Inferring the photometric and size evolution of galaxies from image simulations

Sébastien Carassou

To cite this version:

Sébastien Carassou. Inferring the photometric and size evolution of galaxies from image simulations. Astrophysics [astro-ph]. Université Pierre et Marie Curie - Paris VI, 2017. English. NNT : 2017PA066382 . tel-01773454

HAL Id: tel-01773454
https://theses.hal.science/tel-01773454v1

Submitted on 22 Apr 2018

Thèse présentée pour obtenir le grade de docteur de
l'Université Pierre et Marie Curie

Institut d'Astrophysique de Paris
École doctorale Astronomie et Astrophysique d'Île-de-France - ED127

Discipline : Astrophysique

Inferring the photometric and size evolution of galaxies from image simulations

Inférence de l'évolution photométrique et en taille des galaxies au moyen d'images simulées

---

Par : **Sébastien Carassou**

Sous la co-direction de :

Emmanuel Bertin

Valérie de Lapparent

Damien Le Borgne

Membres du jury:

Président du jury : Frédéric Daigne
Rapporteur : Ignacio Trujillo
Rapporteur : Olivier Bienaymé
Directrice de thèse : Valérie De Lapparent
Examinatrice : Laurence Tresse

Date de soutenance : 20 Octobre 2017

Remerciements

Celles et ceux qui me connaissent savent que cette thèse représente pour moi l'aboutissement d'un rêve d'enfant vieux de 13 ans. Et le moins qu'on puisse dire, c'est que ces trois dernières années ont été les plus riches, les plus intenses et les plus épanouissantes de toute ma vie. Cette période se termine comme elle a commencé: par une succession d'événements aussi improbables qu'extraordinaires, qui me mèneront je ne sais où. Et si je pars confiant, c'est que je me sais entouré des meilleurs amis que le multivers puisse générer. Toi qui lis ce message, sache que croiser ta ligne d'univers a été une chance et un privilège incroyable, que les quelques mots couchés ici ne suffiront pas à exprimer. Ce qui ne m'empêchera pas d'essayer ! Un immense merci donc, en vrac : à Charlène L., pour avoir déclenché il y a 3 ans cette série d'événements aussi improbables qu'extraordinaires. Je te dois une reconnaissance éternelle. À mes directeurs, Valérie, Emmanuel et Damien, pour m'avoir toujours soutenu, et pour avoir toujours su vous rendre disponibles pour répondre à mes questions, même les plus naïves. Pour votre patience aussi, envers mes nombreuses incompréhensions sur des sujets qui vous paraissaient couler de source. Et pour m'avoir laissé la liberté d'écrire la thèse dont j'avais rêvé. En prenant un peu de recul sur cette période de ma vie, je réalise à quel point vous m'avez changé[1]. Au delà du fonctionnement des galaxies, vous m'avez appris à être plus rigoureux, plus autonome, et plus exigeant envers moi même. J'ai beau quitter le monde universitaire pour une période indéterminée, ce sont des enseignements que je n'oublierai jamais. À Valérie en particulier, qui pendant ces 3 années a été beaucoup plus qu'une directrice de thèse. Merci d'avoir été une confidente, une amie, toujours ouverte et à l'écoute, même dans les moments les plus difficiles pour nous deux. Ça compte énormément pour moi. À mes rapporteurs, Olivier Bienaymé et Ignacio Trujillo, ainsi qu'aux membres de mon jury, Frédéric Daigne et Laurence Tresse, pour avoir pris le temps d'étudier mon travail en détail. À maman et papa, qui se sont tant sacrifiés pour que je n'aie jamais à me préoccuper d'autres choses que de mes études. Sans votre soutien inconditionnel et permanent dans mes choix de vie, je n'aurais jamais pu y arriver. À Alex, pour ton empathie inspirante et ta sagesse que j'envie beaucoup. Tu es la personne que j'aurais aimé être a ton âge, et je ne peux pas être plus fier de toi. À Cha, pour avoir su détourner mes yeux du cosmos assez longtemps pour que je me perde dans les tiens. *Dans l'immensité de l'espace et l'éternité du temps, c'est une joie pour moi de partager une planète et une époque avec toi.* Aux meilleurs colocs de ce recoin de la Galaxie: Guillaume C., Nicolas M. et Aurélien C. Vous avez minimisé ma souffrance et maximisé mon plaisir. Merci pour votre grande magnanimité envers mes négligences ménagères, et pour tous nos débats quotidiens, qui sont de ceux qui changent des vies, et qui en tout cas ont changé la mienne. À Etienne, pour notre bromance réussie et pour ton infinie patience envers tant de prises ratées et de rushs interminables. Sans toi, il n'y aurait moins de Sense Of Wonder dans ma vie. Aux savants fous du bureau 10, pour avoir été les plus géniaux et les plus improbables des compagnons de thèse dont on puisse rêver. Et en particulier:

- À Tilman H. d'être un role model pour nous tous et d'avoir fait de moi un homme meilleur. Quand je serai grand, je voudrais être comme toi.
- À Federico M. pour le grand voyage au sud, pour les OHPUTAIN et pour les étoiles dans tes yeux à la seule mention de "Pizza Popolare". Tu es et tu resteras le plus italien de mes amis.

---

[1] En mieux, j'entends.

- À Nicolas C. de continuer d'être l'une des personnes les plus brillantes et les plus dévouées que je connaisse (et je peux me vanter d'en connaître un paquet).

Mais aussi à Nicolas M. pour les combats de NERF et la partie Acknowledgement de ta these. Bureau 10, tu garderas à jamais une place dans mes zygomatiques. Aux merveilleux doctorants de l'IAP, et en particulier : À Alba V. D., pour les câlins d'encouragement, les discussions sur le Dixit et surtout pour CE moment avec la cuillère à soupe. À Caterina U., sans qui la vie à l'IAP sera beaucoup plus triste. À Julia G. d'alimenter les clichés sur les parisiennes. À Pierre F., Clément R. et Rebekka B., pour toutes les discussions (scientifiques ou non) aussi passionnées que passionnantes. À Siwei, pour les meilleurs falafels chinois au monde. À Florent L., pour ton aide précieuse pendant cette thèse et pour tes casse-têtes statistiques. Aux autres doctorants que j'aurais aimé prendre le temps de connaître plus (Florian, Céline, Gohar, Claire, Tanguy, Corentin, Oscar, Hugo...). À Erwan C., Guilhem L. et Jesse P., pour avoir donné un sens à mon travail grâce à vos discussions et à vos commentaires. À la famille de Cha, pour votre accueil chaleureux, vos incroyables tomates et vos soirées Quiplash. À la team Gato, pour les merveilleux moments (g)astronomiques, et tant d'autres folies. À Mathéo, pour ton accueil et ton enthousiasme. À Erwan C. (l'autre), pour tes constants encouragements et tous ces moments intimes dans des endroits exigus[2], À Christophe G., Pierre F., Maelle B., Claire P., Jordan B. et Victor E., pour les parties de loup garou. À la Potûre, pour les soirées, les potins, le lulz et les films. À Pierre Kerner… Parce que c'est Pierre Kerner. À Noémie de S. V. et à toute l'équipe de la Méthode Scientifique, pour votre légèreté et pour m'avoir permis de faire mes débuts radiophoniques dans la meilleure émission de science du service public. À Ali E. A. et Mathieu S., pour n'être pas morts à l'intérieur. À Andreas M., pour m'avoir sauvé la vie plusieurs fois quand je galérais sur du code. À Bruce B., Germain O., David L., Florence P., Renaud J., Arnaud T., Vince H. et Cyril P., pour m'avoir donné l'opportunité de partager ma passion au plus grand nombre. À Lucile G., pour m'avoir montré comment vaincre le Kraken et prouvé que la couleur des sentiments, c'est le bleu. À Cécile B., pour l'invention d'un nouveau genre cinématographique qui fera parler de lui, j'en suis sur. À Nicolas G., dont le potentiel à devenir le prochain Thomas Pesquet n'a d'égal que la portée de ta voix ! À Baur J. d'avoir été le parfait Jar Jar Binks. À Anaïs G., Nadia B., Alexis M. et Gabriel J., pour toutes les soirées à la coloc et sur les quais de Seine. À toute la team du collectif Conscience, pour tant de tolérance face à mon manque d'implication dans l'association ces dernières années. À Valentine D., en particulier, pour les plus exquis des selflous. À Oza et toute la team du café astroparisien, pour les discussions aussi profondes que nos pintes. Je garde un souvenir précieux des soirées dont je me souviens. À Phil Plait, le Bad Astronomer, pour avoir allumé sans le savoir la chandelle de la science dans ma nuit. À Carl Sagan, pour m'avoir si bien transmis ton sens de l'émerveillement et pour avoir donné un sens à ma vie. À André Brahic, passeur de sciences infatigable jusqu'au bout. À Hans Rosling et Jean Christophe Victor, pour avoir retiré mes oeillères sur le monde et m'avoir inspiré à agir pour le rendre meilleur. À David Grinspoon, Sean Carroll, Gérald Bronner, James Gleick et Geoffrey West (entre autres), qui ont su transformer mes trajets de RER en d'intenses questionnements sur le monde. À Douglas Adams, pour la vie, l'univers, et le reste. À Dounia S., pour ta parfaite maîtrise de l'art du sarcasme, et pour avoir été la parfaite compagne d'aventure lors des Utopiales. À Dominique M., le Meier des profs de physique. À Julien D. T., pour ta passion contagieuse de l'histoire et nos excursions archéologiques. À Johan M., pour avoir été le best guide of

---

[2] Je parle bien entendu d'escape rooms.

Baltimore ever. À Maeva R., Nolwenn G. et Marine D., pour ne pas m'avoir oublié après tant d'années. À Yann C. et Yana K., pour votre âme d'enfant et votre bienveillance infinie. À Riccardo N., pour les chansons éméchées sur la côte de Spetses. À Johann B. M. et Alexia M. B., pour tous ces moments toujours trop brefs, mais toujours intenses. À Christophe Michel… D'exister. À J.P. Goux, d'envoyer du rêve. À Léa J., en souvenir de ta jovialité qui transcende les continents, de Londres à Melbourne. À Louisa B., pour les réflexions sur l'infini, sur l'instant, et sur l'infini dans un instant. À Loki, Colas G. et Cyril P., d'être des artistes aussi talentueux qu'adorables. À Sandra B. pour ta personnalité au delà des tests et pour ton zèle à diffuser les sciences à la Réunion... Même si c'est pas simple. À Tabatha S. et Lucas G., d'être aussi adorables. À Léa B. et Tania L., pour les cuites et le théâtre (mais pas en même temps). À Margaux P. d'être une Dana Scully en vrai et pour ma première gueule de bois. À Théo D. d'être né en 1984. À Patrick B., Florence P. et Nans B., pour les bons moments du festival Imaginascience. Aux membres du Café des Sciences, qui en font plus pour la diffusion de la culture scientifique que beaucoup de médias traditionnels réunis. À Nanou, pour ta gentillesse infinie et tes questions aléatoires. Guillaume a bien de la chance. À Quentin B. F. de représenter l'incarnation française de Dale Cooper sur le Weird Wide Web. À Maxime L. et Maxime D., pour votre swag de requin et pour nous avoir guidé comme des chefs à travers les meilleurs coins de Nantes. À Jacquie A. et Michel B., pour les moments de procrastination malheureusement trop brefs. À Nat et Pilou et pis Lou, et pas seulement pour le jeu de mots. Aux frères Bazz d'être des piles nucléaires à la curiosité sans limite. À Thibault D., pour les imitations de Vincent Lagaf' (entre autres). À Emma H. et Alice pour votre positivité et votre énergie communicative au service de LA SCIEEENCE ! À l'imprimante copy1rv, la seule imprimante de l'IAP qui me comprenne vraiment. Aux abonnés et aux tipeurs, pour votre fidélité (et votre patience !) et pour me permettre de récolter des fruits de ma passion. Et à tous les gens que j'oublie forcément.

Abstract

Current constraints on the luminosity and size evolution of galaxies rely on catalogs extracted from multi-band surveys. However resulting catalogs are altered by selection effects difficult to model and that can lead to conflicting predictions if not taken into account properly. In this thesis we have developed a new approach to infer robust constraints on model parameters. We use an empirical model to generate a set of mock galaxies from physical parameters. These galaxies are passed through an image simulator emulating the instrumental characteristics of any survey and extracted in the same way as from observed data for direct comparison. The difference between mock and observed data is minimized via a sampling process based on adaptive Monte Carlo Markov Chain methods. Using mock data matching most of the properties of a Canada-France-Hawaii Telescope Legacy Survey Deep (CFHTLS Deep) field, we demonstrate the robustness and internal consistency of our approach by inferring the size and luminosity functions and their evolution parameters for realistic populations of galaxies. We compare our results with those obtained from the classical spectral energy distribution (SED) fitting method, and find that our pipeline infers the model parameters using only 3 filters and more accurately than SED fitting based on the same observables. We then apply our pipeline to a fraction of a real CFHTLS Deep field to constrain the same set of parameters in a way that is free from systematic biases. Finally, we highlight the potential of this technique in the context of future surveys and discuss its drawbacks.

Résumé

Les contraintes actuelles sur l'évolution en luminosité et en taille des galaxies dépendent de catalogues multi-bandes extraits de relevés d'imagerie. Mais ces catalogues sont altérés par des effets de sélection difficiles à modéliser et pouvant mener à des résultats contradictoires s'ils ne sont pas bien pris en compte. Dans cette thèse nous avons développé une nouvelle méthode pour inférer des contraintes robustes sur les modèles d'évolution des galaxies. Nous utilisons un modèle empirique générant une distribution de galaxies synthétiques à partir de paramètres physiques. Ces galaxies passent par un simulateur d'image émulant les propriétés instrumentales de n'importe quel relevé et sont extraites de la même façon que les données observées pour une comparaison directe. L'écart entre vraies et fausses données est minimisé via un échantillonnage basé sur des chaînes de Markov adaptatives. A partir de donnée synthétiques émulant les propriétés du Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) Deep, nous démontrons la cohérence interne de notre méthode en inférant les distributions de taille et de luminosité et leur évolution de plusieurs populations de galaxies. Nous comparons nos résultats à ceux obtenus par la méthode classique d'ajustement de la distribution spectrale d'énergie (SED) et trouvons que notre pipeline infère efficacement les paramètres du modèle en utilisant seulement 3 filtres, et ce plus précisément que par ajustement de la SED à partir des mêmes observables. Puis nous utilisons notre pipeline sur une fraction d'un champ du CFHTLS Deep pour contraindre ces mêmes paramètres. Enfin nous soulignons le potentiel et les limites de cette méthode.

Contents

Introduction

To see a world in a grain of sand
And a heaven in a wild flower,
Hold infinity in the palm of your hand,
And eternity in an hour.

William Blake, *Auguries of Innocence*

A long time ago in a galaxy far, far away

George Lucas, *Star Wars*

The study of galaxies confronts us to the largest scales imaginable. In order to grasp the processes at play in their evolution, astronomers must think in terms of billions of years and of millions of parsecs, which fills students of these beautiful stellar megalopolises with a deep sense of wonder and awe[3]. In only one century, through the collective effort of thousands of researchers using some of the greatest intellectual achievements in human history, we are barely beginning to develop a quantitative understanding of their evolution. These complex systems hide the details of their history within their structure, and disentangling the pieces of this giant cosmic puzzle requires a lot of time collecting photons from the distant Universe through the mechanical eyes of our best telescopes, as well as a lot of computing power to generate large simulated universes. Although we are currently standing on the threshold of a golden age of galaxy research, we are still far from explaining the precise distribution of shapes, colors and sizes of galaxies at various epochs of cosmic history.

Our understanding of galaxy evolution has been mainly shaped by the observation of the local and deep Universe by imaging surveys covering large cosmological volumes across the electromagnetic spectrum. Astronomers extract catalogs listing fluxes, effective sizes and shapes of detected sources from the multi-band images obtained in these surveys. Physical properties of different galaxy populations are then derived from these catalogs, and these "observed" physical properties are finally compared to the physical properties predicted by models of galaxy evolution. However, performing the comparison of observations and models in the "physical space" can lead to erroneous predictions, as extracted catalogs face a number of selection effects and observational biases that preferentially extinguish faint, low surface brightness and dusty galaxies below the survey limit, alter observed measurements of flux and size (hence the derived estimates of their distance) and can add a contribution to the noise of survey images. These selection effects (which are survey-dependent) have various origins, and depend on the position of the surveyed field on the celestial sphere, on the sensitivity of the survey, on the instruments used, on the wavelength range probed, on the atmospheric conditions (or lack thereof for space-based instruments), on the properties of dust in galaxies, on the spatial distribution of sources and on their distances.

Current catalogs include basic corrections for the predicted effect of these biases, implementing them in an analytical "selection function". However, as astronomers need ever more

---

[3] Although some of them are *really* good at hiding it...

precise measurements to perform quantitative assessments on the statistical evolution of galaxy populations, these selection functions represent oversimplifications of their true effect. A detailed correction routine for the effects of the selection biases affecting a given survey would require the complete modeling *a posteriori* of the full source extraction process, but such modeling is out of reach because of their non-linearities of these effects and the complex correlations among them.

However, it is relatively easier to simulate realistic galaxy survey images that reproduce the detailed characteristics of a given survey, such as the exposure time, the filters used, the optics of the telescope, the properties of the detector, the Point Spread Function, the small-scale and large-scale noise properties on the image, the amount of dust extinction by the Milky Way in the area of the sky reproduced... In fact many realistic image simulators have already been developed for testing the analysis pipeline of current and future galaxy surveys. Provided that the synthetic images are realistic enough, if sources are extracted from the simulated images in the same way as they are on real images, then both resulting catalogs face the same selection effects and measurement biases and therefore can be compared directly, in the "observable space", in a way that is free from systematic biases. This approach is generally referred to as "forward modeling". If synthetic and observed data can be compared directly, their discrepancy can be minimized if the right set of input model parameters are used. The set of possible parameter values form a multidimensional parameter space that needs to be explored efficiently in order to find the parameters that generate the most realistic distributions of mock data.

During this thesis, we have implemented an approach that couples a forward model to compare simulated and observed distributions of galaxy fluxes and radii from a given survey to a statistical inference pipeline to retrieve the sets of input parameters that most closely reproduce the observed data. We have also tested the potential of this technique using mock images with a set of galaxy populations generated by our model as input data. This thesis is divided into four thematic parts, organized as follows:

Part I summarizes what astronomers can learn from the wealth of data provided by galaxy surveys. Chapter 1 provides a brief review of the cosmological context and the current status of the field of galaxy evolution, as well as a detailed description of the observed properties and supposed formation processes of the main structural components of galaxies. In Chapter 2, we list some major galaxy surveys probing the local and deep universe, and summarize what their analysis has revealed on the luminosity and size distribution of different galaxy populations, and the evolution of these quantities over cosmic times. In particular we focus on the debated question of the size evolution of massive elliptical galaxies, whose estimation is particularly sensitive to selection effects. In Chapter 3, we address the issue of selection effects and observational biases affecting the photometric and morphometric catalogs extracted from galaxy surveys, and detail some of the efforts that have been made to minimize or correct for their effects.

Part II introduces the forward modeling approach for understanding galaxy evolution. In Chapter 4, we provide an outline of the main steps of our method. In particular, we describe how the radial light distribution of galaxies and their dust content can be modeled, a well as the inner workings of the empirical model that we use in this work to test the robustness of our method. In Chapter 5, we provide a detailed description of the image simulation and source extraction softwares.

Part III discusses the key notions and methods used for the statistical inference of model parameters. In Chapter 6, we introduce the general framework in which this inference is performed: the *Bayesian probability theory*, and more specifically the *parametric Bayesian Indirect Likelihood* (pBIL) framework. We also present the Markov Chain Monte Carlo methods for sampling the parameter space, as well as several refinements built upon them that we use to perform this exploration in an efficient way in terms of computing time. Different techniques for quantifying the discrepancy between observed and simulated datasets are presented in Chapter 7. Our approach is based on the comparison of binned observables, that are first preprocessed to reduce their dynamic range and maximize the information content they display in the multidimensional observable space. In Chapter 8, we review a set of techniques commonly used to assess whether the Markov chains have sampled the parameter space efficiently and reached the regions that minimize the difference between the synthetic and observed datasets, as well as some tools to visualize and interpret the results of Bayesian inference.

In Part IV, the robustness, accuracy and self-consistency of our pipeline are tested. In Chapter 9, we perform early tests in idealized situations by inferring the luminosity and size distributions (and evolution with redshift) of two populations of galaxies from a mock image reproducing the characteristics of a Canada-France-Hawaii Telescope Legacy Survey (CFHTLS) Deep field covering one square degree in the sky in two optical bands and one infrared band, and we compare the results of our approach to the results of a more classical (although less computer-intensive) approach, i.e. the use of photometric redshifts based on SED fitting using the same set of observables. We then improve the reliability of our pipeline by upgrading the realism of our synthetic images as well as the efficiency of our sampling scheme in Chapter 10. We finally pave the way for an application to real data in Chapter 11 by performing a test of the updated pipeline in more realistic conditions, using a mock CFHTLS Deep image with 3 galaxy populations approximately reproducing observed number counts, covering ~0.1 square degree in the sky in the same 3 photometric passbands as in previous tests. In Chapter 12, we provide a summary of our results, we discuss potential improvements to our approach and detail its inherent limitations.

Part I

Mining the sky

Chapter 1

The Big Picture

> *Through our eyes, the universe is perceiving itself. Through our ears, the universe is listening to its harmonies. We are the witnesses through which the universe becomes conscious of its glory, of its magnificence.*
>
> Alan Watts

Abstract

Galaxies are incredibly complex and fascinating objects. These gravitationally bound structures composed of dark matter, stars, gas, and dust display a remarkable structural diversity, and the story of their formation and evolution is deeply interconnected with that of the universe itself. In this chapter, I summarize what the last century of research has taught us about the statistical properties of galaxy populations in their cosmological context and their evolution across cosmic times, and I provide an overview of the challenges that lie ahead in the development of a complete theory of galaxy evolution.

Contents

1.1 Cosmological context

Our cosmological framework, the ΛCDM model, or *Concordance model* (Spergel et al. 2003), sets the scene for the formation and evolution of galaxies. This model describes a homogeneous and isotropic expanding universe (at large scales), whose expansion has been accelerating for the past 5 billion years due to a mysterious *dark energy* (the Λ in ΛCDM, Riess et al. 1998; Perlmutter et al. 1999) and whose matter content is largely dominated by an invisible "cold" *dark matter* (the "CDM" in ΛCDM) interacting with baryonic matter (i.e. gas, stars...) only via gravity. Dark matter is supposed to be cold because it is best described in numerical simulations as a collection of non-relativistic (i.e. slow moving) particles.

The ΛCDM model has been able to predict the distribution of large-scale structures with staggering precision (e.g. Tegmark et al. 2004; Samushia et al. 2014) and is now considered standard by the majority of the astronomical community. This framework allows us to build the most probable chronological history of the cosmos consistent with current observational constraints, from the Big Bang to the present universe. This story and its foundations are told in the following sections.

1.1.1 Recombination

The story of galaxy evolution begins at $z \sim 1100$, only 380,000 years after the Big Bang. Up to that time, the universe was a dense and extremely hot plasma of electrons, photons and atomic nuclei of hydrogen and helium resulting from primordial nucleosynthesis (at t = Big Bang + 3 min). But at $z \sim 1100$, the expanding cosmos was large enough to reach a temperature of 3000K, lower than the ionization energy of typical atoms. Below that temperature, the universe was cold enough so that ions and electrons could combine together to form the first neutral atoms. This epoch of cosmic history is referred to as cosmological recombination[1].

The universe, then opaque, became "suddenly"[2] transparent, and photons, finally free from the incessant scattering by electrons, decoupled from the baryonic matter and filled the universe with a blackbody radiation we now observe as the Cosmic Microwave Background (CMB) radiation[3] (Gamow 1948; Penzias and Wilson 1965). The study of the CMB temperature anisotropies (of order 10^{-5}K, see Figure 1-1), a relic of this distant past, reveals that at the time of recombination, the universe was remarkably homogeneous and isotropic. Over this vast and uniform sea of hydrogen and helium, slightly overdense regions could be seen, resulting from quantum fluctuations amplified by primordial inflation (e.g. Kolb et al. 1990). These tiny overdensities were the seeds from which galaxies and other cosmic structures were grown.

---

[1] "Recombination" is a rather odd name when you realize that electrons and atomic nuclei were never actually combined before.

[2] Recombination was fast, but not instantaneous. The universe went from completely ionized to neutral over a range of redshifts $\Delta z \sim 200$ (Kinney 2003).

[3] The CMB has since cooled down to its current temperature of 2.7K (Spergel et al. 2003).

1.1.2 The Dark Ages

Over time, due to the expansion and cooling of the universe, the CMB, once a reddish glow of blackbody radiation, ended up emitting most of its energy in the infrared. The universe would then appear completely dark to the human eye, leading to its "Dark Ages"[4]. No source of light had yet been formed and even if they did, their emitted UV radiation would have been absorbed by the neutral gas of the intergalactic medium. This period of cosmic history has yet to be revealed by observations, and probing this era would potentially allow to constrain the shape of the matter power spectrum on scales smaller than the CMB fluctuations (Loeb and Zaldarriaga 2004). During the vast majority of the Dark Ages, the distribution of atomic gas was still close to homogeneous, and the growth of the fluctuations of baryonic density likely remained linear with time. (Furlanetto et al. 2006).

Through gravitational instabilities, regions of high densities collapsed hierarchically, first on small scales and then on large scales. Dark matter structures formed first, because their gravity was not counteracted by any form of internal pressure. Following the gravitational potential wells of dark matter halos, gas clouds began to collapse as they went beyond Jeans mass ($\sim 10^4$ $M_\odot$), the mass above which their gravity could overcome gas pressure (Loeb and Barkana 2001).

1.1.3 The Epoch of Reionization

The dark ages ended a few hundred million years after the Big Bang with the birth of the first generation of luminous objects, such as what are believed to be massive Population III stars, galaxies and quasars, photoionizing the intergalactic medium (IGM) with their emitted UV light, making it transparent to these wavelengths: this period is known as the Epoch of Reionization (EoR, see Figure 1-2). The EoR probably happened in two phases: a slow phase, where the UV emitting sources were first isolated in their Strömgren spheres. These bubbles of ionizing radiation then expanded and overlapped, resulting in a second, much faster reionization phase (Loeb and Barkana 2001).

The dominant sources and redshift range of the EoR are still ill-constrained. Studies of the large-scale polarization of the CMB (Ade et al. 2016) suggest that the EoR might have started as early as $z \sim 14$, and studies of the spectra of high redshift quasars suggest that the universe is fully reionized by $z \sim 6$ (Fan et al. 2006). The EoR ended when almost all the atoms of the IGM ended up ionized. The characteristics of this epoch are important to unveil, as the metal enrichment and supernovae shocks created by feedback from the first generation of stars set the initial conditions for the first galaxies to form (Bromm and Yoshida 2011). This is the goal of the next generation of telescopes, such as the James Webb Space Telescope or the Square Kilometer Array.

1.1.4 The present universe

After 13.8 billion years of evolution, the present universe is a clumpy mess of galaxies, galaxy clusters and superclusters embedded in a cosmic web of sheets, filaments, and halos of gas and dark matter (cf Figure 1-3), whose dynamics at very large scales is dominated by dark

---

[4] The term "Dark Ages" was coined in the cosmological context by Sir Martin Rees, according to Loeb and Barkana (2001).

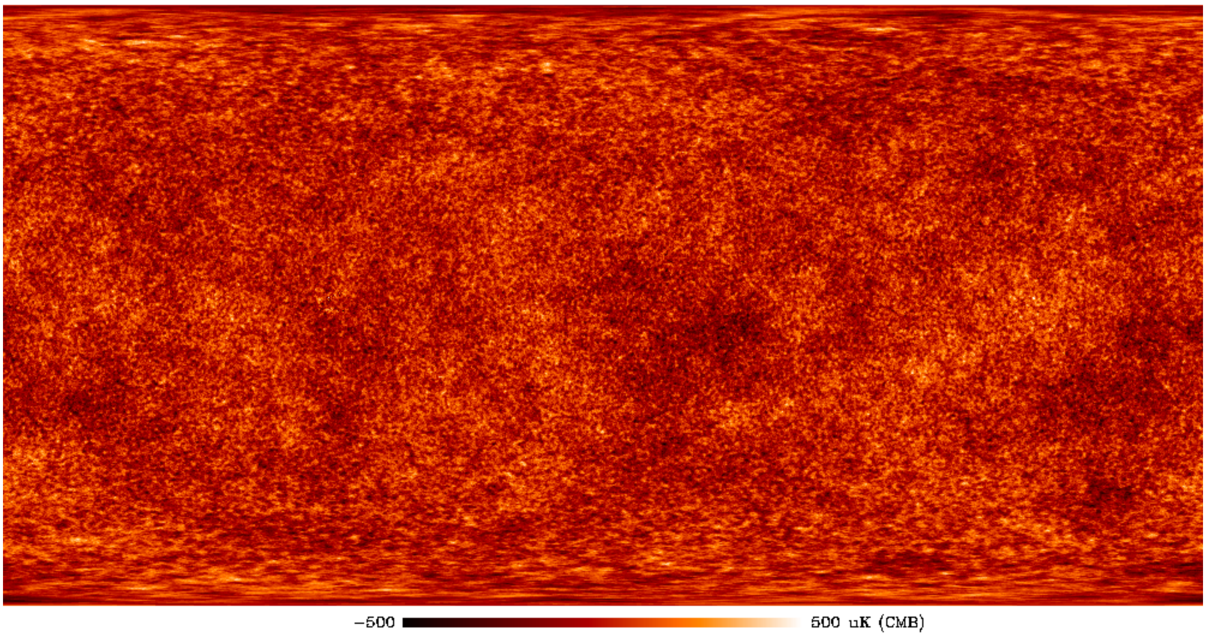


Figure 1-1 – Colorized map of the Cosmic Microwave Background (CMB) in equirectangular projection, as revealed by the Planck satellite (Adam et al. 2016). Credits: ESA and the Planck collaboration.

energy. Only about 10% of the baryonic mass of the universe is currently contained inside galaxies. The other 90% resides in the warm and hot IGM (Read and Trentham 2005). The most recent observations of the CMB are compatible with a spatially flat universe expanding at an accelerating rate that can be described with only 6 parameters: $\Omega_m = 0.31, \Omega_b = 0.05, \Omega_\Lambda = 0.69, H_0 = 100h\text{km s}^{-1}\text{Mpc}^{-1}$ (with $h = 0.68$), $n_s = 0.96$ and $\sigma_8 = 0.8$ (Ade et al. 2016).

1.2 A century of research

The field of galaxy evolution has a rich and fascinating history, characterized by a series of breakthroughs over the last century. Technological breakthroughs of course, as our telescopes have not ceased to grow in size, and our computers in power, but also and most importantly, philosophical breakthroughs. In less than 100 years, our vision of the cosmos has dramatically shifted from the comforting sight of the Milky Way to an ever-expanding universe containing trillions of galaxies. This section is an homage to the work of the great men and women who dedicated their lives to decipher the cosmos, and whose shoulders I stand on. As I am not a science historian, my attempt at summarizing a full century of a field of research will obviously be as incomplete as my knowledge on the matter. The curious reader will therefore have the pleasure to read the more exhaustive review of D'Onofrio et al. (2016) from which this section is inspired.

In 1920, Heber Curtis and Harlow Shapley triggered a fierce debate at the Smithsonian Institute in Washington D.C., the aptly named "Great Debate". The nature of this debate was the

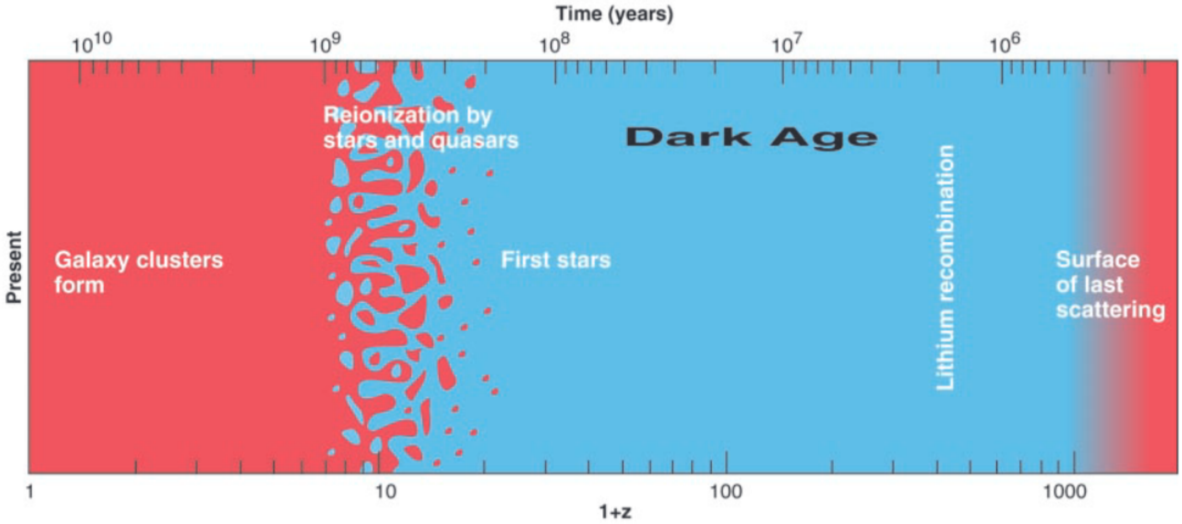


Figure 1-2 – The logarithmic evolution of the universe from $z \sim 1100$ to $z = 0$, according to the hierarchical scenario of structure formation. From Miralda-Escudé (2003).

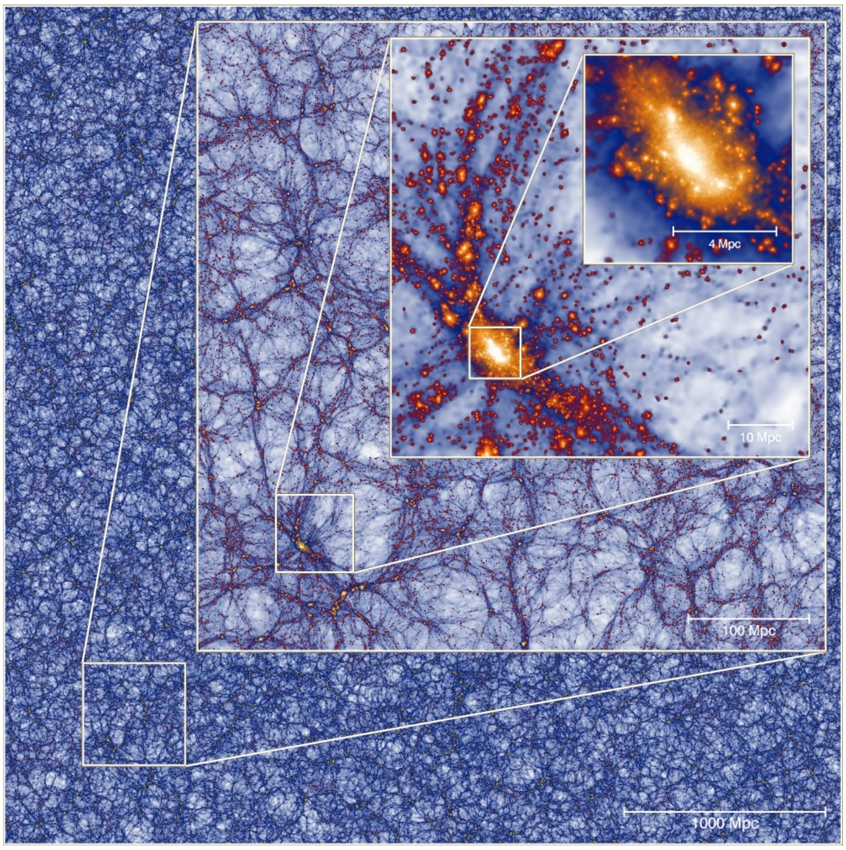


Figure 1-3 – The cosmic web at $z = 0$ from the Millennium-XXL simulation (Angulo et al. 2012). The yellow regions are the sites of galaxy formation.

following question: are the spiral nebulae we see in the sky part of the Galaxy or extragalactic objects? Shapley was in favor of the “Big Galaxy” hypothesis, and Curtis was in favor of the latter hypothesis. The discovery of Cepheid stars in M31 by Hubble in 1923, which were used as a distance measure[5], (almost) settled the debate once and for all: the spiral nebulae were indeed separate "island universes" (Kant 1755), distinct from our own Milky Way. Thanks to Vesto Slipher’s measurements of galaxy redshifts, Hubble (1929) discovered an important correlation between redshift and distance, in what is now called “Hubble’s law”: galaxies recede faster from us as their distance increases. This observation provided the first evidence that we live in an expanding universe.

While Friedman, Robertson, Lemaitre and Walker were independently shaping our standard model of modern cosmology by solving Einstein’s equations of general relativity, Hubble (1926, 1936) devised a system of morphological classification of galaxies according to their visual appearance. The Hubble sequence, or “tuning fork” diagram, divided galaxies into ellipticals (E), lenticulars (S0) and spirals (S), and dividing spirals into barred spirals (SB) and non-barred spirals (S) (cf Figure 1-4). Galaxies that did not fit into this scheme were termed Irregulars (Irr) and peculiars (P). The Hubble sequence is now considered simplistic, and several refinements and revisions of this classification have been proposed since[6] (e.g., Morgan 1958; De Vaucouleurs 1959; Sandage 1961; De Vaucouleurs et al. 1991), but Hubble’s nomenclature still holds to this day. Ellipticals and lenticulars are often referred to as “early-types”, while spirals and irregulars constitute the “late-type” class, as a reference to the early interpretation of the Hubble diagram as an evolutionary sequence. Moreover, galaxies from the left to the right of the spiral sequence are sometimes referred to as “early-type” to “late-type spirals”. The spectral classification first proposed by Morgan and Mayall (1957) provided a good alternative to morphological classification, and spectral types appeared to correlate well with morphological types. Spirals and ellipticals were soon found to exhibit different properties : Holmberg (1958) observed that ellipticals are typically massive and red, and show little star formation, whereas spirals tend to be less massive, bluer and display evidence of ongoing star formation. A first probable causal link between spirals and ellipticals was found by the numerical simulations of Toomre and Toomre (1972): elliptical galaxies could be the end-products of merging spiral galaxies.

Through the study of the velocity dispersion of 8 galaxies of the Coma cluster, Zwicky (1933) remarked that the density of the Coma cluster inferred from its luminous distribution was 400 times smaller than the value inferred from the velocity dispersion. Zwicky’s "missing mass problem" was later confirmed by Rubin and Ford Jr (1970) thanks to the analysis of the rotation curves of spiral galaxies. The nature and properties of this “dark matter” which holds galaxies and galaxy clusters together is one of the greatest unsolved mystery of this century.

For a long time, two visions opposed to explain the formation of galaxies, in what was basically a “nature vs. nurture” debate of cosmic proportions: the monolithic collapse scenario (Eggen et al. 1962) and the hierarchical scenario (White and Rees 1978; Fall and Efstathiou 1980; Blumenthal et al. 1984; White and Frenk 1991; Cole et al. 2000). In the monolithic collapse scenario, galaxies come from the gravitational collapse of a cloud of primordial gas in the early universe, and different galaxy types are born fundamentally different and evolve isolated

---

[5] Cepheid stars are variable stars that exhibit a relation between their luminosity and their period of pulsation, as discovered by American astronomer Henrietta Leavitt, which makes them useful as standard candles to measure extragalactic distances.

[6] For a complete historical review of galaxy morphology classification systems, see Buta (2013).

from their environment. In the hierarchical scenario, galaxies are gradually assembled through successive mergers of smaller structures, and the morphological mix depends on individual merger histories. The model of White and Rees has since become the dominant paradigm for galaxy formation, thanks to, among other things, the observation of increasing merger rates in the high redshift universe (e.g., van Dokkum et al. 1999). In this scenario, galaxy formation is a two-stage process: dark matter halos interacting only via gravity first form hierarchically though mergers, and galaxies then form by cooling and collapse of baryons in the gravitational potential wells of their host dark matter halos (cf Figure 1-8).

The mapping of 2,400 galaxies by the CfA Redshift Survey (Huchra et al. 1983), the first systematic survey of this kind, revealed that the spatial distribution of galaxies was far from random: at the scale of several tens of Mpcs, they were lying on the "surfaces of bubble-like structures" surrounding vast cosmic voids (De Lapparent et al. 1986). The invention of charge coupled devices (CCDs) marked the beginning of the golden age of extragalactic astronomy in the mid-1980s. Astronomical surveys such as the Sloan Digital Sky Survey (York et al. 2000), the first fully digital map of the local universe, provided astronomers with an unprecedented wealth of data that only recently allowed to make reliable, quantitative statements on the distribution of different galaxy populations. The morphology of hundreds of thousands of galaxies was quickly becoming accessible, either by automated classification (e.g. Nair and Abraham 2010), by visual classification from citizen scientists around the world (Lintott et al. 2010), or by detailed visual classification from groups of experts on smaller samples (Baillard et al. 2011).

Galaxies in the local universe were found to be separated into two robust classes, revealed in their color-magnitude distributions: a “blue cloud” of star-forming late-type galaxies, and a “red sequence” of quiescent ellipticals and lenticulars (e.g. Kauffmann et al. 2003; Baldry et al. 2004, cf Figure 1-5). The morphology of galaxies was also linked to their environment: elliptical galaxies were preferably found in regions of high density, such as the core of clusters, whereas spiral galaxies were the dominant population in regions of low-density (Dressler 1980; Goto et al. 2003, cf Figure 1-5).

With its Deep Field (Williams et al. 1996) and Ultra Deep field (Beckwith et al. 2006), the Hubble Space Telescope opened a small window on the early universe, and paved the way for our understanding of galaxy evolution across cosmic times. The high-redshift universe turned out to be a much more active place than the local universe, with a higher fraction of irregular and peculiar types (Abraham et al. 1996, cf Figure 1-6). The analysis of a collection of subsequent deep UV, IR and radio galaxy surveys lead to the observation that the cosmic star formation rate peaked around $z \sim 2$ (an epoch often referred to as the *cosmic high noon*), with stars forming at a rate of $\sim$ 10 times higher than what is observed today, and this rate has been decreasing exponentially ever since (Lilly et al. 1995; Madau et al. 1996; see Madau and Dickinson 2014 for a recent review). Multiwavelengths surveys from ground-based and space-based instruments over the past two decades shaped our understanding of the evolution of the cosmos at all scales, with the light, shape and spectrum of millions of galaxies probed across most of the electromagnetic spectrum. Current and future surveys, such as the Dark Energy Survey (DES, Collaboration et al. 2005), Euclid (Laureijs et al. 2011), and the Large Synoptic Survey (LSST, Ivezic et al. 2008) might bring us new insights about galaxy clustering, as well as the nature of dark matter and dark energy. And with the Square Kilometer Array (SKA, Carilli and Rawlings 2004), which is expected to generate 1 exabyte (10^9 GB) of data every day after completion in 2024, radioastronomy is about to enter the “exascale era”, which presents new technological challenges for data storage, processing and analysis. Astronomy has come a long way since Messier's (1781) catalog of 110 fuzzy objects in the sky.

In parallel with the development of astronomical surveys, cosmological simulations allowed us to model the non-linear dynamics of large-scale structures and their interplay with the small scale physics of galaxy evolution with ever-increasing complexity. Starting from the tiny density fluctuations of the early universe observed in the CMB (now tightly constrained thanks to the results of the WMAP and Planck satellites), these simulations replayed the history of the universe, according to our best theoretical models. The numerical simulation method in the field of galaxy evolution dates back to the first N-body simulations of interacting galaxies of Holmberg (1941) performed on an analog computer[7], and its use has become ubiquitous since the late 1970's, with the pioneering simulation of Miyoshi and Kihara (1975) performed on 400 "galaxies" (particles) in an expanding universe[8], achieving the prediction of the spatial distribution of the large-scale structures with remarkable accuracy, and ushering us in the era of "precision cosmology" (Primack 2005).

For a long time, large-scale cosmological simulations such as the Millenium run (Springel et al. 2005) or the Bolshoi simulation (Klypin et al. 2011) were limited to N-body simulations involving only dark matter, and therefore could not trace the evolution of baryonic matter (i.e. gas and stars), due to computational limitations. Only recently does computational power (along with more efficient numerical methods) allow for the complexity of fluid dynamics to enter the picture, and to simulate statistically significant populations of galaxies. Following Moore's law[9], cosmological simulations have been doubling in size every $\sim$ 16.5 months (Springel et al. 2005). The last generation of cosmological hydrodynamic simulations such as Illustris (Vogelsberger et al. 2014), EAGLE (Crain et al. 2015) and Horizon-AGN (Dubois et al. 2014) are finally able to simulate cosmological volumes (boxes of hundreds of Mpc in comoving size) with enough precision to resolve the internal structure of individual galaxies (of kpc size). But up to now the differences in numerical implementations of unresolved ("sub-grid") physical processes make the comparison of the results of these simulations difficult (Scannapieco et al. 2012). The direct comparison between cosmological hydrodynamical simulations and observations has also been seen as a hazardous undertaking, as they trace different physical phenomena (light for astronomical surveys and mass for simulations). The advent of synthetic observatories built from the results of simulations, mimicking the effects of telescope resolution and noise, such as, e.g., the work of Snyder et al. (2015) on the Illustris simulation (cf Figure 1-7), might have a role to play to better estimate the reliability of simulations.

With the rise of digital surveys and supercomputers, astronomers have definitely turned their backs on the eyepieces of telescopes and switched to computer screens. The data acquired is so vast that fascinating discoveries await in our hard drives. We are still learning.

1.3 What is a galaxy?

According to the physically-motivated definition of Willman and Strader (2012), galaxies are gravitationally bound collections of stars whose properties cannot be explained by a combina-

---

[7] Holmberg's ingenious experiment involved 37 mass points represented by light bulbs, and the inverse square behavior of light was used as an analog of gravity.

[8] A good historical introduction to the early evolution of cosmological simulations can be found in Suto (2005).

[9] Moore's law is the empirical assertion that the number of transistors in integrated circuits doubles approximately every two years (Moore 2006).

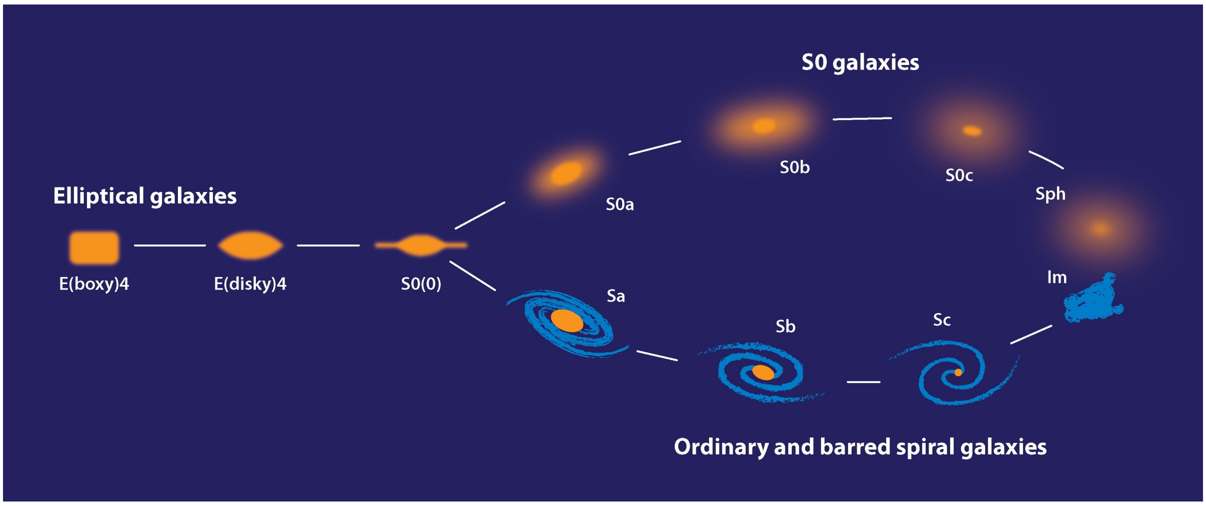


Figure 1-4 – A modern form of the Hubble sequence, extended from the original version to include dwarf spheroidal galaxies and irregular galaxies. From Kormendy and Bender (2011).

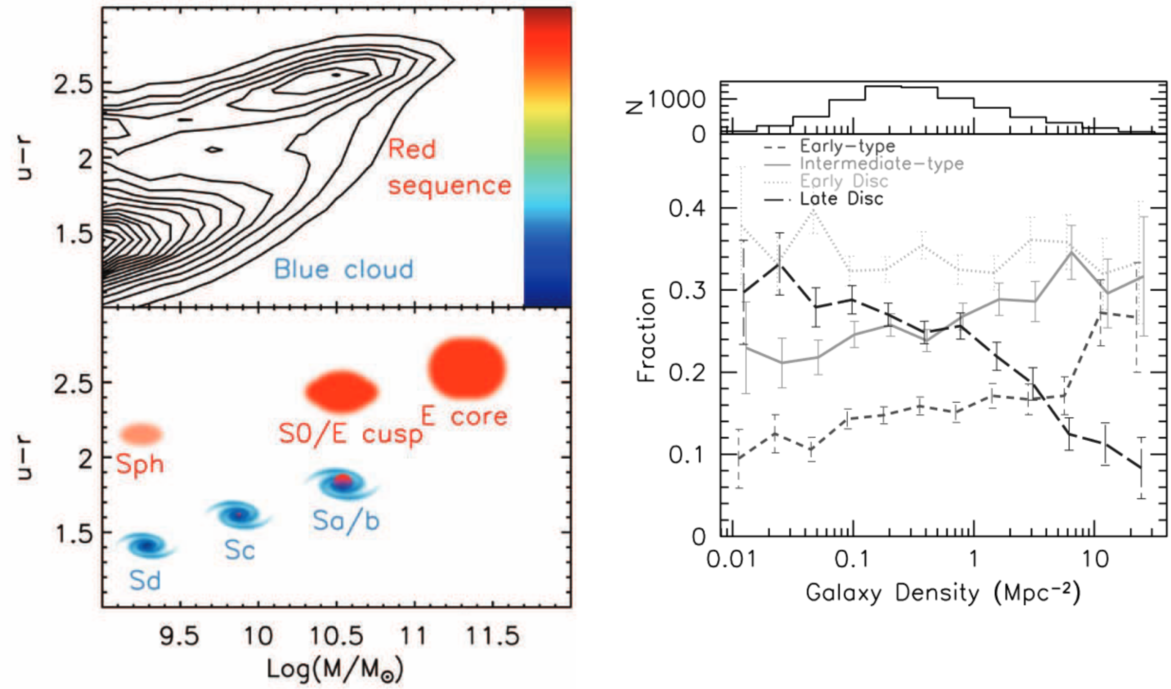


Figure 1-5 – *Left*: The color bimodality of galaxy populations, as revealed in a diagram representing $(u - r)$ SDSS colors as a function of baryonic mass, using the results of Baldry et al. (2004) on SDSS data. The bottom panel shows the morphological types that dominate in different regions of the top panel (from Kormendy 2013). *Right*: Morphology-density relation for a sample of 7,938 SDSS galaxies from $0.05 < z < 0.1$ and $M_r < -20.5$ (from Goto et al. 2003).

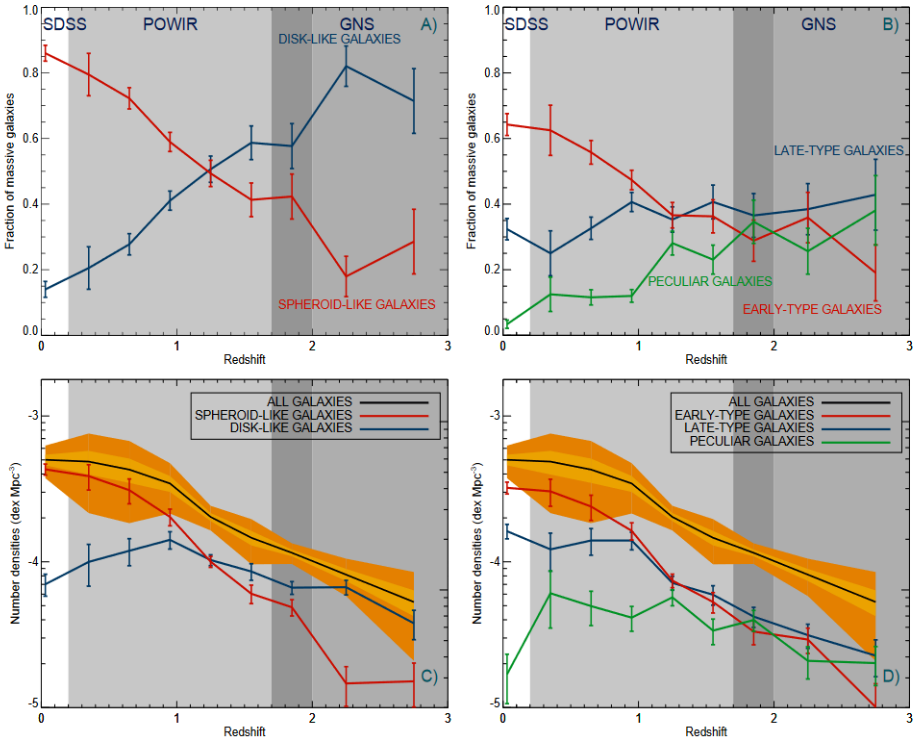


Figure 1-6 – Evolution of the morphological mix and relative number densities of galaxies across cosmic times, as revealed by several astronomical surveys. *Left*: galaxies are classified by Sérsic index (cf Section 4.4). *Right*: galaxies are classified visually. From Conselice 2014.

Figure 1-7 – *Left*: 2.8 × 2.8 arcmin image of the Hubble Ultra Deep Field (HUDF, Beckwith et al. 2006). *Right*: Mock HUDF observation from the Illustris cosmological simulation (Vogelsberger et al. 2014). From Snyder et al. (2015).

tion of baryons and Newton's laws of gravity. These structures span a wide range of masses (from $\sim 10^6$ M$_\odot$ for the least massive dwarf galaxies to $\sim 10^{13}$ M$_\odot$ for the most massive giant ellipticals) and sizes (from $\sim 10^2$ to $\sim 10^5$ parsecs), and by integrating the number counts derived from our deepest images of the universe to date, it is estimated that $\sim 2 \times 10^{12}$ galaxies more massive than 10^6 M$_\odot$ inhabit the observable universe (up to $z \sim 8$), and could theoretically be observed using current technology (Conselice et al. 2016).

Most galaxies have two main structural components: a disk and a bulge (or/and pseudo-bulge), each with a distinct formation history and evolution, and with different observational properties, as detailed in the following sections. Of course this description is very crude, as many galaxies display a large variety of intricate structural components, including bars, spiral arms, lenses, nuclear, inner and outer rings, and stellar halos. But in this thesis we will restrict ourselves to the description of the bulge and disk. The reader is referred to Buta (2013) and Kormendy 2013 for a more detailed description of other galaxy components.

1.3.1 Bulges and pseudo-bulges

Bulges (also referred to as *classical bulges* or *spheroids*) consist in densely packed featureless spheroidal (or mildly triaxial) groups of stars found at the center of spiral and lenticular galaxies. They can be distinguished from other galaxy components by their rounder smooth isophotes, or by their characteristic radial surface-brightness distribution that creates an excess of light at the center of their host galaxy relative to the inward extrapolation of the outer disk component (cf Section 4.4) (Gadotti 2012). These structures are kinematically hot, which means that their kinematics is dominated by the random motion of their stars. It is estimated that $\sim 1/3$ of the stellar mass in the local universe resides in bulges (e.g., Driver et al. 2007; Gadotti 2009), and the stellar populations within bulges display a wide range of ages and metallicities (e.g., Pérez and Sanchez-Blazquez 2011). They share many properties with elliptical galaxies of intermediate luminosity, such as similar color-magnitude relations (Balcells and Peletier 1993) or metallicity-luminosity relations (e.g., Jablonka et al. 1996). However, defining bulges as merely "smaller ellipticals surrounded by a disk" is oversimplistic, as both structures can have different formation histories (Gadotti 2009). Most bulges host a supermassive black hole (SMBH) in their center, whose mass correlates with bulge luminosity/stellar mass and velocity dispersion (e.g., Younger et al. 2008; Gültekin et al. 2009). This suggests that SMBH and bulge formation are tightly connected. The favored explanation for the growth of classical bulges is via major and minor mergers [10] in the early universe (Aguerri et al. 2001; Robertson et al. 2006), but the coalescence of migrating giant star-forming clumps to the inner regions of galaxies at high redshift also constitutes a viable scenario (e.g., Bournaud et al. 2007; Elmegreen et al. 2009). However, the details of this process, such as the number of mergers involved, or the fraction of minor to major mergers are still not well understood.

Classical bulges are to be distinguished from *pseudo-bulges*, which are are rotationally supported structures (i.e. disk-like) mainly found in Sbc and later-type galaxies. First discovered by Kormendy (1993), pseudo-bulges have flatter shapes than classical bulges, contain younger stellar populations and often display boxy or peanut shapes[11] when observed close to edge-on.

---

[10] A merger event that occurs with a mass ratio between 1:1 and 1:4 is called a *major merger*. A merger event that occurs with a mass ratio between 4:1 and 10:1 is called a *minor merger*.

[11] The bulge of the Milky Way is probably a pseudo-bulge, as its peanut shape suggests in near-IR images (Dwek et al. 1995).

Their birth is thought to be attributed to the secular evolution of disks, namely by transport of angular momentum from bars to the center of galaxies via gravitational instabilities (e.g., Combes et al. 1990). The properties of pseudo-bulges are reviewed in Kormendy and Kennicutt Jr (2004).

To make things even more complex, it has recently been shown that bulges and pseudo-bulges can co-exist within the same galaxy: classical bulges can be found within pseudo-bulges, and pseudo-bulges inside classical bulges are found to be relatively common (e.g., Gadotti 2009; Nowak et al. 2010, Erwin et al. 2014; Méndez-Abreu et al. 2014). To this date, the very definition of bulges and pseudo-bulges is still a matter of debate, and the community relies on various working definitions. Good recent discussions on terminology, respective observational properties and formation processes can be found in Erwin et al. (2014), Fisher and Drory (2016) and Kormendy (2016).

1.3.2 Disks

Disks are rotationally-supported flat structures found in lenticular and spiral galaxies, whose exponential radial surface brightness distribution defines their characteristic scale radius (cf Section 4.4). Most stellar disks can be decomposed into a thin disk, populated with younger metal-rich stars and a thick disk, with older metal-poor stars, larger scale heights and longer scale lengths (Burstein 1979; Van der Kruit 1981; Yoachim and Dalcanton 2006). Most of the disk stellar mass resides in the thin disk in massive spiral galaxies (Yoachim and Dalcanton 2006). Moreover, galaxy disks exhibit a radial gradient of stellar populations, with inner regions hosting older and more metal-rich stars than outer regions (Bell and De Jong 2000). Observations of the 21-cm wavelength line of hydrogen gas have also revealed the existence of a warped neutral hydrogen (HI) disk, much more extended than the optical stellar disk and misaligned to its plane, suggesting different formation scenarios for the two components (Sancisi 1976; Rogstad and Shostak 1971). Disks show many correlations between their observed properties: for example brighter and more massive disks tend to be bigger (Shen et al. 2003) and to rotate faster (Tully and Fisher 1977) than less massive ones. At high redshift ($z \sim 2$), most disks display a characteristic clumpy structure, with large (several kpc in size) massive (10^7 to 10^9 $M_\odot$) star-forming clumps, likely formed by gravitational instabilities (e.g., Elmegreen et al. 2005; Bournaud et al. 2007) or by mergers (e.g., Overzier et al. 2008). A theory of disk formation must therefore link these early clumpy galaxies to the exponential disks of the local universe.

The birth of galaxy thin disks is well described within the current paradigm of galaxy formation (White and Rees 1978; Fall and Efstathiou 1980; Mo et al. 1998; Cole et al. 2000). In this paradigm, big dark matter halos form hierarchically through the accretion and merging of smaller structures. Neighboring large-scale structures exert a tidal torque on these halos, causing them to spin. Baryonic gas clouds heated by shocks then collapse in the potential wells created by the dark matter halos, creating a halo of hot gas, whose internal pressure prevents it from further collapsing. As gas cools thanks to various radiative processes (discussed in Kauffmann et al. 1994), pressure decreases and gas clouds start to fall to the center of their host dark halo. Assuming conservation of angular momentum during the collapse, this results in a rotationally-supported disk, with scale lengths that are in good agreement with observations (e.g., de Jong and Lacey 2000). Recent theoretical development supported by numerical simulations have since shown that the accretion of cold streams of gas (not heated by shocks)

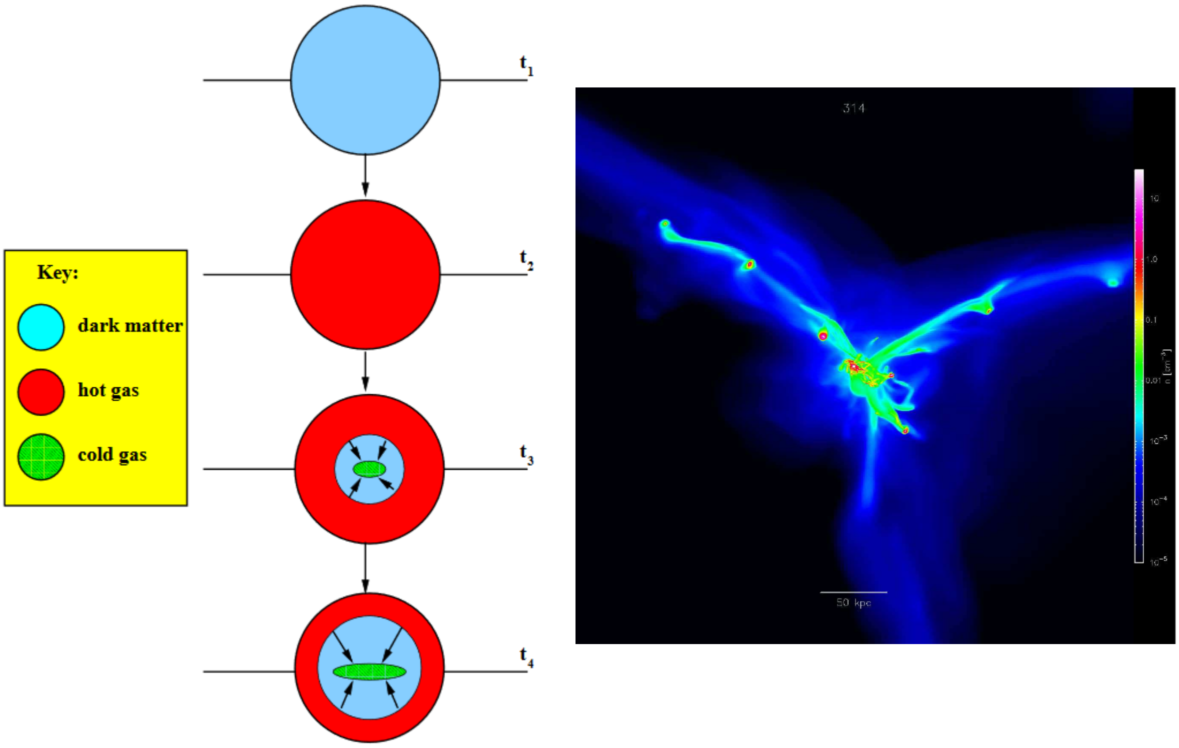


Figure 1-8 – *Left*: Formation of a disk galaxy, according to the two-stage scenario of White and Rees (1978) (from Baugh 2006). *Right*: Cold flows feeding a galaxy in a cosmological simulation produced by D. Tweed (from Danovich et al. 2012).

from the filaments of the cosmic web by disks could be the dominant mode of gas accretion in galaxies, as well as the main driver of cosmic star formation (e.g., Birnboim and Dekel 2003; Kereš et al. 2009, cf Figure 1-8). These cold flows, by interacting with the hot halo, could even result in the formation of the warped outer disks observed in HI (e.g., Roškar et al. 2010; Pichon et al. 2011). However, cold flows have not yet been observed and remain speculative for now. Disks can be destroyed in violent mergers events, resulting in a spheroidal remnant (Toomre and Toomre 1972), but in some gas-rich major mergers, hydrodynamical simulations show that they can be rebuilt (Robertson et al. 2006). The properties of galaxy disks are reviewed extensively in Van Der Kruit and Freeman (2011).

1.4 Ingredients of galaxy evolution

The recipe of galaxy evolution involves many interconnected ingredients, as illustrated in Figure 1-9. Galaxies form in the potential wells of their dark matter halo, and evolve in the environment of the cosmic large-scale structures, whose geometry and dynamics is constrained by the cosmological parameters. They contain various stellar populations, formed in cold and dense giant molecular clouds, and which evolve on separate paths according to their initial mass and metallicity. The most massive stars end in violent explosions known as supernovae, ejecting vast amounts of gas and heating the local interstellar medium, preventing (or "quenching") further star formation, a phenomenon known as "supernova feedback" (cf Figure 1-10). The cycle of life and death of stars enriches the interstellar medium with metals, and the enriched gas, after a cooling phase, is in turn used to form stars with higher metallicities. The interstellar medium also contains interstellar dust, mainly grown in dense molecular clouds, which reddens and dims the light of galaxies by absorption and scattering of incident photons out of the line of sight. In some galaxies, the presence of an active galactic nucleus (AGN) associated to its central supermassive black hole can create relativistic jets and winds that clear interstellar gas out and prevent it from cooling, therefore quenching star formation (this is known as "AGN feedback", cf Figure 1-10). Galaxies can live isolated lives, developing features associated with secular processes, such as bars and spiral arms, or they can interact more or less violently with one another via gas-rich ("wet") mergers, associated with bursts of star formation, or gas-poor ("dry") mergers with no star formation involved (e.g. Lin et al. 2008). Galaxies moving in a cluster environment can also have their cold gas stripped by the hot intracluster medium, a phenomenon known as "ram pressure stripping", which can result in a slow decline of star formation rate (a process called "starvation" or "strangulation").

Understanding how all these elements interact with one another quantitatively to form the distribution of galaxies we observe in astronomical surveys is a daunting task, notably because these complex systems involve physical processes on a remarkably large range of spatial scales, from stellar black holes to the large-scale structures of the cosmic web (cf Table 1.1), with the small scales acting on the large scales and inversely. Our current attempt at modeling the processes at play in galaxy evolution relies on two main approaches: cosmological simulations and semi-analytic models.

- *Cosmological simulations* follow explicitly the evolution of dark matter and gas in the non-linear regime of a ΛCDM universe. In these simulations, dark matter is modeled as a collection of collisionless massive particles whose dynamics follows the Vlasov-Poisson equations in a box of cosmological (comoving) volume with periodic boundary conditions. To reproduce the conditions of the early universe, they use a Gaussian random field of density perturbations with a power spectrum constrained by the CMB (e.g., Aghanim et al. 2016) as their initial conditions. These N-body simulations can be completed with the much more computationally intensive baryonic physics, which in practices comes down to solving the Euler equations on a collisional ideal fluid[12] via Lagrangian or Eulerian methods. In this scheme these simulations are referred to as *cosmological hydrodynamical simulations*. Sub-grid models that reproduce the effects of star formation, turbulence, supernovae and AGN feedback on the surrounding gas, supermassive black hole growth, chemistry of heavy elements and dust are unavoidable because galaxy evolution span a wide range of scales. These sub-grid models usually consist in a mix of empirical (or ad hoc) rules whose calibration is open to fine tuning to fit the data, as the detailed physics of these processes are not yet well understood. State-of-the-art cosmological hydrodynamical simulations include Illustris (Vogelsberger et al. 2014), EAGLE (Crain et al. 2015) and Horizon-AGN (Dubois et al. 2014).

- *Semi-analytic models* (or *SAMs*) model small-scale baryonic processes as simplified analytical descriptions within the framework of hierarchical structure growth. The physical

---

[12] A collisional fluid is a fluid that can be heated and cooled radiatively.

