.

A main-sequence star burns hydrogen in the core, converting it into helium. The higher the initial mass of the star, the higher the energy required to remain in hydrostatic equilibrium, balancing the gravitational force. The higher the energy produced, the brighter the star, which, in this way, burns all the hydrogen in the core faster than a low-initial-mass star. As a consequence, the lifetime of a star with high initial mass is shorter than that of a low-initial-mass star.[3] We can summarize the different evolutionary paths of stars of different initial mass as follows (the numerical values below correspond to stars with solar metal composition).

Objects with mass $m < 0.08\ M_\odot$ never reach the temperature required to burn hydrogen in their core. Such objects slowly contract until free electrons

[3] $L \propto E/t$, where L is the luminosity of the star, E is energy and t is time. Inverting this relation, $t \propto E/L \propto M/L \propto M/M^{3.5} \propto M^{-2.5}$, where M is the mass of the star. $L \propto M^{3.5}$ is an empirical relation which applies to main-sequence stars.

become degenerate and the surface temperature reaches an equilibrium with the surrounding interstellar radiation field (*brown dwarfs*).

Very low-mass stars ($0.08\ M_\odot < m < 0.30\ M_\odot$) can ignite hydrogen-burning processes, but the core temperature always remains below the temperature required for helium ignition. A star of this type never leaves the main sequence, it becomes degenerate and cools down after the end of the hydrogen-burning phase, becoming a *white dwarf*.

Low-mass stars ($0.3\ M_\odot < m < 2.0\ M_\odot$) leave the main sequence when hydrogen in their core is completely exhausted. The helium core contracts, converting gravitational potential energy into thermal energy. Then, hydrogen starts burning in a shell around the core, causing the outer layers of the star to expand. The radius gets larger and the surface temperature lower. These stars are called *red giants*. The helium core increases in mass because of the infall of helium produced in the outer shell. At this point, electrons in the core become degenerate, the core contracts and the star starts burning helium. This increases the temperature in the core, enhancing the rate of helium-burning reactions (helium flash). The temperature in the core keeps increasing until the gas becomes non-degenerate and can cool down. The outer layers contract and the star steadily burns helium in the core and hydrogen in the outer shell. The star settles down in this regime until all the helium in the core has been converted into carbon and oxygen. At this stage, helium starts burning in an outer shell and the star expands evolving toward the *asymptotic giant branch* (AGB, green track on Figure 2.7). This phase can be broken into two parts: early AGB (E-AGB) and thermally pulsing AGB (TP-AGB) [Iben and Renzini, 1983]. In the early-AGB phase, the hydrogen-burning shell is extinct, and the helium-burning shell accounts for most of the emission by the star. The duration of this phase is roughly 10 Myr, although this depends on mass and composition of the star. When helium in the inner shell is nearly exhausted, hydrogen is reignited in a thin outer shell and the star begins to thermally pulse. Expansion is caused by the energy liberated by violent burning processes occurring in the helium shells. This happens because helium is highly sensitive to temperature changes. The star contracts when helium cools down. The duration of the thermally-pulsing AGB phase is essentially determined by the mass-loss rate and ranges between 0.2 and 2 Myr [Vassiliadis and Wood, 1993, Table 1]. Eventually, the strong stellar winds provide enough energy for the outer layers to be ejected, forming a planetary nebula. When helium in the shell is also exhausted, the envelope collapses and the electron gas in the core becomes degenerate. At this point, the star never reaches high enough temperatures to ignite carbon and thus becomes a white dwarf.

Intermediate-mass stars ($2\ M_\odot < m < 8\ M_\odot$) evolve like low-mass stars but have sufficient mass for quiet helium ignition in their core. Helium burns at a higher rate compared to low-mass stars, and the evolution is faster. An intermediate-mass star evolves towards the asymptotic giant branch and passes the thermally-pulsing

AGB phase, ejecting the outer envelop in the interstellar medium (planetary nebula). The lower and upper mass limits of this category of stars strongly depend on metallicity and other model details. This issue is discussed by Marigo [2001], who suggests that the maximum initial mass for a star to develop a degenerate helium-core (helium-flash phase) is comprised between 1.7 and 2.5 $M_{\odot}$. The upper limit instead is defined as the critical stellar mass over which carbon ignition occurs and is comprised between 5 and 8 $M_{\odot}$. It is interesting to note that, without mass loss, the mass of the carbon-oxygen core in these stars would easily reach ~ 1.4 $M_{\odot}$, the effective Chandrasekhar limit, at which point carbon ignition (carbon-flash phase) would lead to a supernova explosion. The rate of supernova explosions in our Galaxy is much smaller than the amount of intermediate-mass stars [Iben, 1991], thus it is evident that not all intermediate-mass stars evolve to a supernova stage. If the temperature in the core is too low to undergo a carbon flash, the star loses its outer envelop and cools down as a white dwarf.

High-mass stars ($m > 8$ $M_{\odot}$) are extremely bright and hot on the zero-age main sequence. Hydrogen burns very fast. As soon as it is completely converted into helium in the core, hydrogen keeps burning in a shell and the temperature starts decreasing (blue track on Figure 2.7). Unlike for low- and intermediate-mass stars, helium is ignited in a non-degenerate core and hydrogen burns in an outer shell. A star of this type becomes a *blue supergiant*. When helium is exhausted in the core, the temperature and density are high enough to start quiet carbon ignition, and the star goes through all nuclear burning phases until an iron core is formed. A star with mass higher than 40 $M_{\odot}$ loses mass at a high rate (*Wolf-Rayet* star, see purple track on Figure 2.7). A high-mass star ends its life when the iron core collapses to form a neutron star or a black hole, after a supernova explosion has ejected the outer shells into the interstellar medium.

The fate of a star depends not only on initial mass, but also on the chemical composition of the gas from which it formed. Figure 2.8 [Heger et al., 2003] summarizes the evolution of stars in different mass groups at solar (*top panel*) and zero (*bottom panel*) metallicity. The blue line represents the mass of the remnant after the helium-burning phase. The departure from the black dotted line shows the mass loss caused by stellar winds. The red line shows the mass of the remnant after the asymptotic-giant-branch phase (for low- and intermediate-mass stars) or the supernova explosion (for high-mass stars). The mass of metals returned to the interstellar medium is highlighted in green. Zero-metallicity stars can reach higher initial masses than solar-metallicity stars and loose less mass. Their evolution is typically very fast compared to that of solar-metallicity stars.

Hence, we have seen that stars with different initial masses follow different evolutionary tracks in the Hertzsprung-Russell diagram, more massive stars evolving faster than low-mass stars. For a set of stars all born at a given time (single-age stellar population), the line joining the loci of luminosity and temperature of stars of different masses at a later time t is called an *isochrone*. The number of stars in each

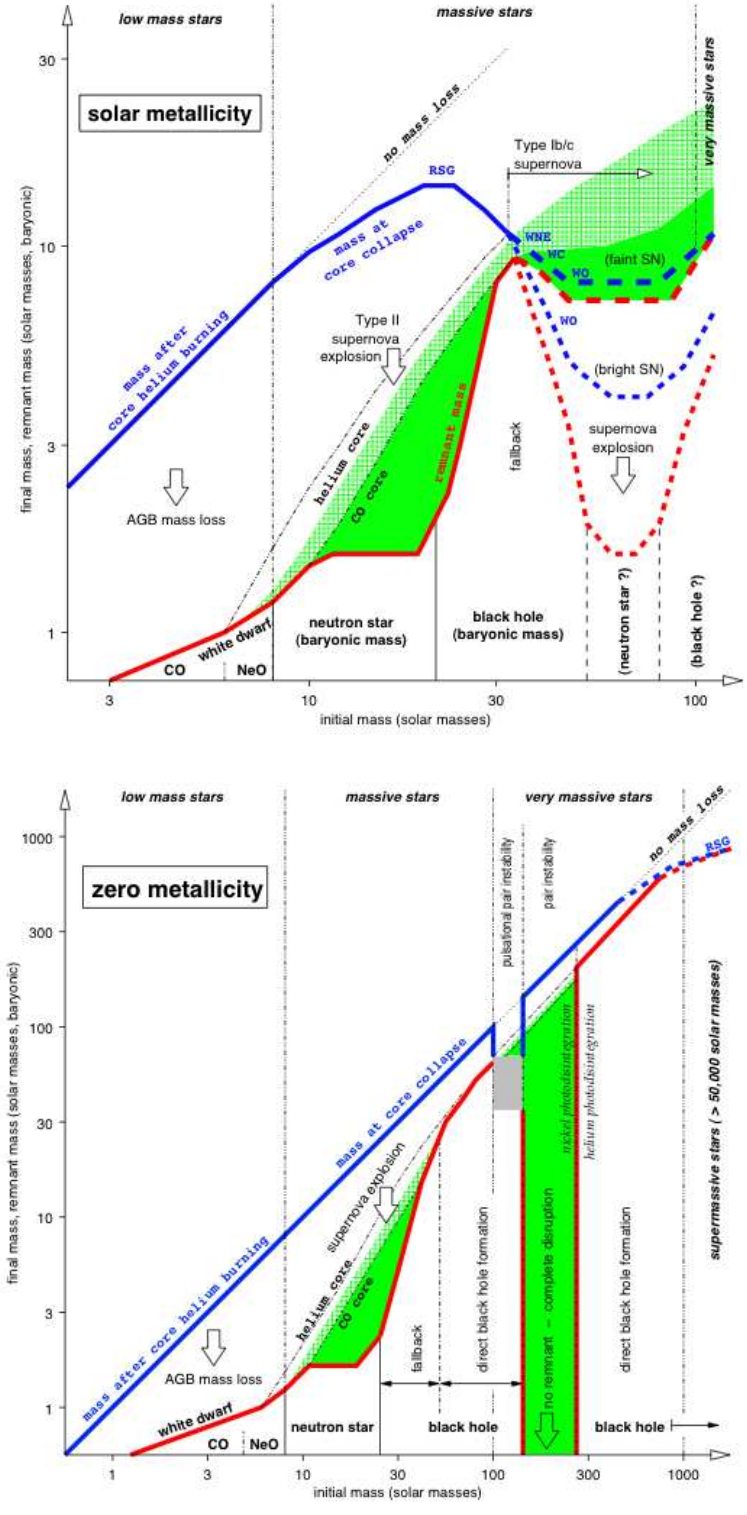

Figure 2.8: Fig. 2 and 3 from Heger et al. [2003]. Stellar mass at the time of final explosion or remnant formation (blue line), remnant mass (red line) and metals released in the interstellar medium (green fill and hatching) as a function of initial mass of the stars for solar metallicity (top panel) and for zero metallicity (bottom panel).

position along the isochrone can be inferred from the stellar initial mass function. To build the spectral energy distribution emitted by a simple stellar population of this type, we require a library of individual stellar spectra spanning complete ranges of stellar parameters.

2.2.3 Library of stellar spectra

To describe the observable properties of a star at any point in the Hertzsprung-Russell diagram, we require a library of individual stellar spectra encompassing a full range of stellar fundamental parameters: bolometric luminosity, $L_{\rm bol}$, radius, R, effective temperature, $T_{\rm eff}$, mass, m, and surface gravity, g. $L_{\rm bol}$, R and $T_{\rm eff}$ are related by the Stefan-Boltzmann law,

$$L_{\rm bol} = 4\pi R^2 \sigma {T_{\rm eff}}^4 \,, \tag{2.2.5}$$

where $\sigma = 5.6708 \cdot 10^{-5}$ erg cm^{-2} K^{-4} s^{-1} is the Stefan-Boltzmann constant. The radius is related to the mass and the surface gravity through

$$g = \frac{Gm}{R^2} \,, \tag{2.2.6}$$

where $G = 6.67 \times 10^{-8}$ is the gravitational constant. A stellar spectrum can usually be approximated as a black-body spectrum.[4] The general shape of the stellar spectrum is driven by the effective temperature.[5] The temperature can be estimated from particular features of the stellar spectrum, such as the *Balmer discontinuity* (this feature is caused by the variation with wavelength of the bound-free cross-section of hydrogen in the stellar atmosphere). Furthermore, at fixed temperature, the strengths of absorption features in the stellar spectrum, which depend on the surface gravity, can help determine the *luminosity class* (i.e. whether the star is a dwarf or a giant).

Libraries of stellar spectra can be either empirical (i.e. observed) or theoretical. A main advantage of theoretical spectra is that they can be computed over complete ranges of stellar types and wavelengths. However, a disadvantage is that even the most state-of-the-art model atmospheres, used to produce theoretical stellar spectra, do not account for all possible atomic and molecular transitions (and the parameters of the models, such as line wavelengths, energy levels and oscillator strengths, are also uncertain). Empirical stellar spectral libraries have their own limitations. They usually rely on observations of nearby stars, thus tied to the specific chemistry and star formation history of the Milky Way and the Magellanic Clouds. For example, empirical spectral libraries do not include massive, low-metallicity stars, which are

[4] When this is the case, a temperature can be estimated by computing the ratio of fluxes at different wavelengths. This is very close to the effective temperature and is called the *color temperature*.

[5] According to Wien's law, $\lambda_{\rm max} = b/T$, where $\lambda_{\rm max}$ is the peak wavelength, T is the absolute temperature of the black body, and b is a constant of proportionality called Wien displacement constant, equal to 2.898×10^{-3} m K.

Name	Type	Wavelength range	FWHM	Reference
BaSeL	T	91Å - 160μm	$\sim$ 20Å	Westera et al. [2002]
Pickles	E	1150Å - 2.5μm	11Å	Pickles [1998]
STELIB	E	3200-9300Å	3Å	Le Borgne et al. [2003]
MILES	E	3500-7430Å	2.3Å	Sánchez-Blázquez et al. [2006]
ELODIE	E	4100-6800Å	$<$ 0.55Å	Prugniel and Soubiran [2001]

T = theoretical, E = empirical

Table 2.2: Examples of theoretical and empirical stellar libraries. We report name, type, wavelength range, FWHM around 5500 Å and reference.

expected to be present in large numbers at high redshift. Thus, empirical spectral libraries do not sample the full range of fundamental stellar parameters (temperature, mass, radius, surface gravity, metal composition). Another disadvantage of such libraries is that they are usually assembled by patching together observations obtained with different telescopes in different wavelength ranges (and often at different resolving powers). Even so, observed stellar spectra do not generally cover the full wavelength range, and theoretical models are required to complement the spectral energy distributions. We list in Table 2.2 the main characteristics of a few widely used stellar spectral libraries.

2.2.4 Modeling the spectral energy distribution of stellar populations

The spectral energy distribution of a simple stellar population at any age is the superposition of the individual spectra of stars at all points along the isochrone, weighted by the initial mass function. Several codes have been developed to compute the spectral evolution of stellar populations as a function of the star formation and chemical enrichment histories, using the principles described in the previous sections: PÉGASE [Fioc and Rocca-Volmerange, 1997, Le Borgne et al., 2004], GALAXEV [Bruzual and Charlot, 2003], and the codes of Maraston [1998] (see also Maraston 2005, Maraston and Strömbäck 2011) and Vazdekis [1999] (see also Vazdekis et al. 2010) differ in the algorithms used to compute the evolution of stars in the Hertzsprung-Russell diagram and in the prescriptions adopted for stellar evolution and stellar spectra; the GALEV evolutionary synthesis model [Kotulla et al., 2009] has been designed specifically to follow the chemical enrichment of the gas associated with stellar populations; Starburst99 [Leitherer et al., 1999] is optimized to model the spectral evolution of young stellar populations; BPASS [binary population and spectral synthesis; Eldridge and Stanway, 2009] has been developed to explore the influence of massive binary stars on the evolution of stellar populations.

We describe here in slightly more detail the GALAXEV code [Bruzual and Char-

lot, 2003], which we will use in Chapter 3. In this code, the spectral energy distribution at time t of a composite stellar population can be written as

$$L_{\lambda,\mathrm{stars}}(t) = \int_0^t dt' \, \psi(t-t') \, S_\lambda[t', Z(t-t')] \, , \tag{2.2.7}$$

where $\psi(t-t')$ is the star formation rate at time $t-t'$ and $S_\lambda[t', Z(t-t')]$ is the luminosity emitted per unit wavelength per unit mass by a simple stellar population of age t' and metallicity $Z(t-t')$, assuming that the initial mass function is constant in time [e.g. Chabrier, 2003]. The code allows one to compute the evolution of stellar populations at ages between 1×10^5 and 2×10^{10} years and metallicities between 0.005 and 4 $Z_\odot$. In the version of the model used in this thesis, the stellar evolution prescription (see Section 2.2.2) relies on the calculations of Bertelli et al. [2008] for stars with masses between 0.15 and 2.5 $M_\odot$ and Bertelli et al. [2009] for stars of mass between 2.5 and 20 $M_\odot$. The evolution of stars more massive than 20 $M_\odot$ is the same as in Bruzual and Charlot [2003]. These tracks are extended with the prescription of Marigo et al. [2008] to describe the evolution of stars through the AGB phase and using calculations from various sources to describe the subsequent evolution down to the white-dwarf cooling sequence (see Bruzual and Charlot 2003 for details). For reference, we show in Figure 2.9 (top panel) isochrones derived from the original Bertelli et al. tracks, at ages between 0.5 and 12 Gyr (solid lines). Dashed lines represent other models at young ages from the Padova database [Bertelli et al., 1994]. The bottom panel shows a zoom-in of the yellow region in the top panel, which includes isochrones for the E-AGB and TP-AGB phases from Marigo et al. [2008]. All phases until the end of the E-AGB are drawn in green, while blue and red colors distinguish the O-rich and C-rich phases, respectively, on the TP-AGB. In this example, the isochrones are plotted at ages between 0.06 and 16 Gyr. The version of Bruzual and Charlot [2003] model used in this thesis also includes the MILES library of observed stellar spectra from Sánchez-Blázquez et al. [2006]. As a result, at wavelengths between 3525 and 7500Å, the stellar population spectra have a native spectral resolution of 2.3Å (full width at half-maximum), corresponding to a resolving power R $\approx$ 2200 at 5000Å (the resolution is coarser outside this wavelength range), and the signal-to-noise ratio is around 95. Dust attenuation can be included in the computation of the composite spectral energy distributions, using the prescriptions by Charlot and Fall [2000] and da Cunha et al. [2008] (see Section 2.4 below).

2.3 Nebular emission

Stars form in gas-rich environments. When this gas is ionized by the radiation from young massive stars, it emits light (nebular emission), which contributes to the spectral energy distribution of galaxies. The inclusion of nebular emission in the spectral analysis of galaxies should improve estimates of the parameters pertaining to the interstellar medium and the young stars ionizing the gas. In this Section, we

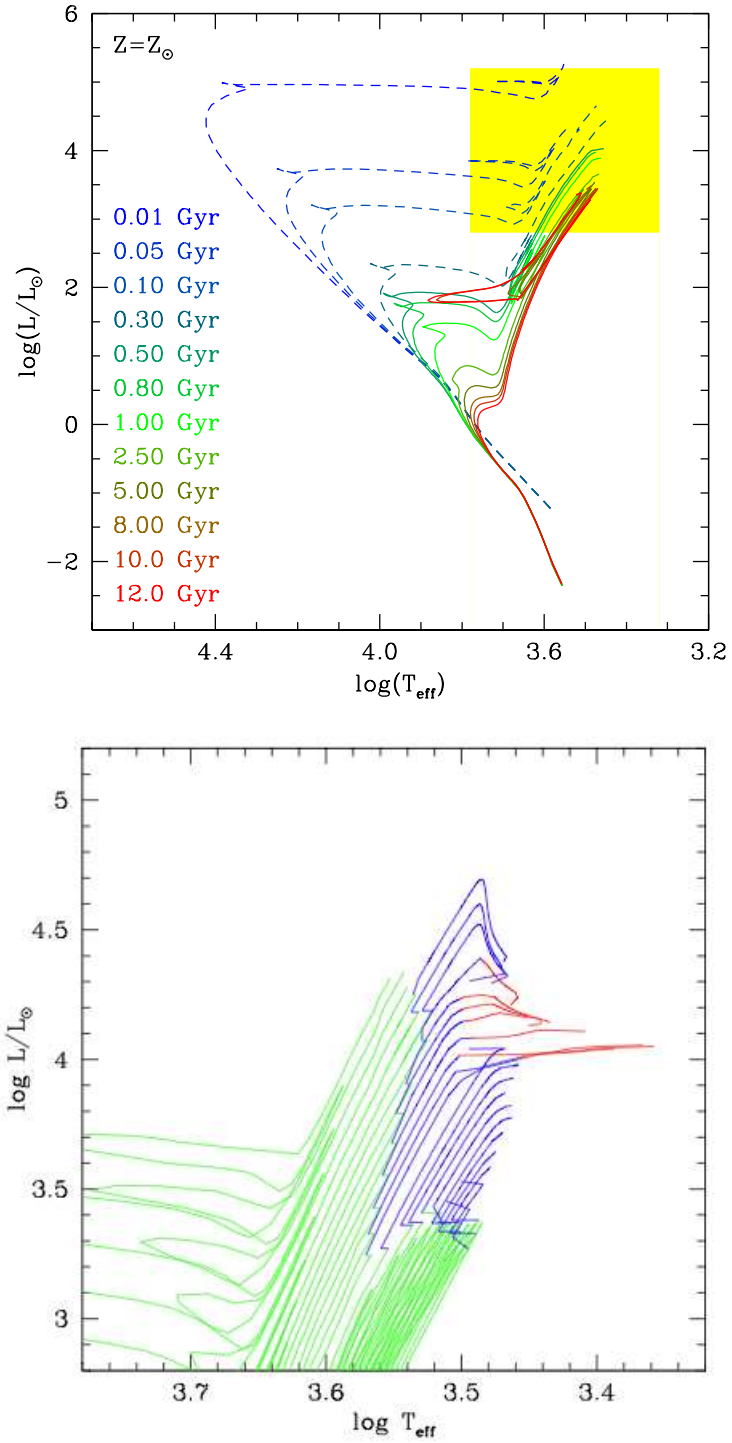

Figure 2.9: Top panel: isochrones from Bertelli et al. for different ages at solar metallicity, solid lines; isochrones from other models in the Padova database at solar metallicity, dashed lines. Bottom panel: zoom of the yellow region in the top panel, showing isochrones (ages between 0.06 and 16 Gyr) form Marigo et al. [2008] (Fig. 1) for pre-TP-AGB phase in green, and O-rich and C-rich configurations of the TP-AGB phase respectively in blue and red.

Description	Phase	T (K)	n (cm^{-3})
Cold molecular	H_2	$\sim$15	> 100
Cool atomic	H I	$\sim$80	$\sim$30
Warm atomic	H I	10^4	$\sim$0.5
Diffuse ionized	H II	10^5	10^{-2}
Hot ionized	H II region	10^4	$10^3 - 10^4$

Table 2.3: Average temperatures and densities of the gas in the interstellar medium in different phases.

first describe the different phases of the gas in the interstellar medium of galaxies, focusing on the origin of ionized gas. Then, we briefly summarize the main features of nebular emission, and at the end, we present different approaches to include nebular emission in the modeling of galaxy spectral energy distributions.

2.3.1 The interstellar medium

The interstellar medium in galaxies presents a variety of phases. These phases are characterized by the state of hydrogen – the most abundant element – within them. McKee and Ostriker [1977] define three phases of the interstellar medium (see below): cold and dense clouds filled mainly by H_2 molecules; diffuse clouds (either cool or warm) filled mainly by H I; and hot ionized gas, which can be divided into a diffuse component and a dense component. These three phases are in rough pressure equilibrium, because the interchange of material among them is fairly rapid, of the order of 10^6 years. The properties of interstellar gas are described by the pressure, p, the number density, n, and the temperature, T. For an ideal gas,

$$p = nkT\,, \tag{2.3.8}$$

where k is the Boltzmann constant. Typical values of these parameters in the interstellar medium are given in Table 2.3 and represented in Figure 2.10. The volume of the interstellar medium is mostly filled by the diffuse hot component, while the mass is dominated by the cooler and denser atomic and molecular phases, i.e. the clouds. We now describe the main features of the three phases of the interstellar medium.

Molecular (or dense) clouds

Typical molecular clouds range from small, not very massive clouds ($\sim$ 1 pc, $< 50\ M_\odot$) to giant molecular clouds ($\sim$ 50 pc, $\sim 10^6\ M_\odot$). The density in these clouds is greater than 10^2 cm^{-3}. At such densities, the formation of H_2 is allowed on dust grains (see Section 2.4.1). Molecules of heavier elements can also be detected in these clouds, such as CH, CH^+ and CN (identified by absorption lines between 3000 and 4000 Å [McKellar, 1940]) and OH and H_2O (detectable in the radio). These molecules form by hydrogenation or oxidation reactions on grain

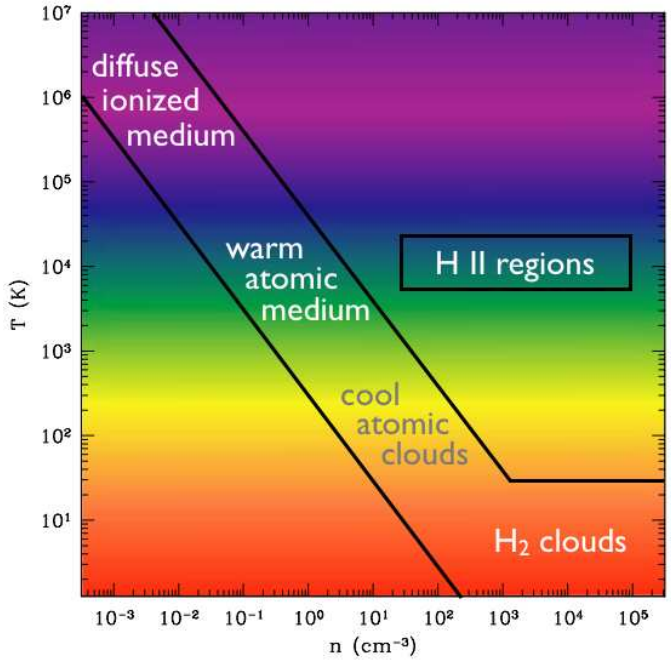

Figure 2.10: Representation of the different phases of the interstellar medium in terms of temperature and density.

surfaces. One of the most abundant molecules, CO, is actually easier to identify than H_2. Since CO molecules are excited by collisions with H_2, radio observations of CO are generally used to trace H_2. Other molecules called *polycyclic aromatic hydrocarbons* (PAH, benzene rings, with sizes ~ 1 nm) are detected in dense clouds, through their characteristic vibrational modes in the infrared [Omont, 1986]. The transition zone between molecular and atomic gas is called the *photodissociation region* [Hollenbach and Tielens, 1997].

Atomic (or diffuse) clouds

Diffuse clouds are characterized by moderate densities ($n \sim 1$ cm^{-3}) and contain mainly neutral hydrogen. They may typically have sizes around 5 pc and masses around $30\,M_\odot$. Neutral hydrogen can be detected through observations of the 21-cm line, which corresponds to the transition between the two hyperfine levels of the fundamental state of H I (i.e. the change in the spin configuration of the electron and proton from parallel to antiparallel). In diffuse interstellar clouds, neutral hydrogen is cool in the inner parts and warm in the outer parts. This outer part is partially ionized by the hard ultraviolet photons of the interstellar radiation field. Along with hydrogen, neutral heavy elements can also be found in diffuse clouds, especially oxygen and carbon, which are important cooling sources.

Hot ionized gas

Ionized gas can be found either in the form of diffuse intercloud gas, or in the form of H II regions. The diffuse component is particularly hot, but it has extremely low density. As suggested by McKee and Ostriker [1977], the state of the interstellar medium is likely to be regulated by supernova explosions, which sweep out the gas, disrupting existing clouds and maintaining the intercloud medium in a hot ionized state.

H II regions

When gas in molecular clouds cools down and condensates, star formation is ignited. The newly formed O- and B-type stars emit photons at wavelength shorter that the hydrogen Lyman limit (912 Å), which corresponds to the energy required to remove a ground state electron from H I. A cloud (nebula) is said to be ionized when the hydrogen in it is nearly completely ionized. The properties of the ionized gas depend on the ionizing source. When the source is a star (or several stars) of O or B type, the ionized gas cloud is called an *H II region*.

Two different types of H II regions can be defined: ionization-bounded and density-bounded. If the cloud is sufficiently large to absorb all the ionizing radiation emitted by the central source, the H II region is surrounded by a shell of neutral atomic gas (H I) and is said to be ionization-bounded. If the cloud is not large enough to absorb all the ionizing radiation, the region is said to be density-bounded. In this case, some ionizing photons can escape into the interstellar medium. Ionization-bounded H II regions are usually fainter than density-bounded H II regions (see Beckman et al. 2000 and references therein).

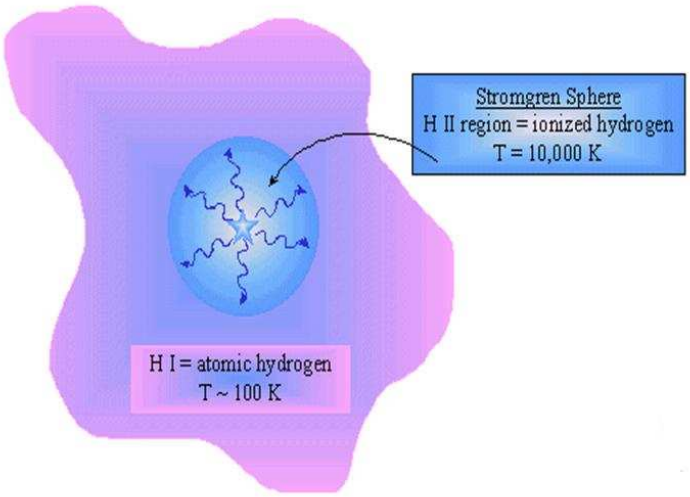

Figure 2.11: A schematic representation of an H II region. The ionizing source is shown in the center. The region occupied by the ionized hydrogen (Stromgren sphere) is shown in blue. The region filled by the neutral hydrogen is shown in pink. (Figure from http://www.astro.cornell.edu/academics/)

Figure 2.11 shows a schematic representation of an ionization-bounded H II region, with the ionizing source at the center and a shell of neutral hydrogen in the outer parts. The effective temperature of stars in the center of an H II region ranges typically between 30,000 and 50,000 K. The hotter the stars, the higher the level of ionization of the cloud. The temperature in the nebula, instead, is set by the equilibrium between heating through photoionization and cooling through recombination and radiation. This temperature ranges between 7000 and 25,000 K, with a typical value around 10,000 K. Typical densities in the ionized part of the nebula range around 10^2 cm^{-3} and can reach up to 10^4 cm^{-3}. The region where all ionizing photons are absorbed and the total number of recombinations equals the total number of photoionizations is called the Strömgren sphere. The Strömgren radius ranges typically between 30 and 150 pc for O- and B-type stars with surface temperature between 30,000 and 50,000 K.

2.3.2 Spectral features of the nebular emission

The spectrum of an H II region is characterized by strong emission lines and a continuum component. This characteristic light arises from recombination processes and the radiative cooling of collisionally excited gas.

Three processes contribute to the continuum radiation: free-bound (recombination), free-free (bremsstrahlung) and two-photon emission processes. The emission coefficients for free-bound and two-photon processes for hydrogen are shown in Fig-

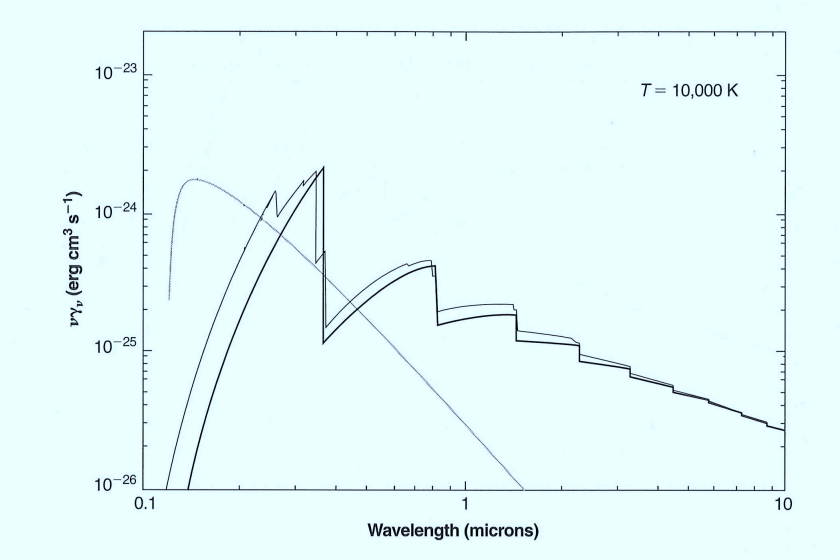

Figure 2.12: Frequency variation of the free-bound continuous-emission coefficient $\gamma_\nu(T)$ for hydrogen (black solid line) and for helium (black thin solid line). Frequency variation of the two-photon-process coefficient (grey solid line). All lines correspond to a low-density regime at $T = 10,000$ K. Figure from Osterbrock and Ferland [2006].

ure 2.12. Free-bound continuum radiation (black solid line) is produced by recombination of free electrons with hydrogen ions. The shape of the continuum radiation as a function of wavelength depends on the distribution of electron energies and the cross-section for the capture of electrons. Free-free emission is generated by the interaction between an electron and a proton or another ion in the gas. It increases with the cube of the wavelength and so dominates the continuum in the infrared and radio regions. The two-photon process (grey solid line) involves transitions from the $n = 2$, $L = 0$ level of hydrogen to the $n = 1$, $L = 0$ ground state. This is forbidden for a single photon (which would contribute to a Lyman-α line) because angular momentum would not be conserved, but it is allowed if two photons are emitted. As the energy can be divided in any way between the two photons, the spectrum is continuous at all wavelength larger than the wavelength of Lyman-α (0.1216 μm).

Emission-line radiation can be caused by recombination of ions and radiative de-excitation of collisionally excited molecules, atoms and ions. Recombination lines of hydrogen are emitted when protons capture electrons onto exited levels and these electrons cascade by downward radiative transitions to the ground level. These lines are identified in series, depending on the level reached by the electron (Lyman: $n = 1$, Balmer: $n = 2$; Paschen: $n = 3$; etc.). Direct recombinations to the ground state produce ionizing photons, which can be reabsorbed elsewhere in the nebula. The *on-the-spot* approximation [Osterbrock and Ferland, 2006] consists in neglecting the possible spatial separation between the emission and reabsorption of

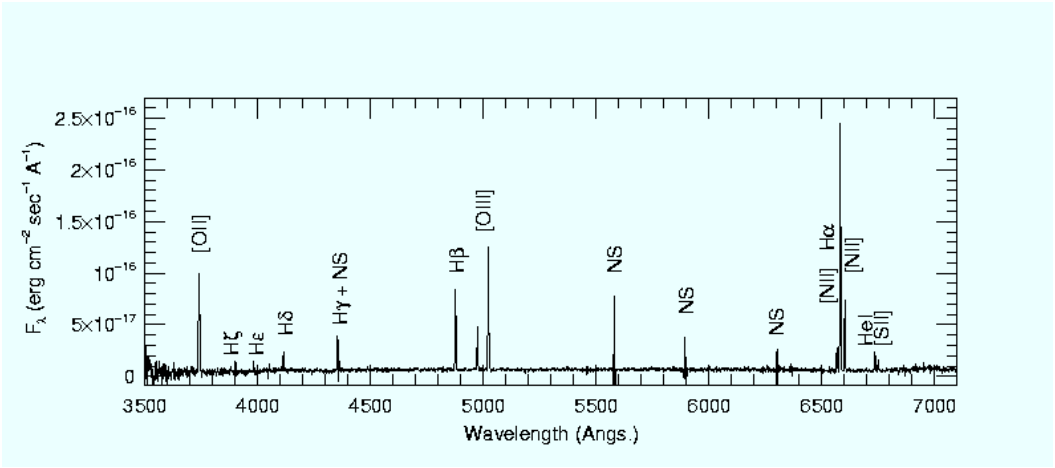

Figure 2.13: Typical spectrum of an H II region with the strongest lines labeled. (htp://frigg.physastro.mnsu.edu)

such photons, so that they have no net effect on the ionization balance. If the cloud is optically thin, all photons issued from recombination can escape the nebula (*case-A recombination*). Otherwise, all Lyman-line photons issued from recombination cascades to the ground state are immediately reabsorbed by a neutral atom, so that all downward transitions to the $n = 1$ level can be ignored (*case-B recombination*). In this case, every recombination must eventually yield a Balmer photon (transition to the $n = 2$ level) at optical wavelengths. In case-B recombination, the luminosity ratio between the Hα line (transition $n = 3 \rightarrow 2$) at 6563 Å and the Hβ line ($n = 4 \rightarrow 2$) at 4861 Å is 2.87 for a nebula of temperature $T = 10,000$. The variation of this ratio, called *Balmer decrement*, is widely used to quantify the presence of dust in the ionized clouds (see Section 2.4).

The other main source of emission-line radiation in H II regions is the radiative de-excitation of collisionally excited ions, such as O^+, O^{++}, N^+ and S^+. Some strong optical transitions are shown in Figure 2.13. The most prominent ones are [O II]$\lambda\lambda$3727, [O III]λ5007, [N II]λ6548, [N II]λ6584 and [S II]$\lambda\lambda$6716, 6731. In the limit of low density, every collisional excitation is followed by the emission of a photon. If the density is high,[6] ions can also be efficiently de-excited through collisions, and the radiative cooling rate is reduced. In laboratory (i.e. high-density) conditions, the excited levels corresponding to the ionic transitions in Figure 2.13 would be collisionally de-excited before radiative decay could take place. This is the reason why these lines are said to be *forbidden*. The luminosity ratios of the strong emission lines in Figure 2.13 provide valuable insight into the properties of the ionized gas, such as metal content and ionization parameter [i.e. the ratio of ionizing-photon to gas densities; e.g., Charlot and Longhetti, 2001, Brinchmann et al., 2004, Dopita

[6]Collisional de-excitation is effective at densities $> n^4$ cm^{-3}, Loeb et al. [2008].

et al., 2006a].

2.3.3 Nebular emission models

Two main approaches have been adopted to include nebular emission in the spectral modeling of galaxies: the adoption of empirical relations and the use of photoionization codes.

The simplest way to account for the nebular emission in spectral models of galaxies is to compute the ionizing rate implied by the emission from young stars and then derive analytically the intensifies of the nebular continuum and hydrogen emission lines, assuming photoionization equilibrium [Osterbrock and Ferland, 2006], for given values of the temperature and hydrogen density. Then, empirical relations may be used to compute the intensities of helium, carbon, nitrogen, oxygen and sulphur emission lines relative to those of H-Balmer lines (see Anders and Fritze-v. Alvensleben 2003). In this approach, therefore, the absolute strength of the continuum and line emission depends primarily on the total number of Lyman continuum photons. This varies with metallicity, since stars of different metallicities have different temperatures [e.g. Krueger et al., 1995]. This approach has been used, for example, in PÉGASE [Fioc and Rocca-Volmerange, 1997] and by Schaerer and de Barros [2009].

The other way to model nebular emission in galaxy spectra is to appeal to a sophisticated photoionization code. We describe here the approach developed by Charlot and Longhetti [2001], which we will use in Chapter 3. In this approach, the luminosity produced by stars and the ionized gas in a galaxy is expressed, by analogy with Equation 3.2.1, as

$$L_{\lambda,\mathrm{stars+neb}}(t) = \int_0^t dt'\, \psi(t-t')\, S_\lambda[t', Z(t-t')] T_\lambda^+(t,t')\,, \tag{2.3.9}$$

where $\psi(t-t')$ is the star formation rate at time $t-t'$ and $S_\lambda[t', Z(t-t')]$ is the luminosity emitted per unit wavelength per unit mass by a simple stellar population of age t' and metallicity $Z(t-t')$. $T_\lambda^+(t,t')$ is the transmission function of the ionized gas, defined as the fraction of the radiation, produced at wavelength λ at time t by a generation of stars of age t', that is transferred by the ionized gas. $T_\lambda^+(t,t')$ will be close to zero at wavelength blueward of the Lyman limit if the H II regions are assumed to be ionization bounded. It will be greater than unity at wavelengths where the ionized gas radiates (continuum and line emission). One may assume for simplicity that the transmission function depends only on the age t' of the ionizing stars, and not on the age t of the galaxy, since the nebular emission depends only on the current gas content of the galaxy. Moreover, only the ionization by O- and B-type stars is significant. Combining these two assumptions, the transmission function can be written as

$$\begin{aligned} T_\lambda^+(t,t') &= 1 \qquad\quad \text{for} \quad t' > t_{OB} \\ T_\lambda^+(t,t') &= T_\lambda^+(t') \qquad \text{for} \quad t' \leqslant t_{OB}\,. \end{aligned} \tag{2.3.10}$$

The lifetime of OB stars, t_{OB}, can range between 1 and 100 Myr. Charlot and Longhetti [2001] adopt a typical value of 10 Myr, corresponding to the age by which the production of ionizing photons by an evolving stellar population drops by over 99 percent. They compute the transmission function using the standard photoionization code CLOUDY [Ferland, 1996]. This simultaneously solves the equations of thermal and statistical equilibrium for a model nebula, including ionization and recombination processes. The code requires input parameters to describe the properties of the gas in the nebula. Charlot and Longhetti [2001] adopt *effective* parameters to describe the ensemble of H II regions and the diffuse gas ionized by young stars in a galaxy:

- the *effective ionization parameter*, defined as the ratio of ionizing-photon to gas densities. This depends on the distance r from the ionizing source and the age t' of the central stars, $\tilde{U}(t', r)$, and can be simply parameterized in terms of a zero-age effective ionization parameter (see below).
- the *effective metallicity* of the gas
- the *effective dust-to-heavy element ratio*, defined as the fraction of heavy elements locked into dust grains in the ionized gas, $\tilde{\xi}_{\rm d}$

The effective ionization parameter is a function of the effective rate of ionizing photons, $\tilde{Q}(t')$, and the effective hydrogen density, $\tilde{n}_{\rm H}$. It can be expressed as

$$\tilde{U}(t', r) = \frac{\tilde{Q}(t')}{4\pi r^2 \tilde{n}_{\rm H} c}\,, \tag{2.3.11}$$

where c is the speed of light and

$$\tilde{Q}(t') = \frac{\tilde{M}_*}{hc} \int_0^{\lambda_L} d\lambda\, \lambda S_\lambda[t', Z(t - t')]\,. \tag{2.3.12}$$

In this equation, $\tilde{M}_*$ is the effective mass of the ionizing star clusters in the galaxy and λ_L is the Lyman limit. The ionized gas is described as a spherical shell around the ionizing source, with innermost radius $r_{\rm in}$. The Stromgren radius fixes the ionization profile of the gas and it is defined by

$$R_{\rm S}^3 = \frac{3\tilde{Q}}{4\pi \tilde{n}_{\rm H}^2 \tilde{\varepsilon} \alpha_{\rm B}}\,. \tag{2.3.13}$$

where $\tilde{\varepsilon}$ is the effective volume-filling factor, i.e. the ratio of the volume-averaged hydrogen density to the effective hydrogen density $\tilde{n}_{\rm H}$. In the approximation of $r_{\rm in} \ll R_{\rm S}$, the geometry is truly spherical and

$$\langle \tilde{U} \rangle (t') \approx 3\tilde{U}(t', R_{\rm S})\,. \tag{2.3.14}$$

Combining 2.3.11, 2.3.13 and 2.3.14, we obtain

$$\langle \tilde{U} \rangle (t') \approx \frac{\alpha_{\rm B}^{2/3}}{c} \left[\frac{3\tilde{Q}(t') \tilde{\varepsilon}^2 \tilde{n}_{\rm H}}{4\pi} \right]\,. \tag{2.3.15}$$

The reason for parametrizing nebular emission in terms of the ionization parameter is that different combinations of $\tilde{Q}$, $\tilde{\varepsilon}$ and $\tilde{n}_{\rm H}$ can give rise to similar values of $\langle\tilde{U}\rangle$. It is convenient also to fix $\tilde{M}_*$ and hence $\tilde{Q}(t')$ and use as a variable the *zero-age effective ionization parameter*, $\langle\tilde{U}_0\rangle = \langle\tilde{U}\rangle(0)$. In this case, a grid of models can be computed for different values of $\log\langle\tilde{U}_0\rangle$, e.g. $-3.5, -3.0, -2.5, -2.0$ and -1.5. The second main input model parameter is the *effective metallicity* of the gas, $\tilde{Z}$. This affects both the rate of ionizing photons from young stars and the relative strengths of different emission lines. Charlot and Longhetti [2001] adopt the default cosmic abundances of CLOUDY at solar metallicity, except for the following element abundances (by number relative to hydrogen): $({\rm N/H}) = 5\times10^{-5}$, $({\rm O/H}) = 6.6\times10^{-4}$ and $({\rm S/H}) = 1.3\times10^{-5}$. These choices were motivated in part by the desire to match the observations of nearby galaxies and H II regions. A grid of models can be computed at different values of the metallicity between 0.02 and 3 times $Z_\odot$. The third main input model parameter is the *effective dust-to-heavy element ratio*. It is a measure of the fraction of refractory elements depleted onto dust grains in the ionized gas. It is convent to compute a grid of modes for at least three values of $\tilde{\xi}_{\rm d}$, e.g. 0.1, 0.3 and 0.5.

Dopita et al. [2005, 2006b,a] propose an alternative model to compute the nebular emission from star-forming galaxies, which is based on the MAPPINGSIII photoionization code [Dopita et al., 2002, Groves et al., 2004]. The main difference with CLOUDY is that MAPPINGSIII includes the dynamical evolution of H II regions. In this model, H II regions expand until the internal pressure caused by stellar winds exceeds the ambient pressure of the interstellar medium. When this happens, the H II region is assumed to stop growing and fade. In this framework, the ionization parameter is not a free variable, but it is replaced by a parameter tracing the ratio of the star-cluster mass to the pressure in the surrounding interstellar medium. This ratio is expressed as

$$\mathcal{R} = \frac{M_{\rm cl}/M_\odot}{P_0/k}\,, \tag{2.3.16}$$

where $M_{\rm cl}$ is the mass of the star cluster inside an H II region and P_0/k is the pressure in the interstellar medium (k is the Boltzmann constant). Although this model should be more appropriate than models with static H II regions, it does not completely agree with observations of local H II regions. In particular, interstellar clouds tend to expand too much, presumably because of a too high internal pressure. Dopita et al. [2005] suggest that their one-dimensional model many not be accurate enough to describe evolving H II regions. Considering the dynamics of star cluster and the hydrodynamics of both the ionized and the molecular gas might lead to better results.

2.4 Attenuation by dust

Dust absorbs and scatters radiation from stars and gas and reemits this radiation mainly at infrared wavelengths. Several prescriptions have been proposed to imple-

ment absorption and scattering by dust grains of the light emitted by stars and gas in galaxies. Differences among these prescriptions arise primarily from uncertainties in the composition and optical properties of dust grains, the way in which dust is distributed in the various phases of the interstellar medium, and the treatment of geometric effects.

In this Section, we first provide an overview of the general properties of interstellar dust. Then, we present different prescriptions which have been proposed to model the effects of dust on the emission from stars and gas, both in our Galaxy and in external galaxies.

2.4.1 General properties of interstellar dust

Dust grains are solid particles composed of elements heavier than hydrogen. As seen in Section 2.2.2, heavy elements are produced in thermonuclear reactions during the evolution of stars. Stellar winds and supernova explosions cause these metals to be spread into the interstellar medium, where they condense to form dust grains. The main observational evidence for the depletion of heavy elements onto dust grains is that the gas-phase abundances of metals are generally lower in interstellar clouds than in stellar atmospheres (where hydrogen and heavy elements are in the gas phase).

The elements contributing the most to the mass of interstellar grains are carbon, nitrogen, oxygen, magnesium, silicon and iron. Grains can be divided into *graphitic grains*, *ices* and *silicates* [Whittet, 2003]. Graphitic (carbon) grains are particularly small in size ($< 2\mu$m). They produce a strong absorption feature in the ultraviolet (the bump at 2175 Å, see Section 2.4.2 below). Ices refer to any volatile molecular solid composed of primarily the CHON group of elements. They require high densities to condense and survive and produce strong emission features by vibrational transitions in the near-infrared ($\sim 3\mu$m). At longer wavelength ($\sim 10 - 20\mu$m), emission and absorption features are generally attributed to silicate grains, which have a wide range of possible chemical compositions based on SiO_4 units.

Dust grains mix with the interstellar medium and experience chemical and isotopic changes through interactions with the gas and radiation in interstellar space. For this reason, different types of dust grains form and survive in different regions of the interstellar medium. Ices tend to form in dense molecular clouds (see Section 2.3.1), where they contribute to the formation of H_2.[7] Graphitic grains and silicates instead can form in both low-density and high-density environment.

2.4.2 Modeling attenuation by dust in galaxies

The ability to constrain galaxy physical parameters (e.g. stellar mass, star formation rate, mean stellar age) largely depends on the accuracy to which the effects of

[7] At the typical densities of molecular clouds, H atoms are trapped onto dust grains. When two hydrogen atoms recombine, the binding energy is partly absorbed into the grain lattice and the resulting molecule is ejected from the surface and returned to the gas (grain surface catalysis).

interstellar dust on the light emitted by stars and gas can be estimated.

In the next paragraphs, we first briefly review the basics of extinction of radiation by dust grains and mention measurements of the extinction curve in our Galaxy and the Magellanic Clouds. Then, we describe how attenuation by dust can be quantified and modeled in external galaxies. In this thesis, we follow the general convention that *extinction* refers to the combined effects of absorption and scattering by dust grains along single lines of sight. The term *attenuation* refers to the net, angle-averaged absorption of photons emitted in all directions by all stars within a galaxy (and eventually scattered before they are absorbed).

2.4.2.1 Basics of interstellar extinction

Extinction occurs when radiation propagates through a medium containing small particles. Such particles absorb and scatter light, reducing the intensity of the radiation. In the approximation that dust grains are small particles with cross-section $\sigma_\lambda^{\rm ext}$, we can write

$$\frac{dI_\lambda}{I_\lambda} = -n_{\rm d}\,\sigma_\lambda^{\rm ext}\,dL\,, \tag{2.4.17}$$

where I_λ is the intensity of radiation at a particular wavelength λ and $n_{\rm d}$ is the number density of dust grains in a cylindrical column of length L. Integrating along the column length, we obtain

$$\frac{I_\lambda}{I_{0,\lambda}} = e^{-\tau_\lambda^{\rm ext}}\,, \tag{2.4.18}$$

where $I_{0,\lambda}$ is the initial intensity at $L = 0$ and $\tau_\lambda^{\rm ext}$ is the optical depth of extinction at wavelength λ caused by dust.[8] The total extinction expressed in magnitudes is

$$A_\lambda = -2.5\log\left(\frac{I_\lambda}{I_{0,\lambda}}\right) = 1.086\,\tau_\lambda^{\rm ext}\,. \tag{2.4.19}$$

$A_\lambda/A_{\rm V}$ is the *absolute extinction* or *extinction curve* (as a function of λ).