Table 1.1 – The spatial scales of galaxy evolution

Object	Order of magnitude in size [pc]
Stellar black hole	$10^{-13} \rightarrow 10^{-12}$
Star	$10^{-13} \rightarrow 10^{-5}$
Supermassive black hole	$10^{-9} \rightarrow 10^{-5}$
Globular cluster	10^{1}
Molecular cloud	$10^{1} \rightarrow 10^{2}$
Galaxy	$10^{2} \rightarrow 10^{5}$
Galaxy cluster	$10^{6} \rightarrow 10^{7}$
Large Scale Structures	$10^{8} \rightarrow 10^{9}$ (comoving)
Observable universe	10^{10} (comoving)

phenomena included in a SAM and the accuracy of their description differ depending on the models. The first self-consistent semi-analytic model of galaxy evolution of White and Frenk (1991) already included the effects of star formation, stellar populations and feedback. The main advantage of SAMs over cosmological hydrodynamical simulations is that they are much less computationally expensive, which allows them to explore the physical parameter space more easily. They are usually superimposed on N-body simulations, relying on their well-calibrated "merger trees" (recordings of the mass growth history of dark matter halos), and using the halo distribution as the sites of galaxy formation. Comparison between the predictions of SAMs and of cosmological hydrodynamical simulations show surprising agreement between the two approaches, at least down to the resolution elements of simulations, but these comparisons are usually limited to the hydrodynamics and cooling of gas (Lu et al. 2011), or to the scale of single galaxies (Stringer et al. 2010). State-of-the-art semi-analytical models include the SAGE model of Croton et al. (2016), GALACTICUS (Benson 2012) and GALFORM (Lacey et al. 2016), among many others.

Recent methodological reviews on cosmological simulations and semi-analytic models can be found in Baugh (2006) and Somerville and Davé (2015).

1.5 Unanswered questions on galaxy evolution

The evolution of galaxies in the universe is still an incomplete story. While cosmological simulations have been very successful at predicting the properties of the large-scale structures of the universe in the ΛCDM paradigm, the modeling of the small-scale baryonic physics of galaxy evolution is still at its early stage and will require many refinements to be able to reproduce the observed properties of galaxy populations in sky surveys. Besides the billion-dollar questions of the nature of elusive dark matter, that dominates the matter content of the universe, and dark energy, which causes the expansion of the universe to accelerate, a lot of physical phenomena associated to galaxy evolution are still poorly understood:

- *How and when did the first stars and galaxies form?* The birth of the first luminous objects, which constitute the building blocks of present-day galaxies, signed the end of the cosmic "dark ages", but so far our understanding of these structures has been

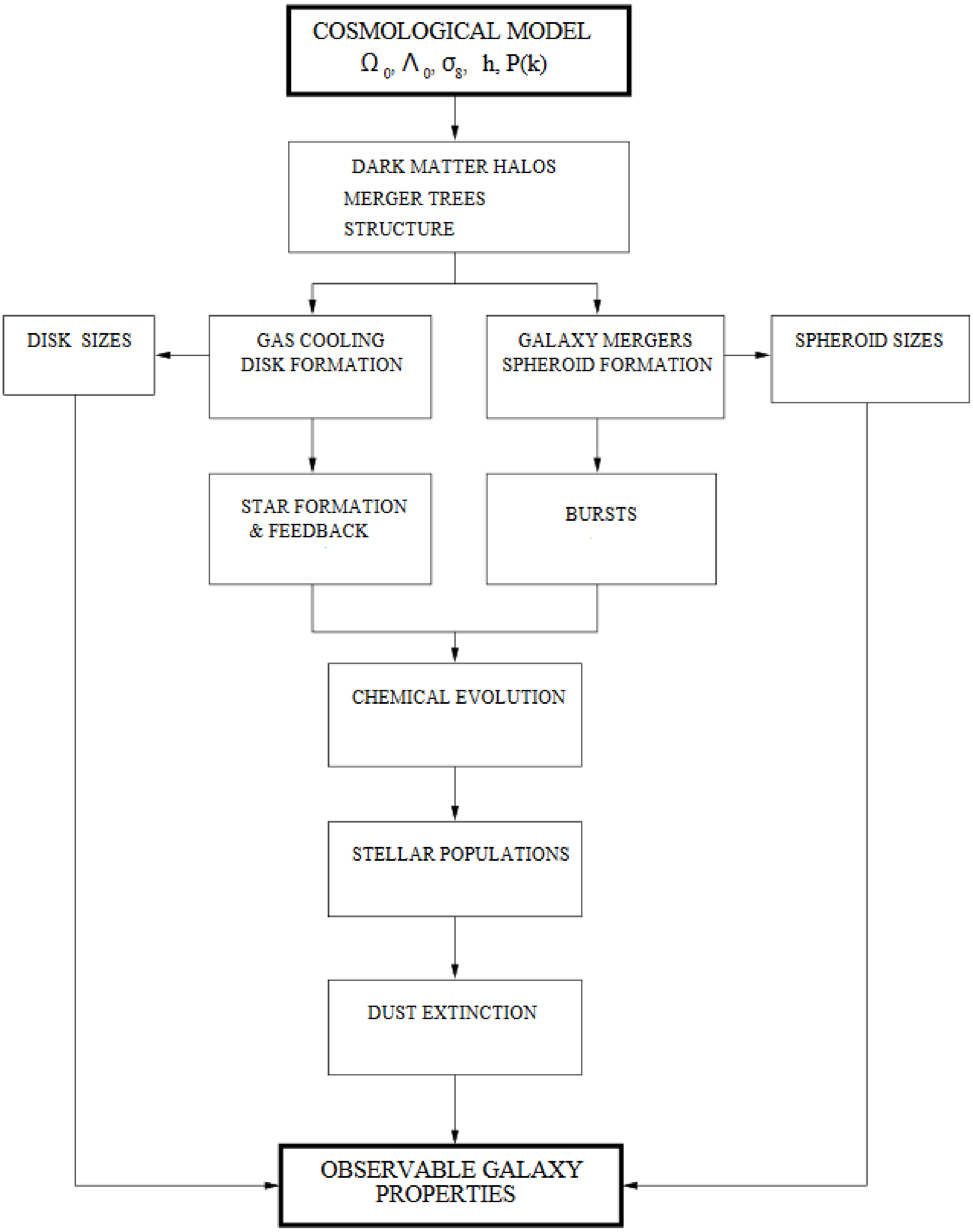


Figure 1-9 – Ingredients of galaxy evolution models and their interactions. From Cole et al. (2000).

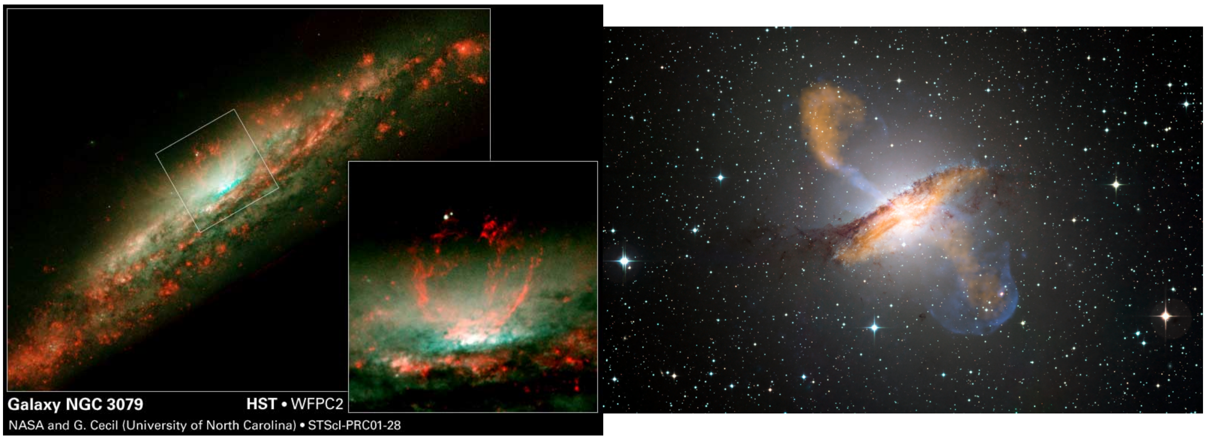


Figure 1-10 – *Left*: Supernova feedback in NGC 3079 (Credits: NASA and G. Cecil (U. North Carolina, Chapel Hill)). *Right*: AGN feedback in Centaurus A (Credits: ESO / WFI (Optical); MPIfR / ESO / APEX / A.Weiss et al. (Sub-millimeter); NASA / CXC / CfA / R.Kraft et al. (X-ray)).

largely theoretical. Imaging the first galaxies to reveal the complete story of the Epoch of Reionization is the next observational frontier, and the new generation of radiotelescopes such as LOFAR and SKA should be able to probe this era (Loeb 2010). In this context, what defines a "first galaxy" is currently open to debate (Bromm and Yoshida 2011).

- *What is the origin of the Hubble sequence?* It all comes down to the question of how galaxies acquire their mass. At the center of the debate is the relative roles of mergers and secular processes in the buildup of galaxy structure. A complete theory of galaxy formation should be able to explain the whole spectrum of structural properties observed in the cosmic zoo, from pure disks to pure spheroids, and in particular to predict the physical origin of galaxy bimodality.

- *What shuts down star formation in galaxies?* This is *probably* related to the fact that we do not currently fully understand how stars form. Observations show that only ∼10% of the baryonic mass of the universe is contained within stars or in cold gas (Fukugita and Peebles 2004). What makes the star formation process so inefficient, and why is the cosmic star formation rate declining from $z \sim 2$? One of the factors to explain this decline is the observed decrease of merger rate of galaxies over cosmic time (e.g. Lin et al. 2004; Conselice et al. 2008; De Ravel et al. 2009). It is also supposed that a number of mechanisms prevents the hot interstellar gas from cooling and forming stars, the so-called "feedback" mechanisms. The two favored candidates are supernova feedback at the low-mass end, and AGN feedback at the high-mass end (as illustrated in Figure 1-10). The small sizes involved in these mechanisms confine them in sub-grid models for now, and different implementations of feedback in cosmological simulations generate a wide scatter in terms of galaxy properties.

- *Why did the most massive galaxies formed their stars first and over a shorter time span?* Hierarchical models predict that small structures form first, and big structures form last.

But the most massive elliptical galaxies in the local universe seem to have formed their stars first and have stopped forming stars first, as attested by their redder colors (De Vaucouleurs 1961) and their higher metallicities (Faber 1973) associated with older stellar populations. This "downsizing" (Cowie et al. 1996) of massive galaxies cannot be explained within the current hierarchical framework.

- *How do supermassive black holes form?* Supermassive black holes, in the form of quasars, might have a role to play in the reionization of the cosmos (albeit ill-constrained for now). Moreover, as we have seen in Section 1.3.1, the formation of galactic bulges is thought to be associated with the formation of supermassive black holes. But how these black holes acquire their mass is still unknown. Hypotheses to explain their growth include the coalescence of stellar remnants from the first generation of stars (Madau and Rees 2001) or the direct collapse of pristine gas clouds (Bromm and Loeb 2003; Hartwig et al. 2016).

A recent review on the current status of the galaxy evolution field detailing these open questions is available in Silk and Mamon (2012). To conclude, there has been tremendous progress in the theoretical and observational sides in the last few decades, but we still have a long way to go to fully understand how these magnificent structures grow and evolve. This is an exciting time to be alive.

Chapter 2

Exploring astronomical surveys

> *Physics is really nothing more than a search for ultimate simplicity, but so far all we have is a kind of elegant messiness.*
>
> Bill Bryson, *A Short History of Nearly Everything*

Abstract

Astronomical surveys constitute our primary source of information on galaxy formation and evolution, by providing us with a panchromatic view of the universe through different slices of cosmic time. In this chapter[1], I review some of the main existing spectro-photometric galaxy surveys probing the local and deep universe, and summarize what they indicate about the size and luminosity distribution of galaxies, as well as the evolution of these observables with redshift and the biases involved in their estimation.

Contents

---

[1] Note: This chapter uses extracts from Carassou et al. (2017) in some of its sections.

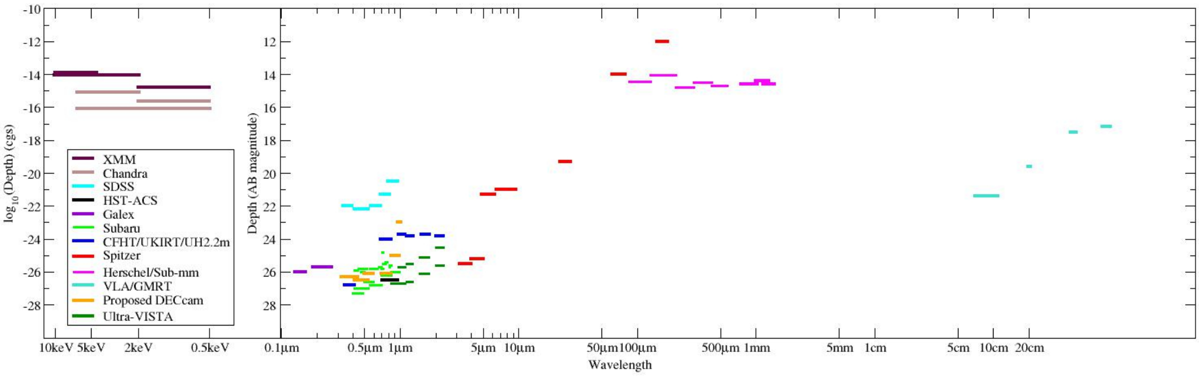


Figure 2-1 – Sensitivity (or depth) of a selection of telescopes used for galaxy surveys as a function of their wavelength range. From the COSMOS field website: `http://cosmos.astro.caltech.edu/page/astronomers`.

2.1 Galaxy surveys

Galaxy surveys provide us with a panchromatic view of the low and high-redshift universe, covering most of the electromagnetic spectrum from X-ray to radio in specific regions of the celestial sphere, although the optical, near-infrared, mid-infrared, far-infrared and sub-mm ranges represent the wavelength coverage of most surveys (cf Figure 2-1). As astronomers do not have the time and resources[2] to follow the same galaxy for billions of years, they must infer the evolution of different galaxy populations by comparing their properties at different epochs along cosmic time, namely their shapes, colors and sizes.

Assuming that the Milky Way does not occupy a special place in the universe (Copernician principle), that the spatial distribution of matter in the universe is homogeneous and isotropic on large scales (cosmological principle), and that the distribution of galaxies in well-separated regions of the universe constitute independent realizations of the same process (ergodic assumption), one can expect that the finite galaxy samples accessible to observations, if probed in a sufficiently large cosmological volume, are fairly representative of the whole universe[3]. This is known as the *fair sample hypothesis* (Peebles 1973), and it is especially important to verify it when studying the clustering properties of galaxies and its distribution at large scales, or in order to make statistically relevant predictions on the evolution of galaxy distributions.

2.1.1 Survey characteristics

There are two main kinds of galaxy surveys:

- *Deep surveys* benefit from large exposure times to retrieve faint galaxies and probe large redshift ranges, but only a tiny area fraction of the sky can be covered, which is why they are usually referred to as “pencil-beam surveys”.

- *Wide surveys* cover significant fractions of the celestial sphere but are shallower in terms of magnitude limit (hence redshift interval). Wide surveys allow to trace the properties

---

[2] And graduate students.

[3] As the American statistician George Gallup once said about public opinion polls, "sampling public opinion is like sampling soup: one spoonful can reflect the taste of the whole pot if the soup is well-stirred.”

of rare populations of galaxies (such as local compact red galaxies, e.g., Trujillo et al. 2009) , hence providing representative snapshots of the universe at low-redshift.

Because of limited exposure time, a compromise between angular area and depth[4] must be found when designing a survey for a specific scientific purpose. Each kind of survey can be purely photometric (probing the morphology and magnitudes of galaxies), purely spectroscopic (probing their spectral energy distribution (SED)), or spectro-photometric. Multi-band photometric surveys (i.e. surveys using multiple well-defined filters) can also measure the colors of galaxies, obtained by subtraction of their apparent magnitude in different passbands.

A galaxy survey is usually characterized by its area on the sky, its wavelength range, its sensitivity, its spectral resolving power $R = \lambda/\Delta\lambda$ (for spectroscopic surveys) and its completeness limits. The *sensitivity* of a survey is the magnitude of the faintest unambiguously detected source, usually defined so that its flux is $> k\sigma$, where σ is the flux measurement uncertainty (Léna et al. 2012). *Completeness limits* (or levels) can be determined for each filter by injecting artificial realistic galaxies generated using an empirical model with random positions, sizes and photometric properties into real images. The completeness limits are then defined as the ratio of the number of detected artificial sources over the number of input sources (cross-identified using their position) in bins of source apparent magnitudes (e.g., Papovich et al. 2016). If a sample extracted from a given survey only contains objects brighter than its 100% completeness limit, it is said to be “complete” to that limit.

In deep surveys in particular, each area of the sky is surveyed many times. A field image is then a collection of these different exposures (or *stack*), that have to be overlapped (the colloquial term is *co-added*) to increase the signal to noise ratio, after carefully designed preprocessing steps that are survey-dependent. In general, deep surveys are conducted in several regions of the sky to limit the effect of cosmic variance[5] on the precision of the derived measurements.

Different wavelengths ranges can probe different physical phenomena involved in galaxy evolution: the morphologies, colors and spectral energy distributions of galaxies can be direct or indirect tracers of their stellar populations and dust extinction. Hence, they provide estimates of their stellar mass, their star formation rate, their distance and their age. Correlating these observables with the galaxies spatial distribution allows astronomers to infer the environmental dependence of galaxy color and morphology, and even to constrain cosmological parameters.

The resolution of ground-based instruments at optical wavelengths is limited by atmospheric turbulence above the observation site that distorts incoming wavefronts, hence generating short-term flux variations (commonly referred to as “scintillation”) and blurring of astronomical images (an effect called “seeing”, Hardy 1998), and some regions of the electromagnetic spectrum are blocked by atmospheric absorption, such as parts of the mid-IR and essentially all of the far-IR regions. To avoid these inconveniences, space-based telescopes are required, such as the Hubble Space Telescope (HST), which can resolve details down to less than 0.12 arcsecond across, a resolution unmatched by most optical ground-based instruments, unless using adaptive optics that measure and compensate the distortions of incident wavefronts in real time.

---

[4] Depending on the context, the depth of a survey can also refer to its sensitivity.

[5] Cosmic variance is the inherent uncertainty in galaxy statistics due to the density fluctuations from large-scale structures, limiting the precision of our inferences on galaxy populations. (e.g., Moster et al. 2011)

2.1.2 Surveying the local universe

Our best vision of the local universe mainly comes from:

- The **Sloan Digital Sky Survey** (SDSS, York et al. 2000), an on-going spectro-photometric survey that started in 1998, conducted using the 2.5m telescope of the Apache Point observatory in New Mexico. The SDSS covers an area of $\sim 14,555\ deg^2$ on the sky (over 1/3 of the celestial sphere) in five optical bands (*ugriz*, Fukugita et al. 1996). The last data release (DR12, Alam et al. 2015) contains $\sim 4.7 \times 10^8$ cataloged objects (detected in the optical r-band) and more than 1.8×10^6 galaxy spectra. Its main galaxy sample has a median redshift of z = 0.1. A slice of the 3D distribution of galaxies as revealed by the SDSS is shown in Figure 2-2.

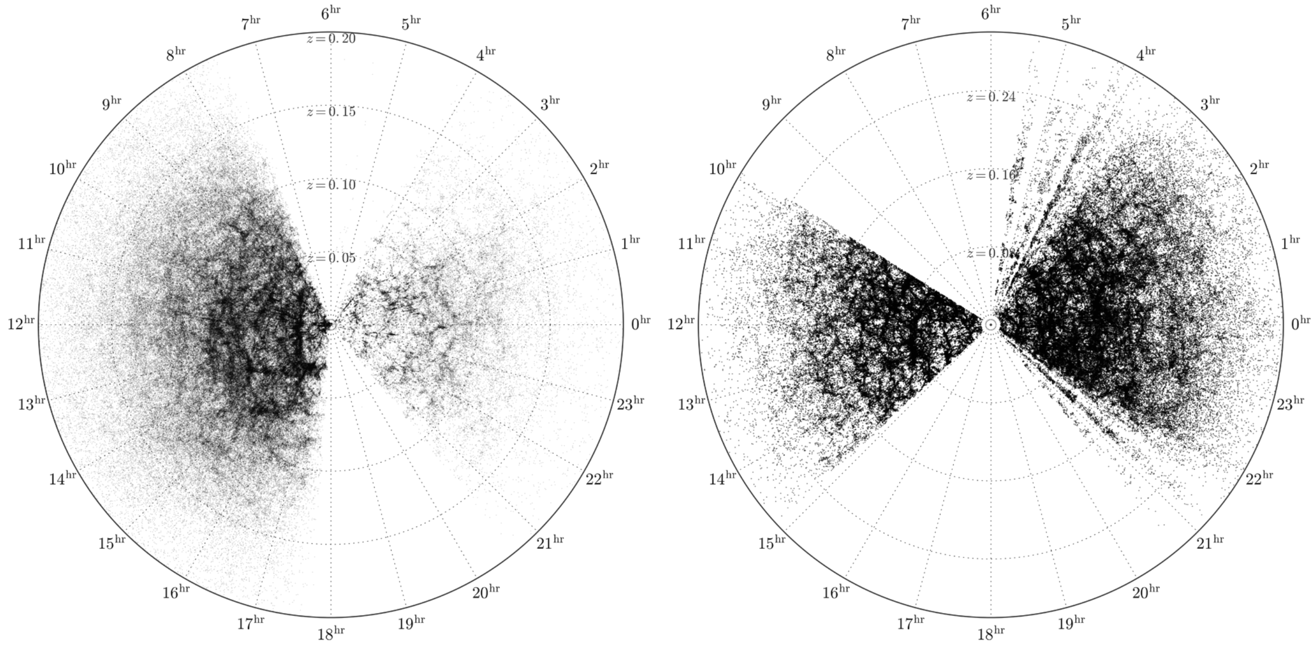
- The **2-degree Field Galaxy Redshift Survey** (2dFGRS, Colless et al. 2001), conducted between 1997 and 2002 using the 3.9m Anglo-Australian Telescope, covers $\sim 2000\ deg^2$ (~5% of the celestial sphere). Its catalog contains the spectra and spectroscopic redshifts of more than 221,000 galaxies with magnitudes brighter than 19.3 in the optical b_J-band up to a redshift of $z \sim 0.3$ (with a median redshift of z = 0.1). A slice of the 3D distribution of galaxies as revealed by the 2dFGRS is shown in Figure 2-2.

- The **Two Micron All-Sky Survey** (2MASS, Jarrett et al. 2000) was conducted between 1997 and 2001 using two 1.3m telescopes in the northern and southern hemispheres (the Fred Lawrence Whipple Observatory in Arizona and the Cerro Tololo Inter-American Observatory in Chile respectively). Its extended source catalog (the largest all-sky photometric catalog of extended sources so far) contains 1.5 million galaxies observed in the near-IR JHK_s bands, of which 1 million are brighter than $K_s = 13.9$ mag, the completeness limit of the survey. A subsample of this catalog contains the spectroscopic data of 44,000 galaxies with a median redshift of $z \sim 0.03$, which constitutes the 2MASS Redshift Survey (2MRS, Huchra et al. 2012). Recently, Bilicki et al. (2014) released the 2MASS Photometric Redshift catalog (2MPZ), containing the distance of 940,000 galaxies with a median redshift $z = 0.07$.

- The **Wide-field Infrared Survey Explorer** (WISE, Wright et al. 2010) is an all-sky survey conducted in 2010 by NASA's eponymous space-based instrument in four mid-IR bands (3.4, 4.6, 12, and 22 μm). Benefiting from an uniform sky coverage, it is the most sensitive mid-IR survey to date. However its poor angular resolution makes its unable to resolve objects beyond the depth already probed by 2MASS. Its publicly released source catalog contains the photometric data of more than 500 million objects (asteroids, planets, stars, nebulae and galaxies) with a signal-to-noise ratio of 5 or higher.

2.1.3 Surveying the deep universe

Here below is a non-exhaustive list of some of the most studied modern deep galaxy surveys, which constitute our main database for the high-redshift universe. A comparison of the relative size over the sky of those of these surveys that use the Hubble Space Telescope is displayed in Figure 2-3. These surveys include:



(a) SDSS

(b) 2dFGRS

Figure 2-2 – A slice of the 3D distribution of galaxies, as revealed by the SDSS and 2dF-GRS, which provide our most complete picture to date of the large-scale structure of the local universe out to $z \sim 0.1 - 0.2$. Maps generated by S. Hinton.

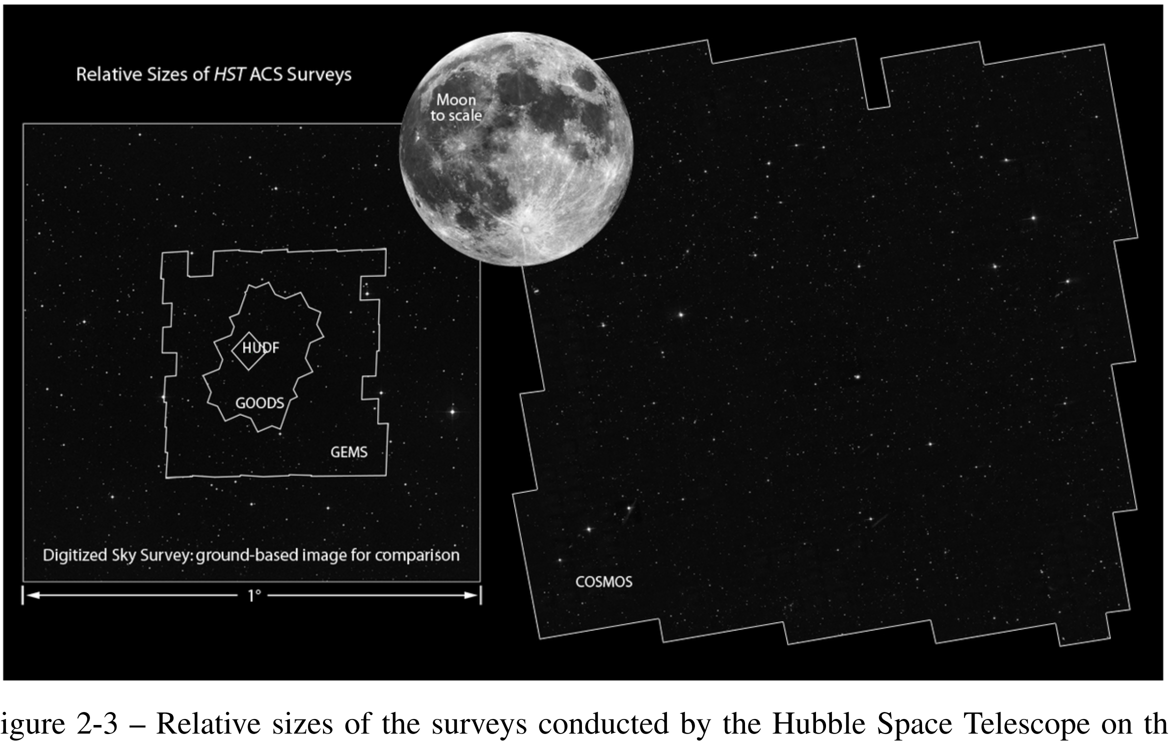


Figure 2-3 – Relative sizes of the surveys conducted by the Hubble Space Telescope on the sky. Credit: NASA, ESA, and Z. Levay (STScI)

- The **VIMOS-VLT Deep Survey** (VVDS, Le Fèvre et al. 2005), which consists in 3 complementary surveys (VVDS-Wide, VVDS-Deep and VVDS-UltraDeep) conducted using the VIRMOS multi-object spectrograph on the 8.2m ESO-VLT telescopes covering 8.7 deg^2, 0.74 deg^2 and 512 arcmin^2 of the sky respectively (in 5 different fields), from 1996 to 2013. Photometric properties were acquired in the UBVRIK near-UV to near-IR bands using the CFHT-CFH12K camera (BVRI filters), the ESO-NTT-SOFI (K' filter), and the ESO-2.2m-WFI (U filter). Its final data release (Le Fèvre et al. 2013) contains a total of $\sim 34,500$ spectroscopic redshifts of galaxies brighter than 22, 24 and 24.75 in $I_{AB}-$band up to a redshift of $z \sim 6.7$ (more than 12 Gyrs of cosmic time), with median redshifts $z = 0.55, 0.92$ and 1.38 for its Wide, Deep and UltraDeep surveys respectively.

- The **Cosmic Evolution Survey** (COSMOS, Scoville et al. 2007) is the largest survey undertaken by the Hubble Space Telescope. HST observations of this 2 deg^2 field are completed with multi-λ observations covering X-ray, UV, optical/IR, mid-IR, mm/submm up to radio by many ground-based and space-based observatories, down to a magnitude $I_{AB} = 27$. Benefiting from its relatively large field, COSMOS samples a cosmological volume almost equivalent to that of the SDSS for the local universe, detecting $\sim 2 \times 10^6$ objects over the redshift range $0.5 < z < 6$, with a median redshift $z \sim 0.66$ at $22 < I_{AB} < 22.5$ and $z \sim 1.06$ at $24.5 < I_{AB} < 25$ (Ilbert et al. 2008).

- The **DEEP2 Galaxy Redshift Survey** (Newman et al. 2013) is a spectroscopic survey of high spectral resolution (R $\sim$ 6000) conducted using the DEIMOS spectrograph on the Keck II telescope and covering an area of 2.8 deg^2 on the sky. Its fourth public data release contains ~50,000 distant galaxies in the redshift range $0 < z < 1.4$ down to $R_{AB} = 24.1$, making it the largest high precision redshift survey of galaxies at $z \sim 1$ completed to date.

- The **Hubble Deep Field** (HDF, Williams et al. 1996) is a survey of the early universe conducted using the Advanced Camera for Surveys (ACS) and the Wide Field Camera (WFC) instruments aboard the Hubble Space Telescope, in 4 filters ranging from near-UV to near-IR. Two fields, the HDF-North and HDF-South, of 5.3 arcmin^2 and 0.7 arcmin^2, were surveyed in 1995 and 1998 respectively. A deeper survey, the Hubble Ultra Deep Field (HUDF, Beckwith et al. 2006), was then observed and revisited several times between 2002 and 2010, in a field of 11 arcmin^2 down to an equivalent magnitude of $\sim$ 30 in the $V-$band. But the Hubble eXtreme Deep Field (HXDF, Illingworth et al. 2013) provides the deepest observations to date, with a field of 4.7 arcmin^2 and a total exposure time of ~2 million seconds, reaching an unprecedented sensitivity of 31.2 $V_{AB}-$band magnitude. Galaxy catalogs of the HXDF provide optical+near-IR photometry of 14,140 galaxies with redshifts up to $z \sim 11$ (~400 Myrs after the Big Bang in cosmic time).

- The **Great Observatories Origins Deep Survey** (GOODS, Dickinson et al. 2003) combines deep observations from NASA's Great Observatories (Spitzer, Hubble, and Chandra) ESA's Herschel and XMM-Newton, and from the most powerful ground-based facilities (ESO-VLT, Keck, Gemini, VLA...), covering ~320 arcmin^2 in regions centered on the HDF-North and Chandra Deep Field South (CDF-S). The extensive coverage of these regions in X-ray, radio, sub-mm, near-IR and optical make them the most studied deep survey areas of the sky.

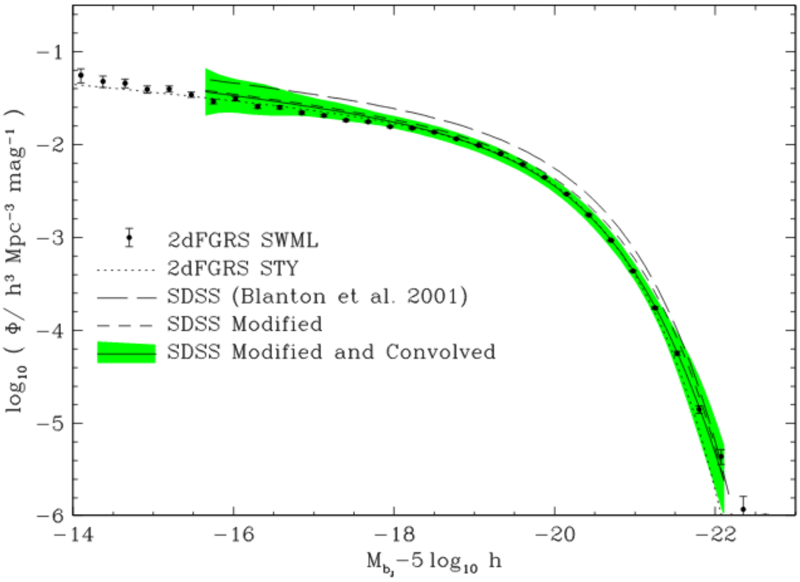


Figure 2-4 – The global luminosity function in the local universe. This plots compares the 2dFGRS b_J-band luminosity function (Norberg et al. 2002) with the LF estimates from the SDSS (Blanton et al. 2003). From Norberg et al. (2002).

- The **Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey** (CANDELS, Grogin et al. 2011) is a multiwavelength survey conducted using the WFC3/IR and ACS instruments of the Hubble Space Telescope. Its "Wide" program covers 668 arcmin^2 in 3 well-separated fields, and its "Deep" program covers 0.04 deg^2 within the GOODS-North and GOODS-South regions, imaging more than 250,000 galaxies in the redshift range $1.5 < z < 8$ in total.

For recent historical and methodological reviews on galaxy surveys and what we can learn from them, the reader is referred to Djorgovski et al. (2013) and Madau and Dickinson (2014). In the next sections, I summarize the properties of the light and size distributions of different populations of galaxies (two main indicators of evolution) that we inferred from the exploration of these surveys, with an emphasis on the problem of the size evolution of massive early-type galaxies (E/S0s) with cosmic time.

2.2 The luminosity distribution of galaxies and its evolution

The luminosity function (LF) of galaxies is a fundamental tool for characterizing the global properties of galaxy populations, and it is often used for constraining models of galactic evolution. As input data, these analyses use catalogs containing the photometric properties, such as apparent magnitudes (or flux), of a selected galaxy sample. Estimating the luminosity function requires the knowledge of the absolute magnitude of the sources, which itself depends upon the determination of their distance/redshift. The number density per luminosity bin can be determined by a variety of methods, parametric or non-parametric, that are described in details in Binggeli et al. (1988), Willmer (1997) and Sheth (2007). The resulting distribution is usually fitted by a Schechter function (Schechter, 1976), but other functions are sometimes required (e.g., Driver and Phillipps 1996; Blanton et al. 2005). The Schechter function is defined as:

$$\Phi(L)dL = \Phi^* \left(\frac{L}{L^*}\right)^{\alpha} e^{-\frac{L}{L^*}} \frac{dL}{L^*} \tag{2.1}$$

where L is the galaxy luminosity and $\Phi(L)dL$ is the number of galaxies with luminosity between L and L+dL per Mpc^3. It is characterized by 3 parameters: ϕ^* is the normalization density, α represents the faint end slope and L^* is a characteristic galaxy luminosity. But the Schechter LF is more conveniently expressed in terms of absolute magnitudes. Using the following transformation:

$$M - M^* = -2.5 log_{10} \left(\frac{L}{L^*}\right) \tag{2.2}$$

one obtains:

$$\Phi(M)dM = 0.4\ln(10)\Phi^* 10^{0.4(\alpha+1)(M-M^*)} \exp\left[-10^{0.4(M-M^*)}\right] dM \tag{2.3}$$

where M^* is a characteristic galaxy absolute magnitude. The global LF at z $\sim$ 0.1 is now well constrained thanks to the analysis of spectroscopic surveys, such as the 2dFGRS (Norberg et al. 2002) and the SDSS (Blanton et al. 2003), as illustrated in Figure 2-4. There is also clear evidence that the global LF evolves with redshift, and that the LFs for different populations of galaxies evolve differently (Lilly et al. 1995, Zucca et al. 2006). The determination of the LF evolution is nevertheless a challenge, as high-z galaxies are faint, and therefore generally unsuitable for a spectroscopic redshift determination, which would require prohibitive exposure times. The current solution to this problem is to use the information contained within the fluxes of these sources in some broad-band filters to derive their distance, an estimate known as *photometric redshift*. This procedure has a number of biases in its own right, because the precision of photometric redshifts relies on the templates and the training set used, that must be assumed as representative of the galaxy populations. These biases are described extensively in MacDonald and Bernstein (2010). In turn, distance uncertainties typically result in an increase of the estimated number of low and high luminosity galaxies (Sheth 2007).

2.3 The size distribution of galaxies and its evolution

2.3.1 Size estimators

Galaxy sizes can be expressed in terms of angular extent (in arcsec) or in terms of physical sizes (in kpc). Angular sizes can be converted into physical sizes if the distance of the galaxy is known. A popular size estimator is the *half-light radius* R_e (also called *effective radius*) of the Sersic (1963) surface brightness profile, defined in Section 4.4. Another is the *Petrosian radius* R_P (Petrosian 1976), defined as the radius at which the local surface brightness averaged over an annulus corresponds to 20% of the mean surface brightness within R_P, i.e.:

$$\frac{\int_{\alpha R_P}^{\beta R_P} I(r) 2\pi r\, dr / [\pi(\beta^2 - \alpha^2) R_P^2]}{\int_0^{R_P} I(r) 2\pi r\, dr / [\pi R_P^2]} = 0.2 \tag{2.4}$$

where $I(r)$ is the surface brightness profile (usually a Sérsic profile). The SDSS team set α and β (the extent of the annulus) to 0.8 and 1.25 respectively (Blanton et al. 2001; Shen et al. 2003).

The *Petrosian flux* F_P is defined as the flux observed within a circle of radius kR_P (the SDSS team set $k = 2$, as suggested by Yasuda et al. (2001) and Shimasaku et al. (2001)):

$$F_P = \int_0^{kR_P} I(R) 2\pi R \, dr \tag{2.5}$$

This allows one to define another size estimator, the *Petrosian half light radius*, which is the radius that contains half of the Petrosian flux.

2.3.2 The mass-size relation

The size distribution of each galaxy type at given luminosity (or stellar mass) can be well described by a log-normal function of mean $\overline{R_e}$ and of standard deviation $\sigma_{\ln R_e}$ (de Jong and Lacey 2000). From a sample of 140,000 local galaxies ($0 \leq z \leq 0.3$) from the SDSS, Shen et al. (2003) showed that galaxy size is correlated with stellar mass: more massive galaxies (early-types as well as late-types) tend to have larger sizes, as seen in Figure 2-5. This so-called *mass-size relation* (or *luminosity-size relation*) has been shown to be strongly dependent on galaxy morphology. For early-type galaxies (as defined by their Sérsic index $n > 2.5$[6]), the mass-size relation is best fitted by a single power law:

$$\overline{R_e} \propto M^a \tag{2.6}$$

where M is the stellar mass and $a = 0.55$ is the best-fit coefficient. For late-type galaxies ($n < 2.5$), it is best described by a double power law:

$$\overline{R_e} \propto M^\alpha \left(1 + \frac{M}{M_0}\right)^{\beta-\alpha} \tag{2.7}$$

where $M_0 = 10^{1.6} M_\odot$ is a characteristic mass at which $\sigma_{\ln R_e}$ changes significantly, and α, β are the fitting coefficients. Late-type galaxies more massive than M_0 follow $\overline{R_e} \propto M^{0.4}$, and those less massive than M_0 follow $\overline{R_e} \propto M^{0.15}$. Figure 2-5 shows that at fixed stellar mass, early-type galaxies tend to have smaller sizes than late-type galaxies (except for the most massive ones, i.e. with $M > 10^{11} M_\odot$), and have a steeper mass-size relation.

Whether the mass-size relation depends on galactic environment or not is still a matter of debate, with several studies finding no environmental dependence (e.g., Rettura et al. 2010; Maltby et al. 2010; Nair et al. 2010; Huertas-Company et al. 2013; Kelkar et al. 2015; Saracco et al. 2017), and others concluding that early-type galaxies tend to be larger in average in a high density environment, such as galaxy clusters, than in the field (e.g., Papovich et al. 2012; Bassett et al. 2013; Lani et al. 2013; Strazzullo et al. 2013; Delaye et al. 2014; Yoon et al. 2017). Lately, Zhang and Yang (2017) found from a sample of 639,555 galaxies with $0.01 \leq z \leq 0.2$ from the SDSS DR7 (Abazajian et al. 2009) that spiral and elliptical galaxies have larger sizes in low density regions than in high density regions.

---

[6] See details in Section 4.4.

2.3.3 The size evolution of massive early-type galaxies

The size evolution of massive early-type galaxies (ellipticals and lenticulars) is one of the highly debated topic in the field of galaxy evolution (e.g., Shankar et al. 2012; Delaye 2013). The debate started when Daddi et al. (2005) discovered in the Hubble Ultra Deep field (UDF) a population of 7 high-redshift ($1.4 \lesssim z \lesssim 2.5$) massive ($M > 10^{11} M_{\odot}$) E/S0 galaxies much more compact than their counterparts in the local universe (by a factor $\sim 2 - 3$, i.e. $R_e \lesssim 1$ kpc). 10 more extremely compact galaxies of this type were also found by Trujillo et al. (2006) at $1.2 < z < 1.7$ in the Munich Near-IR Cluster Survey (MUNICS, Drory et al. 2001), this time 4 to 5 times smaller than their equally massive local analogs. The apparent lack of massive compact elliptical galaxies in the local universe (Trujillo et al. 2009) suggests that the high redshift compact population has evolved and grown into the present-day massive population, in accord to the hierarchical scenario of structure formation. Hypotheses to explain this growth include *in situ* mechanisms, such as AGN activity (e.g., Fan et al. 2008) and SN winds (e.g., Damjanov et al. 2009) acting to "puff up" galaxies, or *ex situ* mechanisms, such as dry minor mergers (e.g., Ciotti and Van Albada 2001; Khochfar and Silk 2006; Hopkins et al. 2008; Nipoti et al. 2009) and major mergers (e.g., Khochfar and Burkert 2006; Oser et al. 2010).

This size evolution trend has since been confirmed by a variety of independent studies (e.g., Buitrago et al. 2008; Van Dokkum et al. 2008; Van Der Wel et al. 2008; Damjanov et al. 2011; Cimatti et al. 2012), as shown in Figure 2-6, and thanks to dynamical mass measurements[7] (Trujillo et al. 2009, Martinez-Manso et al. 2011; Van de Sande et al. 2011; Newman et al.), there appears to be a scientific consensus on the compact nature of a population of high redshift galaxies. However, comparing the results of all these studies with one another rigorously is a difficult task, because different groups use different datasets, different sample selection criteria for their galaxy populations, different data analysis methods, and face different observational biases. Galaxy populations can be selected by stellar mass, morphology, color, or star formation activity. Cosmological surface brightness dimming for example (see Section 3.1.1) is known to produce underestimates of galaxy sizes at high redshifts, and stellar masses estimates are usually provided by fitting spectral energy distribution templates to galaxy broadband photometric data, and therefore depend on the validity of the templates at these redshifts. Moreover, the population of compact galaxies found at high redshift might be unrepresentative of the true underlying population due to Malmquist bias (see Section 3.3), as hinted by, e.g., Mancini et al. (2009) and Valentinuzzi et al. (2010). Finally, linking these high-redshift compact galaxies to their local early-type counterparts is prone to the *progenitor bias* (e.g., Dokkum and Franx 1996; Saglia et al. 2010): by only selecting early-type galaxies in the high-redshift sample, one rules out the high-redshift late-type galaxies that might evolve into early-types at low redshift. Therefore early-type galaxies might have undergone a greater amount of transformation that it may seem. The combination of these effects leads to problematic discrepancies between the conclusions of these works: for example, Fan et al. (2010) assert that size evolution of massive early-types occurs rapidly at $z \gtrsim 1$ and becomes negligible at $z < 1$. Damjanov et al. (2011), however, find a continuous size evolution from $z = 2.5$ to $z = 0$, whereas Saracco et al. (2010) find no strong evidence for size growth from $z = 2$ to $z = 0$. To be able to circumvent these selection effects and measurement biases in order to infer

---

[7] Dynamical mass measurements are obtained via the velocity dispersion of stars within the galaxy, using the virial theorem: $M_{dyn} = \beta(n)\sigma^2 R_e / G$, where $\beta(n)$ depends on the Sérsic index of the galaxy. Cappellari et al. (2006) showed that $\beta = 5$ is a good approximation for elliptical galaxies.

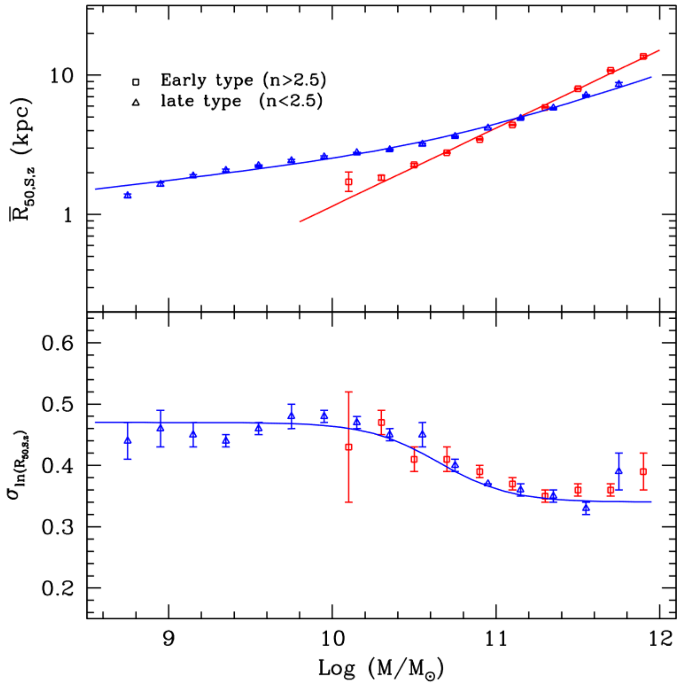


Figure 2-5 – Median mass-size relation in from the 140, 000 galaxies of the SDSS. The size is estimated by the Sérsic half-light radius. From Shen et al. (2003).

robustly the evolution of galaxy structure over cosmic time is the main objective of the work described in this thesis.

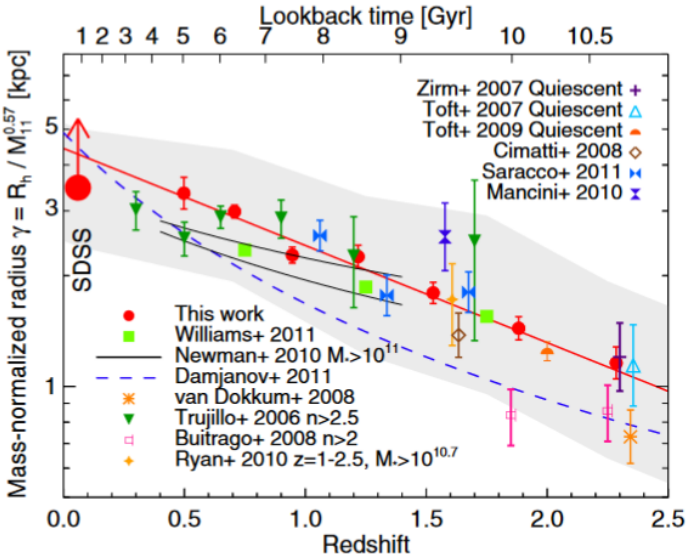


Figure 2-6 – Comparison of the results of different studies on the evolution of the average size of massive early-type galaxies. The half-light radius is normalized to $10^{11} M_{\odot}$ on the plot, using the slope $R_e \propto M^{0.57}$. From Newman et al. (2012).

Chapter 3

From images to catalogs: selection biases in galaxy surveys

> *L'essentiel est invisible pour les yeux*
>
> Antoine de Saint-Exupéry, le Petit Prince

Abstract

There's more to the universe than meets the eye. Because of various selection effects of cosmological, physical and instrumental origins, no source catalog has yet contained the totality of galaxies in a given cosmological volume. Some of them affect the number of samples retrieved in a given population, others affect our measurements of luminosity, color, surface brightness, and morphological properties. Worse, some of them cannot be corrected for analytically and correlate with each other in a non-linear way. This inevitably leads to systematic errors in inferring parameters models of galaxy evolution. In this chapter, I review some of the selection biases that are affecting most galaxy surveys, and some of the attempts that have been designed for dealing with them.

Contents

3.1 Effects of redshift

3.1.1 Cosmological dimming

The bolometric surface brightness (i.e. the surface brightness integrated over all wavelengths) of all cosmologically distant sources is altered by the expansion of the universe. This phenomenon was first predicted by Tolman (1930; 1934) and is now referred to in the literature as *cosmological surface brightness dimming*, or *Tolman dimming*. It can be theoretically derived briefly in the following way[1]:

Let F be the bolometric flux that an observer receives from a source of intrinsic luminosity L. We define its *luminosity distance* D_L as:

$$F = \frac{L}{4\pi D_L^2} \tag{3.1}$$

For a solid angle Ω subtended by the extended source of projected area A, we define its *angular diameter distance* D_A as:

$$\Omega = \frac{A}{D_A^2} \tag{3.2}$$

Etherington (1933) derived a relation between the angular distance and the luminosity distance, valid for any expanding universe described my a metric theory of gravity, as long as the number of photons is conserved and photons travel on null geodesics. This relation is termed as the *cosmic distance duality relation* (CDDR):

$$D_L = (1+z)^2 D_A \tag{3.3}$$

The apparent *bolometric surface brightness* I_{obs} is defined as the bolometric flux density per unit solid angle:

$$I_{\text{obs}} = \frac{F}{\Omega} = \frac{L}{4\pi A}\frac{1}{(1+z)^4} = \frac{I_e}{(1+z)^4} \tag{3.4}$$

The surface brightness of the source therefore strongly decreases with increasing redshift. This dimming effect is the result of 3 different phenomena:

- In an expanding universe, as light is emitted farther and farther from the observer (in terms of proper distances), successive waves of light arrive at the observer with increasing delay, reducing *de facto* the rate of reception of photons by a factor $(1+z)$.

- Each photon is redshifted and therefore loses energy as the universe expands, reducing the flux density by a factor $(1+z)$.

- The expansion of the universe also causes angular sizes to increase by a factor $(1+z)$ in every directions, as the source was closer to the observer when the light was emitted. Therefore the angular surface is increased by a factor $(1+z)^2$.

---

[1] For all the theoretical details, the reader is referred to Chapter 1 of Weinberg (2008)

This unavoidable loss of surface brightness becomes especially apparent at high redshifts: sources are dimmed by a factor 16 at $z = 1$, and by a factor of more than 625 at $z = 4$! Tolman dimming therefore gives us a biased view of the early universe, where only the most compact objects and the highest surface brightness regions of star formation can be detected (Keel 2007; Calvi et al. 2014). In the latter case, these regions can appear as separate sources, which affects number counts.

3.1.2 K-correction

Due to the expansion of the universe, we do not observe the light from high-redshift sources in the same rest-frame that it was emitted. At redshift z, a photon emitted at a frequency ν_e will be observed at a frequency ν_o, where $\nu_e = \nu_o(1 + z)$. When performing broadband photometric measurements, the apparent magnitude[2] of a source will be "redshifted" from its absolute magnitude[3]. Both quantities can therefore end up in different bands. For example, the flux of the most distant galaxies observed in the Hubble Deep Field (Williams et al. 1996), which sees the universe at UV and optical wavelengths, comes from the far-UV range of their emission spectrum. Converting the apparent magnitude of a source in band p to its true absolute magnitude in band q involves a term known as the "K-correction" (Humason et al. 1956; Oke and Sandage 1968; Hogg et al. 2002):

$$m_p = M_q + DM + K(z) \tag{3.5}$$

where DM is the *distance modulus*, or the difference (in magnitudes) between the observed bolometric flux of the source and bolometric flux observed at a distance of 10 pc, which is defined as:

$$DM = 5\log_{10}\left(\frac{D_L}{10\mathrm{pc}}\right) \tag{3.6}$$

and $K(z)$ is the K-correction term:

$$K(z) = -2.5\log_{10}\left[(1+z)\frac{L(\nu_e)}{L(\nu_o)}\right] \tag{3.7}$$

where $L(\nu_o)$ and $L(\nu_e)$ are respectively the observed and intrinsic SED of the source.

There are two main approaches for determining the K-correction for a given galaxy: the first approach involves the direct measurement of the galaxy spectrum. But in most cosmological surveys the only information available is the apparent flux in a set of broadband filters, i.e. a discretized and poorly resolved version of its underlying SED. The galaxy SED can be inferred by fitting the flux measurements to a set of empirical SED templates built from local galaxies, but the reliability of the resulting K-correction will then be limited by the number and width of the passbands used, as well as by the good match between the data and the fitted template (Westra et al. 2010).

As the determination of the luminosity function of a galaxy sample requires the estimation

---

[2] The *apparent magnitude* of an object in a given photometric passband (in the Vega system) is $-2.5\log_{10}(\mathrm{R})$, where R is the ratio of the observed flux of that source to the observed flux of the Vega star (Hogg, 1999).

[3] The *absolute magnitude* of an object is the apparent magnitude that the object would have if it were at a distance of 10 pc (Hogg, 1999).

of the absolute magnitude of the sources (cf Chapter 2.2), poor estimates of the K-correction can affect number counts by causing unphysical distortions to the luminosity function (Dalcanton 1998; Gabrielli et al. 2006).

3.2 Dust extinction and inclination

Interstellar dust affects the light of galaxies in the optical, as well as in UV and IR bands, as a result of the complex interplay between the distribution of stars and dust clouds, radiative transfer and interstellar chemistry[4]. Most of the dust of a “typical” galaxy is located within its disk. Bulges, in fact, might contain little to no dust at all[5] (Graham and Worley 2008). Due to the obscuring dust lanes, incoming light from a distant galaxy can be absorbed and scattered out before reaching the observer, an effect termed *dust extinction* or *dust attenuation* (see details on terminology in Section 4.5). Different morphological types of galaxies are affected differently by extinction: spirals are in fact more dimmed than ellipticals and lenticulars (Nagayama et al. 2004).

Dust attenuation is also correlated to the angle at which the galaxy is viewed, i.e. its inclination. The stronger the inclination, the longer the path length of incident photons through the disk of the galaxy, and the greater the extinction by the obscuring material. Extinction is therefore minimal when the galaxy is viewed face-on, and maximal when the galaxy is viewed edge-on, as illustrated in Figure 3-1. Observations confirm that edge-on galaxies tend to appear dimmer and redder than their face-on counterparts (Holmberg et al. 1975; Burstein et al. 1991; Driver et al. 2007).

As most observed quantities depend on inclination, cuts performed on these quantities can produce samples that are biased by inclination. For example, magnitude-limited surveys will inevitably miss highly inclined faint dusty galaxies (Masters et al. 2010; Devour and Bell 2016). Moreover, photometric quantities of galaxies affected by dust attenuation include optical apparent magnitudes, colors (e.g. Tully et al. 1998; Maller et al. 2009; Masters et al. 2010), and surface brightness profiles (e.g. Byun et al. 1994; de Jong 1996). Because the inner regions of disks contain more dust than the outer regions (Boissier et al. 2004; Popescu and Tuffs 2005), a radial opacity gradient arises which makes the observed disk scale lengths greater than their intrinsic values (Möllenhoff et al. 2006; Graham and Worley 2008, Gadotti et al. 2010). Yip et al. (2011) also showed that inclination effects result in overestimates of the photometric redshifts inferred from the observed magnitudes and colors of highly inclined sources, in order to match their redder color with empirical templates.

Over the last decades, a lot of work has been done on adjusting for the effects of dust attenuation quantitatively (e.g. Valentijn 1990; De Vaucouleurs et al. 1991; Giovanelli et al. 1995; Tully et al. 1998; Graham 2001; Masters et al. 2003; Tuffs et al. 2005; Bailin and Harris 2008; Unterborn and Ryden 2008), but the parameters governing dust attenuation are ill-constrained, and no consensus has yet been reached. More details on attempts at modeling the effects of dust extinction in various studies, including our own approach, are provided in Section 4.5.

---

[4] The reader is referred to Section 4.5 for a more detailed introduction on the physical properties of dust and its effect on the light and color of galaxies.

[5] However the light emitted in the bulge region can be obscured by the dust lanes in the disk when the galaxy is viewed edge-on.

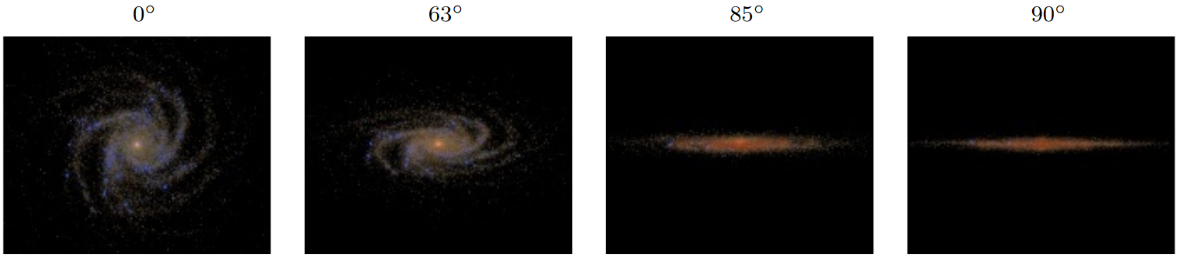


Figure 3-1 – Simulated *urz* color composite image of a dust attenuated Sbc galaxy at different inclinations, from face-on (i=0°) to edge-on (i=90°), generated using the Sunrise radiative transfer code (Jonsson 2006). As the inclination increases, reddening effects due to dust lanes become more and more apparent. Reproduced from Jonsson et al. (2009).

3.3 Malmquist bias

In apparent magnitude-limited surveys, intrinsically brighter sources can be observed at larger distances (i.e. at higher redshifts) than fainter ones. As a result, the observed mean absolute magnitude of the sample will be brighter than its intrinsic mean absolute magnitude, as illustrated in Figure 3-2, and the observed sample will not necessarily be representative of the real underlying population. This selection effect is referred to as *Malmquist bias*[6] (Malmquist 1920, 1922). It was originally discovered for stars spread in Euclidean space, but applies equally to galaxies in an expanding Friedman universe (Teerikorpi 2015). Malmquist bias operates at all magnitudes (Lynden-Bell 1992) and “only” affects the composition of the sample, not the photometric measurements performed on it. In cosmological studies, not taking it into account can lead to a significant overestimate of the average luminosity of standard candles, which in return leads to underestimate of their distance, and therefore to an overestimate of the Hubble constant (Sandage and Tammann 1975; Tammann 1987).

When the luminosity function of the studied population is a simple Gaussian distribution, an analytical correction correction for Malmquist bias can be derived, which depends on the square of the intrinsic dispersion in the absolute magnitude distribution of the sample σ_M (e.g., using the notations of Butkevich et al. 2005):

$$< M >_{lim} = M_0 - \sigma_M^2 \frac{\mathrm{d} \ln N(m_{lim})}{\mathrm{d} m_{lim}} \tag{3.8}$$

where $< M >_{lim}$ is the observed mean absolute magnitude of the magnitude-limited sample, M_0 is the intrinsic mean absolute magnitude of the population, and $N(m_{lim})$ is the total number of sources whose apparent magnitude is brighter than m_{lim}.

But correcting for this bias happens to be more highly complex in practice. In fact, the level of Malmquist bias depends on the shape of the intrinsic luminosity distribution of the population, which is unknown in general, and increases with the width of the intrinsic scatter in the source luminosities. This scatter can be difficult to determine, and must not be confused

---

[6] Malmquist bias bias was actually first discovered by Eddington (1914), in a special case where the spatial density of sources is homogeneous. Therefore, Malmquist bias is sometimes referred to in the literature as the Eddington-Malmquist bias, as suggested by Lynden-Bell (1992). But this terminology is likely to perpetuate the confusion between the Malmquist bias and the Eddington bias described in Section 3.4 that is already too common in the astronomical community.

with the observed scatter, which is rendered smaller than the intrinsic scatter by the Malmquist bias (Tammann 1987).

Methods for correction or minimization of the Malmquist bias include the creation of volume-limited samples, i.e. the sampling of every galaxies in the survey volume (e.g. Kannawadi et al. 2015), the weighting of galaxy properties by the maximum volume over which they can be detected (this is called the V_{max} method, cf Mo et al. 2010), or the comparison of realistic mock catalogs that explicitly include magnitude limits to mock catalogs that don't (e.g. Conley et al. 2010; Saulder et al. 2016). The curious reader is referred to Verter (1993) and Teerikorpi (1997) for historical and methodological reviews of various attempts at implementing Malmquist bias corrections.

3.4 Eddington bias

No measuring instrument is perfect. A set of sources with a single luminosity will always be spread out upon observation due to statistical fluctuations in measurement errors. This scatter gives rise to the *Eddington bias* (Eddington 1913, 1940), which distorts number counts in a systematic way. Thus, except for some pathological cases where the distribution is constant or linear, as galaxy number counts are rising steeply with decreasing flux, the number of low fluxes scattered up by the noise is likely to be greater than the number of large fluxes scattered down, which leads to a steepening of the slope of the number counts. Moreover, a fraction of the population is likely to be scattered below the detection limit of the survey, and another fraction below the detection limit to be scatter up upon observation (Jeffreys 1938). Eddington (1940) proposed a way to estimate analytically the true number counts from the distribution of observed values, provided that measurements errors are Gaussian-distributed with a constant width σ in magnitude, i.e. if the probability of observing a magnitude m' for a source of true magnitude m is:

$$P(m', m) = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-(m-m')^2/2\sigma^2} \tag{3.9}$$

Following the notations of Peterson (1997), let $A(m)$ be the true differential number count per magnitude per unit solid angle. $A(m)dm$ therefore represents the true number of sources between magnitudes m and $m + dm$, and $A_{\rm obs}(m)dm$ the observed number. The true number of sources brighter than m is then:

$$N(m) = \int_{-\infty}^{m} A(m)dm \tag{3.10}$$

and the observed value of $A_{obs}(m')$ is the convolution of the true number counts with the measurement PDF:

$$A_{obs}(m') = \frac{1}{(2\pi\sigma^2)^{1/2}} \int_{-\infty}^{\infty} A(m) e^{-(m-m')^2/2\sigma^2} dm \tag{3.11}$$

Assuming that σ is small and the cumulative number count is a power law:

$$N(m) = C10^{\kappa m} \tag{3.12}$$

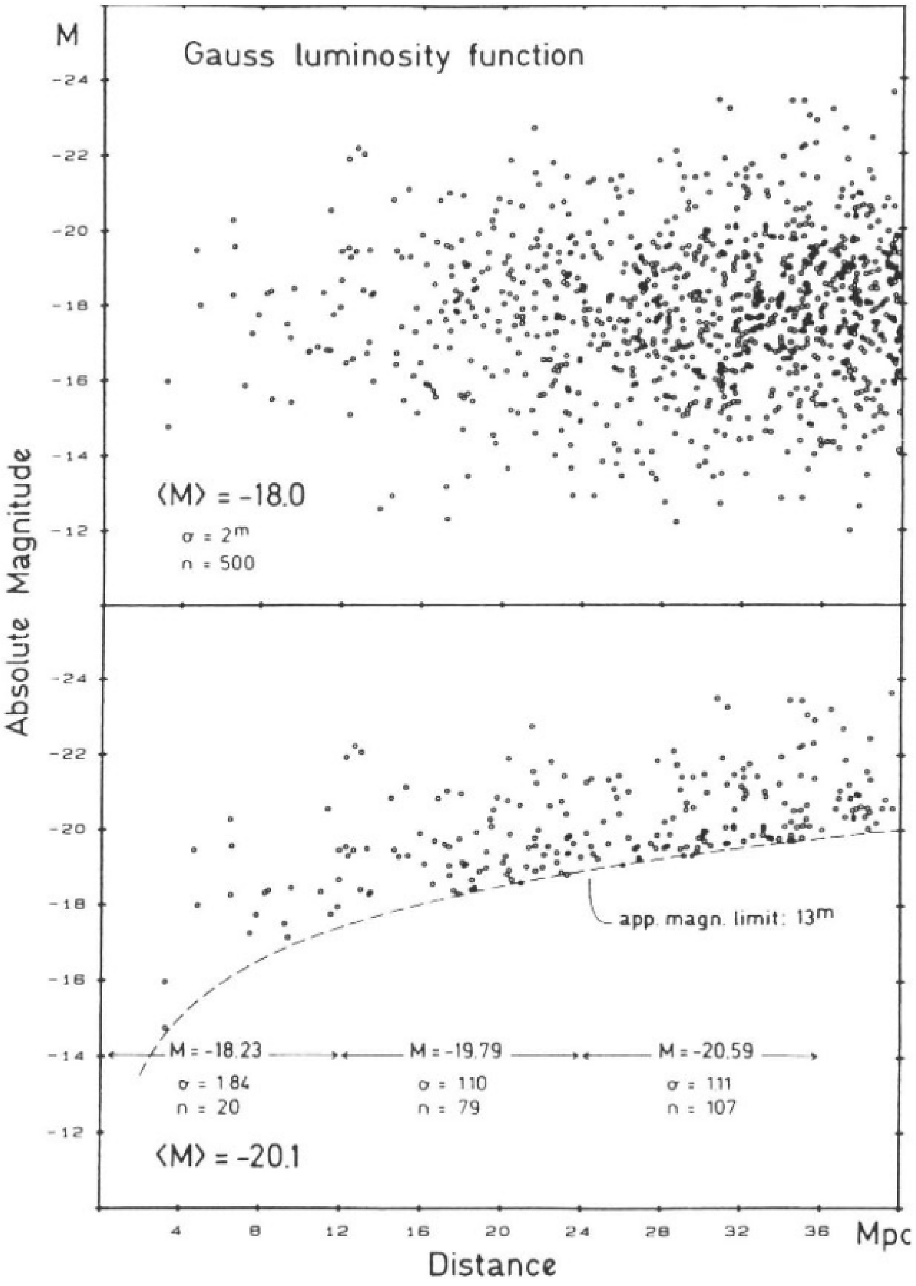


Figure 3-2 – Evolution of absolute magnitude as a function of distance from a simulated galaxy sample with a Gaussian luminosity function. As the volume surveyed expands with increasing distance, more and more objects are seen as distance increases. The survey limit in apparent magnitude gives rise to the so-called Malmquist bias, which leads to a increase of mean absolute magnitude of the sample with distance. Reproduced from Tammann (1987).

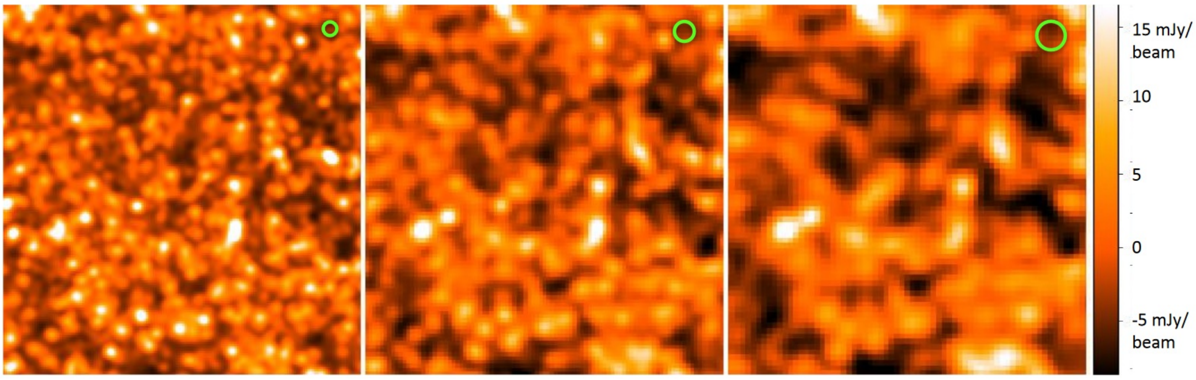


Figure 3-3 – Patch of the GOODS-N field (Dickinson et al. 2003) observed at 250, 350 and 500 μm (left to right). The green circles indicate the beam FWHM in each band. The field is plagued by source confusion, and the size of the background fluctuations is of the order of the beam size. Reproduced from Nguyen et al. (2010).

where C and κ are constant, Eddington showed that[7]:

$$A(m) = A_{obs}(m)\left[1 - \frac{1}{2}\left(\frac{\sigma\kappa}{\log_{10} e}\right)^2\right] \tag{3.13}$$

This correction must be applied at all magnitudes (Baryshev and Teerikorpi 2012). If the measurement scatter increases with fainter magnitudes, as it is often the case in galaxy surveys, Equation 3.11 cannot be used anymore, and an analytical correction becomes tricky to derive (van der Burg et al. 2010). In this situation, numerical experiments by Teerikorpi (2004) have shown that the steepening of the number counts gets even stronger than the theoretical prediction of Eddington (1940).

3.5 Confusion noise

The sky background is not empty. It consists of the accumulated light of faint unresolved sources below the detection limit of the survey, and of extended structures at scales of the order of the telescope beam[8] size (Guiderdoni et al. 1997). *Confusion noise* (or *source confusion*) is the corresponding uncertainty in the determination of the background level. These stochastic fluctuations of the background generate an uncertainty in the determination of the flux of resolved faint sources, which must be added to the flux error estimates: $\sigma^2_{tot} = \sigma^2_{sky} + \sigma^2_{conf}$, where σ^2_{sky} is the classic noise resulting from the combination of Poisson noise and Gaussian read-out noise. Source confusion also increases the positional uncertainty of sources (Hogg 2001), boosts the flux of some sources and can lead to spurious detections by the merging of several sources near the detection limit. It cannot be overcome by longer exposure times, and therefore poses a fundamental limit to the depth of astronomical surveys.

This problem was first mathematically approached by Scheuer (1957) and Condon (1974)

---

[7] The full derivation is available in Chapter 10 of Peterson (1997).

[8] The "beam" here refers to an element of resolution in the images. It can be defined as the solid angle of the 1σ radius circle of the Gaussian point spread function (Hogg 2001).

in the context of early radio surveys. In practice, the background sky brightness defines a limit beyond which individual sources cannot be clearly detected: the *confusion limit*. Near this limit, the distinction between faint objects and the background objects is blurred. The confusion limit is mainly determined by the beam size (as illustrated in Figure 3-3), the wavelength and the position on the sky. Using the notations of Dole et al. (2003), the confusion noise (at a given wavelength) can be defined in the general case as:

$$\sigma^2_{conf} = \int f^2(\theta,\phi) d\theta d\phi \int_0^{S_{lim}} S^2 \frac{dN}{dS} dS \tag{3.14}$$

where $f(\theta,\phi)$ is the two-dimensional beam profile[9], S is the source flux density and $\frac{dN}{dS}$ is the differential number count. The integration includes all sources below the confusion limit, S_{lim}. Moreover, studies by Negrello et al. (2004) and Takeuchi and Ishii (2004) have found that the level of source confusion is probably increased by the clustering of unresolved sources. The confusion noise can be evaluated by two criteria, described extensively in Dole et al. (2003): the "photometric criterion", which is related to the quality of the photometry of detected sources, and the "source density criterion", which is related to the completeness of detected sources above the confusion limit directly linked to the probability to lose sources that are too close to each other to be separated.

Confusion noise is not a problem in optical surveys, as the confusion limit is generally not reached, except in very crowded fields such as star clusters. But it can be the dominant noise contribution in mid and far-infrared, because the sky background is very bright in these wavelengths. In far-IR especially, the thermal emission of dusty Galactic cirrus clouds strongly contributes to confusion noise (Low et al. 1984; Gautier et al. 1992; Jeong et al. 2005). And at sub-mm wavelengths, due to the particular shape of the emission spectrum of dust, some high-redshift ultra luminous infrared galaxies (ULIRGs) actually compensate the combined effects of decrease in flux density with distance and surface brightness dimming, causing a "negative K-correction" which, coupled to the low instrumental angular resolution, also adds to source confusion (Blain et al. 2002).

There are two main approaches for correcting for the effects of source confusion, which both have their advantages and disadvantages (cf Roseboom et al. 2010). One approach involves the individual source counterparts identification in other wavelengths (e.g. Pope et al. 2005; Ivison et al. 2007). Another estimates statistically the galaxy number counts at faint magnitudes via a maximum likelihood fit (assuming a model of number counts, usually a simple power law) on the pixel intensity distribution (e.g. Patanchon et al. 2009).

---

[9] The beam profile can be approximated by a Gaussian profile with the same FWHM as the expected PSF, or by an Airy pattern (Dole et al. 2003).

Part II

The forward modeling approach to galaxy evolution

Chapter 4

Modeling the light, shape and sizes of galaxies

If you wish to make an apple pie from scratch, you must first invent the universe.

Carl Sagan, *Cosmos*

What I cannot create, I do not understand.

Richard Feynman

Abstract

In this chapter[1], I briefly review the tools used by the astronomical community to model the size, shape, color, brightness, and dust content of galaxies. This will allow me to outline our forward-modeling approach in this thesis and to introduce the STUFF empirical model that we are going to use to test the validity of our method. STUFF generates mock photometric and morphometric catalogs in an arbitrary number of filters, from empirical prescriptions about the light and size distributions of various galaxy populations.

Contents

---

[1] Note: This chapter uses extracts from Carassou et al. (2017) in some of its sections.

4.1 The forward modeling approach

The traditional approach when comparing the results of models to data is sometimes referred to as *backward modeling* (e.g., Marzke 1998; Taghizadeh-Popp et al. 2015). In this scheme, physical quantities are derived from the observed data, and are then compared with the physical quantities predicted from simulations, semi-analytical models (SAM), or semi-empirical models. A more reliable technique is the *forward modeling* approach: a distribution of modeled galaxies are passed through a virtual telescope with all the observing process reproduced (filters, exposure time, telescope characteristics, seeing properties, as well as the cosmological and instrumental biases described above), and a direct comparison is made between simulated and observed datasets. The power of this approach comes from the fact that theory and observation are compared in the observational space: ideally, the same systematic errors and selection effects affect the simulated and observed data.

The "mock catalogs" approach, as it is sometimes called, is already being used to test the processing pipeline of galaxy surveys and characterize their systematics (e.g., Yan et al. 2004; McBride et al. 2009), or to test models of galaxy formation (e.g., Simard and Somerville 2000). Blaizot et al. (2005) were the first to introduce realistic mock telescope images from light cones generated by semi-analytical models. Overzier et al. (2013) extended this idea by constructing synthetic images and catalogs from the Millenium Run cosmological simulation (Springel et al. 2005) including detailed models of ground-based and space telescopes. Recently, Bernyk et al. (2016) unveiled the Theoretical Astrophysical Observatory (TAO), an online virtual observatory that enables users to generate mock survey images and galaxy catalogs based on selectable combinations of cosmological simulations or semi-analytical model, as well as stellar population synthesis models. And Taghizadeh-Popp et al. (2015) used semi-empirical modeling to simulate Hubble Deep Field (HDF) images, from cutouts of real SDSS galaxies with modified sizes and fluxes, and compared them to observed HDF images. Lately, Kacprzak et al. (2017) used a forward model to infer the redshift distribution of a cosmological sample. In this thesis, we make the case that forward modeling can be used to perform reliable inferences on the evolution of the galaxy luminosity and size functions. I describe our approach in the following section.

4.2 Outline of our approach

In order to infer the physical properties of galaxies from observed survey images without having to describe the complex selection effects the latter contain, we propose the following pipeline. We start from a set of physical input parameters, drawn from the prior distribution defined for each parameter. These parameters describe the luminosity and size distribution of the various populations of modeled galaxies. From this set of parameters, our forward model generates a catalog of galaxies modeled as the sum of their bulge and disk components, each with a different profile. The projected light profiles of the galaxies are determined by their inclination, the relative fraction of light contained within the bulge, and the galaxy redshift as well as the extinction of the bulge and disk components. The galaxies are randomly drawn from the luminosity function of their respective population. The catalog assumes that galaxies are uniformly distributed on a virtual sky that includes the cosmological effects of an expanding universe with a cosmological constant. The survey image is simulated in every band covered by the observed survey, and reproduces all of its characteristics, such as filters transmission, exposure time, point spread function (PSF) model, and background noise model.

Then, a large number of "simulated" images are generated via an iterative process (a Markov chain) generating new sets of physical parameters at each iteration. Some basic flux and shape parameters are extracted in the same way from the observed and simulated images: after a pre-processing step (which is identical for observed and simulated data) where observables are decorrelated and their dynamic range reduced, the multidimensional distributions of simulated observables are directly compared to the observed distributions using a custom distance function on binned data.

The chain moves through the parameter space towards regions of high likelihood, that is, regions that minimize the distance between the modeled and observed datasets. The pathway of the chain is finally analyzed to reconstruct the multidimensional posterior probability distribution and infer the sets of parameters that most likely reproduce the observed catalogs, as well as the correlations between these parameters. The main steps of this approach are detailed in the following sections and chapters, and the whole pipeline is sketched in Figure 4-1.

4.3 Magnitudes and colors

We need to build mock galaxies whose apparent magnitudes and colors closely match current UV, optical and IR observations. In the AB system[2] (Oke and Gunn 1983[3]), the apparent magnitude of a source in a given band is:

$$m_{\mathrm{AB}} = -2.5\log_{10}\frac{\int f_\nu T_\nu d\nu}{\int T_\nu d\nu} - 48.6 \tag{4.1}$$

where f_ν is the spectral flux density of the source in erg s^{-1} cm^{-2} Hz^{-1} and T_ν is the total response of the system, which is the product of the filter transmission curve by the quantum efficiency of the detector, the response of the telescope (mirror and optics) and the atmospheric transmission, as illustrated in Figure 4-3. Color indices are then derived by subtracting the magnitudes in different bands.

We see from Equation (4.1) that our estimate of the apparent magnitude of a source relies on our knowledge of its spectral energy distribution (SED), which contains all the information about the integrated emission and absorption of stars, gas and dust. This requires the use of a set of templates corresponding to every galaxy spectral types, and to every galaxy components

---

[2] In the AB photometric system, the zero point is 3631 Jansky for all bands, so a source with a flat spectral energy distribution (f_ν = constant) has the same magnitude in all bands, and all colors = 0.

[3] Note that there is a sign error in the main formula of this paper.

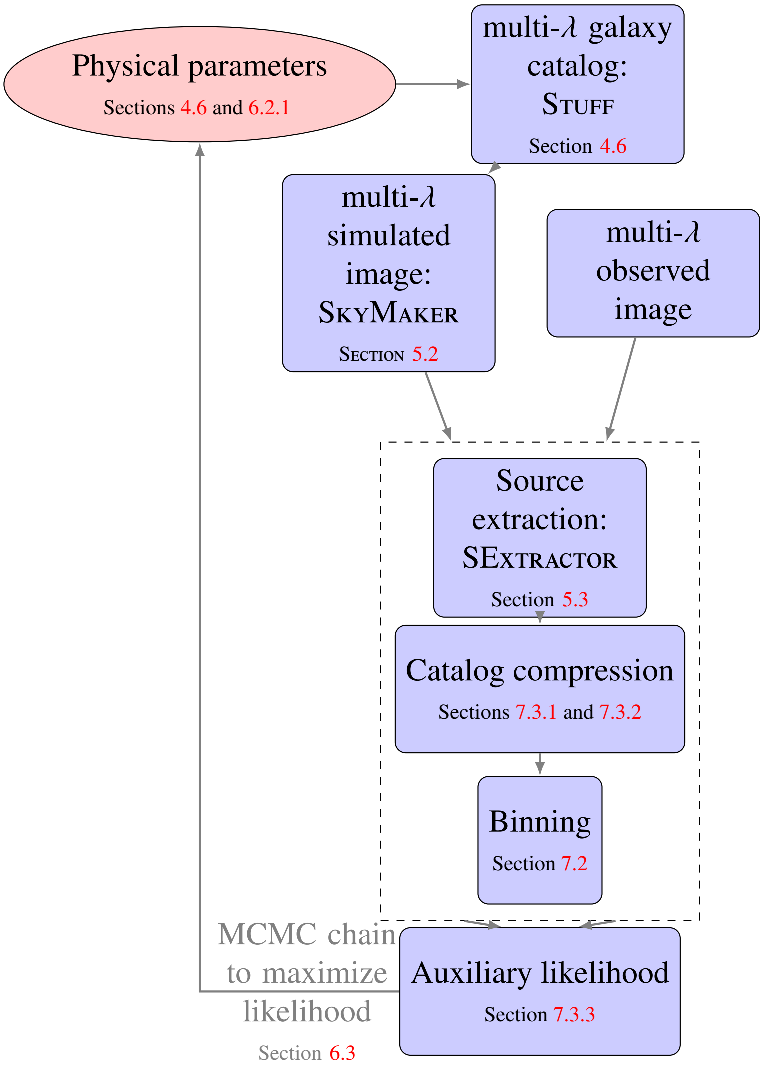


Figure 4-1 – Summary of the workflow

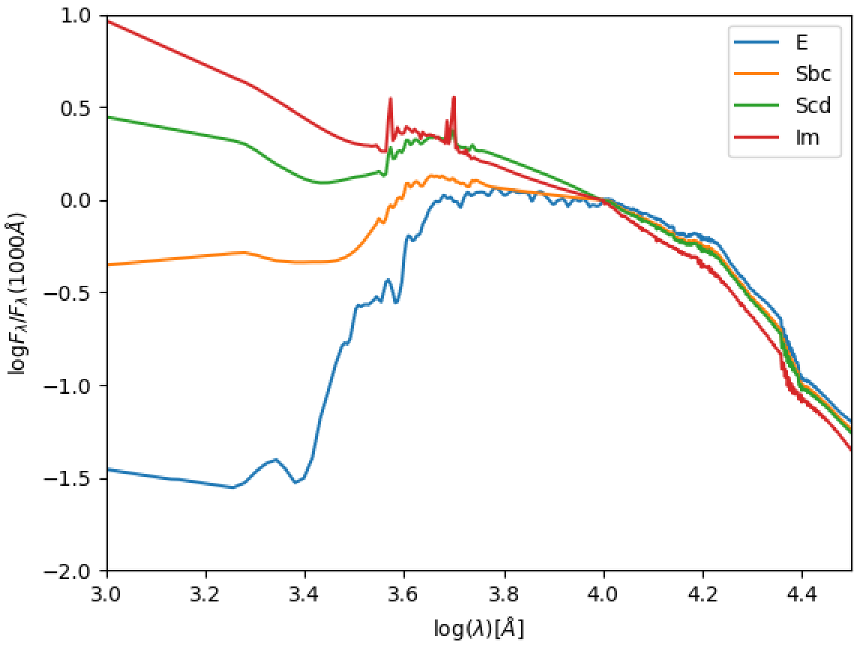


Figure 4-2 – SED templates of Coleman et al. (1980)

(bulge and disk). SED templates can be derived empirically from observations of local bright galaxies[4], such as the widely used set of Coleman et al. (1980) (illustrated in Figure 4-2) or Kinney et al. (1996), or they can be generated by models via stellar population synthesis (e.g., Fioc and Rocca-Volmerange 1999; Bruzual and Charlot 2003). The reader is referred to Walcher et al. (2011) for a recent review on SED modeling.

4.4 Light profiles

Galaxies display a large variety of sizes and shapes. A good parametrization of their structural and photometric properties is therefore crucial for classification and image simulation purposes. The light profile of a galaxy, which measures how its intensity varies as a function of the distance from its center, has proven a useful tool for that purpose. The most notable early attempt[5] at providing analytical fits to the radial distribution of stellar light was the work of de Vaucouleurs (1948, 1959), who found empirically that the intensity profile of bright elliptical galaxies and of the bulges of lenticular and spiral galaxies could be well described by:

$$I(r) = I_e \exp\left[-7.67\left(\left(\frac{r}{r_e}\right)^{1/4} - 1\right)\right] \tag{4.2}$$

---

[4] Empirical templates are computationally advantageous to use, but since they are built on local galaxies, they might not be representative of the spectra of high-z galaxies.

[5] An exhaustive historical review of early models to describe the light profile of elliptical galaxies can be found in Graham (2013).

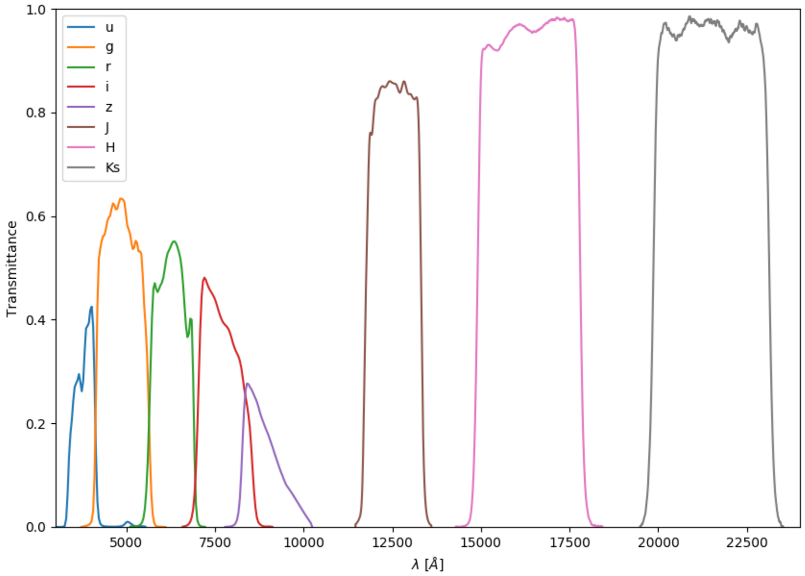


Figure 4-3 – Filter set of the CFHTLS+WIRDS survey (Cuillandre and Bertin 2006; Bielby et al. 2010). This plot displays the filter response curves (multiplied by the mirror+optics+CCD response) of the MegaCam *ugriz* optical bands, as well as of the WIRCam JHK_s infrared bands.

where r_e is the *effective radius*, i.e. the radius that encloses half of the total luminosity, and I_e is the intensity at r_e. The so-called "de Vaucouleurs profile" was later generalized by Sersic (1963), whose profile was so well suited for different types of galaxies and for different galaxy components that it is still widely used to this day. The Sérsic profile is defined as:

$$I(r) = I_e \exp\left[-b_n\left(\left(\frac{r}{r_e}\right)^{1/n} - 1\right)\right] \tag{4.3}$$

where n is the *Sérsic index* (the only free parameter of the model) and b_n a parameter dependent on n chosen so that r_e is the radius that encloses half of the total luminosity. b_n is defined such that $\Gamma(2n) = 2\gamma(2n, b_n)$, where Γ is the gamma function and γ is the incomplete gamma function, and it is usually approximated with good accuracy as $b_n \simeq 2n - 0.324$ for $1 \leq n \leq 15$ (Ciotti and Bertin 1999; Trujillo et al. 2001). The Sérsic index determines the shape of the light profile (cf Figure 4-4). The larger n is, the steeper the central core becomes, and the more extended are the outer wings. Setting $n = 4$ yields the de Vaucouleurs profile, and $n = 1$ yields an exponential disk profile (Freeman 1970).

In practice, galaxies are fitted with a sum of the Sérsic profile of their bulge and disk component, each with a different shape (cf Figure 4-5). This method allows for precise measurements of their flux, and of their bulge to total (B/T) flux ratio. The B/T ratio is strongly correlated to the morphological type of galaxies, at least in the local universe (e.g. de Lapparent et al. 2011, cf Figure 4-6), therefore the Sérsic profile offers a basis for automated classification. But it becomes complicated to fit Sérsic models at $z > 1$ because galaxies beyond that redshift tend to present mostly irregular and peculiar shapes (Abraham et al. 1996).

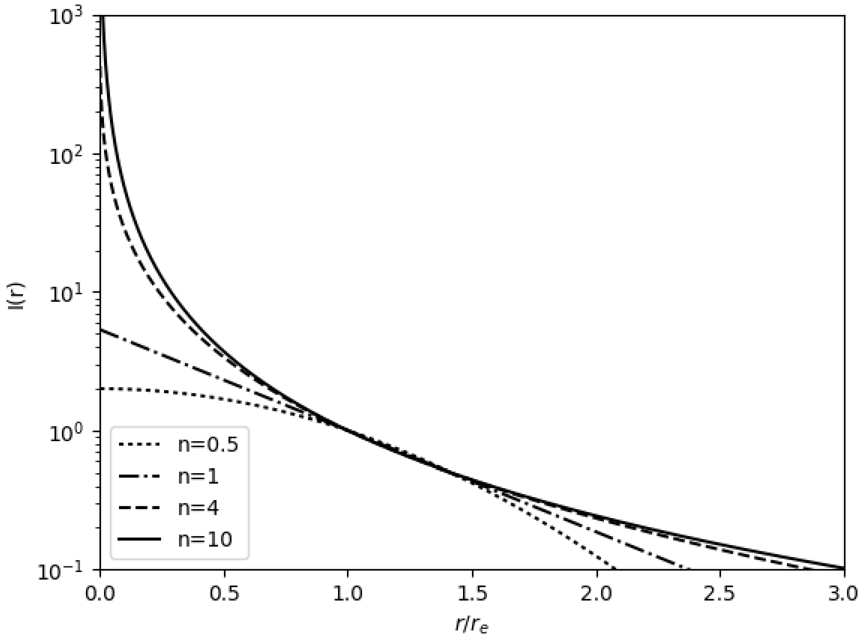


Figure 4-4 – Sérsic curves for various values of n. $n = 4$ corresponds to the de Vaucouleur profile.

For more accurate fits, the one-dimensional structural analysis provided by the Sérsic profile can be completed by a two-dimensional analysis including non-axisymmetric structures, such as bars, rings or spiral arms. For example, bars can be modeled as Sérsic profiles, spiral rotation functions can be used to model spiral structure, and ellipsoidal isophotes can be distorted using Fourier modes to model the shape of irregular galaxies (Peng et al. 2010). But fitting all these components at the same time can be computationally prohibitive when applying the fitting routine to a whole survey, due to the high number of free parameters (for each galaxy) of the resulting model.

Other popular measurements of galaxy structure include the CAS system (Concentration, Asymetry, Clumpiness, e.g. Conselice 2003), the shapelets decomposition (Refregier 2003) and the Gini/M20 parameters (e.g., Abraham et al. 2003; Lotz et al. 2004). The reader is referred to Conselice (2014), Graham (2013) and Peng et al. (2010) for historical and methodological reviews of parametric and non parametric galaxy structure analysis approaches.

4.5 Dust

Interstellar dust strongly affects the light we receive from galaxies. In fact, photons originating from distant galaxies have to go through their interstellar medium (ISM) and through the ISM of the Milky Way before reaching our telescopes. Dust was originally discovered via its obscuring effect on starlight, an effect termed *extinction* or *reddening*, as blue wavelengths are more dimmed than red wavelengths (Trumpler 1930). It constitutes only a small fraction of the total mass of galaxies: in the dust-rich ISM of the Milky Way, its mass represents ~ 1 % of the total HI mass (Draine et al. 2007), and this ratio is supposed to be fairly constant in external

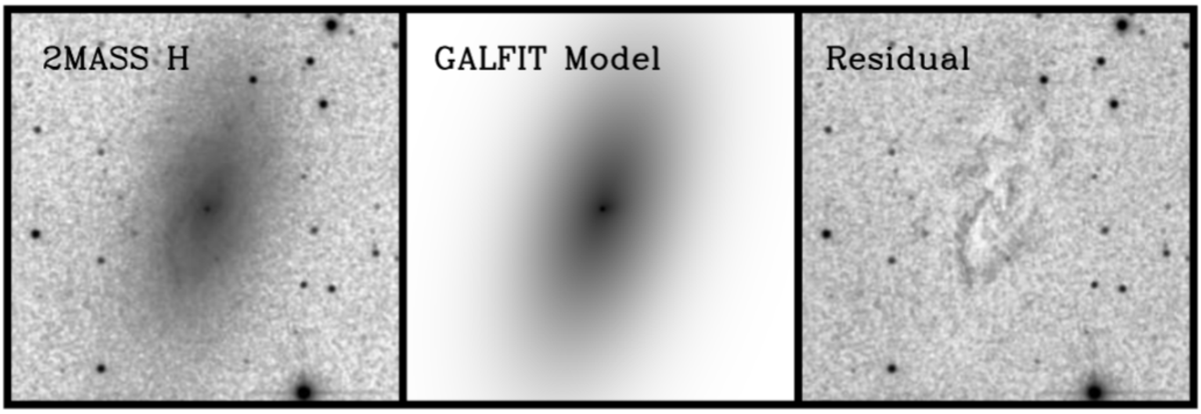


Figure 4-5 – Sérsic bulge+disk fit performed on a 5′ × 5′ H-band image of NGC 3621 with GALFIT (Peng et al. 2010). From Barth et al. (2008).

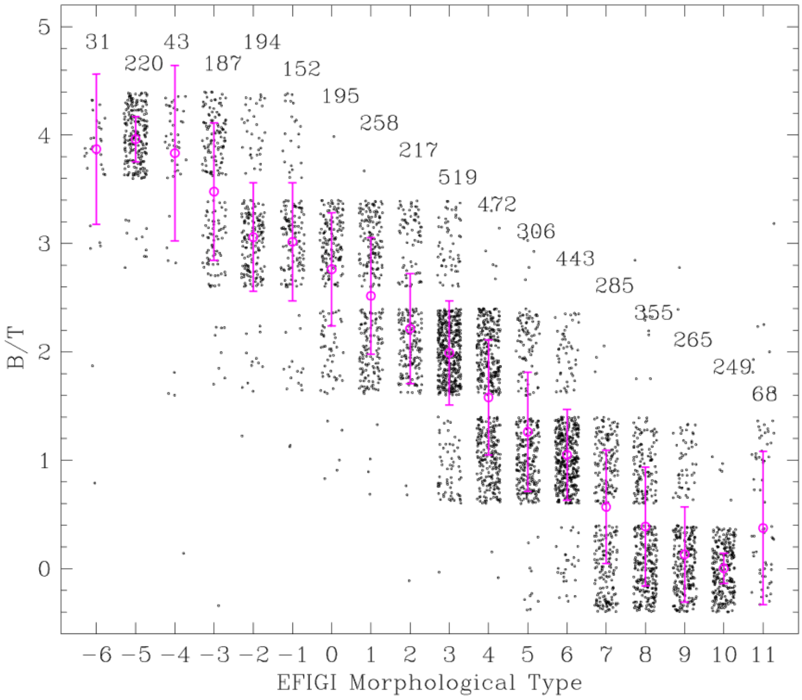


Figure 4-6 – Distribution of the B/T ratio attribute as a function of de Vaucouleurs morphological type (De Vaucouleurs et al. 1991) for the 4458 galaxies in the EFIGI catalog (Baillard et al. 2011). This plot reveals a strong correlation of morphological type with B/T, with a large dispersion of B/T for each type. From de Lapparent et al. (2011).

galaxies[6]. But the contribution of dust to the observed light emitted by galaxies is crucial to understand their evolution, as it shapes their spectra (and therefore their photometry) in various ways. It also plays a great role in the chemical evolution of galaxies, as the surface of dust grains constitute the craddle of molecular hydrogen (H2) formation and complex astrochemistry (e.g. Gould and Salpeter 1963; Cazaux and Spaans 2009). However, as we will see in this section, the dust evolution cycle in galaxies is still poorly constrained, and this constitutes a significant source of uncertainties in current models of galaxy evolution.

4.5.1 Composition, size and evolution

The physical properties of dust, from its chemical composition to its size distribution, are mostly constrained by the study of interstellar extinction, completed by observations in the infrared. In fact, dust grains absorb a fraction of the UV emission from young stars and re-emit in IR (Dale et al. 2009).

Interstellar dust is created from the elements ejected from low- and intermediate-mass stars in their asymptotic giant branch (AGB) phase and from type II supernovae (SNe), however stardust only account for a small fraction (a few %) of the total interstellar dust mass in galaxies (e.g. Draine and Salpeter 1979; Zhukovska and Henning 2014). Dust grains mainly grow by accretion and coagulation in the dense environment of molecular clouds, but can also be destroyed in the shock waves of SNe explosions, or shattered in collisions with other grains in the diffuse ISM (e.g. Asano et al. 2013; Aoyama et al. 2017).

The physical properties of dust grains vary significantly across galactic environments, which suggests various possible evolutionary tracks. But overall, its known chemical composition is a mixture of small grains of polycyclic aromatic hydrocarbons (PAHs), amorphous silicates and large grains of graphite. PAHs and silicates in particular display strong spectral features in the emission spectrum of dust. Dust grains come with all sizes, with a power law distribution for sizes between 0.02 μm and $\geq 0.2\,\mu m$ (e.g. Weingartner and Draine 2001; Draine 2009).

4.5.2 Galactic extinction

The combined effects of light absorption of a point source and scattering of photons out of the line of sight by the ISM of the Milky Way is referred to as *Galactic extinction* (or *interstellar extinction*). Given a uniform cloud of dust with an optical depth τ_λ^{ext} along the line of sight, an observer would see the original flux of the source $F_0(\lambda)$ reduced to:

$$F(\lambda) = F_0(\lambda)10^{-\tau_\lambda^{ext}} \tag{4.4}$$

The extinction A_λ (expressed in magnitudes) is defined as:

$$A_\lambda = -2.5\log_{10}\left(\frac{F(\lambda)}{F_0(\lambda)}\right) = -2.5\log_{10}(e)\tau_\lambda^{ext} = 1.086\tau_\lambda^{ext} = m(\lambda) - m_0(\lambda) \tag{4.5}$$

---

[6] However it is important to note that the dust-to-gas ratio of a galaxy depends on its star formation history because it correlates with metallicity (Draine et al. 2007).

In short, more dust along the line of sight leads to more dimming[7]. Extinction can also be quantified from the amount of reddening or *color excess* that the observed source is subject to. In the Johnson system, the color excess (also called *selective extinction*) is defined as the difference between the measured (B-V) color of the source, and its intrinsic color $(B-V)_0$ (i.e. not affected by extinction):

$$E_{B-V} = (B-V) - (B-V)_0 = (B-B_0) - (V-V_0) = A_B - A_V \tag{4.6}$$

where A_B is the total extinction in band B (450 nm), and A_V is the total extinction in band V (550 nm). In practice, galactic extinction is usually probed via *the pair method*, where a comparison is performed between the spectra of 2 stars with the same spectral type, one reddened and the other unreddened. The shape of the extinction law (the function describing how extinction evolves with wavelength) with different environments can be compared by normalizing A_λ by A_V. Another widely used normalization uses the quantity $k(\lambda) = A_\lambda / E_{B-V}$. Cardelli et al. (1989) derived an average extinction curve for $0.125\ \mu\text{m} \leq \lambda \leq 3.5\ \mu\text{m}$ parametrized by a single variable R_V, which is sometimes referred to in the literature as *the CCM relationship*:

$$< A_\lambda / A_V >= a(x) + b(x)/R_V \tag{4.7}$$

where $x = 1/\lambda\ \mu\text{m}^{-1}$, a and b are functions of wavelength and R_V is the *total-to-selective extinction ratio*, the slope of the extinction law at V-band, used to link extinction to the color excess:

$$R_V = \frac{A_V}{E_{B-V}} \tag{4.8}$$

The value of R_V is determined by the environment along the line of sight: a standard value of $R_V = 3.1$ is observed in the diffuse ISM of the Milky Way (e.g. Schultz and Wiemer 1975), but bigger values (~ 5.5) are found in dense molecular clouds (e.g. Fitzpatrick 2004). The CCM extinction curves for different values of R_V is plotted on Figure 4-7.

To compute Galactic extinction for a given line of sight and a given photometric band, we rely on the use of "dust maps". Schlegel et al. (1998) for example combined IRAS and COBE 100 microns data (shown in Figure 4-8) to derive an all-sky map of Galactic reddening, E_{B-V}. This map reveals the complex filamentary structure of the ISM of the Milky Way, even at high galactic latitudes, a structure known as *galactic cirrus*. An online service has been developed by NASA/IPAC Infrared Science Archive[8] to generate cutouts of this map from a given patch of the sky and return extinction values for most photometric bands.

4.5.3 Attenuation curves for local galaxies

When we look at a distant galaxy, we see the integrated emission of all its constituents over its surface. As a result of the relative position between the stars and the dust distributions, a fraction of the light is scattered by dust clouds into the line of sight, in addition to the effect of dust absorption and scattering out of the line of sight. In order to distinguish this phenomenon

---

[7] Hopefully we live in a relatively optically thin region of the Milky Way. What would astronomy be like if we were shrouded by a cloud of interstellar dust? We would only see the closest stars in the optical, and I wouldn't have a job!

[8] `http://irsa.ipac.caltech.edu/applications/DUST/`

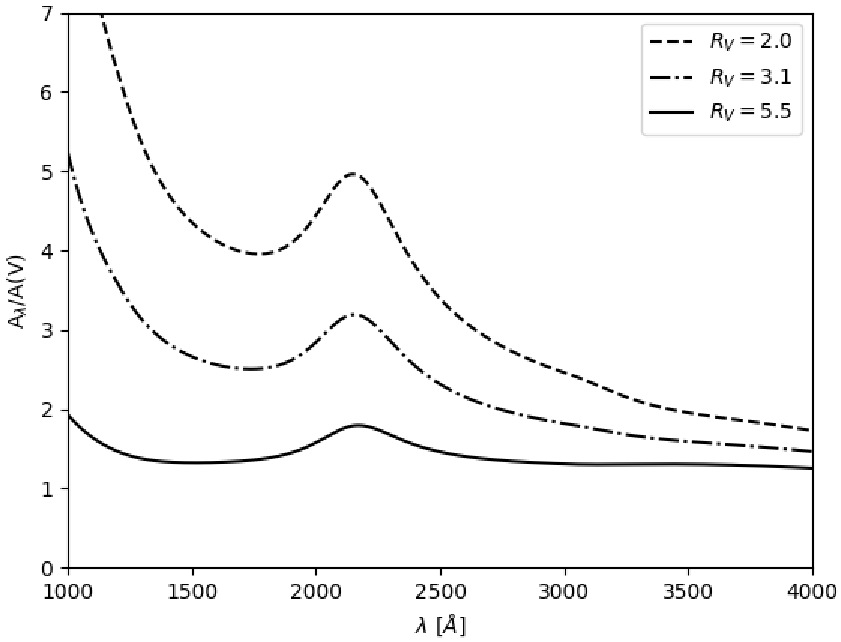


Figure 4-7 – Extinction curves for the Milky Way of Cardelli et al. (1989) for different values of R_V corresponding to different galactic environments. This plot illustrates how R_V controls the slope of the extinction curve.

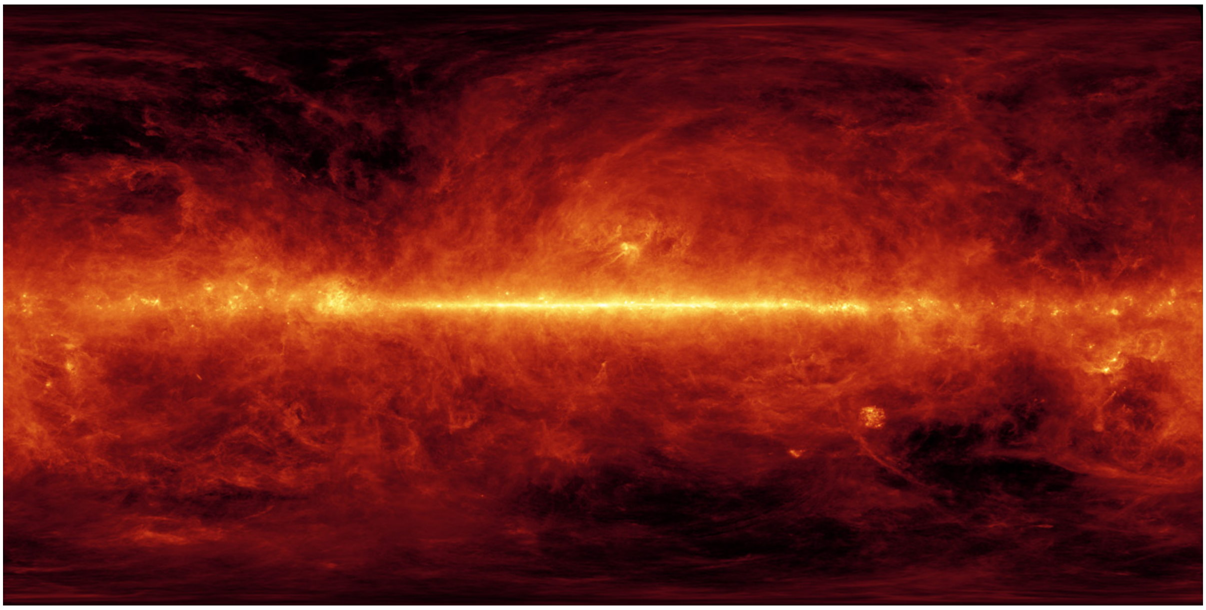


Figure 4-8 – All-sky view of the IRAS+COBE 100 μm imaging data in false colors shown in equirectangular projection, generated by Schlegel et al. (1998). It is used to return the values of Galactic dust reddening for a given line of sight.

from interstellar extinction, we refer to it as *dust attenuation*, although we will use the same formalism to describe attenuation and extinction. The difference between the 2 processes is illustrated in Figure 4-9. In practice, the scattering of photons into the line of sight dampens the effects of reddening compared to the extinction we observe within the Milky Way: an attenuated source appears greyer than an source subject to extinction only.

The shape of the attenuation law for a given galaxy depends on many factors: it is determined by the composition and size distribution of dust grain within that galaxy, and by the geometry of the extended distribution of stars and dust, which can be approximated by a simple screen of dust between the galaxy and observer (e.g. Charlot and Fall 2000) or by a clumpier distribution (e.g. Calzetti 2001). Attenuation also depends on the stellar populations of the host galaxy: in fact, young stars are strongly attenuated by their birth cloud, while older stars have cleared their birth cloud and are only attenuated by the diffuse ISM (Calzetti 1997). Some models, like CIGALE (Noll et al. 2009), use a parametric model to fit the attenuation curve $k(\lambda)$ empirically via SED fitting. More complex models, like GRASIL (Silva et al. 1998) or MAGPHYS (da Cunha et al. 2008), include a physically motivated treatment of dust with 2 components: a distribution of molecular clouds with a characteristic dissipation timescale ($\sim 10^7$ years) and a diffuse ISM.

If we want to derive extinction curves for extragalactic sources, we can only use the pair method in galaxies that are so close to us that we can resolve individual bright stars. This is the case for our closest neighbors, hence the most well-known UV extinction curves in the local universe besides the Milky Way (MW, Seaton 1979; Fitzpatrick and Massa 1986) are those of the Large Magellanic Cloud (LMC, Fitzpatrick and Massa 1986), the Small Magellanic Cloud (SMC, Prevot et al. 1984; Bouchet et al. (1985)) and M31 (Bianchi et al. 1996), and each curve has a distinct shape, as shown in Figure 4-10. In particular, the Magellanic Clouds curves cannot be fitted by the Galactic extinction laws of Cardelli et al. (1989), which suggests different dust properties in these galaxies than in the MW.

Some extinction curves present a characteristic bump at 2175 Å. This feature seems to hint that carbon is a major component of interstellar dust (e.g. Mathis et al. 1977). The bump can be observed in the MW curve, and in a lesser way in the LMC curve, but it has not been found in the SMC. Many uncertainties surround this feature. Its very presence (or not) in most extragalactic sources is still under debate: while radiative transfer dust models suggest that the 2175 Å feature should be common in star forming galaxies (Conroy 2013), other studies, like Battisti et al. (2016), do not find any evidence of a significant 2175 Å feature in their sample of 10^4 local galaxies from the Galaxy Evolution Explorer and the SDSS. Neither do we know whether the presence or not of the bump if an indicator of the spatial distribution of stars and dust, or that of peculiar dust properties.

In the general case, quantifying dust attenuation in external galaxies is a challenge because, unlike stars, there are no dust-free galaxies with the exact same properties to perform a direct comparison of spectra. However, in star-forming galaxies, HII regions heated by young stars produce some characteristic nebular emission lines, which are good tracers of the ionizing flux of the host stars, and which can be used to perform an indirect comparison of galaxy spectra. The flux ratio of the (easily detectable) Hα (6563 Å) and Hβ (4861 Å) recombination lines, called the *Balmer line ratio* (or *Balmer decrement*), is particularly useful because its value is set theoretically from quantum physics. Therefore, any departure from the expected value is attributed to dust attenuation: galaxies with high Balmer line ratios are considered dusty, whereas galaxies with low Balmer line ratios are considered dust-free. This method has been

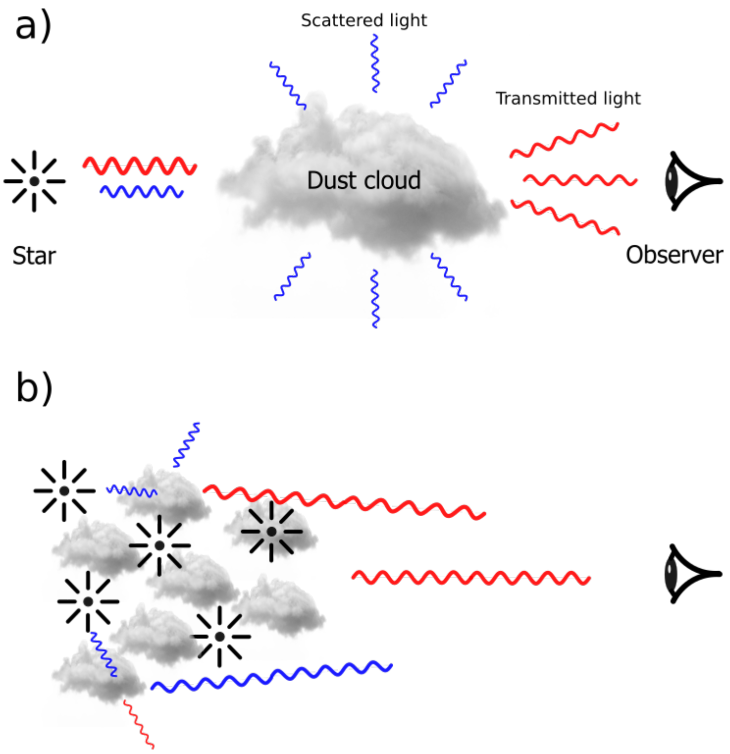


Figure 4-9 – Difference between dust extinction (a) and dust attenuation (b).

used for the first time by Calzetti et al. (1994, 2000) to derive an average attenuation law over the range 1200 Å ≤ λ ≤ 8000 Å from the UV-optical spectra of 39 local starburst and blue compact galaxies. This "Calzetti law" (which does not display any 2175 Å feature) is now used ubiquitously in galaxy evolution models, even if the small sample and the peculiar nature of these galaxies makes its use debatable.

4.5.4 Attenuation curves for high-z galaxies

The situation is even more complex at high redshift as dust, stellar evolution and metallicity give rise to elaborate degeneracies when trying to constrain the reddening of galaxies. Buat et al. (2011) have found from the study of a sample of $\sim 10^4$ galaxies with $0.9 < z < 2.2$ from the Chandra Deep Field-South Survey that the average attenuation law for high redshift galaxies can be well fitted by a modified Calzetti law with a steeper slope and a small 2175 Å bump (35% of the size of the MW bump). Buat et al. (2012) later found from a sub-sample of 751 galaxies at the same redshift range a UV bump amplitude similar to the one found in the super-shell region of the LMC. This claim has been confirmed by other studies, like Noll et al. (2007), who imply the presence of the 2175 Å bump from a sample of 108 massive, star-forming galaxies at $1 < z < 2.5$. But more recently, Scoville et al. (2015) analyzed a sample of 266 galaxies from the COSMOS survey with $2 < z < 6.5$. The resulting average attenuation curve is consistent with the LMC, MW or Calzetti laws, but does not display any significant 2175 Å bump.

Other studies tend to suggest that a universal dust attenuation curve might not be enough to explain the observations. In fact, there is now evidence of strong variations of attenuation curves with the spectral type (Kriek and Conroy 2013; Salmon et al. 2016). With these studies

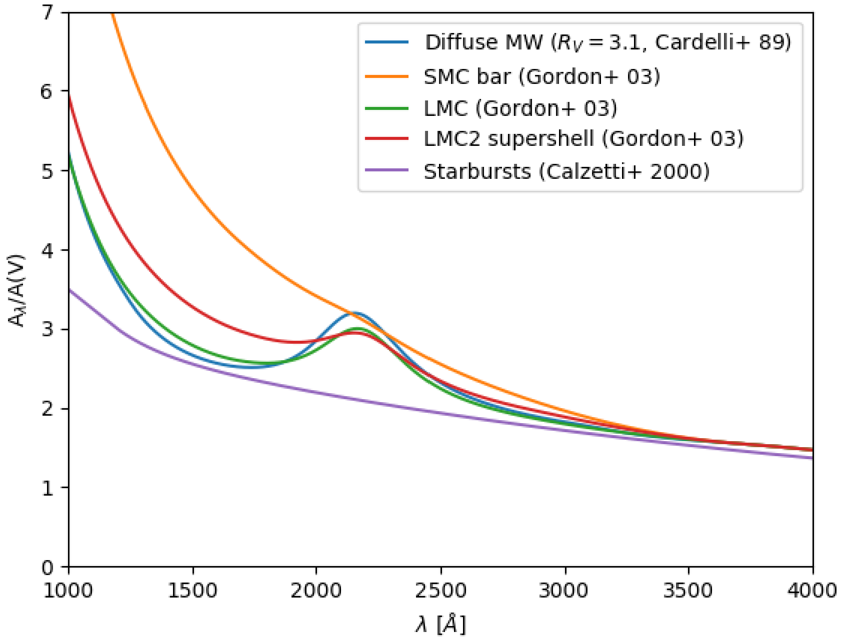


Figure 4-10 – Comparison of the most well known extinction and attenuation curves from the local universe.

in mind, we can safely say that there is no consensus yet reached over an appropriate attenuation law to choose when simulating the luminous properties of galaxies. To this date, that choice is left to the modeler, with the knowledge that this choice can lead to biases in his/her predictions.

4.6 The STUFF model

For this work, we will generate artificial catalogs with the STUFF package of Bertin (2009) in simulated fields of a given size. STUFF relies on empirical scaling laws applied to a set of galaxy “types”, which it uses to draw galaxy samples with photometric properties computed in an arbitrary number of observation passbands. Each galaxy type is defined by its Schechter (1976) luminosity function parameters, its spectral energy distribution (SED), as well as the bulge-to-total luminosity ratio B/T[9] and rest-frame extinction properties of each component of the galaxy through a “reference” passband (i.e. the filter where the LF is measured).

The photometry of simulated galaxies is based on the composite SED templates of Coleman et al. (1980) (cf Figure 4-2) extended by Arnouts et al. (1999). Any of the six “E”, “S0”, “Sab”, “Sbc”, “Scd”, and “Irr” SEDs can be assigned to the bulge and disk components separately, for a given galaxy type. The version of STUFF used in this work does not allow the SEDs to evolve with redshift; instead, following Gabasch et al. (2004), galaxy evolution is modeled as a combination of density (Schechter’s ϕ^*) and luminosity (Schechter’s M^*) evolution with

---

[9] In STUFF, B/T=0 corresponds to a bulgeless irregular galaxy, and B/T=1 corresponds to an elliptical galaxy.

redshift z:

$$M^*(z) = M^*(0) + M_e \ln(1+z) \tag{4.9}$$

$$\phi^*(z) = \phi^*(0)(1+z)^{\phi_e}, \tag{4.10}$$

where M_e and ϕ_e are constants. The reference filter will be set to the MegaCam g-band in the present thesis. Bulges and elliptical galaxies have a de Vaucouleurs (1953) profile:

$$\mu_b(r) = M_b + 8.3268\left(\frac{r}{r_b}\right)^{\frac{1}{4}} + 5\log_{10} r_b + 16.6337, \tag{4.11}$$

where $\mu_b(r)$ is the bulge surface brightness in mag.pc^{-2}, $M_b = M - 2.5\log_{10}(B/T)$ is the absolute magnitude of the bulge component and M the total absolute magnitude of the galaxy, both in the reference passband. As a projection of the fundamental plane[10], the average effective radius $\langle r_b \rangle$ in pc follows an empirical relation we derive from the measurements of Binggeli et al. (1984)):

$$\langle r_b \rangle = \begin{cases} r_{knee} 10^{-0.3(M_b - M_{knee})} & \text{if } M_b < M_{knee} \\ r_{knee} 10^{-0.1(M_b - M_{knee})} & \text{otherwise} \end{cases} \tag{4.12}$$

where $r_{knee} = 1.58h^{-1}$kpc and $M_{knee} = -20.5$. The intrinsic flattening q of bulges, defined as the ratio between the semi minor axis and the semi major axis of the ellipse, follows a normal distribution with $\langle q \rangle = 0.65$ and $\sigma_q = 0.18$ (Sandage et al., 1970), which we convert to the apparent aspect-ratio $\sqrt{q^2 \sin^2 i + \cos^2 i}$, where i is the inclination of the galaxy with respect to the line of sight.

Disks have an exponential profile:

$$\mu_d(r) = M_d + 1.8222\left(\frac{r}{r_d}\right) + 5\log_{10} r_d + 0.8710, \tag{4.13}$$

where $\mu_d(r)$ is the disk surface brightness in mag.pc^{-2}, $M_d = M - 2.5\log_{10}(1 - (B/T))$ is the absolute magnitude of the disk in the reference passband, and r_d is the effective radius. Semi-analytical models where disks originate from the collapse of the baryonic content of dark-matter-dominated halos (Dalcanton et al., 1997, Mo et al., 1998) predict useful scaling relations. Assuming that light traces mass and that there is negligible transport of angular momentum during collapse, one finds $r_d \propto \lambda L_d^{-\beta}$, where λ is the dimensionless spin parameter of the halo, $L_d = 10^{-0.4M_d}$ the total disk luminosity, and $\beta \simeq -1/3$ (de Jong and Lacey, 2000). The distribution of λ, as seen in N-body simulations, can well be described by a log-normal distribution (Warren et al., 1992), and is very weakly dependent on cosmological parameters (Steinmetz and Bartelmann, 1995), hence the distribution of r_d at a given M_d should behave as:

$$n(r_d|M_d) \propto \frac{1}{r_d}\exp\left[-\frac{\left(\ln(r_d/r_d^*) - 0.4\beta_d(M_d - M_d^*)\right)^2}{2\sigma_\lambda^2}\right]. \tag{4.14}$$

In de Jong and Lacey (2000), a convincing fit to I-band catalog data of late-type galaxies

---

[10] The fundamental plane is a narrow logarithmic plane linking the main observables of elliptical galaxies: their stellar mass, their size, and their velocity dispersion (Dressler 1987; Djorgovski and Davis 1987).

corrected for internal attenuation is obtained, with $\beta_d = -0.214$, $\sigma_\lambda = 0.36$, $r_d^* = 5.93$ kpc, and $M_d^* = -22.3$ (for $H_0 = 65\mathrm{km.s}^{-1}.\mathrm{Mpc}^{-1}$). Both bulge and disk effective radii are allowed to evolve (separately) with redshift z using simple $(1+z)^\gamma$ scaling laws (see, e.g., Trujillo et al., 2006, Williams et al., 2010). The original values from Trujillo et al. (2006) are modified to those in Table 9.6 based on the Hubble Space Telescope Ultra Deep Field (UDF, Williams et al. 2010, Bertin, private communication).

Dust attenuation is applied (separately) to the bulge and disk SEDs $S(\lambda)$ using the attenuation law from Calzetti et al. (1994), extended to the UV and the IR assuming an LMC law (Charlot, private communication):

$$S(\lambda) = S_0(\lambda) e^{-\kappa\tau(\lambda)}, \tag{4.15}$$

where $S_0(\lambda)$ is the face-on, unattenuated SED and $\tau(\lambda)$ the uncalibrated attenuation law. The normalization factor κ is computed by integrating the effect of attenuation A_{ref}, expressed in magnitudes, within the reference passband $p_{\mathrm{ref}}(\lambda)$:

$$A_{\mathrm{ref}} = -2.5\log_{10}\frac{\int p_{\mathrm{ref}}(\lambda) S_0(\lambda) e^{-\kappa\tau(\lambda)} d\lambda}{\int p_{\mathrm{ref}}(\lambda) S_0(\lambda) d\lambda}. \tag{4.16}$$

As the variation of $\tau(\lambda)$ is small within the reference passband, we take advantage of a second order Taylor expansion of both the exponential and the logarithm:

$$\begin{aligned} A_{\mathrm{ref}} &\approx -2.5\log_{10}\left(1 - I_1\kappa + \frac{1}{2}I_2\kappa^2\right) && (4.17)\\ &\approx 1.086\left(I_1\kappa + \frac{I_1^2 - I_2}{2}\kappa^2\right), && (4.18)\end{aligned}$$

with

$$I_1 = \frac{\int p_{\mathrm{ref}}(\lambda) S_0(\lambda)\tau(\lambda) d\lambda}{\int p_{\mathrm{ref}}(\lambda) S_0(\lambda) d\lambda}, \quad I_2 = \frac{\int p_{\mathrm{ref}}(\lambda) S_0(\lambda)\tau^2(\lambda) d\lambda}{\int p_{\mathrm{ref}}(\lambda) S_0(\lambda) d\lambda}. \tag{4.19}$$

Solving the quadratic equation (4.18) we obtain:

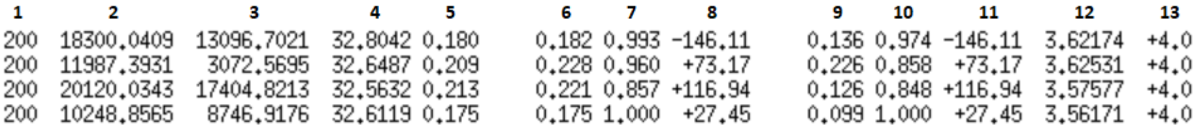
$$\kappa \approx \frac{-2A_{\mathrm{ref}}}{1.086\left(I_1 + \sqrt{I_1^2 - \frac{2}{1.086}(I_1^2 - I_2)A_{\mathrm{ref}}}\right)}. \tag{4.20}$$

We adopt the parametrization of the attenuation from the RC3 catalog (de Vaucouleurs et al., 1991):

$$A_{\mathrm{ref}} = -\alpha(\mathrm{T})\log_{10}(\cos i), \tag{4.21}$$

where i is the disk inclination with respect to the line-of-sight, and $\alpha(\mathrm{T})$ (not to be confused with Schechter's α) is a type-dependent "attenuation coefficient" that quantifies the amount of extinction+diffusion in the blue passband. The attenuation coefficient evolves with de Vaucouleurs (1959) revised morphological type as:

$$\alpha(T) = \begin{cases} 1.5 - 0.03(\mathrm{T}-5)^2 & \text{for T} \geq 0 \\ 0 & \text{for T} < 0 \end{cases} \tag{4.22}$$



```
1      2          3          4       5        6     7      8         9     10      11      12      13
200  18300.0409 13096.7021 32.8042 0.180    0.182 0.993 -146.11    0.136 0.974 -146.11 3.62174  +4.0
200  11987.3931  3072.5695 32.6487 0.209    0.228 0.960  +73.17    0.226 0.858  +73.17 3.62531  +4.0
200  20120.0343 17404.8213 32.5632 0.213    0.221 0.857 +116.94    0.126 0.848 +116.94 3.57577  +4.0
200  10248.8565  8746.9176 32.6119 0.175    0.175 1.000  +27.45    0.099 1.000  +27.45 3.56171  +4.0
```

Figure 4-11 – Example of a Stuff output catalog, with 4 galaxies generated in a given photometric band. 1 : object type index (200=galaxy, 100=star). 2-3 : galaxy position in pixels (x,y). 4 : apparent magnitude. 5 : B/T ratio. 6 : bulge radius (arcsec). 7 : projected bulge axis ratio. 8 : bulge position angle (degrees, with respect to x-axis). 9 : disk scale-length (arcsec). 10 : disk aspect ratio (disk inclination). 11 : disk position angle. 12 : redshift. 13 : de Vaucouleurs type.

Stuff applies to SEDs the mean intergalactic attenuation curve at the given redshift following Madau (1995) and Madau et al. (1996), using the list of Lyman wavelengths and absorption coefficients from the Xspec code (Arnaud, 1996). Galaxies are Poisson distributed in $5h^{-1}$ Mpc redshift slices from $z = 20$ to $z = 0$. For now the model does not include clustering properties, therefore the galaxies positions are uniformly distributed over the field of view. Ultimately Stuff generates a set of mock catalogs (one per filter) to be read by the image simulation software, containing source position, apparent magnitude, B/T, bulge and disk axis ratios and position angles, and redshift (cf Figure 4-11). We note that for consistency, we kept most of the default values applied by Stuff to scaling parameters, although many of them come from slightly outdated observational constraints dating back to the mid-2000's (and even earlier). This of course does not affect the conclusions of this work.

Chapter 5

Painting imaginary skies

> *We all know that Art is not truth. Art is a lie that makes us realize truth at least the truth that is given us to understand. The artist must know the manner whereby to convince others of the truthfulness of his lies.*
>
> ---
>
> Pablo Picasso

Abstract

Being able to reproduce realistic images of astronomical surveys and to retrieve accurate photometric and morphometric catalogs from them is at the heart of the method we are laying out. In this chapter, I draw up an inventory of currently used image simulators, and provide a detailed description of the SkyMaker and SExtractor packages, that we will use for image generation and source extraction respectively.

Contents

5.1 Comparison of image simulators

Simulations of astronomical images from surveys are an essential component in the development of any survey. A lot of simulation softwares have been developed to reproduce images of present and future surveys for various purposes, from testing data analysis pipelines to reproducing the expected subtle effects of weak gravitational lensing. They all differ in their rendering techniques of the source light profiles and morphologies, as well as their modeling of the instrumental and noise properties. This affects the realistic nature of the images generated along with the speed of execution of the softwares. However, no comparative study has yet been published to assess their relative strengths and weaknesses. The list below offers a first attempt of a qualitative comparison of most of the survey image simulators. These simulators include:

- **SkyMaker** (Bertin 2009, open source[1]), which was first introduced as a testing tool for SExtractor (Bertin and Arnouts 1996). SkyMaker uses the PSF from observed data to generate realistic images containing galaxies that are modeled as a combination of their bulge and disk Sérsic profiles, with realistic instrumental, atmospheric and noise properties (cf Section 5.2).

- The **GalSim** software (Rowe et al. 2015, open-source[2]) can render sources via several methods (drawing in real space, via discrete Fourier transform or via photon shooting using the surface brightness model as a probability distribution function). Contrarily to SkyMaker, GalSim can simulate weak lensing effects and realistic galaxy morphologies using cutouts from HST data, however other effects such as pixel saturation are not modeled. According to Emmanuel Bertin (private communication), GalSim is 10× slower than SkyMaker when run in the same conditions.

- The **Ultra Fast image generator** (UFig) (Bergé et al. 2013) produces the fastest simulations to date[3], at the cost of modeling galaxies as pure Sérsic profiles only (i.e. no bulge+disk decomposition). However its speed depends on the exposure time of the survey reproduced because of the photon shooting technique used.

- The **Simage** IDL pipeline (Dobke et al. 2010) generates simulated images able to reproduce the complex morphologies of high redshift galaxies in the B, V, i and z bands using shapelet decomposition (Refregier 2003), with the possibility to add a weak gravitational lensing signal. According to Bergé et al. (2013), Simage is slower in execution than Skymaker and UFig.

- The **SHERA** software (Mandelbaum et al. 2012) simulates images of ground based instruments using space based data of higher resolution (galaxy postage stamp images) from the COSMOS survey (Scoville et al. 2007), with the option of including weak lensing shear using shapelet decomposition.

- **SkyLens** (Meneghetti et al. 2008) generates realistic morphologies rendered via shapelet decompositions of galaxy images from GOODS data (Giovanelli et al. 1995), and incorporates gravitational lensing using ray tracing methods.

---

[1] `http://www.astromatic.net/software/skymaker`

[2] `https://github.com/GalSim-developers/GalSim`

[3] According to Emmanuel Bertin (private communication), Skymaker is probably as fast as UFig when compiled with the Intel(R) C++ compiler (icc).

- The **Photon Simulator** (PhoSim) (Peterson and Jernigan 2013, open source[4]) reproduces images of the Large Synoptic Survey Telescope (LSST, Ivezic et al. 2008), but the code could be extended to other surveys in the future. The virtual photons are collected by the telescope via Monte Carlo sampling, and galaxies are modeled as bulge+disk Sérsic profiles. PhoSim is the most realistic simulator to date, as it implements detailed a physical description of atmospheric effects such as dispersion, cloud absorption and molecular scattering. However it is also the slowest simulator in the list.

Considering its ability to produce realistic images in a relatively short time and the access to the source code for easy bug fixing and updates implementations, we decide to use the SkyMaker software for our goals. In the next section, I describe the inner workings of the code.

5.2 The SkyMaker software

SkyMaker can generate simulated images of the universe as if passed through a virtual telescope. It therefore has the ability to reproduce the observed data from most astronomical surveys with great accuracy (see e.g., Figure 5-1). The user eager to reproduce a given survey can provide its characteristics, which include the exposure time, the filters used[5], the background level, the pupil features such as the diameter of the primary and secondary mirror, and the detector characteristics (gain, image size, full-well capacity of pixels, zero-point, saturation level). These parameters can be found in the technical documentation of the survey reproduced.

The light and shape characteristics of input galaxies are provided in the form of an ASCII catalog of sources (one for each band) generated by an empirical model such as Stuff (cf Section 4.6) or EGG (Schreiber et al. 2017). Bulges follow a de Vaucouleurs light profile (Equation (4.11)), and disks an exponential profile (Equation (4.13)). Stars are generated by convolution of the PSF with the pixel grid. Blooming artifacts at the center of bright stars (creating vertical trails on the images), characteristic of pixel saturation, are also reproduced. *In fine*, the simulated images in each band are provided in the FITS format.

5.2.1 Modeling the Point Spread Function

The PSF is the response of an optical instrument to an impulse. Its good estimation is crucial for accurate measurements of the photometry and shape of galaxies. The PSF model can be generated internally by SkyMaker, as a convolution of the instrumental point spread function with the atmospheric seeing, tracking errors and optical aberrations. For more realistic renders, the PSF can be extracted from images of stars in observed data, modeled as a two-dimensional combination of basis vectors and fitted by χ^2 minimization using the PSFEx software of Bertin (2011a), which returns a FITS file. This operation can be carried out in several equally-spaced areas of the field to accurately characterize the PSF variations over the field (cf Figure 5-2).

Beyond $\sim$ 10 FWHM of the center, the wings of the PSF are dominated by diffusion, adding an "aureole" component to it, as illustrated in Figure 5-3 (Racine 1996). This component, first discovered by de Vaucouleurs (1958) and probably due to micro-ripples on optical surfaces,

---

[4] `https://bitbucket.org/phosim/phosim_release/wiki/Home`

[5] A wide variety of photometric bands can be chosen within the range $[0.29, 87.74831]$ μm (Annunziatella et al. 2013).

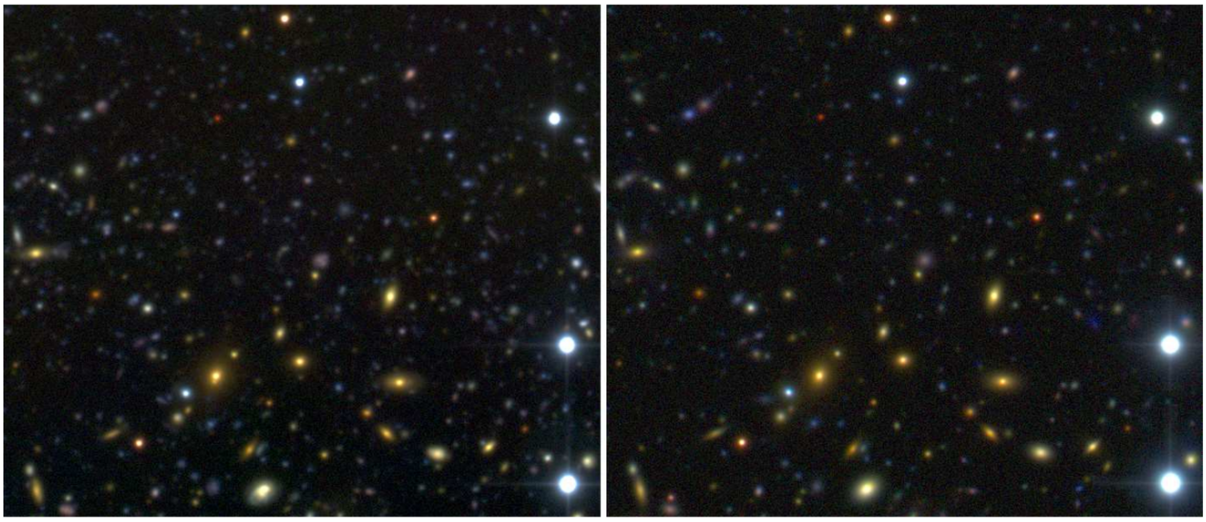


Figure 5-1 – *Left*: a small region of the Canada-France-Hawaii Telescope Legacy Survey (CFHTLS, Cuillandre and Bertin 2006) D1 field (stack from the 85% best seeing exposures) in *gri* bands. *Right*: Full reconstruction of the image on the left using STUFF+SKYMAKER with the same filters, exposure time, and telescope properties as the CFHTLS data. Both images are shown with the same color coding. From Bertin (2009).

scattering on atmospheric aerosols, and dust on mirrors (e.g. Beckers 1995), follows a King radial intensity profile in r^{-2} (King 1971). However, including the aureole component to the PSF can significantly increase the image simulation time due to the large angular sizes involved, especially when generating very bright sources.

5.2.2 Modeling the noise properties

The sky background in a SKYMAKER-generated image is uniform, and its surface brightness is user-defined via the BACK_MAG input parameter (in mag.arcsec^{-2}). The background noise is a combination of Gaussian read-out noise and Poissonian shot noise, defined below. Correlation of the noise between adjacent pixels is neglected for now. This feature will be released in a future version of the code (cf Section 10.2.1).

Read-out noise

In any CCD camera, incident photons knock-out electrons on the sensor via photoelectric effect (Einstein 1905). After passing through an amplifier, the charge of the free electrons is measured in every pixel and converted to a voltage, which is itself converted to Analog-to-Digital Units (ADU). But our instruments are not perfect. Gaussian random fluctuations affect the amplification, measurement and conversion processes, adding a noise component to the images: the *read-out noise*. Read-out noise depends on the camera, therefore its value usually appears in the technical documentation of astronomical surveys. It is also independent of exposure time and is applied uniformly across the detector.

Shot noise

Photon arrival on a CCD detector is a random and independent process. The brighter the

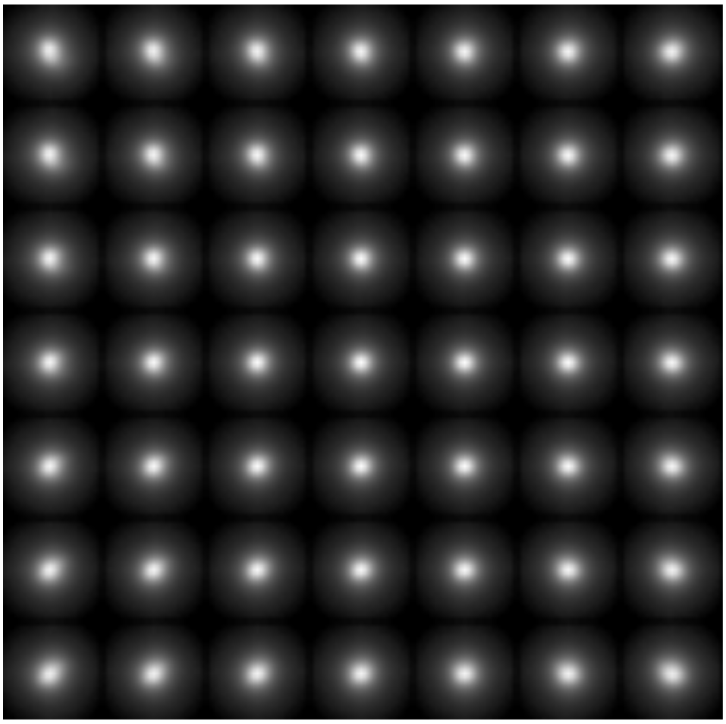


Figure 5-2 – PSF variations over the CFHTLS D4 field (1×1 deg^2) observed in r-band, as output by PSFEx. From Bertin and Moneti (2017).