2.4.2.2 The extinction curves of the Milky Way and the Magellanic Clouds

Observationally, the extinction curve of the Milky Way can be determined by comparing the colors of stars of identical spectral type and luminosity class, but unequal reddening. The apparent magnitude at a particular wavelength is defined as the absolute magnitude, $M(\lambda)$, scaled with the distance, d, and increased with the total extinction

$$m_\lambda = M_\lambda + 5\log d + A_\lambda\,. \tag{2.4.20}$$

Since stars of identical spectral type and luminosity class have the same absolute magnitude, differences in apparent magnitude can be attributed to distance and

[8] $\tau_\lambda^{\rm ext} = \tau_\lambda^{\rm abs} + \tau_\lambda^{\rm scatt}$.

extinction. If we assume that one of the stars has negligible extinction (i.e. $A_{1,\lambda} = A_\lambda$ and $A_{2,\lambda} = 0$), we can write

$$\Delta m_\lambda = 5 \log(d_1/d_2) + A_\lambda \,. \tag{2.4.21}$$

The dependence on distance can be removed by computing the magnitude difference at two different wavelengths (extinction curves are commonly normalized to the value in the B or V band),

$$E_{\lambda-\mathrm{V}} = \Delta m_\lambda - \Delta m_\mathrm{V} = A_\lambda - A_\mathrm{V} \,. \tag{2.4.22}$$

$E_{\lambda-\mathrm{V}}$ is the *color excess* and is a measurable quantity once the intrinsic colors of stars of different spectral types are known. In the Johnson system, the color excess in the B and V bands and the total extinction, A_V, are related by R_V, the total-to-selective visual extinction:

$$E_{\mathrm{B}-\mathrm{V}} = \frac{A_\mathrm{V}}{R_\mathrm{V}} \,. \tag{2.4.23}$$

The total-to-selective visual extinction depends on the chemical composition and physical properties of dust grains. It is also a proxy for the average grain size: low-R_V lines of sight have smaller grains than high-R_V ones [Noll and Pierini, 2005]. On average, $R_\mathrm{V} = 3.05 \pm 0.15$. This value has been empirically derived [Whitford, 1958, Johnson, 1968] for the diffuse interstellar medium of the Milky Way. Normalizing 2.4.22 to $E_{\mathrm{B}-\mathrm{V}}$ and combining with 2.4.23, we can write the expression for the empirical extinction curve:

$$\frac{A_\lambda}{A_\mathrm{V}} = \frac{1}{R_\mathrm{V}} \frac{E_{\lambda-\mathrm{V}}}{E_{\mathrm{B}-\mathrm{V}}} + 1 \,. \tag{2.4.24}$$

Measuring $E_{\lambda-\mathrm{V}}$ along different lines of sight towards stars in the Milky Way has been the first step in the analysis of the effects of interstellar dust. Cardelli et al. [1989] propose an empirical measure of an average extinction curve of the Milky Way from the ultraviolet to the infrared. More recent studies have shown that the extinction curve in the Magellanic Clouds (derived applying the same method) is different from the average extinction curve in our Galaxy. Gordon et al. [2003] present a comparison of different extinction curves (i.e. measured along different lines of sight) of the Small and Large Magellanic Clouds with the average extinction curve in the Milky Way. The results reveal quite different behaviors, especially in the ultraviolet.

We plot in Figure 2.14 the average extinction curves of the Milky Way (red line), Small Magellanic Cloud (short-dashed line) and Large Magellanic Cloud (long-dashed line). Although the slopes of these extinction curves are different, the general trend is the same: the extinction is stronger in the ultraviolet than in the infrared. Light at a certain wavelength λ is strongly absorbed by grains with cross section of the order of λ. In the interstellar medium, there is an appreciable number of grains with sizes between 0.1 and 1 μm, thus light emitted at ultraviolet and optical wavelengths is the most absorbed. The extinction curves of the Milky Way and

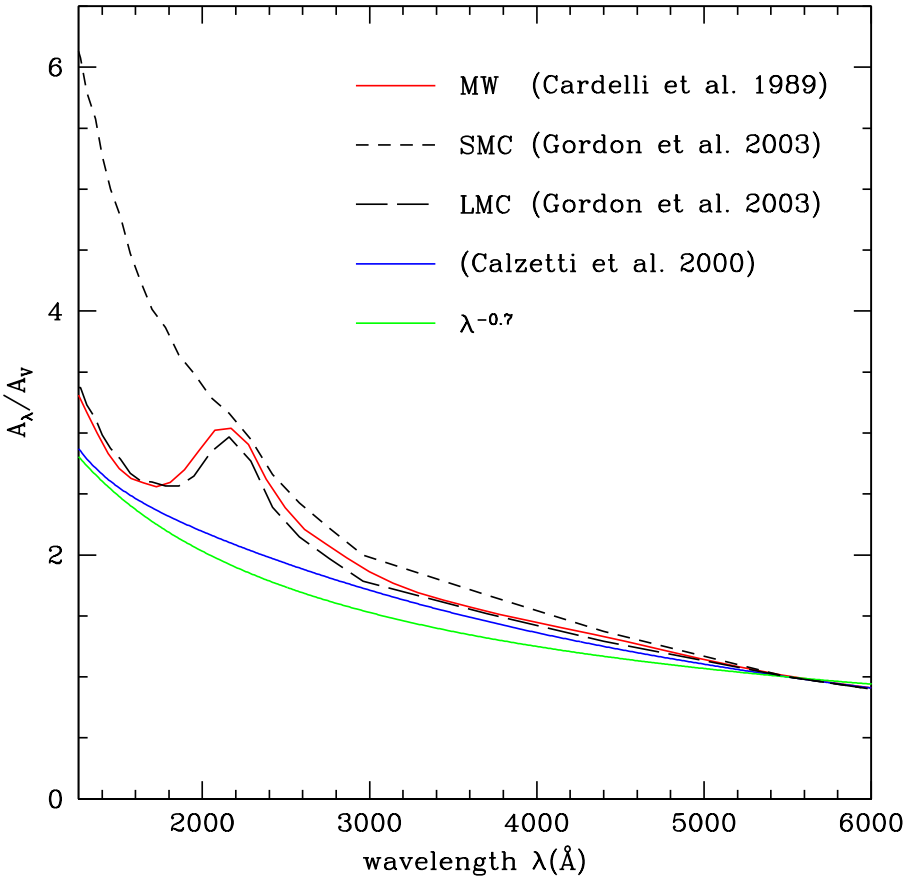

Figure 2.14: Extinction curves (A_λ/A_V) as a function of wavelength in the ultraviolet and optical rest-frame ranges. The red line shows the average extinction curve of the Milky Way as in Cardelli et al. [1989] ($R_V = 3.1$). The short-dashed line shows the extinction curve of the Small Magellanic Cloud bar, while the long-dashed line shows the average extinction curve of the Large Magellanic Cloud, both as in Gordon et al. [2003]. The blue line represents the extinction curve by Calzetti et al. [2000] and the green line shows a curve of the form $\lambda^{-0.7}$ normalized at 5500Å.

the Large Magellanic cloud are also characterized by a prominent feature in the ultraviolet, the *2175 Å bump*. This feature is not observed in the Small Magellanic Cloud, suggesting a different composition of the dust. It is worth mentioning that the 2175 Å bump is purely an absorptive feature and is mainly detectable in lines of sight with $E_{\mathrm{B-V}} > 0.05$. Graphitic grains of dimensions $< 0.01\mu$m are the most likely cause for this absorption. To a good approximation, a fit of this feature is obtained with a Drude profile, as suggested by Fitzpatrick and Massa [1986]:

$$D(x) = \frac{x^2}{(x^2 - x_0^2)^2 + \gamma^2 x^2}\,, \tag{2.4.25}$$

where $x = \lambda^{-1}$. $x_0 = \lambda_0^{-1}$ sets the central wavelength and $1/\gamma$ is the width of the feature (FWHM). The strength of the bump, A_{bump}, is defined as the area under the profile in the extinction curve. The central wavelength tends to be fairly constant in observations through different lines of sight. Instead, γ and A_{bump} vary. In particular, dense molecular clouds are characterized by broad bumps, with a large range in strength. Diffuse regions more likely produce strong bumps, while hot ionized regions tend to have weak and narrow bumps.

2.4.2.3 The attenuation curve in nearby galaxies

When observing the integrated light from distant galaxies, a few complications arise in the estimates of attenuation by dust: the definition of templates of dust-free galaxies; the way in which dust grains are distributed in the interstellar medium; and the role of geometric (and orientation) effects. Empirically, the attenuation curves of distant galaxies can be derived in a similar way to the extinction curve of the Milky Way, by comparing spectral energy distributions of galaxies of similar intrinsic type, but attenuated in different ways. The difficulty is that the spectral energy distribution of a galaxy results from the emission by several stellar generations, and hence, the notion of intrinsic spectral type is not straightforward. Large samples of galaxies are required to average the properties of galaxies with similar spectral energy distributions and overcome this complication. This approach, described in Kinney et al. [1994], has been applied by Calzetti et al. [1994] to measure the attenuation by dust in nearby starburst galaxies. Calzetti et al. [1994] (see also Calzetti 1997, Calzetti et al. 2000) suggest that the attenuation curve in these galaxies is grayer than that observed in the Milky Way, with no 2175 Å-bump detection. For reference, the curve is represented by the blue line in Figure 2.14. The assumption behind this empirical curve is that extended regions of nearby starburst galaxies can be characterized by an average geometrical distribution of dust.

Another complication in the description of attenuation by dust in distant galaxies is the unknown distribution of dust grains. The discrepancy between the attenuation derived from Balmer lines and from the ultraviolet stellar continuum in nearby starburst galaxies suggests that, in reality, the dust is not uniformly distributed (see Fanelli et al. 1988, Calzetti 1997). The finite lifetimes of the H II regions can account for this discrepancy [Charlot and Fall, 2000]. In practice, line photons produced in

the H II regions and the non-ionizing continuum photons emitted by young stars are absorbed in the same way by dust in the thick envelopes of molecular clouds and the diffuse interstellar medium. The non-ionizing ultraviolet and optical photons from stars that live longer than H II regions are absorbed only by the diffuse interstellar medium. This is one of the features of the model by Charlot and Fall [2000]. We now describe this two-component dust model in slightly more detail, as we will use it in the next chapters.

In the Charlot and Fall [2000] model, the interstellar medium is represented by two main components: giant molecular clouds, in which new stars form and ionize H II regions (birth clouds); and a diffuse interstellar component. The stellar birth clouds have a finite lifetime $t_{\rm BC} = 10$ Myr (see Section 2.3.3), after which they are dispersed. The transmission function of this medium can be written as

$$\begin{aligned} T_\lambda^0(t') &= T_\lambda^{\rm BC} T_\lambda^{\rm ISM} \qquad \text{for} \quad t' \leqslant t_{\rm BC}\,, \\ T_\lambda^0(t') &= T_\lambda^{\rm ISM} \qquad\quad\;\; \text{for} \quad t' > t_{\rm BC}\,, \end{aligned} \tag{2.4.26}$$

where $T_\lambda^{\rm BC}$ and $T_\lambda^{\rm ISM}$ are the transmission functions of the birth clouds and the ambient interstellar medium, respectively. The transmission function can also be expressed in terms of the 'effective absorption' ($\equiv$ attenuation) optical depth of the dust, $\hat{\tau}_\lambda(t')$,

$$T_\lambda^0(t') = \exp[-\hat{\tau}_\lambda(t')]\,. \tag{2.4.27}$$

The quantity $\hat{\tau}_\lambda(t')$ depends both on the optical properties and on the spatial distribution of the dust. Similarly, one can define the effective absorption optical depths in the birth clouds and in the ambient interstellar medium,

$$\begin{aligned} T_\lambda^{\rm BC} &= \exp(-\hat{\tau}_\lambda^{\rm BC})\,, \\ T_\lambda^{\rm ISM} &= \exp(-\hat{\tau}_\lambda^{\rm ISM})\,. \end{aligned} \tag{2.4.28}$$

Charlot and Fall [2000] show that observations of starburst galaxies favor $\hat{\tau}_\lambda^{\rm ISM} \propto \lambda^{-0.7}$ in the ambient ISM. For simplicity, they adopt the same wavelength dependence for $\hat{\tau}^{\rm BC}$.[9] It is important to note that, even with these choices, the wavelength dependence of $\hat{\tau}_\lambda(t')$ in Equation 2.4.27 is a function of time (see figure 5 of Charlot and Fall [2000]). Wild et al. [2007] and da Cunha et al. [2008] suggest that a wavelength dependence $\hat{\tau}_\lambda^{\rm BC} \propto \lambda^{-1.3}$ may be more appropriate for stellar birth clouds, which can be assimilated to optically thick foreground screens around newly born stars (the slope of -1.3 corresponds to the middle range of the optical properties of dust grains between the Milky Way, the Large and the Small Magellanic Clouds; see Charlot and Fall 2000). In this thesis we write the attenuation optical depths in the stellar birth clouds and the ambient interstellar medium as

$$\hat{\tau}_\lambda^{\rm BC} = (1-\mu)\hat{\tau}_V(\lambda/5500)^{-1.3} \tag{2.4.29}$$

[9] The dependence of $\hat{\tau}^{\rm BC}$ on wavelength is not well constrained by observations, because stellar birth clouds tend to be optically thick, and hence, stars in these clouds contribute very little to the emergent radiation (except in the emission lines).

$$\hat{\tau}_{\lambda}^{\rm ISM} = \mu \hat{\tau}_{V} (\lambda/5500)^{-n}\,, \tag{2.4.30}$$

where $\mu = \hat{\tau}_{\rm V}^{\rm ISM}/(\hat{\tau}_{\rm V}^{\rm ISM}+\hat{\tau}_{\rm V}^{\rm BC})$ is the fraction of the total effective V-band absorption optical depth contributed by the dust in the ambient interstellar medium. We consider a range of slopes n in the ambient interstellar medium. This two-component model agrees remarkably well with observations of nearby galaxies.

To end this Section, it is worth mentioning that a recent empirical study by Wild et al. [2011] has revealed systematic variations of the dust attenuation curve with galaxy spectral type and orientation. This was achieved by comparing the ultraviolet, optical and near-infrared spectral energy distributions of pairs of nearby galaxies with fixed specific star formation rate, gas-phase metallicity, axial ratio and redshift but different dust contents (as measured from the Balmer decrement). The galaxy pairs were extracted from a sample of 23,000 galaxies at a median redshift of 0.07, with available photometry from GALEX, SDSS and UKIDSS. An interesting outcome of this work is the evidence for the 2175 Å bump in the ultraviolet attenuation curves of these low-redshift galaxies[10]. A main result is that the optical attenuation curve appears to vary strongly with galaxy inclination, in the sense that it is steeper in face-on than in edge-on galaxies. Such a variation is supported by the models of Rocha et al. [2008] (see also Tuffs et al. 2004 and Pierini et al. 2004). In these models, the effects of dust on the emission from galaxies is analyzed by coupling the Monte Carlo radiative transfer code SUNRISE [Jonsson, 2006] with hydrodynamic galaxy simulations. The models predict that the dependence of attenuation on inclination can be described by a quadratic law (edge-on galaxies having larger optical depth than face-on ones) and that the attenuation curve flattens when the optical depth increases (see also Chevallard et al. in preparation).

2.4.2.4 Dust in intermediate- and high-redshift galaxies

Galaxies at intermediate redshifts ($z \sim$2-3) present a wide variety of dust contents, often similar to local starburst galaxies [Reddy et al., 2012]. Accurate spectroscopic observations of ultraviolet-luminous galaxies at $z = 2$ suggest that the most massive galaxies are also the most metal-rich and dusty [van Dokkum et al., 2003, Noll and Pierini, 2005]. Noll and Pierini [2005] and Noll et al. [2009] find that galaxies at intermediate redshifts exhibit attenuation curves similar to the extinction curves of the Milky Way and the Magellanic Clouds, with in some cases a strong 2175 Å bump. The strength of this bump does not seem to depend on metallicity, but rather on the spatial distribution of interstellar dust.

At even higher redshift, little is known about the attenuation curve of the dust. Galaxies at $z > 4$ tend to have fairly blue spectral energy distributions (see for example, Erb et al. 2010; Bouwens et al. 2011). This is usually interpreted in terms

[10]This has also been detected in a different sample of low-redshift galaxies by Conroy et al. [2010a].

of a dominant population of young and low-metallicity stars, combined with an almost dust-free environment. Although this interpretation can be reasonable when observing the Universe at very young ages, selection effects and uncertainties in the photometric measurements could bias these results [Dunlop et al., 2012].

Theoretical expectations are also quite uncertain. In recent simulations of Lyman-alpha emitters at $3 < z < 5$, Garel et al. [2012] introduce a redshift dependence of the dust-to-gas ratio in galaxies. Following Hatton et al. [2003] (see also Guiderdoni and Rocca-Volmerange 1987), they write the attenuation curve as

$$\tau_\lambda = \left(\frac{A_\lambda}{A_V}\right)_{Z_\odot} \left(\frac{Z}{Z_\odot}\right)^{1.35} \left(\frac{N_\mathrm{H}}{2.1 \times 10^{21}}\right) f(z)\,, \tag{2.4.31}$$

where $(A_\lambda/A_V)_{Z_\odot}$ is the extinction curve for solar metallicity from Mathis et al. [1983], Z is the metallicity of interstellar gas, N_H is the H I column density and $f(z) = (1+z)^{-1/2}$ is the scaling of the dust-to-gas ratio with redshift. This prescription however is speculative.

2.5 Absorption by the Intergalactic Medium

The intergalactic medium is filled by a gas composed mainly of hydrogen at very low density (of the order of 0.1 atoms m^{-3}, Tytler et al. 2004). Throughout the epoch of structure formation, this gas becomes clumpy and can be attracted and captured by the gravitational fields of clusters and galaxies. Matter from galaxies can also be returned to the intergalactic medium (by stellar winds, supernova explosions or galaxy interactions), contributing to the enrichment of this gas.

The ionizing radiation emitted by stars and active galactic nuclei in galaxies causes the intergalactic medium to be mainly ionized. This ionized component cannot be observed directly, because the emitted radiation is too weak. In contrast, the minor neutral component of the intergalactic medium, which still remains, can be detected because it contributes to the obscuration of galaxies at cosmological distances. The main evidence for a clumpy and neutral intergalactic medium is the presence of absorption features in the far-ultraviolet spectra of distant quasars. These features correspond to Lyman-α absorption in the rest frame of intervening clouds (*absorbers*) of the redshifted ionizing spectrum of the background quasar (photons emitted by the source at $\lambda_\mathrm{em} < \lambda_\alpha$ are redshifted to the Lyman-α wavelength for clouds along the line of sight). Since these clouds are distributed at different distances along the line of sight, the absorption features are detected at different redshifts.

There are two broad populations of absorbers: Lyman-α forest clouds and metal-line systems. The optical depth of these systems depends on the distribution of the absorbers in column density and redshift (see Madau 1995 for details). The combined effect of many absorption lines produces a characteristic staircase profile in the transmitted spectral energy distribution. We show in Figure 2.15 3 examples of the transmission coefficient, $\exp(\tau_\mathrm{eff})$, at redshifts 3.0, 4.5 and 6.0. At each redshift,

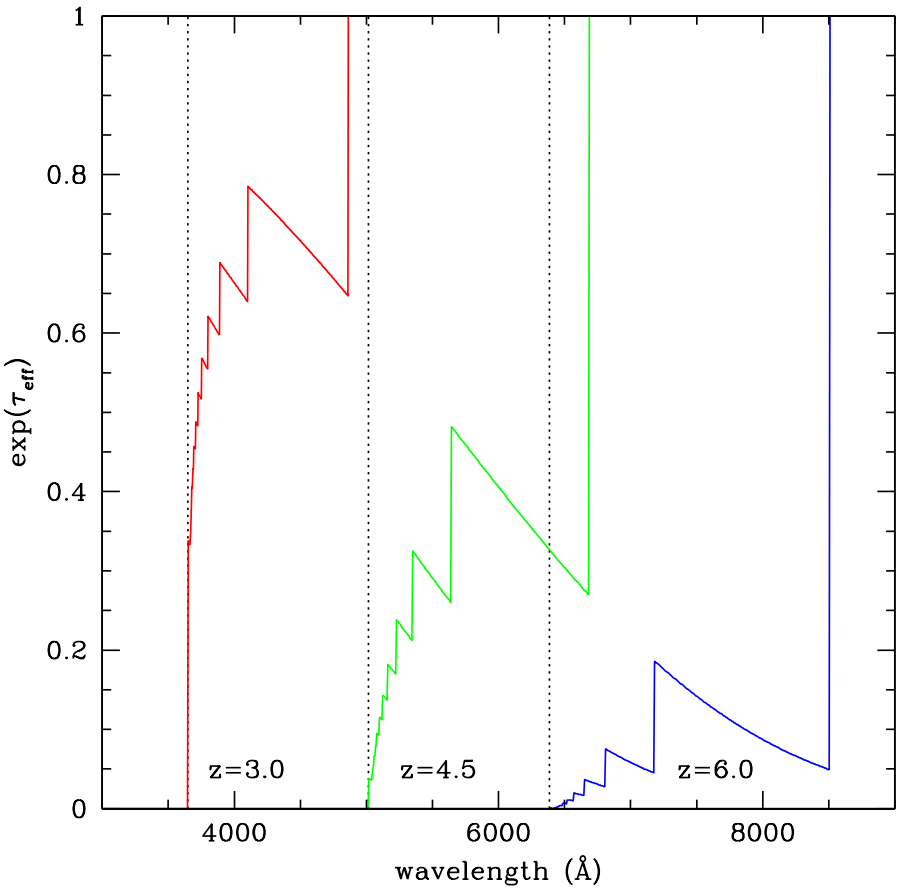

Figure 2.15: Trasmission coefficient in the intergalactic medium at three different redshifts: $z = 3.0$ (red line), $z = 4.5$ (green line), $z = 6.0$ (blue line). Dotted lines represent the Lyman limit (912Å) at the three different redshifts.

the dotted line indicates the Lyman limit (912Å). This cosmic Lyman break is a powerful tool to identify high-redshift galaxies.

2.6 Summary

In this Chapter, we have introduced the different tools involved in the modeling of the spectral energy distributions of galaxies.

In Section 2.1 we presented different approaches to compute star formation and chemical enrichment histories of galaxies. There are two main approaches:

- drawing idealized representations of star formation histories at fixed metallicity (single bursts of star formation, constant star formation rate, exponentially declining star formation rate);

- drawing physically motivated star formation histories and the consistent evolution of the metal content. This can be achieved by combining cosmological N-body simulations (to trace the evolution of dark matter haloes) with either semi-analytical recipes or gas-dynamics codes (to simulate the evolution of baryons in the dark-matter structures).

An advantage of cosmological simulations is the treatment of gas physics. Feedback from supernovae and AGNs is included to regulate gas cooling and the star formation. A disadvantage is that even the most sophisticated simulations present some limitations, such as a too large predicted number of red, non-star forming galaxies.

In Section 2.2, we described the techniques used to model and interpret the emission from stellar populations. The most widely used one is evolutionary population synthesis, which is based on the property that stellar populations can be expanded in series of instantaneous starbursts, called *simple stellar populations*. The main ingredients required for the modeling of stellar populations are:

- a prescription for the mass distribution of newly born stars;
- evolutionary tracks of stars of given initial mass and chemical composition;
- libraries of stellar spectra, to describe the emission from each star in the H-R diagram.

The light produced by stars interacts with gas and dust in the interstellar medium before escaping from the galaxy. To fully interpret the multi-wavelength emission from galaxies, we thus need to also model the influence of gas and dust on the emergent spectral energy distribution. In Section 2.3, we described the different phases of the interstellar medium and in particular the signatures of ionized gas (mainly H II regions) in the spectral energy distributions of galaxies. To model this *nebular emission* we include two components:

- the nebular continuum, which arises from free-bound and free-free interactions between electrons and ions in the ionized gas;
- the nebular emission lines, which are produced in
  - cascade transitions after recombinations of ions and electrons on excited levels,
  - cascade transitions after collisional excitation of metal ions.

In Section 2.4, we described the influence of dust on the spectral energy distributions of galaxies. Dust grains absorb and scatter the ultraviolet and optical radiation from stars and gas and reradiate this in the infrared. The main uncertainties in the modeling of the effects of dust in spectral models of galaxies arise from the unknown composition and spatial distribution of dust grains. We described:

- the extinction curves of the Milky Way and the Small and Large Magellanic Clouds, addressing in particular

  - the different ultraviolet slopes,
  - the presence or absence of the ultraviolet 2175 Å bump;

- the attenuation law in distant galaxies, addressing the main sources of uncertainties;

- dust in intermediate- and high-redshift galaxies.

Finally, in Section 2.5, we addressed the influence of the intergalactic medium on the spectral energy distributions of distant galaxies. Neutral hydrogen in the intergalactic medium absorbs the high-energy photons emitted by stars and active galactic nuclei. These characteristic absorption features depend on the column-density and redshift distributions of absorbing clouds and can be used as a tool to identify high-redshift galaxies.

Chapter 3

Relative merits of different types of rest-frame optical observations to constrain galaxy physical parameters

---

This Chapter is extracted from the paper *Relative merits of different types of rest-frame optical observations to constrain galaxy physical parameters* by Camilla Pacifici, Stèphane Charlot, Jèrèmy Blaizot and Jarle Brinchmann, published in the Monthly Notices of the Royal Astronomical Society.

3.1 Introduction

The light emitted by stars and the interstellar medium in galaxies contains important clues about the physical processes that drive galaxy evolution. Different strategies have been adopted to survey the spectral energy distributions of large samples of galaxies to gather these clues. High-quality photometry from surveys such as the Sloan Digital Sky Survey (SDSS, York et al. 2000), the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS, Goranova et al. 2009) and the Cosmic Evolution Survey (COSMOS, Scoville et al. 2007) has brought important information about the rest-frame optical emission from millions of galaxies out to redshifts of a few. This is supplemented by medium-resolution spectroscopy of over a million local galaxies from the SDSS and the Two-degree Field Galaxy Redshift Survey (2dFGRS, Colless et al. 2001), and low- and medium-resolution spectroscopy of tens of thousands of galaxies at higher redshift from surveys such as the VIMOS VLT Deep Survey (VVDS, Le Fèvre et al. 2005) and the Deep Extragalactic Evolutionary Probe (DEEP2, Davis et al. 2003). The interpretation of the high-quality data gathered by these surveys has triggered the development of sophisticated models of galaxy spectral evolution (e.g., Bruzual and Charlot 2003, González Delgado et al. 2005, Maraston 2005, Coelho et al. 2007, Walcher et al. 2009, Vazdekis et al. 2010, Thomas et al. 2011; see also Charlot and Longhetti 2001, Kewley and Dopita 2002). These models can help us constrain the star formation and chemical enrichment histories of galaxies through the analysis of different types of photometric and spectroscopic observations. So far, however, the relative merits of different observational approaches to constrain basic physical properties of galaxies have not been quanti-

fied in detail.

Recent studies have brought important insight into the efficiency of different types of observations to constrain physical quantities such as star formation history, metallicity and dust content from observations of the rest-frame optical emission of galaxies. Multiband photometry has proved valuable to constrain galaxy redshifts and basic quantities such as stellar mass-to-light ratio (provided a stellar initial mass function; hereafter IMF) and, when combined with rest-frame ultraviolet and near-infrared constraints, star formation rate and attenuation by dust (e.g., Bell and de Jong 2001, Salim et al. 2005, 2007, Ilbert et al. 2006, Schaerer and de Barros 2010; but see Wuyts et al. 2009, Conroy et al. 2010b). Medium-resolution spectroscopy allows separate analyses of nebular-emission and stellar-absorption features, and hence, more accurate constraints on the star formation history, stellar and interstellar metallicities, attenuation by dust and nuclear activity (e.g., Kauffmann et al. 2003a, Brinchmann et al. 2004, Heckman et al. 2004, Tremonti et al. 2004, Cid Fernandes et al. 2005, Gallazzi et al. 2005, Panter et al. 2007, Wild et al. 2007). At low spectral resolution, the difficulty of distinguishing between the signatures of nebular-emission and stellar-absorption features, especially around H-Balmer lines, renders analyses more complicated [e.g., Lamareille et al., 2006]. This difficulty could be overcome with the ability to interpret simultaneously the stellar and nebular emission from galaxies. In this case, the relative merits of multiband photometry and low- and medium-resolution spectroscopy to constrain galaxy physical parameters would need to be reexamined.

In this Chapter, we develop a new approach to interpret the combined stellar and nebular emission from galaxies in wide ranges of multi-wavelength observations. To achieve this, we appeal to a set of state-of-the-art models of galaxy star formation and chemical enrichment histories, stellar population synthesis, nebular emission and attenuation by dust. We use these models to generate a comprehensive library of several million galaxy spectral energy distributions. This library can be used to retrieve probability density functions of physical parameters from the Bayesian analysis of any observed spectral energy distribution. We focus here on the constraints set by 5-band *ugriz* photometry and low- and medium-resolution spectroscopy at rest wavelengths λ=3600–7400 Å on a few basic physical parameters of galaxies: the observer-frame absolute r-band stellar mass-to-light ratio (M_*/L_r); the fraction of a current galaxy stellar mass formed during the last 2.5 Gyr ($f_{\rm SFH}$); the specific star formation rate ($\psi_{\rm S}$); the gas-phase oxygen abundance $[12+\log({\rm O/H})]$; the total effective V-band absorption optical depth of the dust ($\hat{\tau}_V$); and the fraction of this arising from dust in the ambient interstellar medium (hereafter ISM; μ). These parameters cannot be known a priori for any galaxy sample. To assess the accuracy to which they can be retrieved from different types of observations, we therefore simulate 'pseudo-observations' by convolving the spectral energy distributions of models with known parameters with appropriate instrument responses and then applying artificial noise to mimic true observations. This approach allows us to make accurate predictions for the optimistic case in which the models we rely on are good approximations of true galaxies. A most notable outcome of our study is that the

combined analysis of stellar and nebular emission in low-resolution (50 Å FWHM), high-quality (median signal-to-noise ratio per pixel $\overline{\mathrm{S/N}} = 20$) galaxy spectra provides valuable constraints on all the physical parameters mentioned above. We also explore how the constraints from different types of photometric and spectroscopic observations depend on signal-to-noise ratio and the inclusion of nebular emission in the analysis. Our main results are summarized in Table 3.2 (Section 3.5).

We describe our approach to constrain galaxy physical parameters from the combined analysis of stellar and nebular emission in Section 3.2 below. In Section 3.3, we investigate the accuracy to which the parameters M_*/L_r, f_{SFH}, ψ_{S}, $12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and μ can be constrained from different types of photometric and spectroscopic observations. Then, in Section 3.4, we compare the constraints derived from the analysis of the medium-resolution spectra of 12,660 SDSS star-forming galaxies using our approach with those obtained in previous separate analyses of the stellar and nebular emission. We also investigate the influence of the prior distributions of physical parameters on the retrieved probability density functions. Our conclusions are summarized in Section 3.5. Throughout this Chapter, we adopt the standard cosmology $\Omega_{\mathrm{M}} = 0.3$, $\Omega_\Lambda = 0.7$ and $h = 0.7$.

3.2 Modeling approach

In this section, we describe our approach to assess the retrievability of galaxy physical parameters from the combined analysis of stellar and nebular emission in wide ranges of multi-wavelength observations. We first build a comprehensive library of star formation and chemical enrichment histories of galaxies (Section 3.2.1). We appeal to state-of-the-art models of spectral evolution to compute the emission from stars and gas and the attenuation by dust in these galaxies (Section 3.2.2). This allows us to assemble a large library of galaxy spectral energy distributions (Section 3.2.3), which we can use to retrieve likelihood distributions of physical parameters from the Bayesian analysis of any type of observed spectral energy distribution (Section 3.2.4).

3.2.1 Library of star formation and chemical enrichment histories

We build a comprehensive library of galaxy star formation and chemical enrichment histories by performing a semi-analytic post-treatment of the Millennium cosmological simulation of Springel et al. [2005]. This large-scale cosmological simulation follows the growth, interaction and merging history of dark matter haloes from redshift $z = 127$ to the present time.[1] We adopt the simple semi-analytic recipes of De Lucia and Blaizot [2007, see also Croton et al. 2006] to follow star formation and the associated metal production by gas falling into these dark matter haloes. We

[1] In the Millennium simulation, dark matter is represented as a set of about 10^{10} point particles, which interact gravitationally in a cubic region of side $500h^{-1}$ Mpc. Individual particles have a mass of $8.6 \times 10^8 h^{-1} M_\odot$, and the detection threshold for a dark matter halo is 20 particles, i.e. about $1.7 \times 10^{10} h^{-1} M_\odot$ [see Springel et al., 2005, for detail].

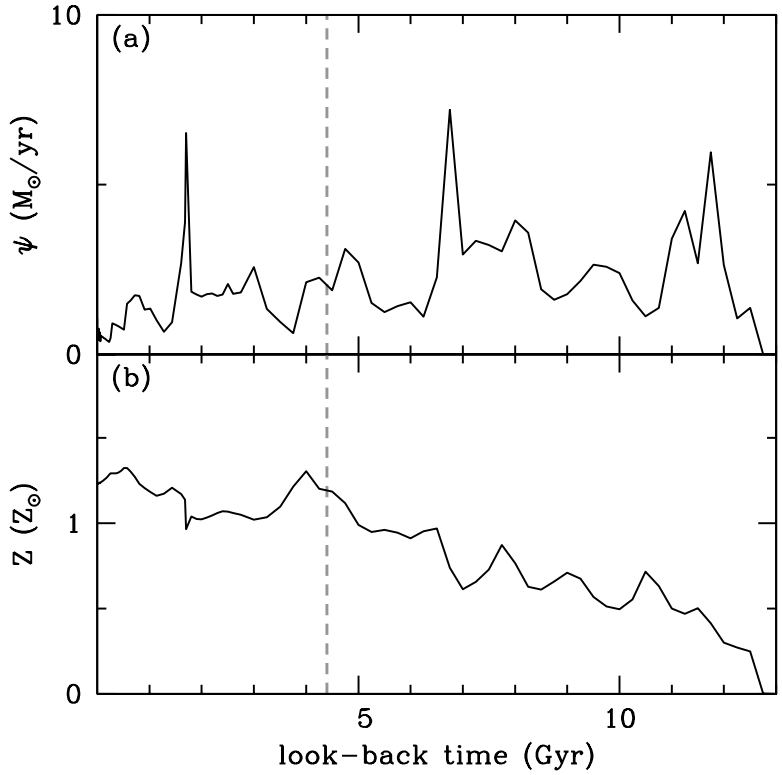

Figure 3.1: Example of galaxy star formation and chemical enrichment histories inferred from the semi-analytic post-treatment of the Millennium cosmological simulation: (*a*) star formation rate, ψ, and (*b*) interstellar metallicity, Z, plotted as a function of look-back time, for a galaxy with a present-day stellar mass of $1.5 \times 10^{10} M_{\odot}$. The vertical dashed line marks the evolutionary stage at which the galaxy is looked at in Fig. 3.3 (corresponding to a galaxy age of 8.4 Gyr).

compute in this way the star formation and chemical enrichment histories of 500,000 galaxies with present-day stellar masses comprised between $M_{\star} \approx 5 \times 10^{9} h^{-1} M_{\odot}$ (corresponding roughly to the mass resolution of the simulation; Blaizot et al. 2004) and $5 \times 10^{11} h^{-1} M_{\odot}$. Figure 3.1 shows an example of star formation and chemical enrichment histories of a galaxy with present-day stellar mass $1.5 \times 10^{10} M_{\odot}$. The spikes in both curves reflect the various episodes of interactions and mergers that shaped the evolution of the galaxy. As shown by Wild et al. [2008], star formation histories computed in this way account well for the observed properties of galaxies at different redshifts in the VVDS ($0.5 \lesssim z \lesssim 1$) and SDSS ($0.05 \lesssim z \lesssim 0.1$) samples.

About 35 percent of the 500,000 galaxies generated from the semi-analytic post-treatment of the Millennium simulation do not form stars at $z = 0$. These contain a negligible amount of interstellar gas and correspond to the population of present-

day early-type galaxies. The remaining galaxies span a relatively narrow range of star formation properties. Their specific star formation rates peak around $\psi_{\rm S} \approx 0.1\,{\rm Gyr}^{-1}$, with a dispersion of only a factor of 2. Furthermore, the interstellar metallicity Z in these galaxies correlates significantly with $\psi_{\rm S}$, in the sense that the most metal-poor galaxies form stars most actively. The galaxies also have roughly similar recent histories of star formation, as quantified by the star formation history parameter $f_{\rm SFH}$, which we define as the fraction of the current galaxy stellar mass formed during the last 2.5 Gyr (this time lag corresponds to the main-sequence lifetime of A-F stars with strong Balmer absorption features). This parameter has a narrow distribution centred on $f_{\rm SFH} \approx 0.15$, with a dispersion of only 0.1. The distributions of $\psi_{\rm S}$ and $f_{\rm SFH}$ are compatible with roughly constant (overall) star formation over the past 2.5 Gyr in galaxies more massive than $5 \times 10^9 h^{-1} M_\odot$, given that 40 percent of the stellar mass formed during this period is returned to the ISM in the form of winds and supernovae [e.g. Bruzual and Charlot, 2003]. The above distributions of $\psi_{\rm S}$, Z and $f_{\rm SFH}$ are consistent with the observed properties of nearby, star-forming SDSS galaxies [e.g. Kauffmann et al., 2003a, Brinchmann et al., 2004, Tremonti et al., 2004].

The relatively limited ranges of star formation histories and interstellar metallicities described above, while appropriate to interpret observations of nearby SDSS galaxies more massive than $5 \times 10^9 h^{-1} M_\odot$, could bias statistical estimates of $\psi_{\rm S}$, Z and $f_{\rm SFH}$ in galaxies drawn from other samples or in different mass ranges (e.g. Weinmann et al. 2006; see also Section 3.4.2 below). To avoid such biases, we broaden the original prior distributions of these parameters predicted by the semi-analytic recipes of De Lucia and Blaizot [2007] in 2 ways. Firstly, we widen the range in $f_{\rm SFH}$ by randomly drawing the evolutionary stage (i.e. galaxy age) at which a galaxy is looked at in the library. In practice, we draw this stage uniformly in redshift between $z = 0$ and 1.7, i.e. the upper limit beyond which $f_{\rm SFH}$ reaches unity for a substantial fraction of the galaxies (Fig. 3.4b below; this resampling makes the proportion of early-type galaxies drop from 35 to about 10 percent in the library). Secondly, for galaxies with ongoing star formation at the selected stage, we broaden the distributions of the current star formation rate and the current interstellar metallicity, as follows (we define here as 'current' the average of a quantity over a period of 10 Myr before a galaxy is looked at). We redistribute the specific star formation rate over the interval $-2 < \log(\psi_{\rm S}/{\rm Gyr}^{-1}) < 1$ using the probability density function $p[\log(\psi_{\rm S}/{\rm Gyr}^{-1})] = 0.05[1.5^5 - \left|\log(\psi_{\rm S}/{\rm Gyr}^{-1}) + 0.5\right|^5]$, which is approximately uniform over the logarithmic range from $\psi_{\rm S} = 0.03$ to $3\,{\rm Gyr}^{-1}$ and drops to zero at $\psi_{\rm S} = 0.01$ and $10\,{\rm Gyr}^{-1}$ (Fig. 3.4c). We redistribute the current interstellar metallicity Z logarithmically between 0.1 and 3.5 times solar ($Z_\odot = 0.017$).

The random resampling of $\psi_{\rm S}$ and Z for star-forming galaxies in the library is justified in part by the stochastic nature of star formation and chemical evolution in hierarchical scenarios of galaxy formation. Also, since the changes impact only the last 10 Myr of evolution, they have a negligible influence on the global star formation and chemical enrichment histories derived from the semi-analytic post-treatment of the Millennium simulation. In return, the use of wider and uncorrelated prior

distributions of $\psi_{\rm S}$ and Z, together with the broader sampling in galaxy age (and $f_{\rm SFH}$), greatly enhances the usefulness of the model library to explore and interpret the spectral properties of observed galaxy populations.

3.2.2 Galaxy spectral modeling

We now describe the way in which we compute the emission from stars and gas and the attenuation by dust in the model galaxies generated in the previous section. The luminosity per unit wavelength $L_\lambda(t)$ emerging at time t from a galaxy can be expressed as [see, e.g., Charlot and Longhetti, 2001]

$$L_\lambda(t) = \int_0^t dt' \, \psi(t-t') \, S_\lambda[t', Z(t-t')] \, T_\lambda(t, t') \, , \tag{3.2.1}$$

where $\psi(t-t')$ is the star formation rate at time $t-t'$, $S_\lambda[t', Z(t-t')]$ is the luminosity emitted per unit wavelength and per unit mass by a stellar generation of age t' and metallicity $Z(t-t')$, and $T_\lambda(t, t')$ is the transmission function of the interstellar medium, defined as the fraction of the radiation produced at wavelength λ at time t by a generation of stars of age t' that is transferred by the ISM. Following Charlot and Longhetti [2001], we write $T_\lambda(t, t')$ as the product of the transmission functions of the ionized gas, $T_\lambda^+(t, t')$, and the neutral ISM, $T_\lambda^0(t, t')$,

$$T_\lambda(t, t') = T_\lambda^+(t, t') T_\lambda^0(t, t') \, . \tag{3.2.2}$$

In the next paragraphs, we describe the prescriptions we adopt for the functions S_λ, T_λ^+ and T_λ^0.

3.2.2.1 Stellar emission

We compute the luminosity $S_\lambda(t, Z)$ emitted per unit wavelength and per unit mass by a stellar generation of age t and metallicity Z using the latest version of Bruzual and Charlot [2003] stellar population synthesis code (Charlot & Bruzual, in preparation). This code predicts the spectral evolution of stellar populations at wavelengths from 91 Å to 160 μm and at ages between 1×10^5 and 2×10^{10} yr, for different metallicities (from about $0.005 Z_\odot$ to $4 Z_\odot$), initial mass functions and star formation histories. We use the most recent version of the code, which incorporates a new library of observed stellar spectra [Sánchez-Blázquez et al., 2006] and new prescriptions for the evolution of stars less massive than $20 \, M_\odot$ [Bertelli et al., 2008, 2009] and for the thermally pulsating asymptotic giant branch (TP-AGB) evolution of low- and intermediate-mass stars [Marigo et al., 2008]. The main effect of the revised TP-AGB prescription is to improve the predicted near-infrared colours of intermediate-age stellar populations [e.g., González-López lira et al., 2010]. In all applications throughout this Chapter, we adopt the Galactic-disc IMF of Chabrier [2003].

3.2.2.2 Nebular emission

A main feature of our study is that we account for the contribution by ionized gas to the emission from galaxies. We follow the prescription of Charlot and Longhetti [2001] to compute the transmission function $T_\lambda^+(t,t')$ of the gas photoionized at time t by stars of age t' in a galaxy. This consists in adopting effective (i.e. galaxy-wide) parameters to describe the ensemble of H II regions and the diffuse gas ionized by young stars throughout the galaxy. The main adjustable parameters are the interstellar metallicity, Z, the zero-age ionization parameter, U_0 (which characterizes the volume-averaged ratio of ionizing-photon to gas densities at age $t'=0$), and the dust-to-metal (mass) ratio, ξ_d (which characterizes the depletion of metals on to dust grains) of the photoionized gas.[2]

Following Charlot and Longhetti [2001], we neglect the contribution by stars older than 10 Myr to nebular emission, i.e., we write

$$T_\lambda^+(t,t') = 1, \qquad \text{for} \quad t' > 10\,\mathrm{Myr}\,. \tag{3.2.3}$$

We use the standard photoionization code CLOUDY [Ferland, 1996] to compute the transmission function at earlier ages, assuming that galaxies are ionization bounded. For $t' < 10^7$yr, therefore, $T_\lambda^+(t,t')$ is close to zero at wavelengths blueward of the H-Lyman limit and greater than unity at wavelengths corresponding to emission lines (we record the luminosities of the 109 most prominent emission lines at all wavelengths; we also include the recombination continuum radiation). When computing $T_\lambda^+(t,t')$ in equations (3.2.1)–(3.2.2), we take the metallicity of the photoionized gas to be the current metallicity $Z(t)$ of the galaxy. We randomly draw the dust-to-metal ratio from a uniform prior distribution between $\xi_\mathrm{d}=0.1$ and 0.5, and the zero-age ionization parameter from a logarithmic prior distribution between $\log U_0 = -3.5$ and -1.5. We fix all other parameters of the photoionization code at the standard values favored by Charlot and Longhetti [2001, see their paper for detail].

In Fig. 3.2, we compare the luminosities of prominent optical emission lines obtained in this way for galaxies with different Z, U_0 and ξ_d, with high-quality observations of a sample of 28,075 star-forming galaxies from the SDSS Data Release 7 [DR7; Abazajian et al., 2009]. The different panels show the relations between different luminosity ratios constructed with the [O II]$\lambda\lambda$3726, 3729 (hereafter [O II]λ3727), Hβ, [O III]λ5007, Hα, [N II]λ6584 and [S II]$\lambda\lambda$6716, 6731 lines. We selected the star-forming galaxies of Fig. 3.2 from the SDSS DR7 by requiring high signal-to-noise ratio (S/N > 10) in all lines and excluding galaxies in which the emission could be contaminated by an active galactic nucleus (AGN; we applied the conservative criterion of Kauffmann et al. 2003b to reject AGNs from the standard Baldwin et al. 1981 line-diagnostic diagram reproduced in Fig. 3.2a). The excellent agreement between models and SDSS observations in Fig. 3.2 is similar to that obtained by Brinchmann et al. [2004] using the original Charlot and Longhetti [2001] models. This allows the derivation of independent constraints on Z, U_0 and ξ_d from the different line luminosity ratios. For simplicity, in what follows, we adopt a fixed intrinsic line velocity

[2] These effective parameters were denoted by $\tilde{Z}$, $\langle\tilde{U}_0\rangle$ and $\tilde{\xi}_\mathrm{d}$ in Charlot and Longhetti [2001].

dispersion of $100\,\mathrm{km\,s^{-1}}$ (typical of SDSS star-forming galaxies; e.g. Brinchmann et al. 2004) when computing the contribution by ionized gas to the emission from galaxies.

3.2.2.3 Attenuation by dust

The attenuation by dust of the emission from a galaxy depends on several factors, including the amount and structure of the interstellar gas, the metallicity and physical conditions of this gas and the orientation of the galaxy. Two galaxies with identical star formation and chemical enrichment histories but different sizes (and hence different total gas masses), ISM structures and orientations can have widely different dust attenuation optical depths, even though their dust-free spectral energy distributions are the same (modulo a mass scaling factor). We assume here for simplicity that the dust attenuation optical depth of a galaxy does not depend on other global galaxy parameters, such as the age t, the specific star formation rate ψ_{S} and the gas metallicity Z. We therefore write the transmission function of the neutral ISM in equations (3.2.1)–(3.2.2) as simply a function of stellar age t',

$$T_\lambda^0(t,t') = T_\lambda^0(t')\,. \tag{3.2.4}$$

In the ISM, stars are born in giant molecular clouds, which dissipate on a time-scale of the order of 10 Myr [e.g., Blitz and Shu, 1980, Kawamura et al., 2009]. Thus the emission from stars younger than this time-scale and from the photoionized gas is more attenuated than that from older stars. We account for this effect using the simple dust model of Charlot and Fall [2000], who express the transmission function of the neutral gas as[3]

$$T_\lambda^0(t') = \exp\left[-\hat{\tau}_\lambda(t')\right]\,. \tag{3.2.5}$$

Here $\hat{\tau}_\lambda(t')$ is the 'effective absorption' (or attenuation) optical depth of the dust seen by stars of age t'. This represents the average absorption of photons emitted in all directions by stars in all locations within the galaxy, including the effects of scattering on the path lengths of the photons before they are absorbed. Following Charlot and Fall [2000], we write

$$\hat{\tau}_\lambda(t') = \begin{cases} \hat{\tau}_\lambda^{\mathrm{BC}} + \hat{\tau}_\lambda^{\mathrm{ISM}} & \text{for} \quad t' \leqslant 10\,\mathrm{Myr}\,, \\ \hat{\tau}_\lambda^{\mathrm{ISM}} & \text{for} \quad t' > 10\,\mathrm{Myr}\,, \end{cases} \tag{3.2.6}$$

where $\hat{\tau}_\lambda^{\mathrm{BC}}$ and $\hat{\tau}_\lambda^{\mathrm{ISM}}$ are the effective absorption optical depths of the stellar birth clouds and the ambient (i.e. diffuse) ISM, respectively. We note that, since the intrinsic spectral energy distributions of young and old stars evolve in time, equation (3.2.6) implies that the global attenuation curve of a galaxy containing several generations of stars is complicated function of galaxy age, even though $\hat{\tau}_\lambda$ is a simple function of t' (see, e.g., fig. 5 of Charlot and Fall 2000).

[3]The function denoted here by $T_\lambda^0(t')$ is equivalent to the function $T_\lambda(t')$ defined by equation (6) of Charlot and Fall [2000], for $f = 0.0$ in their notation.

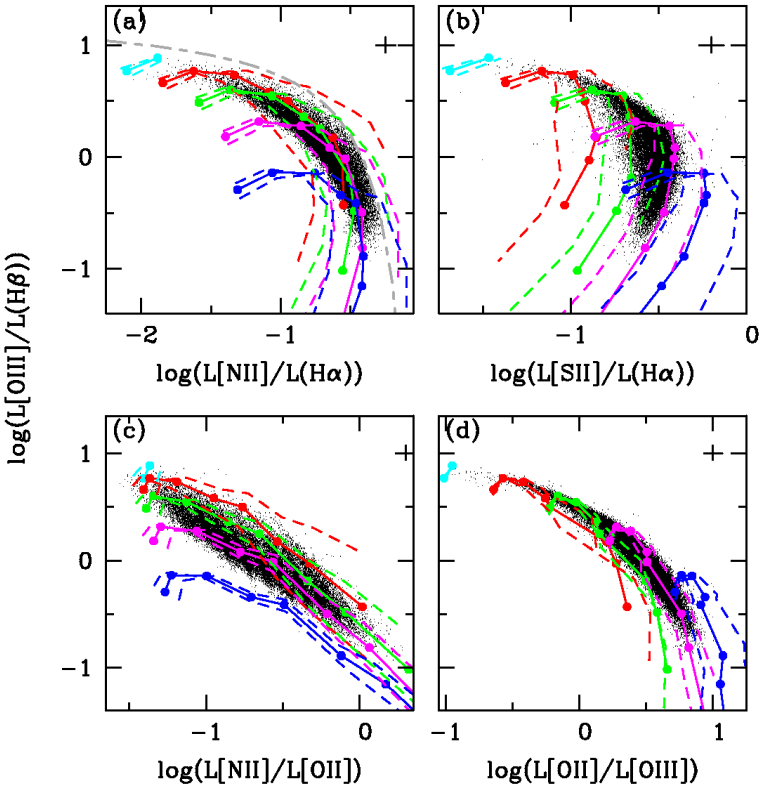

Figure 3.2: Emission-line luminosities of galaxies computed using the models described in Section 3.2.2.2 (lines), compared with high-quality observations of a sample of 28,075 star-forming galaxies from the SDSS DR7 (black dots). The models assume constant star formation over the past 10 Myr. The data are corrected for attenuation by dust as described in Brinchmann et al. [2004]. (*a*) $L([\mathrm{O\,III}]\lambda5007)/L(\mathrm{H}\beta)$ versus $L([\mathrm{N\,II}]\lambda6584)/L(\mathrm{H}\alpha)$. (*b*) $L([\mathrm{O\,III}]\lambda5007)/L(\mathrm{H}\beta)$ versus $L([\mathrm{S\,II}]\lambda\lambda6716,6731)/L(\mathrm{H}\alpha)$. (*c*) $L([\mathrm{O\,III}]\lambda5007)/L(\mathrm{H}\beta)$ versus $L([\mathrm{N\,II}]\lambda6584)/L([\mathrm{O\,II}]\lambda3727)$. (*d*) $L([\mathrm{O\,III}]\lambda5007)/L(\mathrm{H}\beta)$ versus $L([\mathrm{O\,II}]\lambda3727)/L([\mathrm{O\,III}]\lambda5007)$. In each panel, lines of different colours refer to models with different zero-age ionization parameter $\log U_0$ (cyan: -1.5; red: -2.0; green: -2.5; magenta: -3.0; blue: -3.5). At fixed $\log U_0$, the lower dashed line corresponds to models with dust-to-metal ratio $\xi_\mathrm{d} = 0.1$, the solid line to models with $\xi_\mathrm{d} = 0.3$ and the upper dashed line to models with $\xi_\mathrm{d} = 0.5$. Along each line, dots mark the positions of models with different gas metallicity Z (0.10, 0.20, 0.50, 0.75, 1.00, 1.50, 2.00 and 3.00 times $Z_\odot$, from left to right; only low-metallicity models are considered for $\log U_0 = -1.5$). In (*a*), the gray long-short-dashed line shows the Kauffmann et al. [2003b] criterion to separate star-forming galaxies from AGNs.