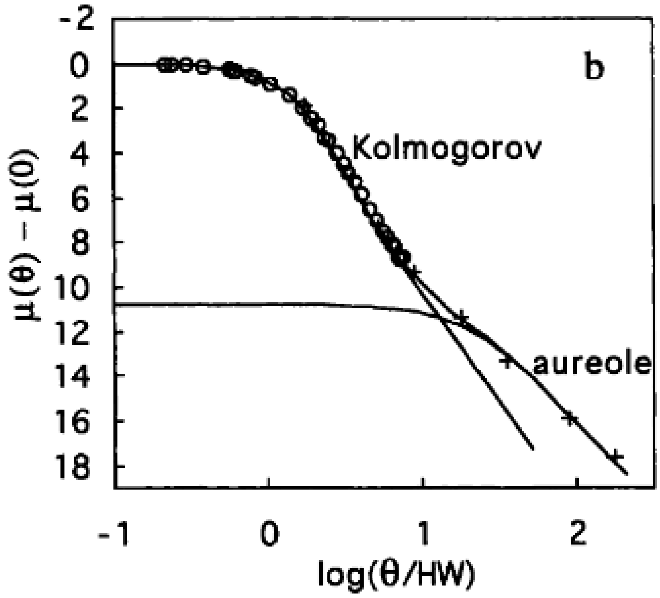


Figure 5-3 – Observed PSF surface brightness profile from CCD data fitted by the sum of a Kolmogorov profile (Fried 1965), which provides an accurate fit for the inner regions of the PSF dominated by atmospheric turbulence, and an aureole profile. The aureole component dominates the surface density profile beyond ~ 10 FWHM of the center. From Racine (1996).

emitting source, the higher the number of photons per second that reaches the sensor. A given pixel of the detector will perceive an average flux with some statistical fluctuations following a Poisson law (Poisson 1837) of parameter p, where p is the number count of the expected signal in a given pixel. This added noise component is called shot noise (or photon noise), and depends on the signal from the sources or the background.

5.3 Extracting sources from images: the SExtractor software

In order to derive photometric and morphometric catalogs from observed and simulated survey images, we use the popular SExtractor package of Bertin and Arnouts (1996). While other softwares have been developed for this purpose such as, e.g., DAOPHOT (Stetson 1987) or the more recent Tractor code (Lang et al. 2016), SExtractor's ability to return accurate photometry of galaxies up to very faint magnitudes make it the most widely used source extractor in the astronomical community (Annunziatella et al. 2013).

In SExtractor, sources are detected in four main steps, as illustrated in Figure 5-4: first, a smooth model of the image background is computed and subtracted. Second, a convolution mask, acting as matched filter, is applied to the background-subtracted image for improving the detection of faint sources. Third, a segmentation algorithm identifies connected regions of pixels with a surface brightness in the filtered image higher than the detection threshold. Finally, the same segmentation process is repeated at increasing threshold levels to separate partially blended sources that may share light at the lowest level.

Once a source has been detected, SExtractor performs a series of measurements according to a user-defined parameter list. This includes various position, shape, and flux estimates. For this work we rely on FLUX_AUTO photometry (cf Figure 5-5). FLUX_AUTO is based on Kron's "first moment" algorithm (Kron, 1980) and gives reasonably robust photometric estimates for all types of galaxies. For object sizes we choose the half-light radius estimation provided by the FLUX_RADIUS parameter, which is the radius of the aperture that encloses half of the FLUX_AUTO source flux. We note that this size estimate includes the convolution of the galaxy light profile by the PSF. In order to retrieve properties such as color, SExtractor is run in the so-called double image mode, where detection is carried out in one image and measurements in another. By repeating source extraction with the same "detection image", but with "measurement images" in different filters, we ensure that the photometry is performed in the exact same object footprints in all filters.

SExtractor flags all issues occurring during the detection and measurements processes in the output catalogs. These issues are summarized in Table 5.1. Flag indices can add up if the source extraction encounters more than a problem. In this work, we consider only detections with a SExtractor FLAG parameter < 4, which excludes sources that are saturated or truncated by the frame boundaries. It is important to stress again again that simulated images must go through the same source extraction process as observed images to ensure their comparability.

Table 5.1 – SExtractor flags

Flag index	Meaning
1	The object has neighbors, bright and close enough to significantly bias the photometry, or bad pixels (more than 10% of the integrated area affected)
2	The object was originally blended with another one
4	At least one pixel of the object is saturated (or very close to)
8	The object is truncates (to close to an image boundary)
16	Object's aperture data are incomplete or corrupted
32	Object's isophotal data are incomplete or corrupted
64	A memory overflow occurred during deblending
128	A memory overflow occurred during extraction

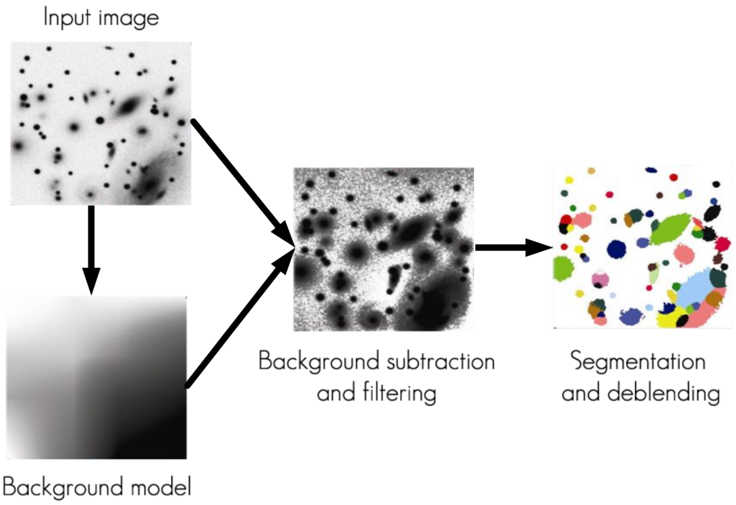


Figure 5-4 – The four steps of source detection in SExtractor. From a talk given by Emmanuel Bertin.

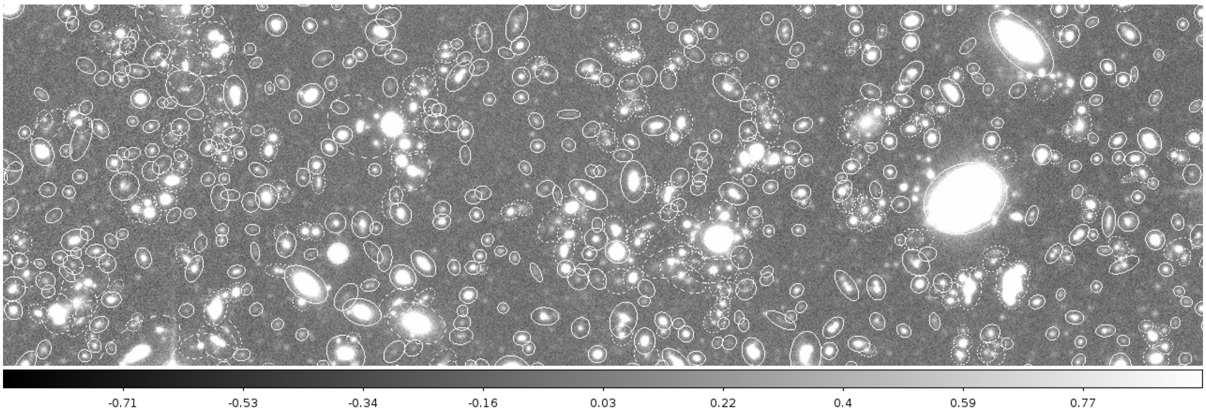


Figure 5-5 – Sources detected by SExtractor in a portion of the CFHTLS D1 field in g-band. The ellipses show the limits of integration for the FLUX_AUTO property of extracted sources.

Part III

Elements of Bayesian indirect inference

Chapter 6

Exploring the parameter space efficiently

> *You have to get lost before you find yourself*
>
> John Green, *Paper Towns*

Abstract

Statistical inference is the process of estimating the properties of a population from a noisy sample of this population (i.e. from observed data). In this chapter[1], I introduce the Bayesian framework, whose logical coherency and conceptual simplicity make it attractive for inference problems. Then I review some techniques that I will later use to explore the parameter space efficiently, i.e. in a way that can can ensure the fast sampling of the regions of the parameter space that are the most compatible with observed data. Finally, I explore the Bayesian indirect likelihood framework, a set of alternatives to traditional Bayesian techniques that have proven useful when the computation of the probability of the data given the model is impossible or impractical.

Contents

---

[1] Note: This chapter uses extracts from Carassou et al. (2017) in some of its sections.

6.1 Scientific motivation

Our goal in this thesis is to provide the most probable model(s) for the observed properties of galaxies. Our Stuff empirical model, described in Section 4.6, generates simulated populations of galaxies from a set of input parameters describing their luminosity and size distributions. The dynamic range of the values of these parameters defines a multidimensional *parameter space* that needs to be explored to find the parameters that reproduce observed data (i.e. photometric catalogs extracted from survey images) with the greatest accuracy. The most natural framework to explore this parameter space is the Bayesian framework, which I describe in the next section.

6.2 The Bayesian framework

What is the nature of probability? For more than two centuries, this philosophical question has split the field of statistics into two main schools of thought: *Bayesianism* and *frequentism*, each with a distinct set of tools. For frequentists, probability is seen as the long-run frequency of occurrence of a random event in a repeated experiment. This interpretation relies on the assumption that if the experiment is reproduced indefinitely, the relative frequencies will converge to a limit, due to the law of large numbers. Uncertainty here is a measure of variability, due to errors in measurement. The model parameters are generally (but not always) kept fixed during the repeated experiment, therefore no probability statement can be made on them. Bayesians, however, see probability as a measure of the degree of belief in the occurrence of an event, which is a broader and more subjective interpretation. The model parameters are treated as random variables, and we can therefore quantify their probable values by sampling their probability distribution function (PDF). This philosophical distinction can create situations where the conclusions of the frequentist and Bayesian approaches to the same problem are contradictory[2] (see e.g. the examples provided in VanderPlas 2014). But in real life problems, both methods can be complementary and are often used in concert. Frequentist tools, such as point estimates and maximum likelihood methods for example, are commonly used to summarize the results of Bayesian inference (Babu 2012).

Bayesian statistics relies on Bayes' theorem (or rule), which provides a logical framework for updating a prior belief about the world when faced with new objective information.

Theorem 1 *(Bayes' theorem) Given a model about the state of the world with parameters θ, the probability of this model explaining the data is:*

$$P(\theta|\mathcal{D}) = \frac{P(\mathcal{D}|\theta)P(\theta)}{P(\mathcal{D})} \tag{6.1}$$

---

[2] For simple problems though, the two approaches often yield similar results.

where $P(\mathcal{D}|\theta)$ *is the likelihood (or the likelihood function) of the data, which gives the probability of the data given the model,* $P(\theta)$ *is the prior, or the probability of the model with the parameters* θ*, and* $P(\mathcal{D})$ *is the evidence, which acts as a normalization constant and can often be ignored in inference problems.*

Bayes' theorem was first intuited in 1740 by English reverend and amateur mathematician Thomas Bayes (Bayes et al. 1763, published posthumously by Richard Price). It was then later formalized in its modern mathematical form by Pierre-Simon Laplace (1774)[3], who applied his "law of inverse probability" to demographics, judicial reforms and celestial mechanics. As Bayes' theorem usually means the computation of complex integrations, it was largely supplanted by frequentist methods, easier to compute, and which dominated the statistical landscape during the first half of the 20th century. But the rise of powerful computers and Bayesian computation techniques, such as Markov Chain Monte Carlo (introduced in Section 6.3), have rendered the use of Bayesian methods very attractive in many scientific areas. In astronomy, Bayesian methods became popular only in the early 2000s (Sharma 2017). The reader is referred to the excellent technical and philosophical introduction to Bayesianism and frequentism of VanderPlas (2014), and for a more complete treatement of the tools of Bayesian inference, to the book of Wasserman (2013).

6.2.1 The choice of prior

The prior distribution encapsulates our current knowledge about the model parameters, thereby setting initial constraints on the parameter space. Priors can be any probability distribution, which makes their choice a notoriously hard task, especially in high-dimensional problems, where it is almost impossible to have a prior distribution that accurately represents our prior beliefs about the parameters. In fact, determining the right prior is an intrinsically subjective process, and this represents the main topic of criticism from frequentists, as the shape of the posterior will be irremediably affected by the shape of the prior.

Priors essentially come in two types: *informative* and *uninformative*[4]. Uninformative priors are often used if no information a priori is known about the model parameter. The term "uninformative" is actually a misnomer, because all priors carry some amount of information. In that case, they are chosen so that their shape minimizes their influence over the posterior. A large variety of methods for constructing uninformative priors have been developed, that have been mostly cataloged by Yang and Berger (1996).

A simple general rule for choosing uninformative prior distributions is the application of Bernoulli's (1713) and Laplace's (1814) "principle of insufficient reason", which states that *if there is no reason to believe that a set of outcomes are not equally likely, it is reasonable to assume that they are equally likely*. Therefore the simplest uninformative prior is the *uniform* (or *flat*) *prior*, which for multivariate problems is expressed as:

$$P(\theta) = \begin{cases} \prod_{i=1}^{n} \frac{1}{u_i - l_i} & \text{if } l_i \leq \theta_i \leq u_i \ \forall i \in [1, n] \\ 0 & \text{if } \exists\, i \in [1, n] \mid \theta_i < l_i \text{ or } \theta_i > u_i \end{cases} \tag{6.2}$$

---

[3] Fun fact: Laplace also invented the frequentist school of thought.

[4] Alternative terms in the literature for uninformative priors include *non-informative priors*, *reference priors*, *vague priors*, *diffuse priors* and *objective priors* (Lunn et al. 2012).

where $\theta = \{\theta_1, \theta_2, ..., \theta_n\}$ is the parameter values vector, and l_i and u_i are the lower and upper limit of the support of the prior PDF for parameter i respectively. When using the uniform distribution as prior, the support of the distribution has to be chosen carefully and with high confidence, because no one want wants to lose information on a potentially important region of the parameter space that will never be explored.

Another popular uninformative prior is the *Jeffreys prior* (Jeffreys 1946; 1961), defined as:

$$P(\theta) = \sqrt{\det(I(\theta))} \tag{6.3}$$

Where $I(\theta))$ is the $n \times n$ Fisher information matrix, i.e. the expected values of the second partial derivatives of the log-likelihood, which measures its curvature:

$$I_{i,j}(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial\theta_i \partial\theta_j} \log P(D|\theta)\right] \tag{6.4}$$

The most important property of the Jeffreys prior is that it is invariant by reparametrization of the model, i.e. if h is a transformation of the parameter θ so that $\phi = h(\theta)$, then $P(\phi) = P(\theta)$. This property, now referred to as *Jeffreys invariance*, ensures that the prior is independent of the parametrization chosen[5]. Despite the appeal of this property however, the application of the Jeffreys prior to multidimensional models is still controversial (Kass and Wasserman 1996). In practice, the Jeffreys prior will give more weight to the parameters that create big changes in the log-likelihood when being subject to small changes (Lunn et al. 2012).

When specific information is known about the parameter, informative priors can be used to strongly restrict the parameter space. According to the problem, the choice of informative priors can be based on the posterior distribution from previous published studies, or merely on intuition on the plausible values of the parameters.

Priors with the same parametric form as the posterior are called *conjugate priors*. First formalized by Schlaifer and Raiffa (1961), conjugate priors can be mathematically convenient if no information contradicts their use. For example, when inferring the parameters of a normal distribution, using a normal prior is advantageous, except if a priori knowledge indicates a skewed distribution. An exhaustive catalog of conjugate priors has been made available by Fink (1997).

Prior distributions whose integral is infinite are called *improper priors*. Improper priors must be treated with great care, because they can lead to improper posteriors, whose interpretation can be awkward. The uniform prior on an infinite interval is a good example of an improper prior.

There is an extensive collection of historical and methodological reviews on the choice of priors. The reader is especially referred to the works of Jaynes (1968; 1985), Kass and Wasserman (1996), and Simpson et al. (2014).

---

[5] The uniform prior is not Jeffreys invariant.

6.3 Sampling the posterior: Markov Chain Monte Carlo

Monte Carlo[6] methods are a class of numerical methods for solving mathematical problems using large sets of randomly generated numbers. The groundworks of these methods have been first laid out by Metropolis and Ulam (1949)[7], using the unpublished idea of John von Neumann for tracking the probable paths taken by neutrons through different materials in a chain reaction explosion (Metropolis 1987). They have become some of the most popular sampling techniques in various scientific areas thanks to the rise of electronic computers.

Markov Chain Monte Carlo (or MCMC) is a class of methods to sample a PDF using Markov chains. More precisely, it is a set of iterative processes which perform a random walk in the parameter space to approximate a posterior distribution with the help of Markov chains (Markov 1913). A Markov chain is a sequence of random variables $\{\theta_{(0)}, \theta_{(1)}, \theta_{(2)}, ..., \}$ in the parameter space (called *states*) that verifies the *Markov property* (Markov 1954): the conditional distribution of $\theta_{(t+1)}$ given $\{\theta_{(0)}, ..., \theta_{(t)}\}$ (called *transition probability* or *kernel*) only depends on $\theta_{(t)}$. In other words, the probability distribution of the next state only depends on the current state. After a period (whose length depends on the starting point and the random path taken by the chain) where the chain travels from low to high probability regions of the parameter space, the MCMC samples ultimately converge to a stationary distribution in such a way that the density of samples is proportional to the posterior PDF, also called *target distribution*, as illustrated in Figure 6-1. The portion of the chain which is not representative of the target distribution (i.e. the first iterations where the chain has not yet reached stationarity) is called *burn-in*, and is usually discarded from the analysis a posteriori. Well optimized MCMC methods provide an efficient tool to avoid wasting a lot of computing time sampling regions of very low probability. There is a great variety of MCMC algorithms, and the choice of a specific algorithm is problem-dependent. The reader is driven to Sharma (2017), who provides a recent review of the history and growing diversity of MCMC based Bayesian methods applied to astronomical problems.

The Metropolis-Hastings algorithm

The Metropolis-Hastings algorithm (Metropolis et al. 1953; Hastings 1970) is one of the most general MCMC methods. In this algorithm, given a state $\theta_{(t)}$, a proposed state θ^* is generated using a user-defined transition kernel $Q(\theta^*|\theta_{(t)})$, which represents the probability of moving from $\theta_{(t)}$ to θ^*. The proposition is accepted with probability:

$$a = \min\left\{\frac{P(\theta^*|D)}{P(\theta_{(t)}|D)}\frac{Q(\theta_{(t)}|\theta^*)}{Q(\theta^*|\theta_{(t)})}, 1\right\}. \tag{6.5}$$

If the proposed sample is accepted, then $\theta_{(t+1)} = \theta^*$ and the chain jumps to the new state. Otherwise, $\theta_{(t+1)} = \theta_{(t)}$. This acceptance rule allows the chain to climb to the high probability regions while descending to lower probability regions from time to time.

The choice of the transition kernel $Q(\theta^*|\theta_{(t)})$ is crucial to guarantee the rapid convergence of

---

[6] The term “Monte Carlo” was first coined by Metropolis from the analogy between statistical simulations and the famous casino of the capital of Monaco where Stanislaw Ulam’s uncle would gamble regularly (Metropolis 1987).

[7] According to Metropolis (1987), Enrico Fermi actually independently invented and used the (yet unnamed) Monte Carlo method using mechanical and mental calculations to study the slowing down of neutrons decades before Metropolis et al., but his contribution is often neglected in history books.

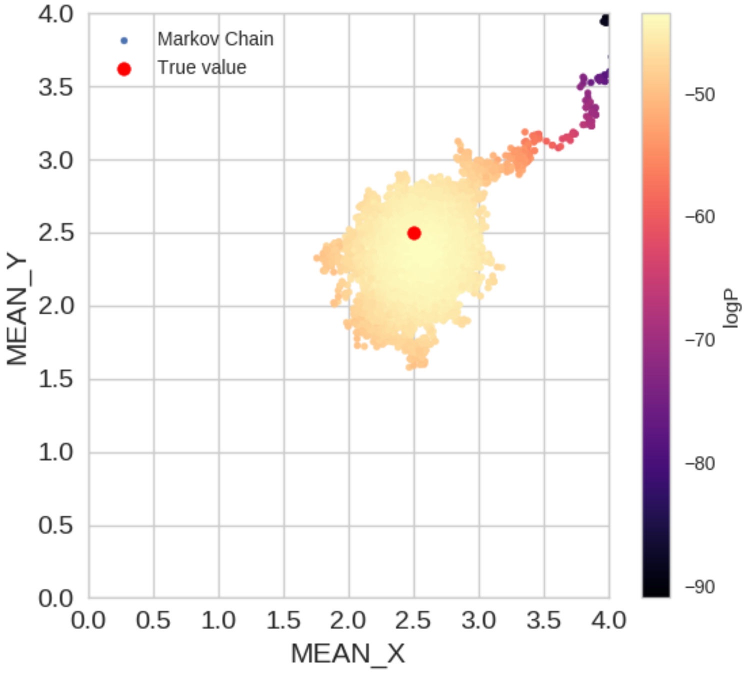


Figure 6-1 – Path of an MCMC run inferring the mean of a bivariate normal distribution after 10^4 iterations. The Markov chain traveled from low (dark) to high (yellow) likelihood regions of the parameter space and finally converged around the true mean of the Gaussian. The dark and violet points can be considered burn-in, i.e. not a representative sample of the posterior distribution.

the chain. If the kernel is symmetric, i.e. if it verifies $Q(\theta^*|\theta_{(t)}) = Q(\theta_{(t)}|\theta^*)$, then the acceptance probability reduces to:

$$a = \min\left\{\frac{P(\theta^*|D)}{P(\theta_{(t)}|D)}, 1\right\}. \tag{6.6}$$

The kernel is usually chosen among distributions that are defined by location and scale parameters. Popular choices include the multivariate Student's t distribution $t_\nu(0,\Sigma)$[8] or the multivariate normal distribution $\mathcal{N}(0,\Sigma)$ centered on the current state and with a covariance matrix Σ (and in the case of the t distribution, with a degree of freedom ν, scalar and positive) which determines the size and orientation of the jumps, so that:

$$\theta^* = \theta_{(t)} + \zeta_{(t+1)}, \tag{6.7}$$

where $\zeta_{(t+1)}$ follows $t_\nu(0,\Sigma)$ or $\mathcal{N}(0,\Sigma)$ (Chib and Greenberg 1995). We opt for the latter solution for our problem, due to its intuitive parametrization and computational ease.

6.4 Adaptive MCMC

The shape and size of the proposal distribution have a great influence on the convergence and fast traveling of the Markov chain across the parameter space (Gelman et al. 1994). But how to find the proposal distribution(s) that maximize the efficiency of a given MCMC run?

A good way to assess convergence speed for a given kernel is to monitor the acceptance rate, that is, the fraction of accepted samples over previous iterations. The acceptance rate is mainly influenced by the covariance matrix of the transition kernel Σ. If the jump sizes are too large, the acceptance rate is too low, and the chain stays still for a large number of iterations. If the jump sizes are too small, the acceptance rate is very high but the chain needs a high number of iterations to move from one region of the parameter space to another. These situations are illustrated in Figure 6-2. The desired acceptance rate depends on the target distribution, and there is no universal criterion for its optimization, but Roberts et al. (1997) proved that for any d-dimensional target distributions (with $d \geq 5$) with independent and identically distributed (i.i.d.) components, optimal performance of the Random Walk Metropolis algorithm is attained for an asymptotic acceptance rate of 0.234[9]. Therefore, the optimal kernel can be found by empirical tuning of its size so that the acceptance rate gets close to 0.234. However, empirical tuning is difficult and time consuming exercise when dealing with multidimensional problem with correlations between parameters and a target distribution of unknown shape. To tackle this issue, a set of methods called *adaptive MCMC* have been developed, that use past information about the chain to automatically tune the kernel. There is a great variety of such methods, and the reader is referred to Roberts and Rosenthal (2009) for a technical review.

As adaptive methods, by definition, destroy the Markovian nature of the chain, the choice of an adaptive scheme must ensure correct ergodic properties with respect to the stationary distribution, in the following sense:

Definition (Ergodicity) A Markov chain $\{\theta_{(0)}, \ldots, \theta_{(t)}\}$ taking its values in a state space $\mathcal{X}$ and with transition kernel P_γ (where γ is the tuning parameter of the kernel) is said to be *ergodic*

---

[8] When $\nu \to \infty$, the t distribution converges to a normal distribution.

[9] For univariate targets, the optimal asymptotic acceptance rate increases to 0.44 (Roberts et al. 2001).

with respect to a target distribution π if:

$$\lim_{t\to\infty} ||P_\gamma^t(\theta,\cdot) - \pi(\cdot)||_{TV} = 0,\ \forall\theta \in \mathcal{X} \tag{6.8}$$

Where $||\cdot||_{TV}$ is the total variation distance[10].

This means that the Markov chain always converges to the same stationary distribution, regardless of the starting point.

Definition (Adaptive MCMC) An adaptive MCMC scheme $\{\theta_t, \Gamma_t\}$ is defined given a family of transition kernels $\{P_{\Gamma_t}\}$, where Γ_t is the tuning parameter at iteration t (for example, the covariance matrix of a multidimensional Gaussian proposal kernel), chosen according to the adaption scheme (usually from the past history of the chain) and each having the target distribution π as their stationary distribution.

In order to be suitable, an adaptive MCMC scheme must verify the following conditions:

Definition (Diminishing adaption) $\lim_{t\to\infty} D_t = 0$ in probability, where

$$D_t := \sup_{\theta\in\mathcal{X}} ||P_{\Gamma_{t+1}}(\theta,\cdot) - P_{\Gamma_t}(\theta,\cdot)||_{TV} \tag{6.9}$$

This means that as $t \to \infty$, the parameters of the proposal distribution must depend less and less on earlier states of the chain, i.e. they must change less as time increases.

Definition (Bounded convergence[11]) Let $M_\epsilon(\theta,\gamma) = \inf\{t \geq 1 : ||P_\gamma^t(\theta,\cdot) - \pi(\cdot)||_{TV} \leq \epsilon\}$ be the convergence time of kernel P_γ when starting in state θ. An adaptive Markov chain satisfies the bounded convergence criterion if the sequence

$$\{M_\epsilon(\theta_{(t)},\Gamma_t)\}_{t=0}^{\infty} \text{ is bounded in probability, } \forall\epsilon > 0 \tag{6.10}$$

This means that the time that transition kernels get close to the target distribution within ϵ is bounded. In the light of these definitions, Roberts and Rosenthal (2007) proved the following theorem:

Theorem 2 *For an MCMC algorithm with adaptive proposal distributions, if it satisfies the bounded convergence and diminishing adaptation conditions, then it is ergodic and it will converge asymptotically to the correct target distribution.*

While diminishing adaption can be relatively easy to measure for a given adaption scheme, bounded convergence is harder to prove in practice (Bai 2010). In the following sections, I review two adaptive methods that will be used in this study: the *Adaptive Proposal algorithm* and the *Adaptive Metropolis algorithm*.

---

[10] The total variation distance can be informally defined as the largest possible distance between two probability distributions.

[11] The Bounded convergence criterion can also be called the *Containment condition* in the literature.

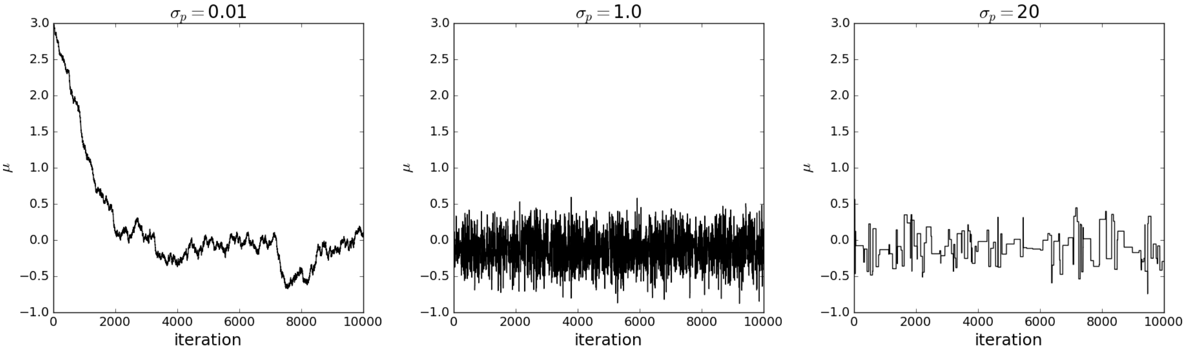


Figure 6-2 – Traceplots depicting three typical situations that can arise in a standard MCMC chain with the (non-adaptive) Metropolis-Hastings algorithm. The input data is a set of 20 points normally distributed with mean 0 and standard deviation 1. The parameter to infer is the mean μ of the input data distribution. The prior is a normal distribution with mean 0 and standard deviation 1, and the transition kernel is a normal distribution centered on the current state and width σ_p. In each case the chain starts from $\mu = 3$ and is run for 10^4 iterations. The target distribution sampled is the same, but the width of the proposal distribution, that is, the jump size, is different for each case. Left panel: The jump size is too small. The burn-in phase is very long and a much longer chain is needed to sample the target distribution. Central panel: The jump size is optimal, therefore the target distribution is well sampled. Right panel: The jump size is too big. Hence the chain spends a lot of iterations in the same position, which makes the sampling of the target distribution inefficient.

6.4.1 The Adaptive Proposal algorithm

In the *Adaptive Proposal* (AP) algorithm proposed by Haario et al. (1999), the proposal distribution (a Gaussian centered on the current state) is tuned every U iterations from the covariance computed on a fixed number H (where H ≥ 2) of previously accepted states. In this situation, the chain learns about the target distribution from its path across the parameter space. The AP algorithm requires three tuning parameters: the *update frequency* U, the *memory parameter* H, and an empirical *scaling factor* c_d for d-dimensional problems.

The scaling factor controls the size of the covariance matrix of the transition kernel. It can be tuned manually, or based on theoretical justifications. For the latter, Roberts et al. (1997) designed a popular scaling factor, defined as $c_d = (2.38)^2/d$. In this case, the proposal distribution is $\mathcal{N}(0, c_d\Sigma_H)$, where Σ_H is the $d \times d$ covariance matrix of the proposal distribution computed using the subset of accepted states over the last H iterations. The scaling factor of Roberts et al. optimizes the efficiency of the MCMC process for multivariate Gaussian target distributions. Numerical experiments by Haario et al. have shown that the efficiency of the Adaptive Proposal algorithm is more sensible to the choice of scaling factor than to the choice of update frequency and memory parameter.

The AP algorithm simulates slightly biased samples of the target distribution, and the stationary distribution reached by the algorithm is not strictly ergodic. However, the authors claim that there is small difference between the simulated distribution and the target distribution for unimodal posteriors.

6.4.2 The Adaptive Metropolis algorithm

The Adaptive Metropolis (AM) algorithm of Haario et al. (2001) is an updated version of the AP algorithm. Instead of tuning the covariance matrix of the Gaussian proposal distribution from a fixed number of previous states determined by the memory parameter, it is now computed at every steps (or every H steps) from the entire path of the chain. After an initial period of length t_0, where it is initialized, the covariance matrix at iteration t is regularly updated according to:

$$\Sigma_t = \begin{cases} \Sigma_0 & \text{if } t \leq t_0 \\ c_d \mathrm{Cov}(\theta_{(0)}, \ldots, \theta_{(t-1)}) + c_d \epsilon I_d & \text{if } t > t_0 \end{cases} \tag{6.11}$$

where Σ_0 is an arbitrary first guess for the empirical covariance matrix, c_d is the empirical scaling factor defined in the previous section, $\mathrm{Cov}(\theta_{(0)}, \ldots, \theta_{(t-1)})$ is the empirical covariance matrix computed from all the previously accepted iterations, ϵ is a user-defined factor > 0 chosen to be small compared to the size of the posterior to prevent the covariance matrix from becoming singular, and I_d is the d-dimensional identity matrix. The proposal distribution is therefore $\mathcal{N}(0, \Sigma_t)$.

With this updating rule, dependent on the history of the chain, the stochastic process used for sampling is no longer Markovian but, contrary to AP algorithm, Haario et al. have proven that the AM algorithm has the correct ergodic properties, assuming the target distribution support S $\subset \mathbb{R}^d$ ($P(\theta|D) = 0$ outside S) is bounded and the density is bounded from above ($\exists M \leq \infty$ so that $P(\theta|D) < M$ for $\theta \in S$). Which means that the simulations provided by the AM algorithm will be non-biased samples of the target distribution.

6.5 Simulated Annealing

In optimization problems, the goal is to find the global minimum of a given cost function. But in situations where this function presents a complex shape, iterative methods used to achieve that goal such as MCMC can be trapped in sub-optimal solutions called *local minima*. To prevent these situations from happening, Kirkpatrick (1983) devised the *simulated annealing algorithm*. Rooted in statistical mechanics, this optimization algorithm exploits the analogy between how a metal cools and the search for the minimum of a cost function. In metallurgy, the annealing process is the action of heating a metal to its point of fusion and then allowing it to cool slowly. Heating allows the atoms in the molten metal to move freely. When the temperature decreases, the thermal mobility also decreases and the atoms form new bonds. If the cooling process is slow enough to let the atoms spontaneously rearrange themselves, annealing results in the formation of a stable crystalline structure with minimal internal energy, or at least with a more stable atomic configuration than the initial configuration. The resulting metal has its ductility increased (it is more easily bent), and presents fewer defaults.

The system can be modelled by a set of atomic configuration $\{r_i\}$ defined by the positions of the atoms of the piece of metal, at a given temperature T. The internal energy of the configuration is denoted $E(\{r_i\})$. If the system is in thermal equilibrium, the probability distribution of these energies E is the Boltzmann distribution (Boltzmann 1877):

$$P(E) \propto \exp[-E(\{r_i\})/k_B T] \tag{6.12}$$

where $E(\{r_i\})$ is the internal energy of the configuration and k_B is the Boltzmann constant. In other words, at a given temperature, the system has a low probability of being in a high-energy configuration, and a high probability of being in a low-energy configuration.

In the simulated annealing algorithm, the cost function is analogous to the energy of the configuration, and $k_B T$ is replaced by a notion of *effective temperature* τ, a variable in the same units as the cost function. In this setting, the probability of moving from a state of energy E_1 to a state of energy E_2 is:

$$P(E_1 \rightarrow E_2) = \exp[-(E_2 - E_1)/\tau] \tag{6.13}$$

At the beginning of the process, the system is "melted" by setting an arbitrarily high initial temperature τ_0. Then, at each step of the iterative process (or every fixed number of steps), the effective temperature is lowered by a user-defined *cooling schedule*, which depends on the problem. By decreasing the temperature, the acceptance probability also decreases progressively, which means that the system moves less and less often towards "worse" solutions, i.e. towards states with a more energetic configuration a iteration i+1 than at iteration i. At the end of the process, the system should wander around the global minimum of the cost function. The whole process is illustrated in Figure 6-3.

Using the formalism introduced in previous sections, when the simulated annealing algorithm is used in concert with the Metropolis Hastings algorithm, the acceptance probability becomes:

$$a = \begin{cases} \exp - \frac{\ln P(\theta_{(t)}|\mathcal{D}) - \ln P(\theta^*|\mathcal{D})}{\tau} & \text{if } \ln P(\theta^*|\mathcal{D}) - \ln P(\theta_{(t)}|\mathcal{D}) < 0 \\ 1 & \text{if } lnP(\theta^*|\mathcal{D}) - \ln P(\theta_{(t)}|\mathcal{D}) \geq 0 \end{cases} \tag{6.14}$$

where $\ln P(\theta_{(t)}|\mathcal{D})$ and $\ln P(\theta^*|\mathcal{D})$ are respectively the log of the posterior density at the current and the proposed state. In other words, if a propositio is considered more probable, it is accepted. Otherwise, it is accepted with probability a (defined in Equation (6.14)). To perform the latter operation in practice, a uniformly distributed random number R_N is drawn in the interval $[0, 1]$. If $R_N < a$, the jump is accepted. As expected, for $\tau = 1$, the acceptance probability is the same as that of the Metropolis-Hastings algorithm in Equation 6.5 for the particular case of a symmetric proposal distribution.

The simulated annealing algorithm is adaptable to a wide variety of complex optimization problems and requires little tuning from the user, merely the initial temperature, the cooling schedule and the frequency at which the temperature is updated.

6.6 Likelihood-free inference

There are multiple cases where the likelihood is intractable or unknown, for mathematical or computational reasons, which renders classical Bayesian approaches unfeasible. In our case, it is the selection effects of the modeled galaxy samples (instrumental, calibration, Malmquist and Eddington bias, etc... cf Chapter 3) that is impractical to include in the likelihood. To tackle

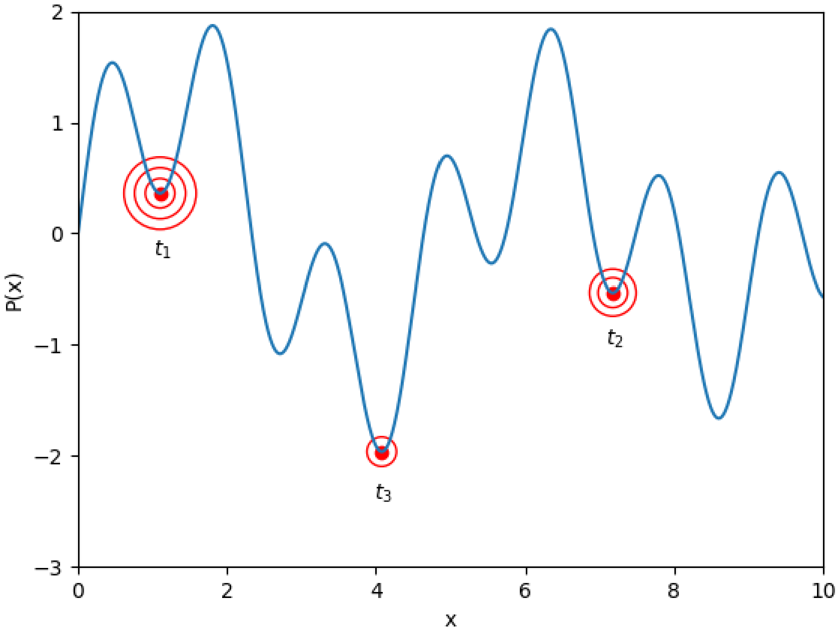


Figure 6-3 – This plot illustrates the simulated annealing algorithm for a cost function P(x) with multiple local minima. The situation is the following: an MCMC particle, depicted as a red dot, performs a random walk across the support of the cost function. The temperature follows a decreasing trend with time. Red concentric circles around the particle illustrate the scale of its transition kernel for a given temperature. The more concentric circles, the more freedom to move for the particle. At t_1, the particle has found a local minimum, but the temperature is high, so the particle can move freely over the support of the function. At t_2, the particle is found within a lower value local minimum, but the temperature is still enough for the particle to get out of it. Finally, at t_3, the particle has found the global minimum of the function. As the temperature is low, its freedom of movement is restrained and its motion stabilizes around the minimum.

this issue, a class of methods, called “likelihood-free”, have been developed in the *Bayesian indirect likelihood* (BIL) general framework to infer posterior distributions without explicit computation of the likelihood. Drovandi et al. (2015) divides the BIL framework into two main classes of methods: *approximate Bayesian computation* and *parametric Bayesian indirect likelihood*. I provide a detailed description of these techniques in the following sections.

6.7 Approximate Bayesian Computation

One of the “likelihood-free” techniques is called Approximate Bayesian Computation (ABC), and was introduced in the seminal article of Pritchard et al. (1999) for population genetics. ABC is based on repeated simulations of datasets generated by a forward model, and replaces the likelihood estimation by a comparison between the observed and synthetic data. Its ability to perform inferences under arbitrarily complex stochastic models, as well as its well established theoretical grounds, have lead to its growing popularity in many fields, including ecology, epidemiology, and stereology (see Beaumont 2010 for an overview).

The classic ABC Rejection sampling algorithm, introduced in its modern form by Pritchard et al. (1999), is defined in Algorithm 1:

```
for $t$ = $1$ to $T$ do
    Repeat
        Generate $\theta^*$ from the prior distribution ;
        Simulate data $\mathcal{D}^*$ from parameters $\theta^*$;
    until $\rho(\eta(\mathcal{D}^*), \eta(\mathcal{D})) \leq \epsilon$ ;
    set $\theta_{(t)} = \theta^*$ ;
end
```

Algorithm 1: ABC Rejection sampling algorithm

where ρ is a distance metric built between the simulated and observed datasets, usually based on some summary statistics η, which are parameters that maximize the information contained within the datasets (for example, normally distributed datasets can be characterized using the mean and standard deviation of the underlying Gaussian distribution), and ϵ is a user-defined tolerance level > 0. Using the ABC algorithm with a good summary statistic and a small enough tolerance ultimately leads to a fair approximation of the posterior distribution (Sunnaker et al. 2013). The choices of ρ, η and ϵ are highly non-trivial though, and they constitute the fundamental difficulty in the application of ABC methods as they are problem-dependent (Marin et al. 2011). Moreover rejection sampling is notorious for its inherent inefficiency, as sampling directly from the prior distribution results in spending computing time simulating datasets in low-probability regions. Therefore, several classes of sampling algorithms have been developed to explore the parameter space more efficiently. Three of the most popular of them are outlined below.

- In the *ABC-MCMC* algorithm (Marjoram et al. 2003), a point in the parameter space called a particle performs a random walk (defined by a proposal distribution or transition kernel) across the parameter space, and is only moving if the simulated dataset generated by these parameters matches better the observed dataset, until it converges to a stationary distribution. As in standard MCMC procedures, the efficiency of the algorithm is largely determined by the choice of the scale of the kernel.

- In the *ABC Sequential Monte Carlo* parallel algorithm (ABC-SMC, Toni et al. 2009), samples are drawn from the prior distribution until N particles are accepted, that is, those with a distance to the data $< \epsilon_0$. All accepted particles are attributed a statistical weight ω_0. The weighted particles then constitute an intermediate distribution from which another set of samples is drawn and perturbed with a fixed transition kernel, until N particles satisfy the acceptance criterion: $\rho < \epsilon_1$, with $\epsilon_1 < \epsilon_0$. They are then weighted with ω_1 and the process is repeated with a diminished tolerance at each step. After T iterations of this process, the particles are sampled from the approximated posterior distribution. The performance of ABC-SMC scales as N, where N is the number of particles. Different variations of ABC-SMC algorithms have been published, each with a different weighting scheme for particles.

- *ABC Population Monte Carlo* (ABC-PMC, Beaumont et al. 2008) is similar to ABC-SMC, but differs in its adaptive weighting scheme: its transition kernel is Gaussian and based on the variance of the accepted particles in the previous iteration. This scheme requires the fewest tuning parameters of the three algorithms discussed here (Turner and Van Zandt 2012). But ABC-PMC is also more computationally costly than ABC-SMC, as its performance scales as N^2 (caused by its adaptability).

The reader is referred to Csilléry et al. (2010), Marin et al. (2011), Turner and Van Zandt (2012), Sunnaker et al. (2013), and Gutmann and Corander (2016) for a set of historical, methodical, and theoretical reviews of the ABC approach, as well as a complete description of the algorithms mentioned above.

Application of likelihood-free inference to astrophysics

The application of likelihood-free methods to astrophysics is still rare, as noted by Cameron and Pettitt (2012) in their review. Only lately has the potential of such techniques been considered. Schafer and Freeman (2012) praised the use of likelihood-free inference in the context of quasar luminosity function estimation. Cameron and Pettitt (2012) explored the morphological transformation of high-redshift galaxies and derived strong constraints on the evolution of the merger rate in the early Universe using an ABC-SMC approach. Weyant et al. (2013) also used SMC for the estimation of cosmological parameters from type Ia supernovae samples, and could still provide robust results when the data was contaminated by type IIP supernovae. Robin et al. (2014) constrained the shape and formation period of the thick disk of the Milky Way using MCMC as their sampling scheme, based on photometric data from the SDSS and the Two Micron All Sky Survey (2MASS). Finally Hahn et al. (2016) demonstrate the feasibility of using ABC to constrain the relationship between galaxies and their dark matter halo. The recent birth of Python packages providing sampling algorithms in an ABC framework, such as astroABC (Jennings and Madigan 2016) and ELFI (Kangasrääsiö et al. 2016), which implement SMC methods, and COSMOABC (Ishida et al. 2015) which implements the PMC algorithm, will probably facilitate the rise of likelihood-free inference techniques in the astronomical community.

6.8 Parametric Bayesian indirect likelihood

Another class of likelihood-free techniques is called parametric Bayesian indirect likelihood (pBIL). First proposed by Reeves and Pettitt (2005) and Gallant and McCulloch (2009), pBIL transforms the intractable likelihood of complex inference problems into a tractable one using an auxiliary parametric model that describes the simulated datasets generated by the forward model. In this scheme, the resulting auxiliary likelihood function quantifies the discrepancy between the observed and simulated data. It is used in Bayes' theorem and the parameter space is explored using a user-defined sampling procedure, in an equivalent way to a classical Bayesian technique. While sharing similarities with the previous technique, pBIL is not an ABC method in the strict sense, as it does not require an appropriate choice of summary statistics and tolerance level to compare the observed and synthetic datasets. The accuracy of the inference in the pBIL scheme is determined by how well the auxiliary model describes the data (observed and simlated). The theoretical foundations of this scheme are described extensively in Drovandi et al. (2015).

6.9 Our approach for this work

In this thesis, we will use a pBIL approach to explore the parameter space, whose auxiliary likelihood and its construction will be described in the next chapter (cf Section 7). We will preferentially consider flat priors on our parameters, using wide boundaries around commonly accepted values in the literature to avoid contamination of our results by previous studies. We will rely on an adaptive MCMC sampler for a fast exploration of the parameter space. In the first tests of the inference pipeline that we are designing (cf Chapter 9), the sampling process will be performed using the Adaptive Proposal algorithm of Haario et al. (1999). The pipeline will be then improved to include the more reliable Adaptive Metropolis algorithm of Haario et al. (2001) that will be used for further tests and application to real data (cf Chapter 11). To take into account the loss of precision in the inference process generated by the inherent stochasticity of our model (i.e. variations of the auxiliary likelihood values for a given set of input model parameters), a temperature notion from the simulated annealing algorithm of Kirkpatrick (1983) will be superimposed on the Metropolis Hastings acceptance probability, and updated every fixed number of iterations along the MCMC run, in a scheme that will be detailed in Section 9.4.1.

Chapter 7

Comparing observed and simulated datasets

> *Illusions can be pleasant, but the rewards of truth are enormously better.*
>
> Sean Carroll, *The Big Picture*

Abstract

Our goal is to find the distributions of apparent luminosity and size of sources extracted from simulated multi-band images generated by our model that maximize the similarity with the observed distributions. In this chapter, I explore various distance metrics that have been designed to compare the discrepancy between two arbitrary multivariate distributions. We finally choose to bin our datasets, that are first preprocessed in order to maximize the efficiency of the binning scheme, and to compare their discrepancy using a "binned maximum likelihood" approach. In this context, I detail some data-based methods found in the literature to yield optimal binning schemes for multidimensional datasets.

Contents

We find ourselves in the following situation: we need to minimize the discrepancy between observed source catalogs from astronomical surveys, and simulated ones generated by our model. Therefore, we need a measure of similarity between our simulated and observed datasets. In our case, the dimensionality of the observable space is determined by the number of photometric and size parameters (i.e. the FLUX_AUTO and the FLUX_RADIUS) in every passbands extracted from the survey images.

Estimating the discrepancy between the observed and simulated binned datasets in a multidimensional space is highly non-trivial, as the choice of a good distance metric is problem dependent (Weber et al. 1998; Aggarwal et al. 2001). Some studies even suggest that the concept of "proximity" in a high dimensional space may even be meaningless (Beyer et al. 1999). Distance metrics can be *parametric*, i.e. relying on assumptions about the shape of the underlying PDF of the two datasets, or *non-parametric*, i.e. relying on no such assumptions (Hoskin 2012). The distributions of our observables may be multimodal and skewed, however many metrics rely on the assumption of normality. Others, such as the Kullback-Leibler divergence (Kullback and Leibler 1951) or the Jensen-Shannon distance (Lin 1991), cannot be used without estimating an analytical underlying PDF, which can be very computationally expensive in a high-dimensional observable space. Hence the need for a non-parametric distance estimator.

7.1 Non-parametric distance metrics

Here is a non-exhaustive list of non-parametric distance metrics found in the literature that can be used on multivariate data in the Approximate Bayesian Computation framework. A more complete review is available in Pardo and Menéndez (2006) and Palombo (2011). However, the literature is astonishingly scarce on guiding the choice of an appropriate distance metric for a given problem, and no study to quantify their relative power has been performed so far. These metrics include:

- the χ^2 **test** (Chardy et al. 1976) is a simple and widely used way of determining whether observed frequencies are significantly different from expected frequencies. The main drawback of this approach is that χ^2 test results are dependent on the binning choice (Aslan and Zech 2002). For example, Kurinsky and Sajina (2014) use the χ^2 distance to compare color-color histograms.

- the **Kolmogorov-Smirnov (KS) test** (Chakravarti et al. 1967) estimates the maximum absolute difference between the empirical distribution functions (EDF) of two samples. A generalization of this test for multivariate data has been proposed (Justel et al. 1997). However, as there is no unique way of ordering data points to compute a distance between two EDF, it is not as reliable as the one-dimensional version without the help of resampling methods such as bootstraping (Babu and Feigelson, 2006).

- the **Anderson-Darling (AD) test** (Stephens 1974) is a modification of the KS test. This method uses a weight function that gives more weight to the tails of the distributions. It is therefore considered more sensitive than the KS test, but it also suffers from the same problems in the multivariate case.

- the **Mahalanobis distance** (Mahalanobis 1936) is similar to the Euclidean norm but has the advantage of taking into account the correlation structure of multivariate data. The

Mahalanobis statistics, coupled with an univariate KS test, are used by Akeret et al. (2015) to compare photometric parameters for cosmological purposes. However, this distance only works for unimodal data distributions.

- the **Bhattacharyya distance** (Bhattacharyya 1946) is related to the Bhattacharyya coefficient, which measures the quantity of overlap between the two samples. It is considered more reliable than the Mahalanobis distance in the sense that its use is not limited to cases where the standard deviations of the distributions are identical.

- the **Earth Mover's distance (EMD)** (Rubner et al. 1998) is based on a solution to the Transportation problem. The distributions are represented by a user-defined set of clusters called signatures, where each cluster is described by its mean and by the fraction of the distribution encapsulated by it. The EMD is defined as the minimum cost of turning one signature into the other, the cost being linked to the distance between the two. A computationally fast approximate version of this distance using the Hilbert space-filling curve can be found in Bernton et al. (2017).

All these metrics rely on the estimation of the PDF of the datasets. Following the idea of Robin et al. (2014) and Rybizki and Just (2015), who grouped their data representing stellar photometry into bins of magnitude and color, we choose to bin our datasets for comparing them, for reasons I detail in the following section.

7.2 Binning multidimensional datasets

7.2.1 Advantages and caveats of binning

Histograms are the oldest and the most widely-used of density estimators. In fact, if normalized, they can be considered as piecewise-constant models of the underlying PDF of a given dataset. As our pipeline requires fast comparisons between simulated and observed distributions of observables, binning offers an attractive solution for density estimation, considering its relative simplicity of implementation and an advantageous computational cost. However, binning comes with some inevitable drawbacks: first, the number of bins increases exponentially with the number of dimensions. For a fixed-size dataset, multivariate histograms are also sparser than their univariate counterparts and display more complex shapes. Also, the choice of the size and location of bins can significantly influence the information content of the dataset. If the bins are too big, the histogram will be likely to blend important features of the dataset. And if the bins are too small, the number count in each bin will largely be determined by statistical fluctuations, and most bins will end up empty. The choice of optimal binning is even less trivial in high-dimensional spaces (Cadez et al. 2002). This class of problems is known as "the curse of dimensionality" (Bellman 1972).

Some data-based binning rules exist to return the optimal size of bins for a given dataset, such as the widely used Scott's rule (Scott 1979) or the Freedman and Diaconis rule (Freedman and Diaconis 1981) for univariate datasets. In the following section I explore three rules that have been developed to tackle multivariate datasets: Hogg's rule, Knuth's rule and Bayesian blocks, whose differences on the same dataset is illustrated in Figure 7-2. Finally, alternatives to binning for density estimation are discussed in the conclusion of this thesis.

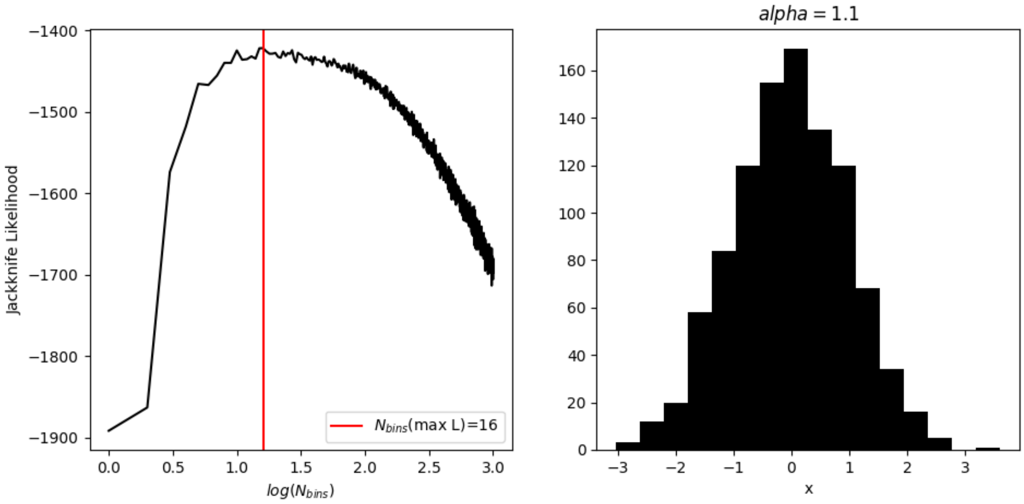


Figure 7-1 – Dataset consisting of 10^3 points normally distributed with mean 0 and unit standard deviation binned using Hogg's rule with a smoothing parameter $\alpha = 1.1$.

7.2.2 Hogg's rule

Hogg (2008) proposed a method based on "Jackknife" or "leave one out cross-validation", where a sub-sample of the dataset is constructed and used to predict the complementary sub-sample. The binning structure for a given number of bins represents a model whose predictive power is encapsulated within a likelihood function. Bins that are too fine or too coarse have a low predictive power, therefore one can find the optimal number of bins by maximizing the likelihood.

Hogg starts from assuming that the data is a sample drawn from a PDF that can be approximated by a piece-wise constant model $\hat{f}(\mathbf{x})$ where the number of data points that falls in bin i is proportional to the probability in that bin. For M equal bins of volume $V_m = \Delta_{\mathbf{x}_1}\Delta_{\mathbf{x}_2}...\Delta_{\mathbf{x}_D}$ where $\Delta_{\mathbf{x}_k} = \frac{max(\mathbf{x}_k)-min(\mathbf{x}_k)}{M}$ is the width of the bin in dimension k, the probability that a data point ends up in bin i is:

$$p(i) = \frac{N_m + \alpha}{\sum_{m=1}^{M}[N_m + \alpha]} \tag{7.1}$$

where N_m is the number of counts in bin m, and α is a "smoothing" parameter of order 1 to prevent zero probabilities in empty bins. The approximate PDF is then:

$$\hat{f}(\mathbf{x}) = \frac{p(i(x))}{V_{i(\mathbf{x})}} \tag{7.2}$$

Hogg's "Jackknife" log-likelihood for a given number of bins can be defined as the weighted sum of bin probabilities over the bins:

$$\ln L = \sum_{m=1}^{M} N_m \ln\left(\frac{N_m + \alpha - 1}{V_m\left[\sum_{m'}(N_{m'} + \alpha) - 1\right]}\right) \tag{7.3}$$

The optimal number of bins for a given α is finally found by maximizing this log-likelihood over M, as we illustrate with a simple univariate example in Figure 7-1. Note that the more empty bins the dataset will be likely to contain, the more the resulting optimal binning will depend on the choice of α.

7.2.3 Knuth's rule

The method of Knuth (2006) uses Bayesian model selection to find the optimal binning scheme with constant bin width. Like Hogg's method, Knuth starts by considering the histogram as a piece-wise constant model of the PDF of the dataset, assuming no measurement error.

Given a data array $\underline{d}$ containing N D-dimensional points, Knuth's rule segments the dataset into $\underline{M} = (M_1, .., M_D)$ bins, where M_i is the number of bins along the i-th dimension. The total number of bins is therefore $M = M_1 \times ... \times M_D$. In this framework, the probability mass of bin k containing n_k data points within volume v_k is:

$$\pi_k = n_k v_k \tag{7.4}$$

The piece-wise constant model of the PDF of the dataset is then:

$$h(x_1, ..., x_D; M_1, ..., M_D) = \frac{M}{V} \sum_{i_1=1}^{M_1} \cdots \sum_{i_D=1}^{M_D} \pi_{i_1,...,i_D} \Pi_{i_1,...,i_D}(x_1, ..., x_D) \tag{7.5}$$

where V is the volume of the entire D-dimensional histogram, and $\Pi_{i_1,...,i_D}(x_1, ..., x_D)$ is the boxcar function, defined as:

$$\Pi_{i_1,...,i_D}(x_1, ..., x_D) = \begin{cases} 1 & \text{if } (x_1, ..., x_D) \text{ falls within the bin} \\ & \text{indexed by the coordinates } (i_1, \ldots, i_D) \\ 0 & \text{otherwise} \end{cases} \tag{7.6}$$

The prior on $\underline{M}$ is a uniform PDF:

$$P(\underline{M}) = \begin{cases} C^{-1} & \text{if } \underline{1} \le \underline{M} \le \underline{C} \\ 0 & \text{otherwise} \end{cases} \tag{7.7}$$

where $\underline{1}$ is the unit vector and $\underline{C}$ is the vector containing the maximum number of bins for each dimension, set by the user. The likelihood of finding a data point d_n from the bin indexed by coordinates $(i_1, \ldots, i_D)$ is:

$$p(d_n|\underline{M}) = \frac{M}{V} \pi_{i_1,...,i_D} \tag{7.8}$$

The posterior probability of this histogram model is then given by Bayes' theorem (Equation (6.1)):

$$p(\underline{M}|\underline{d}) \propto \left(\frac{M}{V}\right)^N \frac{\Gamma(\frac{M}{2})}{\Gamma(\frac{1}{2})^M} \frac{\prod_{i_1=1}^{M_1} \cdots \prod_{i_D=1}^{M_D} \Gamma(n_{i_1,...,i_D} + \frac{1}{2})}{\Gamma(N + \frac{M}{2})} \tag{7.9}$$

where Γ is the Gamma function. For computational simplicity, we use the logarithm of this

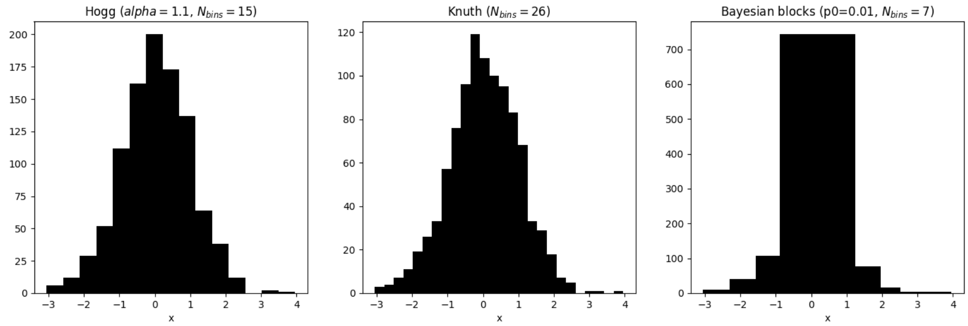


Figure 7-2 – Comparison of the 3 binning rules described in this chapter, using the same univariate dataset of 1000 points normally distributed of mean 0 and unit standard deviation.

posterior to find the optimal number of bins $\underline{M}$ by maximizing the following quantity:

$$\ln p = N \ln M + \ln \Gamma(\frac{M}{2}) - M \ln \Gamma(\frac{1}{2}) - \ln \Gamma(N + \frac{M}{2}) + \sum_{i_1=1}^{M_1} \dots \sum_{i_D=1}^{M_D} \Gamma(n_{i_1,\dots,i_D} + \frac{1}{2}) + K \quad (7.10)$$

A 1D version of Knuth's rule has been implemented in MatLab [1], as well as in the AstroML[2] (Vanderplas et al. 2012) and Astropy[3] (Astropy Collaboration et al. 2013) Python modules. I also implemented Knuth's method for multidimensional datasets by modifying the Python code used in Astropy, in order to test its effect on our pipeline (cf Section 10.6).

7.2.4 Bayesian blocks

The Bayesian blocks method differs from the previous approaches as it generates adaptive bin sizes, whose locations and widths depend on the local structure of the dataset at hand. Developed by Scargle (1998) (and later refined in Scargle et al. (2013)) for 1D histograms, then extended to multidimensional partitions of datasets in Scargle (2002), it was originally thought as a way of accurately recovering the pulse shape of gamma ray bursts in time series data without any explicit model.

Bayesian blocks segments a D-dimensional dataset into subsets called *blocks*, in which each data point is supposed to be independent, and the number counts in blocks are assumed to follow a Poisson distribution. In this context, the free parameters of the model are the bin width (in 1D) or block volume, and the number of points in each block. For a given block with volume V enclosing N data points, the posterior probability of this model is:

$$\Phi(N, V) = \frac{\Gamma(N+1)\Gamma(V-N+1)}{\Gamma(V+2)} = \frac{N!(V-N)!}{(V+1)!} \quad (7.11)$$

---

[1] http://knuthlab.rit.albany.edu/index.php/Products/Code

[2] http://www.astroml.org/modules/generated/astroML.density_estimation.knuth_bin_width.html

[3] http://docs.astropy.org/en/stable/api/astropy.stats.knuth_bin_width.html

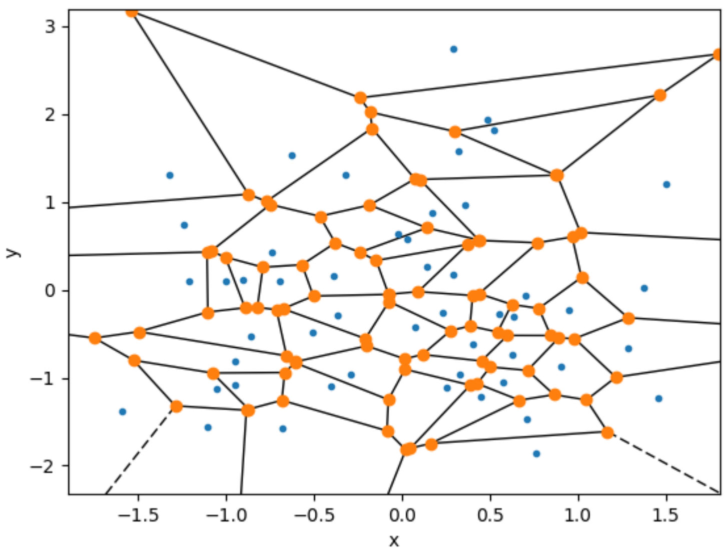


Figure 7-3 – Voronoi tessellation of 50 2-dimensional data points normally distributed

The optimal segmentation will be the one with the largest total posterior probability:

$$P = \Pi_{i=1}^{N_{blocks}} \Phi(N_i, V_i) \tag{7.12}$$

The exploration of the parameter space is done via adjacent block merging. The Maximum A Posteriori (i.e. the global maximum of the posterior distribution) is reached with the help of the *Bayesian merge factor*, which is the ratio of the block posteriors for a pair of adjacent blocks 1 and 2 merged and not merged respectively:

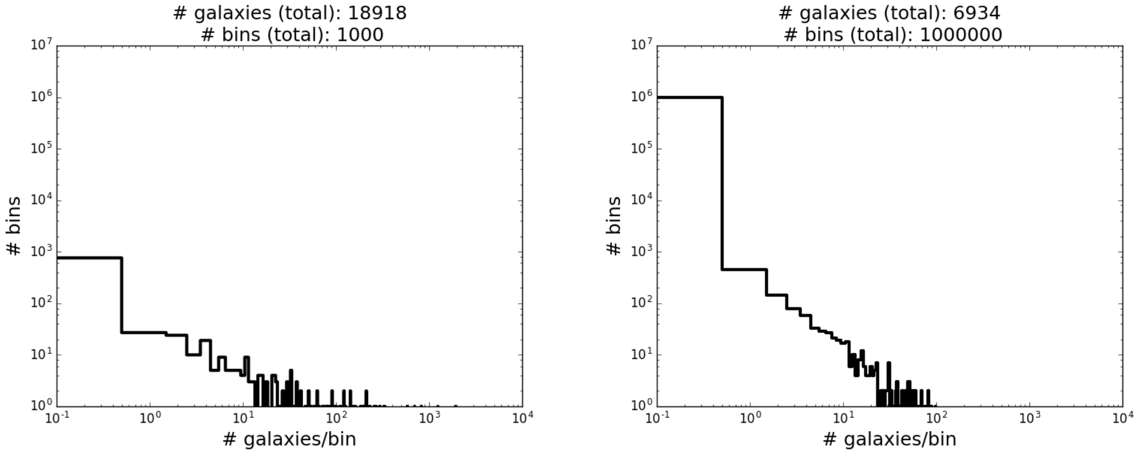
$$\mathrm{P(Merge)} = \frac{\Phi(N_1 + N_2, V_1 + V_2)}{\Phi(N_1, V_1)\Phi(N_2, V_2)} \tag{7.13}$$

The algorithm starts by partitioning the space in the most refined way possible, where each data point is contained within a cell. In the one-dimensional case, these cells are bins whose bin edges are the midpoint between adjacent data points. In multi-dimensional situations, the Voronoi tessellation (e.g. Okabe et al. 2000) provides a natural way of partitioning the space that encloses the dataset, as illustrated in Figure 7-3. In practice, there are an infinite number of ways of partitioning a multi-dimensional volume. But if we only allow the blocks to be collections of adjacent cells created by the Voronoi tessellation, then the problem becomes tractable. Then, through an iterative process, bigger *blocks* (defined as cells or collections of cells) are built from smaller blocks, depending on the posterior gradient that results from block merging. Finally, the algorithm stops when the Maximum A Posteriori has been found. There are many ways of reaching optimal segmentation in this framework that are described extensively in Scargle (1998) and Scargle (2002). In this section we focus on one method, *the cell coalescence algorithm*, described in Algorithm 2.

Unfortunately, as Scargle (2002) points out himself, the cell coalescence algorithm does not guarantee convergence to the global optimal partition in high dimensional spaces. Nevertheless, a 1D version of the Bayesian blocks method based on the MatLab algorithm described

```
Perform Voronoi tesselation on the dataset ;
Identify each Voronoi cell as a block ;
Compute P(Merge) for each pair of adjacent blocks ;
while largest P(Merge) $\geq$ 1 do
    Merge the pair with largest P(Merge) and go to (3) ;
end
```

Algorithm 2: Bayesian Blocks: the cell coalescence algorithm



(a) Multi-type case

(b) Fattening E case

Figure 7-4 – Histogram of the number of sources extracted per bin for the pre-processed input data of the test cases presented in Chapter 9. In the left panel, three observables are considered: the FLUX_AUTO in uiK_s. In the right panel, six observables are considered: the FLUX_AUTO in uiK_s and the FLUX_RADIUS in uiK_s. Using 10 bins per dimension, between the "Multi-type" case and the "Fattening E" case the number of bins increases by a factor 10^3, and the number of empty bins is increased by roughly the same amount. This illustrates the curse of dimensionality we face in this method, and puts computational limits on the number of observables we can use.

in Scargle et al. (2013) has been implemented in the AstroML[4] (Vanderplas et al. 2012) and Astropy[5] (Astropy Collaboration et al. 2013) Python modules.

Despite the appeal of its adaptive nature, we consider the use of the Bayesian blocks method to be impractical "by construction" for our goal: the difference of geometry (due to the Voronoi tesselation) between the two observables datasets partitioned by this method would make the quantification of their discrepancies highly non trivial.