We must also specify the dependence of $\hat{\tau}_\lambda^{\rm BC}$ and $\hat{\tau}_\lambda^{\rm ISM}$ on wavelength. For stellar birth clouds, we adopt a simple power-law of index -1.3 appropriate for optically thick clouds with dust properties in the middle range of the Milky Way, the Large and the Small Magellanic Clouds [da Cunha et al., 2008]. In the ambient ISM, the shape of the effective absorption curve is more uncertain, as the spatial distribution of dust relative to the stars can vary widely from galaxy to galaxy. Also, while $\hat{\tau}_\lambda^{\rm ISM}$ represents an average optical depth along rays emanating in all directions from all stars, in practice, the attenuation will depend to some degree on the angle under which a galaxy is observed. Thus we write

$$\hat{\tau}_\lambda^{\rm BC} = (1-\mu)\hat{\tau}_V \, (\lambda/5500\,)^{-1.3} \, , \tag{3.2.7}$$

$$\hat{\tau}_\lambda^{\rm ISM} = \mu\hat{\tau}_V \, (\lambda/5500\,)^{-n} \, . \tag{3.2.8}$$

In these expressions, $\hat{\tau}_V = \hat{\tau}_V^{\rm BC} + \hat{\tau}_V^{\rm ISM}$ is the total effective V-band absorption optical depth of the dust seen by young stars inside birth clouds, and $\mu = \hat{\tau}_V^{\rm ISM}/(\hat{\tau}_V^{\rm BC} + \hat{\tau}_V^{\rm ISM})$ is the fraction of this contributed by dust in the ambient ISM. A value of μ around 0.3 and a typical slope $n = 0.7$ in equation (3.2.8) account remarkably well for several observed mean relations between the ultraviolet, optical and infrared properties of nearby galaxies (Charlot and Fall 2000, Kong et al. 2004, Salim et al. 2007, da Cunha et al. 2008; Wild et al. 2011a). More refined investigations using sophisticated dust models suggest that, in detail, the attenuation curve should flatten when the optical depth increases [Pierini et al., 2004, Tuffs et al., 2004, Rocha et al., 2008]. This is supported by the typically greyer optical attenuation curves observed in nearby edge-on galaxies compared to face-on ones (Wild et al. 2011b). A recent analysis shows that the trend favored by current models and observations can be reasonably well approximated by a simple expression of the slope of the attenuation curve of the form

$$n \approx \frac{2}{1+\mu\hat{\tau}_V}, \tag{3.2.9}$$

with a scatter of about 20 percent at fixed $\mu\hat{\tau}_V$ (Chevallard et al., in preparation).

To compute $T_\lambda^0(t,t')$ in equations (3.2.1)–(3.2.2), we therefore proceed as follows. We first draw $\hat{\tau}_V$ between 0.01 and 4 from the probability density function $p(\hat{\tau}_V) = 0.18\arctan[5(4-\hat{\tau}_V)]$, which is approximately uniform over the range from $\hat{\tau}_V = 0.01$ to 3 and drops to zero at $\hat{\tau}_V = 4$ (see Fig. 3.4e below). We further draw μ from a uniform prior distribution between 0.1 and 0.7, and we draw n uniformly over an interval $\Delta n = 0.3$ centred on the mean value predicted by equation (3.2.9). Then, we report the selected $\hat{\tau}_V$, μ and n in equations (3.2.4)–(3.2.8). This prescription accounts for a much broader range of plausible dust attenuation laws than the simple recipes usually adopted in galaxy spectral analyses [e.g., Kauffmann et al., 2003a, Brinchmann et al., 2004, Cid Fernandes et al., 2005, Panter et al., 2007, Wuyts et al., 2009]. We therefore expect larger typical uncertainties arising from attenuation by dust than in these previous studies.

3.2.3 Library of galaxy spectral energy distributions

We use the library of star formation and chemical enrichment histories described in Section 3.2.1 and the spectral models described in Section 3.2.2 to simulate a large library of galaxy spectral energy distributions. This will be used in the remainder of this Chapter to assess the retrievability of galaxy physical parameters from various types of observations. To ensure that all kinds of spectral energy distributions are properly sampled, for each of the 500,000 star formation and chemical enrichment histories in the original library, we draw 10 different realizations of the evolutionary stage (and hence of the parameter f_{SFH}) and, for star-forming galaxies, of the current physical properties (parameters ψ_{S}, Z, $\log U_0$, ξ_{d}, $\hat{\tau}_V$ and μ), as summarized in Table 3.1. The final library therefore contains 5 million models, which we can use to generate catalogs of galaxy spectra.

Table 3.1: Prior distributions of the current (i.e. averaged over the last 10 Myr) physical properties of star-forming galaxies in the library of star formation and chemical enrichment histories assembled in Sections 3.2.1 and 3.2.2.

Par.	Description	Range	Probability Density Function
$\psi_{\rm S}$	Specific star formation rate	$-2 < \log(\psi_{\rm S}/{\rm Gyr}^{-1}) < 1$	$0.05\left[1.5^5 - \left\lvert\log(\psi_{\rm S}/{\rm Gyr}^{-1}) + 0.5\right\rvert^5\right]$
Z	Metallicity of the star-forming gas	$-1 < \log(Z/Z_\odot) < 0.54$	uniform
$\log U_0$	Zero-age ionization parameter	$-3.5 < \log U_0 < -1.5$	uniform
$\xi_{\rm d}$	Dust-to-metal ratio in the ionized gas	$0.1 < \xi_{\rm d} < 0.5$	uniform
$\hat{\tau}_V$	Total V-band attenuation optical depth of the dust	$0.01 < \hat{\tau}_V < 4$	$0.18\arctan\left[5(4-\hat{\tau}_V)\right]$
μ	Fraction of $\hat{\tau}_V$ contributed by dust in diffuse ISM	$0.1 < \mu < 0.7$	uniform
n	Slope of the attenuation curve in the diffuse ISM	$1.6 < n\,(1+\mu\hat{\tau}_V) < 2.4$	uniform

In the present study, we focus primarily on the rest-frame optical properties of galaxies. Our main goal is to compare the constraints that can be derived on galaxy physical parameters from different types of observations in this wavelength range. At wavelengths between 3525 and 7500 Å, the stellar population spectra computed using the models presented in Section 3.2.2.1 have a native spectral resolution of 2.3 Å (full width at half-maximum, hereafter FWHM), corresponding to a resolving power $R \approx 2200$ at 5000 Å (the resolution is coarser outside this wavelength range). This allows us to investigate the retrievability of galaxy physical parameters from 4 different types of observations at optical wavelengths:

- *Medium-resolution spectroscopy (λ=3600–7400 Å).* We adopt a fiducial resolution of 5 Å FWHM, corresponding to $R = 1000$ at $\lambda = 5000$ Å, and a fixed pixel size of 2.5 Å.
- *Low-resolution spectroscopy (λ=3600–7400 Å).* We adopt a fiducial resolution of 50 Å FWHM, corresponding to $R = 100$ at $\lambda = 5000$ Å, and a fixed pixel size of 25 Å.
- *Equivalent-width measurements of strong emission lines.* We also investigate the constraints set purely by the equivalent widths of the following emission lines, at both low and medium spectral resolution: [O II]λ3727; Hβ; [O III]λ4959, [O III]λ5007; [N II]λ6548; Hα; [N II]λ6584; [S II]λ6716 and [S II]λ6731. An advantage of equivalent widths is that they can provide useful information even in spectra which are not flux-calibrated.
- *Broadband photometry.* We adopt the SDSS *ugriz* filter response functions.

As an example, Fig. 3.3 shows the spectral energy distribution of the same model galaxy as in Fig. 3.1 at the age of 8.4 Gyr (with current physical parameters resampled using the distributions in Table 3.1), computed alternatively at medium (panel a) and low (panel b) spectral resolution and convolved with the SDSS *ugriz* filter response functions (panel c).

3.2.4 Retrievability of galaxy physical parameters

The galaxy spectral library generated in the previous section can be used to assess the retrievability of physical parameters from different types of observations. To achieve this, we first draw a random subsample of 10,000 spectral energy distributions from the library, to which we apply artificial noise to mimic true observations. For low- and medium-resolution spectroscopy, we apply noise independent of wavelength (i.e. white noise), parametrized in terms of the median S/N per pixel, $\overline{\mathrm{S/N}}$. We consider the cases $\overline{\mathrm{S/N}} = 5$ and 20 (for reference, the stellar population spectra computed using the models presented in Section 3.2.2.1 have a native $\overline{\mathrm{S/N}}$ of about 95). For multiband photometry, instead, we take the signal-to-noise ratio to be the same in all wavebands (S/N = 10 or 30). In the following, we refer to this subsample of 10,000 noisy spectral energy distributions as 'pseudo-observations', and to the 4,990,000 remaining spectral energy distributions in the library as simply 'models'.

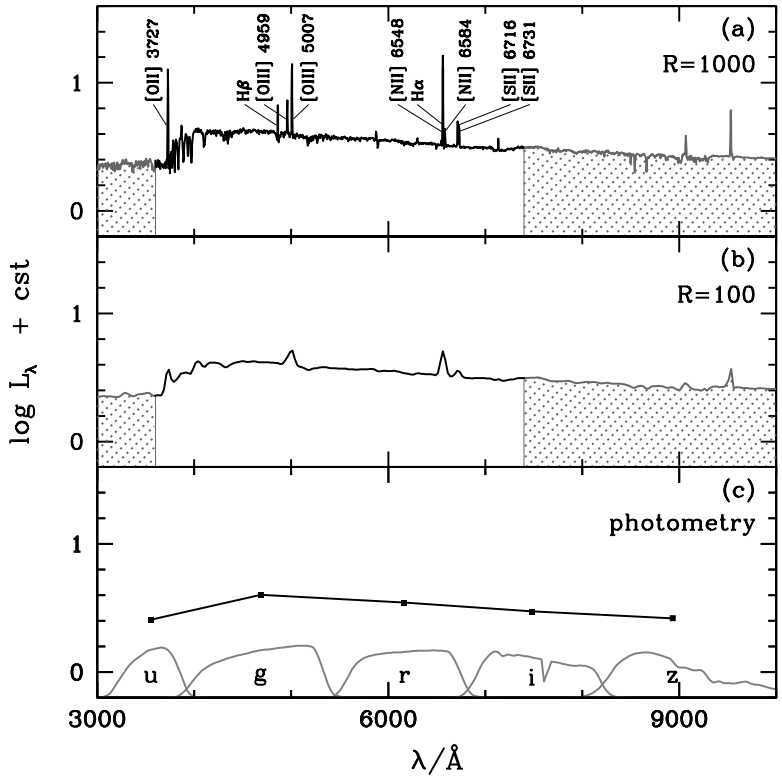

Figure 3.3: Spectral energy distribution of the same model galaxy as in Fig. 3.1 at the age of 8.4 Gyr, computed using the models described in Sections 3.2.1 and 3.2.2. (*a*) At a spectral resolution of 5 Å FWHM ($R = 1000$ at 5000 Å). The black portion of the spectrum (λ=3600–7400 Å) is that used to retrieve galaxy physical parameters at this resolution. (*b*) Same as in (*a*), but at a spectral resolution of 50 Å FWHM ($R = 100$ at 5000 Å). (*c*) Same as in (*a*), but convolved with the SDSS *ugriz* filter response functions (shown at the bottom). The model galaxy has $f_{\rm SFH} = 0.3$ and current parameters (resampled using the distributions in Table 3.1) $\psi_{\rm S} = 0.08\,{\rm Gyr}^{-1}$, $Z = 0.4 Z_{\odot}$, $\log U_0 = -2.8$, $\xi_{\rm d} = 0.3$, $\hat{\tau}_V = 1.0$, $\mu = 0.6$ and $n = 0.7$. In (*a*) and (*b*), the adopted pixel size is half the size of a resolution element.

We use a Bayesian approach to quantify the accuracy to which several physical parameters of these pseudo-galaxies can be retrieved from different types of observables. The likelihood of the j th model, noted $\mathcal{M}_j$, given the spectral energy distribution of a pseudo-galaxy can be written

$$\ln P(\mathcal{M}_j|\{\mathcal{O}_i\}) = -\frac{1}{2}\sum_i \left(\frac{\mathcal{O}_i - w_j\mathcal{M}_{j,i}}{\sigma_i}\right)^2 , \tag{3.2.10}$$

where $\mathcal{O}_i$ and $\mathcal{M}_{j,i}$ are the spectral energy distributions (i.e., either the spectral flux density at the i th wavelength, or the equivalent width of the i th emission line, or the flux in the i th photometric band; see Section 3.2.3) of the pseudo-galaxy and the j th model, respectively, σ_i is the uncertainty in $\mathcal{O}_i$ corresponding to the adopted signal-to-noise ratio, and

$$w_j = \left(\sum_i \frac{\mathcal{O}_i\mathcal{M}_{j,i}}{\sigma_i^2}\right)\left[\sum_i \left(\frac{\mathcal{M}_{j,i}}{\sigma_i}\right)^2\right]^{-1} \tag{3.2.11}$$

is the model scaling factor that maximizes $P(\mathcal{M}_j|\{\mathcal{O}_i\})$ (this parameter is unity in the case of scale-free equivalent widths). In practice, in the case of medium-resolution spectroscopy, the summation on the right-hand side of expression (3.2.10) extends over 1520 wavelength points and hence generally yields values greater than 2000, even for a good-fit model.[4] The absolute likelihood in this case cannot be evaluated numerically, as it is always rounded to 0. For medium-resolution spectroscopy, therefore, we rather compute the probability of the j th model relative to the best-fitting model, noted $\mathcal{M}_0$,

$$\ln \tilde{P}(\mathcal{M}_j|\{\mathcal{O}_i\}) = \ln P(\mathcal{M}_j|\{\mathcal{O}_i\}) - \ln P(\mathcal{M}_0|\{\mathcal{O}_i\}) . \tag{3.2.12}$$

We have checked that in the case of low-resolution spectroscopy, where both $P(\mathcal{M}_j|\{\mathcal{O}_i\})$ and $\tilde{P}(\mathcal{M}_j|\{\mathcal{O}_i\})$ can be evaluated numerically, the likelihood distributions of galaxy physical parameters obtained using equations (3.2.10) and (3.2.12) are similar.

In this Chapter, we focus on the retrievability of 6 physical parameters from the observed spectral energy distributions of galaxies:

- the observer-frame absolute r-band (stellar) mass-to-light ratio, M_*/L_r. This is the ratio of the galaxy stellar mass to the absolute luminosity in the observed r band;

[4] We note that only a small contribution to these high values arises from our neglect of intrinsic correlations between spectral pixels in equation (3.2.10). A detailed investigation (see Appendix A) reveals that such correlations are significant only for pixels associated with emission lines, which reflects the fact that all emission lines arise from the same ionizing radiation. We have used a test sample of 3000 pseudo-galaxies to check the effect of accounting for pixel-to-pixel correlations on the calculation of model likelihoods. This requires the introduction of the covariance matrix in equation (3.2.10) and is computationally very demanding. The result is a negligible effect on the computed model likelihoods. Hence equation (3.2.10) should be appropriate to compute model likelihoods.

- the fraction of the current galaxy stellar mass formed during the last 2.5 Gyr, $f_{\rm SFH}$;
- the specific star formation rate, $\psi_{\rm S}$;
- the gas-phase oxygen abundance in units of $12 + \log({\rm O/H})$, where O/H is the abundance by number of oxygen relative to hydrogen. In the models of nebular emission adopted in Section 3.2.2.2, the gas-phase oxygen abundance corresponding to solar metallicity ($Z = Z_{\odot}$) is 8.81, 8.73 and 8.63 for dust-to-metal ratios $\xi_{\rm d} = 0.1$, 0.3 and 0.5, respectively.
- the total effective V-band absorption optical depth of the dust, $\hat{\tau}_V$ (Section 3.2.2.3);
- the fraction of $\hat{\tau}_V$ arising from dust in the ambient ISM, μ.

Fig. 3.4 shows the prior distributions of these parameters in the galaxy spectral library generated in Section 3.2.3. In each panel, the shaded histogram shows the distribution including all galaxies, while the solid histogram shows the contribution by star-forming galaxies alone. Galaxies without star formation have typically high M_*/L_r in Fig. 3.4a and $f_{\rm SFH} \approx 0$ in Fig. 3.4b. These galaxies are off scale (at $\log \psi_{\rm S} = -\infty$) in Fig. 3.4c. They do not contribute to the histograms in Figs 3.4d–f, which pertain to interstellar parameters. As emphasized in Sections 3.2.1 and 3.2.2, the prior distributions of Fig. 3.4 are as flat as possible to avoid unwanted biases in the retrieval of physical parameters from observed spectral energy distributions of galaxies. This choice is motivated by the desire to make the spectral library appropriate for the analysis of galaxies in wide ranges of physical properties (we note that the finite extent of the prior distributions in Fig. 3.4 could still bias analyses of galaxies with extreme parameters). Alternative prior distributions can be designed to help reduce the uncertainties in the parameters derived for galaxies with known specific physical properties (e.g., actively star-forming, early-type, metal-poor). We return to the influence of prior distributions on parameter determinations in Section 3.4.2 below, where we investigate the results obtained using the library of galaxy star formation and chemical enrichment histories derived from the original semi-analytic recipes of De Lucia and Blaizot [2007] .

In Fig. 3.5, we show examples of the probability density functions (likelihood distributions) of physical parameters retrieved, using equations (3.2.10)–(3.2.11) and the prior distributions of Fig. 3.4, from the low-resolution, high-S/N spectrum of a pseudo-galaxy in our library. This pseudo-observation was obtained by adding artificial noise ($\overline{\rm S/N} = 20$) to the model spectrum of Fig. 3.3b. We take the 'best estimate' of each parameter in Fig. 3.5 to be the median of the retrieved probability density function and the associated 'uncertainty' to be half the 16th–84th percentile range (this would be equivalent to the $\pm 1\sigma$ range in the case of a Gaussian distribution). To quantify potential biases in the retrieved likelihood distribution, we further define the 'accuracy' as the absolute difference between the retrieved best estimate and true parameter value. In this example, all the parameters are well retrieved.

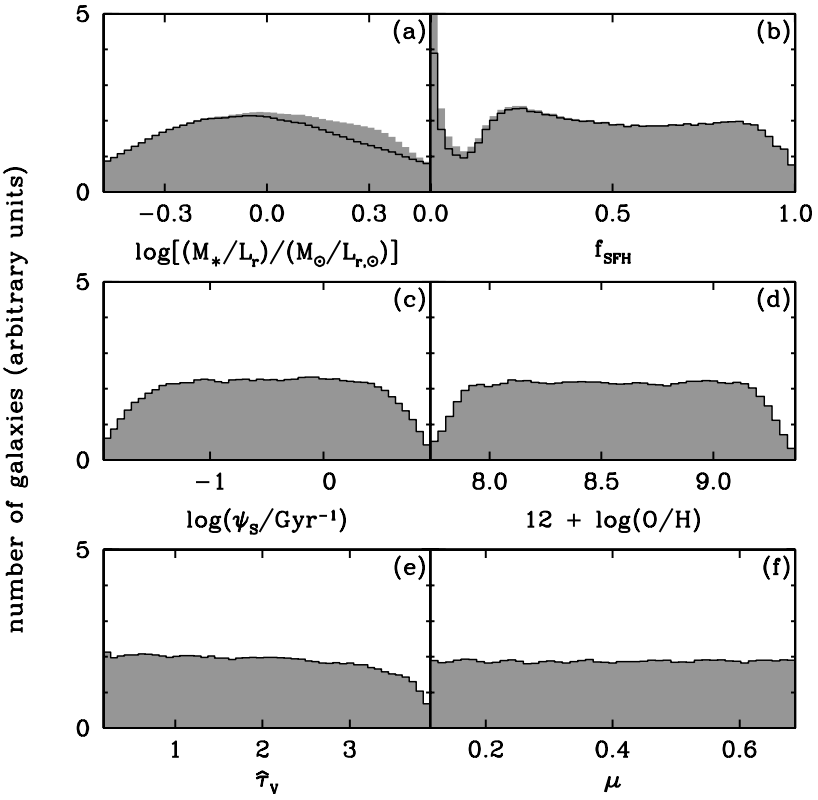

Figure 3.4: Prior distributions of selected physical parameters of the 5 million galaxies in the spectral library generated in Section 3.2.3: (*a*) observer-frame absolute r-band stellar mass-to-light ratio, M_*/L_r; (*b*) fraction of current stellar mass formed during the last 2.5 Gyr, $f_{\rm SFH}$; (*c*) specific star formation rate, $\psi_{\rm S}$; (*d*) gas-phase oxygen abundance, $12 + \log({\rm O/H})$; (*e*) total effective V-band absorption optical depth of the dust, $\hat{\tau}_V$; (*f*) fraction of $\hat{\tau}_V$ arising from dust in the ambient ISM, μ. In each panel, the shaded histogram shows the distribution for all galaxies, while the solid histogram shows the contribution by star-forming galaxies alone. Non-star-forming galaxies are off scale (at $\log\psi_{\rm S} = -\infty$) in panel (*c*) and do not contribute to the distributions of interstellar parameters in panels (*d*)–(*f*).

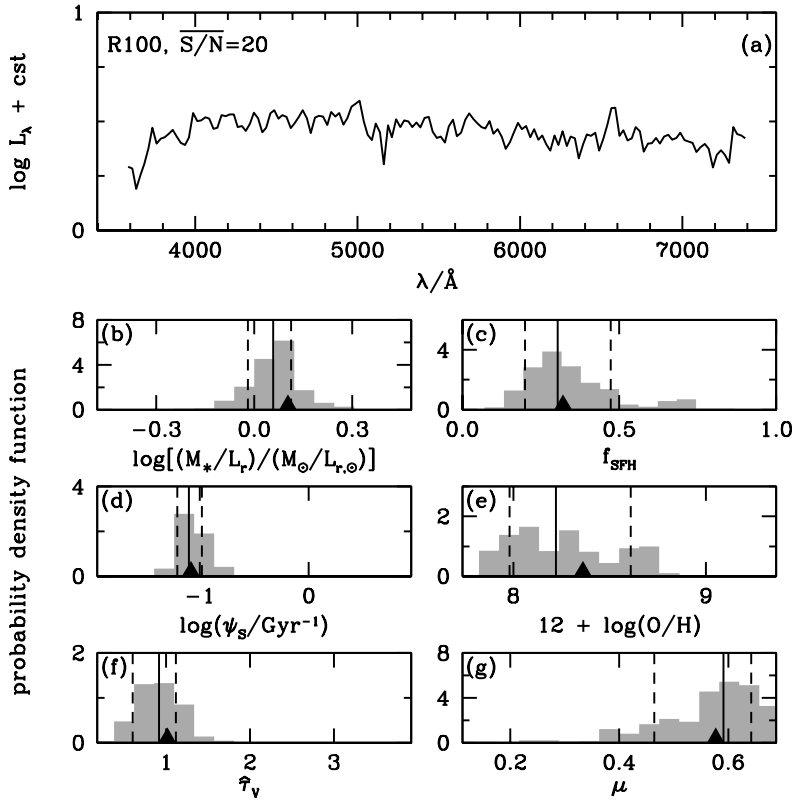

Figure 3.5: Probability density functions of the same physical parameters as in Fig. 3.4 retrieved, using the Bayesian approach described in Section 3.2.4, from the pseudo-galaxy spectrum shown at the top. The spectrum was obtained by adding artificial noise with $\overline{\mathrm{S/N}} = 20$ to the $R = 100$ model spectrum of Fig. 3.3b. In each panel, the black triangle indicates the true value of the parameter of the pseudo-galaxy, the solid line the best estimate (50th percentile of the retrieved likelihood distribution) and the dashed lines the associated 68-percent confidence interval (16th and 84th percentiles).

The success in recovering the fraction $f_{\rm SFH}$ of intermediate-age stars (Fig. 3.5c) and the interstellar parameters $12 + \log({\rm O/H})$, $\hat{\tau}_V$ and μ (Figs 3.5e–3.5g) is remarkable. It is the first illustration of the power of our approach to exploit strong absorption- and emission-line signatures in noisy, low-resolution spectra. We examine this property in detail in Section 3.3.2 below.

It is important to stress that the approach adopted here to assess the retrievability of galaxy physical parameters from different types of observations, while idealized, is as realistic as can be achieved using modern techniques. Pseudo-observations are required to test the retrievability of physical parameters from observables, because the physical parameters of observed galaxies cannot be known a priori. The fact that we draw pseudo-observations from the same spectral library as the models used to analyze them makes the constraints derived on the retrievability of physical parameters optimistic: our results correspond to the best possible case in which the models are a realistic representation of true galaxies. In Section 3.4 below, we also illustrate the constraints that can be derived by applying our approach to interpret a sample of observed galaxy spectra.

3.3 Constraints on galaxy physical parameters from different types of observations

In this section, we use the modeling approach described in the previous section to quantify the accuracy to which physical parameters of galaxies can be retrieved from different types of observations. We focus on the retrievability of the 6 parameters introduced in Section 3.2.4: the observer-frame absolute r-band stellar mass-to-light ratio, M_*/L_r; the fraction of the current galaxy stellar mass formed during the last 2.5 Gyr, $f_{\rm SFH}$; the specific star formation rate, $\psi_{\rm S}$; the gas-phase oxygen abundance, $12 + \log({\rm O/H})$; the total effective V-band absorption optical depth of the dust, $\hat{\tau}_V$; and the fraction of $\hat{\tau}_V$ arising from dust in the ambient ISM, μ. In Section 3.3.1 below, we first examine the constraints set on these parameters by multi-band photometric observations. We consider the influence of several factors, such as the signal-to-noise ratio, the number of photometric bands, the contamination of broad-band fluxes by nebular emission lines and the uncertainty introduced by the lack of redshift information. We find that photometric observations can provide useful constraints mainly on M_*/L_r and $\psi_{\rm S}$. Then, in Section 3.3.2, we quantify the improvement that can be achieved in the constraints on all parameters by appealing to different types of spectroscopic observations.

3.3.1 Constraints from multi-band photometry

3.3.1.1 Nearby galaxies

We first explore the extent to which physical parameters of nearby galaxies can be constrained using 5-band *ugriz* photometry. We extract 10,000 spectral energy distributions from the library assembled in Section 3.2.3 to compute pseudo-

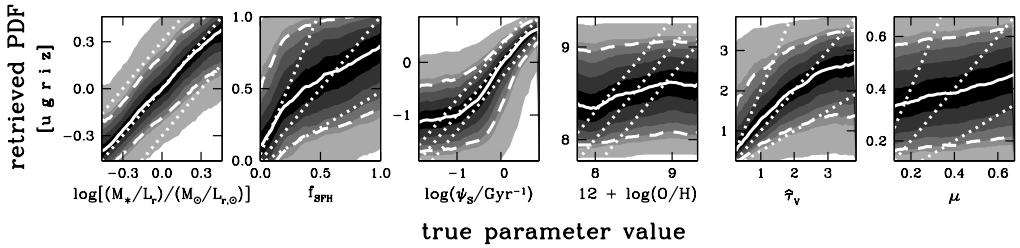

Figure 3.6: Average probability density functions of the same 6 physical parameters as in Fig. 3.4 retrieved, using the Bayesian approach described in Section 3.2.4, from 5-band *ugriz* photometry with S/N = 30 of a sample of 10,000 pseudo-galaxies (standard case). Each panel corresponds to a different parameter [from left to right: M_*/L_r, $f_{\rm SFH}$, $\psi_{\rm S}$, $12+\log({\rm O/H})$, $\hat{\tau}_V$ and μ]. In each case, average probability density functions in 50 narrow bins of true parameter value were obtained by coadding and then renormalizing the probability density functions retrieved for the 10,000 pseudo-galaxies (see Section 3.3.1.1 for detail). Grey levels locate the 2.5th, 12th, 16th, 22nd, 30th, 40th, 60th, 70th, 78th, 84th, 88th and 97.5th percentiles of the average likelihood distribution in each bin. The solid line locates the associated median (best estimate) and the 2 dashed lines the 16th and 84th percentiles (68-percent confidence interval). For reference, 3 dotted lines indicate the identity relation and deviations by a factor of 2 (± 0.3 dex) between the retrieved and true parameter values.

observations of galaxies in these bands, as described in Section 3.2.4. We adopt for the moment a signal-to-noise ratio S/N = 30 in all bands and place all spectral energy distributions at $z = 0$. The rest of the 5 million models in the library allow us to retrieve likelihood distributions of M_*/L_r, $f_{\rm SFH}$, $\psi_{\rm S}$, $12+\log({\rm O/H})$, $\hat{\tau}_V$ and μ for each of these 10,000 pseudo-galaxies, by means of equations (3.2.10)–(3.2.11). In the following, we refer to this setting as the 'standard' case.

Fig. 3.6 summarizes the constraints that can be set on the physical parameters of nearby galaxies using high-quality, 5-band *ugriz* photometry. For each parameter, we plot the average probability density function retrieved from the above analysis as a function of true parameter value. This average was obtained by coadding and then renormalizing the probability density functions retrieved in narrow bins of true parameter value for the 10,000 pseudo-galaxies in our sample. The solid line in each panel of Fig. 3.6 locates the median (i.e. best estimate; Section 3.2.4) of the average likelihood distribution in each bin, the dashed lines the 16th and 84th percentiles (68-percent confidence interval) and the outer edges of the lightest grey level the 2.5th and 97.5th percentiles (these would correspond to the $\pm 2\sigma$ levels in the case of a Gaussian distribution). For reference, 3 dotted lines indicate the identity relation

and deviations by a factor of 2 (± 0.3 dex) between the retrieved and true parameters.

Fig. 3.6 shows that the r-band stellar mass-to-light ratio of nearby galaxies can be reasonably well constrained using 5-band *ugriz* photometry. The retrieved best estimate is in excellent agreement with the true parameter value over most of the explored range, and the associated uncertainty, defined as half the 16th–84th percentile range (Section 3.2.4), is only about 0.19 dex (this number would be larger if we allowed for IMF variations). At the extremities of the range, the relative deficiency of galaxies with very small and very large M_*/L_r in the spectral library (Fig. 3.4a) slightly biases the retrieved likelihood distributions toward intermediate M_*/L_r values. Fig. 3.6 also shows that, in contrast to M_*/L_r, the fraction $f_{\rm SFH}$ of stellar mass formed in the past 2.5 Gyr cannot be well retrieved from 5-band *ugriz* photometry. This is because the strong absorption features of intermediate-age stars have a negligible influence on broadband fluxes. The signatures of $f_{\rm SFH}$ on broadband colours are otherwise difficult to disentangle from those of stellar metallicity and attenuation by dust.

The constraints on specific star formation rate from 5-band *ugriz* photometry in Fig. 3.6 are meaningful only for actively star-forming galaxies, i.e. for $\psi_{\rm S} > 1\,{\rm Gyr}^{-1}$, and the associated uncertainty is at best 0.30 dex. In such galaxies, the contamination of broadband fluxes by strong emission lines produces identifiable signatures ([O II] in the u band; Hβ and [O III] in the g band; Hα, [N II] and [S II] in the r band). Even then, the similar colours of galaxies with both low $\psi_{\rm S}$ and $\hat{\tau}_V$ and those with both high $\psi_{\rm S}$ and $\hat{\tau}_V$ give rise to long tails toward low $\psi_{\rm S}$ in the retrieved likelihood distributions. This degeneracy is illustrated by the overlap of galaxies with different $\psi_{\rm S}$ in Fig 3.7a, which shows the $g-i$ versus $u-g$ colours of a subset of 50,000 models in the spectral library. Together with the fact that, even in galaxies with strong emission lines, broadband colours do not allow the distinction between different line ratios, this also implies poor constraints on the gas-phase oxygen abundance and the dust parameters $\hat{\tau}_V$ and μ in Fig. 3.6. The retrieved likelihood distributions of these parameters are largely driven by the prior distributions of Fig. 3.4.

The results of Fig. 3.6 are consistent with those of Wuyts et al. [2009], who explored the recovery of physical parameters from broadband spectral energy distributions of model galaxies taken from hydrodynamical merger simulations. Our study differs from theirs in that we consider much wider ranges of star formation and chemical enrichment histories and dust properties to retrieve likelihood distributions of physical parameters. Also, Wuyts et al. [2009] neglect the contribution by nebular emission to broadband fluxes and include 11 bands to sample the rest-frame spectral energy distributions of galaxies over a wide range of wavelengths, between roughly 0.15 and 3 μm. Finally, they analyze galaxy simulations in the redshift range $1.5 \leq z \leq 2.9$, which span a much narrower range of stellar population ages than the pseudo-galaxies in our sample. For all these reasons, Wuyts et al. [2009] find smaller typical uncertainties in the retrieved stellar mass (~ 0.10 dex) and star formation rate (~ 0.31 dex) than can be inferred from the constraints on M_*/L_r and $\psi_{\rm S}$ in Fig. 3.6. They also find much tighter constraints on the dust attenuation

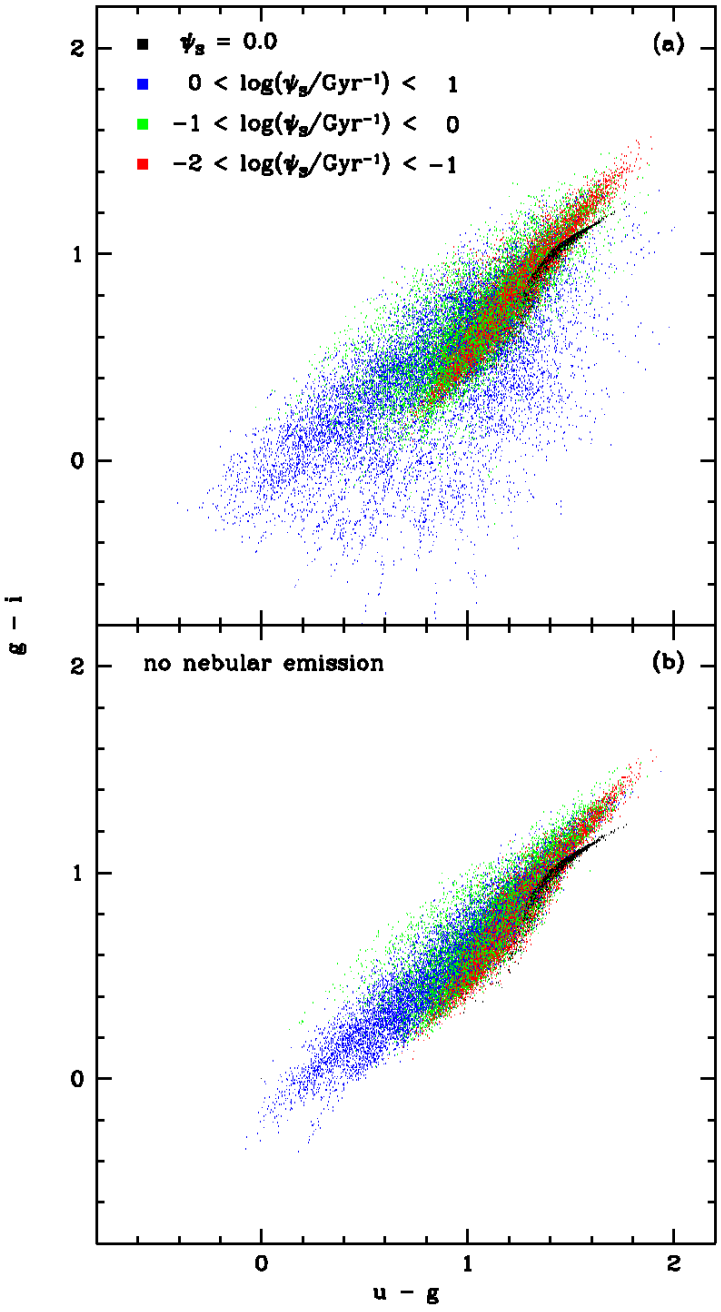

Figure 3.7: (*a*) $g-i$ colour plotted against $u-g$ colour for a subset of 50,000 models from the galaxy spectral library assembled in Section 3.2.3. The models are colour-coded according to specific star formation rate, as indicated. (*b*) Same as (*a*), but without including the contribution by nebular emission to broadband fluxes.

optical depth (~ 0.31). Wuyts et al. [2009] note that using simple (single-burst, constant and exponentially declining) star formation histories to fit the results of more realistic hydrodynamical merger simulations can induce significant deviations between the retrieved maximum-likelihood and true parameters values. Our adoption of the sophisticated galaxy spectral library assembled in Section 3.2.3 should minimize this type of limitation when applying our approach to the interpretation of real observations (Section 3.4).

We now return to our investigation and examine the sensitivity of the results of Fig. 3.6 to the adopted signal-to-noise ratio, the photometric wavelength coverage and the inclusion of nebular emission when interpreting broadband observations of galaxies. To quantify potential differences in the retrieved likelihood distributions relative to the standard case, we define the improvement factor as the ratio

$$I_\sigma = \frac{\text{uncertainty in the standard case}}{\text{uncertainty}}\,. \tag{3.3.13}$$

This quantity can be computed for any true parameter value. It exceeds unity when the constraints are tighter (i.e. the uncertainty smaller) than in the standard case. We further define the gain in accuracy as the difference

$$\Delta = (\text{accuracy in the standard case}) - (\text{accuracy})\,. \tag{3.3.14}$$

This quantity is positive when the retrieved best estimate is more accurate (i.e. the absolute difference between the retrieved best estimate and true parameter value smaller) than in the standard case. In Fig. 3.8, we show the analog of Fig. 3.6 for 3 distinct alternatives to the standard case: adopting a signal-to-noise ratio $\mathrm{S/N} = 10$ instead of 30 (top row); removing the constraints from the reddest 2 photometric bands, i and z (middle row); and not including nebular emission in the model library used to analyze pseudo-observations (bottom row). In all cases, we show at the bottom of each panel the improvement factor I_σ and the gain in accuracy Δ as a function of true parameter value.

Fig. 3.8 (top row) shows that reducing the signal-to-noise ratio from $\mathrm{S/N} = 30$ to 10 makes the uncertainty in the retrieved M_*/L_r larger by about 15 percent and, at the extremities of the explored range, the accuracy between 10 and 25 percent worse. A more dramatic difference relative to the standard case arises from the loss of the emission-line signal for actively star-forming galaxies. This causes a sharp drop in I_σ at high ψ_S, implying poor constraints on the specific star formation rate over the entire explored range. Interestingly, Fig. 3.8 (middle row) shows that the effect of removing the i- and z-band constraints at $\mathrm{S/N} = 30$ is very similar to that of lowering the signal-to-noise ratio of 5-band *ugriz* photometry. This is because the signatures of dust attenuation can be less well delineated from those of star formation from the *ugr* colours alone, even if the contamination of these broadband fluxes by strong emission lines can in principle be detected at high S/N (Fig. 3.6). As expected, the other parameters, f_SFH, $12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and μ, can be even less well retrieved than in the standard case when the signal-to-noise ratio is lowered and when the i- and z-band constraints are removed.

It is also of interest to quantify the importance of including nebular emission when interpreting broadband observations of galaxies. A comparison between the top and bottom panels of Fig. 3.7 shows that the contribution by [O II] emission to the u-band flux and that of Hβ and [O III] emission to the g-band flux can change the broadband $u-g$ and $g-i$ colours of galaxies with $\psi_{\rm S} > 1\,{\rm Gyr}^{-1}$ by several tenths of a magnitude. In Fig. 3.8 (bottom row), we show the results of attempting to interpret the same pseudo-observations as in Fig. 3.6 (which include the contribution by nebular emission to the *ugriz* fluxes) using a spectral library which neglects nebular emission. This spectral library was constructed in the same way as that assembled in Section 3.2.3, but ignoring the prescription for nebular emission in Section 3.2.2.2. The retrieved M_*/L_r in this case is very similar to that obtained in the standard case, except for the slightly worse accuracy ($\Delta < 0$) at very low M_*/L_r, consistent with the contamination of the r-band flux by Hα, [N II] and [S II] emission in actively star-forming galaxies. As anticipated from Fig. 3.7, models without nebular emission do not sample broadband colours well enough to allow the retrieval of meaningful constraints on $\psi_{\rm S}$ for actively star-forming galaxies. The other parameters, $f_{\rm SFH}$, $12+\log({\rm O/H})$, $\hat{\tau}_V$ and μ, are as poorly constrained as in the standard case. We conclude from Fig. 3.8 that high S/N, good wavelength coverage and accounting for nebular emission are all important to tighten the constraints on specific star formation rate from broadband observations of galaxies, while the inclusion of nebular emission is less crucial to constrain stellar mass-to-light ratio.

3.3.1.2 Galaxies at unknown redshift

We have tested so far the retrievability of physical parameters from 5-band *ugriz* photometry only for nearby galaxies. We now wish to explore the constraints achievable for galaxies at unknown redshift. Ilbert et al. [2006] have shown that redshift can be best estimated from 5-band *ugriz* photometry out to $z \sim 1.5$. Photometric redshift measurements of more distant galaxies require observations at longer wavelengths [e.g. Ilbert et al., 2009], which are beyond the scope of the present Chapter (see Pacifici et al., in preparation). We therefore restrict ourselves here to the retrievability of physical parameters of galaxies from 5-band *ugriz* photometry at redshifts in the range $0 < z < 1$.

We use an approach similar to that described in Section 3.2.3 to generate a new library of 5 million spectral energy distributions of galaxies at random redshifts between 0 and 1. For each of the 500,000 star formation and chemical enrichment histories in the original library of Section 3.2.1, we first draw an observation redshift $z_{\rm obs}$ between 0 and 1. Then, we proceed as outlined in Section 3.2.3 and draw 10 different realization of the evolutionary stage (i.e. galaxy age) at which the galaxy is looked at (uniformly in redshift between $z_{\rm obs}$ and $z = 1.7$, to better sample the parameter $f_{\rm SFH}$; see Section 3.2.1) and, for star-forming galaxies, of the current parameters $\psi_{\rm S}$, Z, $\log U_0$, $\xi_{\rm d}$, $\hat{\tau}_V$ and μ (using the distributions in Table 3.1). The resulting spectral library contains about 8 percent early-type galaxies, i.e. slightly less than the 10 percent in the $z = 0$ library assembled in Section 3.2.3.

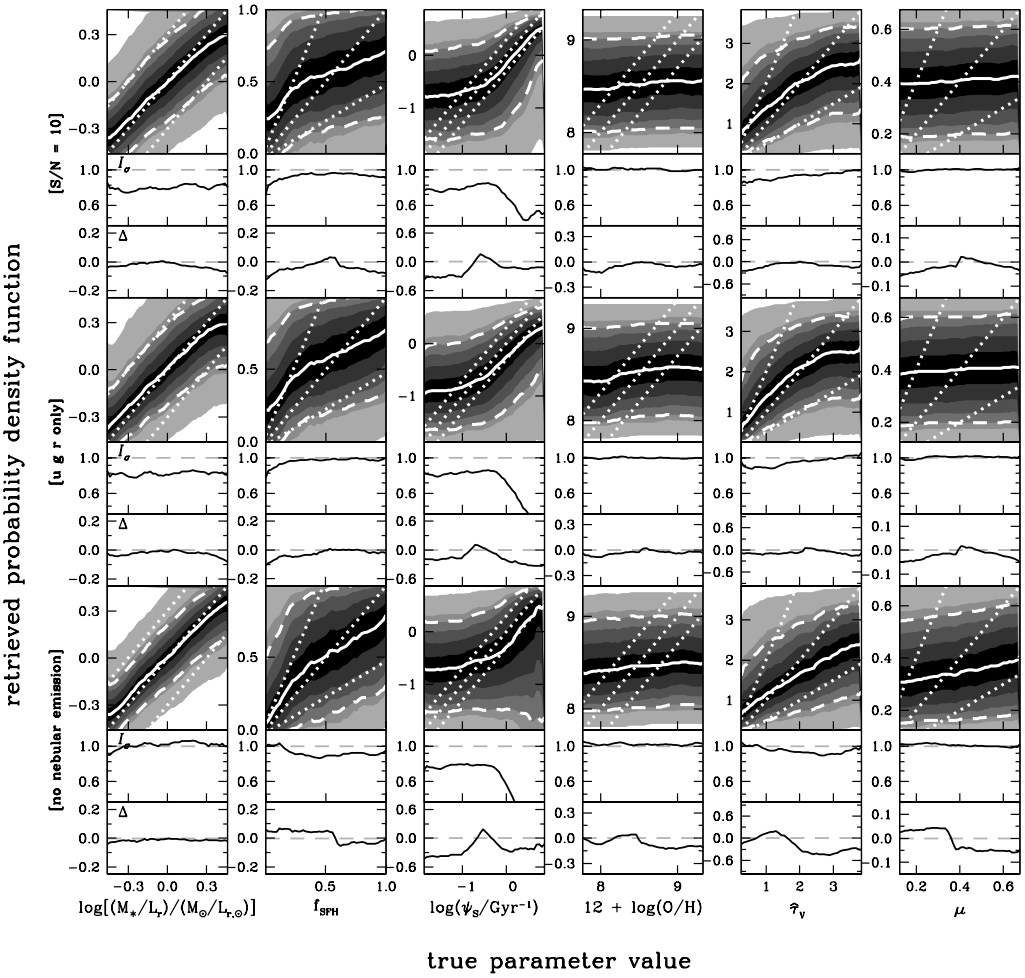

Figure 3.8: Same as Fig. 3.6, but for 3 distinct alternatives to the standard case: (*top row*) adopting a signal-to-noise ratio S/N = 10 instead of 30; (*middle row*) using constraints from only the *ugr* photometric bands instead of *ugriz*; and (*bottom row*) not including nebular emission in the model library used to analyze the sample of 10,000 pseudo-galaxies. In all cases, the improvement factor I_σ and the gain in accuracy Δ (equations 3.3.13–3.3.14) are shown as a function of true parameter value at the bottom of each panel to quantify differences in the retrieved likelihood distributions relative to the standard case of Fig. 3.6.

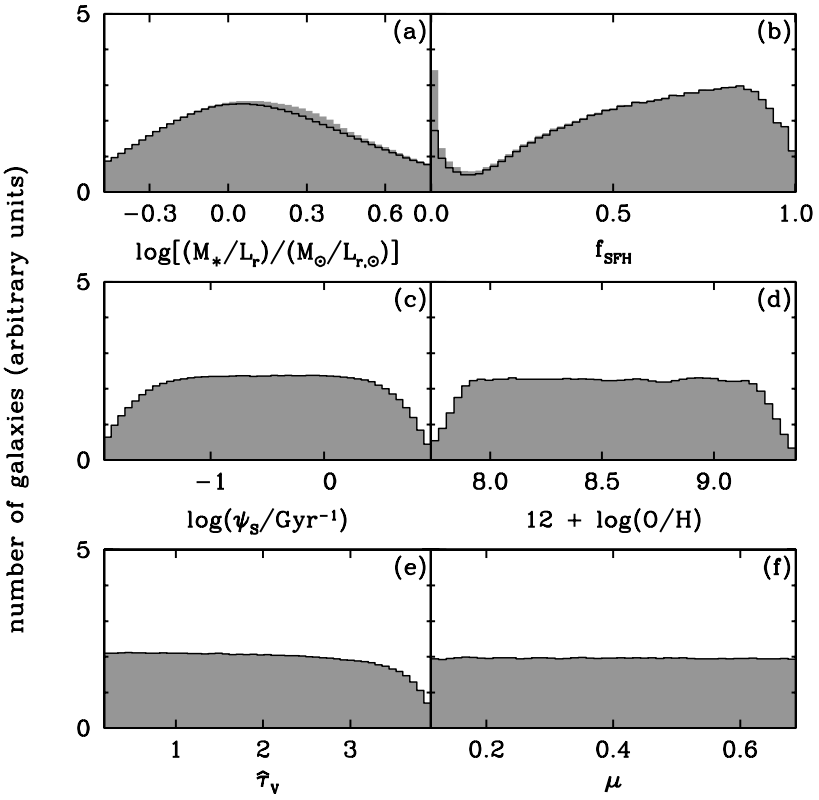

Figure 3.9: Prior distributions of the same physical parameters as in Fig. 3.4 for the 5 million galaxies at redshifts between 0 and 1 in the spectral library assembled in Section 3.3.1.2. In each panel, the shaded histogram shows the distribution for all galaxies, while the solid histogram shows the contribution by star-forming galaxies alone. Non-star-forming galaxies are off scale (at $\log \psi_{\rm S} = -\infty$) in panel (*c*) and do not contribute to the distributions of interstellar parameters in panels (*d*)–(*f*).

Fig. 3.9 shows the prior distributions of the physical parameters M_*/L_r, $f_{\rm SFH}$, $\psi_{\rm S}$, $12 + \log({\rm O/H})$, $\hat{\tau}_V$ and μ in this library. As in Fig. 3.4, in each panel, the shaded histogram shows the distribution including all galaxies, while the solid histogram shows the contribution by star-forming galaxies alone. The prior distribution of the observer-frame absolute r-band stellar mass-to-light ratio in Fig. 3.9a reaches larger values than in Fig. 3.4a because of the contribution by (dusty) high-redshift galaxies with large K corrections. The early evolutionary stage of high-redshift galaxies also implies a larger typical $f_{\rm SFH}$ in Fig. 3.9b than in Fig. 3.4b. By construction, the distributions of the current physical parameters $\psi_{\rm S}$, $12 + \log({\rm O/H})$, $\hat{\tau}_V$ and μ are the same in Figs 3.4 and 3.9 (from Table 3.1).

By analogy with Section 3.3.1.1, we generate pseudo-observations in the *ugriz* photometric bands of 10,000 galaxies in the new library. These pseudo-observations

sample different wavelength ranges of the rest-frame spectral energy distributions of galaxies at different redshifts, from roughly 3000–10,000 Å at $z_{\rm obs} = 0$ to 1500–5000 Å at $z_{\rm obs} = 1$. We consider both high-quality (S/N = 30) and low-quality (S/N = 10) observations. Using the rest of the 5 million models in the library, we can retrieve likelihood distributions of $z_{\rm obs}$ for each of the 10,000 pseudo-galaxies, as well as likelihood distributions of M_*/L_r, $f_{\rm SFH}$, $\psi_{\rm S}$, $12 + \log({\rm O/H})$, $\hat{\tau}_V$ and μ.

In Fig. 3.10a, we plot the distribution of the difference between the retrieved best estimate (i.e. median) and true value of $z_{\rm obs}$, in units of $1 + z_{\rm obs}$, for the 10,000 pseudo-galaxies in our sample. The results for S/N = 30 (shaded histogram) are in remarkable agreement with the $\sigma = 0.03$ Gaussian uncertainty in this quantity quoted by Ilbert et al. [2006], who analyzed CFHTLS (S/N $\gtrsim$ 30) $u^*g'r'i'z'$ photometry of about 3000 galaxies with spectroscopic redshifts between 0.2 and 1.5 (dashed line). For S/N = 10, the retrieved best estimate of $z_{\rm obs}$ in Fig. 3.10a remains in good general agreement with the true value (solid histogram). To better illustrate the origin of the histograms in Figs 3.10a, we show in Figs 3.10b and 3.10c the detail of the average retrieved probability density functions as a function of $z_{\rm obs}$ for S/N = 30 and 10, respectively.

The retrieved probability density functions of the other physical parameters of interest to us are difficult to interpret when averaged over galaxies at all redshifts between 0 and 1, because of the different effects impacting the recovery of parameters in different redshifts ranges. For this reason, it is more instructive to explore the retrievability of M_*/L_r, $f_{\rm SFH}$, $\psi_{\rm S}$, $12 + \log({\rm O/H})$, $\hat{\tau}_V$ and μ for galaxies in different redshift ranges, for example, $0.2 < z_{\rm obs} < 0.4$ and $0.6 < z_{\rm obs} < 0.8$. We redraw 10,000 spectral energy distributions in each of these redshift ranges from the new galaxy spectral library, which we use to generate pseudo-observations in the *ugriz* photometric bands, for S/N = 30. As before, we assume that the redshift $z_{\rm obs}$ of a galaxy is unknown a priori and use all models at redshifts between 0 and 1 in the spectral library to retrieve likelihood distributions of physical parameters. According to Fig. 3.10b, $z_{\rm obs}$ can be retrieved with an uncertainty of roughly 0.05 in both redshift ranges at the adopted signal-to-noise ratio.

Fig. 3.11 shows the average probability density functions of physical parameters retrieved in this way for galaxies at redshifts $0.2 < z_{\rm obs} < 0.4$ (top row) and $0.6 < z_{\rm obs} < 0.8$ (bottom row). At low redshift, the constraints are similar to those obtained in Fig. 3.6 for galaxies at $z = 0$. The main difference is the extension of the M_*/L_r probability density functions to larger values, induced by the tail of high-redshift galaxies with large K corrections in the prior distribution (Fig. 3.9a). This tends to worsen the typical uncertainty in M_*/L_r (from 0.19 to 0.22 dex) over the range sampled by Fig. 3.6, i.e. for $\log[(M_*/L_r)/(M_\odot/L_{r,\odot})] < 0.4$, while at larger M_*/L_r, the smaller number of models implies a significant bias in the retrieved likelihood distribution. Another noticeable difference caused by the absence of redshift information between Fig. 3.11 (top row) and Fig. 3.6 is the slightly worse uncertainty and bias in the retrieved specific star formation rate of actively star-forming galaxies. Fig. 3.11 (bottom row) also shows that, at higher redshift, the loss of Hα information further worsens $\psi_{\rm S}$ estimates in these galaxies. The constraints

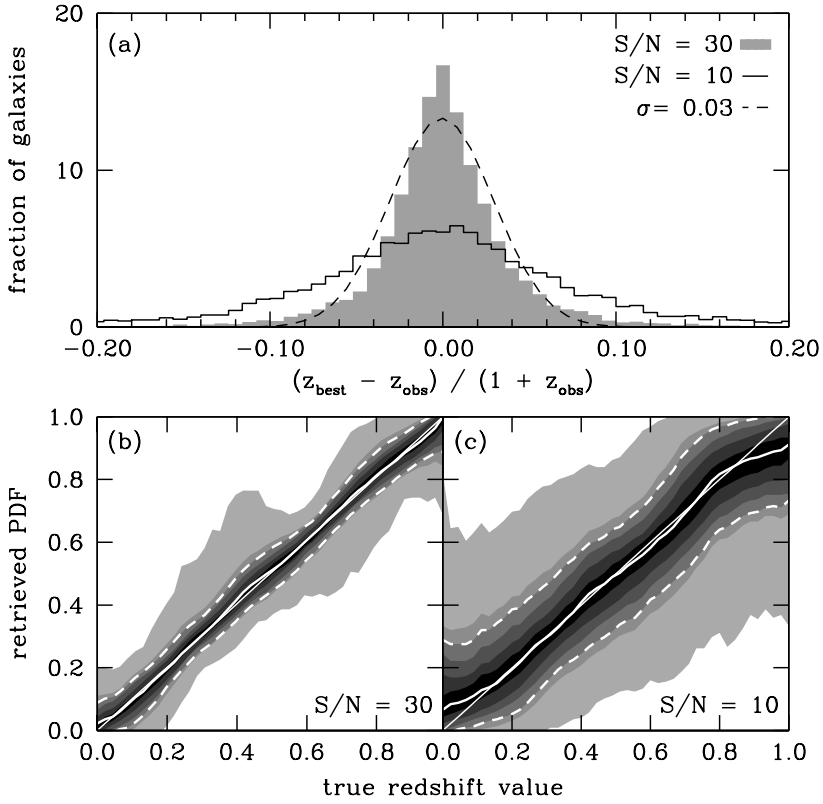

Figure 3.10: (*a*) Distribution of the difference between the retrieved best estimate (i.e. median) and true value of $z_{\rm obs}$, in units of $1+z_{\rm obs}$, for a sample of 10,000 pseudo-galaxies at redshifts between 0 and 1 observed in the *ugriz* photometric bands. The shaded and solid histograms show the distributions obtained for S/N = 30 and S/N = 10, respectively. For reference, a dashed line indicates the $\sigma = 0.03$ Gaussian accuracy quoted by Ilbert et al. [2006], who analyzed CFHTLS (S/N $\gtrsim$ 30) $u^*g'r'i'z'$ photometry of about 3000 galaxies with spectroscopic redshifts between 0.2 and 1.5. (*b*) Detail of the average retrieved probability density function plotted against $z_{\rm obs}$, for S/N = 30. The lines and shading have the same meaning as in Fig. 3.6. (*c*) Same as (*b*), but for S/N = 10.

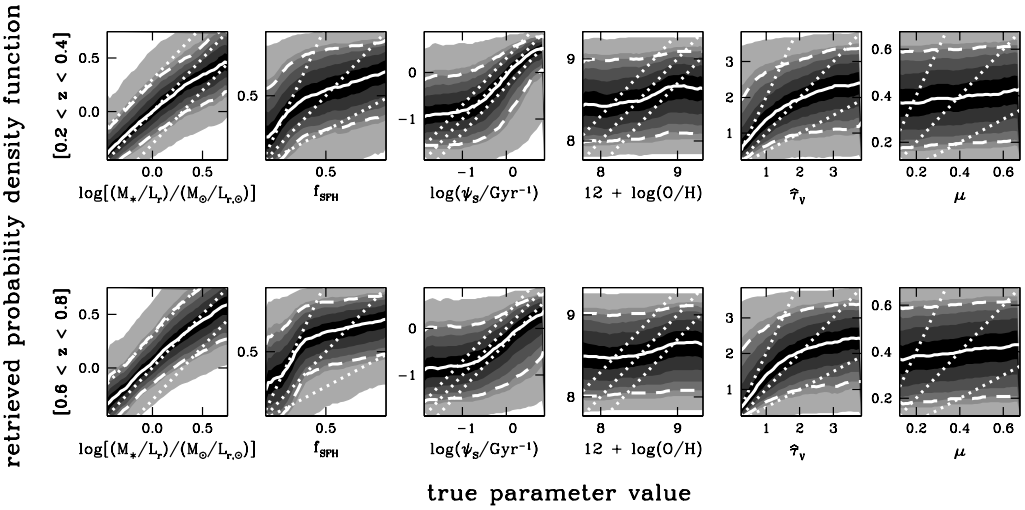

Figure 3.11: Average probability density functions of the same 6 physical parameters as in Fig. 3.9 retrieved, using the Bayesian approach described in Section 3.2.4, from 5-band *ugriz* photometry with S/N = 30 of samples of 10,000 pseudo-galaxies at random redshifts: (*top row*) drawn in the redshift range $0.2 < z_{\rm obs} < 0.4$; and (*bottom row*) drawn in the redshift range $0.6 < z_{\rm obs} < 0.8$. In each case, the redshift $z_{\rm obs}$ of a galaxy is assumed not to be known a priori, and the probability density functions are computed using models at all redshifts between 0 and 1 in the spectral library assembled in Section 3.3.1.2. In all panels, the lines and shading have the same meaning as in Fig. 3.6.

on M_*/L_r are similar to those for low-redshift galaxies. In both redshift ranges in Fig. 3.11, the parameters $f_{\rm SFH}$, $12+\log({\rm O/H})$, $\hat{\tau}_V$ and μ are as poorly constrained as for $z=0$ galaxies in Fig. 3.6.

We conclude that high-quality photometry in the *ugriz* bands can provide useful constraints on the mass-to-light ratio and specific star formation rate of galaxies, especially at low redshift, but less so on the recent history of star formation and the enrichment in metals and dust of the interstellar medium. The lack of redshift information can significantly weaken these constraints.

3.3.2 Spectroscopic constraints

We now investigate how the constraints obtained above from multi-band photometry on the parameters M_*/L_r, $f_{\rm SFH}$, $\psi_{\rm S}$, $12+\log({\rm O/H})$, $\hat{\tau}_V$ and μ can be improved by appealing to different types of spectroscopic observations. In Section 3.3.2.1 below, we first show that the equivalent widths of strong optical emission lines alone can provide useful clues about the specific star formation rate and enrichment of the interstellar medium in star-forming galaxies. Then, in Sections 3.3.2.2 and 3.3.2.3, we explore in detail the constraints retrievable from low- and medium-resolution spectroscopy, exploiting the ability with our approach to interpret simultaneously the stellar and nebular emission from galaxies. Throughout this section, we assume that the galaxies have well-determined spectroscopic redshifts and perform all calculations at rest-frame wavelengths in the range between 3600 and 7400 Å (Fig. 3.3).

3.3.2.1 Equivalent widths of strong optical emission lines

The equivalent widths of optical emission lines reflect the ratio of young stars, which produce ionizing radiation, to older stars, which dominate the optical continuum radiation. These equivalent widths should therefore provide valuable information about the specific star formation rate and, if several lines are available, the gas-phase oxygen abundance and dust attenuation optical depth in star-forming galaxies. Since line equivalent widths do not provide any information about the absolute stellar continuum emission, they cannot be used to constrain the parameters M_*/L_r nor $f_{\rm SFH}$. Similarly, line equivalent widths should not provide any useful constraint on the dust attenuation optical depth in the ambient ISM, $\mu\hat{\tau}_V$, which affects in similar ways emission lines emerging from stellar birth clouds and the continuum radiation from old stars (equation 3.2.8). We therefore focus here on the retrievability of the parameters $\psi_{\rm S}$, $12+\log({\rm O/H})$ and the dust attenuation optical depth in stellar birth clouds, $(1-\mu)\hat{\tau}_V$, from observations of the equivalent widths of strong optical emission lines in star-forming galaxies.

To proceed, we select a sample of 10,000 spectral energy distributions from the library assembled in Section 3.2.3 to compute pseudo-observations of the net emission equivalent widths of the following emission lines: [O II]λ3727; Hβ; [O III]λ4959; [O III]λ5007; [N II]λ6548; Hα; [N II]λ6584; [S II]λ6716 and [S II]λ6731. We con-

sider both the cases of low (50 Å FWHM; $R = 100$ at $\lambda = 5000$ Å) and medium (5 Å FWHM; $R = 1000$ at $\lambda = 5000$ Å) spectral resolution and fix the median signal-to-noise ratio per pixel at $\overline{\mathrm{S/N}} = 20$. In the case of low spectral resolution, we treat as single emission features the line blends: [O III]$\lambda\lambda$4959, 5007; [N II]λ6548 + Hα + [N II]λ6584 and [S II]$\lambda\lambda$6716, 6731. We also require that, at a given spectral resolution, the net emission equivalent widths of all features be measured at better than the 3σ level.[5] We apply the same requirement to the model spectra used to compute likelihood distributions of ψ_{S}, $12 + \log(\mathrm{O/H})$ and $(1-\mu)\hat{\tau}_V$ for each pseudo-galaxy. These requirements are filled by roughly 2.7 million out of the 5 million original models at low resolution, and 1.7 million models at medium resolution (where lines are not blended).

Fig. 3.12 shows the average retrieved probability density functions of ψ_{S}, $12 + \log(\mathrm{O/H})$ and $(1 - \mu)\hat{\tau}_V$ for the 10,000 star-forming pseudo-galaxies in our sample, for low (top row) and medium (bottom row) spectral resolution. We find that, as expected, the equivalent widths of strong optical emission lines set tight constraints on the specific star formation rate, ψ_{S}. The uncertainty in the retrieved best estimate is only slightly worse at low (0.23 dex) than at medium (0.18 dex) spectral resolution. An advantage of medium spectral resolution is the possibility to constrain ψ_{S} in more quiescent star-forming galaxies than at low spectral resolution, for fixed $\overline{\mathrm{S/N}}$. According to Fig. 3.12, for $\overline{\mathrm{S/N}} = 20$, the limit under which the equivalent widths of all lines of interest cannot be measured at the 3σ level is $\psi_{\mathrm{S}} \approx 0.1\,\mathrm{Gyr}^{-1}$ at $R = 100$ and $\psi_{\mathrm{S}} \approx 0.04\,\mathrm{Gyr}^{-1}$ at $R = 1000$.

Fig. 3.12 (bottom row) further shows that, at medium spectral resolution, the equivalent widths of optical emission lines provide tight constraints on the gas-phase oxygen abundance in star-forming galaxies. The uncertainty in the retrieved best estimate of $12 + \log(\mathrm{O/H})$ is only 0.02 at the high-metallicity end. It reaches 0.17 at low $12 + \log(\mathrm{O/H})$, because the rise in intrinsic emission-line intensity with decreasing metallicity [e.g., figure 4d of Charlot and Longhetti, 2001] allows the inclusion of more dusty galaxies in the likelihood distributions. At low spectral resolution, the blending of the Hα and [N II] lines worsens considerably these constraints (Fig. 3.12, top row). Moreover, the drop in emission-line intensity with increasing metallicity implies that $12 + \log(\mathrm{O/H})$ can be retrieved in more metal-rich galaxies at medium than at low spectral resolution, for fixed $\overline{\mathrm{S/N}}$. The constraints on the dust attenuation optical depth in stellar birth clouds in Fig. 3.12 are also much tighter at medium than at low spectral resolution. At low spectral resolution, the median retrieved uncertainty in $(1 - \mu)\hat{\tau}_V$ is only about 20 percent larger than at medium spectral resolution, but the best estimate is much more severely biased toward the prior expectation.

[5] To compute net emission equivalent widths in low-resolution (medium-resolution) spectra, we measure the continuum level as follows. We first de-spike the spectrum with a 9-pixel (11-pixel) median smoothing. We then mask 5-pixel-wide (9-pixel-wide) regions centred on the emission lines, through which we interpolate the spectrum. Finally, we smooth the spectrum again and adopt the result as the continuum level.

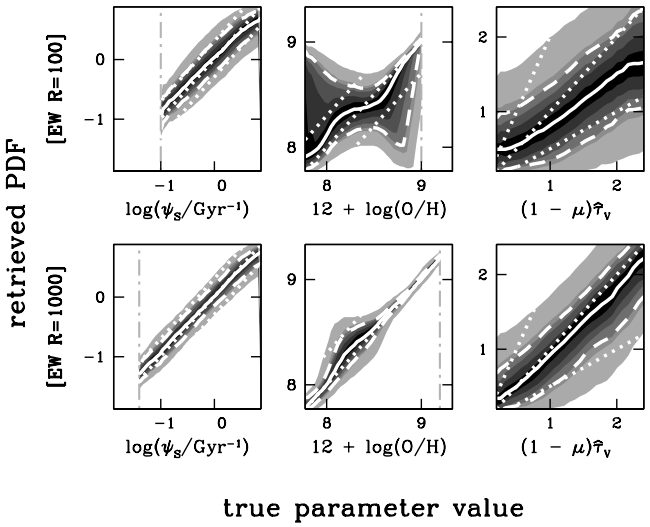

Figure 3.12: Average probability density functions of the specific star formation rate, $\psi_{\rm S}$, gas-phase oxygen abundance, $12 + \log({\rm O/H})$, and dust attenuation optical depth in stellar birth clouds, $(1-\mu)\hat{\tau}_V$, retrieved, using the Bayesian approach described in Section 3.2.4, from the equivalent widths of optical emission lines in a sample of 10,000 star-forming pseudo-galaxies. The sample was extracted from the library of galaxy spectral energy distributions assembled in Section 3.2.3, assuming a median signal-to-noise ratio per pixel $\overline{\rm S/N} = 20$ and requiring 3σ measurements of all line equivalent widths. (*Top row*) using the equivalent widths of [O II]λ3727; Hβ; [O III]$\lambda\lambda$4959, 5007; [N II]λ6548 + Hα + [N II]λ6584; and [S II]$\lambda\lambda$6716, 6731 at a spectral resolution of 50 Å FWHM ($R = 100$ at $\lambda = 5000$ Å). (*Bottom row*) using the equivalent widths of [O II]λ3727; Hβ; [O III]λ4959; [O III]λ5007; [N II]λ6548; Hα; [N II]λ6584; [S II]λ6716 and [S II]λ6731 at a spectral resolution of 5 Å FWHM ($R = 1000$ at $\lambda = 5000$ Å). In all panels, the lines and shading have the same meaning as in Fig. 3.6.

3.3.2.2 Low-resolution spectroscopy

We now turn to one of the most innovative parts of the present study enabled by the approach developed in Section 3.2: the retrieval of physical parameters from the simultaneous interpretation of strong absorption and emission features in low-resolution spectral energy distribution of galaxies over the entire rest-wavelength range from $\lambda = 3600$ to 7400 Å. As before, we extract 10,000 spectral energy distributions from the library assembled in Section 3.2.3 to compute pseudo-observations of galaxies at the resolution of 50 Å FWHM ($R = 100$ at $\lambda = 5000$ Å) in this wavelength range. We adopt for the moment a median signal-to-noise ratio per pixel $\overline{\mathrm{S/N}} = 20$. We use the rest of the 5 million models in the library to retrieve likelihood distributions of physical parameters for each of these 10,000 pseudo-galaxies.

Fig. 3.13 (top row) shows the average retrieved probability density functions of M_*/L_r, f_{SFH}, ψ_{S}, $12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and μ for the 10,000 pseudo-galaxies in our sample. Also shown at the bottom of each panel are the factor of improvement in the retrieved uncertainty, I_σ, and the gain in accuracy, Δ, relative to the standard case of Fig. 3.6 (equations 3.3.13–3.3.14). Fig. 3.13 (top row) reveals that a major improvement of low-resolution spectroscopy over rest-frame *ugriz* photometry is the ability to retrieve meaningful (i.e. unbiased) constraints not only on the stellar mass-to-light ratio and the specific star formation rate, but also on the recent star formation history (f_{SFH}) and the enrichment of metals and dust in the interstellar medium of galaxies [$12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and, to a lesser extent, μ]. The gain in accuracy (Δ) reaches up to 30 percent of the explored dynamic range for $12 + \log(\mathrm{O/H})$, 15 percent for $\hat{\tau}_V$ and μ and up to 20 percent for f_{SFH}. This results from the ability to detect and interpret, even at low resolution, the strong Balmer absorption features of intermediate-age stars and the emission-line signatures of interstellar gas. The median uncertainty in the retrieved fraction of current stellar mass formed in the last 2.5 Gyr amounts to about 0.23, that in $12 + \log(\mathrm{O/H})$ about 0.28 and that in the total dust attenuation optical depth, $\hat{\tau}_V$, about 0.64. For M_*/L_r and ψ_{S}, already constrained in a meaningful way by multiband photometry, the improvement in uncertainty relative to the standard case amounts to a factor (I_σ) of between 1.5 and 3.

A more in-depth investigation reveals that, in all panels in Fig. 3.13 (top row), the largest uncertainties pertain to galaxies with emission lines too weak to allow a distinction between the signatures of star formation history, metallicity and dust in the spectra. These are mainly galaxies with low specific star formation rate, high metallicity and strong attenuation of emission lines by dust in stellar birth clouds. In this context, it is interesting to explore how much the constraints tighten if we restrict the sample to low-resolution galaxy spectra exhibiting unambiguous emission lines. We appeal again to the library assembled in Section 3.2.3 to generate a sample of 10,000 pseudo-observations of low-resolution galaxy spectra, for $\overline{\mathrm{S/N}} = 20$, with the requirement that the net Hα+[N II] emission equivalent width be greater than 5 Å (throughout this Chapter, we adopt the convention of positive equivalent widths for emission lines). Fig. 3.13 (bottom row) shows that, in this case, the constraints

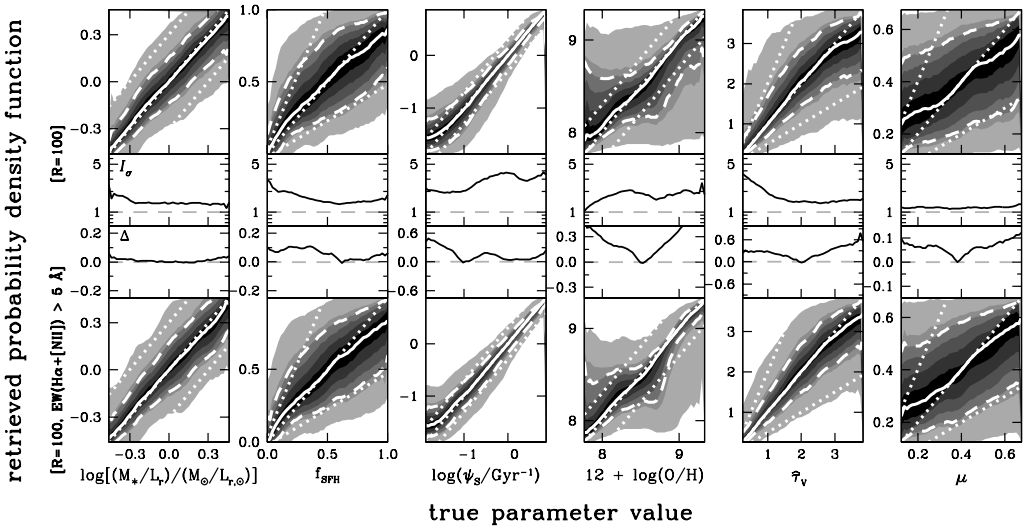

Figure 3.13: Average probability density functions of the same 6 physical parameters as in Fig. 3.4 retrieved, using the Bayesian approach described in Section 3.2.4, from low-resolution optical spectra of a sample of 10,000 pseudo-galaxies. The spectra cover the wavelength range from $\lambda = 3600$ to 7400 Å at the resolution of 50 Å FWHM ($R = 100$ at $\lambda = 5000$ Å) with median signal-to-noise ratio per pixel $\overline{\mathrm{S/N}} = 20$. (*Top row*) for a sample of 10,000 pseudo-galaxies extracted randomly from the spectral library assembled in Section 3.2.3. The improvement factor I_σ and the gain in accuracy Δ (equations 3.3.13–3.3.14) are shown as a function of true parameter value at the bottom of each panel to quantify differences in the retrieved likelihood distributions relative to the standard case of Fig. 3.6. (*Bottom row*) for a sample of 10,000 star-forming pseudo-galaxies extracted from the same spectral library, with the requirement that the net Hα+[N II] emission equivalent width be greater than 5 Å. In all panels, the lines and shading have the same meaning as in Fig. 3.6.

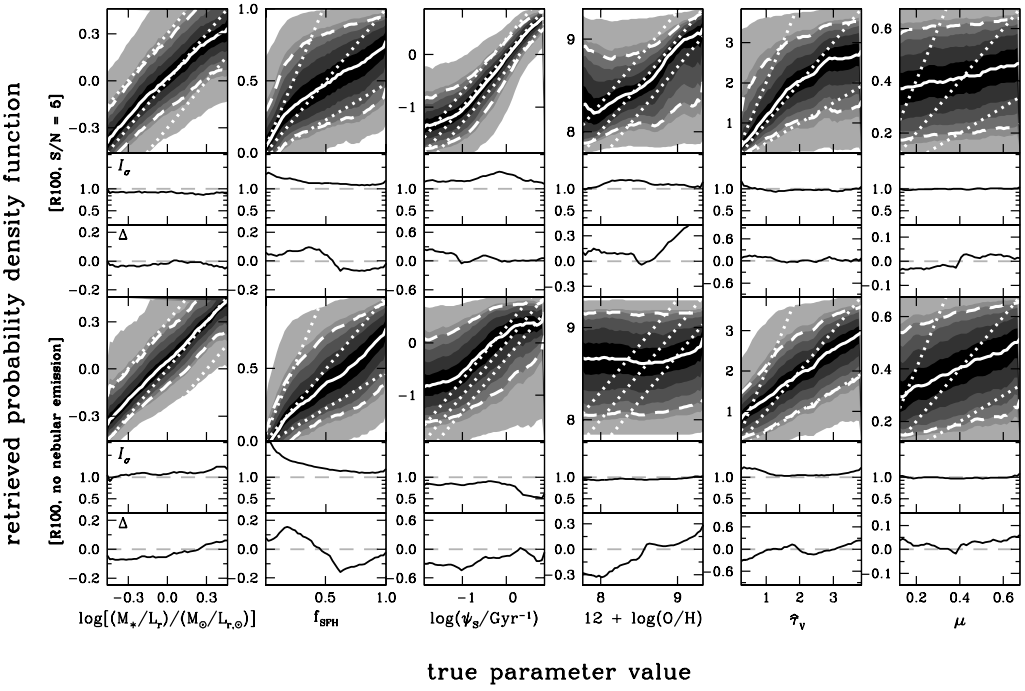

Figure 3.14: Same as Fig. 3.13, but: (*top row*) adopting a median signal-to-noise ratio per pixel $\overline{\mathrm{S/N}} = 5$ instead of 20; (*bottom row*) not including nebular emission in the model library used to analyze the sample of 10,000 pseudo-galaxies. In both cases, the improvement factor I_σ and the gain in accuracy Δ (equations 3.3.13–3.3.14) are shown as a function of true parameter value at the bottom of each panel to quantify differences in the retrieved likelihood distributions relative to the standard case of Fig. 3.6.

on the interstellar parameters $12 + \log(\mathrm{O/H})$ and $\hat{\tau}_V$ improve significantly. The corresponding median uncertainties both drop by about 20 percent. The constraints on ψ_S improve mainly at the low end of the explored range (the uncertainty dropping by about 15 percent at $\psi_\mathrm{S} \lesssim 0.1\,\mathrm{Gyr}^{-1}$), where the omission of galaxies with weak emission lines is most significant. The improvement is only marginal for M_*/L_r, f_SFH and μ, which are constrained primarily by stellar continuum and absorption-line features rather than by nebular emission lines.

The ability with our approach to exploit strong absorption and emission features in low-resolution galaxy spectra depends critically on the quality of the observations. To illustrate this, we show in Fig. 3.14 (top row) the analog of Fig. 3.13 (top row) when adopting a median signal-to-noise ratio per pixel $\overline{\mathrm{S/N}} = 5$ instead of 20. The reduced ability to interpret spectral features at $\overline{\mathrm{S/N}} = 5$ makes the median uncertainties in the different retrieved parameters increase by 30 percent (for M_*/L_r,

$f_{\rm SFH}$, μ) up to 90 percent (for $\psi_{\rm S}$) relative to the case $\overline{\rm S/N} = 20$. It also introduces substantial biases in the retrieved likelihood distributions. A comparison of Fig. 3.14 (top row) with Fig. 3.6 reveals that, in fact, low-resolution spectroscopy with $\overline{\rm S/N} = 5$ over the wavelength range from $\lambda = 3600$ to 7400 Å provides hardly better constraints on the different parameters than high-quality *ugriz* photometry (as confirmed by the fact that I_σ is close to unity and Δ close to zero in Fig. 3.13, top row). The constraints on $\psi_{\rm S}$ and $12 + \log({\rm O/H})$ are somewhat tighter because of the influence of actively star-forming galaxies with strong emission lines on the average retrieved probability density functions. The star formation history parameter is only slightly better constrained, as even strong Balmer absorption lines are difficult to interpret at low $\overline{\rm S/N}$. The constraints on $\hat{\tau}_V$ are as poor as inferred from multiband photometry, reflecting the inability to extract useful information from the Hα/Hβ ratio in low-quality spectra. In Fig. 3.14 (top row), the constraints on M_*/L_r are actually weaker and more severely biased at the extremities of the explored range than those obtained from rest-frame *ugriz* photometry in Fig. 3.6. This suggests that high-quality photometry including the near-infrared *iz* bands can better constrain galaxy stellar masses than low-resolution optical spectroscopy with $\overline{\rm S/N} = 5$.

The inclusion of nebular emission also has a crucial influence on the ability to constrain galaxy physical parameters from low-resolution optical spectroscopy. Fig. 3.14 (bottom row) shows the results obtained when attempting to interpret the same pseudo-observations as in Fig. 3.13 (top row; which include nebular emission) with the library of models not including nebular emission already used in Fig. 3.8 (bottom row).[6] In this case, the average retrieved likelihood distributions of all parameters are severely degraded relative to Fig. 3.13 (top row). In fact, for most parameters, the constraints are worse than those derived in Fig. 3.6 from high-quality *ugriz* photometry when accounting for the effect of nebular emission on photometric fluxes. Fig. 3.14 (bottom row) shows that, for example, using models which neglect the contamination of the *r*-band flux by Hα+[N II] and [S II] line emission biases M_*/L_r estimates upward for star-forming galaxies. Moreover, ignoring the contamination of stellar Balmer absorption features by nebular emission in low-resolution spectra causes systematic underestimates of $f_{\rm SFH}$ in star-forming galaxies. For actively star-forming galaxies, the constraints on $\psi_{\rm S}$ in Fig. 3.14 (bottom row) are weaker than the rough constraints derived in Fig. 3.6 from the influence of nebular emission on *ugriz* photometric observations. Not surprisingly, models without nebular emission do not allow the extraction of useful constraints on the interstellar parameters $12 + \log({\rm O/H})$, $\hat{\tau}_V$ and μ from low-resolution galaxy spectra.

Hence, we find that important constraints can be obtained on the stellar and interstellar parameters of galaxies from low-resolution spectra at optical wavelengths.

[6] To avoid artefacts in the resulting probability density functions, we remove prominent emission lines from the pseudo-galaxy spectra before computing the likelihood of models without nebular emission (equation 3.2.10). To achieve this, we first measure the continuum level across the whole range from $\lambda = 3600$ to 7400 Å, as described in footnote 5. Then, we reset to the continuum level the flux of every pixel exceeding the continuum flux by more than 5 times the noise level.

This however requires high-quality observations as well as the inclusion of nebular emission in the spectral analysis.

3.3.2.3 Medium-resolution spectroscopy

Medium-resolution spectroscopy should provide even more insight into the physical properties of galaxies than low-resolution spectroscopy, because composite features such as the [O III] doublet, the Hα+[N II] blend and the [S II] doublet can be resolved into individual emission lines. To investigate this, we extract 10,000 spectral energy distributions from the library assembled in Section 3.2.3 and compute pseudo-observations of galaxies at the resolution of 5 Å FWHM ($R = 1000$ at $\lambda = 5000$ Å) over the wavelength range from $\lambda = 3600$ to 7400 Å. We adopt a median signal-to-noise ratio per pixel $\overline{\mathrm{S/N}} = 20$. As before, we use the rest of the 5 million models in the library to retrieve likelihood distributions of physical parameters for each of these 10,000 pseudo-galaxies.

In Fig. 3.15 (top row), we show the average probability density functions of M_*/L_r, f_{SFH}, ψ_{S}, $12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and μ retrieved from the medium-resolution spectra of the 10,000 pseudo-galaxies in our sample. The factor I_σ and the gain Δ quantifying the improvements in the retrieved uncertainty and accuracy relative to the standard case of Fig. 3.6 (equations 3.3.13–3.3.14) are also shown at the bottom of each panel. A comparison of Fig. 3.15 (top row) with Fig. 3.13 (top row) shows that an increase in spectral resolution by a factor of 10 produces significantly tighter constraints on all parameters. The improvement is most significant for the specific star formation rate, the gas-phase oxygen abundance and the total dust attenuation optical depth, for which the median retrieved uncertainty drops by about 45 percent from low to medium spectral resolution. The gain is more modest, of the order of 25 percent, for M_*/L_r, f_{SFH}, and μ. The refined spectral information available at medium resolution also help reduce some biases in the retrieved best estimates of f_{SFH} (for young galaxies), ψ_{S} (for quiescent star-forming galaxies) and μ (across most of the explored range) in Fig. 3.15 (top row) relative to Fig. 3.13 (top row).

By analogy with Fig. 3.13, we further show in Fig. 3.15 (bottom row) the average probability density functions obtained when selecting a sample of 10,000 pseudo-observations of medium-resolution galaxy spectra, for $\overline{\mathrm{S/N}} = 20$, with the requirement that the net Hα emission equivalent width be greater than 5 Å. The retrieved likelihood distributions of $12 + \log(\mathrm{O/H})$ and $\hat{\tau}_V$ are significantly narrower than when including also galaxies with weak emission lines in the sample, with median retrieved uncertainties of about 0.08 for $12 + \log(\mathrm{O/H})$ and 0.30 for $\hat{\tau}_V$. As at low spectral resolution, restricting the sample to galaxies with prominent emission lines reduces the uncertainties in ψ_{S} mainly at the low end of the explored range. Changes in the retrieved likelihood distributions of M_*/L_r, f_{SFH} and μ are marginal, because these parameters are constrained primarily by stellar continuum and absorption-line features.

We have also investigated how the constraints obtained in Fig. 3.15 from the analysis of medium-resolution galaxy spectra depend on the assumed signal-to-noise

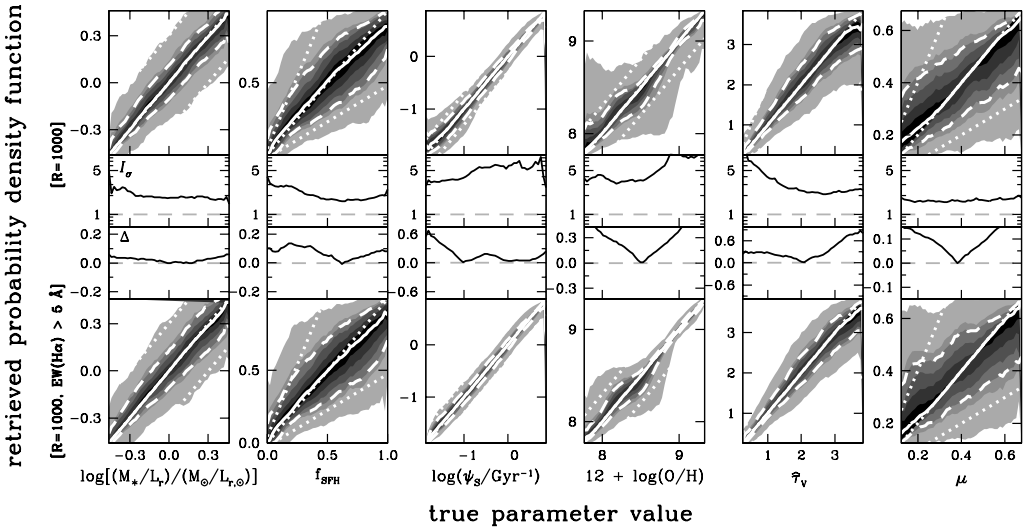

Figure 3.15: Average probability density functions of the same 6 physical parameters as in Fig. 3.4 retrieved, using the Bayesian approach described in Section 3.2.4, from medium-resolution optical spectra of a sample of 10,000 pseudo-galaxies. The spectra cover the wavelength range from $\lambda = 3600$ to 7400 Å at the resolution of 5 Å FWHM ($R = 1000$ at $\lambda = 5000$ Å) with median signal-to-noise ratio per pixel $\overline{\mathrm{S/N}} = 20$. (*Top row*) for a sample of 10,000 pseudo-galaxies extracted randomly from the spectral library assembled in Section 3.2.3. The improvement factor I_σ and the gain in accuracy Δ (equations 3.3.13–3.3.14) are shown as a function of true parameter value at the bottom of each panel to quantify differences in the retrieved likelihood distributions relative to the standard case of Fig. 3.6. (*Bottom row*) for a sample of 10,000 star-forming pseudo-galaxies extracted from the same spectral library, with the requirement that the net Hα emission equivalent width be greater than 5 Å. In all panels, the lines and shading have the same meaning as in Fig. 3.6.

ratio of the observations and the inclusion of nebular emission in the modeling (not shown). We find that adopting a median signal-to-noise ratio per pixel $\overline{\mathrm{S/N}} = 5$ instead of 20 for the 10,000 pseudo-galaxy spectra makes the median retrieved uncertainties in most parameters increase by 30 to 50 percent (only 15 percent for M_*/L_r) relative to Fig. 3.15 (top row). The biases in the retrieved best estimates are less severe than found in Fig. 3.14 (top row) at low spectral resolution. As expected from Fig. 3.14 (bottom row), accounting for nebular emission also has a critical influence on the retrievability of the specific star formation rate and the interstellar parameters from medium-resolution galaxy spectra. It is worth pointing out that, at medium (and high) spectral resolution, prominent nebular emission lines can be excised from the spectra and analyzed separately from the stellar continuum and absorption-line signatures if no tool is available to analyze the stellar and nebular emission simultaneously [e.g., Kauffmann et al., 2003a, Brinchmann et al., 2004, Gallazzi et al., 2005]. This is not the case at low spectral resolution, where the analysis of the stellar and nebular emission must be combined to obtain meaningful constraints on galaxy physical parameters.

3.4 Application to a real sample

The conclusions drawn in the previous section about the retrievability of galaxy physical parameters from different types of observations at optical wavelengths rely on sophisticated simulations of galaxy spectral energy distributions. As emphasized in Section 3.2.4, such pseudo-observations are required to test the retrievability of physical parameters from observables, because the physical parameters of real galaxies cannot be known a priori. In this section, we apply our approach to the interpretation of observed medium-resolution spectra of galaxies from the SDSS DR7. We compare the results obtained through the simultaneous interpretation of stellar and nebular emission (at the resolving power $R \approx 1000$ considered in Section 3.3.2.3) with results obtained for the same galaxies from previous separate analyses of nebular emission-line and stellar absorption-line features (at the native resolving power $R \approx 2000$ of the SDSS observations). We also investigate the influence of the adopted prior distributions on the retrieved physical parameters of the galaxies.

3.4.1 Physical parameters of SDSS galaxies

The SDSS DR7 contains about 1 million galaxy spectra covering the wavelength range from 3800 to 9200 Å (observer frame) at a spectral resolution of roughly 3 Å FWHM. Of these, we extract a subsample of 12,660 star-forming galaxies with high-quality spectra as follows. We first reject any duplicate spectrum of a same galaxy, all spectra of merger candidates (i.e. corresponding to spatial separations less than $1.5''$ and redshift separations less than $200\,\mathrm{km\,s^{-1}}$) and all galaxies at redshift less than $z = 0.03$ (to ensure that the spectrum includes the [O II]λ3727 doublet) from the original sample. Among the remaining 829,570 galaxies, we select

all emission-line galaxies with signal-to-noise ratio greater than 10 in [O II]λ3727, Hβ, [O III]λ5007, Hα, [N II]λ6584 and [S II]λ6716. We reject AGNs by applying the conservative criterion of Kauffmann et al. [2003b] in the standard Baldwin et al. [1981] line-diagnostic diagram (see Fig. 3.2a above). The resulting 28,075 star-forming galaxies have a mean redshift of 0.08 ± 0.05. We exclude all galaxies at redshift $z > 0.15$ to keep the sample homogeneous and further select galaxies with median signal-to-noise ratio per pixel in the range $5 < \overline{\mathrm{S/N}} < 25$. This leaves 23,213 spectra of galaxies at a mean redshift of 0.07 ± 0.03, all of which have net Hα emission equivalent widths greater than 5 Å. We identify 12,660 galaxies in this sample for which good measurements of stellar mass, star formation rate, gas-phase oxygen abundance and dust attenuation optical depth are available from previous studies (see next paragraph). For the purpose of illustrating the usefulness of our approach to interpret galaxy spectra, we focus on the galaxy physical properties probed by the 3-arcsec-diameter SDSS fibers and do not consider any photometry-based aperture correction.

Estimates of star formation rate, gas-phase oxygen abundance and dust attenuation optical depth are available for the 12,600 SDSS galaxies in this sample from the work of Brinchmann et al. [2004] and Tremonti et al. [2004]. These estimates were obtained by fitting the fluxes of the [O II]λ3727, Hβ, [O III]λ5007, Hα, [N II]λ6584 and [S II]λ6716 emission lines (measured to high precision by adjusting a stellar population model to the spectral continuum) with a large library of Charlot and Longhetti [2001] nebular emission models. In addition, stellar-mass estimates are available for all galaxies from the work of Gallazzi et al. [2005], who fitted the strengths of the 4000 Å break and the Hβ, Hδ_A+Hγ_A, [MgFe]$'$ and [$\mathrm{Mg_2Fe}$] absorption-line features with a large library of Bruzual and Charlot [2003] stellar population synthesis models. The combination of these stellar-mass estimates with the star-formation-rate estimates of Brinchmann et al. [2004] provides estimates of the specific star formation rate probed by the SDSS fiber.

The above physical parameters were derived from spectral analyses at the original SDSS resolution of about 3 Å FWHM. To exemplify the power of the approach developed in Section 3.2, and for the purpose of comparison with the results of Section 3.3.2.3, we degrade the SDSS spectra to a resolution of 5 Å FWHM, adopting as before a fixed pixel size of 2.5 Å. Also, we focus on the rest-frame wavelength range from 3700 to 7400 Å, which does not extend as much in the ultraviolet as that considered in Section 3.3.2.3 (Fig. 3.3a), but which does include the [O II]λ3727 emission doublet. By analogy with the analysis of pseudo-observations in Section 3.3.2.3, we use the 5 million models in the spectral library assembled in Section 3.2.3 to retrieve likelihood distributions of physical parameters from the simultaneous analysis of the stellar and nebular emission (equations 3.2.10–3.2.12) for each SDSS galaxy in the sample. We note that, when computing likelihood distributions for a given galaxy, we exclude from the library all models older than the age of the Universe at the redshift of that galaxy.

In Fig. 3.16, we show how the probability density functions of stellar mass, gas-phase oxygen abundance, total V-band attenuation optical depth of the dust and

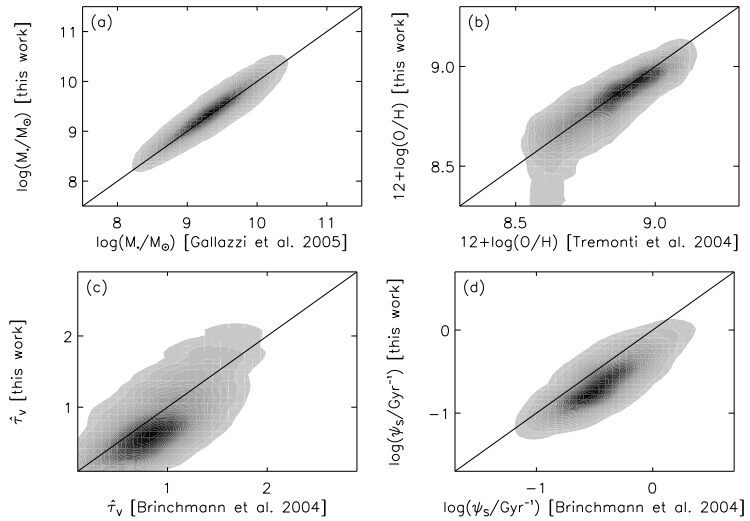

Figure 3.16: Estimates of physical parameters retrieved from the spectra of 12,660 SDSS star-forming galaxies (degraded to a resolution of 5 Å FWHM) using the Bayesian approach described in Section 3.2.4 plotted against estimates of the same parameters from different sources, as indicated. (*a*) Stellar mass, $M_\star$. (*b*) Gas-phase oxygen abundance, $12 + \log(\mathrm{O/H})$. (*c*) Total V-band attenuation optical depth of the dust, $\hat{\tau}_V$. (*d*) Specific star formation rate, ψ_S. In each panel, the likelihood distributions from previous studies were reconstructed from the published 16th, 50th and 84th percentiles and combined with the likelihood distributions obtained in this work to generate 2D probability density functions. The contours depict the normalized co-added 2D probability density function including all 12,660 galaxies in the sample (on a linear scale 25 levels, the outer edge of the lightest grey level corresponding to the 96th percentile of the 2D distribution). The solid line is the identity relation.

specific star formation rate retrieved from our analysis compare with those obtained in previous separate studies of the stellar and nebular emission at higher spectral resolution. Fig. 3.16a shows that, for example, our estimates of $M_\star$ (obtained by multiplying M_*/L_r by the absolute galaxy luminosity in the observed r band) agree well with those derived by Gallazzi et al. [2005] from the analysis of selected absorption-line indices. The slightly lower values derived here (~ 0.1 dex) are attributable to the updated stellar population synthesis prescription (Section 3.2.2.1). We also find good agreement in Fig. 3.16b between our estimates of $12 + \log(\mathrm{O/H})$ and those obtained by Tremonti et al. [2004] from the analysis of selected emission-lines luminosities. The discrepancies at $12 + \log(\mathrm{O/H}) < 8.7$ are consistent with the uncertainty expected at low metallicity from interpretations of medium-resolution galaxy spectra (Fig. 3.15, bottom row). In Figs 3.16c and 3.16d, the likelihood distributions retrieved from our analysis favor systematically lower $\hat{\tau}_V$ and ψ_S than derived by Brinchmann et al. [2004] from the analysis of emission-line luminosities. This is because Brinchmann et al. [2004] adopt a shallower dust attenuation curve ($\hat{\tau}_\lambda \propto \lambda^{-0.7}$ both in stellar birth clouds and in the ambient ISM) than that resulting typically from equations (3.2.7)–(3.2.9). This makes the inferred dust optical depth and the correction for obscured star formation smaller at fixed observed reddening.

Fig. 3.16 therefore shows that the likelihood distributions of $M_\star$, $12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and ψ_S retrieved using our approach agree well (modulo some understood offsets) with those obtained in previous separate analyses of the stellar and nebular emission of SDSS galaxies at twice higher spectral resolution. Such an agreement is expected in part from the fact that these previous analyses rely on earlier versions of the same stellar population synthesis code and nebular emission models as adopted here (Bruzual and Charlot 2003, Charlot and Longhetti 2001; see Sections 3.2.2.1 and 3.2.2.2). What is most remarkable is the demonstrated ability with our approach to interpret simultaneously the stellar and nebular emission from galaxies without having to excise nebular emission lines from noisy spectra. As shown in Section 3.3, this ability extends to low-resolution galaxy spectra and even to multi-band photometry.

It is of interest to see how the fraction of the current stellar mass formed during the last 2.5 Gyr, which has not been investigated previously, relates to the other physical parameters of SDSS galaxies retrieved using our approach. In Fig. 3.17, we show the 2D probability density functions of f_SFH versus $\log M_\star$, $\log \psi_\mathrm{S}$, $12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and μ for the 12,660 SDSS star-forming galaxies in our sample. The Spearman rank correlation coefficients (evaluated using the median estimates of the different parameters) are $r_s = -0.27$ for the relation between f_SFH and $\log M_\star$ and $+0.59$ for that between f_SFH and $\log \psi_\mathrm{S}$. For this sample size, both correlations are significant at better than the 5σ level. In contrast, none of the interstellar parameters $12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and μ correlates significantly with f_SFH (we note in passing that the required high-S/N detection of 6 emission lines tends to select galaxies with low $\hat{\tau}_V$ and high μ). Taken at face value, the results of Fig. 3.17 would suggest that the least massive galaxies, which are the most actively star-forming ones today, have also formed the largest proportion of stars over the

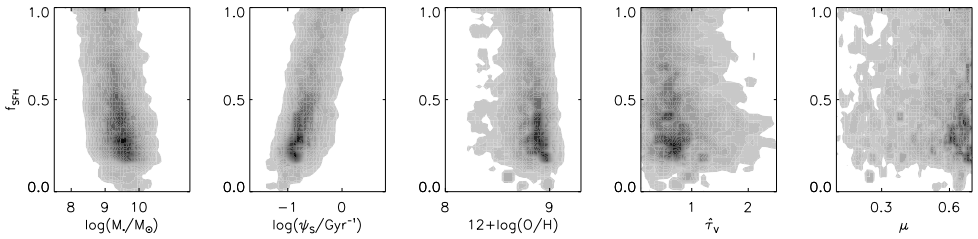

Figure 3.17: Two-dimensional probability density functions of the fraction of the current stellar mass formed during the last 2.5 Gyr, $f_{\rm SFH}$, versus stellar mass, $M_\star$, specific star formation rate, $\psi_{\rm S}$, gas-phase oxygen abundance, $12 + \log({\rm O/H})$, total V-band attenuation optical depth of the dust, $\hat{\tau}_V$, and fraction of $\hat{\tau}_V$ arising from dust in the ambient ISM, μ. All parameters were retrieved from the spectra of the same 12,660 SDSS star-forming galaxies as in Fig. 3.16 (degraded to a resolution of 5 Å FWHM) using the Bayesian approach described in Section 3.2.4. In each panel, the contours depict the normalized co-added 2D probability density function including all 12,660 galaxies in the sample (on a linear scale of 25 levels, the outer edge of the lightest grey level corresponding to the 96th percentile of the 2D distribution). All physical quantities pertain to the regions of galaxies probed by the 3-arcsec-diameter SDSS fibers.

last 2.5 Gyr. Such a trend would be consistent with the general expectation that, as galaxies grow through star formation, the exhaustion of the gas supply makes the rate of star formation drop gradually (see also, e.g., fig. 22 of Brinchmann et al. 2004). However, this conclusion must be taken with caution, as the present sample is not complete and the SDSS fibers probe only the inner parts of galaxies.

3.4.2 Influence of the prior distributions of physical parameters

The constraints derived above on the stellar and interstellar properties of SDSS galaxies, along with the analyses of pseudo-observations carried out in Section 3.3, rely on the adoption of the prior distributions of galaxy physical parameters in Fig. 3.4. As outlined in Section 3.2.4, these distributions are designed to be as flat as possible to make the spectral library appropriate for the retrieval of physical parameters from observed spectral energy distributions of galaxies in wide ranges of physical properties. To illustrate the importance of this choice, we examine here the results obtained when adopting different prior distributions reflecting more specific properties of the galaxy population. As an example, we consider the prior distributions of galaxy physical parameters derived from the original semi-analytic recipes of De Lucia and Blaizot [2007].

In Section 3.2.1, we used the semi-analytic recipes of De Lucia and Blaizot [2007] to generate the star formation and chemical enrichment histories of 500,000 galaxies extracted from the Millennium cosmological simulation of Springel et al. [2005]. The parameters M_*/L_r, $f_{\rm SFH}$, $\psi_{\rm S}$ and the interstellar metallicity Z are computed

by default in this procedure. De Lucia and Blaizot [2007] do not provide any recipe to compute the gas-phase oxygen abundance nor the ionization parameter of the gas heated by young stars. Also, while attenuation by dust is included using the 2-component model of Charlot and Fall [2000, adopting a fixed attenuation curve $\hat{\tau}_\lambda \propto \lambda^{-0.7}$ both in stellar birth clouds and in the ambient ISM], we find that the recipe used by De Lucia and Blaizot [2007] to relate the V-band dust optical depth to the galaxy gas mass, metallicity and orientation produces typically much lower $\hat{\tau}_V$ than observed in SDSS galaxies. To compute prior distributions of interstellar parameters associated to the original semi-analytic recipes of De Lucia and Blaizot [2007], we therefore proceed as follows. For each of the 500,000 galaxies in the library, we draw 10 different realizations of $\log U_0$, ξ_d and μ from the distributions in Table 3.1. We also draw 10 different realizations of $\hat{\tau}_V$ from a Gaussian distribution centred on 1.2 with a standard deviation of 0.6. This roughly represents the global properties of SDSS galaxies for a fixed attenuation curve $\hat{\tau}_\lambda \propto \lambda^{-0.7}$ in stellar birth clouds and in the ambient ISM [Brinchmann et al., 2004]. The new spectral library contains 5 million models, which we can use to interpret observations of galaxy spectral energy distributions in the same way as the library assembled in Section 3.2.3.