---

[4] http://www.astroml.org/modules/generated/astroML.density_estimation.bayesian_blocks.html

[5] http://docs.astropy.org/en/stable/api/astropy.stats.bayesian_blocks.html

7.3 Comparing observed and simulated binned datasets

As we have seen in Section 7.2.1, binning datasets in multidimensional spaces can result in sparsely populated bins. In order to limit this effect, we must maximize the volume taken by the datasets in the multidimensional observable space before even considering to partition them. Therefore, the catalogs have to go through a pre-processing step, which is the result of two operations: a reduction of the dynamic range and a decorrelation of the observables.

7.3.1 Reducing the dynamic range of the observables

Observables such as fluxes may have a large dynamic range that goes up to the saturation level of the chosen survey. This can be problematic for the binning process of our pipeline, in the sense that it will create many sparsely populated bins. We must therefore reduce the dynamic range of the apparent properties of the sources. We cannot simply use the log of the flux arrays, because the noise properties of background-subtracted images can provide faint objects with negative fluxes. We therefore use the following transform $g(X)$, illustrated in Figure 7-5 which has already been applied to model-fitting and machine learning applications[6] (e.g., Bertin 2011b):

$$X_r = g(X) = \begin{cases} \kappa_c \sigma \ln\left(1 + \frac{X}{\kappa_c \sigma}\right) & \text{if } X \geq 0, \\ -\kappa_c \sigma \ln\left(1 - \frac{X}{\kappa_c \sigma}\right) & \text{otherwise,} \end{cases} \tag{7.14}$$

where σ is the baseline standard deviation of X (i.e., the average error of a very low flux sample), and κ_c a user-defined factor which can be chosen in the range from 1 to 100, typically. In all the test cases that we describe in Chapters 9 and 11, we set $\kappa_c = 10$. In practice we apply this compression to each dimension of the observable space, with a different value of σ for each observable. We separate the σ values into two categories for each kind of observable: σ_f for flux-related observables and σ_r for size-related ones. These values are affected by the galaxy populations in the observed field as well as the photometric properties of the field itself, such as the bands used and the noise properties. For fluxes and colors, a root mean square error estimate of the flux measurement is given by SExtractor: FLUXERR_AUTO. We set σ_f to the median value of the distribution of FLUXERR_AUTO values for the sources extracted from input data, and this operation is repeated on each filter. However, SExtractor provides no such error estimate for FLUX_RADIUS. For this kind of observable we rely on the distribution of FLUX_RADIUS of the extracted sources with respect to the corresponding FLUX_AUTO. For each passband, the value of σ_r is set to the approximate FLUX_RADIUS of the extracted sources' distribution when FLUX_AUTO tends to 0. The exact values actually do not matter, because the same compression is applied on the observed and simulated data.

7.3.2 Decorrelating the observables

The choice of the nature and number of observables is a compromise between computational cost and informational content. In fact, memory limitations intrinsic to the computational cluster when binning observed and synthetic data prevent us from using an arbitrary number of observables in the pipeline. Observables such as fluxes or magnitudes in different passbands

---

[6] Another good choice of dynamic range compression function for our flux vectors would be $\arcsin\left(\frac{X}{\kappa_c \sigma}\right)$.

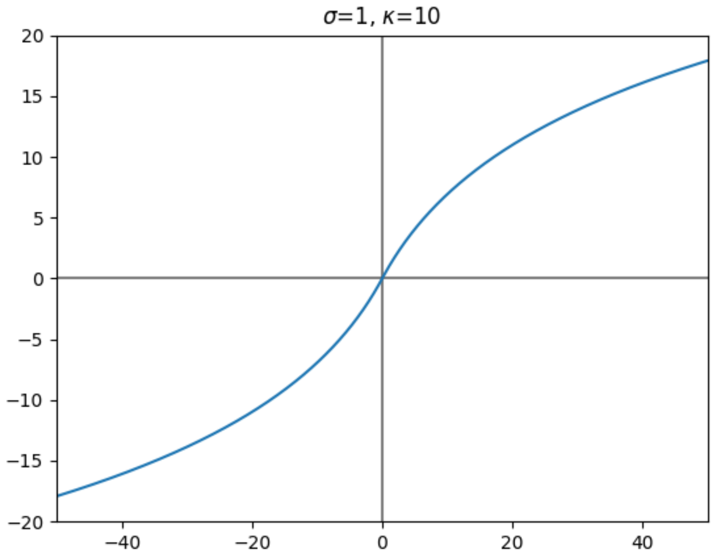


Figure 7-5 – Dynamic range reduction function $g(X)$ defined in Equation 7.14.

also tend to be strongly correlated with one another, as they originate from the same spectrum of a given galaxy from a given population. These correlations can be high if the passbands are too narrow, too close to each other, and not covering a large enough wavelength baseline. One must thoughtfully choose the appropriate set of filters *a priori* in order for the resulting set of observables to be able to disentangle the luminous properties of the different galaxy populations.

Strong correlations between input vector components can also make binning very inefficient, therefore an important pre-processing step is to decorrelate them. In that regard, we apply a linear transformation called *principal component analysis whitening*, or *sphering* (Friedman 1987, Hyvärinen et al. 2009, Shlens 2014, Kessy et al. 2015) to our reduced matrix of observables X_r of size $p \times N_s$, where p is the number of observables and N_s is the number of sources. Principal component analysis (PCA) is an algorithm commonly used in the context of dimensionality reduction. Its goal is to find a set of orthogonal axes in a dataset called principal components that encapsulate most of the variance of the data. This can be performed via a singular values decomposition (SVD) of the covariance matrix of the data:

$$< X_r X_r^T >= U \Lambda V^T, \tag{7.15}$$

where U and V are orthogonal matrices and Λ the diagonal matrix containing the non-negative singular values of the covariance matrix, sorted by descending order.

PCA whitening is the combination of two operations: rotation and scaling. First the dataset (previously centered around zero by subtracting the mean in each dimension) is projected along the principal components, which removes linear correlations, and then each dimension is scaled so that its variance equals to one. The whitening transform can therefore be summarized by:

$$X_w = \Lambda^{-\frac{1}{2}} V^T (X_r - \mu), \tag{7.16}$$

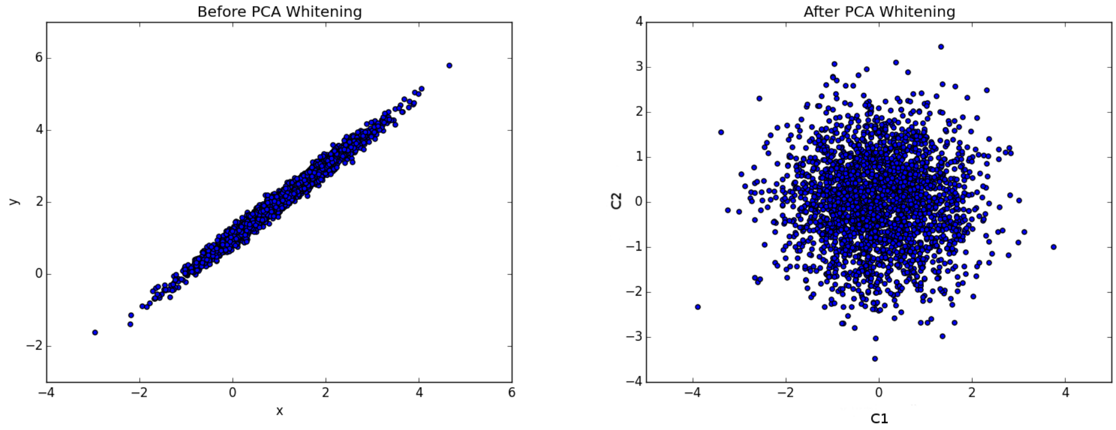


Figure 7-6 – PCA whitening of a highly correlated dataset sampled from a bivariate Gaussian distribution.

where X_w is the whitened version of the observables matrix X_r and μ is the average matrix. The PCA whitening transformation results in a set of new variables that are uncorrelated and have unit variance ($< X_w X_w^T >= I$), as I illustrate on a Gaussian example in Figure 7-6. During the iterations of the MCMC process, the observed and simulated data are centered, rotated, and scaled in the same way to ensure that both distributions can be well superposed and compared.

In practice, the simulated data is whitened using the Λ, V^T, and μ of the observed data. The number of principal components to keep is left to the choice of the user. Retaining only the components with the highest variance and therefore reducing the computational cost of the pipeline may be tempting. Nevertheless, subtle but important features can arise from low variance components, and deleting them comes at a price. In our application, we choose not to reduce the dimensionality of the problem. A step-by-step illustration of our pre-processing pipeline from a three-dimensional distribution of galaxy fluxes is hown in Figure 7-7.

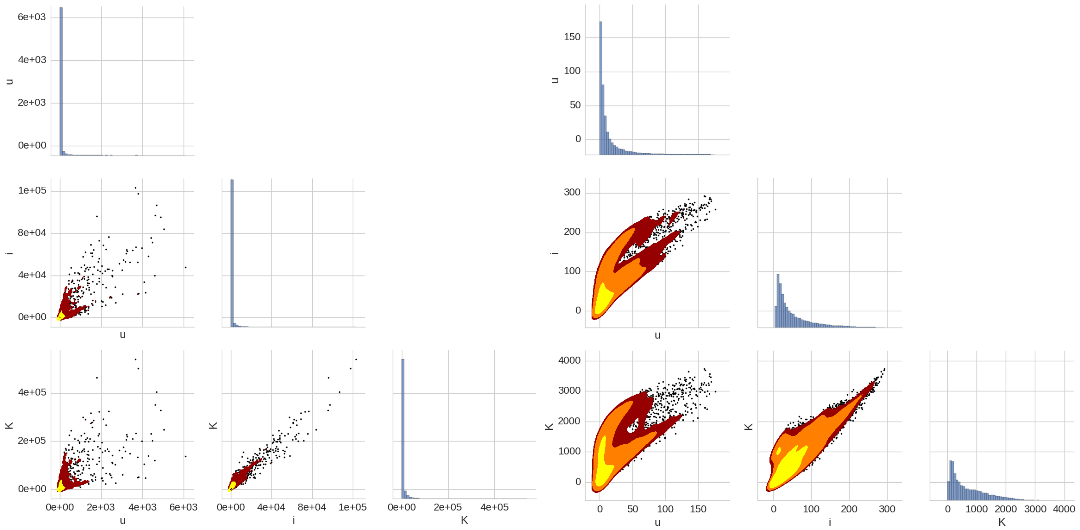
7.3.3 Quantifying the similarities between binned datasets

The simplest idea when dealing with binned datasets is to use a χ^2 fit to quantify their discrepancies. Unfortunately, this method assumes that the number of counts in each bin is large. It turns out to be inappropriate in our case, because the curse of dimensionality creates many empty bins and a sparse data distribution in high dimensional space, as illustrated in Figure 7-4.

Generally, the distances metrics that I described in Section 7.1 can be useful when used in an Approximate Bayesian Computation framework. But as we have placed ourselves within the parametric Bayesian Indirect Likelihood framework described in Section 6.8 to infer the properties of our model, we need to derive an auxiliary likelihood from the comparison of binned observed and simulated datasets.

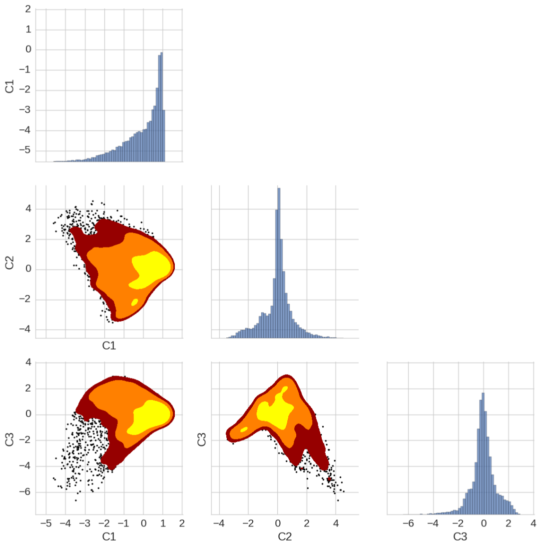
A simple and useful method we can use in this context is the “binned maximum likelihood” method (Barlow and Beeston 1993), that I describe here below.

The binned maximum likelihood method



(a) Initial

(b) Reduced



(c) Whitened

Figure 7-7 – Distribution of observables before and after each step of pre-processing from the mock input data with 2 populations of galaxies (Ellipticals+Spirals) described in Section 9.5.1. The dark red, orange and yellow areas in the contour plots are the regions that enclose 99%, 95% and 68% of the points respectively. *Top left panel*: scatter plot of the FLUX_AUTO of extracted sources (in ADUs) in filters uiK_s and their covariances. *Top right panel*: same plot, but with the dynamic range of the FLUX_AUTO distributions reduced via Equation (7.14). *Bottom panel*: same plot, after whitening of the reduced distribution of observables. The latter distribution is uncorrelated, centered on the mean of the distribution and rescaled, allowing for a much more efficient binning process than on raw fluxes, and a more practical comparison with the simulated observables.

If we assume that the number count in each bin i is described by a Poisson distribution, the probability of o_i given the model s_i is:

$$l_i = \frac{e^{-s_i} s_i^{o_i}}{o_i!} \tag{7.17}$$

The likelihood function for the histogram is then:

$$L = \prod_{i=1}^{b} \frac{e^{-s_i} s_i^{o_i}}{o_i!} \tag{7.18}$$

where b is the total number of bins, s_i is the number count in bin i for the simulated data, and o_i is the number count in bin i for the observed data. Correlations between adjacent bins are neglected here. The log-likelihood is therefore given by:

$$\ln L = \sum_{i=1}^{b} (-s_i + o_i \ln(s_i) - \ln(o_i!)) \tag{7.19}$$

As we want to maximize $\ln L$, $\ln(o_i!)$ is a constant that can be eliminated, so *in fine*, we obtain:

$$\ln L = \sum_{i=1}^{b} (-s_i + o_i \ln(s_i)) \tag{7.20}$$

An additional correction has to be applied on the log-likelihood, because as the logarithm of s_i is used, empty bins cause a problem. In order to avoid singularities, a constant small value ϵ (that I set to 1) is added to every bin up to the edges of the observables space. This process is done in both modeled and observed data so that it does not bias our results too much:

$$\ln L = \sum_{i=1}^{b} (-(s_i + \epsilon) + (o_i + \epsilon) \ln(s_i + \epsilon)) \tag{7.21}$$

The resulting log-likelihood is then introduced in Bayes' theorem (Equation (6.1)) at each iteration of our sampling algorithm.

Equation (7.21) implies that the same binning is applied on synthetic and observed data. In our pipeline, the binning pattern is only computed once and for the observed data only. The same binning is then directly applied to the simulated data to ensure better execution speed and comparability between histograms. Moreover, because the number of counts per bin is directly affected by the model parameters that rule the number density of galaxies, such as the ϕ^* parameters of the luminosity function, our histograms are not normalized to prevent a loss of information during the comparison process. The fact that the comparison does not extend beyond the range of the observed data can potentially limit the precision of our approach in a way that we have not quantified.

To use the terminology of pBIL, the binning structure and the assumption of a Poisson behavior of the number counts in each bin represent the auxiliary model that describes the data. The auxiliary likelihood derived from this structure (Equation 7.20) has been used in previous studies like e.g. Cash (1979), Bienayme et al. (1987), Adye (1998) or Robin et al. (2014).

Chapter 8

Interpreting the results of Bayesian inference

> *I may not have gone where I intended to go, but I think I have ended up where I needed to be.*
>
> Douglas Adams, *The Long Dark Tea-Time of the Soul*

Abstract

The output of an MCMC run is the evolution of the position of each free parameter of the model across iterations. But how can we know whether the MCMC samples are truly representative of the posterior distribution or not? Can we even know for sure? Is there any objective way to determine the length of the burn-in period? How to know when to stop an MCMC run? In this chapter I review some methods that can be used to diagnose the (non-)convergence of a chain. I also discuss some common practices to visualize and summarize the resulting posterior distribution.

Contents

8.1 Convergence diagnostics

Diagnosing the behavior of a Markov chain is a famously hard problem, which gets even more challenging as the number of parameters to infer increases. In fact, it is impossible to know with certainty from a finite MCMC sample that a chain has indeed converged to the target distribution (Cowles and Carlin 1996). This is why all the convergence diagnostics that have been developed can only attest failure of convergence. However, they can provide some clues on the sampling efficiency or the number of iterations after which the starting point of the chain is forgotten. Thus, the use of several diagnostics is recommended to attest convergence with more confidence. As many of these diagnostics rely on the comparison of multiple chains, running more than 3 chains constitutes a good practice. It is common in the statistical community to run between 3 and 10 chains (Sinharay 2003). In general, the more chains are launched and the longer they are run, the more precise will be the estimation of the posterior distribution.

In the following sections, I will describe two graphical and two quantitative convergence diagnostics that will be useful for this work. Of course this list is not meant to be exhaustive. The reader is referred to the excellent reviews of Cowles and Carlin (1996), Brooks and Gelman (1998) and Sinharay (2003) for a more complete overview of these diagnostics and their caveats.

8.1.1 The trace plot

The simplest of convergence diagnostics is the *trace plot* (also called the *time series plot*). It is a plot (or collection of plots) reconstructing the path taken by the Markov chain by displaying the value of each free parameter of the model at each iteration (cf Figure 8-1). If a single chain has converged to the target distribution, its trace ends up stabilizing around the mode(s) of the posterior. If the chain periodically jumps from one region of the parameter space to the other with a stationary behavior in the meantime, it can be an sign of multimodality in the posterior distribution. If multiple chains starting from different regions of the parameter space converge to the same distribution and overlap with each other, it is also a good indication of convergence. It can also be useful to determine the burn-in period visually: for each chain, it is the maximum number of iterations necessary for all the parameters to reach a seemingly stationary distribution.

Trace plots can also provide useful visual clues about the mixing behavior of a chain, which is a good tracer of the efficiency of the exploration of the parameter space. A chain with good (or rapid) mixing will not display any long-term trend in its trace[1]. On the contrary, a badly (or slowly) mixed chain will take a long time to go from one region of the parameter space to the other, either because the step size is too low (which results in a trace that displays a wavy pattern) or too high (which results in a trace that stays on the same state for many iterations). The worse the mixing, the more samples will be needed to estimate the target distribution with good accuracy. These situations are illustrated in Figure 8-2. Bad mixing can generally be solved by tuning the width of the proposal distribution, as detailed in Section 6.4.

---

[1] It is usually said that an ideal trace should look like a hairy caterpillar.

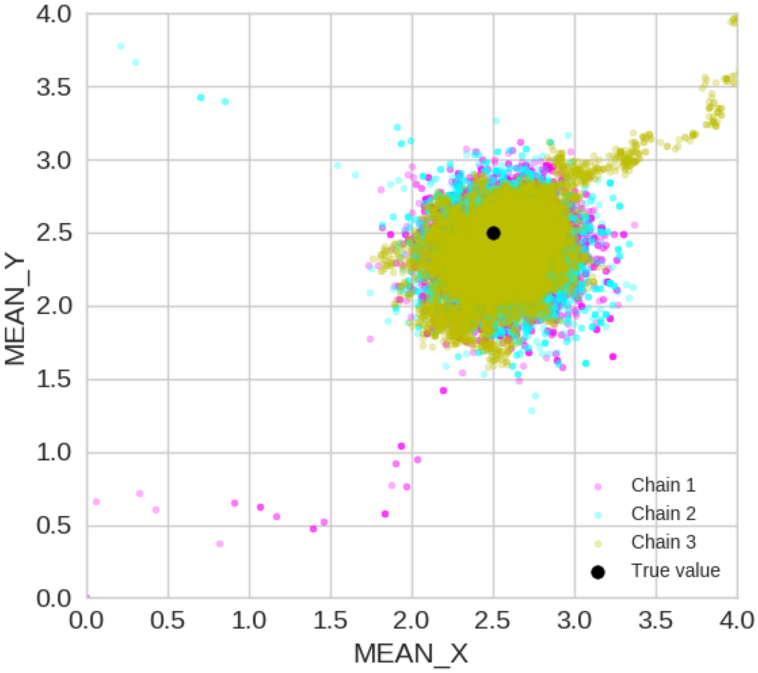


Figure 8-1 – Trace plot of 3 MCMC runs inferring the mean of a bivariate Normal distribution. The data consist in 20 points normally distributed with a mean (2.5,2.5) and a diagonal covariance matrix $\left(\begin{smallmatrix}0.1 & 0\\ 0 & 0.1\end{smallmatrix}\right)$. The chains starting points are overly distributed across the parameter space. The green chain has a step size ten times smaller than the cyan and the magenta chain. After 10000 iterations, all the chains converge to a stationary distribution around the true value.

8.1.2 The autocorrelation plot

Due to its Markovian nature, each iteration of a Markov chain depends on the previous iteration. This introduces a non-zero correlation between distant samples of the same chain. The correlation of a chain of size N with a delayed copy of itself, termed "autocorrelation", is a function of delay (or lag), and is assessed graphically via plotting the autocorrelation function (ACF), denoted ρ_k and defined here below, as a function of lag k:

$$\rho_k = \frac{\sum_{t=1}^{N-k}(\theta_{(t)} - \bar{\theta})(\theta_{(t+k)} - \bar{\theta})}{\sum_{t=1}^{N}(\theta_{(t)} - \bar{\theta})^2} \tag{8.1}$$

By construction, the values of ρ_k are between -1 and 1. At lag 0, the autocorrelation is always 1, which means that the chain is perfectly autocorrelated with itself. The autocorrelation then generally decreases with increasing lag. This diagnostic should be plotted for each free parameter of the model.

A perfect Markov chain should yield independent samples of the posterior distribution. The autocorrelation plot can therefore yield useful information on the mixing rate of the chain: if the ACF rapidly drops to zero, this is a sign of good mixing, which means that convergence to the stationary distribution might be attained quickly. On the contrary, a high degree of autocorrelation with high values of lag means that *on short timescales*, the MCMC samples are unrepresentative of the target distribution. This is an indication of slow mixing: the parameter space is not explored efficiently, and this usually translates into slow convergence. These two situations are illustrated on Figure 8-2.

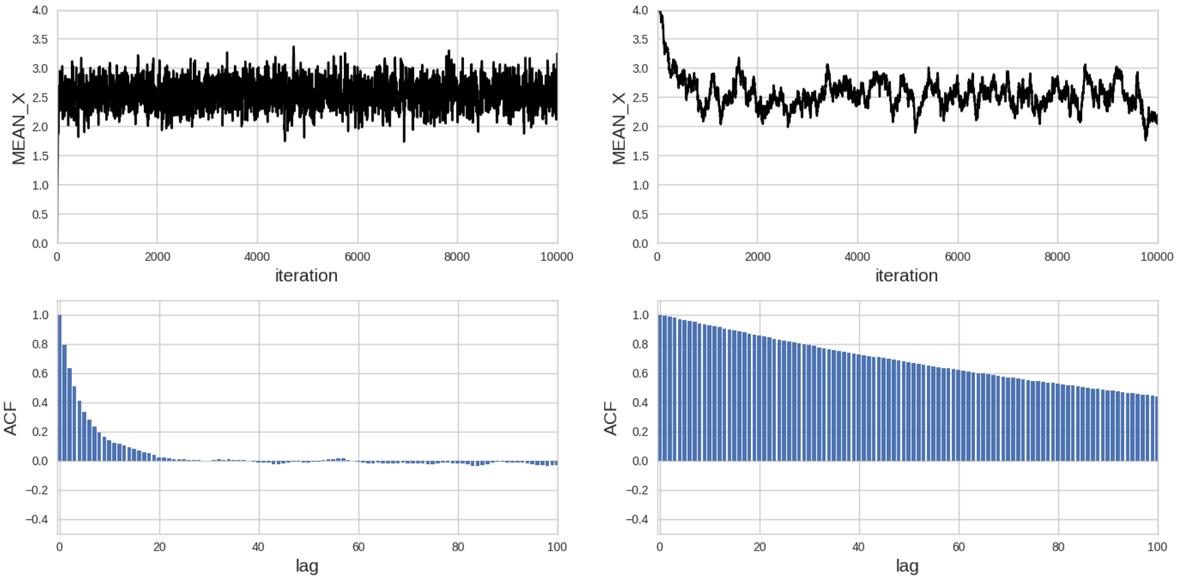
One way of decreasing autocorrelation is to thin the sample, by using only the iterations from a certain lag. Thinning the chain to the lag where the autocorrelation function drops to zero will yield independent samples indeed. But this practice is generally a bad idea, because an unthinned chain always leads to a better estimation of the target distribution than a thinned chain (Maceachern and Berliner 1994). Therefore thinning should be avoided except when facing serious data storage limits. Autocorrelation can also be reduced by tuning the proposal distribution of the Metropolis Hastings algorithm (Blangiardo and Cameletti 2015).

Finally, it is important to note that a high degree of autocorrelation is NOT a sign of a lack of convergence: in fact, the chain can still display a high degree of autocorrelation after having converged to the target distribution. Therefore it is improper to call the autocorrelation plot a "convergence diagnostic" (Sinharay 2003). The PyMC Python package (Patil et al. 2010) provides the autocorrelation plot in its default MCMC diagnostics toolbox.

8.1.3 Geweke's diagnostic

Geweke et al. (1992) devised a convergence diagnostic based on a time series analysis of the Markov chain. Computed for each free parameter of the model, the test divides the chain into 2 intervals a and b of user-defined length called *windows*, one representing the first iterations and the other the later iterations, and compares the empirical mean computed from different segments of the chain using the following statistic, called *Z-score*:

$$Z = \frac{\bar{\theta}_a - \bar{\theta}_b}{\sqrt{Var(\theta_a) + Var(\theta_b)}} \tag{8.2}$$



(a) Chain that mixes well

(b) Chain that mixes slowly

Figure 8-2 – Trace and autocorrelation plots from the output of selected MCMC runs described in Figure 8-1. The left and right plots represent the trace and autocorrelation plot of Chain 1 and 3 respectively. In situation (a), the trace plot indicates that the chain has quickly reached its target distribution. This is confirmed by the shape of the autocorrelation plot, which quickly drops to zero. In situation (b), the chain has also converged to the same distribution, but displays the typical wavy pattern revealing a slow mixing (here due a step size that is too small). This is confirmed by the apparent long trend in the shape of the autocorrelation plot.

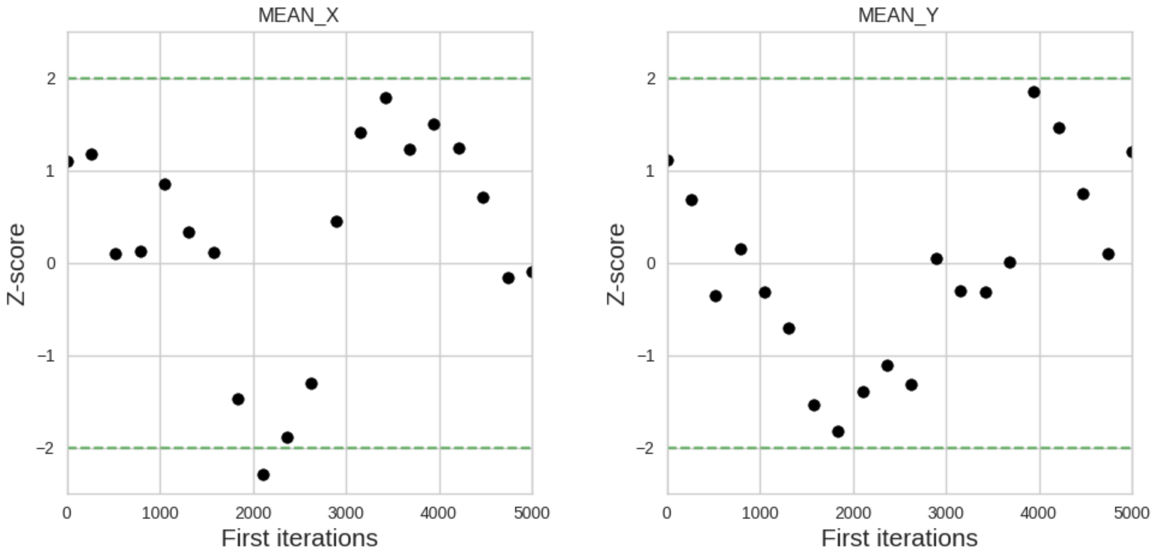


Figure 8-3 – Results of Geweke's diagnostic for the MCMC output of Chain 3, in the example described in Figure 8-1. All the Z-scores computed from the first intervals using more than 2300 iterations fall within 2σ of zero. This suggests that we can consider the first 2300 iterations of the chain as burn-in.

where $\bar{\theta}_a$, $Var(\theta_a)$ and $\bar{\theta}_b$, $Var(\theta_b)$ are the empirical means and asymptotic spectral variances of windows a and b respectively, determined by methods of spectral density estimation (also known as *Fourier analysis*).

If the chain has converged to a stationary distribution, then the two means should be equal and Geweke's Z-score, by construction, should converge asymptotically to a normal distribution with mean 0 and unit standard deviation $N(0, 1)$. Z-scores that are beyond 2σ of zero indicate a failure of convergence, which means that more iterations are needed.

A common practice, as first suggested by Geweke et al. (1992), is to compare the mean of the first 10% and the last 50% of the chain. One can also use Geweke's diagnostic to determine the burn-in period of the chain: the test is performed by comparing various initial windows, and the last 50% of the chain, as illustrated in Figure 8-3. The burn-in period is then considered to be the smallest portion of the chain that passes the test for all the free model parameters. Geweke's diagnostic is implemented in the PyMC Python package (Patil et al. 2010).

8.1.4 The Gelman Rubin test

One of the most popular convergence diagnostics is a test proposed by Gelman and Rubin (1992). Given m chains $\{\theta^j_{(t)}\}$ ($j = 1, ..., m$ and $m \geq 3$, and typically ~ 10), each of length n after discarding burn-in ($t = 1, ..., n$) and with different starting points, the test compares the variance between the mean values of the m chains B and the mean of the m within-chain variances W:

$$B = \frac{n}{m-1}\sum_{j=1}^{m}(\bar{\theta}_{\cdot}^{j} - \bar{\theta}_{\cdot\cdot})^2 \tag{8.3}$$

$$W = \frac{1}{m}\sum_{j=1}^{m}\left[\frac{1}{n-1}\sum_{i=1}^{n}(\theta_{(i)}^{j} - \bar{\theta}_{\cdot}^{j})^2\right] \tag{8.4}$$

Where $\bar{\theta}_{\cdot}^{j} = \frac{1}{n}\sum_{t=1}^{n}\theta_{(t)}^{j}$ is the mean value of chain j, and $\bar{\theta}_{\cdot\cdot} = \frac{1}{m}\sum_{j=1}^{m}\bar{\theta}_{\cdot}^{j}$ is the average value over the m chains.

An overestimate of the true marginal posterior variance is given by the unbiased estimator:

$$\hat{V} = \frac{n-1}{n}W + \frac{1+m}{nm}B \tag{8.5}$$

Finally convergence is estimated using the Potential Scale Reduction Factor (PSRF) $\hat{R}$:

$$\hat{R} = \frac{\hat{V}}{W} \tag{8.6}$$

We consider that convergence has been reached if $\sqrt{\hat{R}} < 1.1$ for all model parameters (Brooks and Gelman 1998). If convergence has not been reached, more iterations are performed until the criterion is met. We use the Gelman Rubin diagnostic implemented in this form in the PyMC package (Patil et al. 2010), to perform our convergence tests.

8.2 The posterior distribution

The goal of Bayesian inference is to reconstruct the shape of the posterior distribution (or simply *posterior*) of a set of parameters, given some observed data. When multiple parameters are involved, the posterior distribution is called the *joint posterior distribution*. The joint posterior, constructed from the combined samples of all the Markov chains taken after burn-in, can be visualized to reveal correlations between parameters.

When inferring the posterior distribution of a d-dimensional random variable $\theta = \{\theta_1, \ldots, \theta_d\}$, the *marginal posterior* for parameter θ_k is found by integrating the joint posterior over all the other parameters:

$$P(\theta_k|D) = \int P(\theta_1, \ldots, \theta_d|D)\mathrm{d}\theta_1 \ldots \mathrm{d}\theta_{k-1}\mathrm{d}\theta_{k+1} \ldots \mathrm{d}\theta_d \tag{8.7}$$

The posterior distribution can also be summarized by point estimates, by computing the mode[2] (i.e. the global maximum of the PDF, or the set of local maxima if the posterior has

---

[2] The mode can also be called the Maximum A Posteriori (MAP).

several modes), the mean and the median of the distribution for example. However, point estimates are usually deprecated by the Bayesian statistics community, because the very essence of Bayesian inference is to use the full information contained within the joint posterior.

The posterior of multiple parameters can be plotted by displaying the correlations between pairs of parameters, as well as the marginal posterior for each parameter. This can be done automatically in Python using the `pairplot` tool of the `seaborn` plotting package[3] (Waskom et al. 2016) for example, or the `corner plot`[4] code of Foreman-Mackey (2016). I use a custom code for plotting joint posteriors from a `.csv` file containing the trace information. An output example is shown on Figure 8-5 from the results of the bivariate Gaussian inference.

8.2.1 Credible regions

A more useful way of summarizing posteriors is to construct intervals (for unidimensional problems) or regions (for multidimensional ones) that capture their variations. A $100(1-\alpha)\%$ *credible region* C_α is the subset of the posterior $P(\theta|D)$ that verifies:

$$\int_{C_\alpha} P(\theta|D)\mathrm{d}\theta \geq 1-\alpha \tag{8.8}$$

There are an infinite number of credibility regions that cover the same probability for a given posterior. The credible region with the smallest volume is called the *highest posterior density region* (HPD or HDR, for *highest density region*). By construction, the probability density of every points inside the HPD region is greater than that of any point outside the region (Box 1992). This point is illustrated in Figure 8-4. HPD regions are especially useful when the posterior distribution is not symmetric. If the posterior is unimodal, the HPD region consists in a single region of the parameter space. But if the posterior is multimodal, it can be a set of disjoint regions whose definition is usually more computationally intensive. Several algorithms have been proposed in the literature to build HPD regions in various cases (cf Hyndman 1996; Turkkan and Pham-Gia 1997), and some utility functions like `hpd`[5] for the `PyMC` package in Python (Patil et al. 2010) or `HPDregionplot`[6] for the `emdbook` package in R (Bolker 2008) can return or plot the boundaries of HPD regions.

Bayesian credible regions are often confused in the literature with frequentist *confidence regions*, essentially because for many simple problems the two regions are indistinguishable numerically. However, it is important to distinguish between these concepts, not only because they can be different in some situations (cf Jaynes and Kempthorne 1976; VanderPlas 2014), but the very question they answer is different. For Bayesians, the observed data are fixed, and the model parameters are considered random variables. Thus, a 95% credible region can be interpreted the following way: “given the data and the prior, there is a 95% probability that the true parameter lies within this region”.

The interpretation of confidence regions is much more counter-intuitive, and is the source of multiple confusions, even for trained statisticians. In fact, for frequentists, the parameters

---

[3] http://seaborn.pydata.org/generated/seaborn.pairplot.html

[4] https://github.com/dfm/corner.py

[5] https://github.com/pymc-devs/pymc3/blob/master/pymc3/stats.py

[6] https://www.rdocumentation.org/packages/emdbook/versions/1.3.9/topics/HPDregionplot

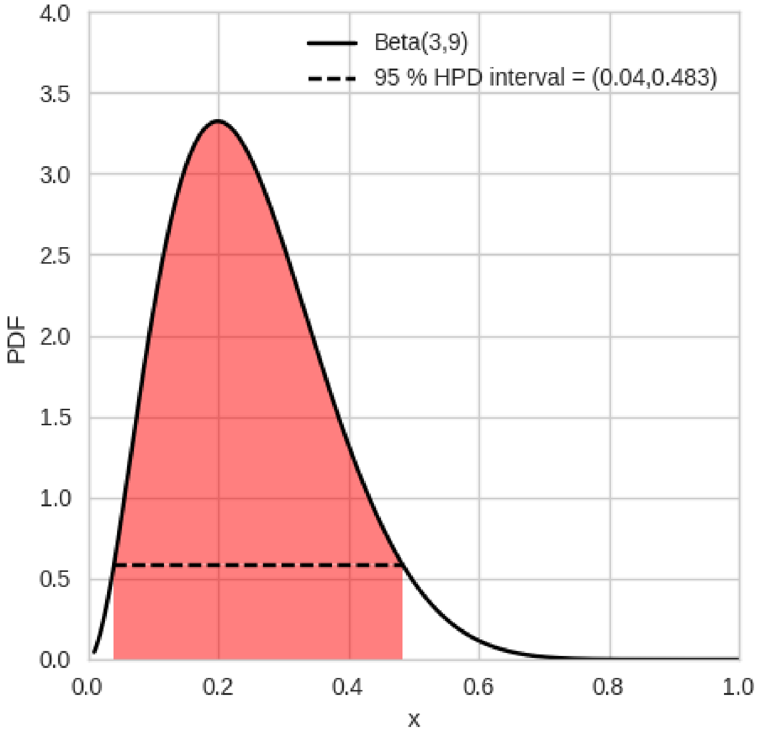


Figure 8-4 – Beta PDF with the 95% HPD interval. The red area under the curve represents the 95% probability contained within the interval.

of the model are fixed, and the data, as well as the confidence regions, are considered random variables. Therefore, a 95% confidence region can be interpreted the following way: "if the experiment is repeated many times and produces a distribution of observed data, 95% of the time the confidence region contains the true parameter". When confronted with a single dataset like in our case, this has real consequences for the interpretation of inference results: as Briggs (2012) remarks, "all we can say about a given confidence interval, calculated from the data at hand, is that a (metaphysical) parameter either lies in the interval or that it does not"[7]. Which lead some Bayesians such as Jaynes and Kempthorne (1976) or VanderPlas (2014), to conclude that in this situation, confidence intervals are useless, and that Bayesian credible regions ought to be preferred.

8.2.2 Constructing credible regions for joint posterior plots

In order to plot user-specified HDP regions on the joint posterior from the output samples of an MCMC run, I coded a script that approximates the bivariate PDF by a normalized two-dimensional histogram, and then computes the contours corresponding to a contained volume of probability of the total volume under the surface of the PDF. An output plot is illustrated in Figure 8-5 for the bivariate gaussian example, along with the 68%, 95% and 99% HDP regions.

---

[7] Which is a tautology

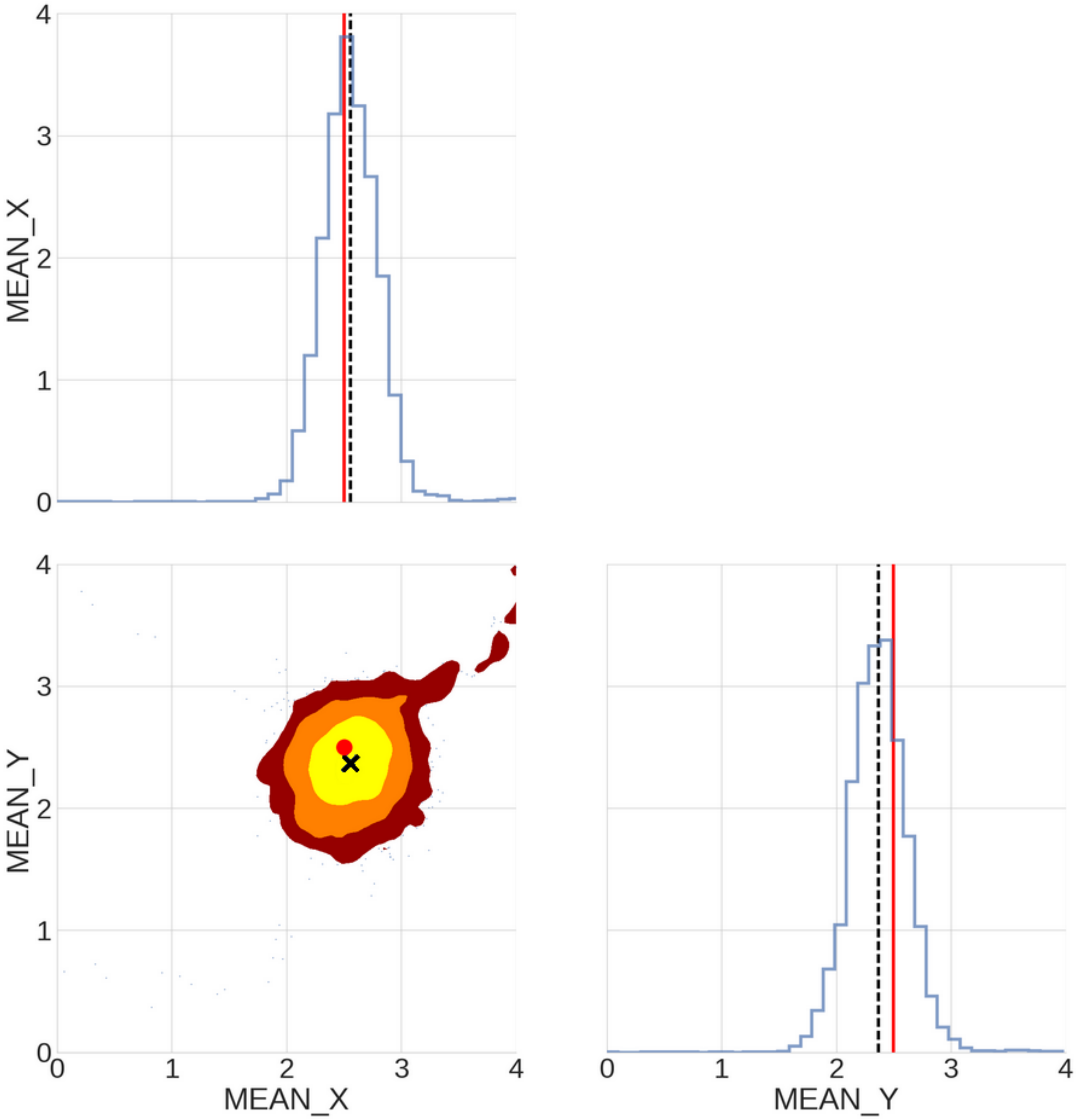


Figure 8-5 – Joint posterior plot from the output of the 3 MCMC run inferring the mean of a bivariate Gaussian, whose setting is described in Figure 8-1. The dark red, orange and yellow areas in the contour plots represent the 99%, 95% and 68% HPD regions respectively. The black crosses and red dots are the mean of the posterior and input true value respectively. In the marginal posterior panels, the black dotted and red lines represent the posterior mean and the true value. At then end of the MCMC run, these two values are very close to each other, which means that the inference has been successful.

Part IV

Inferring the properties of galaxies

Chapter 9

Testing the pipeline

> *The universe is computing its own destiny.*
>
> James Gleick, *The Information: A History, a Theory, a Flood*

Abstract

In this chapter[1], I use the tools and concepts introduced in previous chapters to design a pipeline capable of inferring the size and luminosity of galaxy populations in a simplified situation, using simulated images of a Canada-France-Hawaii Telescope Legacy Survey Deep field as input data where all the input parameters of the model and the selection effects are known. The pipeline turns out to be able to yield robust inferences in all the tests we have run, and even performs better than the classical SED fitting approach applied in the same configuration.

Contents

---

[1] Note: This chapter uses extracts and results from Carassou et al. (2017) in some of its sections.

9.1 Outline of the tests

We need to test the potential of our pipeline (outlined in Section 4.2) in realistic conditions before applying it to real data. This involves generating simulated images of a well-calibrated galaxy survey in several photometric bands, and using them as input data to constrain the luminosity and size properties of multiple galaxy populations generated by Stuff evolving with redshift. As we perform these numerical experiments, we must also take into account the limitations of our computer cluster, as it is difficult to predict a priori the number of iterations of the MCMC sampler required to infer robustly the input parameters of our model. This implies restricting the number of galaxy populations in these mock input images to limit the dimensionality (and therefore the volume) of the parameter space to explore as well as the time spent in each iteration of the sampler.

We consider two test cases in the following sections: in the first test case (cf Section 9.5.1), the input images contains two types of galaxies, a mix between ellipticals and lenticulars, and late-type spirals, which undergo both luminosity and size evolution. But we will limit the inference to the LF shape and redshift evolution parameters (cf Section 4.6) for both populations, which leads to a total of 10 free parameters to infer. The second test case (cf Section 9.5.2) focuses on a single population of pure bulge ellipticals, but this time the inference is performed on both the LF and the distribution of effective radii (both including the redshift evolution parameters), which leads to 5 free parameters to infer.

From now on, unless stated otherwise, we adopt the following cosmological parameters: $H_0 = 100h.\text{km}.s^{-1}.\text{Mpc}^{-1}$ with $h = 1$, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ (Spergel et al. 2003), and magnitudes are given in the AB system.

9.2 The Canada-France-Hawaii Telescope Legacy Survey

Our forward modeling approach can be potentially applied to any astronomical survey reproducible by Skymaker. But in order to test the accuracy of our method, we first need to emulate a realistic survey image containing a statistically significant sample of galaxies to infer the properties of the populations of galaxies spread on it, which are generated with Stuff using known input parameters. As input data image, we choose to reproduce a full-sized stack of the CFHTLS Deep field (e.g., Cuillandre and Bertin, 2006). The CFHTLS Wide and Deep fields offer carefully calibrated stacks in the optical *ugriz* bands, with excellent image quality thanks to the high sensitivity of the MegaCam instrument. Covering 155 deg^2 on the sky in total, the Wide field allows for a detailed study of the large scale distribution of galaxies. As for the Deep field, which covers 4 deg^2 in total, it benefits from long time exposures (33 to 132 hours), which ensure reliable statistical samples of different populations of bright galaxies up to z ~ 1. Each stack of the CFHTLS Deep field (named D1 to D4 according to their location)

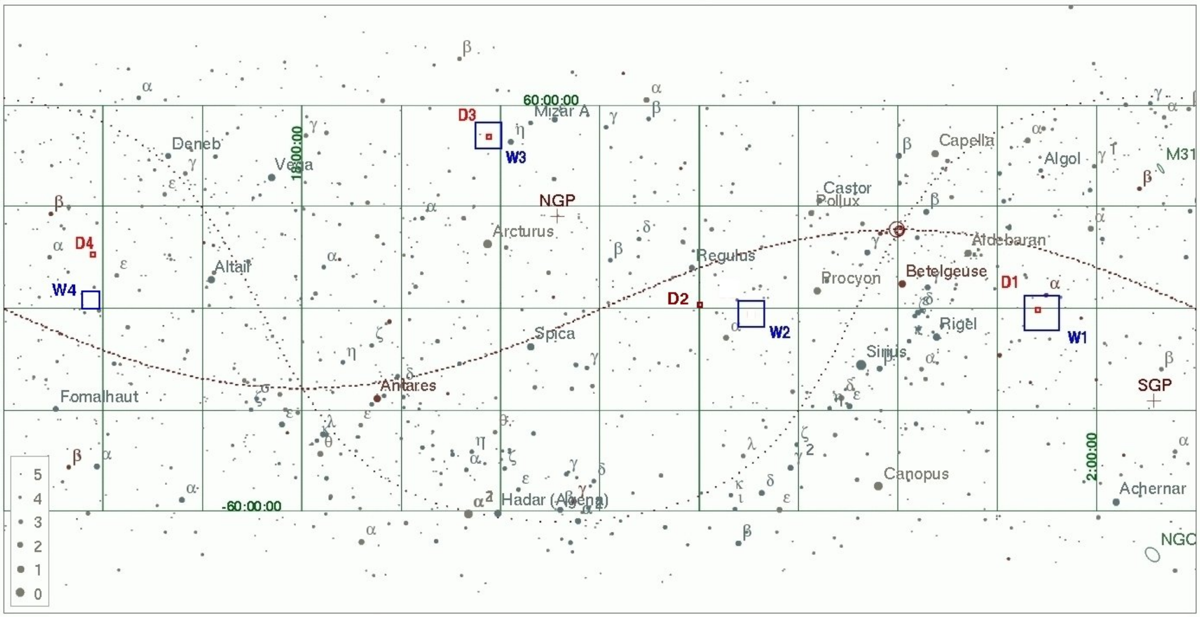


Figure 9-1 – Location of the CFHTLS Deep and Wide fields on the sky. Reproduced from Hudelot et al. (2012).

is a 19,354 × 19,354 pixel image covering 1 deg^2 in total, reaching 80% completeness limits at AB magnitudes of $u = 26.3$, $g = 26.0$, $r = 25.6$, $i = 25.4$ and $z = 23.9$. The final publicly released galaxy catalogs (T0007) contain the photometric redshifts of ~600,000 galaxies down to $i_{AB} = 22.5$ for the Wide survey, and ~240,000 galaxies down to $i_{AB} = 24$ for the Deep survey (Coupon et al. 2009). The locations of the Wide and Deep fields on the sky are illustrated in Figure 9-1. In addition to the observations in optical bands, a fraction of the area of the four Deep fields have been imaged in the JHK_s near-infrared bands by the WIRCam instrument (see Figure 9-2), which represents the WIRCam Deep Survey (WIRDS) (Bielby et al. 2010).

9.3 Providing the input simulated image with realistic galaxy populations

Astronomical survey images contain various galaxy populations. We need to emulate this situation in order to test the behavior of our pipeline in realistic conditions. To do so, we use as input data a simulated CFHTLS Deep image containing two types of galaxies: a population of early-type galaxies (an average between E and S0) of morphological type T=-5 and a population of late-type spirals (Sp) of morphological type T=6. We rely on published results to define these populations. In order to provide Stuff with realistic LF parameters for the simulated populations, we use the results of Faber et al. (2007). Using data from SDSS (York et al. 2000; Blanton et al. 2003), the 2dF Galaxy Redshift Survey (Norberg et al. 2002), COMBO-17 (Wolf et al. 2001; Wolf et al. 2003), and DEEP2 (Davis et al. 2003), Faber et al. (2007) split their distribution of galaxies into two populations by color, using the rest-frame M_B versus $U - B$ color-magnitude diagram: a blue population and a red population. We associate the red and blue populations with our populations of E/S0 and spirals respectively to derive the evolving LF parameters for the two simulated populations. The latter are listed in Table 9.1, and apply

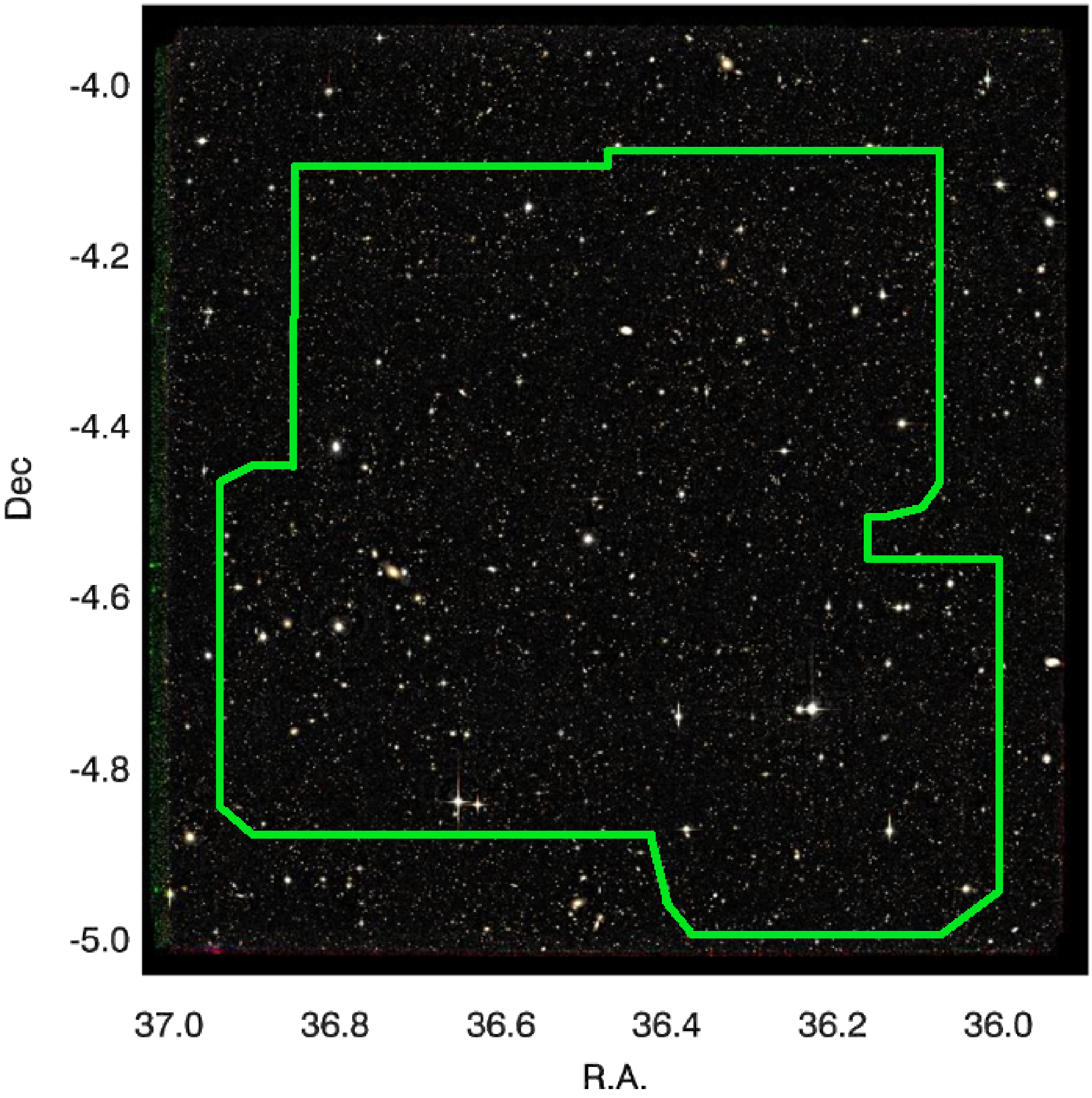


Figure 9-2 – *gri* map of the CFHTLS D1 field representing 1 deg^2 on the sky. The green contour represents the extent of the WIRCam Deep Survey on this field.

Table 9.1 – Luminosity function parameters of the blue and red populations of galaxies at $z = 0.5$ inferred from SDSS, 2dF, COMBO-17, and DEEP2 data, adapted from Tables 3,4, and 6 of Faber et al. (2007).

Population	M_B^*(z=0.5)	$\log_{10}(\phi^*[\mathrm{Mpc}^{-3}])$(z=0.5)	P	Q	α
Red	-20.80	-2.72	-0.46	-1.23	-0.5
Blue	-20.84	-2.55	0.01	-1.35	-1.3

Note: The redshift evolution is fitted by $y = a_0(z = 0.5) + a_1[\log_{10}(1 + z) - \log_{10}(1 + 0.5)]/\log_{10}(2)$, where M_B^* and $\log_{10}(\phi^*)$ are the zero points and Q and P are the slopes respectively.respectively. The LF parameters in Faber et al. (2007) are given for $H_0 = 70$ km.s^{-1}.Mpc^{-1}.

to $z = 0.5$. But several corrections to the values of Faber et al. (2007) are needed before they can get used as input parameters by Stuff, including a magnitude system conversion, a band transformation, and a cosmological correction.

First, in order to obtain the LF parameters for $z = 0$, we use the fitted functions provided by Faber et al. (2007) for each population:

$$M_B^*(z = 0) = M_B^*(z) - \frac{Q\log_{10}(1 + z)}{\log_{10}(2)} \tag{9.1}$$

$$\log_{10}\phi^*(z = 0) = \log_{10}\phi^*(z) - \frac{P\log_{10}(1 + z)}{\log_{10}(2)} \tag{9.2}$$

The absolute magnitude in Equation 9.1 is given in the Johnson system. Because in our simulation Stuff operates in the AB system, we use the AB offset calculated by Frei and Gunn (1994):

$$B_{AB} = B_{Johnson} - 0.163 \tag{9.3}$$

We then apply the transformation equations of Jester et al. (2005) for stars with $R_c - I_c < 1.15$ and $U - B > 0$,

$$B_{AB} = g + 0.39(g - r) + 0.21, \tag{9.4}$$

in order to derive g-band magnitudes:

$$M^*(z = 0)_g = M_B^*(z = 0) - 0.39(g - r) - 0.21 - 0.163. \tag{9.5}$$

We subsequently adopt average colors of $(g-r)_{E/S0} = 0.75$ and $(g-r)_{Sp} = 0.5$ from EFIGI data (de Lapparent, private communication) to derive the value of $M^*(z = 0)_g$ for each population.

In Stuff, the input LF parameters are provided assuming $H_0 = 100h$ km.s^{-1}.Mpc^{-3} with $h = 1$. As Faber et al. (2007) provide their results assuming $h = 0.7$, an additional conversion is needed:

$$M^*_{\text{STUFF}} = M^*(z = 0)_g - 5\log_{10} h \tag{9.6}$$

$$\phi^*_{\text{STUFF}} = \phi^* h^{-3}. \tag{9.7}$$

In STUFF, the LF evolution parameters are defined as:

$$M^*(z) = M^*(z = 0) + M_e \ln(1 + z) \tag{9.8}$$

$$\log_{10} \phi^*(z) = \log_{10} \phi^*(z = 0) + \phi_e \log_{10}(1 + z). \tag{9.9}$$

Combining Equations 9.1 and 9.2 with Equations 9.8 and 9.9 respectively, we obtain:

$$M_e = \frac{Q}{\ln(10) \log_{10}(2)} \tag{9.10}$$

$$\phi_e = \frac{P}{\log_{10}(2)}. \tag{9.11}$$

The values of P and Q listed in Table 9.1 are used to derive the LF parameters for the populations of E/S0 and Sp. *In fine*, the $\phi^*(z = 0)$ of each population is reduced by a factor ten to limit computation time. The final LF parameters are listed in Table 9.5.

9.4 Configuration of the pipeline

To estimate the posterior distribution $P(\theta|D)$ defined in Equation (6.1) in a reasonable amount of time, one must explore the parameter space in a fast and efficient way. For our purposes, we designed a custom sampling procedure, described in Algorithm 3, based on the MCMC AP algorithm (cf Section 6.4.1) coupled with the simulated annealing algorithm (cf Section 6.5). Bayes' theorem (cf Equation 6.1) written in log probability space provides us with the function that we are looking to maximize:

$$\ln P(\theta|\mathcal{D}) \propto \ln P(\theta) + \ln P(\mathcal{D}|\theta), \tag{9.12}$$

where $\mathcal{D}$ is the input data, $P(\theta|\mathcal{D})$ is the posterior, $P(\theta)$ is the prior defined in section 9.4.2, and $P(\mathcal{D}|\theta)$ is the auxiliary likelihood defined in Equation 7.21.

9.4.1 Temperature evolution

Because of the intrinsic stochasticity of our model, many realizations of the model at the same state $\theta_{(t)}$ can lead to many $\ln P(\theta_{(t)}|\mathcal{D})$ values. Therefore, artificial local maxima of the target distribution appear, because each iteration relies on a single realization of the model. The simulated annealing algorithm, as detailed in Section 6.5, was designed to find the global maximum of the target distribution without knowing the posteriori distribution, and this requires us to lower τ in a user-defined scheme. But our goal is distinct as we need to freely explore the parameter space landscape in order to estimate the full posterior distribution. The main constraint for τ is to be comparable to the posterior density difference resulting from the jump. Here we define it as the root mean square (RMS) of the current state, as suggested by Mehrotra et al.

```
Initialize parameters $\theta_{(0)}$ from prior distribution ;
Initialize covariance matrix and temperature ;
for $t = 0$ to $T$ do
    Every $S$ iterations:
        Update covariance matrix and temperature;
    Propose new state $\theta^*$ from proposal distribution;
    while $\theta^*$ is outside the prior bounds do
        Propose another state
    end
    Compute $\ln P(\theta^*|\mathcal{D})$ from proposed state (Equation (9.12))
    if $\ln P(\theta^*|\mathcal{D}) \geq \ln P(\theta_{(t)}|\mathcal{D})$ then
        Accept the jump
    else
        Compute acceptance probability $a$ ;
        Draw uniformly distributed random number $R_N$ in the interval $[0, 1]$ ;
        if $R_N < a$ then
            Accept the jump
        else
            Refuse the jump
        end
    end
end
```

Algorithm 3: Proposed sampling algorithm based on the AP algorithm (Haario et al., 1999).

(1997). In that scheme, a high noise level or a small difference between the proposed and the current state leads to a higher probability of jumping to this state. The temperature is computed every S iterations by running an empirically-defined number of realizations N_R of the model at the current state $\theta_{(t)}$, storing every $\ln P(\theta_{(t)}|\mathcal{D})$ values returned in a vector, and computing the standard deviation of the resulting distribution. In the application below, we find that 20 realizations are sufficient to give a reasonable estimate of the RMS (cf Figure 9-6) and that the temperature quickly reaches a stationary distribution at a relatively low level of $\tau \simeq 30$, after the first few 10^3 iterations (cf Figure 9-7).

9.4.2 Prior

As we have seen in Chapter 6, in any Bayesian inference problem, the choice of the prior distribution $P(\theta)$ is of crucial importance, because different prior choices can result in different posterior distributions from the same data. Without any information on what parameter values most probably explain our data, we adopt non-informative, uniform priors for the free parameters of all the considered models (cf Equation (6.2)), with bounds defined in Table 9.4. The bounds are chosen to prevent the pipeline from exploring non physical domains, such as a very steep LF faint end, which leads to an unreasonably high number of generated galaxies and dramatically increases the computing time, or situations in which no galaxies are generated by the model. We select the least constraining prior possible, which corresponds to a large interval around generally accepted values, such as the values reviewed in de Lapparent et al. (2003) for example.

9.4.3 Initialization of the chains

The initial state $\theta_{(0)}$ of each chain is drawn randomly from the prior distribution. The initial position will only affect the speed of convergence, because the final distribution shall not depend on the initial position, if the chain converges. The initial temperature is then computed from this state. As for the proposal distribution, it is initialized so that no direction in the parameter space is preferred by the sampling algorithm at first. The initial covariance matrix is therefore diagonal, whose non-zero elements are set to:

$$C_{ii} = \frac{u_i - l_i}{E} \; \forall i \in [1, N_p], \tag{9.13}$$

where u_i and l_i are respectively the upper and lower bounds of the prior distribution for parameter i, N_p the number of parameters, and E a value set empirically to 200 in order to ensure reasonable acceptance rates at the beginning of the chain. According to Haario et al. (1999), the adaptive nature of the algorithm implies that the choice of E should not influence the output of the chain.

9.4.4 Image simulation configuration

We simulate one stack of the Deep field in three bands: MegaCam u and i from the CFHTLS, and the WIRCam K_s, well-separated filters chosen in order to maximize the information on the stellar content of the mock galaxies. In fact, the $u-$band is appropriate to trace the emission of young stellar populations, the K_s infrared band is relevant for the emission of old stellar populations, and i is an optical band with a mean wavelength at midpoint between the two other filters (cf Figure 4-3). In accordance with CFHTLS product conventions, the image exposure time is normalized to one second and the AB magnitude zero-point is 30. The SkyMaker PSF model for the CFHTLS image is generated within the software. The *aureole* simulation step (cf Section 5.2.1) is deactivated to speed up the image generation process. For the same reason, we exclude from the Stuff list all galaxies with apparent magnitudes in the reference band brighter than 19 or fainter than 30, in order to avoid simulating both very large and very numerous galaxies. There is no stellar contamination, as Stuff does not yet offer the possibility to simulate realistic star fields. The overall characteristics of the simulated images are summarized in Table 9.2.

9.4.5 Source extraction configuration

SExtractor is configured according to the prescription of the T0007 CFHTLS release documentation (Hudelot et al., 2012). We use it in double image mode, with the i-band image as the detection image, and the background is estimated and subtracted automatically with a 256×256-pixels background mesh size. In order to optimize the detectability of faint extended sources, detection is performed on the images convolved with a 7×7 pixels Gaussian mask having a full width at half maximum (FWHM) of three pixels, that approximates the size of the PSF and acts as a matched filter. Finally, the detection threshold is set to 1.2 times the (unfiltered) RMS background noise above the local sky level.

In order for the results concerning faint sources near the detection limit not to depend too closely on the details of noise statistics, all negative fluxes and radii are clipped to 0 after

Table 9.2 – Imaging characteristics of the CFHTLS+WIRDS used for SkyMaker

Passband	u	i	K_s
Image size [pixels]	19354×19354	19354×19354	19354×19354
Effective gain [e^-/ADU]	74590	6807	2134
Well capacity [e^-]	∞	∞	∞
Saturation level [ADU]	6465	4230	110884
Effective read-out noise [e^-]	4.2	4.2	30
Total exposure time [s]	1	1	1
Zero-point magnitude [“ADU/s”]	30	30	30
Effective wavelength [μm]	0.381	0.769	2.146
Sky level [AB mag/arcsec^2]	22.2	20.0	15.4
Seeing FWHM [arcsec]	0.87	0.76	0.73

Table 9.3 – Parameters of the dynamic range reduction function

Filter	u	i	K_s
$\sigma_{\text{FLUX\_AUTO}}$	3.4	3.6	54.0
$\sigma_{\text{FLUX\_RADIUS}}$	3.5	2.7	2.6
κ_c	10	10	10

extraction.

9.4.6 Dynamic range compression function configuration

To perform the dynamic range compression as defined in Section 7.3.1, we need an estimate of the noise level in the conditions of a CFHTLS Deep field. To that end, we use the population of $\sim 10^4$ pure bulge elliptical galaxies described in Section 9.5.2 and apply the recipe described in Section 7.3.1. The resulting parameters for the dynamic range reduction function in the uiK_s filters are summarized in Table 9.3. For the various cases considered in this article, we use for all galaxy populations the $\sigma_{\text{FLUX\_AUTO}}$ and $\sigma_{\text{FLUX\_RADIUS}}$ values measured for the elliptical galaxies.

9.4.7 Optimizing the run time: parallelization

By construction, our sampling procedure based on MCMC (cf Chapter 6) cannot be parallelized, because the knowledge of the $n-1_{th}$ iteration is required to compute the n_{th} iteration. We can, however, parallelize the process of source extraction and, most importantly, image simulation. In fact, we find in performance tests that the pipeline runtime is largely dominated by the image generation process (cf Figure 9-4), and that the image generation time scales linearly with the area of the simulated image (cf Figure 9-5). Even if SkyMaker is itself largely multithreaded, simulating a single image per band containing all the sources for every iteration would make this problem computationally unfeasible in terms of execution time. In order to limit the runtime of an iteration, the image making step is therefore split into $N_{sub} \times N_f$ parallel small square patches, as illustrated in Figure 9-3, where N_f is the number of filters fixed by the

Table 9.4 – Uniform prior bounds for the parameters of the luminosity and size functions, and their evolution with redshift.

Parameter	ϕ^*	M^*	α	ϕ_e	M_e	M_{knee}	r_{knee}	γ_b
lower bound	10^{-7}	-22	-2.5	-3	-2.5	-21	0	-2
upper bound	10^{-2}	-17	0	2	0	-19	3	0

Note: All the parameters above are given for $H_0 = 100$ km.s^{-1}.Mpc^{-3}.

observed data and N_{sub} the user-defined number of patches per band. Both quantities must be chosen so that their product optimizes the resources used by the computing cluster.

We start with N_f input catalogs generated from the model, each containing a list of sources' positions in a full-sized square field of size L_f, as well as their photometric and size properties. The sources are then filtered according to their spatial coordinates and dispatched to their corresponding patch. Each patch has a size $L_f/\sqrt{N_{sub}}$, where N_{sub} is a square number. In practice, the sources are extracted from a box 150 pixels wider than the patch size in order to include the objects outside the frame that partially affect the simulated image. All the sources of position (x,y) are within a patch of coordinate $(i, j) \in [0, \sqrt{N_{sub}} - 1] \times [0, \sqrt{N_{sub}} - 1]$ if $x \in [i\frac{L_f}{\sqrt{N_{sub}}} - 150, (i+1)\frac{L_f}{\sqrt{N_{sub}}} + 150]$, and $y \in [j\frac{L_f}{\sqrt{N_{sub}}} - 150, (j+1)\frac{L_f}{\sqrt{N_{sub}}} + 150]$.