Fig. 3.18 shows the prior distributions of M_*/L_r, f_SFH, ψ_S, $12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and μ in this library at the mean redshift $z = 0.07$ of the SDSS sample of Section 3.4.1. In each panel, the shaded histogram shows the distribution including all galaxies, while the thin solid histogram shows the contribution by star-forming galaxies alone. As noted in Section 3.2.1, 35 percent of the galaxies in the original library of De Lucia and Blaizot [2007] do not form stars at low redshift, while the distributions of ψ_S and f_SFH for the remaining galaxies are narrow and characteristic of quiescent star formation (Figs 3.18b and 3.18c). As a result, most galaxies have relatively large M_*/L_r (Fig. 3.18a). In Fig. 3.18d, the narrow distribution of $12 + \log(\mathrm{O/H})$ arises from the correlation of the interstellar metallicity Z with ψ_S for star-forming galaxies in this library. Also shown as thick solid histograms in Fig. 3.18 are the distributions of the best estimates of M_*/L_r, ψ_S, $12 + \log(\mathrm{O/H})$ and $\hat{\tau}_V$ from Gallazzi et al. [2005], Brinchmann et al. [2004] and Tremonti et al. [2004] for the 12,660 SDSS star-forming galaxies of Section 3.4.1. The required detection of 6 emission lines with $\mathrm{S/N} > 10$ in these galaxies implies that they are typically more actively star-forming (and hence have lower M_*/L_r) and less dusty than the bulk of SDSS galaxies the Millennium cosmological simulation is expected to represent. This is the origin of the differences between the shaded and thick solid histograms in Fig. 3.18.

In Fig. 3.19, we show the analog of Fig. 3.16 obtained when using the prior distributions in Fig 3.18, instead of those in Fig. 3.4, to retrieve the probability density functions of $M_\star$, $12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and ψ_S for the 12,660 SDSS star-forming galaxies in our sample. The upward bias in the prior distribution of M_*/L_r relative to the observed sample (Fig. 3.18a) causes the estimates of $M_\star$ in Fig. 3.19a to be systematically larger than those derived by Gallazzi et al. [2005]. In Fig. 3.19d, the retrieved ψ_S is about 60 percent smaller than that estimated by Brinchmann et al. [2004], as in Fig. 3.16d. The offset in this case results from the downward bias in

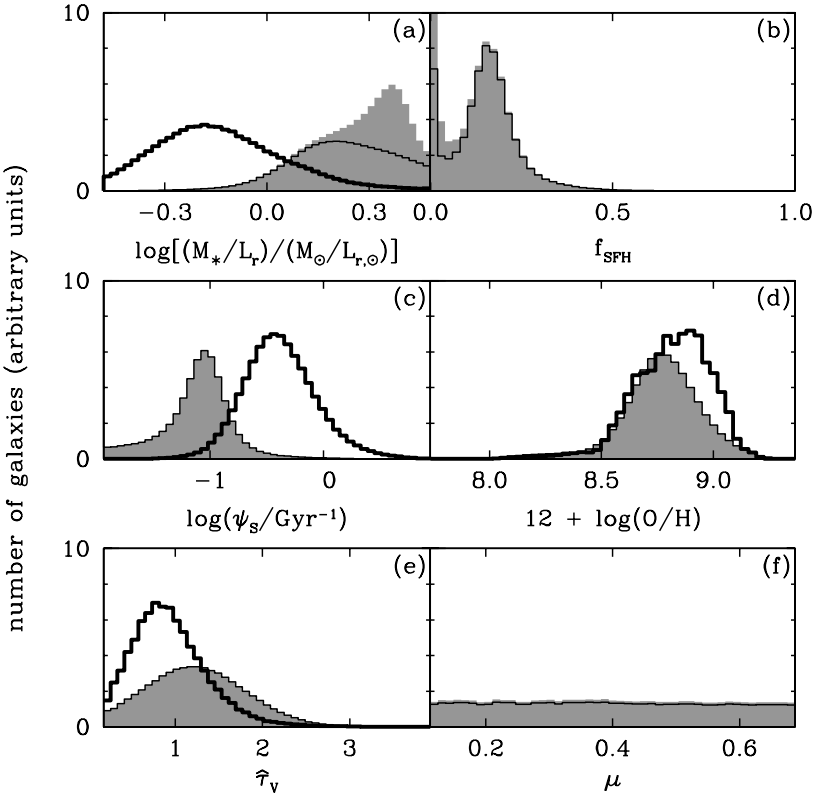

Figure 3.18: Prior distributions of the same physical parameters as in Fig. 3.4 for the 5 million galaxies in the spectral library generated in Section 3.4.2. These distributions are those arising, at the mean redshift $z = 0.07$ of the SDSS sample of Section 3.4.1, from the original semi-analytic post-treatment of the Millennium cosmological simulation of Springel et al. [2005] by De Lucia and Blaizot [2007, see text for detail]. In each panel, the shaded and thin solid histograms have the same meaning as in Fig. 3.4. In panels (*a*), (*c*), (*d*) and (*e*), the thick solid histograms show the distributions of the best estimates of M_*/L_r, $\psi_{\rm S}$, $12 + \log({\rm O/H})$ and $\hat{\tau}_V$ from Gallazzi et al. [2005], Brinchmann et al. [2004] and Tremonti et al. [2004] for the same sample of 12,660 SDSS star-forming galaxies as in Fig. 3.16.

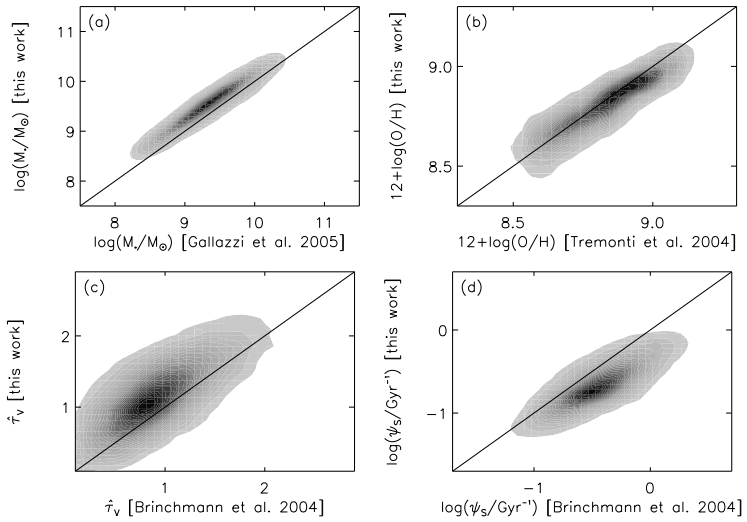

Figure 3.19: Same as Fig. 3.16, but using the prior distributions in Fig 3.18 instead of those in Fig. 3.4 to retrieve the probability density functions of $M_\star$, $12 + \log(\mathrm{O/H})$, $\hat{\tau}_V$ and ψ_S for the 12,660 SDSS star-forming galaxies in the sample.

the prior distribution of ψ_S (Fig. 3.18c) rather than from differences in the dust attenuation curve, which is assumed to be the same by De Lucia and Blaizot [2007] and Brinchmann et al. [2004]. Meanwhile, the upward bias in the prior distribution of $\hat{\tau}_V$ (Fig. 3.18e) causes estimates of this parameter in Fig. 3.19c to be systematically larger than those derived by Brinchmann et al. [2004]. In Fig. 3.19b, the estimates of $12 + \log(\mathrm{O/H})$ retrieved using the prior distribution of Fig. 3.18d agree well with those derived by Tremonti et al. [2004].

The results of Figs 3.16 and 3.19 illustrate the importance of the choice of appropriate prior distributions when applying the Bayesian approach described in Section 3.2.4 to retrieve physical parameters from observed spectral energy distributions of galaxies. While the prior distributions of Fig. 3.18 are expected to be representative of the SDSS sample as a whole, adopting these distributions to analyze subsamples of galaxies with properties different from the bulk can induce unwanted offsets in the retrieved physical parameters. In cases where some specific physical property of the observed galaxies can be known in advance (e.g., actively star-forming, early-type, metal-poor), the use of appropriate restrictive prior distributions can help reduce the uncertainties in the retrieved parameters. When no advance information is available, it is preferable to appeal to flat (maximum-ignorance) prior distributions encompassing wide ranges of galaxy physical properties.

Table 3.2: Summary of the constraints on galaxy physical parameters retrieved from different types of optical observations using the method developed in Section 3.2. For each parameter and each type of observation, the quoted uncertainty (defined as half the 16th–84th percentile range of the retrieved likelihood distribution) and accuracy (defined as the absolute difference between the retrieved best estimate and true parameter value) are median quantities determined from the analysis of 10,000 pseudo-observations in Section 3.3.

Parameter	Constraint	Photometry[a]	Net emission-line EWs[b]		Spectroscopy[b]			
		rest-frame *ugriz*	$R = 100$	$R = 1000$	$R = 100$		$R = 1000$	
		All[c]	SF[d]	SF[e]	All[c]	SF[f]	All[c]	SF[g]
$\log[(M_*/L_r)/(M_\odot/L_{r,\odot})]$	Uncertainty	0.19	...	...	0.13	0.14	0.10	0.11
	Accuracy	0.04	...	...	0.01	0.01	0.00	0.00
$f_{\rm SFH}$	Uncertainty	0.36	...	...	0.23	0.23	0.17	0.19
	Accuracy	0.17	...	...	0.05	0.05	0.02	0.03
$\log(\psi_{\rm S}/{\rm Gyr}^{-1})$	Uncertainty	0.67	0.23	0.18	0.24	0.22	0.12	0.12
	Accuracy	0.15	0.12	0.05	0.01	0.00	0.00	0.00
$12 + \log({\rm O/H})$	Uncertainty	0.50	0.23	0.05	0.28	0.21	0.13	0.08
	Accuracy	0.33	0.11	0.03	0.04	0.03	0.00	0.00
$\hat{\tau}_V$	Uncertainty	0.91	0.48[h]	0.41[h]	0.64	0.52	0.35	0.30
	Accuracy	0.21	0.57[h]	0.14[h]	0.05	0.00	0.00	0.00
μ	Uncertainty	0.20	...	...	0.16	0.15	0.10	0.10
	Accuracy	0.10	...	...	0.04	0.02	0.00	0.00

3.5 Summary and conclusion

We have developed a new approach to constrain galaxy physical parameters from the combined interpretation of stellar and nebular emission in wide ranges of photometric and spectroscopic observations. Our approach relies on the simulation of a comprehensive library of 5 million galaxy spectral energy distributions using a set of state-of-the-art models of star formation and chemical enrichment histories, stellar population synthesis, nebular emission and attenuation by dust. This library can be used to retrieve probability density functions of physical parameters from the Bayesian analysis of any type of multi-wavelength galaxy observation. We focus in this Chapter on the retrievability of the observer-frame absolute r-band stellar mass-to-light ratio (M_*/L_r), the fraction of a current galaxy stellar mass formed during the last 2.5 Gyr ($f_{\rm SFH}$), the specific star formation rate ($\psi_{\rm S}$), the gas-phase oxygen abundance $[12+\log({\rm O/H})]$, the total effective V-band absorption optical depth of the dust ($\hat{\tau}_V$) and the fraction of this arising from dust in the ambient ISM (μ) from different types of observations at optical wavelengths: 5-band *ugriz* photometry; equivalent-width measurements of strong emission lines; low-resolution spectroscopy (50 Å FWHM over the wavelength range λ=3600–7400 Å); and medium-resolution spectroscopy (5 Å FWHM over the same wavelength range).

Since we cannot know the true properties of any sample of observed galaxies, assessing the retrievability of galaxy physical parameters from these different types of observations requires the simulation of pseudo-observations. We therefore simulate pseudo-observations by convolving the spectral energy distributions of models with known parameters with appropriate instrument responses and then applying artificial noise to mimic true observations. For each type of photometric or spectroscopic observation, we simulate 10,000 pseudo-observations in this way and use the other $\sim$ 5 million models in the library to retrieve likelihood distributions of M_*/L_r, $f_{\rm SFH}$, $\psi_{\rm S}$, $12+\log({\rm O/H})$, $\hat{\tau}_V$ and μ for each pseudo-galaxy (using equations 3.2.10–3.2.11). Table 3.2 summarizes the main conclusions of our study. For each parameter and each type of observation, we list the median retrieved uncertainty (defined as half the 16th–84th percentile range of the retrieved likelihood distribution) and accuracy (defined as the absolute difference between the retrieved best estimate and true parameter value) determined from the analysis of the 10,000 pseudo-observations in Section 3.3.

Several interesting conclusions can be drawn from Table 3.2.[7] We find that, for example, rest-frame *ugriz* photometry can provide roughly 50-percent uncertainty on M_*/L_r measurements, i.e. almost as tight as the 25- to 35-percent uncer-

[7] a) For fixed signal-to-noise ratio S/N = 30 in all bands. b) For median signal-to-noise ratio per pixel $\overline{\rm S/N}$ = 20 across the wavelength range λ=3600–7400 Å. c) Averaged over all galaxy types. d) For greater-than-3σ detections of [O II]λ3727; Hβ; [O III]$\lambda\lambda$4959, 5007; [N II]λ6548 + Hα + [N II]λ6584; and [S II]$\lambda\lambda$6716, 6731. e) For greater-than-3σ detections of [O II]λ3727; Hβ; [O III]λ4959; [O III]λ5007; [N II]λ6548; Hα; [N II]λ6584; [S II]λ6716 and [S II]λ6731. f) For galaxies with net Hα+[N II] emission equivalent width greater than 5 Å. g) For galaxies with net Hα emission equivalent width greater than 5 Å. h) Constraint on the birth-cloud component of the attenuation only (i.e. $\hat{\tau}_V^{\rm BC}$ and not $\hat{\tau}_V$).

tainty provided by medium- and low-resolution optical spectroscopy (photometric constraints worsen somewhat if the galaxy redshift is not known; see Fig. 3.11). Meanwhile, optical spectroscopy is required to constrain $f_{\rm SFH}$ through the strong absorption-line signatures of intermediate-age stars. Table 3.2 also shows that observations of the net emission equivalent widths of prominent optical lines are sufficient to constrain $\psi_{\rm S}$ to within 50–70 percent (depending on the spectral resolution) and $12 + \log({\rm O/H})$ to within 0.05–0.23 in star-forming galaxies. Optical spectroscopy can help reduce these uncertainties and the potential small offsets in the retrieved best estimates of $\psi_{\rm S}$ and $12 + \log({\rm O/H})$. In addition, spectroscopy is required to constrain both parameters in quiescent star-forming galaxies with weak emission lines ($\psi_{\rm S} \lesssim 0.07\,{\rm Gyr}^{-1}$; see Fig. 3.12). Medium-resolution spectroscopy provides the best constraints on dust parameters in Table 3.2, with uncertainties of typically 0.35 in $\hat{\tau}_V$ and 0.10 in μ (these include uncertainties in the optical properties and spatial distribution of the dust as well as in the orientation of the galaxy; see Section 3.2.2.3). At lower spectral resolution, the less distinct signatures of emission lines make the uncertainty in $\hat{\tau}_V$ rise by about 80 percent, while the equivalent widths of optical emission lines at any spectral resolution constrain only the part of the attenuation arising from stellar birth clouds.

We stress that the constraints in Table 3.2 rely critically on the ability with our approach to interpret simultaneously the stellar and nebular emission from galaxies. In the case of low-resolution spectroscopy, no reliable constraint can be obtained on any of the physical parameters in Table 3.2 (except for the mass-to-light ratio of early-type galaxies) when nebular emission is not included in the analysis (Fig. 3.14). The same is true at medium spectral resolution, although if the resolution is high enough, prominent nebular emission lines can be excised from the spectra and analyzed separately from the stellar continuum and absorption-line signatures. We find that the likelihood distributions of $M_\star$, $12 + \log({\rm O/H})$, $\hat{\tau}_V$ and $\psi_{\rm S}$ retrieved from the combined analysis of stellar and nebular emission in the spectra of 12,660 SDSS star-forming galaxies (degraded to a resolution of 5 Å FWHM) using our approach agree well (modulo some recent model improvements) with those obtained in previous separate analyses of the stellar and nebular emission of these galaxies at twice higher spectral resolution [Brinchmann et al., 2004, Tremonti et al., 2004, Gallazzi et al., 2005, see Fig. 3.16]. Accounting for nebular emission is less crucial when deriving constraints on M_*/L_r from *ugriz* photometry (Fig. 3.8, bottom row).

Another important factor we can investigate in a straightforward way using our approach is the influence of signal-to-noise ratio on the retrieval of galaxy physical parameters from different types of observations. For multiband photometry, we find that lowering the signal-to-noise ratio from ${\rm S/N} = 30$ to 10 makes the uncertainty in the retrieved M_*/L_r in Table 3.2 larger by about 15 percent (Fig. 3.8). For spectroscopy, adopting a median signal-to-noise ratio per pixel $\overline{\rm S/N} = 5$ instead of 20 makes the median uncertainties in the different retrieved parameters increase by about 30 to 90 percent and introduces substantial biases in the retrieved likelihood distributions (Fig. 3.14). The ability to make such assessments is valuable as it can help decide between different strategies when designing an observational

campaign. For example, as noted in Section 3.3.2.2, low-resolution spectroscopy with $\overline{\mathrm{S/N}} = 5$ over the wavelength range from $\lambda = 3600$ to $7400\,\text{Å}$ provides hardly better constraints on the different parameters in Table 3.2 than high-quality *ugriz* photometry (Figs 3.6 and 3.14). In contrast, the constraints from low-resolution spectroscopy with $\overline{\mathrm{S/N}} = 20$ over the same wavelength range are generally far superior to those obtained from multiband photometry (Table 3.2).

It is important to keep in mind that the results reported in Table 3.2 about the retrievability of galaxy physical parameters from different types of observations correspond to the optimistic case in which the models we rely on are good approximations of true galaxies (Section 3.2.4). We have ensured this in the best possible way by appealing to state-of-the-art models including the most recent progress in galaxy spectral modeling to generate the library of spectral energy distributions on which our approach is based. With this in mind, the tool we have developed should be useful not only to interpret existing galaxy datasets, but also to design future observations. In principle, such observations can be of any photometric or spectroscopic type (or a combination thereof) across the wavelength range covered by spectral evolution models. We have focused in this Chapter on a few examples of photometric and spectroscopic observations at rest-frame optical wavelengths. In future work, we will present applications of our approach to the interpretation of the ultraviolet and infrared emission from galaxies at various redshifts. The tool we have developed is intended to be made available to the general astronomical community. In the meantime, we encourage colleagues interested in the application of our approach to the interpretation of specific galaxy observations to contact us.

Chapter 4

Constraining the physical properties of 3D-HST galaxies through the combination of photometric and spectroscopic data

4.1 Introduction

Studies on the evolution of the cosmic star formation density have revealed that the Universe has undergone a peak of star formation activity at redshifts between $z \sim 1$ and $z \sim 2$ (see Section 2.1.1). Ironically, this redshift range has long been referred to as the *redshift desert*. This is because in this range, the strong rest-frame optical lines allowing redshift determinations are shifted into the infrared and blocked by Earth's atmosphere, while the strong rest-frame ultraviolet features are not redshifted enough to be observable at optical wavelengths. Space observatories have opened a crucial window on this important cosmic epoch.

We present here the first steps of an analysis aimed to illustrate the power of the approach described in Chapter 3 to interpret combined photometric and spectroscopic observations of galaxies at redshift $z \sim 1 - 3$. For the galaxies we consider, photometric data are available in a wide range of wavelengths from the FIREWORKS catalogue (rest-frame ultraviolet to infrared; Wuyts et al. 2008), while spectroscopic information is available only in a narrow spectral window from grism spectroscopy of the 3D-HST Treasury Survey (rest-frame optical; van Dokkum et al. 2011). Both datasets are described in detail in the next Section. Our goal is to maximize the constraints on the physical parameters of these galaxies through the simultaneous interpretation of the emission from stars and dust, which dominates photometry, and nebular emission, which produces characteristic emission-line features.

For the preliminary analysis presented in this Chapter, we select a sample of 12 galaxies with 3D-HST and FIREWORKS data. A full analysis will require that we interpret all data (including the infrared photometry from FIREWORKS) using the models presented in Chapter 3. However, we have not yet combined these models with the dust *emission* models of da Cunha et al. [2008] to compute simultaneously

the ultraviolet, optical and infrared emission from stars, gas and dust. Moreover, the 3D-HST and FIREWORKS data have not been properly inter-calibrated yet. For this preliminary analysis, therefore, we will compare the constraints on physical parameters derived by interpreting, on the one hand, 3D-HST optical spectroscopy using the models of Chapter 3, and on the other hand, FIREWORKS ultraviolet-to-infrared photometry using the models of da Cunha et al. [2008] (which do not include nebular emission and are based on a slightly older version of the Bruzual and Charlot [2003] models than that adopted in this thesis). We also describe how we will incorporate dust emission in our existing models to interpret both datasets simultaneously.

We present the 3D-HST and FIREWORKS datasets in Section 4.2. In Section 4.3, we describe the spectroscopic and photometric approaches to fit the galaxy spectral energy distributions. In Section 4.4, we compare the estimates of stellar mass and specific star formation rate obtained using the two approaches. We discuss these results in Section 4.5. Finally, in Section 4.6, we outline the next steps of the project.

This project started in November 2011, when I visited Dr. Elisabete da Cunha and Professor Hans-Walter Rix at MPIA (Heidelberg). The results presented here are preliminary. I will be spending five months at MPIA over the summer of 2012 to complete the project.

4.2 The data

We select 12 star-forming galaxies at redshifts in the range $1 < z < 3$ from the 3D-HST database, a near-IR spectroscopic survey (248-orbit HST Treasury program) performed with the Wide Field Camera 3 (WFC3) on HST. The 12 spectra were obtained in the GOODS-South field using the WFC3 G141 grism, which provides spatially resolved spectra covering the observed wavelength range $0.9\,\mu$m–$1.9\,\mu$m with point-source spectral resolving power $R \approx 130$ and a fixed pixel size of 22Å. The median signal-to-noise ratio per pixel of the continuum is around 2. Since the spectra are spatially resolved, any particular spectrum may be contaminated by overlapping spectra of neighboring objects. This contamination is quantified in the reduction process and can be corrected for. To select the 12 spectra, we require:

- at least one emission line in the observed range;
- good signal-to-noise ratio of the emission line (S/N ~ 5);
- less than 10 percent contamination by neighboring objects;
- extended nebular emission, to avoid AGNs;
- a counterpart in the FIREWORKS photometric catalogue.

The FIREWORKS catalogue includes 4 HST/ACS filters (B_{435}, V_{606}, i_{775}, z_{850}), 5 ESO filters (U_{38}, B, V, R, I), 3 VLT/ISAAC filters (J, H, K_s), 4 Spitzer/IRAC

ID (FIREWORKS)	ID (3D-HST)	$z_{\rm grism}$
1586	308	1.784
1705	422	1.377
1876	501	1.377
1799	675	2.210
5689	55	1.052
4013	61	1.314
5548	187	1.668
5434	271	1.215
5349	320	1.991
5341	355	1.218
5209	451	1.018
5134	545	1.008

Table 4.1: ID and spectroscopic redshifts of the galaxies in the sample described in Section 4.2.

filters (3.6, 4.5, 5.8, 8.0 μm) and the Spitzer/MIPS-24 μm filter. In a few cases, Herschel/PACS photometry (60-85 μm, 85-125 μm and 125-210 μm) is available too. For the 12 galaxies in our sample, spectroscopic redshifts, $z_{\rm grism}$, estimated using emission lines are available from the 3D-HST catalogue. We report these in Table 4.1.

In Figure 4.1, we plot the G141 spectrum of one of the galaxies in the sample. The spectrum extends from 0.9 to 1.9 μm (red dotted line), but noise and contaminations are strong below 1.1μm and above 1.6μm. We thus consider only the spectrum inside this range (blue line) for the analysis (for every galaxy in the sample). One of the main current issues with 3D-HST data is spectral calibration, which does not always match with the calibration of FIREWORKS photometry. The consortium is working actively on resolving this issue. To ensure at least a consistent overall absolute calibration between 3D-HST and FIREWORKS data for the galaxies in our sample, we decide to rescale the 3D-HST spectra based on the J-band photometric flux from FIREWORKS. The VLT/ISAAC J-band filter is the only one falling in the range sampled by the WFC3/G141 grism. We therefore compute the median flux of the 3D-HST spectrum in a square band of 0.1 μm width centered on the effective wavelength of the J-band filter (1.253 μm, magenta line). We then rescale uniformly the spectrum by multiplying it by the ratio between the FIREWORKS J-band flux and the estimated median spectroscopic J-band flux. We list the scaling factors for the 12 galaxies in our sample in Table 4.2.

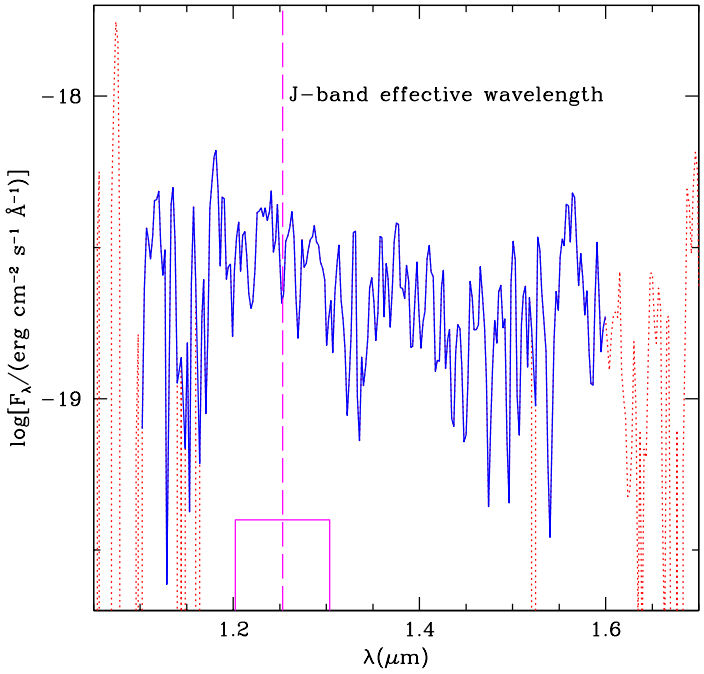

Figure 4.1: Galaxy #1705. The G141 spectrum (red dotted line), the part of the spectrum used in the analysis ($1.1 \leq \lambda \leq 1.6\mu$m, blue solid line), J-band effective wavelength (magenta dashed line) and square band of 0.1 μm width (magenta solid line).

ID (FIREWORKS)	J-band scaling factor
1586	2.212
1705	0.934
1876	0.914
1799	0.883
5689	1.329
4013	0.781
5548	1.407
5434	0.867
5349	0.975
5341	1.012
5209	1.243
5134	2.735

Table 4.2: Scaling factors of the galaxies in the sample defined as the ratio between the median flux in a square band of 0.1μm width, centered on the J-band effective wavelength, and the FIREWORKS mean flux in the J-band.

4.3 Fits of the spectral energy distribution

In this Section, we describe the separate analyses of the 3D-HST spectra and the FIREWORKS photometry to constrain key physical parameters of the 12 galaxies in our sample.

4.3.1 Photometric approach

We interpret the FIREWORKS photometric data using MAGPHYS, a model package to interpret observed spectral energy distributions of galaxies in terms of galaxy-wide physical parameters. The approach is described in da Cunha et al. [2008]. The analysis of the spectral energy distribution of an observed galaxy is done in two steps:

1. the assembly of a comprehensive library of model spectral energy distributions at the same redshift and in the same photometric bands as the observed galaxy;

2. the build-up of the marginalized likelihood distribution of each physical parameter of the observed galaxy, through the comparison of the observed spectral energy distribution with all the models in the library.

The library of spectra is assembled by combining a library of *optical* spectra (describing the ultraviolet, optical and near-infrared emission from stellar populations) and a library of *infrared* spectra (describing the mid- and far-infrared emission from dust). The spectra in the optical library are computed using the stellar population synthesis code of Bruzual and Charlot [2003]. They include the effects of dust attenuation as prescribed by Charlot and Fall [2000]. More specifically, each star formation history in the library is parameterized in terms of an underlying continuous model with exponentially declining star formation rate, on top of which are superimposed random bursts. The models are distributed uniformly in metallicity between 0.2 and 2 times solar. The attenuation by dust is randomly sampled by drawing the total effective V-band absorption optical depth ($\hat{\tau}_V$) between 0 and 6, and the fraction of this contributed by dust in the ambient ISM (μ) between 0 and 1. In the infrared library, the spectra are computed using the prescription of da Cunha et al. [2008], which includes wide ranges of dust temperatures and fractional contributions by different dust components to the total infrared luminosity (see the original paper of da Cunha et al. 2008 for more details on the prior distributions of these parameters). The optical and infrared libraries are linked together by requiring consistency between the total energy absorbed by dust in stellar birth clouds and in the interstellar medium, and the re-distribution of this energy at infrared wavelengths.

The fitting procedure is very similar to that described in Chapter 3. It consists in comparing the observed FIREWORKS fluxes with the fluxes of all models in the combined optical+infrared library to build likelihood distributions of 7 key physical parameters:

- the fraction of total dust luminosity contributed by the diffuse ISM, f_μ;
- the total optical depth seen by young stars in birth clouds, τ_V;
- the optical depth seen by stars in the diffuse ISM, $\mu\tau_V$;
- the star formation rate, ψ;
- the stellar mass, M_*;
- the specific star formation rate, $\psi_S = \psi/M_*$;
- the total dust luminosity, $L_\mathrm{d}^\mathrm{tot}$.

Figure 4.2 shows, as an example, the fit of the spectral energy distribution of the galaxy #1799 using this approach. The top panel shows the FIREWORKS photometric data (red dots) and best-fit model spectral energy distribution (black line), together with the corresponding unattenuated model (cyan line). The middle panel shows a zoom of the (observer frame) optical and near-infrared wavelength range. Also plotted for reference is the 3D-HST spectrum (green line) with error bars (in grey). The 7 panels at the bottom show the probability density functions of the retrieved physical parameters for this galaxy.

4.3.2 Spectroscopic approach

The approach we adopt to interpret the 3D-HST grism spectra is very similar to that presented in Chapter 3. In brief, we perform a semi-analytic post-treatment of the Millennium simulation to build a library of star formation and chemical enrichment histories. We sample uniformly the redshift of observation in the range between 0.7 and 3.4 appropriate for 3D-HST galaxies. By analogy with our approach in Chapter 3, we resample the *current* star formation rate and interstellar metallicity using broad distributions of these parameters (as before, we define as *current* the average of a quantity over a period of 10 Myr before a galaxy is looked at). This resampling removes unwanted correlations between specific star formation rate, gas-phase oxygen abundance and total effective optical depth of the dust in the original library. For the present analysis, we consider gas-phase oxygen abundance down to $12 + \log(\mathrm{O/H}) = 7$, i.e. lower values than considered in Chapter 3, to account for the potentially low metallicities of high-redshift galaxies (see e.g. Pettini et al. 1997). We use the same prescriptions as in Chapter 3 to describe the stellar and nebular emission from galaxies. We also use the same prescription for dust attenuation, but we remove the correlation between effective optical depth and slope of the attenuation curve for dust in the ambient interstellar medium, which was calibrated using low-redshift data. Instead, we account for orientation effects on dust attenuation by varying randomly and uniformly the slope of the attenuation curve in the ambient interstellar medium in the range $0.4 < n < 1.0$ (keeping a fixed slope of 1.3 for the attenuation curve in the birth clouds). We resample the total effective optical depth of the dust in the range $0.01 < \hat{\tau}_V < 3.0$, with a probability density

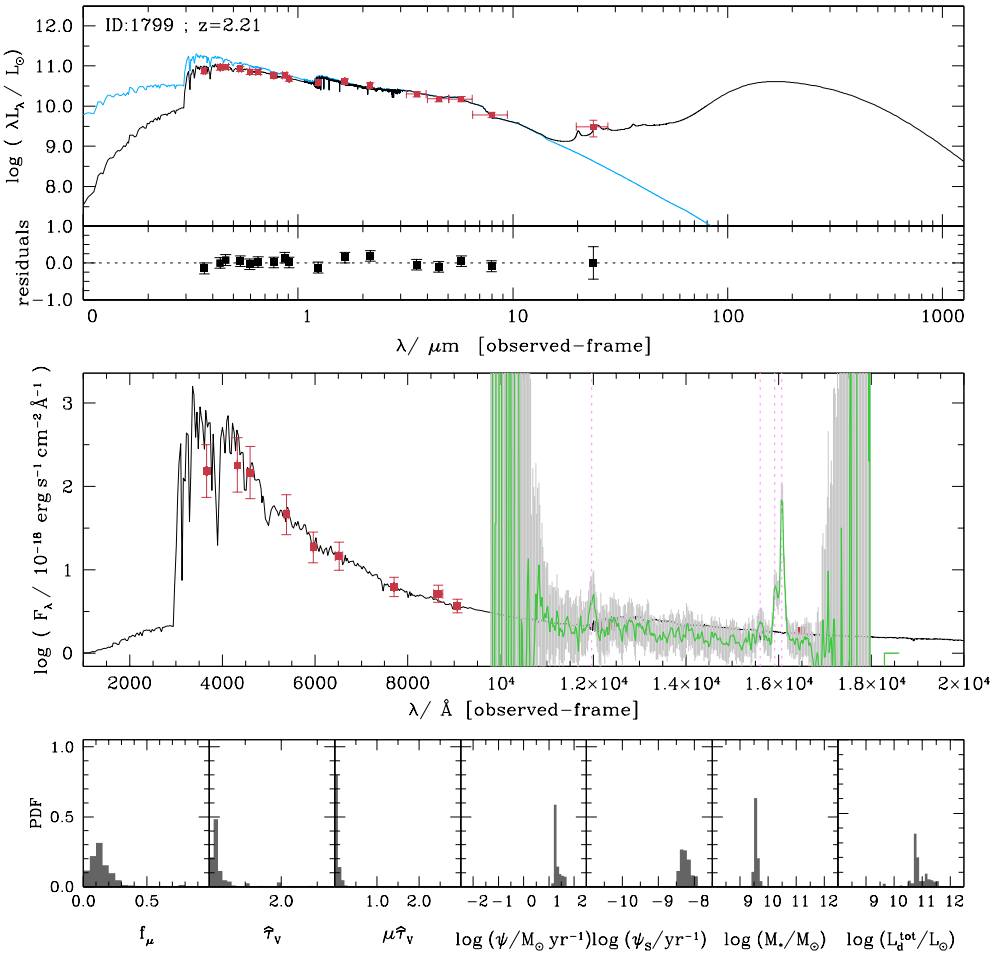

Figure 4.2: Top panel: MAGPHYS fit ($z_{\rm grism} = 2.21$ is the input redshift) to the FIREWORKS photometry (red dots) of galaxy #1799; the black spectrum is the best-fit model and the blue spectrum is the pure stellar spectrum (i.e. unattenuated by dust) corresponding to this best-fit model. Middle panel: zoom of the observed-frame wavelength range from the ultraviolet to the near-infrared; FIREWORKS photometry (red dots), MAGPHYS best-fit model spectrum (black line), 3D-HST spectrum (green line) with error bars (grey). Bottom panels: histograms showing the likelihood distributions of the fraction of total dust luminosity contributed by the diffuse ISM, f_μ, the total optical depth seen by young stars in birth clouds, τ_V, the optical depth seen by stars in the diffuse ISM, $\mu\tau_V$, the star formation rate, ψ, the specific star formation rate, ψ_S, the stellar mass, M_*, and total dust luminosity, $L_{\rm d}^{\rm tot}$.

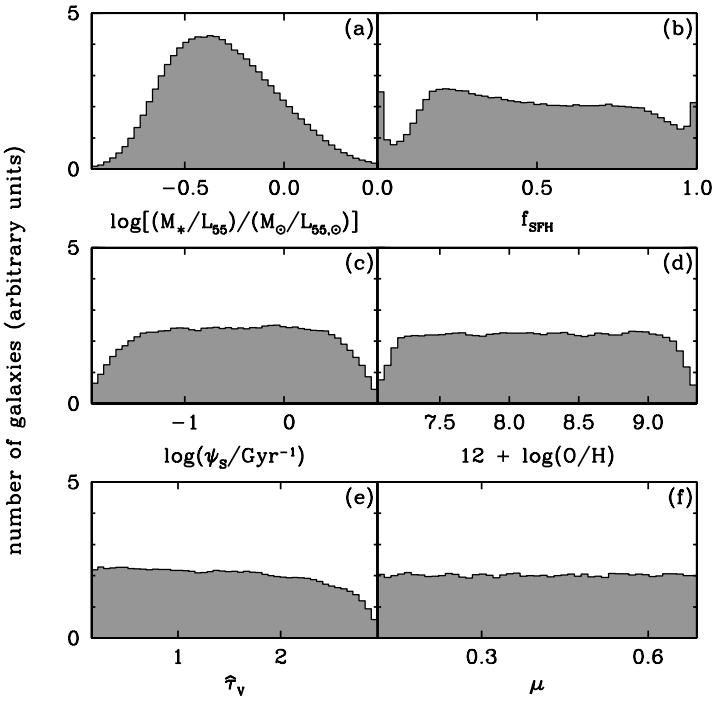

Figure 4.3: Prior distributions of selected physical parameters of the 2,500,000 galaxies in the spectral library: (a) rest-frame absolute V-band stellar mass-to-light ratio, M_*/L_{55}; (b) fraction of the current galaxy stellar mass formed during the last Gyr, $f_{\rm SFH}$; (c) specific star formation rate, $\psi_{\rm S}$; (d) gas-phase oxygen abundance, $12 + \log({\rm O/H})$; (e) total effective V-band absorption optical depth of the dust, $\hat{\tau}_V$; (f) fraction of $\hat{\tau}_V$ arising from dust in the ambient ISM, μ. In each panel, the shaded histogram shows the distribution for all galaxies, while the solid histogram shows the contribution by star-forming galaxies alone. Non-star-forming galaxies are off scale (at $\log(\psi_{\rm S}) = \infty$) in panel (c) and do not contribute to the distributions of interstellar parameters in panels (d)–(f).

function of the form $0.18 \arctan[5(3 - \hat{\tau}_V)]$. This does not extend to values as high as $\hat{\tau}_V = 4.0$, as in the library in Chapter 3, because the galaxies in our sample are all optically bright. We include attenuation by the intergalactic medium following the prescription of Madau [1995, see also Section 2.5 of this thesis].

We consider a library of 2,500,000 different model spectral energy distributions.[1] Only about 1 percent of these models correspond to galaxies currently not-forming stars. Figure 4.3 shows the prior distributions of six physical parameters in the model library. Panel (a) shows the distribution of the rest-frame absolute V-band (5500Å) stellar mass-to-light ratio. This spans a range of 1.5 dex, with the majority of galaxies lying around -0.5. Panel (b) shows the fraction of the current galaxy stellar mass formed during the last Gyr. The (resampled) distributions of specific star formation rate, gas-phase oxygen abundance, total optical depth of the dust and fraction of this contributed by the dust in the interstellar medium are shown in panels (c), (d), (e) and (f).

As in Chapter 3, we use a Bayesian approach to constrain the likelihood distributions of redshift and of the physical parameters shown in Figure 4.3 for each of the 12 galaxies in our sample. This consists in comparing each observed spectrum with all the spectra in the library. The spectra are compared pixel-per-pixel in the observer frame to exploit the entire information from the stellar continuum and nebular lines simultaneously. Since the spectra are spatially resolved, emission lines appear broader than would be expected for a point source at a resolving power of 130. In practice, the median FWHM of emission lines in our sample is roughly 200 Å (observed) at all redshifts. To fit the observed 3D-HST spectra with our models, therefore, we convolve all model spectral energy distributions in the observer frame with a gaussian smoothing function of FWHM= 200 Å. Then, we resample the spectral energy distributions using 22 Å-wide bins corresponding exactly to the observed wavelengths.

Figure 4.4 shows the result of this resampling for the same galaxy #1799 as shown in Figure 4.2. The original spectrum is plotted in red, while the best-fit model and the respective parameters are plotted in blue. Even without a redshift input, the code is able to recognize three emission lines ([O II], Hβ and [O III]) and recover an estimate of redshift in agreement with $z_{\rm grism} = 2.21$ (Table 4.1). The panels at the bottom show the retrieved probability distributions of M_*, $f_{\rm SFH}$, $\psi_{\rm S}$, $12 + \log({\rm O/H})$, $\hat{\tau}_V$ and z.

4.4 Results from photometric versus spectroscopic fits

We compare here the estimates of key physical parameters of the 12 galaxies in our sample obtained through independent fits of FIREWORKS photometry and 3D-HST spectroscopy, as described above. Our ultimate goal will be to combine the

[1] This is half as large as the library described in Chapter 3. In the case of low resolution spectroscopy, a few million models are enough to reach convergence in the uncertainty and accuracy of the retrieved parameters. Using 2,500,000 models instead of 5,000,000 speeds up the fitting procedure by a factor of 2, without biasing the results.

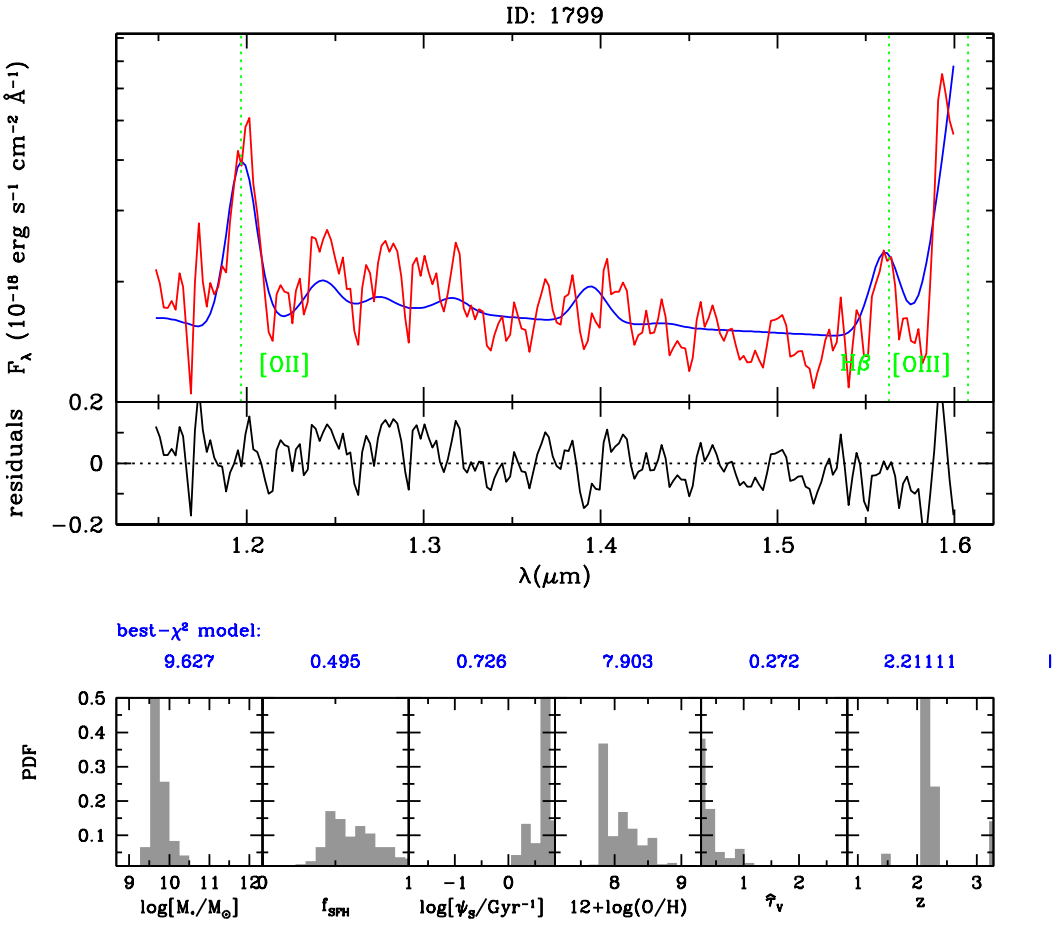

Figure 4.4: Spectral fit of galaxy #1799, using this thesis approach, with no prior information on redshift. Top panel: the red line is the observed 3D-HST spectrum, and the blue line is the best-fit model spectrum. Medium panel: residuals between the observed and the best-fit model spetra. Bottom panels: histograms showing the likelihood distributions of stellar mass, fraction of stellar mass formed in the last Gyr, specific star formation rate, gas-phase oxygen abundance, total effective optical depth of the dust and redshift.

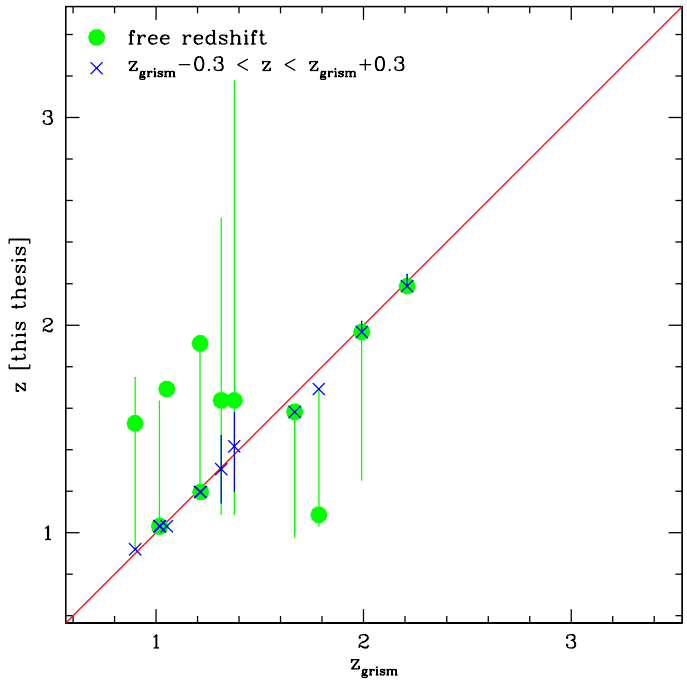

Figure 4.5: Comparison between the redshift estimates obtained with our spectroscopic approach, with the redshift provided in the catalogue ($z_{\rm grism}$). Green dots represent the median estimates when redshift is treated as a free parameter. Blue crosses represent the median estimates obtained fitting each galaxy only with models in the range $z_{\rm grism} \pm 0.3$. Error bars are the 16 and 84 percentiles in the likelihood distributions.

two datasets to improve the constrains on physical parameters. In the meantime, it is important to understand the potential differences and biases between the two approaches. MAGPHYS requires an input redshift, unlike the approach we have developed in Chapter 3. Thus, in what follows, we first compare the redshifts retrieved from our spectroscopic fits with the value of $z_{\rm grism}$ provided in the 3D-HST spectroscopic catalogue. Then, we compare the estimates of mass and specific star formation rate derived independently from the photometric and spectroscopic fits of the 12 galaxies in our sample.

4.4.1 Redshift

We plot in Figure 4.5 the redshift derived from our spectroscopic fit against the redshift $z_{\rm grism}$ provided in the 3D-HST catalogue (green dots). In some cases, the spectroscopic fit favors two possible redshifts. The likelihood distribution exhibits two distinct spikes, and the median of the distribution usually corresponds to one of the two spikes, with a large uncertainty towards the second. This happens when a

single emission line is detectable across the observed wavelength range, which can be interpreted either as an oxygen ([O II] or [O III]) or a hydrogen (Hα) transition. To avoid this problem, we can restrict the parameter range by including only models with redshifts within 0.3 of $z_{\rm grism}$ in the library. The results are shown by blue crosses in Figure 4.5.

Figure 4.6 and 4.7 show two spectroscopic fits of the same galaxy (#5434) obtained when no constraint is applied on the redshift and when the redshift is constrained to be within 0.3 of $z_{\rm grism}$, respectively. In each figure, the rightmost panel at the bottom shows the likelihood distribution of redshift. When the redshift is unconstrained (Figure 4.6), the single observable emission line can be interpreted as either an [O III] or a Hα transition. These different interpretations cause the redshift likelihood distribution to be divided into two spikes, at $z \sim 1.2$ and $z \sim 1.9$. This uncertainty affects also the constraints on the other parameters. When the redshift is constrained to be within 0.3 of $z_{\rm grism}$ (Figure 4.7), the uncertainties in M_* and $\psi_{\rm S}$ improve by about 0.1 dex, while the uncertainty in $f_{\rm SFH}$ improves by 50 percent.

4.4.2 Mass and specific star formation rate

We now want to compare our estimates of mass and specific star formation rate with the ones obtained from FIREWORKS photometry using MAGPHYS. Since MAGPHYS requires an input redshift, the comparison of MAGPHYS estimates with our estimates will be more meaningful if we constrain the redshift to be within 0.3 of $z_{\rm grism}$ in the spectroscopic fits.

We are primarily interested in the comparison of the constraints on stellar mass and specific star formation rate, because these two parameters should be the most easily retrievable by both approaches. Figure 4.8 shows a comparison of M_* and $\psi_{\rm S}$ retrieved from the spectroscopic and photometric fits. As expected, the uncertainty in the spectroscopic estimate for both parameters is smaller when including information about the redshift (blue crosses) than when not including it (green dots). Figure 4.8 shows that, intringuingly, the stellar mass retrieved from 3D-HST spectroscopy is systematically larger, by a factor of 0.5 dex, than that retrieved from FIREWORKS photometry, with a scatter of about 0.2 dex. The specific star formation rate is systematically lower by roughly the same amount, with larger scatter and errors than for stellar mass. We note that the error bars are much larger for spectroscopic than for photometric estimates, because the observed wavelength range used in the constraints is much narrower. We explore in the next Section the possible causes for the important discrepancies revealed by Figure 4.8.

4.5 Possible causes of discrepancy between photometric and spectroscopic estimates

There are at least two possible origins of the differences in the estimates of stellar mass and specific star formation rate obtained from the photometric and spectro-

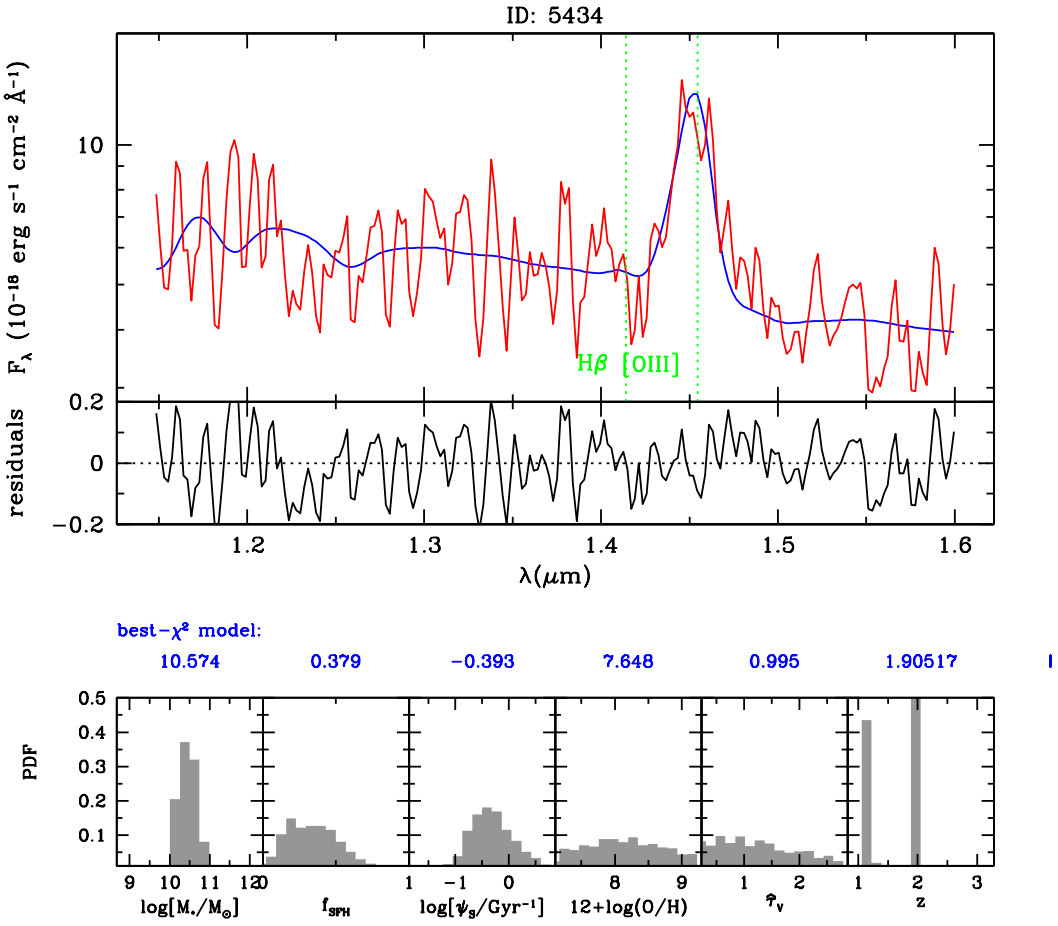

Figure 4.6: Same as Fig. 4.4. Spectral fit of galaxy #5435, using the approach developed in this thesis. No constraint is applied on the redshift.

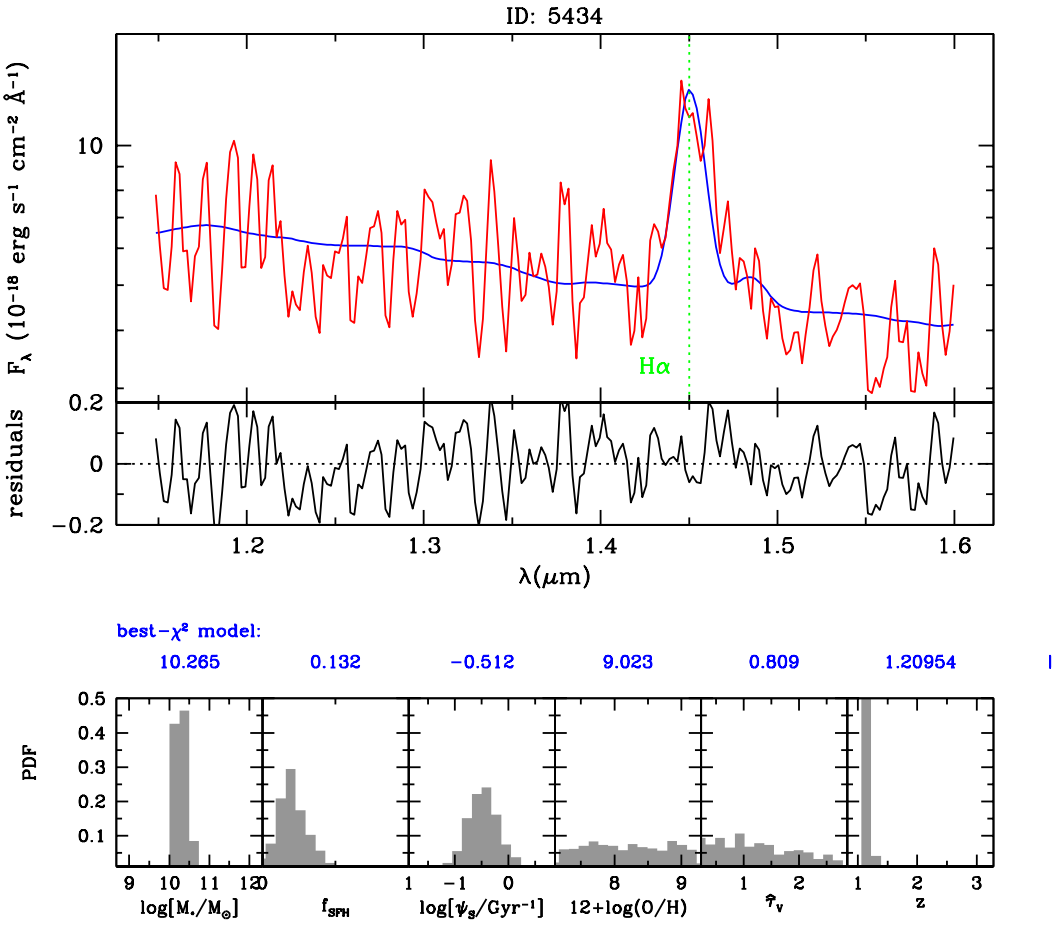

Figure 4.7: Same as Fig. 4.4. Spectral fit of galaxy #5434, using the approach developed in this thesis. The redshift is constrained to be within 0.3 of $z_{\rm grism}$.

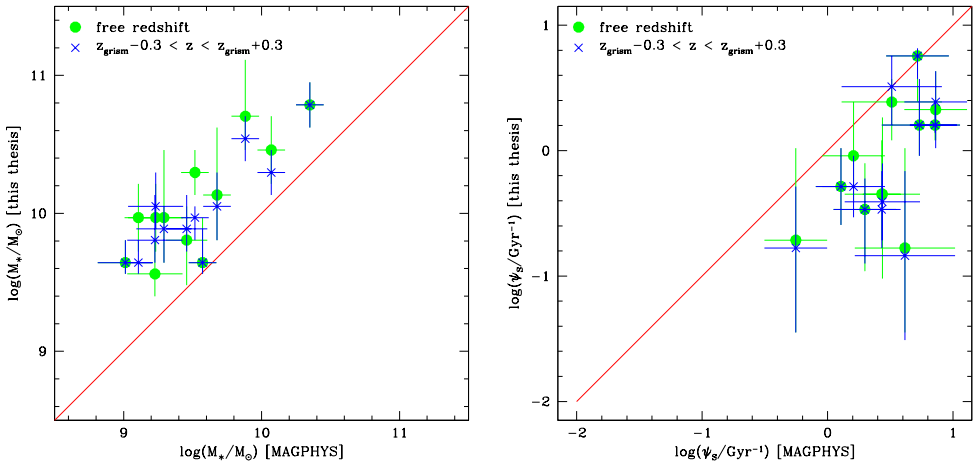

Figure 4.8: Left panel: comparison between stellar mass estimates obtained with our spectroscopic approach and the ones obtained with MAGPHYS. Right panel: comparison between specific star formation rate estimates obtained with our spectroscopic approach and the ones obtained with MAGPHYS. In both panels, green dots represent the median estimates when redshift is treated as a free parameter. Blue squares represent the median estimates obtained when the redshift is constrained to be within 0.3 of $z_{\rm grism}$. Error bars represents the 16th and 84th percentiles of the likelihood distributions.

	MAGPHYS	This thesis
Star formation history	Exponentially declining with random bursts	SAM on Millennium Simulation
Metal enrichment	Fixed for all stars in a given galaxy	Enrichment history from simulation
Stellar population synthesis model	Bruzual and Charlot [2003], version of 2007	Bruzual and Charlot [2003], the most recent version
Nebular emission	Not included	Included
Dust attenuation	Fixed slope in ISM	Range of slopes in ISM
Dust emission	Included	Not included

Table 4.3: Differences in the prescriptions adopted in MAGPHYS and in the spectroscopic approach described in this thesis.

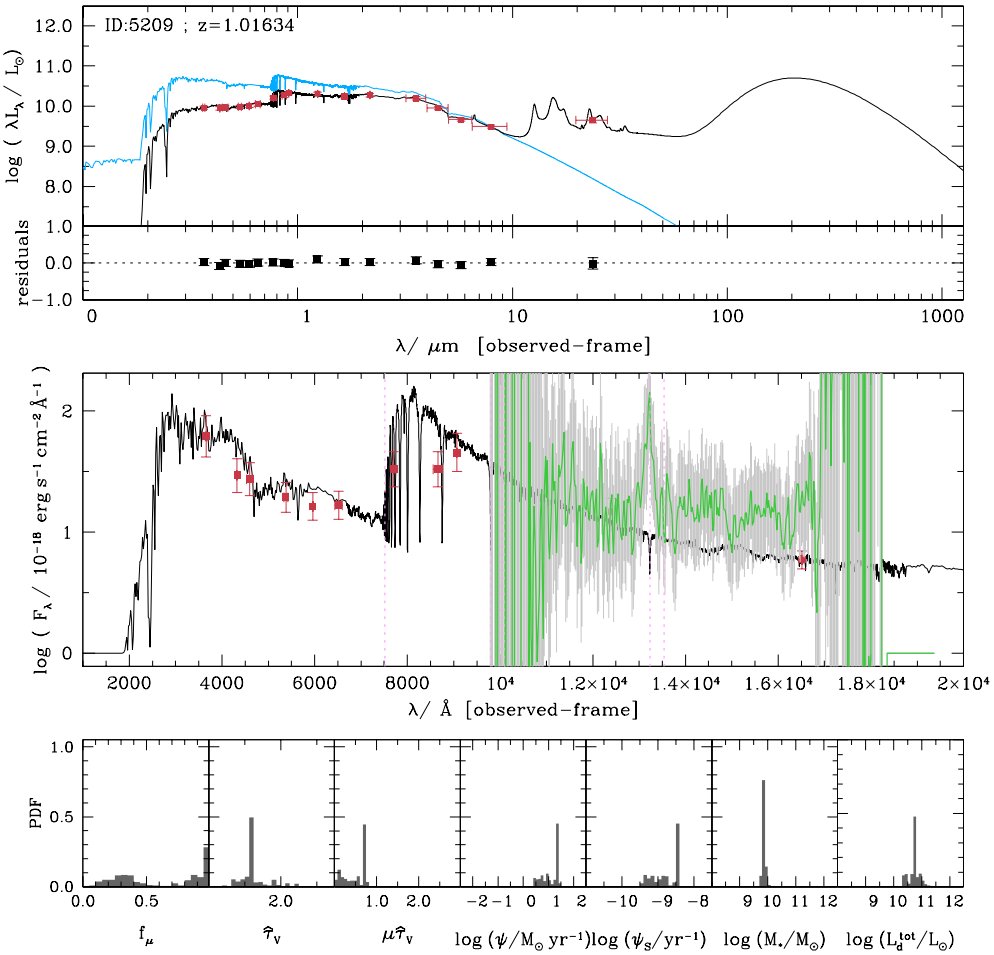

Figure 4.9: Top panel: MAGPHYS fit ($z_{\rm grism} = 1.02$ is the input redshift) to the FIREWORKS photometry (red dots) of galaxy #5209; the black spectrum is the best-fit model and the blue spectrum is the pure stellar spectrum (i.e. unattenuated by dust) corresponding to this best-fit model. Middle panel: zoom of the observed-frame wavelength range from the ultraviolet to the near-infrared; FIREWORKS photometry (red dots), MAGPHYS best-fit model spectrum (black line), 3D-HST spectrum (green line) with error bars (grey). Bottom panels: histograms showing the likelihood distributions of the fraction of total dust luminosity contributed by the diffuse ISM, f_μ, the total optical depth seen by young stars in birth clouds, τ_V, the optical depth seen by stars in the diffuse ISM, $\mu\tau_V$, the star formation rate, ψ, the specific star formation rate, ψ_S, the stellar mass, M_*, and total dust luminosity, $L_{\rm d}^{\rm tot}$.

scopic fits described above.

1. Differences in flux calibration between FIREWORKS photometry and 3D-HST spectroscopy. This is suggested by the fact that the slope of the continuum in the 3D-HST spectra of some galaxies differs significantly from that inferred from J- and H-band FIREWORKS photometry. We show in Figure 4.9 (middle panel, green spectrum compared to black best-fit model) the case of largest discrepancy, where the spectroscopic estimate of mass is 0.7 dex larger than the MAGPHYS estimate. Despite the imposed agreement in J-band flux (see above), the two slopes are largely different. Updates of the flux calibration should be available next summer, when this project will be concluded.

2. Differences in the model prescriptions used in the two approaches. Table 4.3 summarizes the differences in the models used in MAGPHYS and in the approach developed in this thesis. At first sight, none of these differences appears sufficient to account for large offsets in the estimated stellar mass and specific star formation rate. A thorough analysis of the influence of each of the differences in Table 4.3 on the results will be performed next summer. A first test will consists in fitting FIREWORKS photometry using our library of spectral energy distributions (i.e. without including emission by dust in the mid and far infrared, but still including the ultraviolet, optical and near-infrared photometric constraints).

4.6 Summary and next steps

In this Chapter, we have started to investigate the differences in the constraints derived on the physical properties of galaxies at redshifts in the range $1 < z < 3$ from the separate analyses of FIREWORKS photometry (using the MAPGPHYS model package) and 3D-HST spectroscopy (using the models developed in Chapter 3). In Section 4.2, we have presented the characteristics of these two datasets:

- 3D-HST near-infrared grism spectra sample the optical rest frame of the galaxies;
- FIREWORKS photometry includes 17 bands ranging from ultraviolet to far-infrared wavelengths.

In Section 4.3, we have described the separate photometric and spectroscopic fits of the galaxies:

- using the MAGPHYS model package to analyze FIREWORKS photometry, including emission by dust in the infrared;
- using the approach developed in Chapter 3, which is best suited to extract information from the combined stellar and nebular emission in low-resolution ultraviolet and optical spectra.

In Section 4.4, we have compared the estimates of stellar mass and specific star formation rate inferred separately from the two approaches. We found that the stellar masses inferred from the spectroscopic fits are significantly lower (and specific star formation rates significantly larger) than those derived from the photometric fits. In Section 4.5, we have shown that possible causes for these discrepancies could be:

- differences in the flux calibration between the FIREWORKS and 3D-HST data;
- differences in the model prescriptions used to fit the data.

This project is intended to be concluded over the summer 2012. The first step will be to understand the differences between the photometric and spectroscopic estimates of stellar mass and specific star formation rate shown in Figure 4.8. Once this is settled, we will estimate the stellar masses, star formation rates and dust attenuation optical depths of the largest possible sample of FIREWORKS+3D-HST galaxies. Meanwhile, we will combine the models developed in Chapter 3 with the dust *emission* models of da Cunha et al. [2008] to compute simultaneously the ultraviolet, optical and infrared emission from stars, gas and dust. This will allow us to extract the most reliable measurements of physical parameters from the combination of FIREWORKS and 3D-HST data.

Chapter 5

ACS and NICMOS photometry in the Hubble Ultra Deep Field

5.1 Introduction

This thesis work is part of the European project ELIXIR (EarLy unIverse eXploration with nIRspec). ELIXIR is a Marie Curie Initial Training Network funded by the Seventh Framework Programme (FP7) of the European Commission to prepare for the exploration of the early universe with the NIRSpec near-infrared spectrograph on board the James Webb Space Telescope (JWST). In this framework, we have started a collaboration with Bernhard Dorner and Dr. Pierre Ferruit, who are responsible for developing the NIRSpec instrument performance simulator.[1]

The aim of this study is to use models of the type described in Chapter 3 to simulate observations of high-redshift galaxies as will be achievable with JWST. Using such *pseudo-observations*, we wish to compare the constraints on physical parameters that can be derived from low- and medium-resolution spectroscopy to those that can be inferred from currently available photometric observations of high-redshift galaxies. To generate pseudo-galaxies with the most realistic colors, the idea is to select model spectral energy distributions that best reproduce existing photometric observations of the Hubble Ultra Deep Field (HUDF).

In this Chapter, we present preliminary results from this analysis. In Section 5.2, we describe the HUDF photometric sample of distant galaxies that will be used to create *pseudo-observed scenes* for JWST/NIRSpec.[2] In Section 5.3, we generate a library of model spectral energy distributions tailored to studies of high-redshift galaxies. In Section 5.4, we use this library to assess the extent to which the physical parameters of HUDF galaxies can be retrieved from the analysis of available photometry. We also show a first example of pseudo-observed scene generated using model spectral energy distributions and a preliminary version of the JWST/NIRSpec instrument performance simulator. We discuss the results obtained from the analysis of HUDF photometry in Section 5.5 and summarize our conclusions in Section 5.6.

This project is intended to be concluded and submitted for publication as soon as the NIRSpec instrument performance simulator is ready. Our study will quantify the

[1] The NIRSpec instrument performance simulator is a software for studies of the instrument performance, optical and geometrical effects, as well as the creation of realistic calibration and science exposures to develop and test data analysis tools

[2] In the instrument performance simulator, the spatial distribution of pseudo-galaxy spectra on the detector is referred to as a *pseudo-observed scene*.

value of low- and medium-resolution spectroscopy with JWST/NIRSpec to explore the physical parameters of the very first galaxies that formed in the early Universe.

5.2 The data

We select galaxies from the HUDF catalogue of Coe et al. [2006], which contains *B V i' z' J H* photometry and photometric redshifts for 8042 galaxies detected at the 10-σ level. *B V i'* and *z'* data were obtained with the Advanced Camera for Surveys (ACS) on board HST. *J* and *H* data were obtained with the Near Infrared Camera and Multi-Object Spectrometer (NICMOS) on board HST. We show the 6 filter response functions in Figure 5.1: *B* in blue, *V* in magenta, *i'* in cyan, *z'* in green, *J* in orange and *H* in red. Also plotted in grey are examples of spectral energy distributions extracted from the model library of Section 5.3. Flux is expressed in AB magnitudes at observed wavelengths between 0.3 to 2 μm. The five model galaxies are computed at five different redshifts to show how the position of the Lyman break[3] changes in observations between $z \sim 2$ and $z \sim 7$.

We select 29 galaxies with both ACS and NICMOS data and another 26 with ACS data alone. The selection derives from the following constraints:

- requirement that the listed photometric redshift be between 2 and 8;
- either *z'* or *J* magnitude (when available) brighter than 27 in AB system;
- observational requirements imposed by the NIRSpec microshutter array.

We list all 55 objects along with photometric redshifts, magnitudes and errors in Appendix C. The colors are plotted in Figure 5.2. In both panels, dots represent model galaxies in different ranges of redshift: $z < 3.5$ black, $3.5 < z < 4.5$ red, $4.5 < z < 5.5$ yellow, $5.5 < z < 6.5$ green, $6.5 < z < 7.5$ magenta. Grey squares with error bars represent the selected HUDF galaxies. Panel (a) shows $V - i'$ versus $i' - z'$ colors. This plot is commonly used to select *V*-dropout galaxies at $z \gtrsim 5$: objects with redshift greater than $z = 5.5$ do not have reliable photometry in the *V* band, which falls blueward of the Lyman break. This is the likely reason why three galaxies with uncertain photometry (at redshift $z = 5.53$, 5.715 and 6.336) fall outside the model range in Figure 5.2a. Panel (b) shows $i' - z'$ versus $z' - J$ colors, used to identify i'-dropout galaxies at $z \gtrsim 6$. In this case, the three high-redshift objects overlap reasonably well with the range of colors spanned by the models.

5.3 Modeling

In this Section, we generate a library of model spectral energy distributions specifically tailored to studies of high-redshift galaxies. We appeal to prescriptions of

[3]The Lyman break denotes the drop in flux of distant objects blueward of the Lyman limit (912 Å rest frame) caused by absorption by neutral hydrogen in the intergalactic medium (the break can also occur blueward of the Lyman-α line, at 1215 Å rest-frame, for the most distant objects; see Section 2.5).

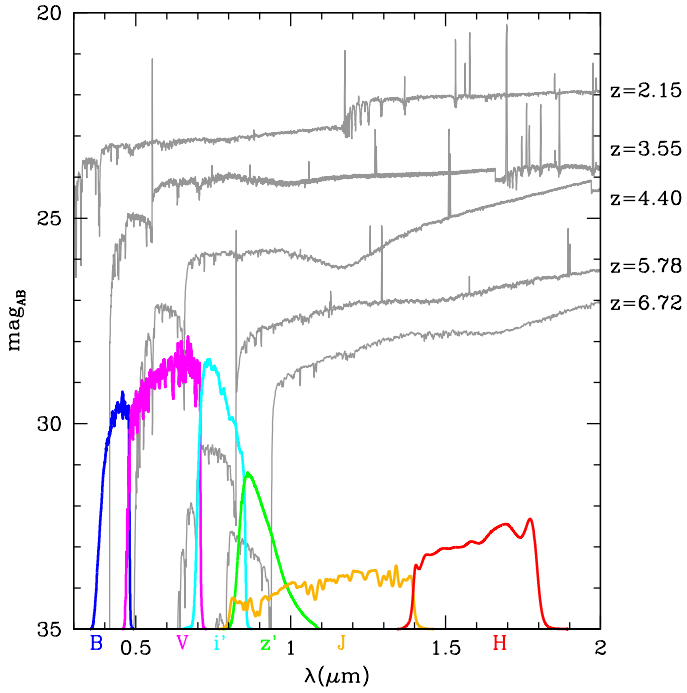

Figure 5.1: ACS and NICMOS filters: B in blue, V in magenta, i' in cyan, z' in green, J in orange and H in red. In gray, five model galaxies at redshifts 2.15, 3.55, 4.40, 5.78 and 6.72. Flux is expressed in AB magnitudes in the observed-range between 0.3 and 2 μm. The increasing effect of the IGM absorption with redshift is visible blueward of the Lyman-α line.

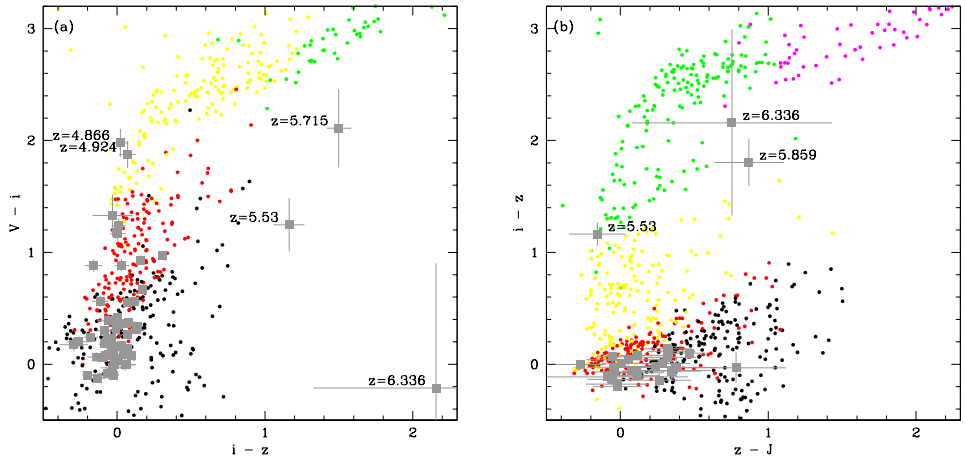

Figure 5.2: Panel *a*: $V - i'$ versus $i' - z'$ colors. Panel *b*: $i' - z'$ versus $z' - J$ colors. In both plots, dots represent model galaxies in different ranges of redshift: $z < 3.5$ black, $3.5 < z < 4.5$ red, $4.5 < z < 5.5$ yellow, $5.5 < z < 6.5$ green, $6.5 < z < 7.5$ magenta. Grey squares with error bars represent the selected HUDF galaxies. Photometric redshifts from the catalogue by Coe et al. [2006] label a few high-redshift galaxies to identify the outliers in panel (a) and check their position in panel (b).

star formation and chemical enrichment histories, stellar and nebular emission and attenuation by dust similar to those described in Chapter 3, but with some key specificities for studies of the (esp. ultraviolet) emission from high-redshift galaxies.

5.3.1 Library of spectral energy distributions

We perform a semi-analytic post-treatment of the Millennium simulation analogous to that described in Chapter 3 to build a library of star formation and chemical enrichment histories of high-redshift galaxies. In this case, we stop the simulation at $z = 2$ instead of $z = 0$. To generate a model library that matches the properties of the observed HUDF sample, we randomly sample the redshift of observation between $z = 2$ and $z = 8$. In practice, the fraction of galaxies decreases slightly at the highest redshifts in the library (Figure 5.3, panel b), because some models do not start forming stars before redshifts smaller than $z = 8$. By analogy with our approach in Chapter 3, we resample the *current* star formation rate and interstellar metallicity using broad distributions of these parameters (as before, we define as *current* the average of a quantity over a period of 10 Myr before a galaxy is looked at). This resampling removes unwanted correlations between specific star formation rate, gas-phase oxygen abundance and total effective optical depth of the dust in the original library. In particular, we extend the range of specific star formation rates up to the limit where all stars were formed in the last 10 Myr ($\log[\psi_S/\mathrm{Gyr}^{-1}] = 2$;

see Figure 5.3, panel c) and the range of gas-phase oxygen abundances down to $12 + \log(\mathrm{O/H}) = 7$ (see Figure 5.3, panel d). These choices are motived by the results of recent observations of high-redshift galaxies characterized by very high specific star formation rates and very metal-poor gas [e.g. Erb et al., 2010, Bouwens et al., 2010].

To compute model spectral energy distributions, we include stellar and nebular emission using the same prescriptions as in Chapter 3. We include attenuation by the intergalactic medium following the prescription of Madau [1995, see also Section 2.5 of this thesis]. Attenuation by dust is computed as in Chapter 4, except at ultraviolet wavelengths. At these wavelengths, we explore the potential influence of the 2175 Å-bump feature on the attenuation curve of high-redshift galaxies (Section 2.4.2.2). This feature has been detected recently in observations of low-redshift galaxies (e.g. Conroy et al. 2010a, Wild et al. 2011). To describe the 2175 Å bump, we introduce a Drude profile in the attenuation curve of the ambient interstellar medium in our models. The profile is parameterized by the central wavelength, the width and the strength of the bump. We keep the central wavelength and width constant, as suggested by observations through different lines of sight in the Milky Way. We randomly vary the strength of the bump between 0 and the average value in the Milky Way (i.e. $A_{\mathrm{bump}} = 5.17$, see below). We can thus write the attenuation curve in the ambient interstellar medium as

$$\hat{\tau}_\lambda = \mu \hat{\tau}_V \left(\frac{\lambda}{5500} \right)^{-n} + \frac{c}{R_\mathrm{V}} D(x) \tag{5.3.1}$$

$$D(x) = \frac{x^2}{(x^2 - x_0^2)^2 + \gamma^2 x^2}\,, \tag{5.3.2}$$

where $x = (\lambda/\mu\mathrm{m})^{-1}$. Following Whittet [2003], $x_0 = (0.2175)^{-1}$ characterizes the central wavelength and $1/\gamma = 0.99\mu\mathrm{m}^{-1}$ the FWHM of the feature. In the Milky Way, the coefficient c is related to the strength A_{bump} of the bump by

$$A_{\mathrm{bump}} = \int_0^\infty c\, D(x)\, dx = \frac{\pi c}{2\gamma} = 5.17\,. \tag{5.3.3}$$

In our models, the quantity c/R_V spans values between 0 and 1.07. The total effective V-band optical depth, $\hat{\tau}_V$, is sampled randomly between 0.01 and 3 (Figure 5.3, panel e), and the fraction of this contributed by dust in the ambient interstellar medium, μ, is sampled randomly between 0.1 to 0.7 (Figure 5.3, panel f).

We draw in this way a library of 500,000 model spectral energy distributions sampling comprehensive ranges of properties of high-redshift galaxies.[4] We note that the observer-frame H-band stellar mass-to-light ratio in Figure 5.3 (panel a) spans almost 2 orders of magnitude, i.e. far more than the r-band mass-to-light ratio in the library used in Chapter 3. This is because the H band filter samples a wide range of rest wavelengths in the present library, from optical wavelengths at $z \sim 2$ to far-ultraviolet wavelengths at $z \sim 8$.

[4] We show in Appendix B that a few hundred thousand models are enough to reach convergence in the uncertainty and accuracy of the retrieved parameters.

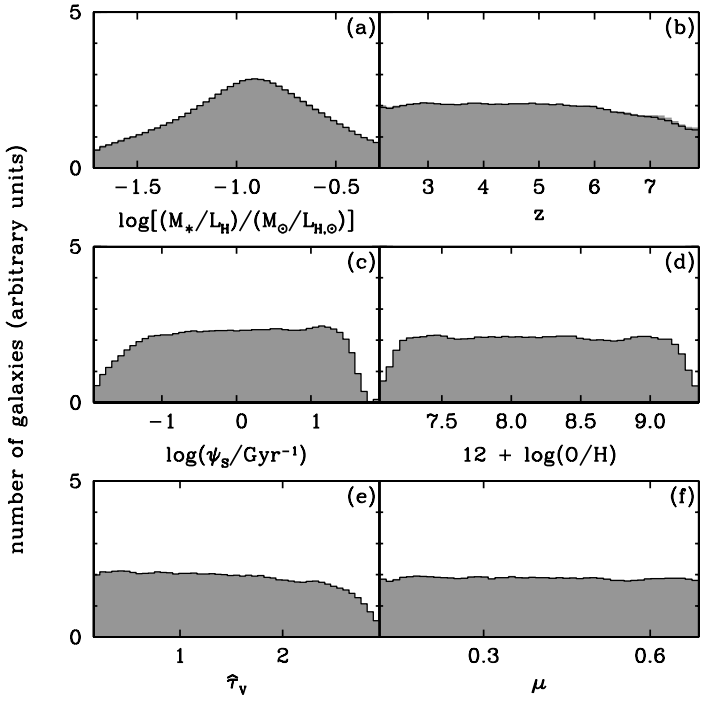

Figure 5.3: Prior distributions of selected physical parameters of the 500,000 galaxies in the spectral library: (a) observer-frame absolute H-band stellar mass-to-light ratio, M_*/L_H; (b) redshift, z; (c) specific star formation rate, $\psi_{\rm S}$; (d) gas-phase oxygen abundance, $12 + \log({\rm O/H})$; (e) total effective V-band absorption optical depth of the dust, $\hat{\tau}_V$; (f) fraction of $\hat{\tau}_V$ arising from dust in the ambient ISM, μ. In each panel, the shaded histogram shows the distribution for all galaxies, while the solid histogram shows the contribution by star-forming galaxies alone. Non-star-forming galaxies are off scale (at $\log(\psi_{\rm S}) = \infty$) in panel (c) and do not contribute to the distributions of interstellar parameters in panels (d)–(f).

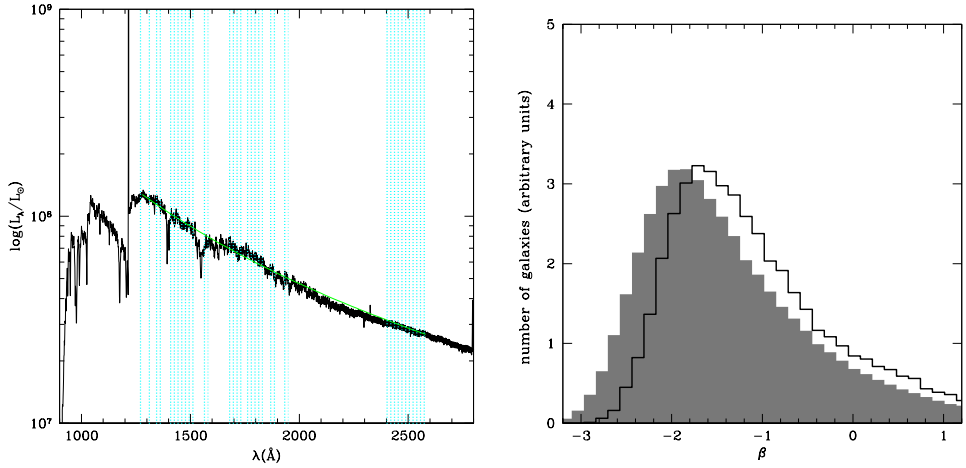

Figure 5.4: Left panel: rest-frame ultraviolet spectral energy distribution of a galaxy ($z = 3.41$) extracted form the model library (black) and the associated power-law best fit (green). The wavelength windows used to perform the fit are highlighted in cyan. Right panel: distribution of the rest-frame ultraviolet slope in the model library for all galaxies (shaded histogram) and for galaxies with negligible 2175Å bump only (solid histogram).

5.3.2 The ultraviolet spectral slope

Optical and near-infrared observations of high-redshift galaxies sample the rest-frame ultraviolet light. It is thus important to investigate the dependence of the ultraviolet light on galaxy physical parameters. For example, the slope of the ultraviolet spectral energy distribution will depend on the age and metallicity of the stars dominating the light, as well as on attenuation by dust [Meurer et al., 1999, Dunlop et al., 2012]. Recently, different authors have debated the steepness of this ultraviolet slope in high-redshift galaxies. Some studies support the idea that high-redshift galaxies are characterized by very steep rest-frame ultraviolet spectral energy distributions [Bouwens et al., 2011], while others suggest that such steep slopes are caused by biases in the data reduction process [flux boosting; Dunlop et al., 2012].

We adopt here the definition of Calzetti et al. [1994] to measure the rest-frame ultraviolet slope of model spectral energy distributions. This amounts to fitting a power law of the form $F_\lambda \propto \lambda^\beta$ to the spectrum sampled in 10 narrow wavelength windows, chosen to avoid strong stellar and interstellar absorption features and the 2175 Å bump (see also Section 5.3.1). Figure 5.4 (left panel) shows an example of model spectrum (in black) and the associated power-law fit (in green). The wavelength windows used to perform the fit are highlighted in cyan. In Figure 5.4 (right panel), the shaded histogram shows the distribution of ultraviolet slopes of all the galaxies in the model library. For reference, the solid histogram shows the same distribution when including only galaxies with negligible ultraviolet bumps

($c/R_{\rm V} < 0.1$).[5] We note that the strength of the bump in our library is not correlated to any other galaxy parameter, such as age, stellar metallicity and optical depth of the dust. Thus, the galaxy populations contributing to the two histograms are comparable. In this context, the fact that galaxies with significant bumps have significantly steeper ultraviolet spectra (by $\Delta\beta \sim 0.4$) than galaxies with negligible bumps in Figure 5.4 (right panel) indicates that the selected narrow wavelength windows do not entirely remove the dependence of β on the 2175 Å dust feature.

5.4 Fitting procedure

In this Section, we use the model library built in Section 5.3.1 to estimate physical parameters of HUDF galaxies extracted from the photometric catalogue described in Section 5.2. We also select the model spectral energy distributions that best fit these HUDF galaxies to generate a pseudo-observed scene for JWST/NIRSpec with a preliminary version of the instrument performance simulator.

5.4.1 Estimates of the physical parameters

To constrain the physical parameters of HUDF galaxies using the model library built in Section 5.3.1, we first compute the photometric properties of the models by convolving the (observer-frame) spectral energy distributions with the response functions of the four ACS filters and the two NICMOS filters. Galaxies at $z \gtrsim 4$ are not detected in the ACS bands that fall blueward the redshifted Lyman break. When this is the case, we assign a photometric flux equal to 0 in these bands and adopt the same error as in the ACS z band (the errors are typically roughly the same for all detected ACS bands and a factor of 10 larger for NICMOS photometry). Then, we use a Bayesian approach similar to that described in Chapter 3 to compare the observed fluxes of each HUDF galaxy with the fluxes of all the models in the library. This allows us to compute marginalized likelihood distributions of six physical parameters for each HUDF galaxy: stellar mass, M_*, specific star formation rate, $\psi_{\rm S}$, gas-phase oxygen abundance, $12 + \log({\rm O/H})$, total effective optical depth of the dust, $\hat{\tau}_V$, rest-frame ultraviolet slope, β, and redshift, z.

We note that, when the redshift is determined (photometrically), the rest-frame ultraviolet slope can be measured directly from the observed colors in the bands redward of the Lyman break (see Hathi et al. 2008 for galaxies at $z = 4 - 5 - 6$ and Dunlop et al. 2012 for galaxies at $z = 7 - 8$). Finkelstein et al. [2011] show that deriving the rest-frame ultraviolet slope from the best-fit model spectrum is more accurate than a single-color method. For this reason, and also because in our sample the Lyman break does not always fall in the observed wavelength range, we include the rest-frame ultraviolet slope, β, among the parameters that we retrieve using the Bayesian approach.

[5] The two histograms are normalized differently. The first is normalized to the total number of models in the library. The second is normalized to the number of models with $c/R_{\rm V} < 0.1$ (about 10 percent of the entire library).

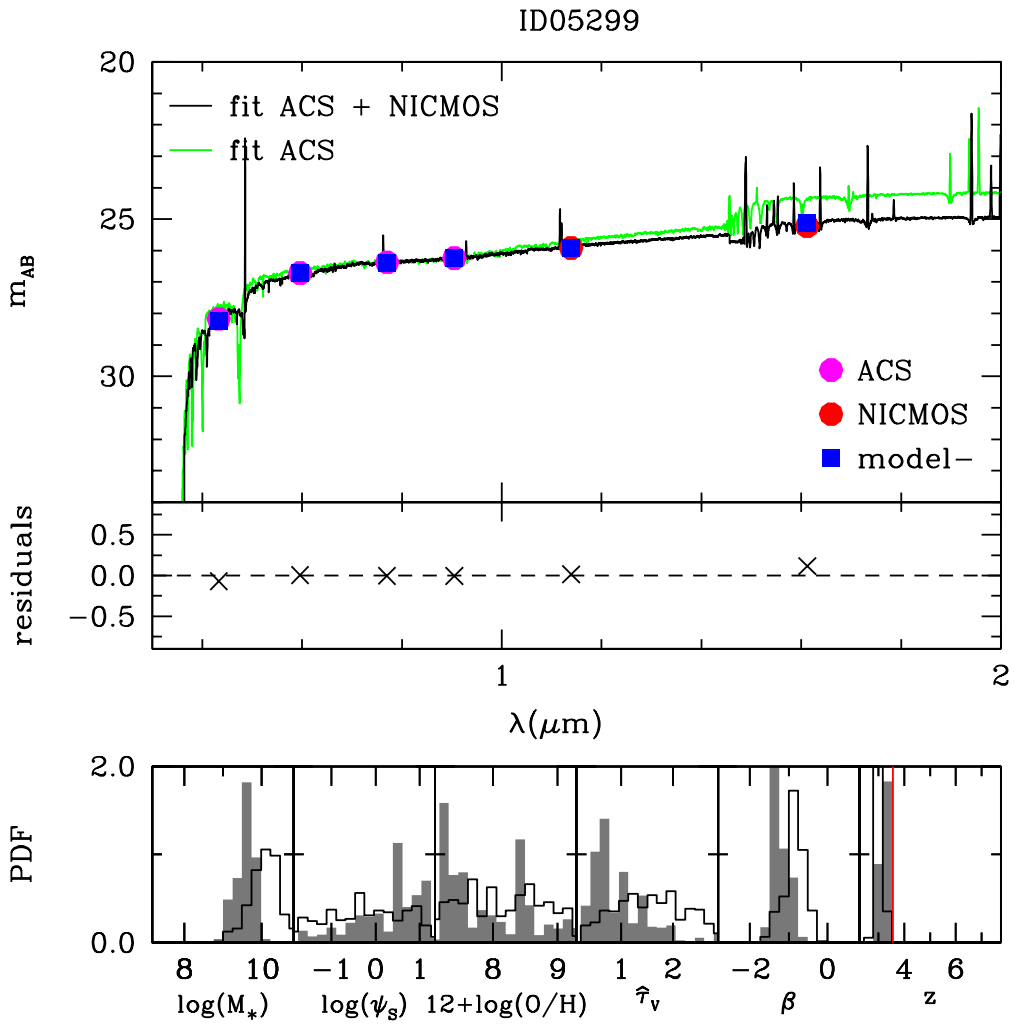

Figure 5.5: ID 5299. Top panel: ACS observed magnitudes (magenta dots), NICMOS observed magnitudes (red dots), ACS and NICMOS magnitudes of the best-fit model galaxy (blue squares), best-fit model spectrum when fitting both ACS and NICMOS magnitudes (black line), best-fit model spectrum when fitting ACS data alone (green line). Middle panel: residuals between best-fit model magnitudes and data (crosses). Bottom panels: probability density functions retrieved for stellar mass, specific star formation rate, gas-phase oxygen abundance, total effective optical depth of the dust, rest-frame ultraviolet slope, redshift. In each panel, the shaded histogram was derived by fitting both ACS and NICMOS data, while the solid histogram was derived using ACS data alone. In the redshift panel, the vertical red line shows the photometric redshift listed in the Coe et al. [2006] catalogue.

Figure 5.5 shows an example of a fit of a HUDF galaxy in our sample and the associated constraints on physical parameters. In the top panel, the dots represent the ACS (magenta) and NICMOS (red) data. Blue squares represent the ACS and NICMOS magnitudes of the best-fit model galaxy. The corresponding best-fit model spectrum is plotted in black. Also shown in green is the model spectrum that best fits the ACS data alone, i.e. when NICMOS observations are ignored. The crosses in the middle panel show the residuals between best-fit model magnitudes and data. In the bottom panels, we show the probability density functions retrieved for (from left to right) stellar mass, specific star formation rate, gas-phase oxygen abundance, optical depth of the gas, rest-frame ultraviolet slope and redshift. In each panel, the shaded histogram was derived by fitting both ACS and NICMOS data, while the solid histogram was derived using ACS data alone. In the redshift panel (far right), the vertical red line shows, for reference, the photometric redshift listed in the Coe et al. [2006] catalogue. As expected, all the parameters are less well constrained (i.e. the likelihood distributions are wider) when only the ACS data are used. For this particular galaxy, neglecting NICMOS constraints causes the rest-frame ultraviolet slope and the stellar mass to be both overestimated.

We return to the physical properties of HUDF galaxies in Section 5.5 below, where we pay particular attention to the accuracy of photometric redshift estimates and the correlation between rest-frame ultraviolet slope and dust attenuation.

5.4.2 Preliminary pseudo-observed scene

We use the best-fit model spectral energy distributions of HUDF galaxies derived in Section 5.4.1 to generate a pseudo-observed scene with the JWST/NIRSpec instrument performance simulator. To present the scene, we first need to briefly describe the design of NIRSpec. NIRSpec is a near-infrared multi-object dispersive spectrograph, which will be capable of simultaneously observing more than 100 sources over a field-of-view larger than $3' \times 3'$. Targets will normally be selected by opening groups of shutters to form multiple apertures in a micro-shutter array. Over 62,000 such micro-shutters are arranged in a waffle-like grid on the array. Six gratings provide high-resolution ($R = 2700$) and medium-resolution ($R = 1000$) spectroscopy over the wavelength range $1 - 5\,\mu$m, while a prism yields lower-resolution ($R = 100$) spectroscopy over the range $0.6 - 5\,\mu$m.[6]

Our simulation of a pseudo-observed scene for JWST/NIRSpec is still preliminary. So far, each galaxy is assumed to be point-like and placed exactly 'behind' a micro-shutter. We show in Figure 5.6 the multi-object spectrograph (MOS) scene obtained in this way with a typical exposure configuration of a NIRSpec deep-field observation. Each light-blue 'line' is a galaxy observed through 3 vertically-adjacent micro-shutters with the low-resolution prism (the galaxy is placed behind the central micro-shutter, while the adjacent shutters provide an estimate the sky background).[7]

[6] NIRSpec also includes fixed slits and an integral field unit, which provide high-contrast spectra of point-like sources and of extended objects, respectively.

[7] In this example, some of our 55 best-fit spectral energy distributions are used more than once

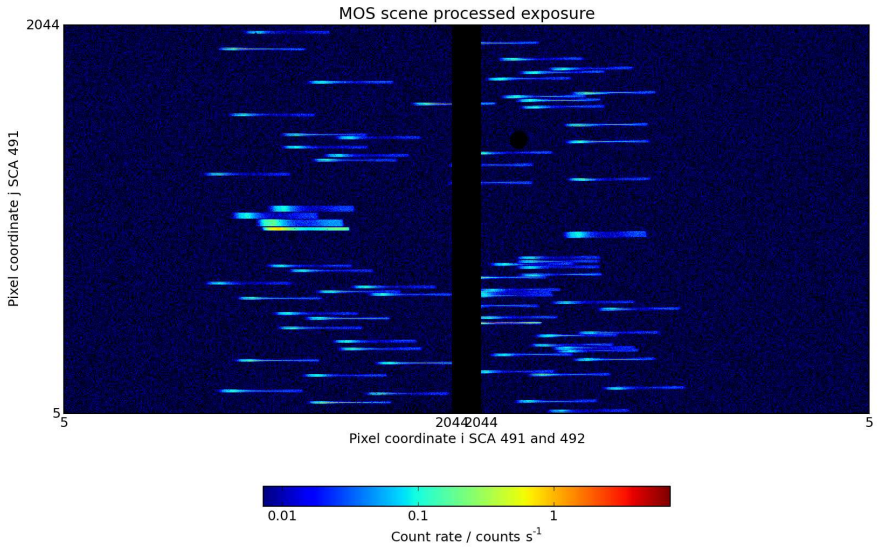

Figure 5.6: Example of the multi-object spectrograph (MOS) scene obtained by placing different point-like galaxies 'behind' different micro-shutters. Light is dispersed using NIRSpec low-resolution prism ($R = 100$).

In this broad overview, all galaxies look very similar to each other because the dominant feature is the background radiation of the Zodiacal light. Zooming closer (Figure 5.7), we can discern a continuum radiation slightly above the background level and some peaks of emission lines.

Once the instrument performance simulator is fully operational, we will be able to mimic spectroscopic observations of spatially-extended galaxies with the low-resolution prism and the medium-resolution gratings. Then, we will use an approach similar to that described in Section 5.4 to derive constraints on the physical parameters of these pseudo-observed galaxies using a comprehensive library of model spectra (generated using the library of spectral energy distribution of Section 5.3.1). This will allow us to quantify the power of JWST/NIRSpec observations relative to currently available photometric observations to explore the physical properties of primordial galaxies.

5.5 Discussion

In this Section, we return to the physical properties of HUDF galaxies derived from the photometric analysis described in Section 5.4. We first compare our estimates

to fill the scene with $\sim$ 70 galaxies. The 5 wide spectra in the center of the scene in Figure 5.6 correspond to the fixed slits used to detect the Zodiacal background.

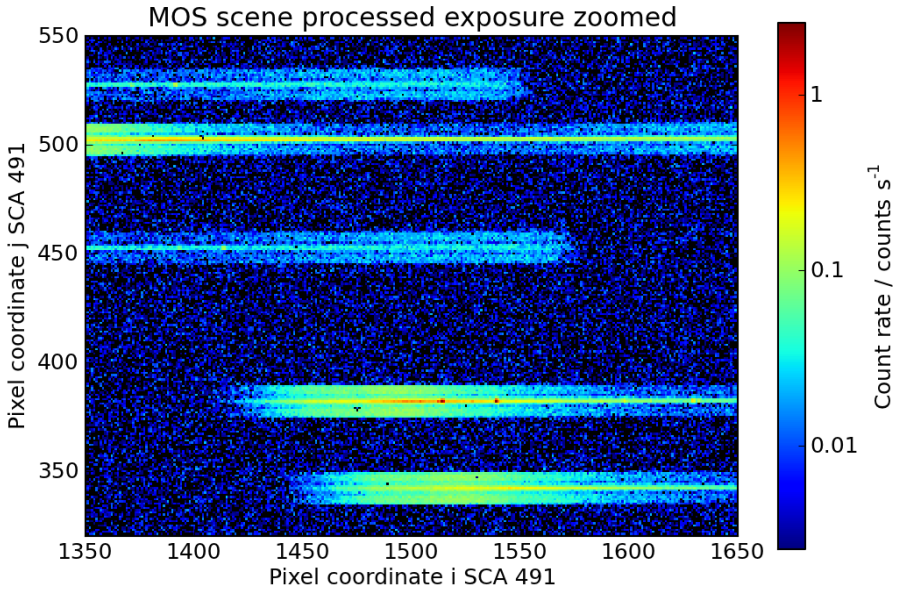

Figure 5.7: Zoom of the scene presented in Figure 5.6. In correspondence of every galaxy, 3 vertically-adjacent micro-shutters are open. In each case, the radiation of the Zodiacal light dominates the background in the upper and lower micro-shutters. The continuum emission from the galaxies and some emission lines (red isolated pixels) can be detected in the central micro-shutter.

of redshift with the photometric redshifts listed in the original Coe et al. [2006] catalogue. Then, we explore the relation between rest-frame ultraviolet spectral slope and total effective optical depth of the dust, which has strong implications for estimates of the star formation rate and dust attenuation in galaxies [e.g., Meurer et al., 1999, Charlot and Fall, 2000].

5.5.1 Comparison of redshift estimates

The photometric redshifts listed in the original Coe et al. [2006] catalogue were derived using an updated version of the Bayesian photometric redshift software BPZ [Benítez, 2000]. The template library used by Benítez [2000] includes templates of E, Sbc, Scd and Im galaxies from Coleman et al. [1980], two starburst-galaxy templates from Kinney et al. [1996] and two young (5 and 25 Myr old) stellar-population templates from [Bruzual and Charlot, 2003]. These last two templates were included to conform the library to the large population of blue galaxies observed in the HUDF.

In Figure 5.8, we compare our estimates of redshift with the photometric redshifts provided by Coe et al. [2006]. Gray dots pertain to galaxies for which both ACS and NICMOS photometry is available, while black dots refer to galaxies with ACS photometry alone. For each measurement, the error bar indicates the sum of the 16th–84th percentile range of the likelihood distribution retrieved from our analysis and the 1-sigma uncertainty in the Coe et al. estimate (≈ 0.3). The red dotted line marks the location of perfect agreement between the two redshift estimates. At $z \lesssim 3$, the median redshift retrieved from our analysis appears to be slightly larger (by $\Delta z \sim 0.3$–0.5) than that listed by Coe et al. [2006], while for $3 \lesssim z \lesssim 4$, it is smaller by a similar amount. At greater redshifts, discrepancies are smaller ($\Delta z \sim -0.15$), and the values favored by our analysis are in good general agreement (within the errors) with those listed by Coe et al. [2006]. Figure 5.8 therefore illustrates the substantial uncertainties affecting photometric redshift estimates, especially when the Lyman break cannot be detected with the ACS filters, i.e. for $z \lesssim 4$. JWST spectroscopic observations of these galaxies will largely improve redshift estimates by sampling strong rest-frame ultraviolet and optical (absorption and emission) features.

5.5.2 Correlation between ultraviolet spectral slope and optical depth of the dust

The rest-frame ultraviolet spectral slope is sensitive to stellar age, stellar metallicity and attenuation by dust. Young, metal-poor stellar populations tend to have steeper slopes than old, metal-rich ones. In this context, observations of high-redshift galaxies with very steep slopes have attracted much attention over the past few years, as these are suggestive of extremely young, metal-poor, dust-free primordial galaxies (see for example, Wilkins et al. 2011, Finkelstein et al. 2010, Bouwens et al. 2010).