As a result, all the sources are scattered through N_{sub} catalog files per band. We then use the HTCondor distributed jobs scheduler (e.g. Thain et al. 2005) on our computing cluster to generate and analyze all the patches at the same time. The flexibility of HTCondor offers many advantages to a pipeline that requires distributed computing over long periods of time. Thanks to its dynamic framework, jobs can be check pointed and resumed after being migrated if a node of the cluster becomes unavailable, and the scheduler efficiently provides an efficient match-making between the required and the available resources. This framework also has its drawbacks, in the form of inherent and uncontrollable latencies when jobs input files are sent to the various nodes.

In our case, each job corresponds to a single patch, and the $N_{sub} \times N_f$ resulting catalogs serve as input files for the jobs. We found that HTCondor latencies represent between 7% and 50% of the run time of each iteration, as illustrated in Figure 9-4 in the context of the application described below (cf Section 9.5).

For each job, the image generation and source extraction procedures are multiprocessed: SkyMaker is first launched simultaneously in every band on the $L_f/\sqrt{N_{sub}}$-sized patch and, when all the images are available, SExtractor is launched in double image mode (cf Section 5.3) in order to extract color information of detected galaxies. HTCondor then waits until all jobs are completed. Finally, the catalog files generated from all the patches are merged into one, so that *in fine*, a single catalog file per band contains all the extracted sources.

9.5 Tests with one and two galaxy populations

9.5.1 Multi-type configuration: luminosity evolution

The simulated image used as input data contains two populations of galaxies generated by Stuff: E/S0 and Sp, whose luminosity function has been defined in Section 9.3. The B/T ratios in the g adopted reference band are determined using the distribution of B/T in g-band as a

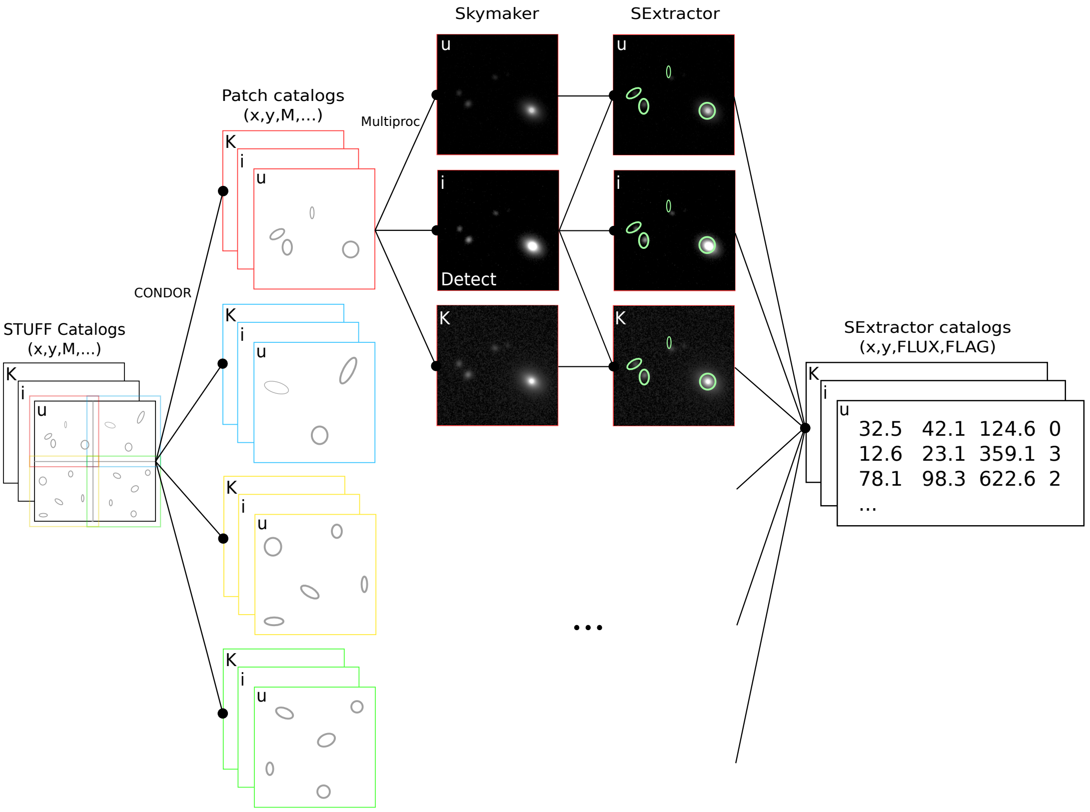


Figure 9-3 – Illustration of the parallelization process of our pipeline, described in detail in Section 9.4.7. Stuff generates a catalog, that is, a set of files containing the properties of simulated galaxies, such as inclination, bulge-to-disk ratio, apparent size, and luminosity. Each file lists the same galaxies in a different passband. The parallelization process is performed on two levels: first, the Stuff catalogs are split into sub-catalogs according to the positions of the sources on the image. These sub-catalogs are sent to the nodes of the computer cluster in all filters at the same time using the HTCondor framework. Each sub-catalog is then used to generate a multiband image corresponding to a fraction of the total field. This step is multiprocessed in order to generate the patches in every band simultaneously. SExtractor is then launched on every patch synchronously, also using multiprocessing. The source detection is done in one pre-defined band, and the photometry is done in every band. Finally, the SExtractor catalogs generated from all the patches are merged into one large catalog containing the photometric and size parameters of the extracted sources from the entire field.

Table 9.5 – Characteristics of the galaxy test populations

Population	SED_b [a]	SED_d	ϕ^* [$h^3 Mpc^{-3}$]	M^*	α	ϕ_e	M_e	B/T	T [b]	$\alpha(T)$	N [c]
Multi-type: E/S0	E	E	0.003	-19.97	-0.5	-1.53	-1.77	0.65	-5	0.0	10447
Multi-type: Sp	E	Scd	1.4e-4	-19.84	-1.3	0.03	-1.95	0.2	6	1.47	28281
Fattening E	E	E	0.0035	-19.97	-0.5	-1.53	-1.77	1.0	-5	0.0	11353

Notes: The LF parameters are given for H_0 = 100 km.s^{-1}.Mpc^{-3}.

[a] The bulge and disk SEDs are Coleman et al. (1980) templates.

[b] de Vaucouleurs (1959) revised morphological type.

[c] Number of sources generated by one realization of Stuff.

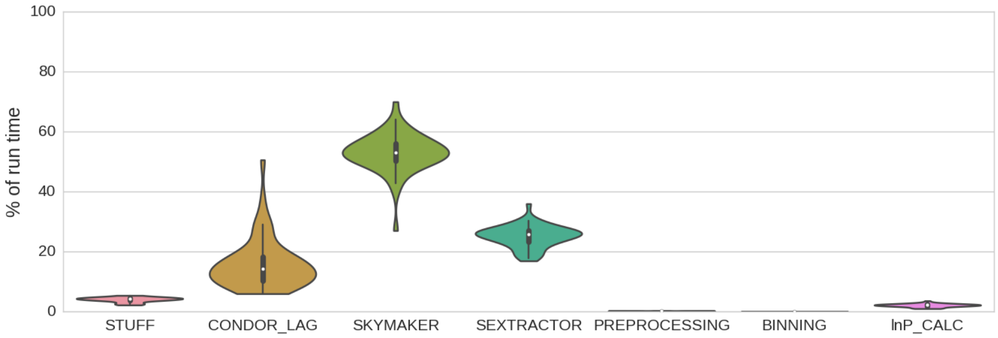


Figure 9-4 – Benchmarking of a full iteration of our pipeline, obtained with 50 realizations of the same iteration. An iteration starts with the STUFF catalog generation (here we consider a case where ~ 55,000 sources spread into two populations of galaxies are produced), and ends with the posterior density computation. The runtime of each subroutine called is analyzed in terms of the fraction of the total runtime of the iteration. In this scheme, the image simulation step clearly dominates the runtime, followed by the source extraction step and the HTCondor latencies. Source generation, pre-processing, binning and posterior density calculation (labeled lnP_CALC), however, account for a negligible fraction of the total runtime.

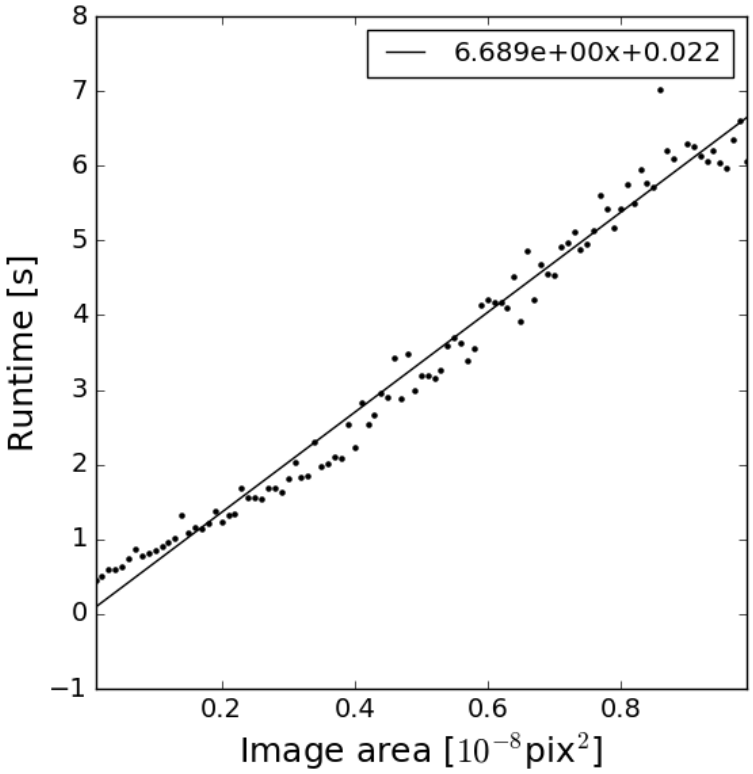


Figure 9-5 – Evolution of the run time of a SKYMAKER run as a function of image area simulated, using the same STUFF parameters for each run.

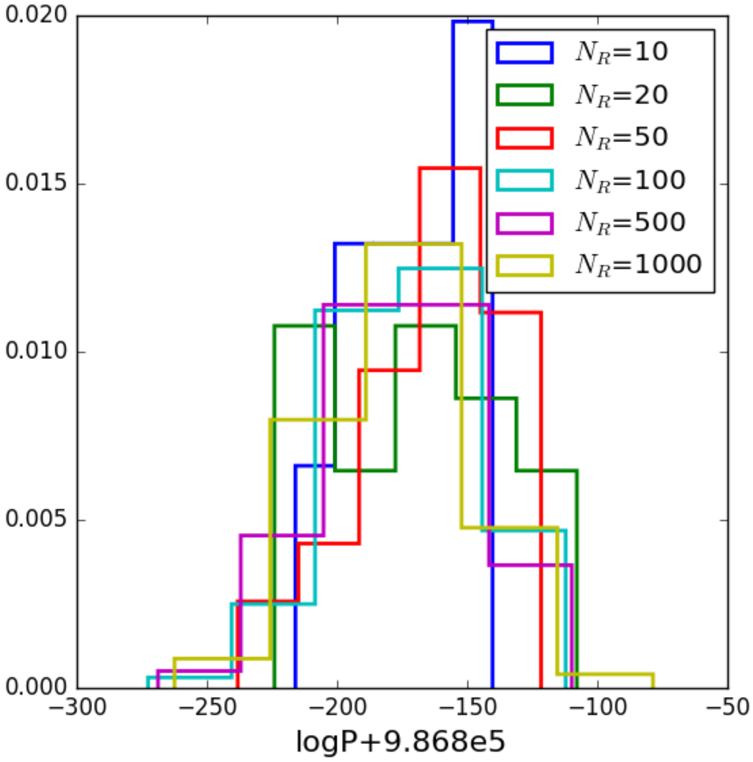


Figure 9-6 – Normed distribution of $\ln P$ for various numbers of realizations N_R of the model. Each distribution is generated in the conditions of the “Fattening E” case, at “true” input values (cf Table 9.5) and with the same seed for galaxy generation in STUFF. Standard deviation of the distributions do not appear to differ significantly. We conclude that 20 realizations of the model are enough to characterize the order of magnitude of RMS.

Table 9.6 – Size parameters for the bulge and disk of each galaxy test population

Disk				Bulge		
β_d	r_d^* [h^{-1}kpc]	γ_d	σ_λ	M_{knee}	r_{knee} [h^{-1}kpc]	γ_b
-0.214	3.85	-0.80	0.36	-20.0	1.58	-1.00

Note: All the parameters above are given for H_0 = 100 km.s^{-1}.Mpc^{-3}.
“b” refers to bulge, and “d” to disk.

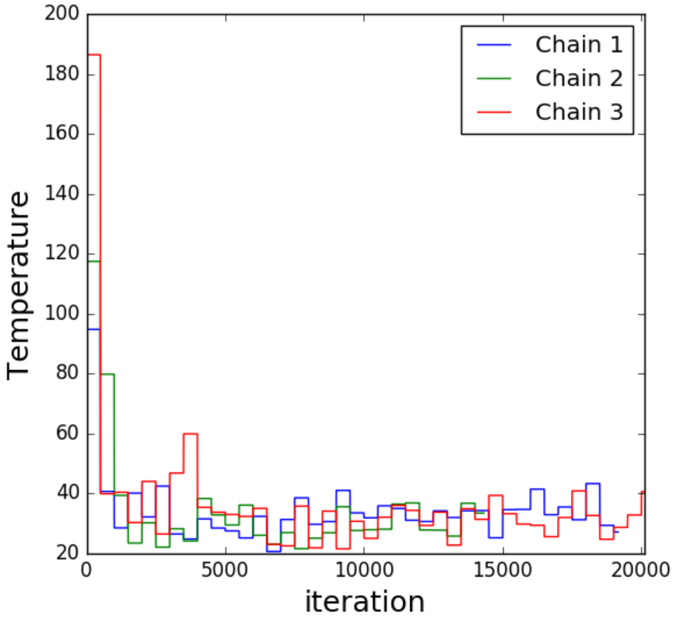


Figure 9-7 – Temperature evolution with the number of iterations of the MCMC process in the "Fattening E" case described in Section 9.5.2. Here the temperature is computed every 500 iterations at current state with 20 realizations of the model. We note that for each chain, the temperature values quickly converge to the level of noise of the model near input values.

function of morphological type from EFIGI data (Baillard et al. 2011, de Lapparent, private communication). To limit run time, the ϕ^* values for each population are set to have $\sim 4 \times 10^4$ galaxies in total generated quickly by Stuff in the field area. In this scheme, we have $\sim 10^4$ E/S0, and $\sim 3 \times 10^4$ Sp. We cannot reproduce the number counts of a real CFHTLS Deep field, because they are dominated by irregular galaxies (Irr) for $18 < i < 26$. Reproducing realistic number counts would imply for Stuff to produce a number of galaxies one order of magnitude higher for E/S0 and Sp, and also to add a population of $\sim 10^5$ Irr, which would lead to unreasonable computing time.

The input parameters used to generate both populations are listed in Table 9.5. The parameters to infer in this case are the 5 evolving LF parameters for each of the populations: ϕ^*, M^*, α, ϕ_e and M_e, that is a total of 10 parameters (note that we do not infer the size distribution and evolution parameters). The observables are the SExtractor FLUX_AUTO in uiK_s, which leads to a three-dimensional observable space. Using 10 bins for each observable, we obtain a total number of 10^3 bins in the observable space. Over the $\sim 4 \times 10^4$ galaxies generated by Stuff, we find that $\sim 2 \times 10^4$ are extracted with SExtractor. The number of extracted galaxies per bin is presented in Figure 7-4.

Results

We run the pipeline on a heterogeneous computing cluster of 7 machines totaling 152 CPU cores. We launched 3 chains in parallel for 18,357, 18,565 and 16,211 iterations respectively with randomly distributed starting points, using 50,400 CPU hours in total. The burn-in phase is estimated by visual examination of the trace plot. All the iterations before the upper and lower envelope of the trace becomes constant for all the chains and for all the parameters simultaneously are discarded as burn-in, which in the case under study corresponds to the first 10^4 iterations. Then convergence on the f last iterations of each chain is assessed based on the Gelman-Rubin test (cf Table 9.7), where f is the minimum length over the 3 chains after burn-in, as the convergence test requires the same number of iterations for all the chains: $f = 6,211$ iterations. Table 9.7 lists the results of the Gelman-Rubin test, which suggest that all the chains have converged to the same stationary distribution.

The final joint posterior distribution is the result of the combined accepted states of all the chains run after burn-in. The posterior PDF plot is shown in Figure 9-8: it contains 3,017 accepted iterations out of 23,132 propositions, corresponding to an overall 13% acceptance rate after burn-in. The graph shows that the "true" input values are all lying within the 68% credible region, which in Bayesian terms means that there is a 68% probability that the model value falls within the credible region, given the data. Summary statistics of the posterior PDF are listed in Table 9.8. As the pipeline generates constraints that are consistent with the input parameters, We therefore conclude that our approach can be used to perform unbiased inference on the photometric parameters of galaxies using two broad classes of galaxy types given non-informative priors.

Moreover, we find in Figure 9-8 some strong correlations or anti-correlation between various pairs of parameters, that are symptomatic of the degeneracies in the parameters for our specific set of observables (fluxes). For example, a strong anti-correlation is found between M^* and M_e in the 2 populations. This can be explained by the fact that a brighter (lower) M^* population at $z = 0$ can be partly compensated by a shallower (higher) redshift evolution.

9.5.2 Fattening ellipticals configuration: size and luminosity evolution

We then test whether our pipeline can also infer the characteristic luminosity and size evolution of galaxies. Because of memory limitations, we perform this test in a simplified framework. We use as input data a CFHTLS image in uiK_s containing $\sim 10^4$ E/S0 (pure bulge) galaxies generated with Stuff. The input photometric parameters are listed in Table 9.5 and those for bulge size are listed in Table 9.6. The parameters to infer are the 5 evolving LF parameters, as well as 3 parameters governing the bulge distribution and evolution: M_{knee}, r_{knee} and γ_b (as defined in Chapter 4). That is a total of 8 parameters. No extinction is included in this case. As the size evolution parameters cannot be retrieved with the photometric information only (FLUX_AUTO), the FLUX_RADIUS parameter of SExtractor for all galaxies in each passband are added to the observables space. This leads to a 6 dimensional observable space. Over the $\sim 10^4$ E generated by Stuff, we find that $\sim 7 \times 10^3$ are found by SExtractor. With 10 bins per observable, this results in a total number of bins of 10^6. The number of extracted galaxies per bin is presented in Figure 7-4.

Results

We run our pipeline with 3 chains in parallel for 18,898, 14,056 and 20,110 iterations respectively, with uniformly distributed starting points, using 19,656 CPU hours in total. The first 10^4 iterations of each chain are discarded as burn-in. Convergence is reached for the last $f = 4,323$ last iterations of each chain, as assessed by the Gelman-Rubin test results displayed in Table 9.7. The resulting posterior distribution is shown in Figure 9-9. It contains 6,287 accepted

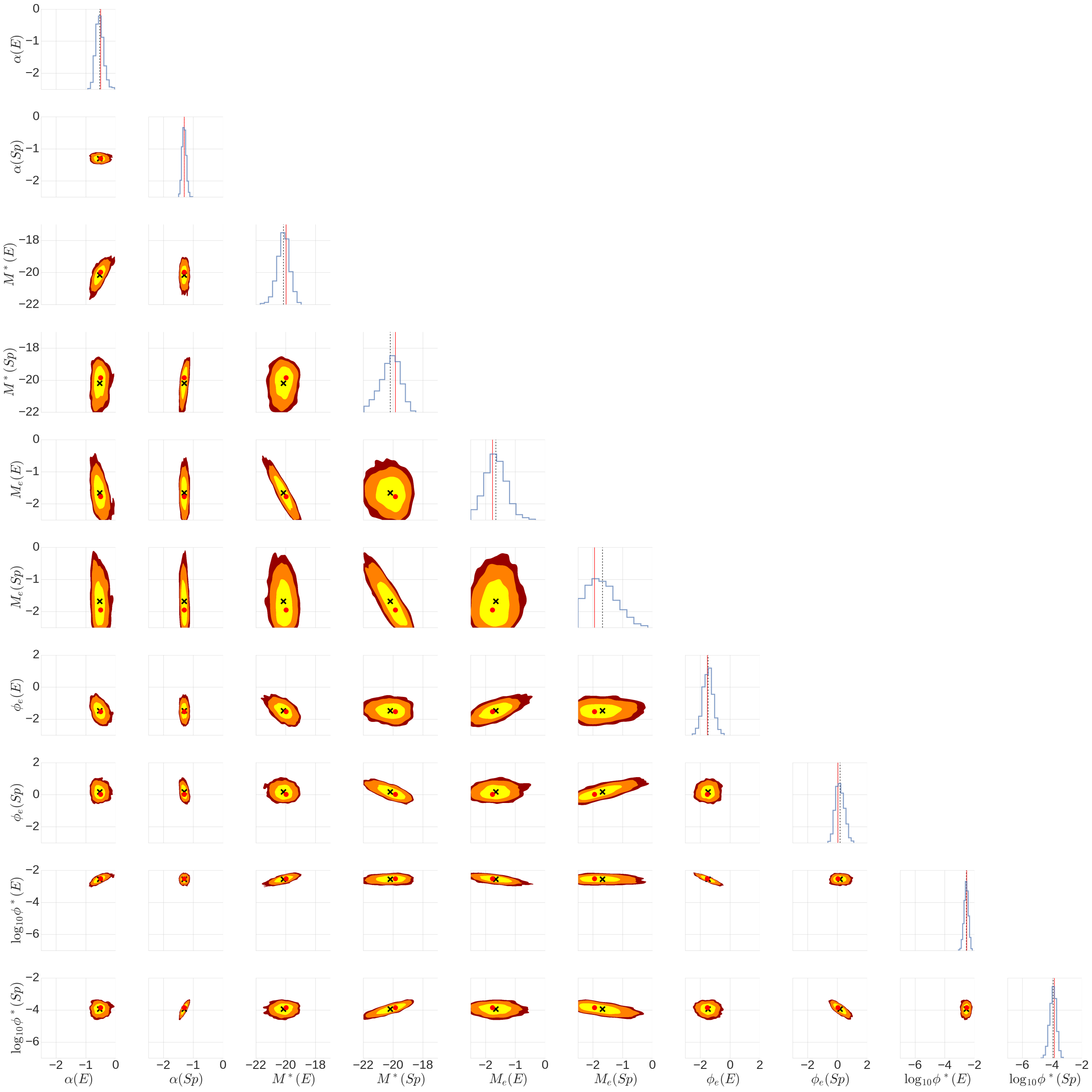


Figure 9-8 – Joint posterior distribution resulting from the "Multi-type" test described in Section 9.5.1. The diagonal plots show the marginal distribution for each parameter (the projection of the posterior onto that parameter). Each panel is bounded by the prior range values. The dark red, orange and yellow areas in the contour plots represent the 99%, 95% and 68% credible regions respectively. The black crosses and red dots are the mean of the posterior and input true value resp. In the marginalized posterior panels, the black dotted and red lines represent the posterior mean and the true value resp.

Table 9.7 – Results of the Gelman-Rubin test

Population	$\log_{10}(\phi^*)$	M^*	α	ϕ_e	M_e	M_{knee}	r_{knee} $[h^{-1}kpc]$	γ_b
Multi-type: E/S0	1.015	1.006	1.014	1.012	1.005	∅	∅	∅
Multi-type: Sp	1.020	1.013	1.028	1.007	1.012	∅	∅	∅
Fattening E	1.013	1.003	1.020	1.003	1.001	1.008	1.010	1.008

Note: The values of $\sqrt{R}$ are obtained using 3 chains for each case, whose burn-in phase for each chain is determined by eye. All values are < 1.1, which is a hint that in each case, all the chains have converged to the same distribution.
The parameters above are given for $H_0 = 100$ km.s^{-1}.Mpc^{-3}.

iterations over $38,064$, which leads to an acceptance rate of 16.5%.

Each marginalized posterior plot exhibits a main mode, with the peak and the mean almost indistinguishable from the input values. The joint posterior distribution shows that the input values are all falling within the 68% credible region. Summary statistics of the posterior PDF are listed in Table 9.8. Here again, our pipeline produces constraints that are consistent with the true parameters. So we conclude that our ABC method can reliably infer the luminosity and size distribution of one population of galaxies without any systematic bias.

The joint posterior PDF also reveals covariances between parameters. For instance, the ϕ^* and ϕ_e parameters are naturally anti-correlated because an increase of ϕ^* (at $z = 0$) can partially be compensated by a steeper decrease of the normalization with redshift, hence a smaller value of ϕ_e.

We also note that the 68% credible region appears to be wide compared to the distance between the input parameters and the mean of the posterior. We suspect that this effect results from the intrinsic stochasticity of the model, estimated by the "temperature" term introduced in the acceptance probability of the Metropolis Hastings, and that contributes to the uncertainty in the posterior.

9.6 Comparison with SED fitting

As demonstrated above, our pBIL method is efficient at recovering the input parameters used to define the luminosity and size evolutions in the mock CFHTLS image. One may wonder how it compares with the classical, less CPU-expensive method for measuring LFs – SED fitting– which provides, from a multi-band photometric catalog, estimates of the photometric redshifts as well as rest-frame luminosities. Luminosity functions can then be derived using independent redshift bins.

A comparison has therefore been carried out by Damien Leborgne. The simulated field used for this comparison is the "Fattening E" sample, with a single population of pure bulges with the Coleman et al. (1980) "E" template. The Z-PEG code (Le Borgne and Rocca-Volmerange, 2002) is applied to the u, i, and K_s photometric catalog obtained with SExtractor in the same configuration as described for the pBIL method. Photometric redshifts are measured together with g-band luminosities for every i-band detected object down to $u_{AB} = 30$. The fits were performed using the whole range of SED templates from Coleman et al. (1980), from E to Irr galaxies. The discrete LFs obtained in each redshift bin were volume weighted with a V_{max} correction at the faint magnitude bins, and a Schechter (1976) function was fitted to the data

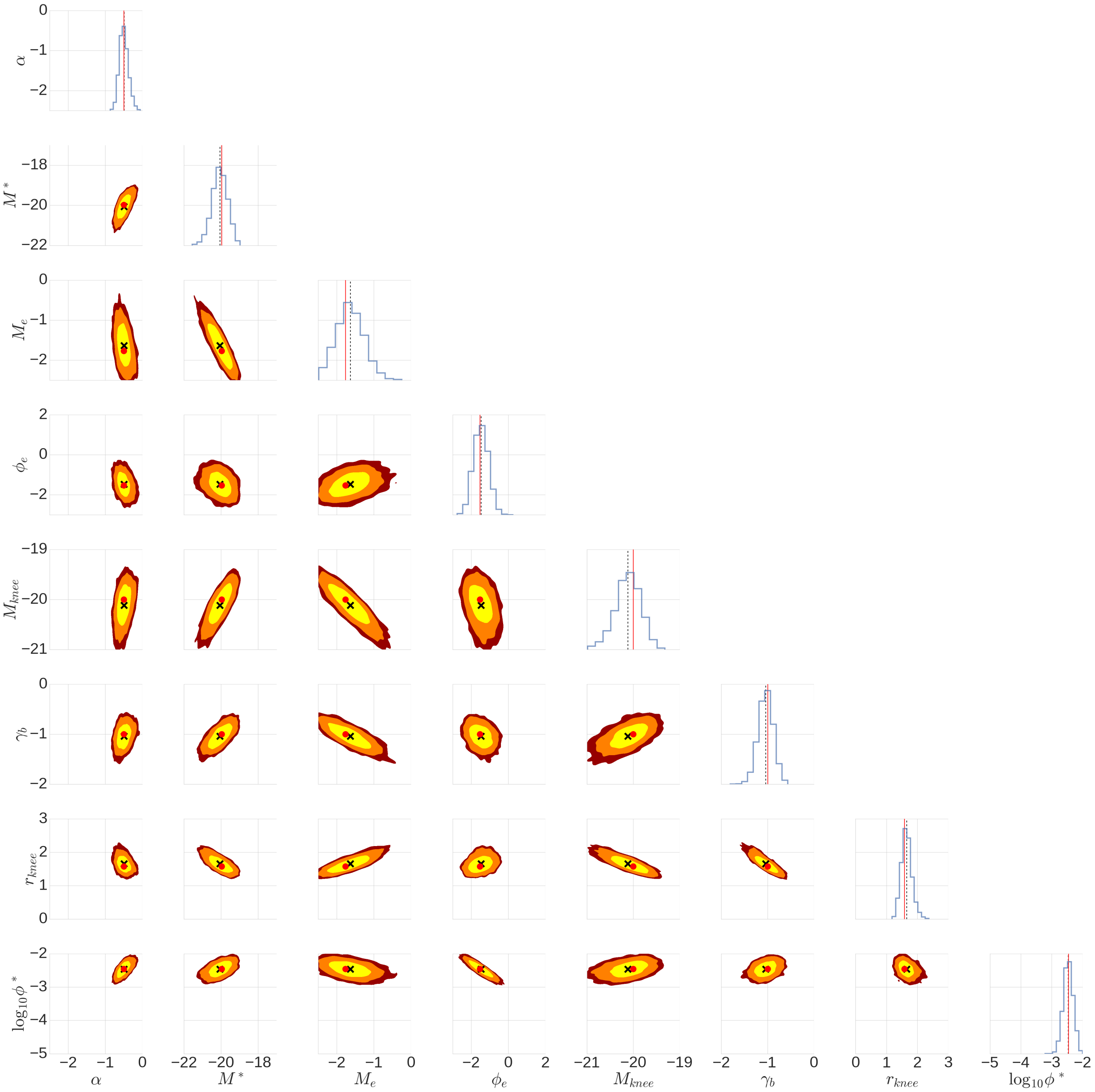


Figure 9-9 – Joint posterior distribution resulting from the “Fattening E” test described in Section 9.5.2. The diagonal plots show the marginal distribution for each parameter. Each panel is bounded by the prior range values. The dark red, orange and yellow areas in the contour plots represent the 99%, 95% and 68% credible regions respectively. The black crosses and red dots are the mean of the posterior and input true value resp. In the marginalized posterior panels, the black dotted and red lines represent the posterior mean and the true value resp.

Table 9.8 – Summary statistics on the marginalized posterior distributions for the galaxy test populations and comparison with the input values.

Population	Parameters	Input value	Mean	MAP [a]	68% interval	95% interval	99% interval
Multi-type: E/S0	$\log_{10}(\phi^*)$	-2.52	-2.56	−2.57	[-2.72,-2.40]	[-2.85,-2.29]	[-2.95,-2.19]
	M^*	-19.97	-20.15	−20.12	[-20.54,-19.74]	[-20.95,-19.38]	[-21.20,-19.08]
	α	-0.5	-0.53	−0.54	[-0.68,-0.43]	[-0.77,-0.24]	[-0.82,-0.06]
	ϕ_e	-1.53	-1.48	−1.37	[-1.82,-1.15]	[-2.16,-0.81]	[-2.35,-0.49]
	M_e	-1.77	-1.66	−1.70	[-1.99,-1.29]	[-2.36,-0.96]	[-2.46,-0.75]
Multi-type: Sp	$\log_{10}(\phi^*)$	-3.85	-3.93	−3.96	[-4.15,-3.67]	[-4.35,-3.54]	[-4.53,-3.40]
	M^*	-19.84	-20.18	−19.81	[-20.68,-19.43]	[-21.59,-19.01]	[-21.93,-18.84]
	α	-1.3	-1.30	−1.33	[-1.37,-1.25]	[-1.42,-1.17]	[-1.44,-1.13]
	ϕ_e	0.03	0.18	0.15	[-0.16,0.47]	[-0.34,0.75]	[-0.48,0.94]
	M_e	-1.95	-1.68	−1.81	[-2.39,-1.29]	[-2.49,-0.87]	[-2.49,-0.19]
Fattening E	$\log_{10}(\phi^*)$	-2.46	-2.46	−2.51	[-2.60,-2.31]	[-2.75,-2.18]	[-2.84,-2.08]
	M^*	-19.97	-20.04	−20.08	[-20.41,-19.63]	[-20.84,-19.23]	[-21.15,-19.12]
	α	-0.5	-0.49	−0.51	[-0.62,-0.40]	[-0.72,-0.25]	[-0.78,-0.13]
	ϕ_e	-1.53	-1.49	−1.56	[-1.84,-1.09]	[-2.22,-0.74]	[-2.41,-0.49]
	M_e	-1.77	-1.65	−1.61	[-2.04,-1.29]	[-2.35,-0.96]	[-2.47,-0.73]
	M_{knee}	-20.00	-20.10	−19.99	[-20.32,-19.79]	[-20.68,-19.58]	[-20.90,-19.46]
	r_{knee}	1.58	1.64	1.56	[1.47,1.77]	[1.34,1.97]	[1.27,2.09]
	γ_b	-1.00	-1.03	−1.07	[-1.18,-0.87]	[-1.31,-0.71]	[-1.50,-0.65]

Notes:The LF parameters are given for $H_0 = 100$ km.s^{-1}.Mpc^{-3}.

[a] Maximum A Posteriori

independently in each redshift bin with a Levenberg-Marquardt algorithm with Φ^*, M^*, and α as free parameters.

Comparison of the evolution with redshift of the LF parameters between the pBIL approach (green dashed line for mean of posterior, and 68% light green shaded region) and the results from SED fitting (red symbols with error bars) are shown in Figure 9-10. As expected, they both roughly follow the trends set by the evolution of the input parameters (blue solid line), with some offsets that can be explained by the fact that SED fitting is done on only three photometric bands. This is clearly a major limiting factor, albeit partly compensated by the choice of the SED templates: the input SED and the templates share a common SED (the "E" SED, even if all SEDs from Coleman et al. (1980) are also used for the SED fitting). We believe that this choice is fair because in the pBIL method, the same set of SEDs was also used for data generation and for the inference of LF parameters.

The significant systematic offset in α from SED fitting compared to the input and pBIL curves in Figure 9-10 shows that the faint-end slope parameters α is poorly estimated, with a significant systematic offset at $z \geq 0.7$. This is caused by the negative input slope (see Table 9.5), which yields few faint galaxies in the sample. Moreover, because of numerous catastrophic outliers in the photometric redshifts (caused by the u,i, K_s-only photometric catalog), there is a mismatch between the true redshift of many faint objects and the redshift bin to which they are assigned. This leads to an underestimate of the error bars on the individual points.

For this comparison of the LF parameters between the two approaches, we had to derive the envelop of the LF parameters as a function of redshift for the trace elements of the MCMC chains within the 68% credible region of the parameters space. Of course, the area appears in Figure 9-10 as much smoother than the individual points derived from SED fitting because the chosen LF model for the inference evolves smoothly with redshift. Still, it is remarkable that the region is tight and almost centered on the values of the true parameters at all redshifts. This is because the various covariances between the five LF parameters of the model tend to narrow down the shaded areas in these graphs, therefore implying that galaxies at all redshifts in the images contribute to constrain the parameters of the model in the pBIL approach.

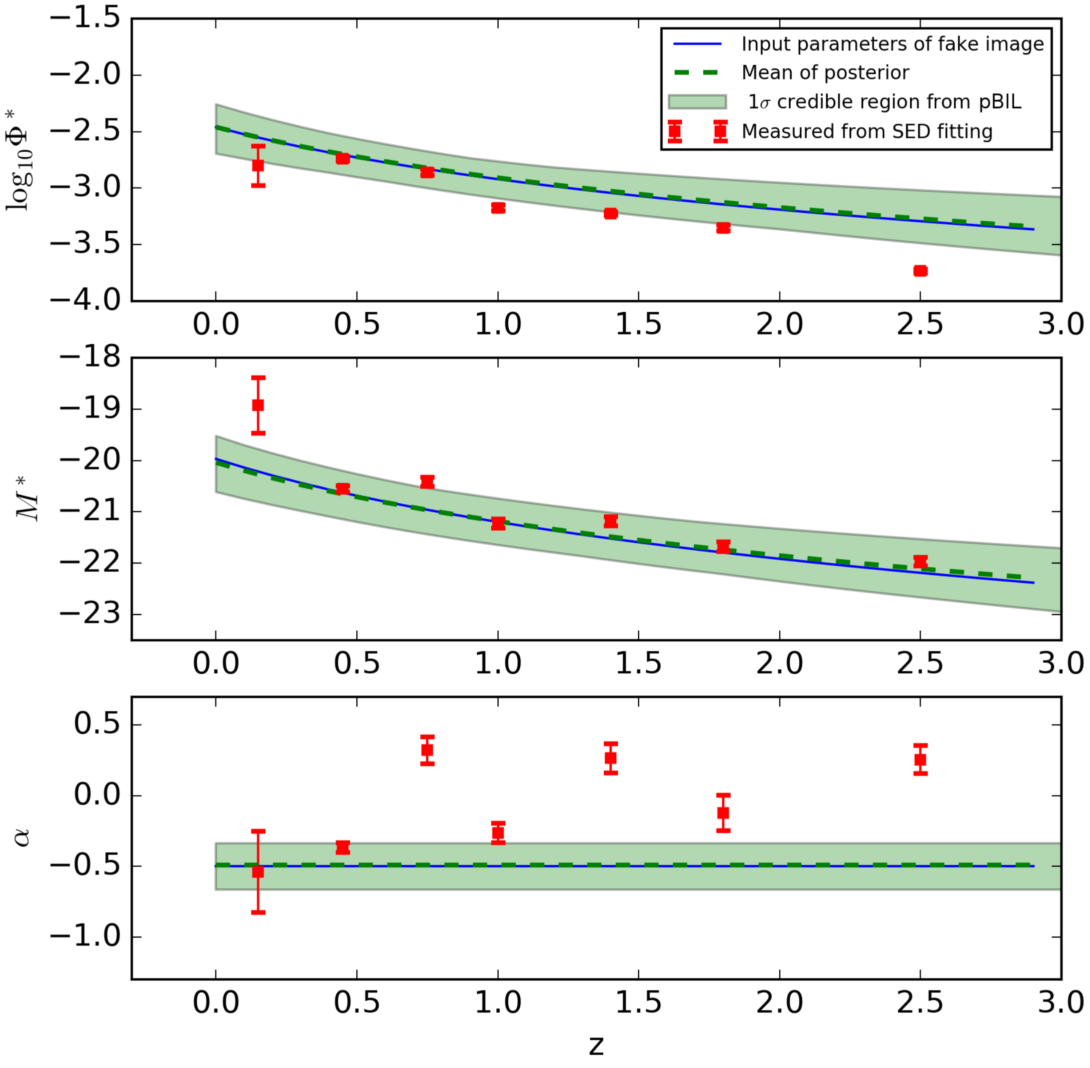


Figure 9-10 – Evolution of the LF parameters as defined in the mock data image (blue solid line) and inferred from the pBIL method in the "Fattening E" population described in Section 9.5.2 (the green dotted line is the mean of the posterior and the shaded area represents the 68% credible region), compared to the direct measurement of the LF obtained per redshift bin and estimated using a V_{max} weighting, after determination of the photometric redshifts from SED fitting (red dots).

Chapter 10

Improving the pipeline

> *We have a method, and that method helps us to reach not absolute truth, only asymptotic approaches to the truth - never there, just closer and closer, always finding vast new oceans of undiscovered possibilities. Cleverly designed experiments are the key.*
>
> Carl Sagan, *'Wonder and Skepticism', Skeptical Enquirer (Jan-Feb 1995)*

Abstract

Due to the high computational cost of running our pipeline over large periods of time, we need to minimize the time spent on each iteration and ensure that the Markov chains quickly reach their target distribution before considering applying it to real observations. We also need to simulate more realistic images by using the PSF measured on observed data, by reproducing the small and large-scales properties of background noise that arise in observed astronomical survey images, and by improving the dust extinction properties of our model. In this chapter, I explore some of the ways we can make the pipeline faster, more reliable and more efficient, and implement some of them.

Contents

10.1 Updating the parallelization strategy

Section 9.4.7 has taught us that the image simulation runtime using SKYMAKER scales linearly with image surface. We have also seen in Figure 9-4 that image simulation represents the biggest fraction of time spent in an full iteration of the pipeline (30-70% of the run time) relative to the other processes. Therefore, it is natural to think that splitting the input STUFF catalogs into smaller patches will ultimately lead to shorter run times. However, the unavoidable latencies introduced by HTCondor and the inherent characteristics of our cluster could make this argument invalid. Hence the need of empirically finding the optimal number of small patches that minimizes the run time spent in each iteration.

To that end, we decide to benchmark the process of simulated observables generation, corresponding to a full iteration of the MCMC run with our pipeline. By “full iteration” I refer to the combined processes of STUFF catalog generation, splitting of the STUFF catalog into N_{sub}^{y} vertical slices[1], dispatching of every job to the cluster, image simulation, source extraction, pre-processing of the observables and binning.

The simulated data consists in $\sim 60 \times 10^4$ sources from 3 galaxy populations (E/S0, Sp and Irr, whose characteristics are outlined in Tables 11.2 and 11.3) generated with STUFF, spread over a simulated image of size $9,538 \times 4,446$ pixels reproducing the characteristics of the CFHTLS Deep field in bands uiK_s, and from which $\sim 50 \times 10^4$ are extracted at the end of the iteration. Then I follow the evolution of run time with respect to number of patches by making N_{sub}^{y} evolve from 1 to 15, and tracking the corresponding wall-clock-time over a full iteration. For each given N_{sub}^{y}, I run 100 realizations of the data generation process with the same parameters, in order to account for the time variations resulting from HTCondor latencies. The resulting evolution of run time is plotted in Figure 10-1. We conclude that given the specificities of our cluster and our parallelization scheme, splitting the simulated STUFF catalog into 3 sub-catalogs (each containing sources within a different area) with the same surface minimizes the run time spent in a given iteration of the pipeline, contrary to what we expected.

10.2 Simulating more realistic images

10.2.1 Including noise correlation

In astronomical images, the background noise can usually be accurately modeled by a combination of Poisson noise on the photon number counts and Gaussian read-out noise (cf Chapter 5). But for images resulting from the co-addition of multiple overlapping individual exposures, the situation is different. In fact, due to pointing errors between exposures resulting

---

[1] The geometry of the slices don’t matter here, as the runtime is only influenced by the area of the simulated image.

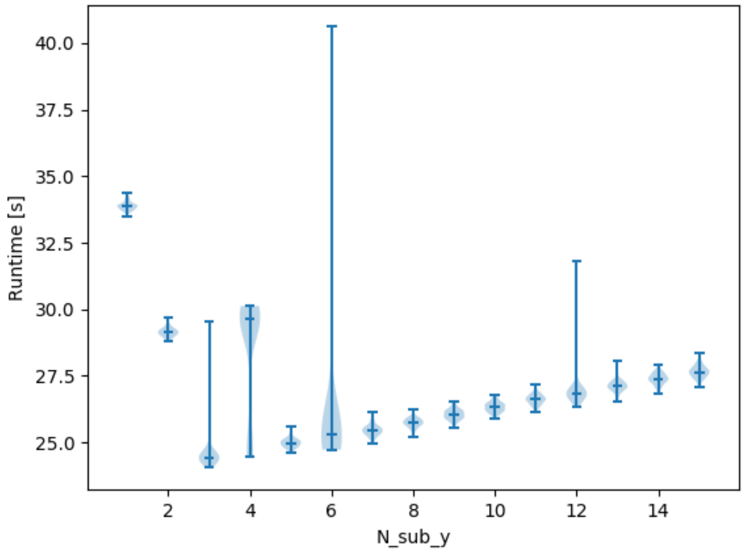


Figure 10-1 – Benchmark of a full iteration of the pipeline (cf Section 10.1) for various numbers of parallelized Stuff sub-catalogs N^y_{sub} (corresponding to equally-sized vertical slices), using ~60000 sources (spread into E/S0,Sp and Irr) generated by Stuff. For each N^y_{sub} I ran 100 realizations of the data generation process. Boundaries in the violin plot correspond to extremal values of run time. Setting N^y_{sub}=3 will yield minimal run times for the pipeline. An bump of runtime of unknown origin appears at N^y_{sub}=4, probably due to the use of computing resources from other jobs running in parallel at the time of the test.

from seeing fluctuations, the pixels from the different input images don't fit perfectly. During the co-addition process, the input images can be shifted, rotated, zoomed, and/or warped to fit the same output pixel grid, as illustrated in Figure 10-2. This requires a new mapping of the input pixel values to the output grid, an operation called *resampling*[2]. In practice, resampling involves an interpolation scheme (as the intensity of a given input pixel is divided into several adjacent output pixels), determined by an interpolation kernel which introduces a correlation between adjacent pixels. The choice of the interpolation kernel is important, as non-ideal kernels can blur the output image and result in aliasing artifacts (Aganj et al. 2013). The size of the kernel will also determine the degree of noise correlation in the output image. This correlation can be seen by eye in the form of clumpy noise patterns on the image (cf Figure 10-5), and can introduce biases in the estimation and uncertainties of galaxy flux and shape measurement, especially useful for weak lensing studies (Rowe et al. 2011; Gurvich and Mandelbaum 2016).

For the CFHTLS Deep T0007 release, all input images are located within a 3' radius with respect to the center of the field, in order to fill the gaps between CCD detectors (Hudelot et al. 2012). To produce the final stacks, individual exposures have been co-added using the SWARP software by Bertin et al. (2002) (with the COMBINE_TYPE "MEDIAN" option). Each image is therefore set on the same grid of 19354 × 19354 pixels of size 0.186", covering 1 deg × 1 deg on the sky in total. The input exposures have been interpolated using the Lanczos3 function $L_3(x, y) = L_3(x)L_3(y)$ (Duchon 1979) with size 6×6 pixels, where:

$$L_3(x) = \begin{cases} \text{sinc}(x)\text{sinc}(x/3) & \text{if } |x| < 3 \\ 0 & \text{otherwise} \end{cases} \tag{10.1}$$

This kernel (illustrated in Figure 10-3), which correlates the background noise up to 6 pixels (assuming no zoom factor is applied from input to output images), offers a good compromise between a good signal-to-noise ratio and limited aliasing artifacts on image discontinuities (Turkowski 2013; Bertin et al. 2002).

In order to generate accurate levels of noise correlation to our simulated images, a new (unreleased) version of SKYMAKER (v4.0), implementing this feature, has been developed by Emmanuel Bertin (private communication). The correlation process is performed analytically in real space by convolution of the input white noise and the interpolation kernel, assuming an infinite number of dithered input images with a uniform distribution in (x,y) position and no spread in position angle. Three new keywords have been added to that end in the SKYMAKER v4.0 configuration file:

- CORREL_TYPE: the pixel correlation can be applied to the noise only (CORREL_TYPE NOISE), or to both the noise and image data (CORREL_TYPE ALL), in which case the PSF may be enlarged by the correlation process. By default, image noise is not correlated (CORREL_TYPE NONE). In our case, we only correlate the background noise.
- CORREL_FUNC_TYPE: the type of correlation function used, selected between NEAREST, BILINEAR, LANCZOS2, LANCZOS3 and LANCZOS4. In our case, the LANCZOS3 kernel is used to reproduce the characteristics of the CFHTLS Deep field.
- CORREL_SCALE: the scaling of the correlation function. For instance, if the simulated image is supposed to be the result of resampling an image with pixels that are twice

---

[2] Sampling is the action of mapping a continuous function to a discrete one.

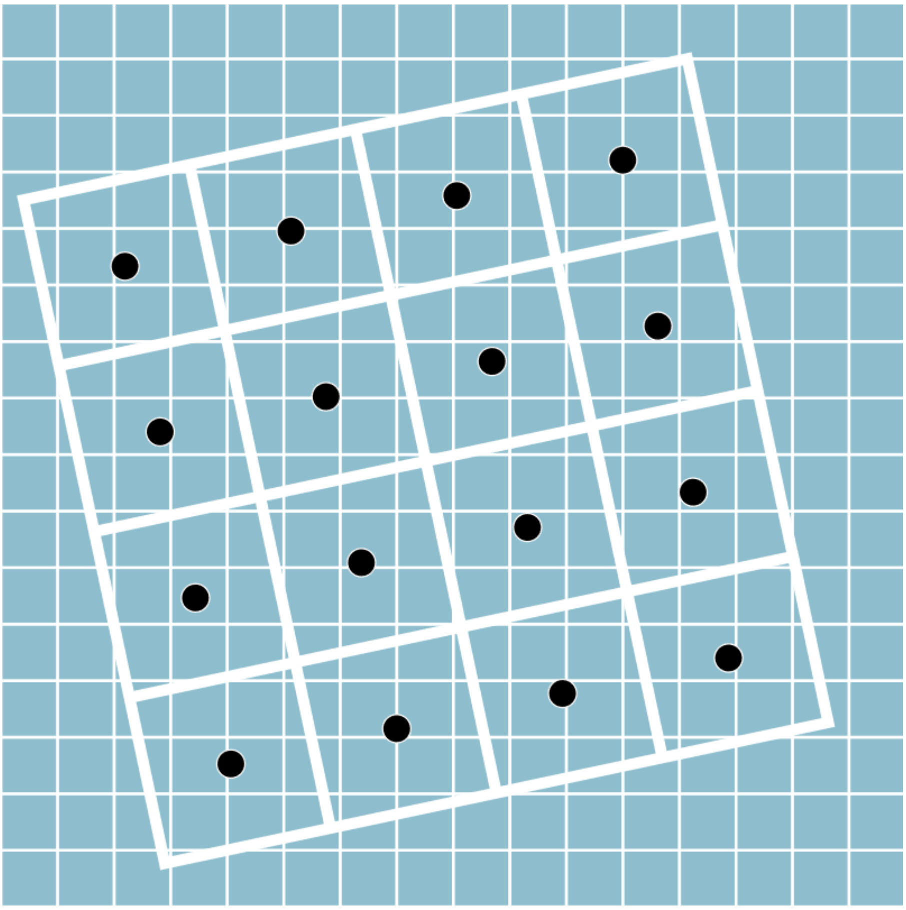


Figure 10-2 – Illustration of the resampling process, reproduced from SWARP user's guide (Bertin 2003). The input grid is shown as a collection of small gray squares, and the output grid is shown as large tilted ones. One interpolation of the input image, as represented by a dark spot, is done at the center of each output pixel.

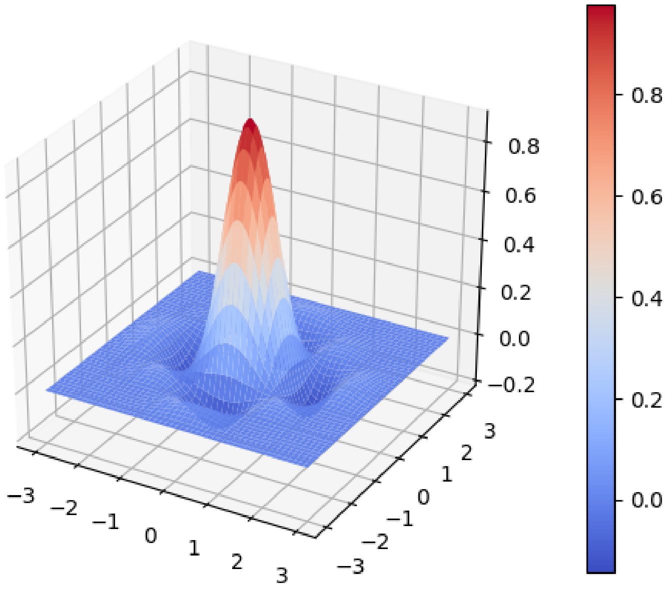


Figure 10-3 – The Lanczos3 2D interpolation kernel

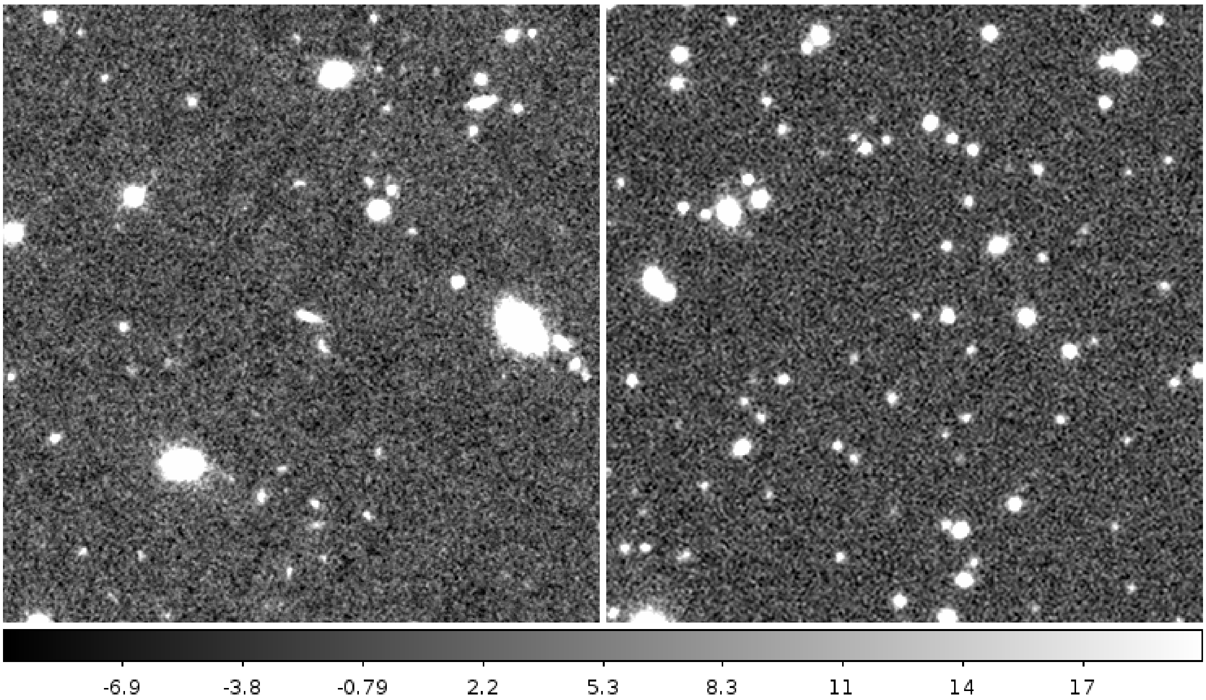


Figure 10-4 – Comparison between an observed image in K_s band from the WIRDS survey (left) and a simulated WIRDS image of the same size (370 × 370 pixels), generated by SkyMaker with the correlated noise generation feature (right). Both images display the same intensity scaling. The same level of granularity can be seen in the background noise of both images.

larger, CORREL_SCALE should be set to 2.0. In CFHTLS stacks, no zoom is applied on individual exposures, so I set CORREL_SCALE to 1.0 in *ugriz*. However, in order to fit WIRDS images obtained with the WIRCAM instrument, whose pixel size is 0.3” (Bielby et al. 2012), to the same pixel grid as CFHTLS Deep images, whose pixel size is 0.186”, CORREL_SCALE should be set to 0.3”/0.186”=1.613 in JHK_s.

As illustrated in Figure 10-4, the noise correlation introduced on simulated images is visually almost indistinguishable from the noise of real data, when attested by the level of noise granularity. We also note that this new feature has a negligible cost on the image simulation process in terms of computing time and memory usage.

10.2.2 Including weight maps

Over a co-added science image, not all pixels are born equal: some are noisier than others. This can be due to inherent discrepancies of quantum efficiency in the pixel matrix of the CCD sensor, gaps between detectors and vignetting caused by unequal effective exposure time, or other external factors such as satellite tracks, cosmic rays, or saturation trails (Erben et al. 2005). The resulting structure of the variable noise level over the co-added image can be complex, especially when using multidetector (mosaic) cameras creating composite images, such as MegaCam. Hence, a co-added survey image is usually provided with a *weight map* for each band, i.e. a FITS image with the same dimensions as the survey image, computed from a normalized flat field and displaying the noise level of each pixel, usually expressed in units of

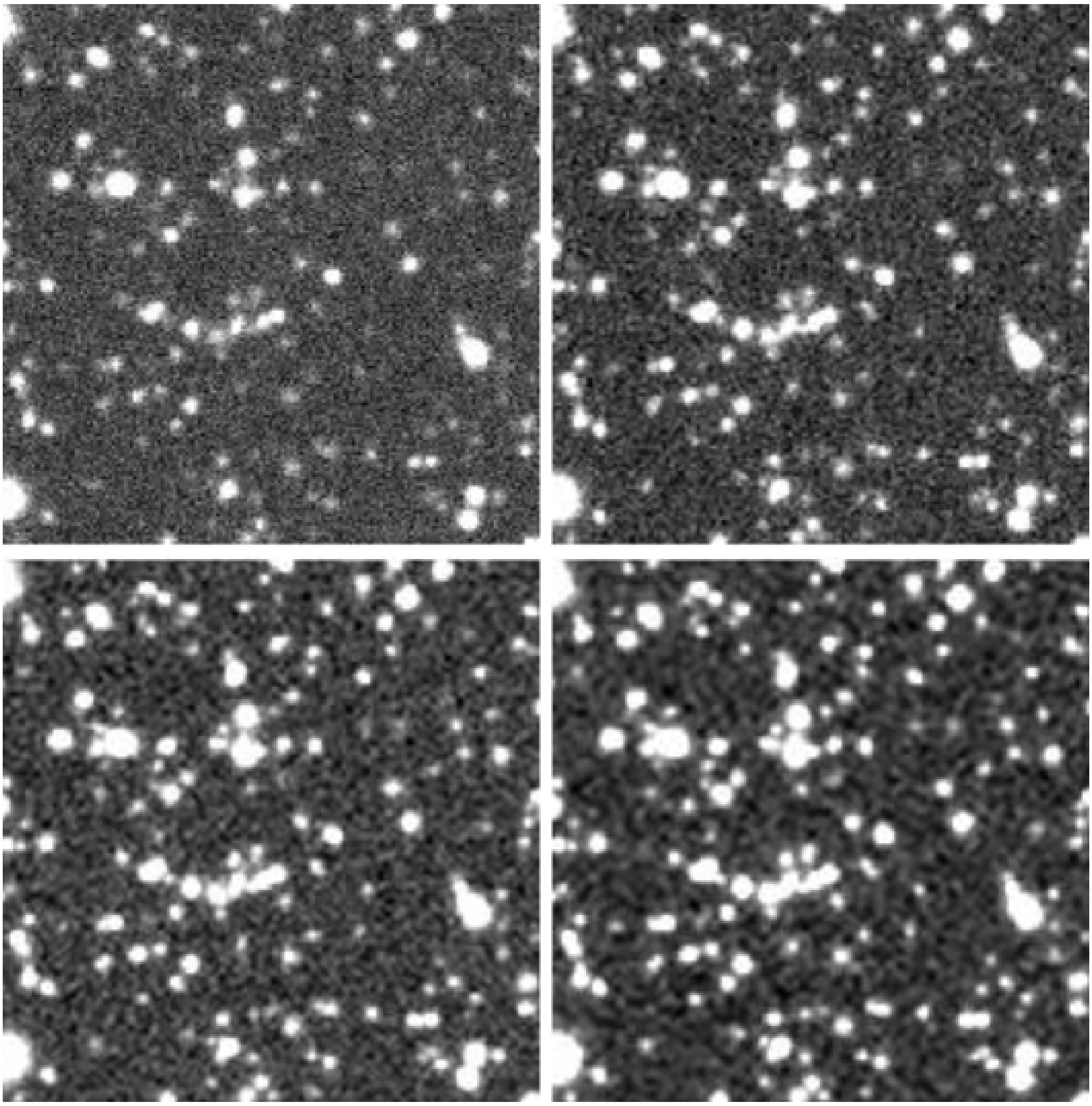


Figure 10-5 – Simulated image in g band of size 300×300 pixels generated using Skymaker v4.0. The noise is correlated using a Lanczos3 kernel, and each image presents the same field with a different scaling of the correlation function. CORREL_SCALE=1,2,3,4 from left to right and from top to bottom respectively. The bigger the scaling, the bigger the size of the background noise clumps.

relative inverse noise variance[3]:

$$\omega_j \propto \frac{1}{\sigma_j^2} \tag{10.2}$$

In such maps, regions of low weight values correspond to regions of high noise level. Bad pixels and other unreliable regions for scientific analysis are set to 0, which corresponds to infinite variance of the noise level. Weight maps for co-added images are usually created using the weight maps from all the input images before co-addition, generated with the WeightWatcher tool of Bertin (1997). All the input weight images are then co-added by SWarp (Bertin et al. 2002). Some examples of the resulting maps for the CFHTLS D1 field are displayed in Figure 10-6.

SExtractor uses weight maps when detecting sources and estimating measurement (fluxes, magnitudes...) uncertainties (Bertin 2006). In fact, bad pixels are discarded from the background estimation, and the detection threshold is defined relative to the estimation of the RMS background noise. Not including weight maps might therefore result in a lot of spurious detections in noisy regions, as illustrated in Figure 10-7. Thus, they have to be included in the source extraction step of our pipeline.

We can also use weight maps to reproduce the large scale properties of the noise level from real data in our simulated images. The version 4.0 of Skymaker (Bertin, private communication) takes weight maps as input files, and returns images with the appropriate variable noise level, provided a small calibration step from the user. The result (using an custom weight map) is illustrated in Figure 10-8. This updated version of Skymaker introduces a new set of parameters in its configuration file:

- WEIGHT_TYPE, which is type of weight map used (weight, variance or RMS map)

- WEIGHT_IMAGE, which is the filename of the input weight image

- WEIGHT_GAIN, which allows (or not) the modulation of the gain (in e^-/ADU) with weights. It is set to Y by default, assuming the changes in noise levels over the images are due to local gain changes, which is the most common case (Bertin 2006)

- WEIGHT_FACTOR, a noise scaling factor which corresponds to the background RMS noise for a weight of 1

The WEIGHT_FACTOR value has to be set in each band so that the background RMS noise level output by SExtractor reproduces the one in observed data. To that end, I first set the WEIGHT_FACTOR to 1 in every filter to compare the resulting background RMS noise levels on a simulated and observed CFHTLS D1 field with the same area. The results are outlined below.

---

[3] This assumes that input images are background noise limited

Filter	WEIGHT_FACTOR	Background RMS (sim)	Background RMS (data)
u	1	0.225091	0.188747
g	1	0.0781753	0.121543
r	1	0.132736	0.165939
i	1	0.220724	0.222246
z	1	0.659721	0.568447
J	1	1.43995	2.35003
H	1	3.03379	3.3733
K_s	1	3.15413	3.4625

By dividing the background RMS computed on the data by the background RMS computed on the simulation, I obtain the appropriate WEIGHT_FACTOR for each band:

Filter	WEIGHT_FACTOR	Background RMS (sim)	Background RMS (data)
u	0.84	0.189575	0.188747
g	1.55	0.120099	0.121543
r	1.25	0.164979	0.165939
i	1	0.220724	0.222246
z	0.86	0.569466	0.568447
J	1.65	2.35296	2.35003
H	1.11	3.36918	3.3733
K_s	1.10	3.45914	3.4625

In the parallelization step of our pipeline described in Section 9.4.7, at each iteration of the MCMC run, the weight images are split into sub-images and each cropped version is sent to the various nodes of the cluster for image simulation and source extraction purposes.

10.2.3 Using the measured PSF on image simulations

Previous tests of the pipeline relied on the internally generated PSF of SKYMAKER with no simulated aureole. The choice of the PSF didn't matter then, because the simulated images we used as "observed" data were also generated using the SKYMAKER internal PSF, so our tests were self-consistent. But if the pipeline were to be applied to real data, our simulated images of the CFHTLS D1 field would have to be convolved with the appropriate PSF in order to accurately reproduce the photometric properties of the survey.

The PSF is measured by Andrea Moneti (private communication) in *ugriz* using catalogs extracted from the T0007bis release of the CFHTLS D1 stacks (private communication), and in JHK_s using those of the T0002 release of the WIRDS stacks. Using the PSFEx software of Bertin (2011a), the PSF is modeled as a two-dimensional Moffat profile fitted to a selection of unsaturated and bright stellar sources for each passband via χ^2 minimization. The PSF is assumed to be constant over the field, an approximation justified by the small size of the survey we are about to simulate (cf Section 11.1). In practice, a 0-degree polynomial on pixel coordinates is used to map the PSF variations.

The resulting PSF consist in a series of data cubes in the .PSF format (one by filter) used by SEXTRACTOR, and a series of check images generated by PSFEX in the form of small patches (called "vignettes") in the FITS format, of size 35 × 35 pixels so that they captures all the features of the non-saturated stars. These vignettes, which are illustrated in bands uiK_s in

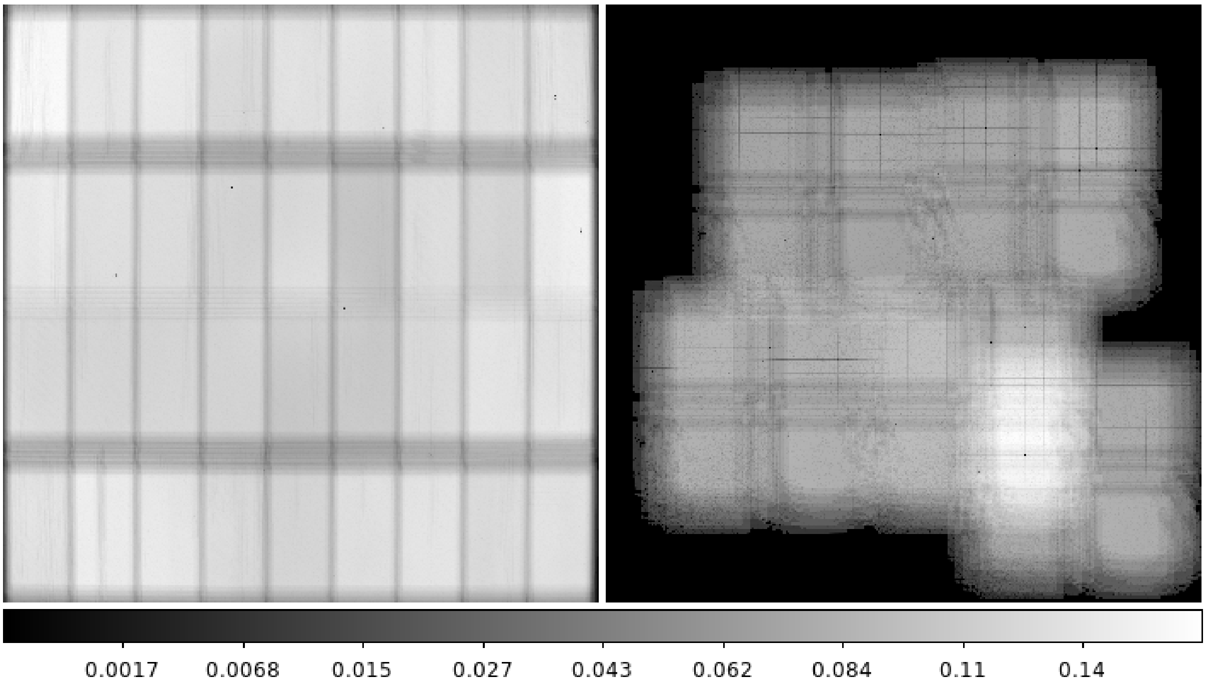


Figure 10-6 – Weight maps of the CFHTLS D1 field in bands u (left) and K_s (right). The same intensity scaling is applied to both maps. Dark regions correspond to areas of high noise level, whereas bright regions indicate low noise levels. Effects of vignetting are apparent in the borders of the K_s map.