In Figure 5.9, we plot rest-frame ultraviolet spectral slope, β, against total effec-

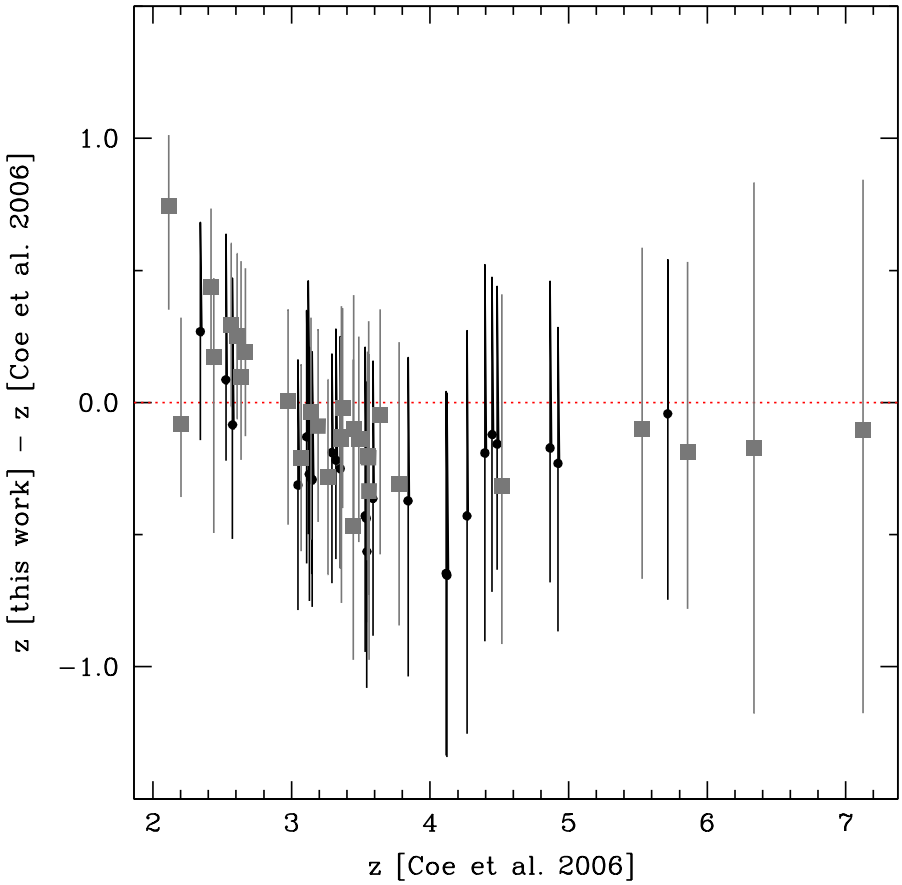

Figure 5.8: Differences in redshift estimates between our fits and the values reported in Coe et al. [2006] catalogue. Gray squares represent the fits using ACS and NICMOS magnitudes. Black dots represent the fits when ACS magnitudes alone are available. For each measurement, the error bar indicates the sum of the 16th–84th percentile range of the likelihood distribution retrieved from our analysis and the 1-sigma uncertainty in the Coe et al. estimate (≈ 0.3). The red dotted line marks the location of perfect agreement between the two redshift estimates.

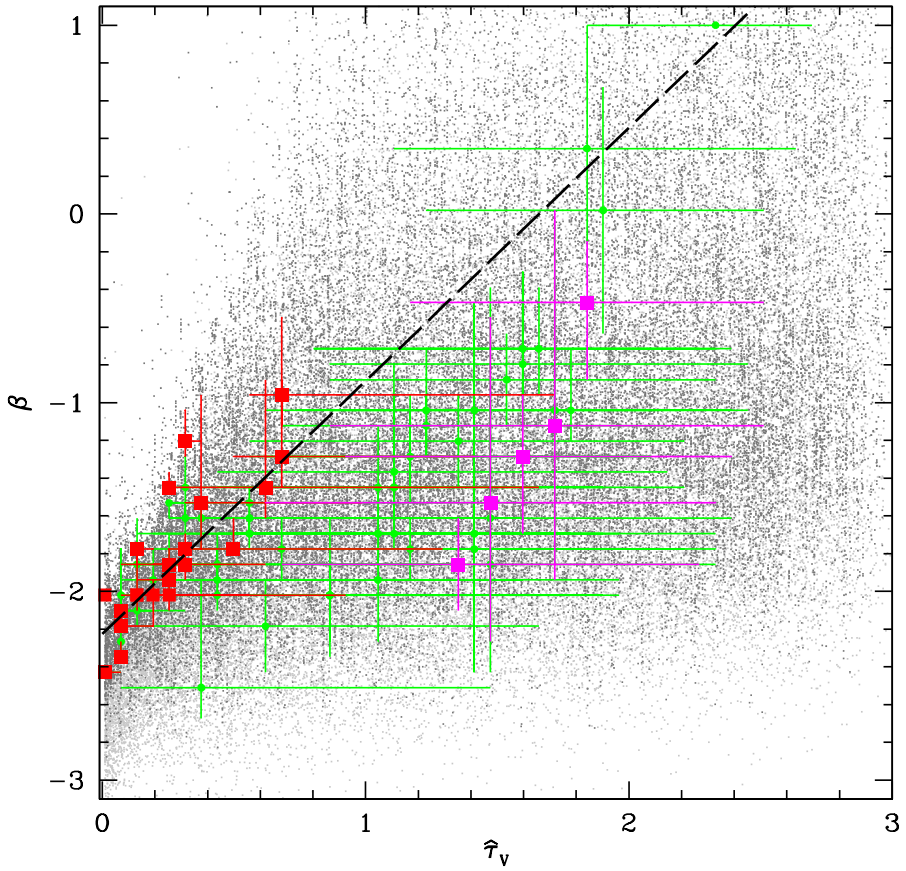

Figure 5.9: Estimates of rest-frame ultraviolet slope, β, versus total effective optical depth of the dust, $\hat{\tau}_V$. Grey dots: 50,000 random galaxies from the model library. Red ($z < 4$) and magenta ($z > 4$) squares: median estimates fitting both ACS and NICMOS data. Green dots: median estimates fitting ACS data alone. The black dashed line shows the relation derived by Meurer et al. [1999] for nearby, ultraviolet-selected starburst galaxies (see text for details).

tive optical depth of the dust, $\hat{\tau}_V$, for different samples. Gray dots show the properties of 50,000 galaxies extracted randomly from the model library of Section 5.3.1. Red and magenta squares show the median properties retrieved from our analysis for HUDF with both ACS and NICMOS data at redshifts smaller and greater than $z = 4$, respectively. Green dots show the median properties of HUDF galaxies with ACS data only. For each estimate, the error bar indicates the 16th–84th percentile range of the likelihood distribution retrieved from our analysis. Finally, the heavy dashed line is the relation derived by Meurer et al. [1999] for nearby, ultraviolet-selected starburst galaxies,

$$A_{1600} = 4.43 + 1.99\beta\,, \tag{5.5.4}$$

where $A_{1600} = 1.086\,\tau_{1600}$ and $\tau_{1600} = 2.46\,\tau_V$ for a Calzetti et al. [1994] law.

The large magenta horizontal error bars in Figure 5.9 indicate that the attenuation optical depth of the dust is not well constrained in objects at redshift above 4. At those redshifts, the Lyman break falls redward of the V band, and thus galaxies are detected only in 2 or 3 bands. This is not sufficient to tightly constrain the optical depth of the dust. The retrieved $\hat{\tau}_V$ in this case corresponds to the median of the prior distribution (Figure 5.3f). Also, the large green error bars (both horizontal and vertical) in Figure 5.9 show that ACS data alone do not allow us to tightly constrain the total effective optical depth of the dust and the ultraviolet spectral slope of galaxies at any of the considered redshifts. At low redshift ($z < 4$), our estimates of $\hat{\tau}_V$ and β for HUDF galaxies are roughly consistent with the Meurer et al. local relation. It is interesting to mention that all HUDF galaxies in our sample are generally best fitted by models with a negligible bump feature in the rest-frame ultraviolet attenuation curve.

A main result of this analysis is that the selected galaxies seem to be characterized by low dust content. This may be a consequence of our requirement to select only galaxies with z' or J magnitude brighter than AB = 27, implying that the selected galaxies tend to be bright and not strongly attenuated by dust. Deeper observations would likely bring more dusty galaxies into the sample.

5.6 Summary and next steps

This work is part of a project to prepare for the exploration of the early Universe with the NIRspec near-infrared multi-object spectrograph on board the James Webb Space Telescope (JWST). In this framework we have started a collaboration with Bernhard Dorner and Dr. Pierre Ferruit, who are responsible for the development of the NIRSpec instrument performance simulator. We have set for this project the following goals:

1. the creation of pseudo-observed scenes for the JWST/NIRSpec simulator using our model galaxy spectral energy distributions;

2. the quantitative assessment of the improvement in the estimates of galaxy physical parameters from future JWST/NIRSpec observations, relative to the constraints derived from current deep photometric observations.

Our results are still preliminary. So far, we have used the approach developed in Chapter 3 (and summarized in Section 5.3) to build a sophisticated library of spectral energy distributions tailored to studies of high-redshift galaxies. We have used this library to generate input spectra for the JWST/NIRSpec instrument performance simulator, by selecting model spectral energy distributions that best reproduce existing HUDF photometric observations of high-redshift galaxies. We have presented a first simulation of a pseudo-observed MOS scene, in which galaxies are treated as point sources and observed in the low-resolution mode of the NIRSpec instrument.

As soon as the instrument performance simulator is finalized, we will be able to use more realistic *pseudo observations* of this type to assess the accuracy and uncertainty to which physical parameters can be derived from future JWST/NIRSpec. This will allow us to quantify the improvement that JWST will bring compared to currently available photometric observations of the high-redshift Universe.

We have also analyzed in this Chapter the constraints that can be derived on the physical parameters of high-redshift galaxies from available HUDF photometry. We find that, not surprisingly, photometric redshift estimates are particularly uncertain when the Lyman break is not bracketed by any of the available ACS filters (i.e. for $z \lesssim 4$). Spectroscopic observations sampling strong ultraviolet and optical (absorption and emission) features with JWST/NIRSpec will greatly improve these constraints. For the small sample of ~ 20 HUDF galaxies with both ACS and NICMOS data at redshift $z < 4$ we have analyzed, the relation between rest-frame ultraviolet spectral slope and attenuation optical depth of the dust is similar to that observed by Meurer et al. [1999] in nearby ultraviolet-selected starburst galaxies. The fact that the galaxies in our sample are generally found to have low dust content and a weak 2175 Å-bump feature may result from the relatively bright magnitude limit (z' or J brighter than AB = 27).

The preliminary results presented here already provide valuable hints on how future large samples of spectroscopic observations with JWST/NIRSpec will help constrain the physical parameters of high-redshift galaxies. As soon as the instrument performance simulator is ready, we will be able to complete this study and optimize the design of future observations.

Chapter 6

Conclusions

In this thesis, we have seen that the light emitted by galaxies at different wavelengths encloses important information about the past history of star formation and chemical enrichment. Constraining the physical properties of galaxies from the analysis of this light requires a consistent model of the radiation emitted by stars and gas and of the effect of dust on this radiation. To this goal, we have developed a new approach to constrain galaxy physical parameters from different types of multi-wavelength observations. A main novelty of our approach is the ability to interpret simultaneously the stellar and nebular emission from galaxies. To achieve this, we rely on a comprehensive library of spectral energy distributions, assembled using state-of-the-art models of

- star formation and chemical enrichment histories from the semi-analytic post-treatment of cosmological simulations;
- stellar population synthesis, nebular emission and attenuation by dust.

This library enables us to estimate galaxy physical parameters from the Bayesian analysis of different types of observations in a wide range of redshifts. The spectra in the library extend from ultraviolet to near-infrared wavelengths and span broad ranges in all physical parameters. The distribution of priors has been designed to avoid unwanted biases in the sampling of the parameter space. This distribution can be tailored to the study of galaxies with specific physical properties (e.g., actively star-forming, early-type, metal-poor). In this thesis, we focus on estimates of redshift, stellar mass, specific star formation rate, gas-phase oxygen abundance and total effective optical depth of the dust. We also explore a new parameter, the fraction of a current galaxy stellar mass formed in the past 2.5 Gyr, which provides information about the recent history of star formation.

The powerful approach we have developed to extract key physical parameters from multi-wavelength galaxy observations also allows us to quantify how the reliability of such estimates depends on the type of observation employed. To achieve this, we simulate *pseudo-observations* by convolving the spectral energy distributions of models with known parameters with appropriate instrument responses and then applying artificial noise to mimic true observations. This approach allows us to make accurate predictions for the optimistic case in which the models we rely on are good approximations of true galaxies. Pseudo-observations allow us to assess the retrievability of physical parameters in terms of accuracy and uncertainty. This allows us in turn to identify how to best constrain physical parameters using different spectral resolutions and signal-to-noise ratios, which is crucial to design future

observations. We have simulated both photometric and spectroscopic observations, obtaining the following results:

- rest-frame *ugriz* photometry can provide roughly 50-percent uncertainty on mass-to-light ratio measurements; meaningful constraints can also be obtained on the specific star formation rate for actively star forming galaxies ($\psi_S > 1$ Gyr^{-1});

- equivalent-width measurements of prominent optical emission lines ([O II], Hβ, [O III], [N II], Hα, [S II]) probe the specific star formation rate and to some extent the gas-phase oxygen abundance and the optical depth of the dust seen by young stars in giant molecular clouds;

- pixel-by-pixel analysis of low-resolution (50 Å FWHM) spectral energy distributions of galaxies at wavelengths 3600-7400 Å enables us to extract valuable information from the most prominent emission and absorption features. This approach (in a regime of median signal-to-noise ratio per pixel around 20) is sufficient to constrain the observer-frame absolute r-band stellar mass-to-light ratio to within 0.13 dex, the fraction of a current galaxy stellar mass formed during the last 2.5 Gyr to within 0.23, the specific star formation rate to within 0.24 dex, the gas-phase oxygen abundance to within 0.28, the total effective V-band absorption optical depth of the dust to within 0.64 and the fraction of this arising from dust in the ambient interstellar medium to within 0.16;

- pixel-by-pixel analysis of medium-resolution spectral energy distributions (5 Å FWHM) at wavelengths 3600-7400 Å enables us to de-blend emission lines and thus improve the constraints on specific star formation rate, gas metallicity and dust optical depth by a factor of about 2 relative to low-resolution spectroscopy, while the constraints on other parameters improve by 20–30 percent.

We find that the retrievability of physical parameters is very sensitive to signal-to-noise ratio. Photometric measurements are generally gathered with good signal-to-noise ratio. Spectroscopy, instead, requires long exposure times to attain good signal-to-noise ratio. The ability with our approach to interpret simultaneously the stellar and nebular emission from galaxies allows us to show that low-resolution spectroscopy actually represents an excellent trade-off between photometry and medium-resolution spectroscopy.

The tool we have developed is useful to interpret currently available datasets and plan for future observations. We have tested our approach on the analysis of a sample of ∼13,000 SDSS spectra of nearby galaxies at a spectral resolution of 5Å FWHM. We find that we can infer likelihood distributions of stellar mass, gas-phase oxygen abundance, total effective optical depth of the dust and specific star formation rate similar to those obtained in previous separate analyses of the stellar and nebular emission at the original (twice higher) SDSS spectral resolution.

We have also started to analyze the spectral energy distributions of a sample of galaxies at redshifts between 1 and 3 (collaboration with Dr. Elisabete da Cunha

and Prof. Hans-Walter Rix, MPIA). This project consists in exploiting our models to interpret a combination of spectroscopic and photometric observations to best constrain galaxy physical parameters, such as stellar mass, specific star formation rate, gas-phase oxygen abundance and dust content. Photometric observations have been extracted from the FIREWORKS catalogue and span a wide wavelength range, from the rest-frame ultraviolet to the infrared. Spectroscopic observations have been extracted from the 3D-HST database and have the advantage to probe the narrow emission and absorption features at rest-frame ultraviolet and optical wavelengths. So far we have been working on understanding the differences in the constraints derived from the separate analyses of FIREWORKS and 3D-HST data. The combination of photometric and spectroscopic approaches will require the inclusion of the da Cunha et al. [2008] prescription for dust emission by in our model library. This should be achieved over the summer of 2012, when this project will be concluded.

This thesis is part of an effort to explore of the formation and evolution of the first galaxies in the early Universe with the NIRSpec near-infrared multi-object spectrograph on board the future James Webb Space Telescope (JWST). In this framework, we have started a collaboration with Bernhard Dorner and Dr. Pierre Ferruit, who are responsible for the development of the NIRSpec instrument performance simulator. The aim of our study is to create a pseudo-observed *scene* of NIRSpec using our model galaxy spectral energy distributions. Pseudo-observed scenes of this type will allow us to precisely simulate NIRSpec observations and quantify the improvement in the estimates of galaxy physical parameters relative to the constraints derived from current photometric observations of the deep Universe. In fact, the ability with JWST/NIRSpec to gather high-quality spectra of the rest-frame ultraviolet and optical emission from primeval galaxies must be exploited as best as possible.

Our results are still preliminary. So far we have used our approach to analyze a sample of galaxies selected among the deepest photometric observations in the Hubble Ultra Deep Field ($2 < z < 8$) and to generate input spectra with the most realistic colors for the NIRSpec instrument performance simulator. In this analysis, we are giving particular attention to the features of the rest-frame ultraviolet radiation. The shape of a galaxy spectral energy distributions in this wavelength region depends on the age and metallicity of the stars dominating the light as well as on the presence of dust. Our preliminary results indicate that the galaxies in our sample have generally low dust content and a weak 2175Å-bump feature. This result could likely derive form the relatively bright magnitude limit of this sample.

We are planning a few extension of our models. As mentioned above, we will first extend the spectral domain of the model library out to mid- and far-infrared wavelengths using the prescription of da Cunha et al. [2008]. We are also planning to extend our models to account for the contribution by active galactic nuclei to the line and continuum emission from galaxies. These improvements will help us derive more reliable constraints on the physical properties of both quiescent and star-forming galaxies from low up to very high redshifts. This should have important implications for our understanding of galaxy formation and evolution.

Appendix A

Intrinsic correlations between spectral pixels

In Chapters 3, 4 and 5 we apply a Bayesian approach to compute likelihood distributions of physical parameters of galaxies from different types of observations. We write the likelihood of the j th model, noted $\mathcal{M}_j$, given the spectral energy distribution of a galaxy, as

$$\ln P(\mathcal{M}_j|\{\mathcal{O}_i\}) = -\frac{1}{2}\sum_i \left(\frac{\mathcal{O}_i - w_j\mathcal{M}_{j,i}}{\sigma_i}\right)^2 , \tag{A.0.1}$$

where $\mathcal{O}_i$ and $\mathcal{M}_{j,i}$ are the spectral energy distributions (i.e., either the spectral flux density at the i th wavelength, or the equivalent width of the i th emission line, or the flux in the i th photometric band) of the pseudo-galaxy and the j th model, respectively, σ_i is the uncertainty in $\mathcal{O}_i$ corresponding to the adopted signal-to-noise ratio, and

$$w_j = \left(\sum_i \frac{\mathcal{O}_i\mathcal{M}_{j,i}}{\sigma_i^2}\right)\left[\sum_i \left(\frac{\mathcal{M}_{j,i}}{\sigma_i}\right)^2\right]^{-1} \tag{A.0.2}$$

is the model scaling factor that maximizes $P(\mathcal{M}_j|\{\mathcal{O}_i\})$ (this parameter is unity in the case of scale-free equivalent widths). This formulation is valid when the measurements are not correlated with one another. Correlated measurements contain the same information and thus do not add constraints to the fit, but instead may cause large χ^2 values. Correlations may derive from particular noise conditions (frequency dependent noise) or from an oversampling of the spectral energy distribution.

To analyze this issue in the case of medium-resolution spectral energy distributions, the best is to compute the mean *pseudo-observational* error covariance matrix of medium-resolution spectra. Pseudo-observations are preferable over real observations (SDSS) here because they allow us to investigate physical correlations between spectral pixels without contamination by the response function of the SDSS spectrograph. To proceed, we must consider a large sample of N pseudo-observations and express every pseudo-observed spectrum ($\mathcal{O}_{k,i}$, where k ranges between 1 and N and i is the index of the wavelength) as the sum of a best-fit model spectrum ($\mathcal{M}_{k,i}^{\rm best}$) and a residual spectrum ($\mathcal{R}_{k,i}$), i.e.,

$$\mathcal{O}_{k,i} = \mathcal{M}_{k,i}^{\rm best} + \mathcal{R}_{k,i}. \tag{A.0.3}$$

The residual spectra computed for this large sample of pseudo-observations will allow us to build the mean pseudo-observational error covariance matrix.

In practice, we select a sample of $N = 3000$ spectral energy distributions from the library of 5 million models assembled in Chapter 3 to compute pseudo-observations at resolving power $R = 1000$ and a median signal-to-noise ratio per pixel $\overline{\mathrm{S/N}} = 20$ (the features of the mean pseudo-observational error covariance matrix already converge with $N = 2000$). To select the best-fit model $\mathcal{M}_{k,i}^{\mathrm{best}}$ corresponding to each pseudo-observed spectrum $\mathcal{O}_{k,i}$, we compute the likelihood of every model spectrum $\mathcal{M}_{j,i}$ in the library given $\mathcal{O}_{k,i}$ as in Equations A.0.1 and A.0.2. For each of the 3000 pseudo-galaxies in the sample, we compute the residual spectrum

$$R_{k,i} = \mathcal{O}_{k,i} - \mathcal{M}_{k,i}^{\mathrm{best}}. \tag{A.0.4}$$

The mean pseudo-observational error covariance matrix can be expressed as

$$Cov_{i,i'} = (1/N) \sum_{k} [(R_{k,i} - \overline{R}_i)(R_{k,i'} - \overline{R}_{i'})] \tag{A.0.5}$$

where i and i' are any 2 spectral pixels, and

$$\overline{R}_i = (1/N) \sum_{k} R_{k,i} \tag{A.0.6}$$

is the mean residual spectrum. It is actually most convenient to investigate correlations between spectral pixels using the mean pseudo-observational error *correlation* matrix, defined as

$$Corr_{i,i'} = \frac{Cov_{i,i'}}{\sqrt{Cov_{i,i} Cov_{i',i'}}}\,. \tag{A.0.7}$$

Each element of the mean pseudo-observational error correlation matrix can take values between $Corr_{i,i'} = -1$ (maximum anti-correlation) and $+1$ (maximum correlation), with 0 corresponding to no correlation at all between the spectral pixels i and i'. All elements on the diagonal are unity, i.e., $Corr_{i,i} = 1$.

Figure A.1 shows a representation of mean pseudo-observational error correlation matrix $Corr_{i,i'}$. The only cases (outside the diagonal) where $Corr_{i,i'}$ takes values close to unity are when i an i' are neighboring pixels of the same emission line (creating small dark squares of ~ 5 pixels on the side around the diagonal). Pixels belonging to different emission lines also correlate with one another, since they are produced by the same H II regions, but to a lesser extent (at the level of ~ 30 percent). Figure A.1 also reveals some very weak anti-correlations (at the level of ~ 1 percent) between emission-line and continuum fluxes, which presumably result from slight imperfections in the fits of emission-line strengths (dark horizontal and vertical lines). For all other spectral pixel pairs in Figure A.1, the correlations are below 1 percent. Hence most of the correlation signal in Figure A.1 is carried by emission-line pixels, which represent only ~ 7 percent of the entire spectral energy distribution. The information provided by these pixels is important, as the relative strengths of different emission lines provide important constraints on the star formation rate, gas-phase metallicity and dust content of the observed galaxies.

We now check that pixel-to-pixel correlations may have non-negligible influence on the calculation of model likelihoods. To quantify this, we rewrite Equation A.0.1

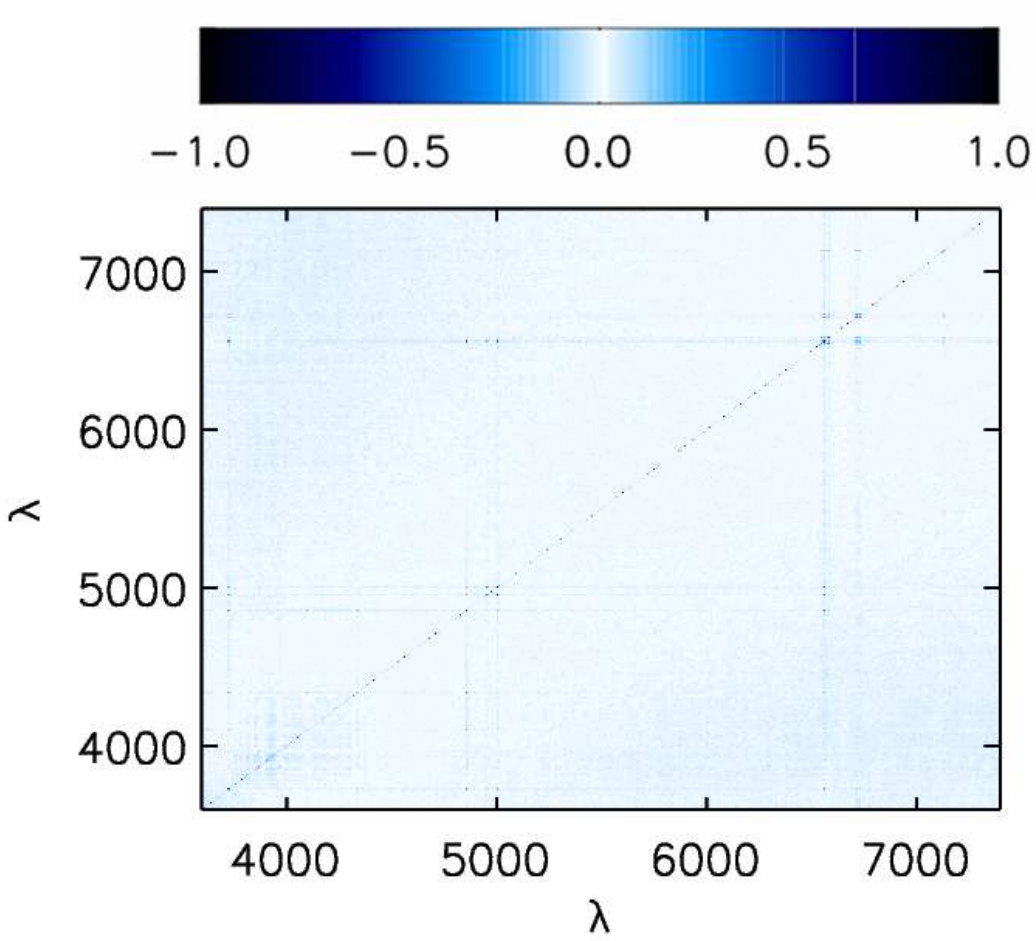

Figure A.1: Representation of the mean pseudo-observational error correlation matrix $Corr_{i,i'}$ (equation A.0.7) across the wavelength range from 3600 to 7400 Å, derived as described in the text from the residuals of the fits of 3000 pseudo-observed spectral energy distributions with the library of 5 million models assembled in Chapter 3. The color scale is indicated at the top.

including the mean pseudo-observational error correlation matrix (and omitting the pseudo-galaxy index k introduced above),

$$\ln P(\mathcal{M}_j|\{\mathcal{O}_i\}) = -\frac{1}{2}\sum_i \left(\frac{\mathcal{O}_i - w_j\mathcal{M}_{j,i}}{\sigma_i}\right) Corr^{-1}_{i,i'} \left(\frac{\mathcal{O}_{i'} - w_j\mathcal{M}_{j,i'}}{\sigma_{i'}}\right). \quad \text{(A.0.8)}$$

Here $Corr^{-1}_{i,i'}$ is the "pseudo-inverse" of the mean pseudo-observational error correlation matrix, which contains the weights of the different elements contributing to the likelihood (correlated pixels having less weight than uncorrelated ones). A pseudo-inverse matrix must be used here because, as is generally the case for matrices so large that some rows (or columns) are likely to be linear combinations of the others, the $Corr_{i,i'}$ matrix has zero determinant (i.e. it is singular). To compute $Corr^{-1}_{i,i'}$, we appeal to a single-value-decomposition algorithm, which performs the inversion of a singular matrix by means of 2 orthogonal matrices and a diagonal matrix with positive or zero elements (the singular values). The difficulty in this procedure is that some level of arbitrariness is introduced by the fact that, to perform the inversion, the algorithm must cut out near-zero singular values (i.e. with inverse tending to infinity). Typically, the more conservative this cut, the smaller the elements of $Corr^{-1}_{i,i'}$, and hence, the larger the likelihood given by equation (A.0.8).

To explore the influence of this choice on the computation of model likelihoods, we have calculated the pseudo-inverse of the mean error correlation matrix using different cuts of near-zero singular values. We find that, when using a conservative cut (e.g. keeping all singular values larger than 0.1), the computation is extremely heavy and the results are nearly identical to those obtained using Equation A.0.1. That is, the likelihoods are generally rounded to zero, we must appeal again to *relative* rather than *absolute* likelihoods, and the constraints on physical parameters are similar. Non-zero absolute likelihoods can be obtained only if we use a much less conservative cut, i.e., by rejecting all singular values smaller than at least 2. This however causes the elements $Corr^{-1}_{i,i'}$ of the pseudo-inverse matrix to be extremely small, of the order of ~ 0.01 on the diagonal and close to 0 everywhere else. The implied drop in the weights of correlated pixels (corresponding to emission lines) makes retrieved physical parameters, such as star formation rate, gas-phase oxygen abundance and optical depth of the dust, deviate strongly from their true values.

We conclude from all this that Equation A.0.1 provides a reasonable way of computing model likelihoods and that neglecting intrinsic correlations between spectral pixels has a negligible influence on our results.

Appendix B

Number of model galaxies in the library

In previous studies, galaxy physical parameters have been usually derived using libraries assembled with about 50,000 model galaxies. Such libraries have always been considered large enough to make estimates of physical parameters converge. We address here the issue related to the minimum number of models required to make estimates of stellar mass-to-light ratio and specific star formation rate converge when fitting medium-resolution spectral energy distributions.

In our study, the inclusion of emission lines expands the parameter space significantly. Because of this, in our approach, 50,000 models are not enough to obtain meaningful constraints on physical parameters, as we would not properly sample the parameter space. We need to sample a comprehensive range of stochastic star formation and chemical enrichment histories extracted from galaxy formation simulations (which probe a much larger range of galaxy spectral energy distributions than most previous studies of this type) and variations in 3 parameters to describe nebular emission (metallicity of the gas, zero-age ionization parameter and dust-to-metal ratio in the ionized gas) and in 3 parameters to describe the attenuation by dust (total V-band attenuation optical depth of the dust, fraction of this contributing by dust in the diffuse interstellar medium and slope of the attenuation curve in the diffuse interstellar medium). When fitting simultaneously the continuum and nebular emission from a galaxy, all possible combinations of these parameters must be taken into account to build the probability density functions of physical parameters. This requires a library containing at least a few million models.

To illustrate this, we show in Figure B.1 how the uncertainty and accuracy (defined as in the caption of Table 3.2) of the stellar mass-to-light ratio M_*/L_r and specific star formation rate $\psi_{\rm S}$ retrieved from 2 types of spectral fits depend on the number of spectral energy distributions in the model library. Panels (a) and (b) show the uncertainty and accuracy in M_*/L_r (black) and $\psi_{\rm S}$ (blue) retrieved from medium-resolution spectroscopy. Panels (c) and (d) those retrieved form multi-band photometry (note that the vertical scales are different from those in Panels a and b). Figure B.1 illustrates once again that medium-resolution spectroscopy provides tighter and more accurate constraints than multi-band photometry on M_*/L_r and $\psi_{\rm S}$. It also demonstrates that a few million models are required to reach these performances. In the case of multi-band photometry, a few hundred thousand models are enough to reach convergence in the uncertainty and accuracy of the retrieved parameters.

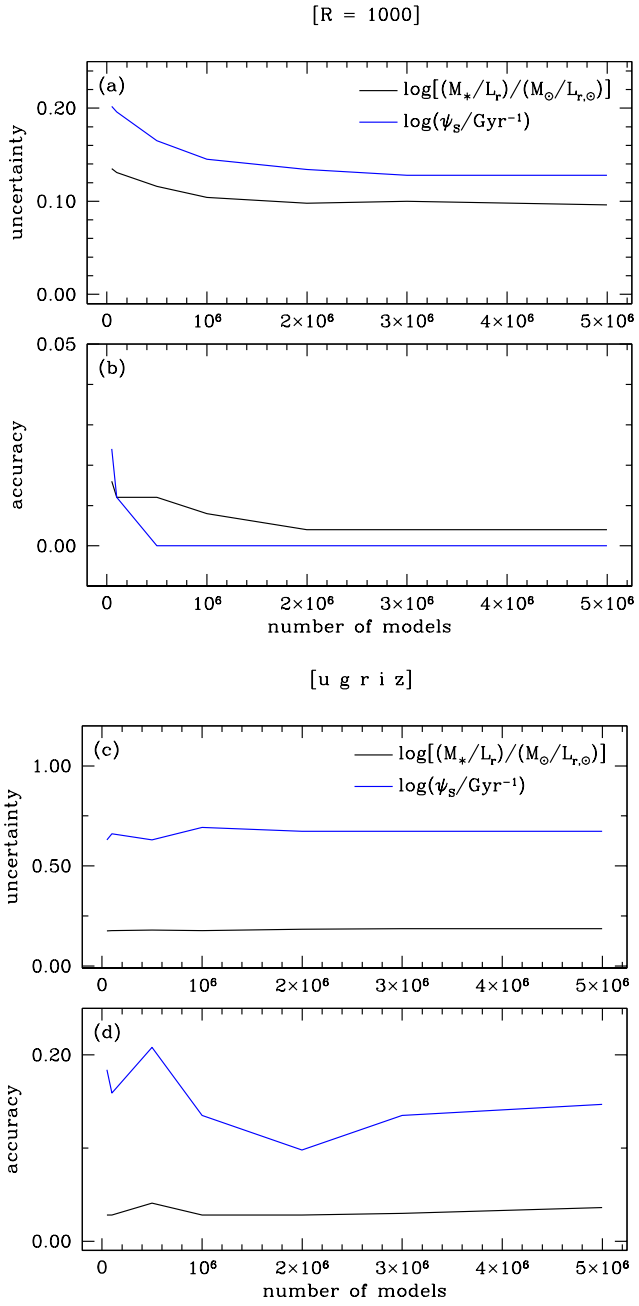

Figure B.1: Uncertainty and accuracy (defined as in the caption of Table 3.2) of the stellar mass-to-light ratio M_*/L_r (black curves) and specific star formation rate ψ_S (blue curves) retrieved from 2 types of spectral fits, plotted against number of spectral energy distributions in the model library. (*a*) & (*b*) As retrieved from medium-resolution spectroscopy. (*c*) & (*d*) As retrieved form multi-band photometry (note that the vertical scales are different from those in Panels *a* and *b*).

Appendix C

UDF data

In the following pages, we report the photometric data of the 55 UDF galaxies used in Chapter 5.

ID	$z_{\rm tab}$	B	V	i'	z'	J	H
41066	7.125	...	...	...	29.531 ± 0.318	27.509 ± 0.234	26.760 ± 0.189
41271	6.336	...	30.007 ± 0.490	30.220 ± 0.624	28.062 ± 0.205	27.307 ± 0.470	26.912 ± 0.384
32408	5.859	...	...	29.144 ± 0.151	27.340 ± 0.060	26.471 ± 0.174	26.337 ± 0.176
3398	5.715	...	30.396 ± 0.297	28.287 ± 0.053	26.788 ± 0.027	...	...
3317	5.530	29.878 ± 0.437	29.162 ± 0.174	27.915 ± 0.064	26.752 ± 0.040	26.908 ± 0.154	26.522 ± 0.146
8664	4.924	...	28.721 ± 0.099	26.845 ± 0.021	26.774 ± 0.036	...	...
7352	4.866	...	28.765 ± 0.105	26.786 ± 0.020	26.764 ± 0.034	...	...
2422	4.518	...	29.223 ± 0.141	27.894 ± 0.048	27.925 ± 0.087	27.142 ± 0.250	26.709 ± 0.223
197	4.484	30.108 ± 0.467	27.899 ± 0.049	26.660 ± 0.018	26.649 ± 0.031	...	...
3620	4.448	...	28.150 ± 0.049	26.982 ± 0.019	26.978 ± 0.034	...	...
6135	4.395	...	27.877 ± 0.038	26.995 ± 0.019	27.154 ± 0.040	...	...
148	4.266	30.368 ± 0.528	27.608 ± 0.034	26.727 ± 0.017	26.697 ± 0.030	...	...
401	4.122	29.149 ± 0.333	26.309 ± 0.019	25.382 ± 0.009	25.222 ± 0.015	...	...
56445	4.116	29.777 ± 0.369	27.112 ± 0.027	26.140 ± 0.011	25.834 ± 0.015	...	...
3148	3.841	29.082 ± 0.195	26.754 ± 0.017	26.085 ± 0.010	25.917 ± 0.017	...	...
5181	3.777	29.354 ± 0.174	27.768 ± 0.029	27.206 ± 0.019	27.319 ± 0.039	27.406 ± 0.499	26.605 ± 0.340
2750	3.641	27.958 ± 0.103	26.397 ± 0.017	26.051 ± 0.014	26.053 ± 0.024	26.321 ± 0.143	25.972 ± 0.141
259	3.587	27.991 ± 0.100	26.326 ± 0.015	25.778 ± 0.010	25.708 ± 0.018	...	...
5299	3.559	28.169 ± 0.090	26.715 ± 0.017	26.377 ± 0.014	26.240 ± 0.022	25.918 ± 0.135	25.231 ± 0.135
7675	3.556	28.547 ± 0.143	26.856 ± 0.022	26.295 ± 0.015	26.180 ± 0.024	25.857 ± 0.117	25.263 ± 0.118
4860	3.549	27.018 ± 0.068	25.689 ± 0.014	25.326 ± 0.011	25.260 ± 0.018	25.316 ± 0.088	24.851 ± 0.094
57117	3.544	25.730 ± 0.041	24.316 ± 0.008	24.049 ± 0.007	23.977 ± 0.013	...	...
7957	3.540	27.169 ± 0.050	25.848 ± 0.010	25.431 ± 0.008	25.428 ± 0.014	...	...

ID	$z_{\rm tab}$	B	V	i'	z'	J	H
53090	3.531	28.407 ± 0.100	26.999 ± 0.019	26.690 ± 0.015	26.698 ± 0.028	...	...
9185	3.486	27.715 ± 0.059	26.511 ± 0.014	26.152 ± 0.011	26.145 ± 0.020	25.860 ± 0.160	26.052 ± 0.197
5006	3.448	27.305 ± 0.035	26.218 ± 0.009	26.157 ± 0.009	26.259 ± 0.019	26.141 ± 0.128	25.997 ± 0.133
9349	3.446	27.906 ± 0.080	26.795 ± 0.021	26.420 ± 0.016	26.341 ± 0.029	26.222 ± 0.181	26.162 ± 0.189
5346	3.368	28.136 ± 0.067	27.074 ± 0.017	26.833 ± 0.015	27.011 ± 0.032	27.029 ± 0.184	26.835 ± 0.186
5322	3.360	28.890 ± 0.123	27.906 ± 0.036	27.517 ± 0.028	27.573 ± 0.053	27.225 ± 0.208	26.924 ± 0.199
5856	3.354	26.993 ± 0.030	25.967 ± 0.008	25.734 ± 0.007	25.802 ± 0.014	25.843 ± 0.132	25.707 ± 0.138
7151	3.352	28.030 ± 0.078	27.020 ± 0.022	26.720 ± 0.018	26.804 ± 0.036	...	...
4450	3.320	27.962 ± 0.091	26.963 ± 0.025	26.762 ± 0.023	27.022 ± 0.051	...	...
9676	3.291	27.080 ± 0.043	26.113 ± 0.012	25.953 ± 0.012	25.994 ± 0.022	...	...
4047	3.262	27.761 ± 0.052	27.106 ± 0.020	27.042 ± 0.021	27.184 ± 0.042	26.922 ± 0.174	27.107 ± 0.197
7874	3.191	27.556 ± 0.044	26.715 ± 0.014	26.636 ± 0.014	26.713 ± 0.028	26.607 ± 0.159	26.176 ± 0.162
4267	3.149	25.469 ± 0.020	24.626 ± 0.007	24.454 ± 0.006	24.422 ± 0.010	...	...
6504	3.147	27.515 ± 0.041	26.727 ± 0.014	26.546 ± 0.013	26.634 ± 0.025	...	...
8354	3.140	27.519 ± 0.048	26.699 ± 0.016	26.622 ± 0.016	26.649 ± 0.030	26.698 ± 0.149	25.995 ± 0.141
9853	3.128	27.306 ± 0.054	26.578 ± 0.019	26.348 ± 0.018	26.389 ± 0.033	...	...
53643	3.121	26.533 ± 0.083	25.889 ± 0.030	25.715 ± 0.029	26.008 ± 0.067	...	...
383	3.110	28.185 ± 0.093	27.025 ± 0.022	26.715 ± 0.019	26.587 ± 0.031	...	...
6082	3.068	23.216 ± 0.003	22.577 ± 0.001	22.486 ± 0.001	22.546 ± 0.003	22.470 ± 0.060	22.497 ± 0.071
9312	3.047	26.851 ± 0.030	26.132 ± 0.011	25.972 ± 0.011	25.970 ± 0.019	...	...
7658	2.975	27.638 ± 0.052	26.874 ± 0.018	26.786 ± 0.019	26.746 ± 0.033	26.694 ± 0.252	26.061 ± 0.196
5964	2.666	26.348 ± 0.024	26.008 ± 0.012	25.976 ± 0.013	25.968 ± 0.023	25.959 ± 0.114	25.444 ± 0.119
8322	2.637	27.967 ± 0.049	27.634 ± 0.025	27.666 ± 0.029	27.695 ± 0.054	27.317 ± 0.222	26.946 ± 0.201

ID	$z_{\rm tab}$	B	V	i'	z'	J	H
7758	2.606	27.341 ± 0.040	27.059 ± 0.021	27.144 ± 0.025	27.206 ± 0.048	26.859 ± 0.144	26.501 ± 0.144
9385	2.574	27.274 ± 0.046	26.992 ± 0.025	26.997 ± 0.028	26.944 ± 0.047	...	...
1738	2.563	26.298 ± 0.015	26.267 ± 0.010	26.366 ± 0.012	26.565 ± 0.026	26.585 ± 0.138	26.168 ± 0.138
841	2.526	26.517 ± 0.018	26.381 ± 0.011	26.479 ± 0.013	26.503 ± 0.023	...	...
3205	2.440	27.308 ± 0.034	27.115 ± 0.019	27.119 ± 0.022	27.057 ± 0.036	26.736 ± 0.156	26.369 ± 0.154
533	2.419	25.610 ± 0.013	25.516 ± 0.008	25.519 ± 0.009	25.539 ± 0.015	25.303 ± 0.135	24.784 ± 0.126
9706	2.343	26.124 ± 0.027	25.618 ± 0.012	25.473 ± 0.011	25.405 ± 0.019	...	...
7681	2.202	25.831 ± 0.015	25.677 ± 0.009	25.597 ± 0.009	25.498 ± 0.015	25.032 ± 0.096	24.597 ± 0.103
3060	2.114	26.344 ± 0.019	26.478 ± 0.015	26.603 ± 0.018	26.738 ± 0.036	26.790 ± 0.136	26.218 ± 0.128

Bibliography

K. N. Abazajian, J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, C. Allende Prieto, D. An, K. S. J. Anderson, S. F. Anderson, J. Annis, N. A. Bahcall, and et al. The Seventh Data Release of the Sloan Digital Sky Survey. *Astrophysical Journal, Supplement*, 182:543–558, June 2009. doi: 10.1088/0067-0049/182/2/543. (Cited on page 59.)

P. Anders and U. Fritze-v. Alvensleben. Spectral and photometric evolution of young stellar populations: The impact of gaseous emission at various metallicities. *Astronomy and Astrophysics*, 401:1063–1070, April 2003. doi: 10.1051/0004-6361:20030151. (Cited on page 39.)

J. A. Baldwin, M. M. Phillips, and R. Terlevich. Classification parameters for the emission-line spectra of extragalactic objects. *Publications of the ASP*, 93:5–19, February 1981. doi: 10.1086/130766. (Cited on pages 59 and 92.)

J. E. Beckman, M. Rozas, A. Zurita, R. A. Watson, and J. H. Knapen. Populations of High-Luminosity Density-bounded H II Regions in Spiral Galaxies: Evidence and Implications. *Astronomical Journal*, 119:2728–2744, June 2000. doi: 10.1086/301380. (Cited on page 35.)

E. F. Bell and R. S. de Jong. Stellar Mass-to-Light Ratios and the Tully-Fisher Relation. *Astrophysical Journal*, 550:212–229, March 2001. doi: 10.1086/319728. (Cited on page 54.)

N. Benítez. Bayesian Photometric Redshift Estimation. *Astrophysical Journal*, 536:571–583, June 2000. doi: 10.1086/308947. (Cited on page 133.)

C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, E. Wollack, E. L. Wright, C. Barnes, M. R. Greason, R. S. Hill, E. Komatsu, M. R. Nolta, N. Odegard, H. V. Peiris, L. Verde, and J. L. Weiland. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results. *Astrophysical Journal, Supplement*, 148:1–27, September 2003. doi: 10.1086/377253. (Cited on page 17.)

G. Bertelli, A. Bressan, C. Chiosi, F. Fagotto, and E. Nasi. Theoretical isochrones from models with new radiative opacities. *Astronomy and Astrophysics, Supplement*, 106:275–302, August 1994. (Cited on page 31.)

G. Bertelli, L. Girardi, P. Marigo, and E. Nasi. Scaled solar tracks and isochrones in a large region of the Z-Y plane. I. From the ZAMS to the TP-AGB end for 0.15-2.5 $M\odot$ stars. *Astronomy and Astrophysics*, 484:815–830, June 2008. doi: 10.1051/0004-6361:20079165. (Cited on pages xiv, 31, 32 and 58.)

G. Bertelli, E. Nasi, L. Girardi, and P. Marigo. Scaled solar tracks and isochrones in a large region of the Z-Y plane. II. From 2.5 to 20 $M_{\odot}$ stars. *Astronomy and Astrophysics*, 508:355–369, December 2009. doi: 10.1051/0004-6361/200912093. (Cited on pages 31 and 58.)

J. Blaizot, B. Guiderdoni, J. E. G. Devriendt, F. R. Bouchet, S. J. Hatton, and F. Stoehr. GALICS- III. Properties of Lyman-break galaxies at a redshift of 3. *Monthly Notices of the RAS*, 352:571–588, August 2004. doi: 10.1111/j.1365-2966.2004.07947.x. (Cited on page 56.)

L. Blitz and F. H. Shu. The origin and lifetime of giant molecular cloud complexes. *Astrophysical Journal*, 238:148–157, May 1980. doi: 10.1086/157968. (Cited on page 60.)

R. J. Bouwens, G. D. Illingworth, P. A. Oesch, M. Stiavelli, P. van Dokkum, M. Trenti, D. Magee, I. Labbé, M. Franx, C. M. Carollo, and V. Gonzalez. Discovery of $z \sim 8$ Galaxies in the Hubble Ultra Deep Field from Ultra-Deep WFC3/IR Observations. *Astrophysical Journal, Letters*, 709:L133–L137, February 2010. doi: 10.1088/2041-8205/709/2/L133. (Cited on pages 125 and 133.)

R. J. Bouwens, G. D. Illingworth, P. A. Oesch, M. Franx, I. Labbe, M. Trenti, P. van Dokkum, C. M. Carollo, V. Gonzalez, and D. Magee. UV-continuum slopes at z˜4-7 from the HUDF09+ERS+CANDELS observations: Discovery of a well-defined UV-color magnitude relationship for $z > 4$ star-forming galaxies. *ArXiv e-prints*, September 2011. (Cited on pages 48 and 127.)

R. G. Bower, A. J. Benson, R. Malbon, J. C. Helly, C. S. Frenk, C. M. Baugh, S. Cole, and C. G. Lacey. Breaking the hierarchy of galaxy formation. *Monthly Notices of the RAS*, 370:645–655, August 2006. doi: 10.1111/j.1365-2966.2006.10519.x. (Cited on page 20.)

M. Boylan-Kolchin, V. Springel, S. D. M. White, A. Jenkins, and G. Lemson. Resolving cosmic structure formation with the Millennium-II Simulation. *Monthly Notices of the RAS*, 398:1150–1164, September 2009. doi: 10.1111/j.1365-2966.2009.15191.x. (Cited on pages xxvii, 17 and 18.)

J. Brinchmann, S. Charlot, S. D. M. White, C. Tremonti, G. Kauffmann, T. Heckman, and J. Brinkmann. The physical properties of star-forming galaxies in the low-redshift Universe. *Monthly Notices of the RAS*, 351:1151–1179, July 2004. doi: 10.1111/j.1365-2966.2004.07881.x. (Cited on pages xvi, xxi, 14, 38, 54, 57, 59, 60, 61, 62, 91, 92, 94, 95, 96, 97, 98 and 101.)

G. Bruzual and S. Charlot. Spectral evolution of stellar populations using isochrone synthesis. *Astrophysical Journal*, 405:538–553, March 1993. doi: 10.1086/172385. (Cited on page 22.)

G. Bruzual and S. Charlot. Stellar population synthesis at the resolution of 2003. *Monthly Notices of the RAS*, 344:1000–1028, October 2003. doi: 10.1046/j.

1365-8711.2003.06897.x. (Cited on pages xiii, xiv, 14, 15, 16, 24, 30, 31, 53, 57, 58, 92, 94, 104, 107, 117 and 133.)

D. Calzetti. Reddening and Star Formation in Starburst Galaxies. *Astronomical Journal*, 113:162–184, January 1997. doi: 10.1086/118242. (Cited on page 46.)

D. Calzetti, A. L. Kinney, and T. Storchi-Bergmann. Dust extinction of the stellar continua in starburst galaxies: The ultraviolet and optical extinction law. *Astrophysical Journal*, 429:582–601, July 1994. doi: 10.1086/174346. (Cited on pages 46, 127 and 136.)

D. Calzetti, L. Armus, R. C. Bohlin, A. L. Kinney, J. Koornneef, and T. Storchi-Bergmann. The Dust Content and Opacity of Actively Star-forming Galaxies. *Astrophysical Journal*, 533:682–695, April 2000. doi: 10.1086/308692. (Cited on pages xv, 45 and 46.)

J. A. Cardelli, G. C. Clayton, and J. S. Mathis. The relationship between infrared, optical, and ultraviolet extinction. *Astrophysical Journal*, 345:245–256, October 1989. doi: 10.1086/167900. (Cited on pages xv, 44 and 45.)

R. Cen and J. Ostriker. A hydrodynamic treatment of the cold dark matter cosmological scenario. *Astrophysical Journal*, 393:22–41, July 1992. doi: 10.1086/171482. (Cited on page 21.)

G. Chabrier. Galactic Stellar and Substellar Initial Mass Function. *Publications of the ASP*, 115:763–795, July 2003. doi: 10.1086/376392. (Cited on pages xiii, xiv, 14, 15, 23, 24, 31 and 58.)

S. Charlot and S. M. Fall. A Simple Model for the Absorption of Starlight by Dust in Galaxies. *Astrophysical Journal*, 539:718–731, August 2000. doi: 10.1086/309250. (Cited on pages 31, 46, 47, 60, 62, 96, 107 and 133.)

S. Charlot and M. Longhetti. Nebular emission from star-forming galaxies. *Monthly Notices of the RAS*, 323:887–903, May 2001. doi: 10.1046/j.1365-8711.2001.04260.x. (Cited on pages xiii, 14, 15, 38, 39, 40, 41, 53, 58, 59, 83, 92 and 94.)