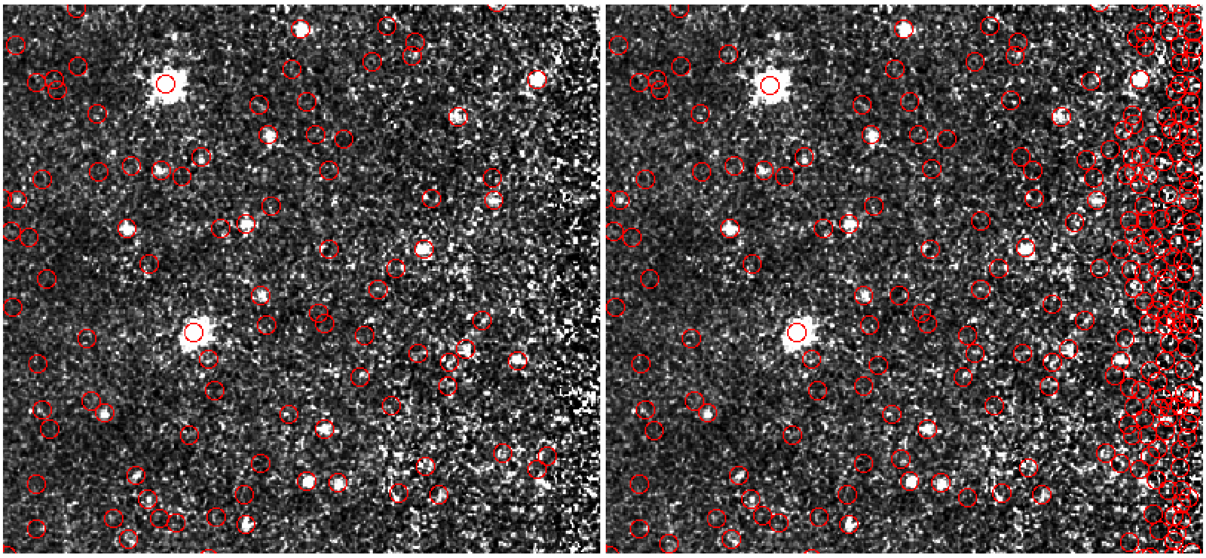


Figure 10-7 – Effect of weight map on source detection using SExtractor (each detection is denoted as a red circle) in a noisy region of the WIRDS field in K_s band. *Left:* the detection is performed without a weight map. *Right:* the detection is performed with a weight map. Without weight maps, a lot of noise peaks are detected as sources.

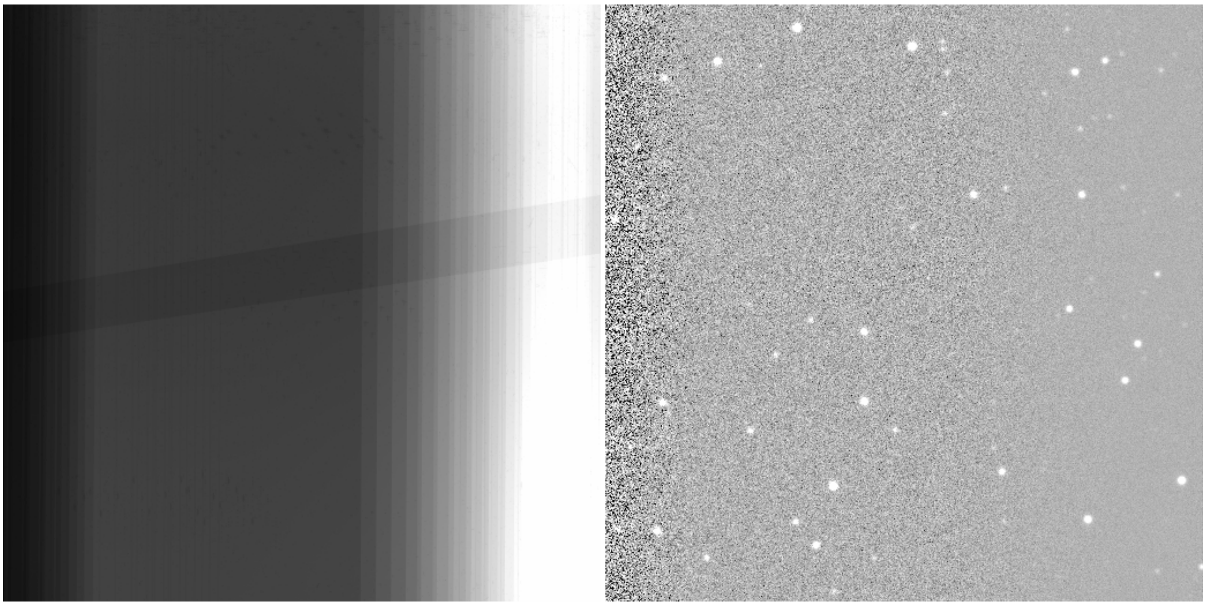


Figure 10-8 – *Left:* a custom weight map used for testing purposes. *Right:* a simulated image generated by SKYMAKER v4.0 using the custom weight map. The large scale structure of the noise defined in the weight image is well reproduced.

Table 10.1 – Characteristics of the PSF modeled on observed data, computed by Andrea Moneti using PSFEx

Survey	Release	Passband	Nb of stars loaded	Nb of stars accepted	FWHM [pixels]	χ^2
CFHTLS D1	T0007bis	u	1138	1009	5.50	3.39
CFHTLS D1	T0007bis	g	1289	1103	5.20	4.08
CFHTLS D1	T0007bis	r	1784	1520	4.84	3.96
CFHTLS D1	T0007bis	i	1426	1196	4.51	4.90
CFHTLS D1	T0007bis	z	1672	1477	4.69	2.24
WIRDS	T0002	J	257	223	4.35	2.52
WIRDS	T0002	H	246	209	3.69	2.69
WIRDS	T0002	K_s	181	164	3.87	4.59

Figure 10-9, will be used by SKYMAKER for image simulation purpose. The characteristics of the measured PSF in every band are detailed in Table 10.1.

10.3 Implementing a more realistic dust treatment

10.3.1 Adding Galactic extinction to STUFF magnitudes

By default, STUFF doesn't handle Galactic extinction internally, so in order to simulate realistic extinction properties (cf Section 10.3.1), we have to manually alter the apparent magnitudes from STUFF catalogs. To that end, we need the value of the total extinction at the coordinates of the survey field that we are reproducing on the celestial sphere.

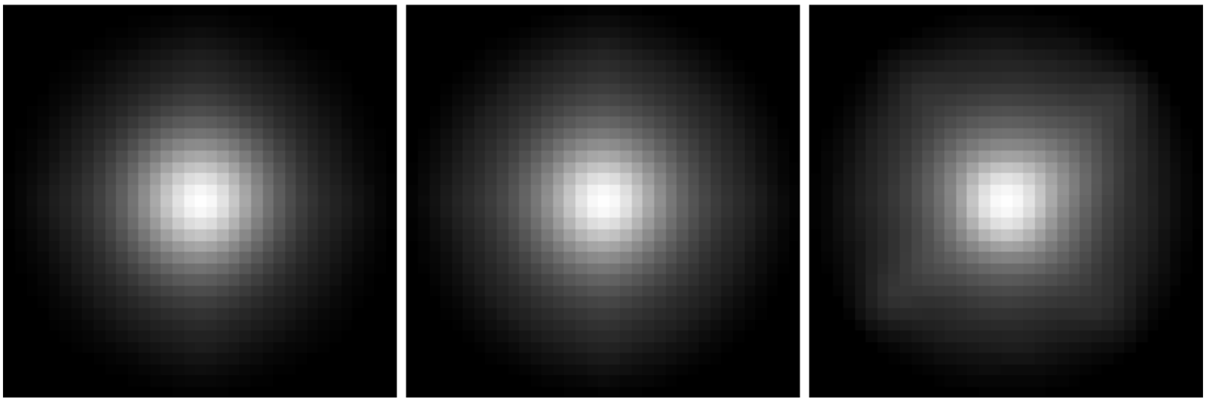


Figure 10-9 – Point spread function measured on CFHTLS D1+WIRDS images in bands uiK_s (from left to right respectively) using PSFEx. The PSF, modeled on a 35×35 vignette, will be considered constant over the field in our future studies. All the images feature the same intensity scaling.

Using the online tool provided by NASA/IPAC Infrared Science Archive[4], we generate a cutout of the Galactic dust reddening map of Schlegel et al. (1998), compiled from IRAS and COBE data that we described in Section 10.3.1. This resulting cutout is a 2 deg^2 square image (the smallest area possible to extract) centered on the CFHTLS Deep field, shown on Figure 10-10.

We consider that the resulting E(B-V) reddening values have negligible spatial variations over the field ($\text{E(B-V)}_{\text{max}} - \text{E(B-V)}_{\text{min}} = 0.01$ mag, which is at least 10× lower (depending on the filter) than the median RMS error for MAG_AUTO estimation with SExtractor). We therefore opt for the extinction value computed within a 5 arcmin radius circle at the center of the D1 field. This area is represented as a green circle on Figure 10-10. The service provides the total extinction values of Schlegel et al. (1998), as well as the more recent values estimated by Schlafly and Finkbeiner (2011) from the analysis of the Sloan Digital Sky Survey. Both results assume an extinction to reddening ratio $A_V/E(B-V) = 3.1$, but Schlafly and Finkbeiner (2011) use a different calibration and extinction law. We use the most recent results for our needs.

Table 10.2 – Total extinction per band at the center of the CFHTLS D1 field

band	u	g	r	i	z	J	H	K_s
A_λ	0.113	0.088	0.061	0.045	0.034	0.019	0.012	0.008

Note: the values of extinction in JHK_s are those in the UKIRT magnitude system.
The extinction values are provided using the extinction law of Schlafly and Finkbeiner (2011) (assuming $R_V = 3.1$).

Using these data, we finally derive the Stuff apparent magnitudes reddened by Galactic extinction m_r via:

$$m_r = m_{\text{Stuff}} + A_\lambda \tag{10.3}$$

---

[4] `http://irsa.ipac.caltech.edu/applications/DUST/`

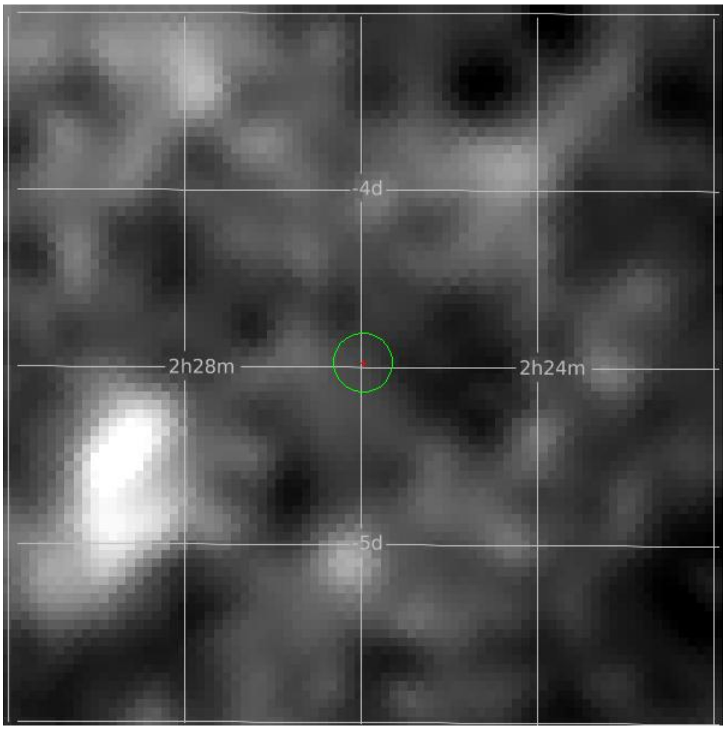


Figure 10-10 – Cutout of the Galactic dust E(B-V) reddening map of Schlegel et al. (1998) from IRAS + COBE 100 μm imaging data centered on the CFHTLS Deep field we want to simulate and covering 2 deg^2 on the sky. The field we are simulating represents the inner square on this image, and is centered on (α, δ)=(36.4926489,-4.486492445) at epoch J2000. The E(B-V) reddening values of Schlafly and Finkbeiner (2011) are 14% smaller than the values of Schlegel et al. (1998).

10.3.2 Updating the attenuation curves of Stuff

In Stuff, the only available attenuation law is a Calzetti (1994,2000) curve extrapolated below 1200 Å by a LMC curve. But as we have seen in Section 4.5.3, the Calzetti law might not be representative of the majority of extragalactic sources. For more flexibility for future analyses, we decide to add other useful attenuation curves to the Stuff package, based on recent analyses on local galaxies such as the Small and Large Magellanic clouds, as well as the Milky Way. Moreover, an appropriate parametrization of the attenuation laws has to be used for an easy conversion into the Stuff format with the correct wavelength range.

As proposed by Fitzpatrick and Massa (1986), the shape of the 2175 Å feature can be well described by the Lorentzian-like Drude profile:

$$D(x,\gamma,x_0) = \frac{x^2}{(x^2 - x_0^2)^2 + \gamma^2 x^2} \tag{10.4}$$

Where $x = \lambda^{-1}$, $x_0 = \lambda_0^{-1}$ is the position of the feature and γ its FWHM. According to Boselli (2011), the LMC and SMC attenuation curves are usually fitted by:

$$k(\lambda) = R_V + c_1 + c_2 x + c_3 D(x,\gamma,x_0) + c_4 F(x) \tag{10.5}$$

with

$$F(x) = \begin{cases} 0.5392(x-5.9)^2 - 0.05644(x-5.9)^3 & \text{for } x \geq 5.9\mu\text{m}^{-1} \\ 0 & \text{for } x < 5.9\mu\text{m}^{-1} \end{cases} \tag{10.6}$$

and $x = \lambda^{-1}$ in μm. As for the MW extinction curve, it is fitted by:

$$k_{\mathrm{MW}}(\lambda) = \begin{cases} R_V + c_1 + c_2 x + c_3 D(x,\gamma,x_0) & \text{for } x \leq c_5 \\ R_V + c_1 + c_2 x + c_3 D(x,\gamma,x_0) + c_4 (x - c_5)^2 & \text{for } x > c_5 \end{cases} \tag{10.7}$$

The values of the best fit coefficients, from the results of Gordon et al. (2003) for LMC and SMC, and Fitzpatrick and Massa (1986) for MW, are provided in Table 10.3.

Table 10.3 – Best fit coefficients for the attenuation curves of local galaxies (reproduced from Boselli (2011))

Galaxy	R_V	c_1	c_2	c_3	c_4	c_5	x_0 [μm^{-1}]	γ [μm^{-1}]
MW	3.00	-0.175	0.807	2.991	0.319	6.097	4.592	0.922
LMC	3.41	-0.890	0.998	2.719	0.40	-	4.579	0.934
LMC2	2.76	-1.475	1.132	1.463	0.294	-	4.558	0.945
SMC (bar)	2.74	-4.959	2.264	0.389	0.461	-	4.600	1.000
SMC (wing)	2.05	-0.856	1.038	3.215	0.107	-	4.703	1.212

Attenuation laws in Stuff take the form of $\frac{2.5}{\ln(10)} \frac{A_\lambda}{A_{4000}}$ as a function of λ, a quantity normalized so that $A(4000\text{Å}) = 1.086$. Each law consists in a list of 240 points with a wavelength range of $100\ \text{Å} \leq \lambda \leq 10^6\ \text{Å}$ [5], and with $A_\lambda = 0$ for $\lambda = 100$ and 10^6Å. The total extinctions are then linearly interpolated between data points for catalog generation.

---

[5] Too many points would result in a significant increase of computation time.

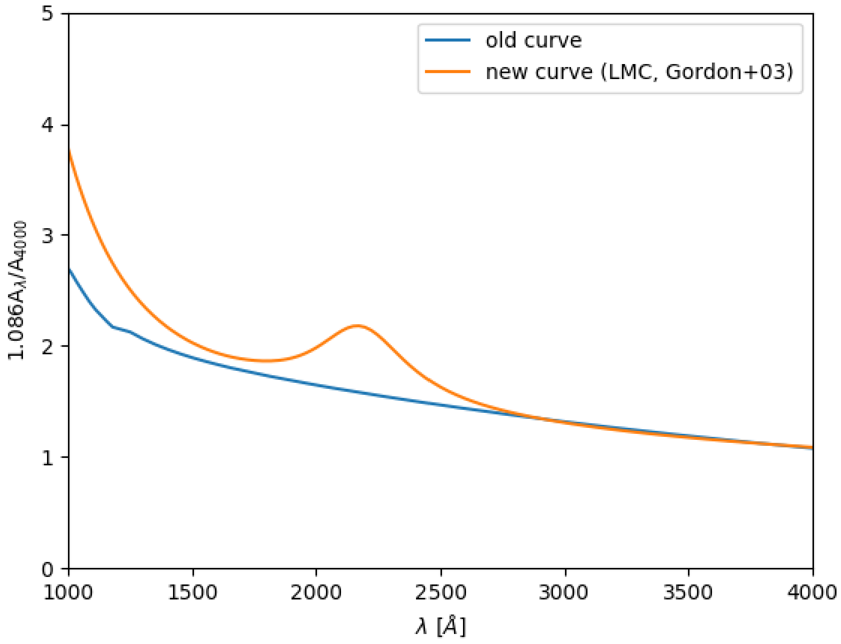


Figure 10-11 – Comparison of Stuff default dust attenuation curves before and after the update. The blue curve is a Calzetti (1994, 2000) law extrapolated below 1200 Å by a LMC curve. The orange curve is the LMC law provided by the best fit of Gordon et al. (2003).

Stuff now uses LMC as its default attenuation curve (cf Figure 10-11). With the new default settings, it generates about 30% less galaxies than before ($\sim 15,000$ instead of $\sim 21,400$).

10.4 Masking the bright stars

Bright stars contaminate the estimated flux of nearby sources, and saturate all the individual exposures, which prevent their removal via co-addition (Coupon et al. 2017). As Stuff doesn't feature (yet) the implementation of a realistic stellar field to simulated images, we decide to mask out all regions around the brightest stars, and only use non-contaminated regions for photometric and morphometric measurements.

The mask files are produced semi-automatically by the IAP TERAPIX team[6], who used the USNO source catalog (Monet et al. 2003) to locate bright stars on the CFHTLS field and drew polygons whose size depends on the stars' magnitudes around them. The masks are provided as FITS files (one for the optical bands, and one for the IR bands), where non-masked regions are assigned a values of 0, and masked regions a value of 1. The shapes of the resulting masked areas are illustrated in Figure 10-12 on a fraction of the CFHTLS D1 field. Mask files (or *flag images*) are recognized by SExtractor via the FLAG_IMAGE keyword in its configuration file, provided that IMAFLAGS_ISO is present in the parameters file. In this way, sources detected in masked regions are assigned an IMAFLAGS_ISO of 1 in the output catalog (0 otherwise). They can therefore be easily ignored in further analyses.

---

[6] `http://terapix.iap.fr/`

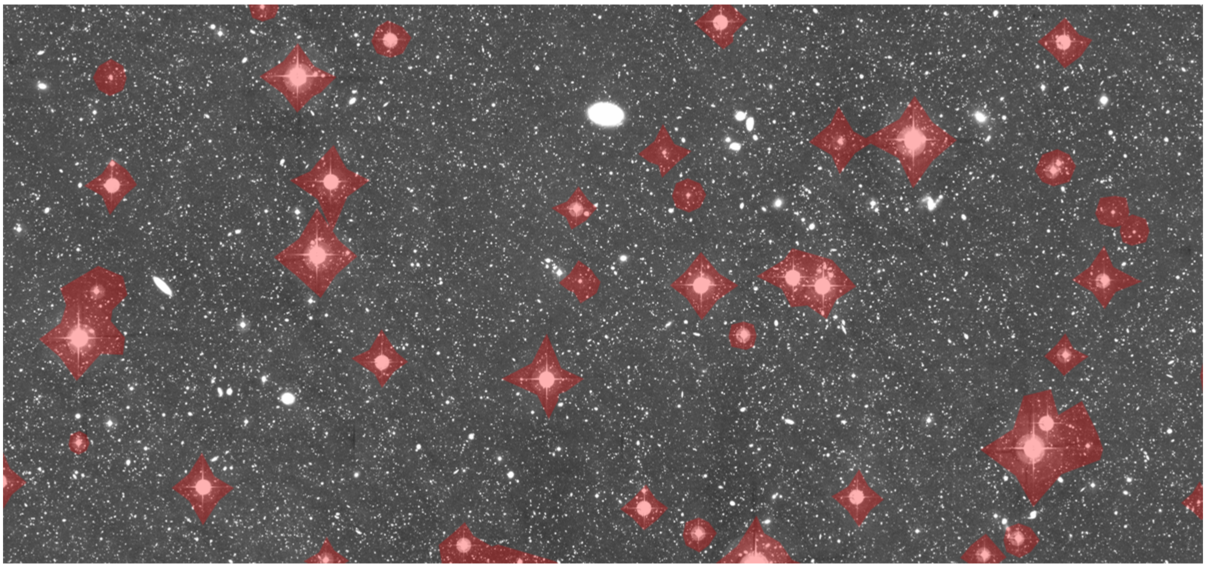


Figure 10-12 – “Proposal 3” region of the CFHTLS D1 field in u band. The red polygons represent the masked regions of the field around bright stars discarded from scientific analyses.

In the parallelization step of our pipeline described in Section 9.4.7, at each iteration of the MCMC run, the mask images are split into sub-images and each cropped version is sent to the various nodes of the cluster for source extraction. I assigned the same mask files for the simulation and observation, so that the same effective area is used for analysis.

10.5 Improving the sampling scheme

Up to now, all the tests of the pipeline we have carried in Chapter 9 have been using the Adaptive Proposal sampling algorithm of Haario et al. (1999). But, as we have seen in Section 6.4.1, this algorithm is prone to generate mildly biased estimates of the target distribution, as it does not possess the right ergodic properties. Our tests using known input parameters have shown that this bias is small enough to ensure reliable inferences on the size and luminosity distributions of simulated galaxies in a simple, controlled environment. But we have to eliminate any known potential sources of biases completely if we want to test our pipeline on the “real” universe. We therefore update our sampling algorithm to the more reliable Adaptive Metropolis algorithm of Haario et al. (2001) described in Section 6.4.2.

It is difficult to predict how many iterations will be needed to achieve convergence of our pipeline. Therefore, in order to make better use of the full potential of the adaptive nature of the sampling algorithm and potentially increase the efficiency of our MCMC runs, we set the update frequency of the proposal kernel covariance matrix to 100 iterations instead of 500 as was done previously.

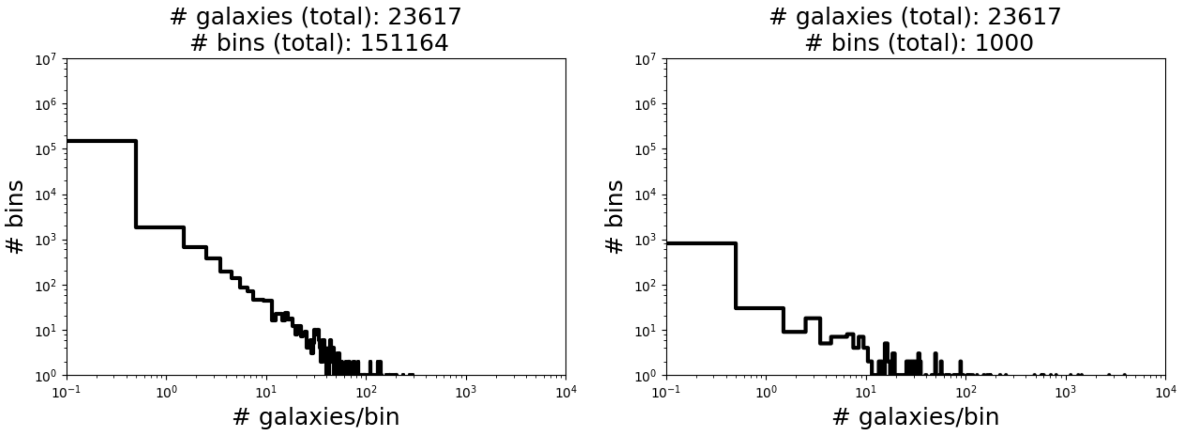
10.6 Determining the appropriate binning rule

The previous tests we have carried on (cf Section 9.5) relied on an *ad hoc* binning rule partitioning the observable space into 10 bins per (pre-processed) observable used (MAG_AUTO and/or FLUX_RADIUS in several photometric bands). We have seen in Chapter 7 that less *ad hoc* binning rules exist to split multidimensional datasets, such as Knuth's rule or Bayesian blocks. In this section, I explore the consequences of the binning choice on the quality of inference, by comparing the use of Knuth's rule to the use of our previous binning rule during a realistic run of our pipeline.

The tests are performed in the "Multi-type" configuration (cf Section 9.5.1), using a simulated image of size 9358×4446 in bands uiK_s with a combined distribution of E/S0 and Sp galaxies, reproducing a fraction of the area of the D1 tile of the CFHTLS Deep field, as our "observed" data. From this image, 23,617 sources with a FLAG < 4 (i.e. unsaturated and uncropped) are extracted in total. For each binning rule, 3 chains are launched from randomly selected starting points for 10^4 iterations, using the AM + simulated annealing algorithm to sample the parameter space efficiently.

Figure 10-13 plots the number of bins partitioning the observable space as a function of the number of galaxies per bin, for each binning rule, applied to the "observed" data. The plot reveals that Knuth's rule introduces $\sim 150\times$ more bins in total than our previous rule, and $\sim 100\times$ more empty bins. When computing the auxiliary likelihood, each empty bin will be artificially filled with 1 source to prevent singularities in log space. Therefore, if Knuth's rule is used as our binning rule, we can expect that a lot of unphysical noise will be added to our distribution of observables, that will potentially affect the accuracy of the inference process via a more random evolution of the acceptance probability (defined in Equation 6.14) across iterations. This hypothesis is confirmed by the erratic behavior of the temperature evolution (which constitutes a good tracer of the noise level of our model) over 10^4 iterations when using Knuth's rule, as plotted in Figure 10-14, in contrast to the relatively stable behavior of the temperature when the previous rule is used.

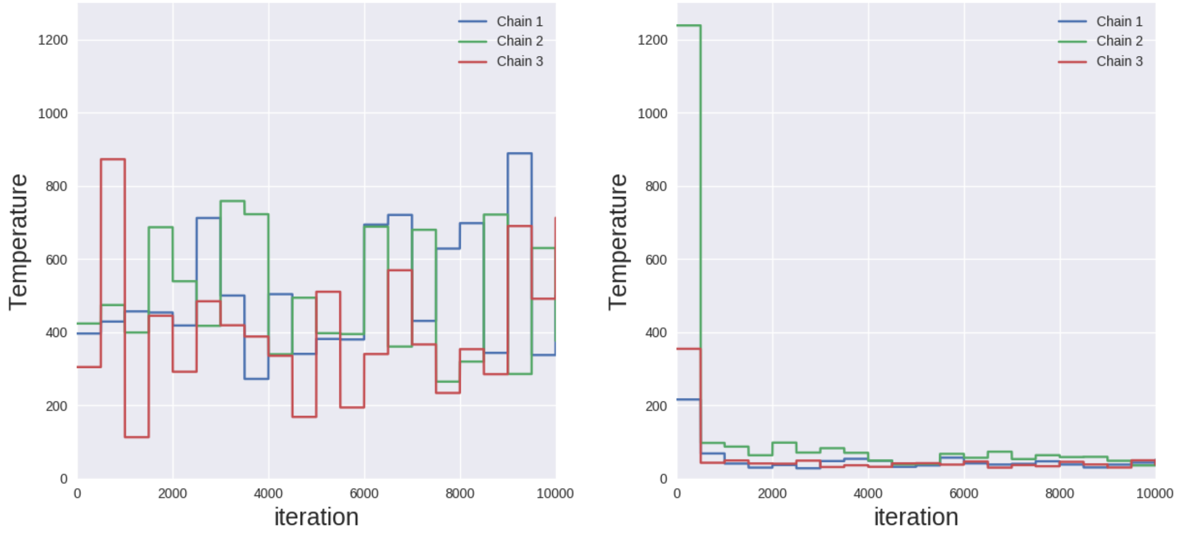
We therefore decide to keep the "10 bins per dimension" binning rule when applying our pipeline to real data, considering its proven efficiency on previous tests.



(a) Knuth's rule

(b) 10 bins per dimension

Figure 10-13 – Comparison of the distribution of number of bins as a function of number of galaxies per bin for different binning rules, using the same dataset consisting of the pre-processed MAG_AUTO of 23617 SExtracted galaxies in uiK_s. We conclude that Knuth's rule generates a lot more empty bins than the "10 bins per dimension" rule, which can ultimately deteriorate the precision of our future inferences if used in our pipeline.



(a) Knuth's rule

(b) 10 bins per dimension

Figure 10-14 – Comparison of the simulated annealing temperature evolution under different binning rules, for the purpose of the test described in Section 10.6. The same simulated image in uiK_s is fed to the pipeline as input data in both configurations, and the same observables are used (i.e. the MAG_AUTO in uiK_s). Using Knuth's rule for partitioning the observable space has the consequence of radically changing the behavior of the temperature evolution, i.e. of the noise level of the model, to a more unstable one (in terms of absolute level and variation across iterations) with respect to the "10 bins per dimension" rule.

Chapter 11

Towards applying the pipeline to real data

> *"THERE IS AS YET INSUFFICIENT DATA FOR A MEANINGFUL ANSWER."*
>
> Isaac Asimov, *The Last Question*

Abstract

Now that our updated pipeline provides more realistic image simulations and has been optimized to limit its runtime, it is finally ready to be run on real images from the CFHTLS Deep. In this chapter, I first test the reliability of our improved pipeline by inferring the luminosity evolution of a set of 3 realistic galaxy populations (E/S0, Sp and Irr) with known properties spread on a simulated image reproducing ∼11% of the area of a CFHTLS Deep field. I then use the same settings to infer the luminosity function (and its evolution with redshift) of these 3 populations using real CFHTLS D1 data. To this day, the Markov chains are still traveling...

Contents

11.1 Choosing a field to analyze

In the proof-of-concept runs with simulated images containing 1 or 2 populations of galaxies that we designed to test the pipeline in Chapter 9, we were able to limit computing time via lowering by 1 order of magnitude the number of sources[1] generated over the area of the

---

[1] Reminder: in STUFF, the number density of sources per galaxy population is determined by the ϕ^* parameter of the luminosity function.

CFHTLS D1 simulated field. As we have seen in benchmarking tests (cf Section 9.4.7) and with the large number of iterations required by the MCMC process (Section 9.5), if we want to test our pipeline on real data within a reasonable amount of user time (i.e. a few weeks), we can only afford to simulate a small portion (∼10 %) of the area of the CFHTLS D1 field. The given region of the D1 field will be chosen according to the following criteria:

- It has to maximize the area of the survey fit for analysis by minimizing the number of very bright stars contaminating the field (i.e. large masked regions),
- The weight maps of this region must avoid large spatial variations, which are revealing of noisy regions unsuitable for a reliable analysis.

The chosen region that meets these two criteria, dubbed "Proposal 3" and found by exploration of the D1 field using the `DS9` visualization tool (Joye 2013), is a 9,538 × 4,446 pixels region centered on $(\alpha, \delta) = (2:25:56, -4:23:07)$ (in the FK5 system[2]) and covering 11.3 % of the surface of the D1 field[3]. The position and area of the region on the D1 field is illustrated in Figure 11-1, and Figure 11-2 displays a color image of the "Proposal 3" field in the *gri* bands. Finally, the weight maps corresponding to "Proposal 3" are shown in Figure 11-3.

The last public release of the CFHTLS D1 T0007 stack suffers from an overestimation of the sky background level around bright objects and regions of high number densities due to insufficient masking of the sources, which makes the sky-subtracted stacks biased in these regions. Lately, the IAP TERAPIX team has performed a re-stacking of the Deep fields, where the background level is estimated including bright object masking with mask files originally designed for the staking of WIRDS data. In this "T0007bis" new stack, the sky subtraction is therefore improved, and we use it for this work.

11.2 Test with 3 populations

Before running our pipeline on CFHTLS Deep data, we need to perform a last test of self-consistency of our updated version of the pipeline in the most realistic way possible considering our limited resources in terms of computing time. In this test, we use a simulated image of a fraction of the CFHTLS D1 field as input data. This image is a mock "proposal 3" field of size $9,538 \times 4,446$ pixels (each pixel representing 0.186" on the sky) reproduced in bands uiK_s containing 3 populations of galaxies generated by Stuff, namely E/S0, Sp and Irr.

Contrary to previous tests, we want to reproduce the observed number counts by introducing a realistic population of Irr, which can be done in this situation as we have reduced the simulated field area. The E/S0 and Sp populations are generated using the LF parameters of Faber et al. (2007), this time without scaling down the number densities. The LF parameters (and evolution parameters) of the Irr population from EFIGI data (Baillard et al. 2011; de Lapparent, private communication) were tuned to closely match visually the observed slope of the uiK_s number counts in the "proposal 3" field from the CFHTLS D1 T0007bis + WIRDS T0002 stacks (described in Section 9.2). The contribution of each population was tracked by cross-identification of input sources and extracted sources using their position on the image. The best visual combination of luminosity functions and their evolution per type is shown in 11-4. This

---

[2] $(x_{\text{center}}, y_{\text{center}})$=(9843,11785) in pixel coordinates.

[3] This corresponds to a surface of 42405948 pixels^2

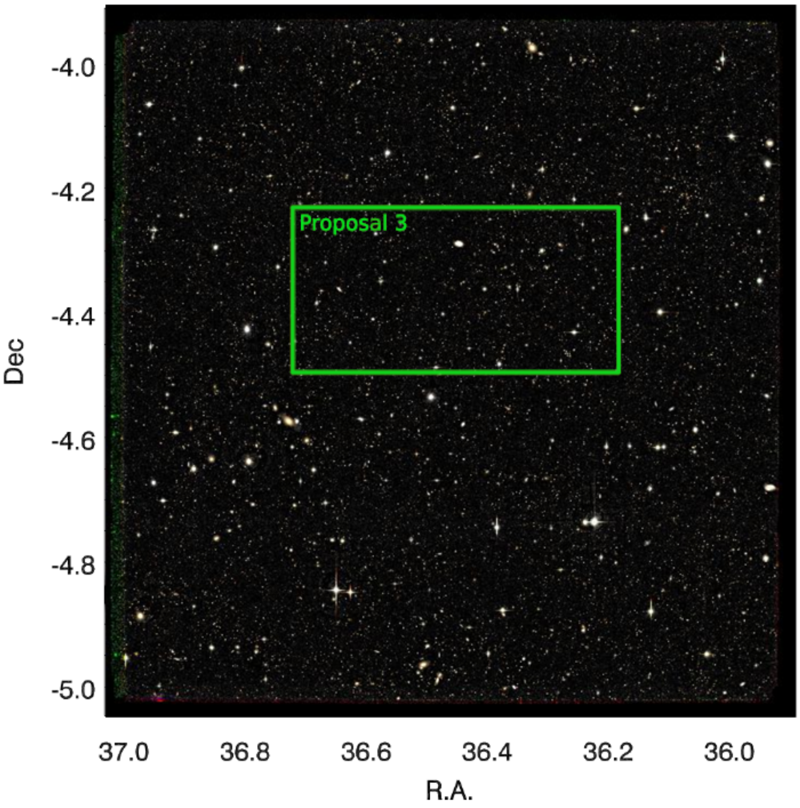


Figure 11-1 – Position and area of the "Proposal 3" region on the CFHTLS D1 (T0007bis stacks) field.

Figure 11-2 – *gri* image of the full "Proposal 3" region of the CFHTLS D1 field (T0007bis stacks).

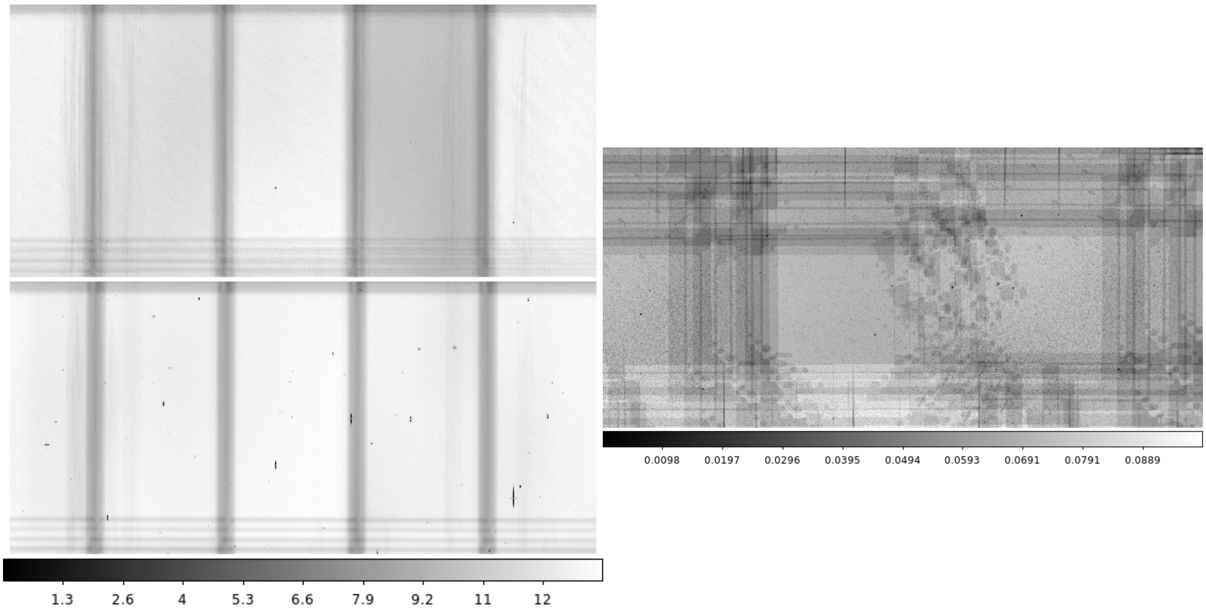


Figure 11-3 – *Left*: weight maps of the “Proposal 3” region of the CFHTLS D1 field (T0007bis stacks) in bands u and i (from top to bottom respectively). *Right*: weight map of the same region in band K_s.

configuration allows to generate images with color-magnitude distributions that are at least comparable to the observed distributions, although the discrete number of galaxy populations introduced by STUFF generates linear features in the distribution due to the fixed SED of each type, as displayed in Figure 11-6. The B/T ratios in the g adopted reference band are determined using the distribution of B/T in g-band as a function of morphological type from EFIGI (Baillard et al. 2011) measurements (de Lapparent, private communication). In particular we consider our Irr population to be bulgeless. The input parameters used to generate the 3 populations are listed in Tables 11.2 and 11.3 for the luminosity and size characteristics respectively. Galactic extinction reddens the apparent magnitude of the modeled galaxy populations following the method detailed in Section 10.3.1, and using total extinction values A_λ = 0.113, 0.045 and 0.008 in bands uiK_s respectively, according to the prescriptions of Schlafly and Finkbeiner (2011) (cf Table 10.2). Moreover, we use an LMC law (cf Figure 10-11) as the dust internal attenuation curve.

The synthetic images use a constant PSF of 35×35 pixels modeled by Andrea Moneti from the CFHTLS D1 T0007bis data (cf Section 10.2.3), and include correlated background noise due to resampling (cf Section 10.2.1), using a LANCZOS3 kernel with a correlation scale of 1 in u and i bands, and of 1.613 in band K_s. Even though there are no stars in the simulated image, this test uses the Terapix bright stars mask files in the source extraction process (cf Section 10-12) to make the test in the exact same conditions as with the real CFHTLS Deep stack (i.e. using the same image area). The weight maps of the “proposal 3” field (cf Figure 11-3) are also included in the image simulation and source extraction steps of the pipeline. The imaging characteristics for this test are summarized in Table 11.4, and the input image generated with these settings is shown in gri colors in Figure 11-5.

Table 11.1 – Uniform prior bounds for the parameters of the luminosity functions, and their evolution with redshift for the "3 populations" test.

Parameter	ϕ^*	M^*	α	ϕ_e	M_e
lower bound	10^{-6}	-22	-2.5	-3	-2.5
upper bound	$10^{-1.5}$	-17	0	2	0

The parameters to infer in this case are the 5 evolving LF parameters (cf Section 4.6) for each of the populations: ϕ^*, M^*, α, ϕ_e and M_e, that is a total of 15 parameters. In this test the size of the bulge and disk of different populations evolve according to the values of Table 11.3, the same values as used in previous tests, but we do not infer the size distribution and evolution parameters. The observables are the SExtractor FLUX_AUTO in bands uiK_s, which leads to a three-dimensional observable space. The pre-processing step of the observables is carried exactly as in previous tests: the observables multidimensional vector is decorrelated and its dynamic range is reduced using Equation (7.14) with the parameters of the reduction function listed in Table 9.3. Dividing our observable space into 10 bins per observable, we obtain a total number of 10^3 bins.

SExtractor is configured according to the prescription of the T0007 CFHTLS release documentation (Hudelot et al., 2012). We use it in double image mode, in which sources are detected in the *i*-band, and the background is estimated and subtracted automatically with a 256×256-pixels background mesh size. In order to optimize the detectability of faint extended sources, detection is performed on the images convolved with a 7×7 pixels Gaussian mask having a full width at half maximum (FWHM) of three pixels, that approximates the size of the PSF and acts as a matched filter. Finally, the detection threshold is set to 1.2 times the (unfiltered) RMS background noise above the local sky level. In order for the results concerning faint sources near the detection limit not to depend too closely on the underdensities compared to the background, all negative fluxes and radii are clipped to 0 after extraction.

The exploration of the parameter space is carried out using the sampling algorithm described in Algorithm 3, but this time based on the more robust Adaptive Metropolis algorithm of Haario et al. (2001) detailed in Section 6.4.2, with an update frequency of 100 iterations for the temperature and covariance matrix of the proposal kernel. The temperature is computed using 20 realizations of the model with the same input parameters, as done in previous tests. We rely on flat priors for the free parameters of our model, with essentially the same bounds as in previous tests, except for the ϕ^* LF parameter, whose bounds are made slightly tighter (from 10^{-6} to $10^{-1.5}$ $h^3\mathrm{Mpc}^3$ instead of from 10^{-7} to 10^{-2} $h^3\mathrm{Mpc}^3$ previously) in order to prevent Stuff from generating unnecessary weakly populated images. The parameters of our priors are listed in Table 11.1.

Results

We run the pipeline on a heterogeneous computing cluster of 7 machines totaling 152 CPU cores. We launched 2 chains in parallel for 59,200 and 63,391 iterations respectively, with uniformly distributed random starting points, using 74,540 CPU hours in total (representing ~ 67 days in wall-clock time). The burn-in is estimated by visual inspection of the trace plot: the trace of all parameters appears to stabilize after the first 10^4 iterations. Convergence is attested by the Gelman-Rubin test results over the last 49,200 iterations of each chain, as shown in Table

Table 11.2 – Luminosity characteristics of the galaxies in the “3 populations” test.

Population	SED_b [a]	SED_d	ϕ^* [h^3 Mpc^{-3}]	M^*	α	ϕ_e	M_e	B/T	T [b]	$\alpha(T)$	N [c]
E/S0	E	E	1.0e-2	-19.97	-0.5	-1.53	-1.77	0.65	-5	0.0	1819
Sp	E	Scd	4.2e-3	-19.84	-1.3	0.03	-1.95	0.25	6	1.47	28602
Irr	E	Im	2.0e-4	-19.00	-1.8	1.50	-2.0	0.00	10	0.75	30263

Notes: The LF parameters are given for H_0 = 100 km.s^{-1}.Mpc^{-3}.

[a] The bulge and disk SEDs are Coleman et al. (1980) templates.

[b] de Vaucouleurs (1959) revised morphological type.

[c] Number of sources generated by one realization of STUFF.

Table 11.3 – Size parameters for the bulge and disk of each galaxy population in the “3 populations” test.

Disk				Bulge		
β_d	r_d^* [h^{-1}kpc]	γ_d	σ_λ	M_{knee}	r_{knee} [h^{-1}kpc]	γ_b
-0.214	3.85	-0.80	0.36	-20.0	1.58	-1.00

Note: All the parameters above are given for H_0 = 100 km.s^{-1}.Mpc^{-3}.
“b” refers to bulge, and “d” to disk.

Table 11.4 – Imaging characteristics of the CFHTLS+WIRDS used for SKYMAKER in the “3 populations” test

Passband	u	i	K_s
Image size [pixels]	9538×4446	9538×4446	9538×4446
Effective gain [e^-/ADU]	74590	6807	2134
Well capacity [e^-]	∞	∞	∞
Saturation level [ADU]	6465	4230	110884
Effective read-out noise [e^-]	4.2	4.2	30
Total exposure time [s]	1	1	1
Zero-point magnitude [“ADU/s”]	30	30	30
Effective wavelength [μm]	0.381	0.769	2.146
Sky level [AB mag/arcsec^2]	22.2	20.0	15.4
Seeing FWHM [arcsec]	0.87	0.76	0.73
Scale of the correlated noise	1	1	1.613

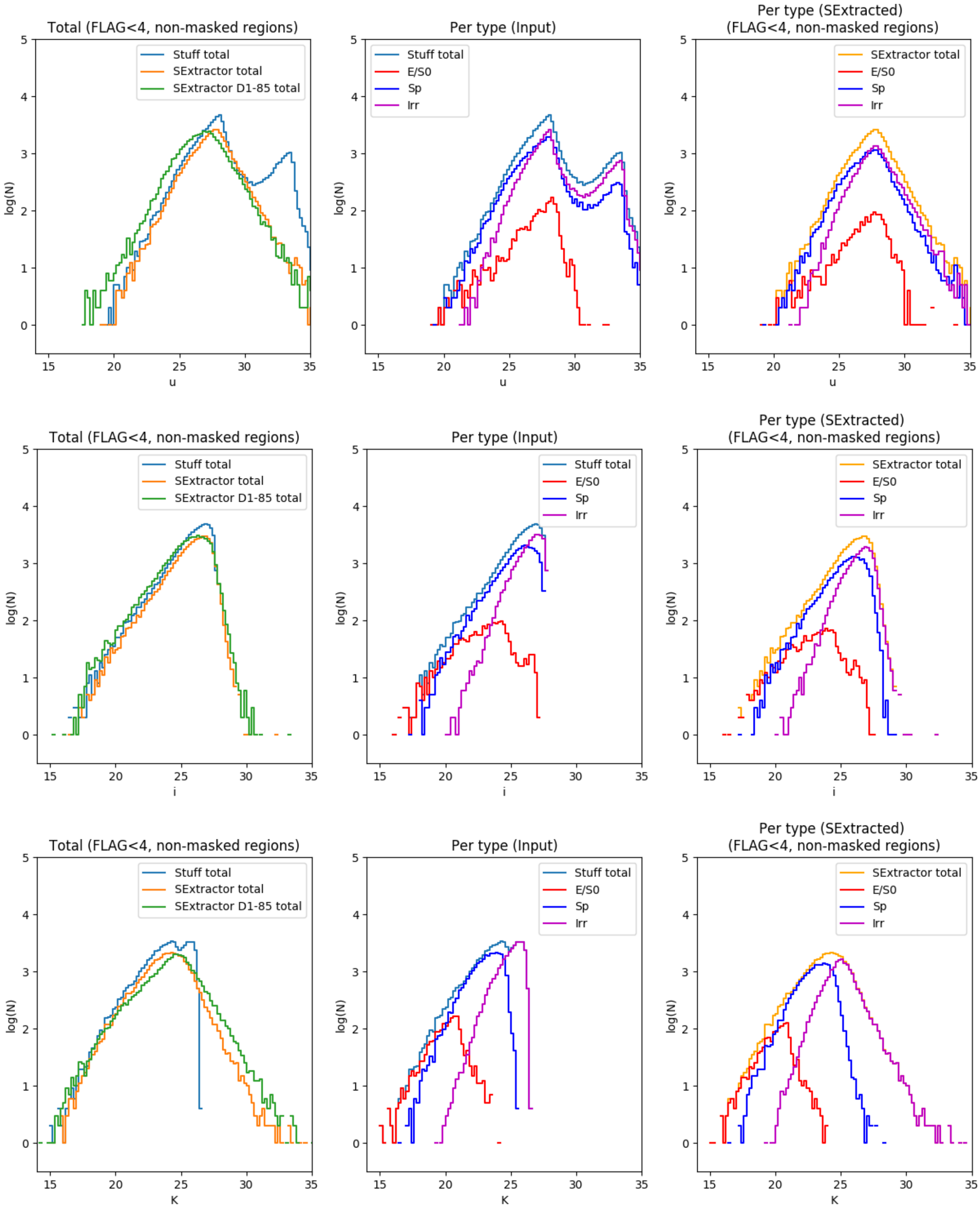


Figure 11-4 – Number counts in uiK_s for the simulated "proposal 3" data with 3 populations used for tests. In $u-$band, a second peak in the number counts can be observed at $u > 30$, probably a result of the interplay between the luminosity evolution and the selection effects implemented in Stuff. However, the peak lies beyond the limiting magnitude of the CFHTLS Deep ($u_{lim} = 26.3$, cf Section 9.2), therefore the sources at these magnitudes do not influence the distribution of observables retrieved by SExtractor.

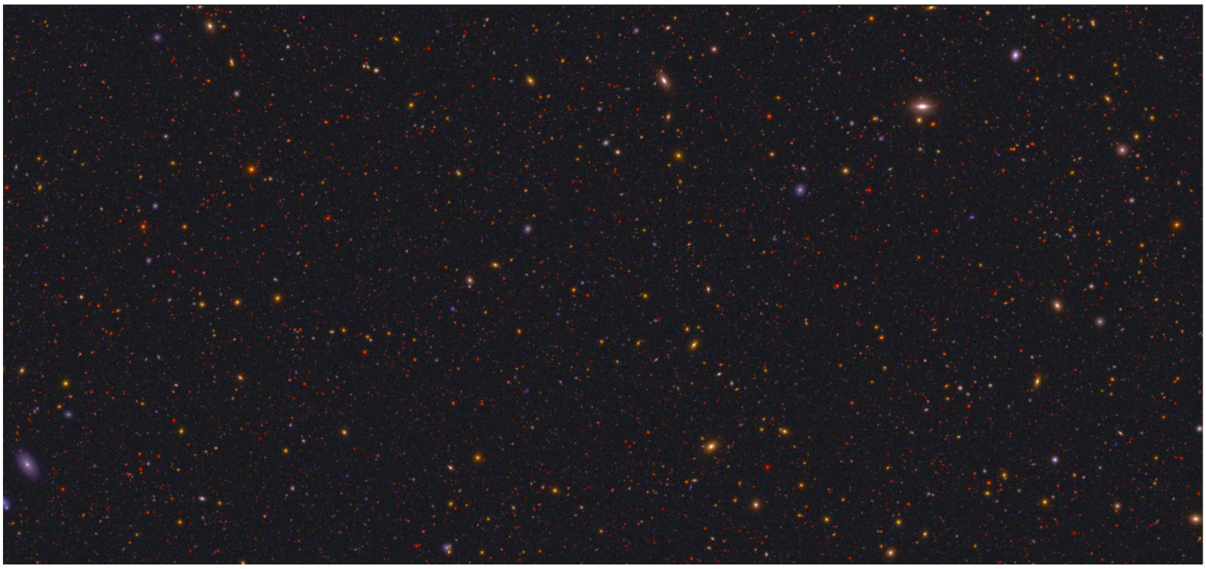


Figure 11-5 – Simulated image in bands *gri* featuring 3 populations of galaxies (E/S0, Sp, Irr) in the area of the “Proposal 3” region of the CFHTLS D1 field, used as observed data for the “3 populations” test described in Section 11.2.

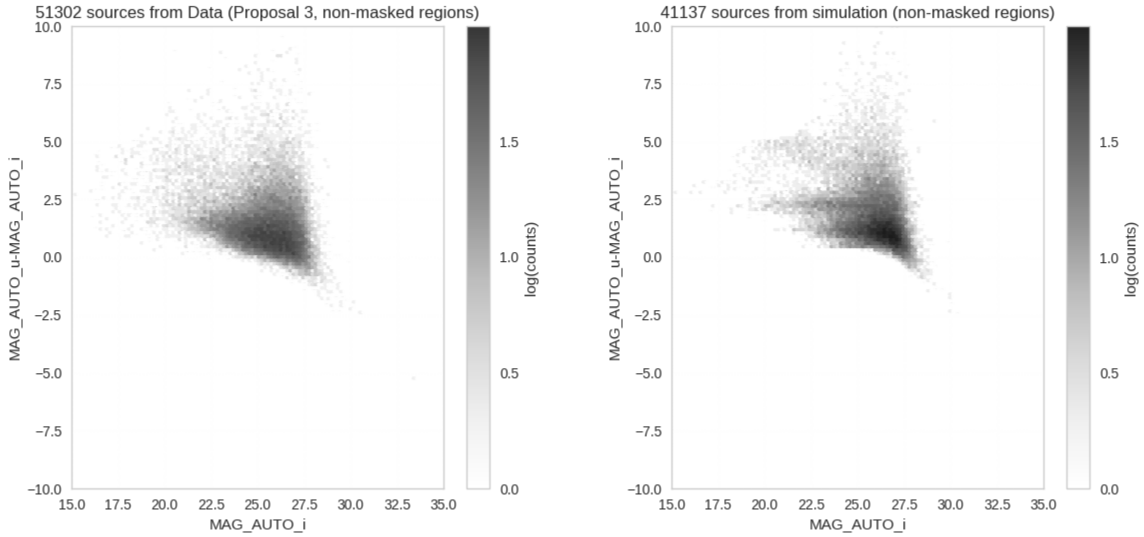


Figure 11-6 – Color-magnitude diagram comparison between the CFHTLS D1 (T0007bis stack) “proposal 3” field and the simulated “proposal 3” field used as data for the “3 populations” test detailed in Section 11.2. The distributions of galaxies occupy roughly the same range in the diagram, although effects of the discretization of galaxy populations can be observed in the simulated case as linear features. The observed color-magnitude distribution appears smoother, as galaxy populations represent a continuum and therefore occupy the color-magnitude space in a more uniform way.

Table 11.5 – Results of the Gelman-Rubin test for the “3 populations” test.

Population	$\log_{10}(\phi^*)$	M^*	α	ϕ_e	M_e
E/S0	1.000	1.001	1.027	1.074	1.000
Sp	1.010	1.000	1.003	1.004	1.004
Irr	1.033	1.040	1.008	1.001	1.004

Note: The values of $\sqrt{R}$ (cf Section 8.1.4) are obtained using 2 chains, whose burn-in phase for each chain is determined by eye. All values are < 1.1, which is a hint that in each case,
all the chains have converged to the same distribution.
The parameters above are given for $H_0 = 100$ km.s^{-1}.Mpc^{-3}.

11.5. The joint posterior PDF resulting from the combined accepted states of all the chains after burn-in contains 6,957 accepted iterations out of 102,591 propositions, corresponding to an overall 6.8% acceptance rate after burn-in. The posterior plots for population E/S0, Sp and Irr are shown in Figure 11-7, 11-8, and 11-9 respectively, and the corresponding summary statistics are displayed in Table 11.6.

The graphs show that all input parameters lie within the 68% credible region of the posterior, which means that the updated pipeline still generates consistent constraints with the input parameters. However, we notice a significant loss of precision compared to the previous tests performed in Section 9.5, as measured by the increased width of the various credible regions. This loss of precision appears to be population-dependent: the widening effect is most striking in the posterior plot of the E/S0 population, where the size of the 99% credible regions almost fills the entire parameter space.

This is not merely a zoom effect arising from the different prior bounds used for this test relative to previous tests. In fact, we compared the projected width of the 68% credible regions for each parameter of each population. The results, shown in Table 11.7, show that the widths of the marginal posteriors in this test are systematically wider than those of the test with 2 populations, with a relative widening ranging from 110% for the M_e parameter of the Sp population to 687% wider for the $\log_{10}\phi^*$ parameter of the E/S0 population. This can be attributed to the relatively low number of galaxies in each population, especially of E/S0 over the input mock image (only 1,819 generated by Stuff, compared to the 10,447 E/S0s in the test with 2 populations), which generates little constraints on the properties of this population. Therefore one might expect that more precise inferences can be performed with our pipeline as the number of extracted galaxies in each population increases in input images (or equivalently as we use input images with larger areas on the sky).

11.3 Application to CFHTLS data

Our pipeline is now ready to be applied on real galaxy survey images. As input data, we use the “proposal 3” field image of the CFHTLS D1 T0007bis stack in the u and i optical bands (cf Section 11.1), as well as the same area in the WIRDS field (T0002 stack) in the K_s infrared band. The pipeline is run in the same configuration as the “3 populations” test detailed in Section 11.2, the only difference being the bounds of the uniform prior distribution for the ϕ^* and α parameters, listed in Table 11.8. Here, we have broadened the prior intervals compared to the “2 populations” test (cf Section 9.5.1) in order to include LF values that,

Table 11.6 – Summary statistics on the marginalized posterior distributions for the galaxy populations in the “3 populations” test and comparison with the input values.

Population	Parameters	Input value	Mean	MAP [a]	68% interval	95% interval	99% interval
E/S0	$\log_{10}(\phi^*)$	-2.00	-3.22	-2.33	[-3.95,-1.75]	[-5.48,-1.53]	[-5.89,-1.53]
	M^*	-19.97	-20.11	-20.83	[-21.87,-18.98]	[-21.97,-17.48]	[-21.97,-17.03]
	α	-0.5	-1.10	-1.06	[-1.76,-0.31]	[-2.19,-0.01]	[-2.36,-0.01]
	ϕ_e	-1.53	-0.60	-0.38	[-2.52,1.32]	[-2.92,1.92]	[-2.92,1.97]
	M_e	-1.77	-1.39	-1.54	[-2.31,-0.81]	[-2.49,-0.19]	[-2.49,-0.04]
Sp	$\log_{10}(\phi^*)$	-2.38	-2.72	-2.52	[-3.37,-2.00]	[-3.97,-1.52]	[-4.45,-1.52]
	M^*	-19.84	-20.51	-20.83	[-21.68,-19.89]	[-21.97,-18.70]	[-21.97,-17.96]
	α	-1.3	-1.39	-1.53	[-1.77,-1.19]	[-2.05,-0.49]	[-2.17,-0.14]
	ϕ_e	0.03	0.31	-0.04	[-0.46,1.18]	[-1.07,1.88]	[-1.49,1.98]
	M_e	-1.95	-1.60	-1.91	[-2.41,-1.19]	[-2.49,-0.54]	[-2.49,-0.17]
Irr	$\log_{10}(\phi^*)$	-3.70	-3.01	-2.82	[-3.53,-1.97]	[-4.47,-1.52]	[-5.18,-1.52]
	M^*	-19	-19.09	-18.87	[-20.33,-17.28]	[-21.52,-17.08]	[-21.82,-17.03]
	α	-1.8	-1.44	-1.61	[-2.06,-0.96]	[-2.36,-0.31]	[-2.41,-0.04]
	ϕ_e	1.5	0.64	1.18	[0.33,1.97]	[-1.92,1.97]	[-2.92,1.97]
	M_e	-2.0	-1.36	-1.69	[-2.41,-0.39]	[-2.49,-0.24]	[-2.49,-0.06]

Notes: The LF parameters are given for H_0 = 100 km.s^{-1}.Mpc^{-3}.

[a] Maximum A Posteriori

Table 11.7 – Comparison of the width of the 68% credible regions between the results of the test with 2 populations (Section 9.5.1) and those of the test with 3 populations (Section 11.2).

Population	Parameter	Width of the 68% interval (test 3 pop)	Width of the 68% interval (test 2 pop)
E/S0	$\log_{10}(\phi^*)$	2.2	0.32
	M^*	2.89	0.8
	α	1.45	0.25
	ϕ_e	3.84	0.67
	M_e	1.5	0.7
Sp	$\log_{10}(\phi^*)$	1.37	0.48
	M^*	1.79	1.25
	α	0.58	0.12
	ϕ_e	1.64	0.63
	M_e	1.22	1.10
Irr	$\log_{10}(\phi^*)$	1.56	∅
	M^*	3.05	∅
	α	1.1	∅
	ϕ_e	1.64	∅
	M_e	2.02	∅

Notes: The LF parameters are given for H_0 = 100 km.s^{-1}.Mpc^{-3}.

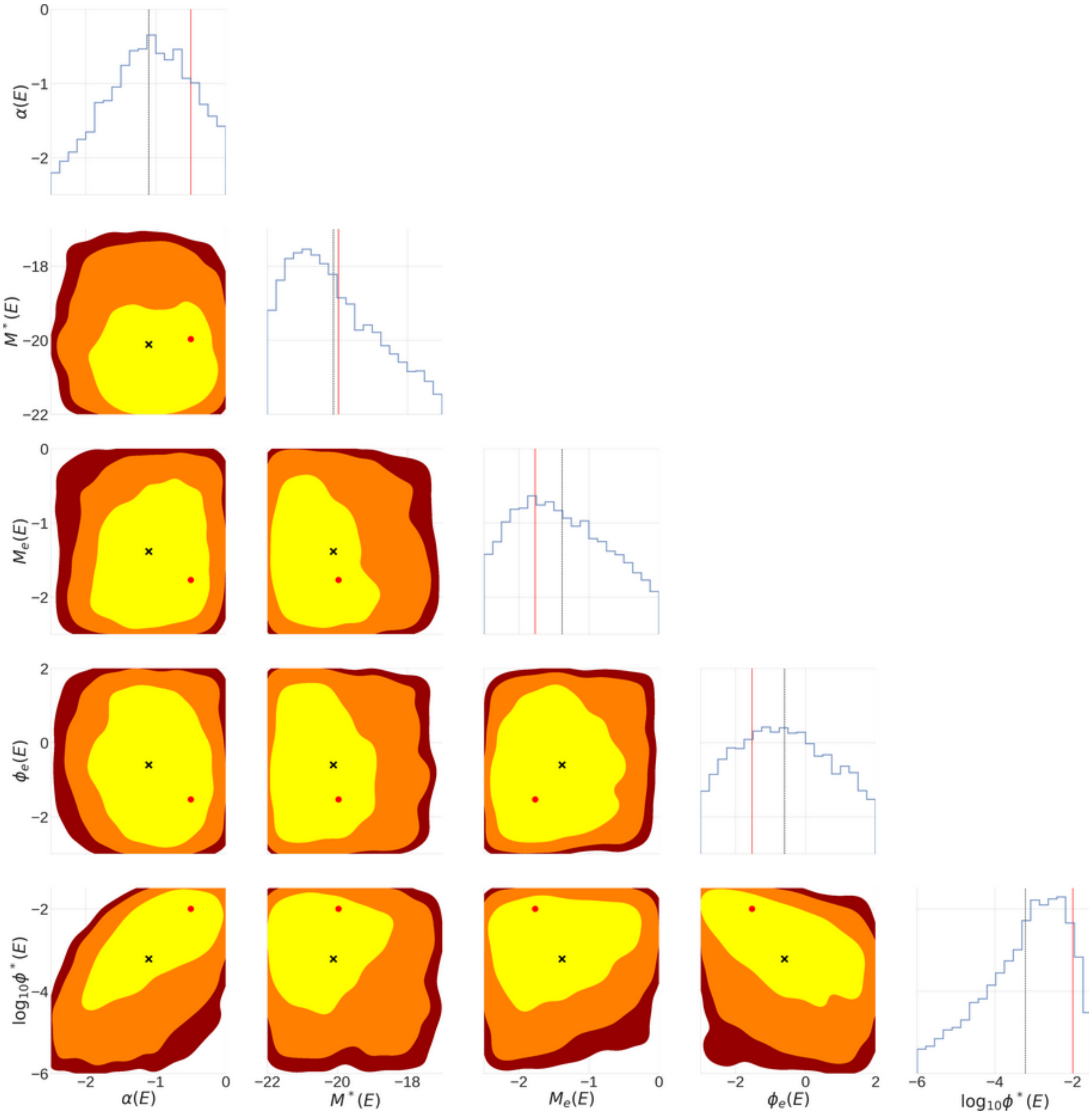


Figure 11-7 – Joint posterior distribution for the parameters of the E/S0 population resulting from the “3 populations” test. Each panel is bounded by the prior range values. The dark red, orange and yellow areas in the contour plots represent the 99%, 95% and 68% credible regions respectively. The black crosses and red dots are the mean of the posterior and input true value resp. In the marginalized posterior panels, the black dotted and red lines represent the posterior mean and the true value resp.

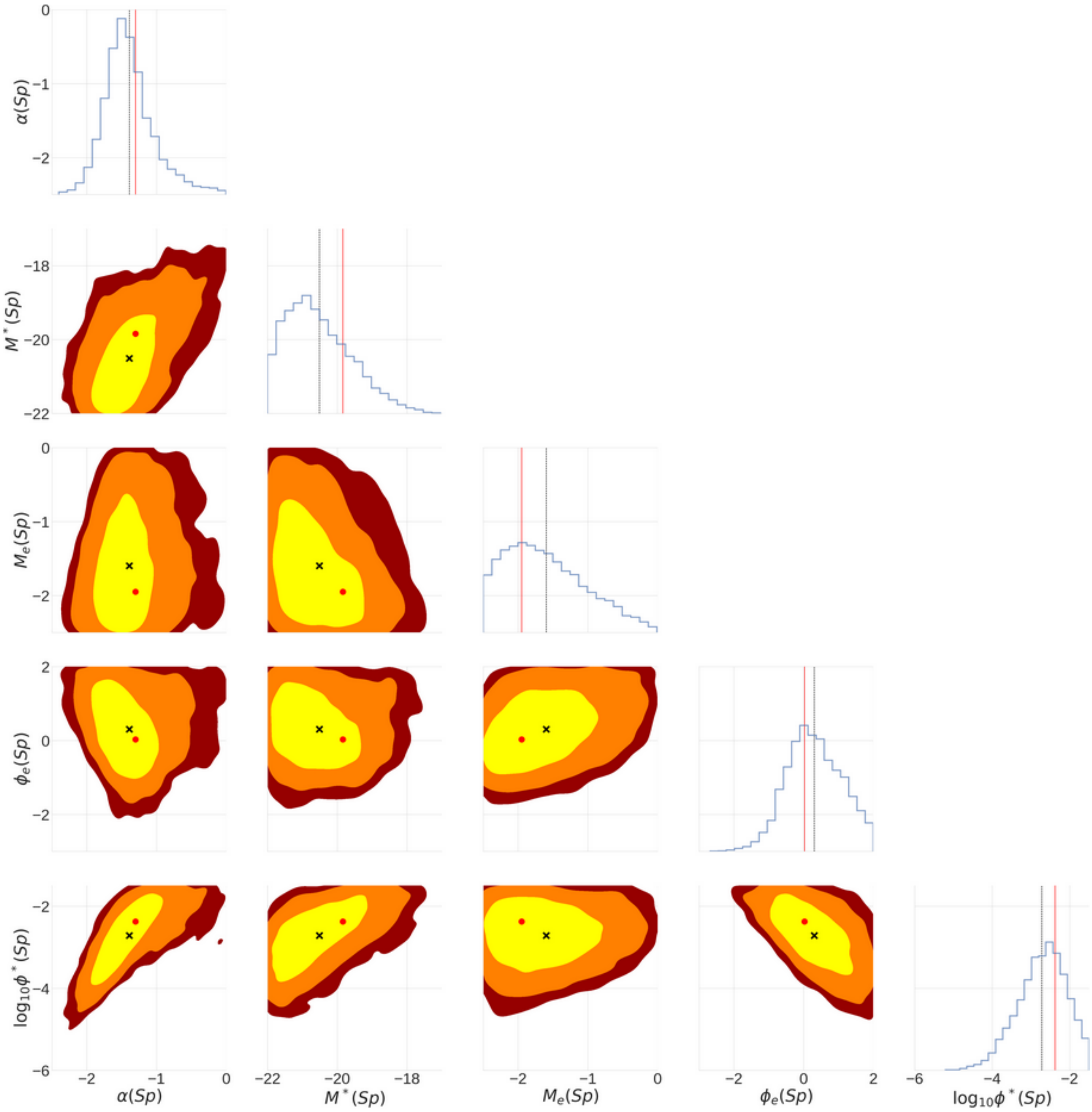


Figure 11-8 – Joint posterior distribution for the parameters of the Sp population resulting from the “3 populations” test. Each panel is bounded by the prior range values. The dark red, orange and yellow areas in the contour plots represent the 99%, 95% and 68% credible regions respectively. The black crosses and red dots are the mean of the posterior and input true value resp. In the marginalized posterior panels, the black dotted and red lines represent the posterior mean and the true value resp.