R. Cid Fernandes, A. Mateus, L. Sodré, G. Stasińska, and J. M. Gomes. Semi-empirical analysis of Sloan Digital Sky Survey galaxies - I. Spectral synthesis method. *Monthly Notices of the RAS*, 358:363–378, April 2005. doi: 10.1111/j.1365-2966.2005.08752.x. (Cited on pages 54 and 62.)

D. Coe, N. Benítez, S. F. Sánchez, M. Jee, R. Bouwens, and H. Ford. Galaxies in the Hubble Ultra Deep Field. I. Detection, Multiband Photometry, Photometric Redshifts, and Morphology. *Astronomical Journal*, 132:926–959, August 2006. doi: 10.1086/505530. (Cited on pages xxiii, xxiv, xxv, 122, 124, 129, 130, 133 and 134.)

P. Coelho, G. Bruzual, S. Charlot, A. Weiss, B. Barbuy, and J. W. Ferguson. Spectral models for solar-scaled and α-enhanced stellar populations. *Monthly Notices of the RAS*, 382:498–514, December 2007. doi: 10.1111/j.1365-2966.2007.12364.x. (Cited on page 53.)

G. D. Coleman, C.-C. Wu, and D. W. Weedman. Colors and magnitudes predicted for high redshift galaxies. *Astrophysical Journal, Supplement*, 43:393–416, July 1980. doi: 10.1086/190674. (Cited on page 133.)

M. Colless, G. Dalton, S. Maddox, W. Sutherland, P. Norberg, S. Cole, J. Bland-Hawthorn, T. Bridges, R. Cannon, C. Collins, W. Couch, N. Cross, K. Deeley, R. De Propris, S. P. Driver, G. Efstathiou, R. S. Ellis, and et al. The 2dF Galaxy Redshift Survey: spectra and redshifts. *Monthly Notices of the RAS*, 328:1039–1063, December 2001. doi: 10.1046/j.1365-8711.2001.04902.x. (Cited on page 53.)

C. Conroy, D. Schiminovich, and M. R. Blanton. Dust Attenuation in Disk-dominated Galaxies: Evidence for the 2175 Å Dust Feature. *Astrophysical Journal*, 718:184–198, July 2010a. doi: 10.1088/0004-637X/718/1/184. (Cited on pages 48 and 125.)

C. Conroy, M. White, and J. E. Gunn. The Propagation of Uncertainties in Stellar Population Synthesis Modeling. II. The Challenge of Comparing Galaxy Evolution Models to Observations. *Astrophysical Journal*, 708:58–70, January 2010b. doi: 10.1088/0004-637X/708/1/58. (Cited on page 54.)

D. J. Croton, V. Springel, S. D. M. White, G. De Lucia, C. S. Frenk, L. Gao, A. Jenkins, G. Kauffmann, J. F. Navarro, and N. Yoshida. Erratum: The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. *Monthly Notices of the RAS*, 367:864–864, April 2006. doi: 10.1111/j.1365-2966.2006.09994.x. (Cited on pages 20 and 55.)

O. Cucciati, L. Tresse, O. Ilbert, O. Le Fevre, B. Garilli, V. Le Brun, P. Cassata, P. Franzetti, D. Maccagni, M. Scodeggio, E. Zucca, G. Zamorani, S. Bardelli, M. Bolzonella, R. M. Bielby, H. J. McCracken, A. Zanichelli, and D. Vergani. The Star Formation Rate Density and Dust Attenuation Evolution over 12 Gyr with the VVDS Surveys. *ArXiv e-prints*, September 2011. (Cited on pages xiii, 12 and 13.)

E. da Cunha, S. Charlot, and D. Elbaz. A simple model to interpret the ultraviolet, optical and infrared emission from galaxies. *Monthly Notices of the RAS*, 388:1595–1617, August 2008. doi: 10.1111/j.1365-2966.2008.13535.x. (Cited on pages 14, 31, 47, 62, 103, 104, 107, 120 and 141.)

M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The evolution of large-scale structure in a universe dominated by cold dark matter. *Astrophysical Journal*, 292:371–394, May 1985. doi: 10.1086/163168. (Cited on page 19.)

M. Davis, S. M. Faber, J. Newman, A. C. Phillips, R. S. Ellis, C. C. Steidel, C. Conselice, A. L. Coil, D. P. Finkbeiner, D. C. Koo, P. Guhathakurta, B. Weiner, R. Schiavon, C. Willmer, N. Kaiser, G. A. Luppino, G. Wirth, and et al. Science Objectives and Early Results of the DEEP2 Redshift Survey. In P. Guhathakurta, editor, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 4834 of *Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference*, pages 161–172, February 2003. doi: 10.1117/12.457897. (Cited on page 53.)

G. De Lucia and J. Blaizot. The hierarchical formation of the brightest cluster galaxies. *Monthly Notices of the RAS*, 375:2–14, February 2007. doi: 10.1111/j.1365-2966.2006.11287.x. (Cited on pages xiv, xxi, 19, 20, 55, 57, 68, 95, 96, 97 and 98.)

G. De Lucia, G. Kauffmann, and S. D. M. White. Chemical enrichment of the intracluster and intergalactic medium in a hierarchical galaxy formation model. *Monthly Notices of the RAS*, 349:1101–1116, April 2004. doi: 10.1111/j.1365-2966.2004.07584.x. (Cited on page 20.)

M. A. Dopita, B. A. Groves, R. S. Sutherland, L. Binette, and G. Cecil. Are the Narrow-Line Regions in Active Galaxies Dusty and Radiation Pressure Dominated? *Astrophysical Journal*, 572:753–761, June 2002. doi: 10.1086/340429. (Cited on page 41.)

M. A. Dopita, B. A. Groves, J. Fischera, R. S. Sutherland, R. J. Tuffs, C. C. Popescu, L. J. Kewley, M. Reuland, and C. Leitherer. Modeling the Pan-Spectral Energy Distribution of Starburst Galaxies. I. The Role of ISM Pressure and the Molecular Cloud Dissipation Timescale. *Astrophysical Journal*, 619:755–778, February 2005. doi: 10.1086/423948. (Cited on page 41.)

M. A. Dopita, J. Fischera, R. S. Sutherland, L. J. Kewley, C. Leitherer, R. J. Tuffs, C. C. Popescu, W. van Breugel, and B. A. Groves. Modeling the Pan-Spectral Energy Distribution of Starburst Galaxies. III. Emission Line Diagnostics of Ensembles of Evolving H II Regions. *Astrophysical Journal, Supplement*, 167:177–200, December 2006a. doi: 10.1086/508261. (Cited on pages 38 and 41.)

M. A. Dopita, J. Fischera, R. S. Sutherland, L. J. Kewley, R. J. Tuffs, C. C. Popescu, W. van Breugel, B. A. Groves, and C. Leitherer. Modeling the Pan-Spectral Energy Distribution of Starburst Galaxies. II. Control of the H II Region Parameters. *Astrophysical Journal*, 647:244–255, August 2006b. doi: 10.1086/505418. (Cited on page 41.)

J. S. Dunlop, R. J. McLure, B. E. Robertson, R. S. Ellis, D. P. Stark, M. Cirasuolo, and L. de Ravel. A critical analysis of the ultraviolet continuum slopes (β) of high-redshift galaxies: no evidence (yet) for extreme stellar populations at $z > 6$. *Monthly Notices of the RAS*, 420:901–912, February 2012. doi: 10.1111/j.1365-2966.2011.20102.x. (Cited on pages 49, 127 and 128.)

J. J. Eldridge and E. R. Stanway. Spectral population synthesis including massive binaries. *Monthly Notices of the RAS*, 400:1019–1028, December 2009. doi: 10.1111/j.1365-2966.2009.15514.x. (Cited on page 30.)

D. K. Erb, C. C. Steidel, A. E. Shapley, M. Pettini, N. A. Reddy, and K. L. Adelberger. The Stellar, Gas, and Dynamical Masses of Star-forming Galaxies at $z \sim 2$. *Astrophysical Journal*, 646:107–132, July 2006. doi: 10.1086/504891. (Cited on page 14.)

D. K. Erb, M. Pettini, A. E. Shapley, C. C. Steidel, D. R. Law, and N. A. Reddy. Physical Conditions in a Young, Unreddened, Low-metallicity Galaxy at High Redshift. *Astrophysical Journal*, 719:1168–1190, August 2010. doi: 10.1088/0004-637X/719/2/1168. (Cited on pages 48 and 125.)

S. M. Faber. Quadratic programming applied to the problem of galaxy population synthesis. *Astronomy and Astrophysics*, 20:361–374, September 1972. (Cited on page 22.)

M. N. Fanelli, R. W. O'Connell, and T. X. Thuan. Spectral synthesis in the ultraviolet. II - Stellar populations and star formation in blue compact galaxies. *Astrophysical Journal*, 334:665–687, November 1988. doi: 10.1086/166869. (Cited on page 46.)

G. J. Ferland. *Hazy, A Brief Introduction to Cloudy 90.* 1996. (Cited on pages 40 and 59.)

S. L. Finkelstein, C. Papovich, M. Giavalisco, N. A. Reddy, H. C. Ferguson, A. M. Koekemoer, and M. Dickinson. On the Stellar Populations and Evolution of Star-forming Galaxies at $6.3 < z < 8.6$. *Astrophysical Journal*, 719:1250–1273, August 2010. doi: 10.1088/0004-637X/719/2/1250. (Cited on page 133.)

S. L. Finkelstein, C. Papovich, B. Salmon, K. Finlator, M. Dickinson, H. C. Ferguson, M. Giavalisco, A. M. Koekemoer, N. A. Reddy, R. Bassett, C. J. Conselice, J. S. Dunlop, S. M. Faber, N. A. Grogin, N. P. Hathi, D. D. Kocevski, K. Lai, K.-S. Lee, R. J. McLure, B. Mobasher, and J. A. Newman. CANDELS: The Evolution of Galaxy Rest-Frame Ultraviolet Colors from z = 8 to 4. *ArXiv e-prints*, October 2011. (Cited on page 128.)

M. Fioc and B. Rocca-Volmerange. PEGASE: a UV to NIR spectral evolution model of galaxies. Application to the calibration of bright galaxy counts. *Astronomy and Astrophysics*, 326:950–962, October 1997. (Cited on pages 30 and 39.)

E. L. Fitzpatrick and D. Massa. An analysis on the shapes of ultraviolet extinction curves. I - The 2175 A bump. *Astrophysical Journal*, 307:286–294, August 1986. doi: 10.1086/164415. (Cited on page 46.)

A. Gallazzi, S. Charlot, J. Brinchmann, S. D. M. White, and C. A. Tremonti. The ages and metallicities of galaxies in the local universe. *Monthly Notices of the*

RAS, 362:41–58, September 2005. doi: 10.1111/j.1365-2966.2005.09321.x. (Cited on pages xxi, 54, 91, 92, 94, 96, 97 and 101.)

J. Gallego, J. Zamorano, A. Aragon-Salamanca, and M. Rego. The Current Star Formation Rate of the Local Universe. *Astrophysical Journal, Letters*, 455:L1, December 1995. doi: 10.1086/309804. (Cited on pages xiii and 13.)

T. Garel, J. Blaizot, B. Guiderdoni, D. Schaerer, A. Verhamme, and M. Hayes. Modelling high redshift Lyman-alpha Emitters. *ArXiv e-prints*, February 2012. (Cited on page 49.)

B. Garilli, O. Le Fèvre, L. Guzzo, D. Maccagni, V. Le Brun, S. de la Torre, B. Meneux, L. Tresse, P. Franzetti, G. Zamorani, A. Zanichelli, L. Gregorini, D. Vergani, D. Bottini, R. Scaramella, M. Scodeggio, G. Vettolani, C. Adami, S. Arnouts, S. Bardelli, M. Bolzonella, A. Cappi, S. Charlot, P. Ciliegi, T. Contini, S. Foucaud, I. Gavignaud, O. Ilbert, A. Iovino, F. Lamareille, H. J. McCracken, B. Marano, C. Marinoni, A. Mazure, R. Merighi, S. Paltani, R. Pellò, A. Pollo, L. Pozzetti, M. Radovich, E. Zucca, J. Blaizot, A. Bongiorno, O. Cucciati, Y. Mellier, C. Moreau, and L. Paioro. The Vimos VLT deep survey. Global properties of 20,000 galaxies in the I_{AB} <22.5 WIDE survey. *Astronomy and Astrophysics*, 486:683–695, August 2008. doi: 10.1051/0004-6361:20078878. (Cited on page 12.)

R. M. González Delgado, M. Cerviño, L. P. Martins, C. Leitherer, and P. H. Hauschildt. Evolutionary stellar population synthesis at high spectral resolution: optical wavelengths. *Monthly Notices of the RAS*, 357:945–960, March 2005. doi: 10.1111/j.1365-2966.2005.08692.x. (Cited on page 53.)

R. A. González-López lira, G. Bruzual-A., S. Charlot, J. Ballesteros-Paredes, and L. Loinard. Tracers of stellar mass loss - I. Optical and near-IR colours and surface brightness fluctuations. *Monthly Notices of the RAS*, 403:1213–1238, April 2010. doi: 10.1111/j.1365-2966.2009.16205.x. (Cited on page 58.)

Y. Goranova, P. Hudelot, F. Magnard, H. McCracken, Y. Mellier, M. Monnerville, M. Schultheis, J.-C. Sémah, J.-C. Cuillandre, and H. Aussel. The CFHTLS T0006 Release. *http://terapix.iap.fr/cplt/T0006-doc.pdf*, October 2009. (Cited on page 53.)

K. D. Gordon, G. C. Clayton, K. A. Misselt, A. U. Landolt, and M. J. Wolff. A Quantitative Comparison of the Small Magellanic Cloud, Large Magellanic Cloud, and Milky Way Ultraviolet to Near-Infrared Extinction Curves. *Astrophysical Journal*, 594:279–293, September 2003. doi: 10.1086/376774. (Cited on pages xv, 44 and 45.)

S. Gottlöber, A. Klypin, and A. V. Kravtsov. Merging History as a Function of Halo Environment. *Astrophysical Journal*, 546:223–233, January 2001. doi: 10.1086/318248. (Cited on pages xxvii and 18.)

S. Gottloeber, G. Yepes, C. Wagner, and R. Sevilla. The MareNostrum Universe. *ArXiv Astrophysics e-prints*, August 2006. (Not cited.)

B. A. Groves, M. A. Dopita, and R. S. Sutherland. Dusty, Radiation Pressure-Dominated Photoionization. I. Model Description, Structure, and Grids. *Astrophysical Journal, Supplement*, 153:9–73, July 2004. doi: 10.1086/421113. (Cited on page 41.)

B. Guiderdoni and B. Rocca-Volmerange. A model of spectrophotometric evolution for high-redshift galaxies. *Astronomy and Astrophysics*, 186:1–21, November 1987. (Cited on pages 22 and 49.)

Q. Guo, S. White, C. Li, and M. Boylan-Kolchin. How do galaxies populate dark matter haloes? *Monthly Notices of the RAS*, 404:1111–1120, May 2010. doi: 10.1111/j.1365-2966.2010.16341.x. (Cited on page 21.)

N. P. Hathi, S. Malhotra, and J. E. Rhoads. Starburst Intensity Limit of Galaxies at z˜=5-6. *Astrophysical Journal*, 673:686–693, February 2008. doi: 10.1086/524836. (Cited on page 128.)

S. Hatton, J. E. G. Devriendt, S. Ninin, F. R. Bouchet, B. Guiderdoni, and D. Vibert. GALICS- I. A hybrid N-body/semi-analytic model of hierarchical galaxy formation. *Monthly Notices of the RAS*, 343:75–106, July 2003. doi: 10.1046/j.1365-8711.2003.05589.x. (Cited on pages xxvii, 18 and 49.)

T. M. Heckman, G. Kauffmann, J. Brinchmann, S. Charlot, C. Tremonti, and S. D. M. White. Present-Day Growth of Black Holes and Bulges: The Sloan Digital Sky Survey Perspective. *Astrophysical Journal*, 613:109–118, September 2004. doi: 10.1086/422872. (Cited on page 54.)

A. Heger, S. E. Woosley, C. L. Fryer, and N. Langer. Massive Star Evolution Through the Ages. In W. Hillebrandt & B. Leibundgut, editor, *From Twilight to Highlight: The Physics of Supernovae*, pages 3–540, 2003. doi: 10.1007/10828549_1. (Cited on pages xiv, 27 and 28.)

D. J. Hollenbach and A. G. G. M. Tielens. Dense Photodissociation Regions (PDRs). *Annual Review of Astron and Astrophys*, 35:179–216, 1997. doi: 10.1146/annurev.astro.35.1.179. (Cited on page 35.)

A. M. Hopkins. On the Evolution of Star-forming Galaxies. *Astrophysical Journal*, 615:209–221, November 2004. doi: 10.1086/424032. (Cited on page 12.)

A. M. Hopkins and J. F. Beacom. On the Normalization of the Cosmic Star Formation History. *Astrophysical Journal*, 651:142–154, November 2006. doi: 10.1086/506610. (Cited on page 12.)

E. Hubble. A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Sciences of the United States of America*, 15:168, 1929. doi: 10.1073/pnas.15.3.168. (Cited on page 1.)

E. P. Hubble. Extragalactic nebulae. *Astrophysical Journal*, 64:321–369, December 1926. doi: 10.1086/143018. (Cited on page 3.)

I. Iben, Jr. Single and binary star evolution. *Astrophysical Journal, Supplement*, 76:55–114, May 1991. doi: 10.1086/191565. (Cited on page 27.)

I. Iben, Jr. and A. Renzini. Asymptotic giant branch evolution and beyond. *Annual Review of Astron and Astrophys*, 21:271–342, 1983. doi: 10.1146/annurev.aa.21.090183.001415. (Cited on page 26.)

O. Ilbert, S. Arnouts, H. J. McCracken, M. Bolzonella, E. Bertin, O. Le Fèvre, Y. Mellier, G. Zamorani, and R. Pellò. Accurate photometric redshifts for the CFHT legacy survey calibrated using the VIMOS VLT deep survey. *Astronomy and Astrophysics*, 457:841–856, October 2006. doi: 10.1051/0004-6361:20065138. (Cited on pages xviii, 54, 76, 79 and 80.)

O. Ilbert, P. Capak, M. Salvato, H. Aussel, H. J. McCracken, D. B. Sanders, N. Scoville, J. Kartaltepe, S. Arnouts, E. Le Floc'h, B. Mobasher, and Y. Taniguchi. Cosmos Photometric Redshifts with 30-Bands for 2-deg^2. *Astrophysical Journal*, 690:1236–1249, January 2009. doi: 10.1088/0004-637X/690/2/1236. (Cited on page 76.)

O. Ilbert, M. Salvato, E. Le Floc'h, H. Aussel, P. Capak, H. J. McCracken, B. Mobasher, J. Kartaltepe, N. Scoville, D. B. Sanders, S. Arnouts, K. Bundy, P. Cassata, J.-P. Kneib, A. Koekemoer, O. Le Fèvre, S. Lilly, J. Surace, Y. Taniguchi, L. Tasca, D. Thompson, L. Tresse, M. Zamojski, G. Zamorani, and E. Zucca. Galaxy Stellar Mass Assembly Between $0.2 < z < 2$ from the S-COSMOS Survey. *Astrophysical Journal*, 709:644–663, February 2010. doi: 10.1088/0004-637X/709/2/644. (Cited on pages xiii and 13.)

H. L. Johnson. *Interstellar Extinction*, page 167. the University of Chicago Press, January 1968. (Cited on page 44.)

P. Jonsson. SUNRISE: polychromatic dust radiative transfer in arbitrary geometries. *Monthly Notices of the RAS*, 372:2–20, October 2006. doi: 10.1111/j.1365-2966.2006.10884.x. (Cited on page 48.)

N. Katz, D. H. Weinberg, and L. Hernquist. Cosmological Simulations with TreeSPH. *Astrophysical Journal, Supplement*, 105:19, July 1996. doi: 10.1086/192305. (Cited on page 21.)

G. Kauffmann, et al., and . Stellar masses and star formation histories for 10^5 galaxies from the Sloan Digital Sky Survey. *Monthly Notices of the RAS*, 341:33–53, May 2003a. doi: 10.1046/j.1365-8711.2003.06291.x. (Cited on pages xiv, 14, 16, 54, 57, 62 and 91.)

G. Kauffmann, et al., and . The host galaxies of active galactic nuclei. *Monthly Notices of the RAS*, 346:1055–1077, December 2003b. doi: 10.1111/j.1365-2966.2003.07154.x. (Cited on pages xvi, 59, 61 and 92.)

A. Kawamura, et al., and . The Second Survey of the Molecular Clouds in the Large Magellanic Cloud by NANTEN. II. Star Formation. *Astrophysical Journal, Supplement*, 184:1–17, September 2009. doi: 10.1088/0067-0049/184/1/1. (Cited on page 60.)

R. C. Kennicutt, Jr. The Global Schmidt Law in Star-forming Galaxies. *Astrophysical Journal*, 498:541–+, May 1998. doi: 10.1086/305588. (Cited on page 20.)

L. J. Kewley and M. A. Dopita. Using Strong Lines to Estimate Abundances in Extragalactic H II Regions and Starburst Galaxies. *Astrophysical Journal, Supplement*, 142:35–52, September 2002. doi: 10.1086/341326. (Cited on page 53.)

J. Kim, C. Park, J. R. Gott, III, and J. Dubinski. The Horizon Run N-Body Simulation: Baryon Acoustic Oscillations and Topology of Large-scale Structure of the Universe. *Astrophysical Journal*, 701:1547, August 2009. doi: 10.1088/0004-637X/701/2/1547. (Cited on pages xxvii, 17 and 18.)

A. L. Kinney, D. Calzetti, E. Bica, and T. Storchi-Bergmann. The Reddening law outside the local group galaxies: The case of NGC 7552 and NGC 5236. *Astrophysical Journal*, 429:172–176, July 1994. doi: 10.1086/174309. (Cited on page 46.)

A. L. Kinney, D. Calzetti, R. C. Bohlin, K. McQuade, T. Storchi-Bergmann, and H. R. Schmitt. Template Ultraviolet to Near-Infrared Spectra of Star-forming Galaxies and Their Application to K-Corrections. *Astrophysical Journal*, 467:38, August 1996. doi: 10.1086/177583. (Cited on page 133.)

A. A. Klypin, S. Trujillo-Gomez, and J. Primack. Dark Matter Halos in the Standard Cosmological Model: Results from the Bolshoi Simulation. *Astrophysical Journal*, 740:102, October 2011. doi: 10.1088/0004-637X/740/2/102. (Cited on pages xxvii, 18 and 19.)

X. Kong, S. Charlot, J. Brinchmann, and S. M. Fall. Star formation history and dust content of galaxies drawn from ultraviolet surveys. *Monthly Notices of the RAS*, 349:769–778, April 2004. doi: 10.1111/j.1365-2966.2004.07556.x. (Cited on page 62.)

R. Kotulla, U. Fritze, P. Weilbacher, and P. Anders. GALEV evolutionary synthesis models - I. Code, input physics and web interface. *Monthly Notices of the RAS*, 396:462–484, June 2009. doi: 10.1111/j.1365-2966.2009.14717.x. (Cited on page 30.)

P. Kroupa. On the variation of the initial mass function. *Monthly Notices of the RAS*, 322:231–246, April 2001. doi: 10.1046/j.1365-8711.2001.04022.x. (Cited on pages xiv, 23 and 24.)

H. Krueger, U. Fritze-v. Alvensleben, and H.-H. Loose. Optical and near infrared spectral energy distributions. of blue compact galaxies from evolutionary synthesis. *Astronomy and Astrophysics*, 303:41, November 1995. (Cited on page 39.)

F. Lamareille, T. Contini, J.-F. Le Borgne, J. Brinchmann, S. Charlot, and J. Richard. Spectrophotometric properties of galaxies at intermediate redshifts (z ˜ 0.2-1.0). I. Sample description, photometric properties and spectral measurements. *Astronomy and Astrophysics*, 448:893–906, March 2006. doi: 10.1051/0004-6361:20053601. (Cited on page 54.)

J. M. Lamarre, J. L. Puget, F. Bouchet, P. A. R. Ade, A. Benoit, J. P. Bernard, J. Bock, P. de Bernardis, J. Charra, F. Couchot, J. Delabrouille, G. Efstathiou, M. Giard, G. Guyot, A. Lange, B. Maffei, A. Murphy, F. Pajot, M. Piat, I. Ristorcelli, D. Santos, R. Sudiwala, J. F. Sygnet, J. P. Torre, V. Yurchenko, and D. Yvon. The Planck High Frequency Instrument, a third generation CMB experiment, and a full sky submillimeter survey. *NewAstronomy Review*, 47:1017–1024, December 2003. doi: 10.1016/j.newar.2003.09.006. (Cited on pages 2 and 17.)

D. Le Borgne, B. Rocca-Volmerange, P. Prugniel, A. Lançon, M. Fioc, and C. Soubiran. Evolutionary synthesis of galaxies at high spectral resolution with the code PEGASE-HR. Metallicity and age tracers. *Astronomy and Astrophysics*, 425: 881–897, October 2004. doi: 10.1051/0004-6361:200400044. (Cited on page 30.)

J.-F. Le Borgne, G. Bruzual, R. Pelló, A. Lançon, B. Rocca-Volmerange, B. Sanahuja, D. Schaerer, C. Soubiran, and R. Vílchez-Gómez. STELIB: A library of stellar spectra at R ˜ 2000. *Astronomy and Astrophysics*, 402:433–442, May 2003. doi: 10.1051/0004-6361:20030243. (Cited on page 30.)

O. Le Fèvre, G. Vettolani, S. Paltani, L. Tresse, G. Zamorani, V. Le Brun, C. Moreau, D. Bottini, D. Maccagni, J. P. Picat, R. Scaramella, M. Scodeggio, A. Zanichelli, C. Adami, S. Arnouts, S. Bardelli, M. Bolzonella, A. Cappi, S. Charlot, T. Contini, S. Foucaud, P. Franzetti, B. Garilli, I. Gavignaud, L. Guzzo, O. Ilbert, A. Iovino, H. J. McCracken, D. Mancini, B. Marano, C. Marinoni, G. Mathez, A. Mazure, B. Meneux, R. Merighi, R. Pellò, A. Pollo, L. Pozzetti, M. Radovich, E. Zucca, M. Arnaboldi, M. Bondi, A. Bongiorno, G. Busarello, P. Ciliegi, L. Gregorini, Y. Mellier, P. Merluzzi, V. Ripepi, and D. Rizzo. The VIMOS VLT Deep Survey. Public release of 1599 redshifts to I_{AB} <24 across the Chandra Deep Field South. *Astronomy and Astrophysics*, 428:1043–1049, December 2004. doi: 10.1051/0004-6361:20048072. (Cited on page 12.)

O. Le Fèvre, G. Vettolani, B. Garilli, L. Tresse, D. Bottini, V. Le Brun, D. Maccagni, J. P. Picat, R. Scaramella, M. Scodeggio, A. Zanichelli, C. Adami, M. Arnaboldi,

S. Arnouts, S. Bardelli, and et al. The VIMOS VLT deep survey. First epoch VVDS-deep survey: 11 564 spectra with 17.5 $\leq$ IAB $\leq$ 24, and the redshift distribution over $0 \leq z \leq 5$. *Astronomy and Astrophysics*, 439:845–862, September 2005. doi: 10.1051/0004-6361:20041960. (Cited on pages 12 and 53.)

C. Leitherer and T. M. Heckman. Synthetic properties of starburst galaxies. *Astrophysical Journal, Supplement*, 96:9–38, January 1995. doi: 10.1086/192112. (Cited on page 22.)

C. Leitherer, D. Schaerer, J. D. Goldader, R. M. González Delgado, C. Robert, D. F. Kune, D. F. de Mello, D. Devost, and T. M. Heckman. Starburst99: Synthesis Models for Galaxies with Active Star Formation. *Astrophysical Journal, Supplement*, 123:3–40, July 1999. doi: 10.1086/313233. (Cited on page 30.)

S. J. Lilly, O. Le Fevre, F. Hammer, and D. Crampton. The Canada-France Redshift Survey: The Luminosity Density and Star Formation History of the Universe to Z approximately 1. *Astrophysical Journal, Letters*, 460:L1, March 1996. doi: 10.1086/309975. (Cited on pages xiii, 12 and 13.)

A. Loeb, A. Ferrara, and R. S. Ellis. *First Light in the Universe*. 2008. (Cited on page 38.)

P. Madau. Radiative transfer in a clumpy universe: The colors of high-redshift galaxies. *Astrophysical Journal*, 441:18–27, March 1995. doi: 10.1086/175332. (Cited on pages 49, 111 and 125.)

P. Madau, H. C. Ferguson, M. E. Dickinson, M. Giavalisco, C. C. Steidel, and A. Fruchter. High-redshift galaxies in the Hubble Deep Field: colour selection and star formation history to $z \sim 4$. *Monthly Notices of the RAS*, 283:1388–1404, December 1996. (Cited on pages xiii, 12 and 13.)

P. Madau, L. Pozzetti, and M. Dickinson. The Star Formation History of Field Galaxies. *Astrophysical Journal*, 498:106, May 1998. doi: 10.1086/305523. (Cited on pages xiii and 13.)

C. Maraston. Evolutionary synthesis of stellar populations: a modular tool. *Monthly Notices of the RAS*, 300:872–892, November 1998. doi: 10.1046/j.1365-8711.1998.01947.x. (Cited on pages 22 and 30.)

C. Maraston. Evolutionary population synthesis: models, analysis of the ingredients and application to high-z galaxies. *Monthly Notices of the RAS*, 362:799–825, September 2005. doi: 10.1111/j.1365-2966.2005.09270.x. (Cited on pages 30 and 53.)

C. Maraston and G. Strömbäck. Stellar population models at high spectral resolution. *Monthly Notices of the RAS*, 418:2785–2811, December 2011. doi: 10.1111/j.1365-2966.2011.19738.x. (Cited on page 30.)

P. Marigo. Chemical yields from low- and intermediate-mass stars: Model predictions and basic observational constraints. *Astronomy and Astrophysics*, 370: 194–217, April 2001. doi: 10.1051/0004-6361:20000247. (Cited on page 27.)

P. Marigo, L. Girardi, A. Bressan, M. A. T. Groenewegen, L. Silva, and G. L. Granato. Evolution of asymptotic giant branch stars. II. Optical to far-infrared isochrones with improved TP-AGB models. *Astronomy and Astrophysics*, 482: 883–905, May 2008. doi: 10.1051/0004-6361:20078467. (Cited on pages xiv, 31, 32 and 58.)

J. S. Mathis, P. G. Mezger, and N. Panagia. Interstellar radiation field and dust temperatures in the diffuse interstellar matter and in giant molecular clouds. *Astronomy and Astrophysics*, 128:212–229, November 1983. (Cited on page 49.)

C. F. McKee and J. P. Ostriker. A theory of the interstellar medium - Three components regulated by supernova explosions in an inhomogeneous substrate. *Astrophysical Journal*, 218:148–169, November 1977. doi: 10.1086/155667. (Cited on pages 33 and 35.)

A. McKellar. Evidence for the Molecular Origin of Some Hitherto Unidentified Interstellar Lines. *Publications of the ASP*, 52:187, June 1940. doi: 10.1086/125159. (Cited on page 33.)

G. R. Meurer, T. M. Heckman, and D. Calzetti. Dust Absorption and the Ultraviolet Luminosity Density at '$z \sim 3$ as Calibrated by Local Starburst Galaxies. *Astrophysical Journal*, 521:64–80, August 1999. doi: 10.1086/307523. (Cited on pages xxv, 127, 133, 135, 136 and 137.)

G. E. Miller and J. M. Scalo. The initial mass function and stellar birthrate in the solar neighborhood. *Astrophysical Journal, Supplement*, 41:513–547, November 1979. doi: 10.1086/190629. (Cited on pages xiv, 23 and 24.)

S. Noll and D. Pierini. Dust properties of UV bright galaxies at $z \sim 2$. *Astronomy and Astrophysics*, 444:137–155, December 2005. doi: 10.1051/0004-6361:20053635. (Cited on pages 44 and 48.)

S. Noll, D. Pierini, A. Cimatti, E. Daddi, J. D. Kurk, M. Bolzonella, P. Cassata, C. Halliday, M. Mignoli, L. Pozzetti, A. Renzini, S. Berta, M. Dickinson, A. Franceschini, G. Rodighiero, P. Rosati, and G. Zamorani. GMASS ultradeep spectroscopy of galaxies at $z \sim 2$. IV. The variety of dust populations. *Astronomy and Astrophysics*, 499:69–85, May 2009. doi: 10.1051/0004-6361/200811526. (Cited on page 48.)

R. W. O'connell. Galaxy spectral synthesis. I - Stellar populations in the nuclei of giant ellipticals. *Astrophysical Journal*, 206:370–390, June 1976. doi: 10.1086/154392. (Cited on page 22.)

A. Omont. Physics and chemistry of interstellar polycyclic aromatic molecules. *Astronomy and Astrophysics*, 164:159–178, August 1986. (Cited on page 35.)

D. E. Osterbrock and G. J. Ferland. *Astrophysics of gaseous nebulae and active galactic nuclei*. 2006. (Cited on pages xv, 37 and 39.)

B. Panter, R. Jimenez, A. F. Heavens, and S. Charlot. The star formation histories of galaxies in the Sloan Digital Sky Survey. *Monthly Notices of the RAS*, 378: 1550–1564, July 2007. doi: 10.1111/j.1365-2966.2007.11909.x. (Cited on pages 54 and 62.)

C. Papovich, L. A. Moustakas, M. Dickinson, E. Le Floc'h, G. H. Rieke, E. Daddi, D. M. Alexander, F. Bauer, W. N. Brandt, T. Dahlen, E. Egami, P. Eisenhardt, D. Elbaz, H. C. Ferguson, M. Giavalisco, R. A. Lucas, B. Mobasher, P. G. Pérez-González, A. Stutz, M. J. Rieke, and H. Yan. Spitzer Observations of Massive, Red Galaxies at High Redshift. *Astrophysical Journal*, 640:92–113, March 2006. doi: 10.1086/499915. (Cited on page 14.)

P. J. E. Peebles. Structure of the Coma Cluster of Galaxies. *Astronomical Journal*, 75:13, February 1970. doi: 10.1086/110933. (Cited on page 17.)

S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and The Supernova Cosmology Project. Measurements of Omega and Lambda from 42 High-Redshift Supernovae. *Astrophysical Journal*, 517:565–586, June 1999. doi: 10.1086/307221. (Cited on page 17.)

M. Pettini, D. L. King, L. J. Smith, and R. W. Hunstead. Dust in High-Redshift Galaxies. *Astrophysical Journal*, 478:536, March 1997. doi: 10.1086/303826. (Cited on page 108.)

A. J. Pickles. A Stellar Spectral Flux Library: 1150-25000 Å. *Publications of the ASP*, 110:863–878, July 1998. doi: 10.1086/316197. (Cited on page 30.)

D. Pierini, K. D. Gordon, A. N. Witt, and G. J. Madsen. Dust Attenuation in Late-Type Galaxies. I. Effects on Bulge and Disk Components. *Astrophysical Journal*, 617:1022–1046, December 2004. doi: 10.1086/425651. (Cited on pages 48 and 62.)

C. Pritchet. Stellar population synthesis of galactic nuclei. *Astrophysical Journal, Supplement*, 35:397–418, December 1977. doi: 10.1086/190485. (Cited on page 22.)

P. Prugniel and C. Soubiran. A database of high and medium-resolution stellar spectra. *Astronomy and Astrophysics*, 369:1048–1057, April 2001. doi: 10.1051/0004-6361:20010163. (Cited on page 30.)

N. Reddy, M. Dickinson, D. Elbaz, G. Morrison, M. Giavalisco, R. Ivison, C. Papovich, D. Scott, V. Buat, D. Burgarella, V. Charmandaris, E. Daddi, G. Magdis, E. Murphy, B. Altieri, H. Aussel, H. Dannerbauer, K. Dasyra, H. S. Hwang, J. Kartaltepe, R. Leiton, B. Magnelli, and P. Popesso. GOODS-Herschel Measurements of the Dust Attenuation of Typical Star-forming Galaxies at High Redshift: Observations of Ultraviolet-selected Galaxies at $z \sim 2$. *Astrophysical Journal*, 744: 154, January 2012. doi: 10.1088/0004-637X/744/2/154. (Cited on page 48.)

A. G. Riess, A. V. Filippenko, P. Challis, A. Clocchiatti, A. Diercks, P. M. Garnavich, R. L. Gilliland, C. J. Hogan, S. Jha, R. P. Kirshner, B. Leibundgut, M. M. Phillips, D. Reiss, B. P. Schmidt, R. A. Schommer, R. C. Smith, J. Spyromilio, C. Stubbs, N. B. Suntzeff, and J. Tonry. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *Astronomical Journal*, 116:1009–1038, September 1998. doi: 10.1086/300499. (Cited on pages 2 and 17.)

M. Rocha, P. Jonsson, J. R. Primack, and T. J. Cox. Dust attenuation in hydrodynamic simulations of spiral galaxies. *Monthly Notices of the RAS*, 383:1281–1291, February 2008. doi: 10.1111/j.1365-2966.2007.12431.x. (Cited on pages 48 and 62.)

S. Salim, S. Charlot, R. M. Rich, G. Kauffmann, T. M. Heckman, T. A. Barlow, L. Bianchi, Y.-I. Byun, J. Donas, K. Forster, P. G. Friedman, P. N. Jelinsky, Y.-W. Lee, B. F. Madore, R. F. Malina, D. C. Martin, B. Milliard, P. Morrissey, and et al. New Constraints on the Star Formation Histories and Dust Attenuation of Galaxies in the Local Universe from GALEX. *Astrophysical Journal, Letters*, 619:L39–L42, January 2005. doi: 10.1086/424800. (Cited on page 54.)

S. Salim, R. M. Rich, S. Charlot, J. Brinchmann, B. D. Johnson, D. Schiminovich, M. Seibert, R. Mallery, T. M. Heckman, K. Forster, P. G. Friedman, D. C. Martin, P. Morrissey, S. G. Neff, T. Small, T. K. Wyder, L. Bianchi, and et al. UV Star Formation Rates in the Local Universe. *Astrophysical Journal, Supplement*, 173: 267–292, December 2007. doi: 10.1086/519218. (Cited on pages 54 and 62.)

E. E. Salpeter. The Luminosity Function and Stellar Evolution. *Astrophysical Journal*, 121:161, January 1955. doi: 10.1086/145971. (Cited on pages xiv, 22, 23 and 24.)

P. Sánchez-Blázquez, R. F. Peletier, J. Jiménez-Vicente, N. Cardiel, A. J. Cenarro, J. Falcón-Barroso, J. Gorgas, S. Selam, and A. Vazdekis. Medium-resolution Isaac Newton Telescope library of empirical spectra. *Monthly Notices of the RAS*, 371:703–718, September 2006. doi: 10.1111/j.1365-2966.2006.10699.x. (Cited on pages 30, 31 and 58.)

J. Scalo. Fifty years of IMF variation: the intermediate-mass stars. In E. Corbelli, F. Palla, & H. Zinnecker, editor, *The Initial Mass Function 50 Years Later*, volume 327 of *Astrophysics and Space Science Library*, page 23, January 2005. (Cited on page 22.)

C. Scannapieco, M. Wadepuhl, O. H. Parry, J. F. Navarro, A. Jenkins, V. Springel, R. Teyssier, E. Carlson, H. M. P. Couchman, R. A. Crain, C. Dalla Vecchia, C. S. Frenk, C. Kobayashi, P. Monaco, G. Murante, T. Okamoto, T. Quinn, J. Schaye, G. S. Stinson, T. Theuns, J. Wadsley, S. D. M. White, and R. Woods. The Aquila comparison Project: The Effects of Feedback and Numerical Methods on Simulations of Galaxy Formation. *ArXiv e-prints*, December 2011. (Cited on page 21.)

D. Schaerer and S. de Barros. The impact of nebular emission on the ages of z≈ 6 galaxies. *Astronomy and Astrophysics*, 502:423–426, August 2009. doi: 10.1051/0004-6361/200911781. (Cited on page 39.)

D. Schaerer and S. de Barros. On the physical properties of z ≈ 6-8 galaxies. *Astronomy and Astrophysics*, 515:A73+, June 2010. doi: 10.1051/0004-6361/200913946. (Cited on page 54.)

N. Scoville, H. Aussel, M. Brusa, P. Capak, C. M. Carollo, M. Elvis, M. Giavalisco, L. Guzzo, G. Hasinger, C. Impey, J.-P. Kneib, O. LeFevre, S. J. Lilly, B. Mobasher, A. Renzini, R. M. Rich, D. B. Sanders, and et al. The Cosmic Evolution Survey (COSMOS): Overview. *Astrophysical Journal, Supplement*, 172: 1–8, September 2007. doi: 10.1086/516585. (Cited on page 53.)

D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R. Nolta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. *Astrophysical Journal, Supplement*, 148:175–194, September 2003. doi: 10.1086/377226. (Cited on pages 2 and 17.)

H. Spinrad and B. J. Taylor. The Stellar Content of the Nuclei of Nearby Galaxies. I. M31, M32, and M81. *Astrophysical Journal, Supplement*, 22:445, April 1971. doi: 10.1086/190232. (Cited on page 22.)

V. Springel. E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. *Monthly Notices of the RAS*, 401:791–851, January 2010. doi: 10.1111/j.1365-2966.2009.15715.x. (Cited on page 21.)

V. Springel, N. Yoshida, and S. D. M. White. GADGET: a code for collisionless and gasdynamical cosmological simulations. *NewAstronomy*, 6:79–117, April 2001. doi: 10.1016/S1384-1076(01)00042-2. (Cited on pages 20 and 21.)

V. Springel, S. D. M. White, A. Jenkins, C. S. Frenk, N. Yoshida, L. Gao, J. Navarro, R. Thacker, D. Croton, J. Helly, J. A. Peacock, S. Cole, P. Thomas, H. Couchman, A. Evrard, J. Colberg, and F. Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature*, 435:629–636, June 2005. doi: 10.1038/nature03597. (Cited on pages xxi, xxvii, 17, 18, 20, 55, 95 and 97.)

C. C. Steidel, M. Giavalisco, M. Pettini, M. Dickinson, and K. L. Adelberger. Spectroscopic Confirmation of a Population of Normal Star-forming Galaxies at Redshifts Z > 3. *Astrophysical Journal, Letters*, 462:L17, May 1996. doi: 10.1086/310029. (Not cited.)

I. Strateva, Ž. Ivezić, G. R. Knapp, V. K. Narayanan, M. A. Strauss, J. E. Gunn, R. H. Lupton, D. Schlegel, N. A. Bahcall, J. Brinkmann, R. J. Brunner, T. Budavári, I. Csabai, F. J. Castander, M. Doi, M. Fukugita, Z. Győry, M. Hamabe, G. Hennessy, T. Ichikawa, P. Z. Kunszt, D. Q. Lamb, T. A. McKay, S. Okamura, J. Racusin, M. Sekiguchi, D. P. Schneider, K. Shimasaku, and D. York. Color Separation of Galaxy Types in the Sloan Digital Sky Survey Imaging Data. *Astronomical Journal*, 122:1861–1874, October 2001. doi: 10.1086/323301. (Cited on page 3.)

R. Teyssier. Cosmological hydrodynamics with adaptive mesh refinement. A new high resolution code called RAMSES. *Astronomy and Astrophysics*, 385:337–364, April 2002. doi: 10.1051/0004-6361:20011817. (Cited on page 21.)

R. Teyssier, S. Pires, S. Prunet, D. Aubert, C. Pichon, A. Amara, K. Benabed, S. Colombi, A. Refregier, and J.-L. Starck. Full-sky weak-lensing simulation with 70 billion particles. *Astronomy and Astrophysics*, 497:335–341, April 2009. doi: 10.1051/0004-6361/200810657. (Cited on pages xxvii, 17 and 18.)

D. Thomas, C. Maraston, and J. Johansson. Flux-calibrated stellar population models of Lick absorption-line indices with variable element abundance ratios. *Monthly Notices of the RAS*, pages 49–+, January 2011. doi: 10.1111/j.1365-2966.2010.18049.x. (Cited on page 53.)

D. P. Todd. *A new astronomy*. 1906. (Cited on page 1.)

C. A. Tremonti, T. M. Heckman, G. Kauffmann, J. Brinchmann, S. Charlot, S. D. M. White, M. Seibert, E. W. Peng, D. J. Schlegel, A. Uomoto, M. Fukugita, and J. Brinkmann. The Origin of the Mass-Metallicity Relation: Insights from 53,000 Star-forming Galaxies in the Sloan Digital Sky Survey. *Astrophysical Journal*, 613:898–913, October 2004. doi: 10.1086/423264. (Cited on pages xxi, 54, 57, 92, 94, 96, 97, 98 and 101.)

R. J. Tuffs, C. C. Popescu, H. J. Völk, N. D. Kylafis, and M. A. Dopita. Modelling the spectral energy distribution of galaxies. III. Attenuation of stellar light in spiral galaxies. *Astronomy and Astrophysics*, 419:821–835, June 2004. doi: 10.1051/0004-6361:20035689. (Cited on pages 48 and 62.)

B. E. Turnrose. The stellar content of the nuclear regions of SC galaxies. *Astrophysical Journal*, 210:33–37, November 1976. doi: 10.1086/154801. (Cited on page 22.)

D. Tytler, D. Kirkman, J. M. O'Meara, N. Suzuki, A. Orin, D. Lubin, P. Paschos, T. Jena, W.-C. Lin, M. L. Norman, and A. Meiksin. Cosmological Parameters

σ_8, the Baryon Density Ω_b, the Vacuum Energy Density $\Omega_?$, the Hubble Constant and the UV Background Intensity from a Calibrated Measurement of H I Lyα Absorption at z=1.9. *Astrophysical Journal*, 617:1–28, December 2004. doi: 10.1086/425226. (Cited on page 49.)

P. G. van Dokkum, N. M. Förster Schreiber, M. Franx, E. Daddi, G. D. Illingworth, I. Labbé, A. Moorwood, H.-W. Rix, H. Röttgering, G. Rudnick, A. van der Wel, P. van der Werf, and L. van Starkenburg. Spectroscopic Confirmation of a Substantial Population of Luminous Red Galaxies at Redshifts $z \geq 2$. *Astrophysical Journal, Letters*, 587:L83–L87, April 2003. doi: 10.1086/375156. (Cited on page 48.)

E. Vassiliadis and P. R. Wood. Evolution of low- and intermediate-mass stars to the end of the asymptotic giant branch with mass loss. *Astrophysical Journal*, 413:641–657, August 1993. doi: 10.1086/173033. (Cited on page 26.)

A. Vazdekis. Evolutionary Stellar Population Synthesis at 2 Å Spectral Resolution. *Astrophysical Journal*, 513:224–241, March 1999. doi: 10.1086/306843. (Cited on page 30.)

A. Vazdekis, P. Sánchez-Blázquez, J. Falcón-Barroso, A. J. Cenarro, M. A. Beasley, N. Cardiel, J. Gorgas, and R. F. Peletier. Evolutionary stellar population synthesis with MILES - I. The base models and a new line index system. *Monthly Notices of the RAS*, 404:1639–1671, June 2010. doi: 10.1111/j.1365-2966.2010.16407.x. (Cited on pages 30 and 53.)

C. J. Walcher, P. Coelho, A. Gallazzi, and S. Charlot. Differential stellar population models: how to reliably measure [Fe/H] and [α/Fe] in galaxies. *Monthly Notices of the RAS*, 398:L44–L48, September 2009. doi: 10.1111/j.1745-3933.2009.00705.x. (Cited on page 53.)

S. M. Weinmann, F. C. van den Bosch, X. Yang, H. J. Mo, D. J. Croton, and B. Moore. Properties of galaxy groups in the Sloan Digital Sky Survey - II. Active galactic nucleus feedback and star formation truncation. *Monthly Notices of the RAS*, 372:1161–1174, November 2006. doi: 10.1111/j.1365-2966.2006.10932.x. (Cited on pages 21 and 57.)

P. Westera, T. Lejeune, R. Buser, F. Cuisinier, and G. Bruzual. A standard stellar library for evolutionary synthesis. III. Metallicity calibration. *Astronomy and Astrophysics*, 381:524–538, January 2002. doi: 10.1051/0004-6361:20011493. (Cited on page 30.)

S. D. M. White and C. S. Frenk. Galaxy formation through hierarchical clustering. *Astrophysical Journal*, 379:52–79, September 1991. doi: 10.1086/170483. (Cited on page 20.)

A. E. Whitford. The law of interstellar reddening. *Astronomical Journal*, 63:201–207, May 1958. doi: 10.1086/107725. (Cited on page 44.)

D. C. B. Whittet, editor. *Dust in the galactic environment*, 2003. (Cited on pages 42 and 125.)

V. Wild, G. Kauffmann, T. Heckman, S. Charlot, G. Lemson, J. Brinchmann, T. Reichard, and A. Pasquali. Bursty stellar populations and obscured active galactic nuclei in galaxy bulges. *Monthly Notices of the RAS*, 381:543–572, October 2007. doi: 10.1111/j.1365-2966.2007.12256.x. (Cited on pages 47 and 54.)

V. Wild, T. Budavári, J. Blaizot, C. J. Walcher, P. H. Johansson, G. Lemson, G. De Lucia, and S. Charlot. Quenching of Star Formation. In C. A. L. Bailer-Jones, editor, *American Institute of Physics Conference Series*, volume 1082 of *American Institute of Physics Conference Series*, pages 119–125, December 2008. doi: 10.1063/1.3059021. (Cited on page 56.)

V. Wild, S. Charlot, J. Brinchmann, T. Heckman, O. Vince, C. Pacifici, and J. Chevallard. Empirical determination of the shape of dust attenuation curves in star-forming galaxies. *Monthly Notices of the RAS*, 417:1760–1786, November 2011. doi: 10.1111/j.1365-2966.2011.19367.x. (Cited on pages 48 and 125.)

V. Wild, B. Groves, T. Heckman, P. Sonnentrucker, L. Armus, D. Schiminovich, B. Johnson, L. Martins, and S. Lamassa. Optical versus infrared studies of dusty galaxies and active galactic nuclei - I. Nebular emission lines. *Monthly Notices of the RAS*, 410:1593–1610, January 2011a. doi: 10.1111/j.1365-2966.2010.17536.x. (Not cited.)

S. M. Wilkins, A. J. Bunker, E. Stanway, S. Lorenzoni, and J. Caruana. The ultraviolet properties of star-forming galaxies - I. HST WFC3 observations of very high redshift galaxies. *Monthly Notices of the RAS*, 417:717–729, October 2011. doi: 10.1111/j.1365-2966.2011.19315.x. (Cited on page 133.)

S. Wuyts, I. Labbé, M. Franx, G. Rudnick, P. G. van Dokkum, G. G. Fazio, N. M. Förster Schreiber, J. Huang, A. F. M. Moorwood, H.-W. Rix, H. Röttgering, and P. van der Werf. What Do We Learn from IRAC Observations of Galaxies at $2 < z < 3.5$? *Astrophysical Journal*, 655:51–65, January 2007. doi: 10.1086/509708. (Cited on page 14.)

S. Wuyts, M. Franx, T. J. Cox, L. Hernquist, P. F. Hopkins, B. E. Robertson, and P. G. van Dokkum. Recovering Stellar Population Properties and Redshifts from Broadband Photometry of Simulated Galaxies: Lessons for SED Modeling. *Astrophysical Journal*, 696:348–369, May 2009. doi: 10.1088/0004-637X/696/1/348. (Cited on pages 14, 17, 54, 62, 73 and 75.)

D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, and et al. The Sloan Digital Sky Survey: Technical Summary. *Astronomical Journal*, 120:1579–1587, September 2000. doi: 10.1086/301513. (Cited on pages 5 and 53.)