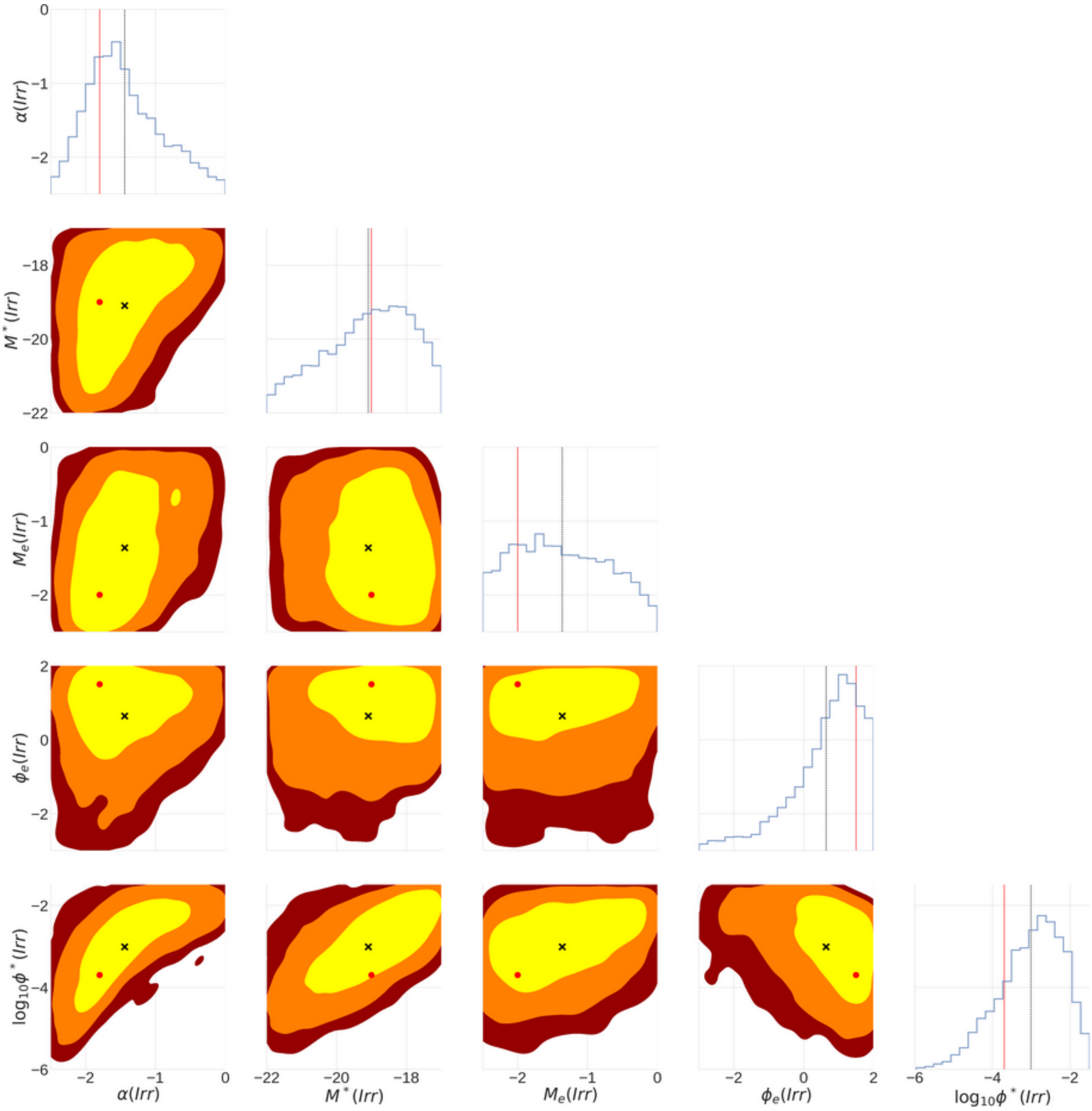


Figure 11-9 – Joint posterior distribution for the parameters of the Irr population resulting from the “3 populations” test. Each panel is bounded by the prior range values. The dark red, orange and yellow areas in the contour plots represent the 99%, 95% and 68% credible regions respectively. The black crosses and red dots are the mean of the posterior and input true value resp. In the marginalized posterior panels, the black dotted and red lines represent the posterior mean and the true value resp.

Table 11.8 – Uniform prior bounds for the parameters of the luminosity functions, and their evolution with redshift for the data application.

Parameter	ϕ^*	M^*	α	ϕ_e	M_e
lower bound	10^{-6}	-22	-2.5	-3	-2.5
upper bound	10^{0}	-17	2	2	0

even though they lead to high numbers of galaxies generated by STUFF, are still physically plausible (hence the likelihood needs to be computed). In this situation, the observables are the SExtractor FLUX_AUTO in bands uiK_s, and the free parameters of the model are the 5 evolving LF parameters for each population (E/S0, Sp, Irr): ϕ^*, M^*, α, ϕ_e and M_e, which leads to a parameter space of 15 dimensions.

Results

We run the pipeline on a heterogeneous cluster of 7 machines totaling 152 CPU cores. We launched 10 chains in parallel, with starting points randomly generated from a Gaussian distribution centered on the LF values of the “3 populations” test and with a variance of 0.1 for each parameter, a first guess that should prevent the chains from starting in completely unrealistic regions of the parameter space. The Markov chains are expected to converge by the day of this thesis defense[4].

---

[4] Hopefully...

Chapter 12

Conclusion and perspectives

In the present thesis, we have laid the basis for a new method to infer robust constraints on galaxy evolution models. In this method, populations of synthetic galaxies are generated by an empirical model, sampled from luminosity functions for each galaxy type, and determined by the SEDs of the bulge and disk components and the bulge-to-total ratio. In order to reproduce the selection effects affecting real catalogs, we use an image simulator to reproduce realistic survey images with the appropriate instrumental characteristics. Real and mock images undergo the same source extraction process. After a pre-processing step, the distributions of extracted observables (fluxes and effective radii) are then compared via a binned maximum likelihood method, and we minimize their discrepancy using an adaptive MCMC sampling scheme in the parametric Bayesian indirect likelihood framework, designed for an efficient exploration of the parameter space.

This is one of the first attempts in the field of galaxy evolution to make image simulation a central part of the inference process. We have tested the self-consistency of this approach using a simulated image of a CFHTLS Deep field covering 1 deg^2 on the sky in three bands: u and i in the optical, and K_s in the near infrared, generated with the Stuff model containing E/S0 and spiral galaxies with evolving size and luminosity.

Starting from non-informative uniform priors, we found that our pipeline can reliably infer the input parameters governing the luminosity and size evolution of each of the galaxy populations in a few 10^4 iterations, using few and disjointed observables, that is, the photometry (fluxes and radii) of the extracted sources in uiK_s. In each test performed, the input parameters lie within the 68% highest posterior density region. We have also compared the results of our method with those of the classical photometric redshifts approach, with measurements from SED fitting on one of the mock sample, and found that when using the same set of observables (uiK_s photometry), our inference pipeline yields more accurate results.

As the validity of our pipeline has been established on mock data, we have applied it to a fraction of the area of an observed CFHTLS Deep field covering ~0.1 deg^2 in the sky, in order to infer the luminosity distributions and their evolution of three galaxy populations (E/S0, Sp and Irr). We are now waiting for the results of the inference, but further tests performed on realistic mock data with the same area containing these three populations suggest that a loss of accuracy in the inference can be expected compared to the results of the first test cases with only two populations.

The next logical step for this project would use larger areas of the CFHTLS (i.e. the full D1+W1 fields or the 4 Deep fields) as input data to better constrain the input parameters of the model. Another possible future work could consist in combining these data with several extragalactic surveys at various depths and with different instrumental setups simultaneously, such as HUDF (Williams et al., 2010) at z ~ 2, and SDSS (Blanton et al., 2003) at z ~ 0.1, in order to constrain the evolution of galaxies in a wide range of redshifts. Nevertheless, such application will raise various modeling issues. In particular, real survey images display a continuum of galaxy populations, and our model only generates a discrete number of galaxy populations, defined by their bulge and disk SEDs and their B/T ratio. In practice, the number of modeled populations will be limited by computing time, as more populations lead to more free parameters to infer, hence to more iterations for the pipeline to find the high probability regions. This will certainly require a compromise between the desired accuracy of the modeled universe and convergence of the chains within a reasonable computing time.

One issue of concern in the posterior distributions that we derived with our pipeline is illustrated by the fact that the 68% shaded region is large compared to the distance between the input parameters and the mean of the posterior. Two causes of this enlargement are suspected: first, the small number (only $\sim 10^3$ for the E/S0 population in the “3 populations” case) of input galaxies spread on input images for tests generates few constrains for the possible parameter values of their luminosity function. Using surveys with large statistics in the number of characteristic population of galaxies is, as always, preferable, and should limit this effect, as attested by the fact that the populations with the most numerous galaxies in input images (i.e. Sp, Irr) have their parameters better constrained than the populations with less galaxies in all our test runs (i.e. E/S0). We also suspect the “temperature” term, that we use in order to circumvent the stochastic nature of each model realization, to play a role in the size of our posteriors. In essence, the model’s stochasticity itself (galaxies are randomly drawn from the distribution functions) inevitably contributes to the uncertainty in the posterior. We have no quantitative estimate of this enlargement and we suspect it might be a limiting factor on the precision of the parameter inference. Estimating this enlargement from simulated data would have required us to generate a very large number of realizations for each step of the chain (hence we could have turned off the “temperature” term). This would, however, be prohibitive in computing time, even in the considered simple tests performed in this thesis.

Moreover, there is room for several technical improvements of our pipeline, in order to guarantee a faster convergence and a more accurate inference:

- As implemented in the present thesis, our method faces the inevitable curse of dimensionality. In fact, as we bin each observable over ten intervals, for each observable added the hyper-dimensional number of bin increases by one order of magnitude. This limits our approach to a restricted number of observables in order to prevent memory errors. In order to adapt this method to higher numbers of observables, we may have to change our strategy and reduce the dimensionality of the problem, either via principal component analysis or by binning only projections of the datasets instead of binning the complete observable space, with the drawback of losing mutual information. More sophisticated methods using unsupervised machine learning, such as autoencoders[1] trained on large simulated datasets, could be used to minimize the loss of information induced by the dimensionality reduction.

- Instead of binning the distribution of observables, whose results depend on the bin edges

---

[1] Autoencoders are neural networks that are trained to learn a representation (or “encoding”) of a dataset for the purpose of dimensionality reduction.

and bin width, a more reliable method for density estimation for multivariate data is Kernel Density Estimation (KDE). KDE transforms the data points into a smooth density field, and alleviates the dependence of the results on the bin edges by centering a unimodal function with mean 0, the kernel, on each data point. In practice, KDE is more computationally expensive than binning, and also requires some level of hand tuning, in the form of the right kernel function and the optimal bandwidth, which in KDE is the analog of bin width in histograms.

- The mean runtime of an MCMC chain in the context of the test cases described in Chapter 9 is approximately two weeks. Up to 50% of this runtime is currently spent in job scheduling latencies for each iteration (as shown in Figure 9-4). A more integrated approach, based, for example, on Message Passing Interface (MPI) and operating only in memory might reduce those latencies. The next step would be to increase computational efficiency by offloading the most time-consuming image rendering and source extraction tasks to graphics processing units (GPUs), especially convolutions and rasterizations.

- We emphasize that on the order of 10^4 iterations is needed to attain convergence in the test cases studied. Considering the high computational cost of this approach, one may wonder how to attain faster convergence in realistic frameworks (i.e. when applied to the full area covered by imaging surveys, deep or shallow). In that regard, Gutmann and Corander (2016), who explored the computational difficulties arising from likelihood-free inference methods, proposed a strategy based on probabilistic modeling of the relation between the input parameters and the difference between observed and synthetic data. This approach would theoretically reduce the number of iterations needed to perform the inference.

- Instead of using the evolving luminosity functions as input parameters of our empirical model to generate simulated populations of galaxies whose SEDs are fixed, a more physically-driven approach would make the SEDs of the bulge and disk of the different populations evolve with redshift via a stellar population synthesis model such as PEGASE.2 (Fioc and Rocca-Volmerange 1999). In this case, the free parameters to constrain would be the parameters governing the evolution of the stellar mass functions of each component (bulge or disk) of each galaxy population with redshift. But such approach would be much more demanding in terms of computing resources than the simple empirical approach developed in this thesis, as it would lead to an increase in the number of input parameters to infer.

Finally, more realistic mock astrophysical images would lead to more precise inferences:

- The addition of a likely stellar field to the simulated images would contaminate the source extraction process in a realistic way. This could be done via the use of photometric catalogs from real or simulated stellar surveys (e.g. Gaia Collaboration et al. 2016, Robin et al. 2012).

- The contribution of clustering and environmental effects influence the colors (e.g. Madgwick et al. 2003, Bamford et al. 2009) and spectral types (Zehavi et al., 2002) of galaxies as a function of the local density and location within the cosmic web: red and quiescent galaxies are mostly distributed in regions of high density, such as the centers of clusters, whereas blue and star forming galaxies are less clustered. Galaxy clustering also has an impact on source blending and confusion. These effects are not implemented in STUFF, and this might bias our results in a way that is difficult to estimate.

- The present application uses as a reference the CFHTLS Deep survey, whose sensitivity is very homogeneous over the field of view. This is, however, not the case for many surveys. A more general application of the method would require simulating each individual raw survey exposure, and performing the very same co-addition as with the observed data to generate stacks, hence reproducing all the observational effects affecting the reduced images, which would require a lot of computing power but could be done for large and shallow surveys such as the Dark Energy Survey (Collaboration et al. 2005).

Because the pipeline in this work makes it possible to constrain not only the galaxy luminosity evolution, but also the evolution of galaxy sizes, it opens interesting perspectives for addressing the current debate on the evolution of galaxy sizes with cosmic time. The contradictory results of, for example, Longhetti et al. (2006), Trujillo et al. (2007), Saracco et al. (2010), and Morishita et al. (2014) on the growth of massive early-type galaxies may be plagued with the varying selection effects in the surveys on which these analyses are based.

The broad features of the technique we detailed in this thesis are model-independent and can be applied to a variety of problems. In fact, similar forward modeling approaches have been used recently to infer the distribution of galaxy redshifts in cosmological galaxy samples (e.g. Herbel et al. 2017), or to constrain cosmological parameters directly from image simulations reproducing the effects of “cosmic shear”, the weak gravitational lensing caused by large scale structure (e.g. Refregier and Amara 2014, Fenech Conti et al. 2017, Hildebrandt et al. 2017). The forward modeling method, although expensive in terms of computing time, could provide astronomers with promising and reliable results in the context of future large-scale surveys such as the Dark Energy Survey (Collaboration et al. 2005) or Euclid (Laureijs et al. 2011).

List of Figures

List of Tables

Bibliography

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., Prieto, C. A., An, D., Anderson, K. S., Anderson, S. F., Annis, J., Bahcall, N. A., et al. (2009). The seventh data release of the sloan digital sky survey. *The Astrophysical Journal Supplement Series*, 182(2):543.

Abraham, R., Tanvir, N., Santiago, B., Ellis, R., Glazebrook, K., and Bergh, S. v. d. (1996). Galaxy morphology to i= 25 mag in the hubble deep field. *Monthly Notices of the Royal Astronomical Society*, 279(3):L47–L52.

Abraham, R. G., Van Den Bergh, S., and Nair, P. (2003). A new approach to galaxy morphology. i. analysis of the sloan digital sky survey early data release. *The Astrophysical Journal*, 588(1):218.

Adam, R., Ade, P., Aghanim, N., Akrami, Y., Alves, M., Argüeso, F., Arnaud, M., Arroja, F., Ashdown, M., Aumont, J., et al. (2016). Planck 2015 results-i. overview of products and scientific results. *Astronomy & Astrophysics*, 594:A1.

Ade, P. A., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A., Barreiro, R., Bartlett, J., Bartolo, N., et al. (2016). Planck 2015 results-xiii. cosmological parameters. *Astronomy & Astrophysics*, 594:A13.

Adye, T. J. (1998). *A Study of J/psi Production at the LEP e+e- Collider; and the Implementation of the DELPHI Slow Controls System.* PhD thesis, Lincoln College, Oxford.

Aganj, I., Yeo, B. T., Sabuncu, M. R., and Fischl, B. (2013). On removing interpolation and resampling artifacts in rigid image registration. *IEEE Transactions on Image Processing*, 22(2):816–827.

Aggarwal, C. C., Hinneburg, A., and Keim, D. A. (2001). On the surprising behavior of distance metrics in high dimensional spaces. In *ICDT*, volume 1, pages 420–434. Springer.

Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A., Barreiro, R., Bartlett, J., Bartolo, N., Battaner, E., et al. (2016). Planck 2015 results-xi. cmb power spectra, likelihoods, and robustness of parameters. *Astronomy & Astrophysics*, 594:A11.

Aguerri, J., Balcells, M., and Peletier, R. (2001). Growth of galactic bulges by mergers-i. dense satellites. *Astronomy & Astrophysics*, 367(2):428–442.

Akeret, J., Refregier, A., Amara, A., Seehars, S., and Hasner, C. (2015). Approximate Bayesian computation for forward modeling in cosmology. *Journal of Cosmology and Astroparticle Physics*, 2015(08):043–043.

Alam, S., Albareti, F. D., Prieto, C. A., Anders, F., Anderson, S. F., Anderton, T., Andrews, B. H., Armengaud, E., Aubourg, É., Bailey, S., et al. (2015). The eleventh and twelfth data releases of the sloan digital sky survey: final data from sdss-iii. *The Astrophysical Journal Supplement Series*, 219(1):12.

Angulo, R., Springel, V., White, S., Jenkins, A., Baugh, C., and Frenk, C. (2012). Scaling relations for galaxy clusters in the millennium-xxl simulation. *Monthly Notices of the Royal Astronomical Society*, 426(3):2046–2062.

Annunziatella, M., Mercurio, A., Brescia, M., Cavuoti, S., and Longo, G. (2013). Inside catalogs: a comparison of source extraction software. *Publications of the Astronomical Society of the Pacific*, 125(923):68.

Aoyama, S., Hou, K.-C., Shimizu, I., Hirashita, H., Todoroki, K., Choi, J.-H., and Nagamine, K. (2017). Galaxy simulation with dust formation and destruction. *Monthly Notices of the Royal Astronomical Society*, 466:105–121.

Arnaud, K. A. (1996). XSPEC: The First Ten Years. In Jacoby, G. H. and Barnes, J., editors, *Astronomical Data Analysis Software and Systems V*, volume 101 of *Astronomical Society of the Pacific Conference Series*, page 17.

Arnouts, S., Cristiani, S., Moscardini, L., Matarrese, S., Lucchin, F., Fontana, A., and Giallongo, E. (1999). Measuring and modelling the redshift evolution of clustering: the Hubble Deep Field North. *Monthly Notices of the Royal Astronomical Society*, 310:540–556.

Asano, R. S., Takeuchi, T. T., Hirashita, H., and Nozawa, T. (2013). What determines the grain size distribution in galaxies? *Monthly Notices of the Royal Astronomical Society*, 432:637–652.

Aslan, B. and Zech, G. (2002). A new class of binning free, multivariate goodness-of-fit tests: the energy tests. *ArXiv High Energy Physics - Experiment e-prints*.

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., Greenfield, P., Droettboom, M., Bray, E., Aldcroft, T., Davis, M., Ginsburg, A., Price-Whelan, A. M., Kerzendorf, W. E., Conley, A., Crighton, N., Barbary, K., Muna, D., Ferguson, H., Grollier, F., Parikh, M. M., Nair, P. H., Unther, H. M., Deil, C., Woillez, J., Conseil, S., Kramer, R., Turner, J. E. H., Singer, L., Fox, R., Weaver, B. A., Zabalza, V., Edwards, Z. I., Azalee Bostroem, K., Burke, D. J., Casey, A. R., Crawford, S. M., Dencheva, N., Ely, J., Jenness, T., Labrie, K., Lim, P. L., Pierfederici, F., Pontzen, A., Ptak, A., Refsdal, B., Servillat, M., and Streicher, O. (2013). Astropy: A community Python package for astronomy. *Astronomy and Astrophysics*, 558:A33.

Babu, G. J. (2012). Bayesian and frequentist approaches. In *Online Proceedings of the Astronomical Data Analysis Conference*.

Babu, G. J. and Feigelson, E. D. (2006). Astrostatistics: Goodness-of-Fit and All That! In Gabriel, C., Arviset, C., Ponz, D., and Enrique, S., editors, *Astronomical Data Analysis Software and Systems XV*, volume 351 of *Astronomical Society of the Pacific Conference Series*, page 127.

Bai, Y. (2010). *Convergence of adaptive Markov chain Monte Carlo algorithms*. University of Toronto.

Bailin, J. and Harris, W. E. (2008). Inclination-independent galaxy classification. *The Astrophysical Journal*, 681(1):225.

Baillard, A., Bertin, E., de Lapparent, V., Fouqué, P., Arnouts, S., Mellier, Y., Pelló, R., Leborgne, J.-F., Prugniel, P., Makarov, D., Makarova, L., McCracken, H. J., Bijaoui, A., and Tasca, L. (2011). The EFIGI catalogue of 4458 nearby galaxies with detailed morphology. *Astronomy and Astrophysics*, 532:A74.

Balcells, M. and Peletier, R. F. (1993). Colors and color gradients in bulges of galaxies. *arXiv preprint astro-ph/9311005*.

Baldry, I. K., Glazebrook, K., Brinkmann, J., Ivezić, Ž., Lupton, R. H., Nichol, R. C., and Szalay, A. S. (2004). Quantifying the bimodal color-magnitude distribution of galaxies. *The Astrophysical Journal*, 600(2):681.

Bamford, S. P., Nichol, R. C., Baldry, I. K., Land, K., Lintott, C. J., Schawinski, K., Slosar, A., Szalay, A. S., Thomas, D., Torki, M., Andreescu, D., Edmondson, E. M., Miller, C. J., Murray, P., Raddick, M. J., and Vandenberg, J. (2009). Galaxy Zoo: the dependence of morphology and colour on environment. *Monthly Notices of the Royal Astronomical Society*, 393:1324–1352.

Barlow, R. and Beeston, C. (1993). Fitting using finite monte carlo samples. *Computer Physics Communications*, 77(2):219–228.

Barth, A. J., Strigari, L. E., Bentz, M. C., Greene, J. E., and Ho, L. C. (2008). Dynamical constraints on the masses of the nuclear star cluster and black hole in the late-type spiral galaxy ngc 3621. *The Astrophysical Journal*, 690(1):1031.

Baryshev, Y. and Teerikorpi, P. (2012). Cosmic distances and selection biases. In *Fundamental Questions of Practical Cosmology*, pages 47–68. Springer.

Bassett, R., Papovich, C., Lotz, J. M., Bell, E. F., Finkelstein, S. L., Newman, J. A., Tran, K.-V., Almaini, O., Lani, C., Cooper, M., et al. (2013). Candels observations of the environmental dependence of the color-mass-morphology relation at z= 1.6. *The Astrophysical Journal*, 770(1):58.

Battisti, A., Calzetti, D., and Chary, R.-R. (2016). Characterizing dust attenuation in local star-forming galaxies: Uv and optical reddening. *The Astrophysical Journal*, 818(1):13.

Baugh, C. M. (2006). A primer on hierarchical galaxy formation: the semi-analytical approach. *Reports on Progress in Physics*, 69(12):3101.

Bayes, T., Price, R., and Canton, J. (1763). *An essay towards solving a problem in the doctrine of chances*. C. Davis, Printer to the Royal Society of London.

Beaumont, M. A. (2010). Approximate Bayesian Computation in Evolution and Ecology. *Annual Review of Ecology, Evolution, and Systematics*, 41(1):379–406.

Beaumont, M. A., Cornuet, J.-M., Marin, J.-M., and Robert, C. P. (2008). Adaptive approximate Bayesian computation. *ArXiv e-prints*.

Beckers, J. (1995). *Scientific and engineering frontiers for 8-10 M telescopes. Instrumentation for large telescopes in the 21st century*, page 303. Iye, Masanori and Nishimura, Tetsuo.

Beckwith, S. V., Stiavelli, M., Koekemoer, A. M., Caldwell, J. A., Ferguson, H. C., Hook, R., Lucas, R. A., Bergeron, L. E., Corbin, M., Jogee, S., et al. (2006). The hubble ultra deep field. *The Astronomical Journal*, 132(5):1729.

Bell, E. F. and De Jong, R. S. (2000). The stellar populations of spiral galaxies. *Monthly Notices of the Royal Astronomical Society*, 312(3):497–520.

Bellman, R. (1972). *Dynamic programming.* Princeton University Press.

Benson, A. J. (2012). Galacticus: A semi-analytic model of galaxy formation. *New Astronomy*, 17(2):175–197.

Bergé, J., Gamper, L., Réfrégier, A., and Amara, A. (2013). An ultra fast image generator (ufig) for wide-field astronomy. *Astronomy and Computing*, 1:23–32.

Bernoulli, J. (1713). *Ars conjectandi.* Impensis Thurnisiorum, fratrum.

Bernton, E., Jacob, P. E., Gerber, M., and Robert, C. P. (2017). Inference in generative models using the Wasserstein distance. *ArXiv e-prints*.

Bernyk, M., Croton, D. J., Tonini, C., Hodkinson, L., Hassan, A. H., Garel, T., Duffy, A. R., Mutch, S. J., Poole, G. B., and Hegarty, S. (2016). The theoretical astrophysical observatory: cloud-based mock galaxy catalogs. *The Astrophysical Journal Supplement Series*, 223(1):9.

Bertin, E. (1997). Weightwatcher 1.2. *User's manual.*

Bertin, E. (2003). Swarp user's guide.

Bertin, E. (2006). Sextractor user's manual.

Bertin, E. (2009). SkyMaker: astronomical image simulations made easy. *Memorie della Societa Astronomica Italiana*, 80:422–.

Bertin, E. (2011a). Automated Morphometry with SExtractor and PSFEx. In Evans, I. N., Accomazzi, A., Mink, D. J., and Rots, A. H., editors, *Astronomical Data Analysis Software and Systems XX*, volume 442 of *Astronomical Society of the Pacific Conference Series*, page 435.

Bertin, E. (2011b). STIFF: Converting Scientific FITS Images to TIFF. Astrophysics Source Code Library.

Bertin, E. and Arnouts, S. (1996). Sextractor: Software for source extraction. *Astronomy and Astrophysics Supplement Series*, 117(2):393–404.

Bertin, E., Mellier, Y., Radovich, M., Missonnier, G., Didelon, P., and Morin, B. (2002). The terapix pipeline. In *Astronomical Data Analysis Software and Systems XI*, volume 281, page 228.

Bertin, E., Mellier, Y., Radovich, M., Missonnier, G., Didelon, P., and Morin, B. (2002). The TERAPIX Pipeline. In Bohlender, D. A., Durand, D., and Handley, T. H., editors, *Astronomical Data Analysis Software and Systems XI*, volume 281 of *Astronomical Society of the Pacific Conference Series*, page 228.

Bertin, E. and Moneti, A. (2017). Psfex documentation 3.18.2.

Beyer, K., Goldstein, J., Ramakrishnan, R., and Shaft, U. (1999). When is "nearest neighbor" meaningful? In *International conference on database theory*, pages 217–235. Springer.

Bhattacharyya, A. (1946). On a measure of divergence between two multinomial populations. *Sankhya: The Indian Journal of Statistics (1933-1960)*, 7(4):401–406.

Bianchi, L., Clayton, G. C., Bohlin, R. C., Hutchings, J., and Massey, P. (1996). Ultraviolet extinction by interstellar dust in external galaxies: M31. *The Astrophysical Journal*, 471(1):203.

Bielby, R., Finoguenov, A., Tanaka, M., McCracken, H., Daddi, E., Hudelot, P., Ilbert, O., Kneib, J., Le Fevre, O., Mellier, Y., et al. (2010). The wircam deep infrared cluster survey-i. groups and clusters at z $\geq$ 1.1. *Astronomy & Astrophysics*, 523:A66.

Bielby, R., Hudelot, P., McCracken, H. J., Ilbert, O., Daddi, E., Le Fèvre, O., Gonzalez-Perez, V., Kneib, J.-P., Marmo, C., Mellier, Y., Salvato, M., Sanders, D. B., and Willott, C. J. (2012). The WIRCam Deep Survey. I. Counts, colours, and mass-functions derived from near-infrared imaging in the CFHTLS deep fields. *Astronomy and Astrophysics*, 545:A23.

Bienayme, O., Robin, A. C., and Creze, M. (1987). The mass density in our Galaxy. *Astronomy and Astrophysics (ISSN 0004-6361), vol. 180, no. 1-2, June 1987, p. 94-110.*, 180:94–110.

Bilicki, M., Peacock, J. A., Jarrett, T. H., Cluver, M. E., and Steward, L. (2014). Mapping the cosmic web with the largest all-sky surveys. *Proceedings of the International Astronomical Union*, 11(S308):143–148.

Binggeli, B., Sandage, A., and Tammann, G. A. (1988). The Luminosity Function of Galaxies. *Annual Review of Astronomy and Astrophysics*, 26(1):509–560.

Binggeli, B., Sandage, A., and Tarenghi, M. (1984). Studies of the Virgo Cluster. I - Photometry of 109 galaxies near the cluster center to serve as standards. *The Astronomical Journal*, 89:64.

Birnboim, Y. and Dekel, A. (2003). Virial shocks in galactic haloes? *Monthly Notices of the Royal Astronomical Society*, 345(1):349–364.

Blain, A. W., Smail, I., Ivison, R., Kneib, J.-P., and Frayer, D. T. (2002). Submillimeter galaxies. *Physics Reports*, 369(2):111–176.

Blaizot, J., Wadadekar, Y., Guiderdoni, B., Colombi, S. T., Bertin, E., Bouchet, F. R., Devriendt, J. E. G., and Hatton, S. (2005). MoMaF: the Mock Map Facility. *Monthly Notices of the Royal Astronomical Society*, 360:159–175.

Blangiardo, M. and Cameletti, M. (2015). *Spatial and Spatio-temporal Bayesian Models with R - INLA*.

Blanton, M. R., Dalcanton, J., Eisenstein, D., Loveday, J., Strauss, M. A., SubbaRao, M., Weinberg, D. H., Anderson Jr, J. E., Annis, J., Bahcall, N. A., et al. (2001). The luminosity function of galaxies in sdss commissioning data. *The Astronomical Journal*, 121(5):2358.

Blanton, M. R., Hogg, D. W., Bahcall, N. A., Brinkmann, J., Britton, M., Connolly, A. J., Csabai, I., Fukugita, M., Loveday, J., Meiksin, A., Munn, J. A., Nichol, R. C., Okamura, S., Quinn, T., Schneider, D. P., Shimasaku, K., Strauss, M. A., Tegmark, M., Vogeley, M. S., and Weinberg, D. H. (2003). The Galaxy Luminosity Function and Luminosity Density at Redshift z = 0.1. *The Astrophysical Journal*, 592:819–838.

Blanton, M. R., Lupton, R. H., Schlegel, D. J., Strauss, M. A., Brinkmann, J., Fukugita, M., and Loveday, J. (2005). The Properties and Luminosity Function of Extremely Low Luminosity Galaxies. *The Astrophysical Journal*, 631:208–230.

Blumenthal, G. R., Faber, S., Primack, J. R., and Rees, M. J. (1984). Formation of galaxies and large scale structure with cold dark matter'.

Boissier, S., Boselli, A., Buat, V., Donas, J., and Milliard, B. (2004). The radial extinction profiles of late-type galaxies. *Astronomy & Astrophysics*, 424(2):465–476.

Bolker, B. M. (2008). *Ecological Models and Data in R*. Princeton University Press.

Boltzmann, L. (1877). On the relationship between the second fundamental theorem of the mechanical theory of heat and probability calculations regarding the conditions for thermal equilibrium. *Sitzungberichte der Kaiserlichen Akademie der Wissenschaften*, pages 373–435.

Boselli, A. (2011). *Dust Extinction*, pages 125–142. Wiley-VCH Verlag GmbH and Co. KGaA.

Bouchet, P., Lequeux, J., Maurice, E., Prevot, L., and Prevot-Burnichon, M. L. (1985). The visible and infrared extinction law and the gas-to-dust ratio in the Small Magellanic Cloud. *Astronomy and Astrophysics*, 149:330–336.

Bournaud, F., Elmegreen, B. G., and Elmegreen, D. M. (2007). Rapid formation of exponential disks and bulges at high redshift from the dynamical evolution of clump-cluster and chain galaxies. *The Astrophysical Journal*, 670(1):237.

Box, G. (1992). Ep, and tiao, gc,". *Bayesian Inference in Statistical Analysis*.

Briggs, W. M. (2012). It is Time to Stop Teaching Frequentism to Non-statisticians. *ArXiv e-prints*.

Bromm, V. and Loeb, A. (2003). Formation of the first supermassive black holes. *The Astrophysical Journal*, 596(1):34.

Bromm, V. and Yoshida, N. (2011). The first galaxies. *Annual Review of Astronomy and Astrophysics*, 49:373–407.

Brooks, S. and Gelman, A. (1998). Some issues for monitoring convergence of iterative simulations. *Computing Science and Statistics*, pages 30–36.

Bruzual, G. and Charlot, S. (2003). Stellar population synthesis at the resolution of 2003. *Monthly Notices of the Royal Astronomical Society*, 344(4):1000–1028.

Buat, V., Giovannoli, E., Heinis, S., Charmandaris, V., Coia, D., Daddi, E., Dickinson, M., Elbaz, D., Hwang, H., Morrison, G., et al. (2011). Goods-herschel: evidence of a uv extinction bump in galaxies at z> 1. *Astronomy & Astrophysics*, 533:A93.

Buat, V., Noll, S., Burgarella, D., Giovannoli, E., Charmandaris, V., Pannella, M., Hwang, H. S., Elbaz, D., Dickinson, M., Magdis, G., Reddy, N., and Murphy, E. J. (2012). GOODS-Herschel: dust attenuation properties of UV selected high redshift galaxies. *Astronomy and Astrophysics*, 545:A141.

Buitrago, F., Trujillo, I., Conselice, C. J., Bouwens, R. J., Dickinson, M., and Yan, H. (2008). Size evolution of the most massive galaxies at 1.7< z< 3 from goods nicmos survey imaging. *The Astrophysical Journal Letters*, 687(2):L61.

Burstein, D. (1979). Structure and origin of s0 galaxies. ii-disk-to-bulge ratios. *The Astrophysical Journal*, 234:435–447.

Burstein, D., Haynes, M. P., and Faber, S. (1991). Dependence of galaxy properties on viewing angle. *Nature*, 353(6344):515–521.

Buta, R. J. (2013). Galaxy morphology. In *Planets, Stars and Stellar Systems*, pages 1–89. Springer.

Butkevich, A., Berdyugin, A., and Teerikorpi, P. (2005). Statistical biases in stellar astronomy: the malmquist bias revisited. *Monthly Notices of the Royal Astronomical Society*, 362(1):321–330.

Byun, Y., Freeman, K., and Kylafis, N. (1994). Diagnostics of dust content in spiral galaxies: Numerical simulations of radiative transfer. *The Astrophysical Journal*, 432:114–127.

Cadez, I. V., Smyth, P., McLachlan, G. J., and McLaren, C. E. (2002). Maximum Likelihood Estimation of Mixture Densities for Binned and Truncated Multivariate Data. *Machine Learning*, 47(1):7–34.

Calvi, V., Stiavelli, M., Bradley, L., Pizzella, A., and Kim, S. (2014). The effect of surface brightness dimming in the selection of high-z galaxies. *The Astrophysical Journal*, 796(2):102.

Calzetti, D. (1997). UV opacity in nearby galaxies and application to distant galaxies. In Waller, W. H., editor, *American Institute of Physics Conference Series*, volume 408 of *American Institute of Physics Conference Series*, pages 403–412.

Calzetti, D. (2001). The Dust Opacity of Star-forming Galaxies. *Publications of the Astronomical Society of the Pacific*, 113:1449–1485.

Calzetti, D., Armus, L., Bohlin, R. C., Kinney, A. L., Koornneef, J., and Storchi-Bergmann, T. (2000). The dust content and opacity of actively star-forming galaxies. *The Astrophysical Journal*, 533(2):682.

Calzetti, D., Kinney, A. L., and Storchi-Bergmann, T. (1994). Dust extinction of the stellar continua in starburst galaxies: The ultraviolet and optical extinction law. *The Astrophysical Journal*, 429:582–601.

Cameron, E. and Pettitt, A. N. (2012). Approximate Bayesian Computation for astronomical model analysis: a case study in galaxy demographics and morphological transformation at high redshift. *Monthly Notices of the Royal Astronomical Society*, 425:44–65.

Cappellari, M., Bacon, R., Bureau, M., Damen, M., Davies, R. L., De Zeeuw, P. T., Emsellem, E., Falcón-Barroso, J., Krajnovic, D., Kuntschner, H., et al. (2006). The sauron project—iv. the mass-to-light ratio, the virial mass estimator and the fundamental plane of elliptical and lenticular galaxies. *Monthly Notices of the Royal Astronomical Society*, 366(4):1126–1150.

Carassou, S., de Lapparent, V., Bertin, E., and Le Borgne, D. (2017). Inferring the photometric and size evolution of galaxies from image simulations. *ArXiv e-prints*.

Cardelli, J. A., Clayton, G. C., and Mathis, J. S. (1989). The relationship between infrared, optical, and ultraviolet extinction. *The Astrophysical Journal*, 345:245–256.

Carilli, C. and Rawlings, S. (2004). Science with the square kilometer array: motivation, key science projects, standards and assumptions. *arXiv preprint astro-ph/0409274*.

Cash, W. (1979). Parameter estimation in astronomy through application of the likelihood ratio. *The Astrophysical Journal*, 228:939–947.

Cazaux, S. and Spaans, M. (2009). HD and H{2} formation in low-metallicity dusty gas clouds at high redshift. *Astronomy and Astrophysics*, 496:365–374.

Chakravarti, M., Laha, R. G., and Roy, J. (1967). *Handbook of Methods of Applied Statistics*, volume I. John Wiley and Sons. pp. 392-394.

Chardy, P., Glemarec, M., and Laurec, A. (1976). Application of inertia methods of Benthic Marine Ecology : Practical implications of the basic options. *Estuarine and Coastal Marine Science*, 4:179–205.

Charlot, S. and Fall, S. M. (2000). A Simple Model for the Absorption of Starlight by Dust in Galaxies. *The Astrophysical Journal*, 539:718–731.

Chib, S. and Greenberg, E. (1995). Understanding the metropolis-hastings algorithm. *The american statistician*, 49(4):327–335.

Cimatti, A., Nipoti, C., and Cassata, P. (2012). Fast evolving size of early-type galaxies at z> 2 and the role of dissipationless (dry) merging. *Monthly Notices of the Royal Astronomical Society: Letters*, 422(1):L62–L66.

Ciotti, L. and Bertin, G. (1999). Analytical properties of the r^(1/m) luminosity law. *arXiv preprint astro-ph/9911078*.

Ciotti, L. and Van Albada, T. S. (2001). The mbh-σc relation as a constraint on the formation of elliptical galaxies. *The Astrophysical Journal Letters*, 552(1):L13.

Cole, S., Lacey, C. G., Baugh, C. M., and Frenk, C. S. (2000). Hierarchical galaxy formation. *Monthly Notices of the Royal Astronomical Society*, 319(1):168–204.

Coleman, G. D., Wu, C.-C., and Weedman, D. W. (1980). Colors and magnitudes predicted for high redshift galaxies. *The Astrophysical Journals*, 43:393–416.

Collaboration, D. E. S. et al. (2005). The dark energy survey. *arXiv preprint astro-ph/0510346*.

Colless, M., Dalton, G., Maddox, S., Sutherland, W., Norberg, P., Cole, S., Bland-Hawthorn, J., Bridges, T., Cannon, R., Collins, C., et al. (2001). The 2df galaxy redshift survey: spectra and redshifts. *Monthly Notices of the Royal Astronomical Society*, 328(4):1039–1063.

Combes, F., Debbasch, F., Friedli, D., and Pfenniger, D. (1990). Box and peanut shapes generated by stellar bars. *Astronomy and Astrophysics*, 233:82–95.

Condon, J. (1974). Confusion and flux-density error distributions. *The Astrophysical Journal*, 188:279–286.

Conley, A., Guy, J., Sullivan, M., Regnault, N., Astier, P., Balland, C., Basa, S., Carlberg, R., Fouchez, D., Hardin, D., et al. (2010). Supernova constraints and systematic uncertainties from the first three years of the supernova legacy survey. *The Astrophysical Journal Supplement Series*, 192(1):1.

Conroy, C. (2013). Modeling the Panchromatic Spectral Energy Distributions of Galaxies. *Annual Review of Astronomy and Astrophysics*, 51:393–455.

Conselice, C. J. (2003). The relationship between stellar light distributions of galaxies and their formation histories. *The Astrophysical Journal Supplement Series*, 147(1):1.

Conselice, C. J. (2014). The Evolution of Galaxy Structure Over Cosmic Time. *Annual Review of Astronomy and Astrophysics*, 52:291–337.

Conselice, C. J., Rajgor, S., and Myers, R. (2008). The structures of distant galaxies–i. galaxy structures and the merger rate to z 3 in the hubble ultra-deep field. *Monthly Notices of the Royal Astronomical Society*, 386(2):909–927.

Conselice, C. J., Wilkinson, A., Duncan, K., and Mortlock, A. (2016). The evolution of galaxy number density at z< 8 and its implications. *The Astrophysical Journal*, 830(2):83.

Coupon, J., Czakon, N., Bosch, J., Komiyama, Y., Medezinski, E., Miyazaki, S., and Oguri, M. (2017). The bright-star masks for the hsc-ssp survey. *arXiv preprint arXiv:1705.00622*.

Coupon, J., Ilbert, O., Kilbinger, M., McCracken, H., Mellier, Y., Arnouts, S., Bertin, E., Hudelot, P., Schultheis, M., Le Fèvre, O., et al. (2009). Photometric redshifts for the cfhtls t0004 deep and wide fields. *Astronomy & Astrophysics*, 500(3):981–998.

Cowie, L. L., Songaila, A., Hu, E. M., and Cohen, J. (1996). New insight on galaxy formation and evolution from keck spectroscopy of the hawaii deep fields. *arXiv preprint astro-ph/9606079*.

Cowles, M. K. and Carlin, B. P. (1996). Markov Chain Monte Carlo Convergence Diagnostics: A Comparative Review. *Source Journal of the American Statistical Association*, 91(434):883–904.

Crain, R. A., Schaye, J., Bower, R. G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C. S., McCarthy, I. G., Helly, J. C., et al. (2015). The eagle simulations of galaxy formation: calibration of subgrid physics and model variations. *Monthly Notices of the Royal Astronomical Society*, 450(2):1937–1961.

Croton, D. J., Stevens, A. R., Tonini, C., Garel, T., Bernyk, M., Bibiano, A., Hodkinson, L., Mutch, S. J., Poole, G. B., and Shattow, G. M. (2016). Semi-analytic galaxy evolution (sage): Model calibration and basic results. *The Astrophysical Journal Supplement Series*, 222(2):22.

Csillėry, K., Blum, M. G. B., Gaggiotti, O. E., and François, O. (2010). Approximate Bayesian Computation (ABC) in practice. *Trends in Ecology & Evolution*, 25(7):410–418.

Cuillandre, J.-C. and Bertin, E. (2006). CFHT Legacy Survey (CFHTLS) : a rich data set. In Barret, D., Casoli, F., Lagache, G., Lecavelier, A., and Pagani, L., editors, *SF2A-2006: Semaine de l'Astrophysique Francaise*, page 265.

da Cunha, E., Charlot, S., and Elbaz, D. (2008). A simple model to interpret the ultraviolet, optical and infrared emission from galaxies. *Monthly Notices of the Royal Astronomical Society*, 388:1595–1617.

Daddi, E., Renzini, A., Pirzkal, N., Cimatti, A., Malhotra, S., Stiavelli, M., Xu, C., Pasquali, A., Rhoads, J., Brusa, M., et al. (2005). Passively evolving early-type galaxies at 1.4 < z < 2.5 in the hubble ultra deep field. *The Astrophysical Journal*, 626(2):680.

Dalcanton, J. J. (1998). Systematic Biases in Galaxy Luminosity Functions. *The Astrophysical Journal*, 495:251–266.

Dalcanton, J. J., Spergel, D. N., Gunn, J. E., Schmidt, M., and Schneider, D. P. (1997). The Number Density of Low-Surface Brightness Galaxies with 23 < mu_0 < 25 V Mag/arcsec^2. *The Astronomical Journal*, 114:635.

Dale, D. A., Cohen, S. A., Johnson, L. C., Schuster, M. D., Calzetti, D., Engelbracht, C. W., Gil de Paz, A., Kennicutt, R. C., Lee, J. C., Begum, A., Block, M., Dalcanton, J. J., Funes, J. G., Gordon, K. D., Johnson, B. D., Marble, A. R., Sakai, S., Skillman, E. D., van Zee, L., Walter, F., Weisz, D. R., Williams, B., Wu, S.-Y., and Wu, Y. (2009). The Spitzer Local Volume Legacy: Survey Description and Infrared Photometry. *The Astrophysical Journal*, 703:517–556.

Damjanov, I., Abraham, R. G., Glazebrook, K., McCarthy, P. J., Caris, E., Carlberg, R. G., Chen, H.-W., Crampton, D., Green, A. W., Jørgensen, I., et al. (2011). Red nuggets at high redshift: structural evolution of quiescent galaxies over 10 gyr of cosmic history. *The Astrophysical Journal Letters*, 739(2):L44.

Damjanov, I., McCarthy, P. J., Abraham, R. G., Glazebrook, K., Yan, H., Mentuch, E., Le Borgne, D., Savaglio, S., Crampton, D., Murowinski, R., et al. (2009). Red nuggets at z $\sim$ 1.5: Compact passive galaxies and the formation of the kormendy relation. *The Astrophysical Journal*, 695(1):101.

Danovich, M., Dekel, A., Hahn, O., and Teyssier, R. (2012). Coplanar streams, pancakes and angular-momentum exchange in high-z disc galaxies. *Monthly Notices of the Royal Astronomical Society*, 422(2):1732–1749.

Davis, M., Faber, S. M., Newman, J., Phillips, A. C., Ellis, R. S., Steidel, C. C., Conselice, C., Coil, A. L., Finkbeiner, D. P., Koo, D. C., Guhathakurta, P., Weiner, B., Schiavon, R., Willmer, C., Kaiser, N., Luppino, G. A., Wirth, G., Connolly, A., Eisenhardt, P., Cooper, M., and Gerke, B. (2003). Science Objectives and Early Results of the DEEP2 Redshift Survey. In Guhathakurta, P., editor, *Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II*, volume 4834 of *Proceedings of SPIE*, pages 161–172.

de Jong, R. S. (1996). Near-infrared and optical broadband surface photometry of 86 face-on disk dominated galaxies. iv. using color profiles to study stellar and dust content of galaxies. *arXiv preprint astro-ph/9604010*.

de Jong, R. S. and Lacey, C. (2000). The Local Space Density of SB-SDM Galaxies as Function of Their Scale Size, Surface Brightness, and Luminosity. *The Astrophysical Journal*, 545:781–797.

de Laplace, P. (1774). Mémoire sur les suites récurro-récurrentes et sur leurs usages dans la théorie des hasards. *Mém. Acad. Roy. Sci. Paris*, 6:353–371.

de Lapparent, V., Baillard, A., and Bertin, E. (2011). The EFIGI catalogue of 4458 nearby galaxies with morphology. II. Statistical properties along the Hubble sequence. *Astronomy and Astrophysics*, 532:A75.

de Lapparent, V., Galaz, G., Bardelli, S., and Arnouts, S. (2003). The ESO-Sculptor Survey: Luminosity functions of galaxies per spectral type at redshifts 0.1-0.5. *Astronomy and Astrophysics*, 404:831–860.

De Lapparent, V., Geller, M. J., and Huchra, J. P. (1986). A slice of the universe. *The Astrophysical Journal*, 302:L1–L5.

De Ravel, L., Le Fèvre, O., Tresse, L., Bottini, D., Garilli, B., Le Brun, V., Maccagni, D., Scaramella, R., Scodeggio, M., Vettolani, G., et al. (2009). The vimos vlt deep survey-evolution of the major merger rate since z˜ 1 from spectroscopically confirmed galaxy pairs. *Astronomy & Astrophysics*, 498(2):379–397.

de Vaucouleurs (1953). The Distribution of Light in Galaxies. *Astronomical Society of the Pacific Leaflets, Vol. 6, No. 296, p.362*, 6:362.

de Vaucouleurs, G. (1948). Recherches sur les Nebuleuses Extragalactiques. *Annales d'Astrophysique*, 11:247.

de Vaucouleurs, G. (1958). Photoelectric photometry of the andromeda nebula in the ubv system. *The Astrophysical Journal*, 128:465.

De Vaucouleurs, G. (1959). Classification and morphology of external galaxies. In *Astrophysik iv: Sternsysteme/astrophysics iv: Stellar systems*, pages 275–310. Springer.

de Vaucouleurs, G. (1959). Classification and Morphology of External Galaxies. *Handbuch der Physik*, 53:275.

De Vaucouleurs, G. (1959). General physical properties of external galaxies. In *Astrophysik IV: Sternsysteme/Astrophysics IV: Stellar Systems*, pages 311–372. Springer.

De Vaucouleurs, G. (1961). Integrated colors of bright galaxies in the u, b, v system. *The Astrophysical Journal Supplement Series*, 5:233.

de Vaucouleurs, G., de Vaucouleurs, A., Corwin, Jr., H. G., Buta, R. J., Paturel, G., and Fouqué, P. (1991). *Third Reference Catalogue of Bright Galaxies. Volume I: Explanations and references. Volume II: Data for galaxies between 0^h and 12^h. Volume III: Data for galaxies between 12^h and 24^h.*

De Vaucouleurs, G., De Vaucouleurs, A., Corwin Jr, H., Buta, R., Paturel, G., and Fouqué, P. (1991). Third reference catalogue of bright galaxies. volume i: Explanations and references. volume ii: Data for galaxies between 0h and 12h. volume iii: Data for galaxies between 12h and 24h. 1.

Delaye, L. (2013). *Evolution des propriétés structurelles des galaxies de type précoce dans différents environnements*. PhD thesis, Université Paris-Diderot-Paris VII.

Delaye, L., Huertas-Company, M., Mei, S., Lidman, C., Licitra, R., Newman, A., Raichoor, A., Shankar, F., Barrientos, F., Bernardi, M., et al. (2014). Larger sizes of massive quiescent early-type galaxies in clusters than in the field at 0.8< z< 1.5. *Monthly Notices of the Royal Astronomical Society*, 441(1):203–223.

Devour, B. M. and Bell, E. F. (2016). Global dust attenuation in disc galaxies: strong variation with specific star formation and stellar mass, and the importance of sample selection. *Monthly Notices of the Royal Astronomical Society*, 459(2):2054–2077.

Dickinson, M., Giavalisco, M., and GOODS Team (2003). The Great Observatories Origins Deep Survey. In Bender, R. and Renzini, A., editors, *The Mass of Galaxies at Low and High Redshift*, page 324.

Djorgovski, S. and Davis, M. (1987). Fundamental properties of elliptical galaxies. *The Astrophysical Journal*, 313:59–68.

Djorgovski, S. G., Mahabal, A., Drake, A., Graham, M., and Donalek, C. (2013). *Sky Surveys*, page 223.

Dobke, B. M., Johnston, D. E., Massey, R., High, F. W., Ferry, M., Rhodes, J., and Vanderveld, R. A. (2010). Astronomical image simulation for telescope and survey development. *Publications of the Astronomical Society of the Pacific*, 122(894):947.

Dokkum, P. G. v. and Franx, M. (1996). The fundamental plane in cl 0024 at z= 0.4: implications for the evolution of the mass-to-light ratio. *Monthly Notices of the Royal Astronomical Society*, 281(3):985–1000.

Dole, H., Lagache, G., and Puget, J.-L. (2003). Predictions for cosmological infrared surveys from space with the multiband imaging photometer for sirtf. *The Astrophysical Journal*, 585(2):617.

D'Onofrio, M., Rampazzo, R., and Zaggia, S. (2016). From the Realm of the Nebulae to Populations of Galaxies. *From the Realm of the Nebulae to Populations of Galaxies*, 435.

Draine, B. T. (2009). Interstellar Dust Models: Extinction, Absorption and Emission. In Boulanger, F., Joblin, C., Jones, A., and Madden, S., editors, *EAS Publications Series*, volume 35 of *EAS Publications Series*, pages 245–268.

Draine, B. T., Dale, D. A., Bendo, G., Gordon, K. D., Smith, J. D. T., Armus, L., Engelbracht, C. W., Helou, G., Kennicutt, Jr., R. C., Li, A., Roussel, H., Walter, F., Calzetti, D., Moustakas, J., Murphy, E. J., Rieke, G. H., Bot, C., Hollenbach, D. J., Sheth, K., and Teplitz, H. I. (2007). Dust Masses, PAH Abundances, and Starlight Intensities in the SINGS Galaxy Sample. *The Astrophysical Journal*, 663:866–894.

Draine, B. T. and Salpeter, E. E. (1979). On the physics of dust grains in hot gas. *The Astrophysical Journal*, 231:77–94.

Dressler, A. (1980). Galaxy morphology in rich clusters-implications for the formation and evolution of galaxies. *The Astrophysical Journal*, 236:351–365.

Dressler, A. (1987). Lynden-bell, d., burstein, d., davies, rl. *L., Faber SM, Terlevich R. J., Wegner G.*

Driver, S. P. and Phillipps, S. (1996). Is the Luminosity Distribution of Field Galaxies Really Flat? *The Astrophysical Journal*, 469:529.

Driver, S. P., Popescu, C. C., Tuffs, R. J., Liske, J., Graham, A. W., Allen, P. D., and De Propris, R. (2007). The millennium galaxy catalogue: the b-band attenuation of bulge and disc light and the implied cosmic dust and stellar mass densities. *Monthly Notices of the Royal Astronomical Society*, 379(3):1022–1036.

Drory, N., Feulner, G., Bender, R., Botzler, C., Hopp, U., Maraston, C., Mendes de Oliveira, C., and Snigula, J. (2001). The munich near-infrared cluster survey–i. field selection, object extraction and photometry. *Monthly Notices of the Royal Astronomical Society*, 325(2):550–562.

Drovandi, C. C., Pettitt, A. N., and Lee, A. (2015). Bayesian indirect inference using a parametric auxiliary model. *Statist. Sci.*, 30(1):72–95.

Dubois, Y., Pichon, C., Welker, C., Le Borgne, D., Devriendt, J., Laigle, C., Codis, S., Pogosyan, D., Arnouts, S., Benabed, K., et al. (2014). Dancing in the dark: galactic properties trace spin swings along the cosmic web. *Monthly Notices of the Royal Astronomical Society*, pages 1453–1468.

Duchon, C. E. (1979). Lanczos filtering in one and two dimensions. *Journal of Applied Meteorology*, 18(8):1016–1022.

Dwek, E., Arendt, R., Hauser, M., Kelsall, T., Lisse, C., Moseley, S., Silverberg, R., Sodroski, T., and Weiland, J. (1995). Morphology, near-infrared luminosity, and mass of the galactic bulge from cobe dirbe observations. *The Astrophysical Journal*, 445:716–730.

Eddington, A. (1913). On a formula for correcting statistics for the effects of a known error of observation. *Monthly Notices of the Royal Astronomical Society*, 73:359–360.

Eddington, A. (1940). The correction of statistics for accidental error. *Monthly Notices of the Royal Astronomical Society*, 100:354.

Eddington, A. S. (1914). *Stellar Movements and the Structure of the Universe*. Macmillan and Company, limited.

Eggen, O., Lynden-Bell, D., and Sandage, A. (1962). Evidence from the motions of old stars that the galaxy collapsed. *The Astrophysical Journal*, 136:748.

Einstein, A. (1905). On a heuristic point of view concerning the production and transformation of light. *Annalen der Physik*, pages 1–18.

Elmegreen, B. G., Elmegreen, D. M., Fernandez, M. X., and Lemonias, J. J. (2009). Bulge and clump evolution in hubble ultra deep field clump clusters, chains and spiral galaxies. *The Astrophysical Journal*, 692(1):12.

Elmegreen, B. G., Elmegreen, D. M., Vollbach, D. R., Foster, E. R., and Ferguson, T. E. (2005). On the origin of exponential disks at high redshift. *The Astrophysical Journal*, 634(1):101.

Erben, T., Schirmer, M., Dietrich, J. P., Cordes, O., Haberzettl, L., Hetterscheidt, M., Hildebrandt, H., Schmithuesen, O., Schneider, P., Simon, P., Deul, E., Hook, R. N., Kaiser, N., Radovich, M., Benoist, C., Nonino, M., Olsen, L. F., Prandoni, I., Wichmann, R., Zaggia, S., Bomans, D., Dettmar, R. J., and Miralles, J. M. (2005). GaBoDS: The Garching-Bonn Deep Survey. IV. Methods for the image reduction of multi-chip cameras demonstrated on data from the ESO Wide-Field Imager. *Astronomische Nachrichten*, 326:432–464.

Erwin, P., Saglia, R. P., Fabricius, M., Thomas, J., Nowak, N., Rusli, S., Bender, R., Vega Beltrán, J. C., and Beckman, J. E. (2014). Composite bulges: the coexistence of classical bulges and discy pseudo-bulges in s0 and spiral galaxies. *Monthly Notices of the Royal Astronomical Society*, 446(4):4039–4077.

Etherington, I. (1933). Lx. on the definition of distance in general relativity. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 15(100):761–773.

Faber, S. (1973). Variations in spectral-energy distributions and absorption-line strengths among elliptical galaxies. *The Astrophysical Journal*, 179:731–754.

Faber, S. M., Willmer, C. N. A., Wolf, C., Koo, D. C., Weiner, B. J., Newman, J. A., Im, M., Coil, A. L., Conroy, C., Cooper, M. C., Davis, M., Finkbeiner, D. P., Gerke, B. F., Gebhardt, K., Groth, E. J., Guhathakurta, P., Harker, J., Kaiser, N., Kassin, S., Kleinheinrich, M., Konidaris, N. P., Kron, R. G., Lin, L., Luppino, G., Madgwick, D. S., Meisenheimer, K., Noeske, K. G., Phillips, A. C., Sarajedini, V. L., Schiavon, R. P., Simard, L., Szalay, A. S., Vogt, N. P., and Yan, R. (2007). Galaxy Luminosity Functions to z˜1 from DEEP2 and

COMBO-17: Implications for Red Galaxy Formation. *The Astrophysical Journal*, 665:265–294.

Fall, S. M. and Efstathiou, G. (1980). Formation and rotation of disc galaxies with haloes. *Monthly Notices of the Royal Astronomical Society*, 193(2):189–206.

Fan, L., Lapi, A., Bressan, A., Bernardi, M., De Zotti, G., and Danese, L. (2010). Cosmic evolution of size and velocity dispersion for early-type galaxies. *The Astrophysical Journal*, 718(2):1460.

Fan, L., Lapi, A., De Zotti, G., and Danese, L. (2008). The dramatic size evolution of elliptical galaxies and the quasar feedback. *The Astrophysical Journal Letters*, 689(2):L101.

Fan, X., Carilli, C., and Keating, B. (2006). Observational constraints on cosmic reionization. *Annu. Rev. Astron. Astrophys.*, 44:415–462.

Fenech Conti, I., Herbonnet, R., Hoekstra, H., Merten, J., Miller, L., and Viola, M. (2017). Calibration of weak-lensing shear in the kilo-degree survey. *Monthly Notices of the Royal Astronomical Society*, 467(2):1627–1651.

Fink, D. (1997). A compendium of conjugate priors. *See http://www. people. cornell. edu/-pages/df36/CONJINTRnew% 20TEX. pdf*, page 46.

Fioc, M. and Rocca-Volmerange, B. (1999). Pegase. 2, a metallicity-consistent spectral evolution model of galaxies: the documentation and the code. *arXiv preprint astro-ph/9912179*.

Fisher, D. B. and Drory, N. (2016). An observational guide to identifying pseudobulges and classical bulges in disc galaxies. In *Galactic Bulges*, pages 41–75. Springer.

Fitzpatrick, E. L. (2004). Interstellar Extinction in the Milky Way Galaxy. In Witt, A. N., Clayton, G. C., and Draine, B. T., editors, *Astrophysics of Dust*, volume 309 of *Astronomical Society of the Pacific Conference Series*, page 33.

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

Foreman-Mackey, D. (2016). corner.py: Scatterplot matrices in python. *The Journal of Open Source Software*, 24.

Freedman, D. and Diaconis, P. (1981). On the histogram as a density estimator: L 2 theory. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 57(4):453–476.

Freeman, K. C. (1970). On the Disks of Spiral and S0 Galaxies. *The Astrophysical Journal*, 160:811.

Frei, Z. and Gunn, J. E. (1994). Generating colors and K corrections from existing catalog data. *Astronomical Journal*, 108:1476–1485.

Fried, D. L. (1965). Statistics of a geometric representation of wavefront distortion. *JOSA*, 55(11):1427–1435.

Friedman, J. H. (1987). Exploratory Projection Pursuit. *Journal of the American Statistical Association*, 82(397):249–266.

Fukugita, M., Ichikawa, T., Gunn, J., Doi, M., Shimasaku, K., and Schneider, D. (1996). The sloan digital sky survey photometric system. *The Astronomical Journal*, 111:1748.

Fukugita, M. and Peebles, P. (2004). The cosmic energy inventory. *The Astrophysical Journal*, 616(2):643.

Furlanetto, S. R., Oh, S. P., and Briggs, F. H. (2006). Cosmology at low frequencies: The 21cm transition and the high-redshift universe. *Physics Reports*, 433(4):181–301.

Gabasch, A., Bender, R., Seitz, S., Hopp, U., Saglia, R. P., Feulner, G., Snigula, J., Drory, N., Appenzeller, I., Heidt, J., Mehlert, D., Noll, S., Böhm, A., Jäger, K., Ziegler, B., and Fricke, K. J. (2004). The evolution of the luminosity functions in the FORS Deep Field from low to high redshift. I. The blue bands. *Astronomy and Astrophysics*, 421:41–58.

Gabrielli, A., Labini, F. S., Joyce, M., and Pietronero, L. (2006). *Statistical physics for cosmic structures*. Springer Science & Business Media.

Gadotti, D. A. (2009). Structural properties of pseudo-bulges, classical bulges and elliptical galaxies: a sloan digital sky survey perspective. *Monthly Notices of the Royal Astronomical Society*, 393(4):1531–1552.

Gadotti, D. A. (2012). Galaxy bulges and elliptical galaxies-lecture notes. *arXiv preprint arXiv:1208.2295*.

Gadotti, D. A., Baes, M., and Falony, S. (2010). Radiative transfer in disc galaxies–iv. the effects of dust attenuation on bulge and disc structural parameters. *Monthly Notices of the Royal Astronomical Society*, 403(4):2053–2062.

Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J., Mignard, F., Drimmel, R., Babusiaux, C., Bailer-Jones, C. A. L., Bastian, U., and et al. (2016). Gaia Data Release 1. Summary of the astrometric, photometric, and survey properties. *Astronomy and Astrophysics*, 595:A2.

Gallant, A. R. and McCulloch, R. E. (2009). On the determination of general scientific models with application to asset pricing. *Journal of the American Statistical Association*, 104(485):117–131.

Gamow, G. (1948). The origin of elements and the separation of galaxies. *Physical Review*, 74(4):505.

Gautier, T., Boulanger, F., Perault, M., and Puget, J. (1992). A calculation of confusion noise due to infrared cirrus. *The Astronomical Journal*, 103:1313–1324.

Gelman, A., Roberts, G., and Gilks, W. (1994). Efficient metropolis jumping rules". *Bayesian Statistics*, 5.

Gelman, A. and Rubin, D. B. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7(4):457–472.

Geweke, J. et al. (1992). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments*, volume 196. Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA.

Giovanelli, R., Haynes, M. P., Salzer, J. J., Wegner, G., da Costa, L. N., and Freudling, W. (1995). Dependence on luminosity of photometric properties of disk galaxies: surface brightness, size, and internal extinction. *The Astronomical Journal*, 110:1059.

Gordon, K. D., Clayton, G. C., Misselt, K. A., Landolt, A. U., and Wolff, M. J. (2003). A Quantitative Comparison of the Small Magellanic Cloud, Large Magellanic Cloud, and Milky Way Ultraviolet to Near-Infrared Extinction Curves. *The Astrophysical Journal*, 594:279–293.

Goto, T., Yamauchi, C., Fujita, Y., Okamura, S., Sekiguchi, M., Smail, I., Bernardi, M., and Gomez, P. L. (2003). The morphology—density relation in the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 346(2):601–614.

Gould, R. J. and Salpeter, E. E. (1963). The Interstellar Abundance of the Hydrogen Molecule. I. Basic Processes. *The Astrophysical Journal*, 138:393.

Graham, A. W. (2001). Ak-band central disc surface brightness correlation with scalelength for early-type disc galaxies, and the inclination correction. *Monthly Notices of the Royal Astronomical Society*, 326(2):543–552.

Graham, A. W. (2013). *Elliptical and Disk Galaxy Structure and Modern Scaling Laws*, page 91.

Graham, A. W. and Worley, C. C. (2008). Inclination-and dust-corrected galaxy parameters: bulge-to-disc ratios and size–luminosity relations. *Monthly Notices of the Royal Astronomical Society*, 388(4):1708–1728.

Grogin, N. A., Kocevski, D. D., Faber, S., Ferguson, H. C., Koekemoer, A. M., Riess, A. G., Acquaviva, V., Alexander, D. M., Almaini, O., Ashby, M. L., et al. (2011). Candels: the cosmic assembly near-infrared deep extragalactic legacy survey. *The Astrophysical Journal Supplement Series*, 197(2):35.

Guiderdoni, B., Bouchet, F. R., Puget, J.-L., Lagache, G., and Hivon, E. (1997). The optically dark side of galaxy formation. *Nature*, 390:257.

Gültekin, K., Richstone, D. O., Gebhardt, K., Lauer, T. R., Tremaine, S., Aller, M. C., Bender, R., Dressler, A., Faber, S., Filippenko, A. V., et al. (2009). The m-σ and ml relations in galactic bulges, and determinations of their intrinsic scatter. *The Astrophysical Journal*, 698(1):198.

Gurvich, A. and Mandelbaum, R. (2016). The impact of correlated noise on galaxy shape estimation for weak lensing. *Monthly Notices of the Royal Astronomical Society*, 457:3522–3534.

Gutmann, M. U. and Corander, J. (2016). Bayesian optimization for likelihood-free inference of simulator-based statistical models. *Journal of Machine Learning Research*, 17(125):1–47.

Haario, H., Saksman, E., and Tamminen, J. (1999). Adaptive proposal distribution for random walk metropolis algorithm. *Computational Statistics*, 14(3):375–395.

Haario, H., Saksman, E., and Tamminen, J. (2001). An adaptive metropolis algorithm. *Bernoulli*, 7(2):223–242.

Hahn, C., Vakili, M., Walsh, K., Hearin, A. P., Hogg, D. W., and Campbell, D. (2016). Approximate Bayesian Computation in Large Scale Structure: constraining the galaxy-halo connection. *ArXiv e-prints*.

Hardy, J. W. (1998). *Adaptive optics for astronomical telescopes*, volume 16. Oxford University Press on Demand.

Hartwig, T., Latif, M. A., Magg, M., Bromm, V., Klessen, R. S., Glover, S. C., Whalen, D. J., Pellegrini, E. W., and Volonteri, M. (2016). Exploring the nature of the lyman-α emitter cr7. *Monthly Notices of the Royal Astronomical Society*, 462(2):2184–2202.

Hastings (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(1):97–109.

Herbel, J., Kacprzak, T., Amara, A., Refregier, A., Bruderer, C., and Nicola, A. (2017). The redshift distribution of cosmological samples: a forward modeling approach. *arXiv preprint arXiv:1705.05386*.

Hildebrandt, H., Viola, M., Heymans, C., Joudaki, S., Kuijken, K., Blake, C., Erben, T., Joachimi, B., Klaes, D., Miller, L., et al. (2017). Kids-450: Cosmological parameter constraints from tomographic weak gravitational lensing. *Monthly Notices of the Royal Astronomical Society*.

Hogg, D. W. (1999). Distance measures in cosmology. *ArXiv Astrophysics e-prints*.

Hogg, D. W. (2001). Confusion errors in astrometry and counterpart association. *The Astronomical Journal*, 121(2):1207.

Hogg, D. W. (2008). Data analysis recipes: Choosing the binning for a histogram. *ArXiv e-prints*.

Hogg, D. W., Baldry, I. K., Blanton, M. R., and Eisenstein, D. J. (2002). The K correction. *ArXiv Astrophysics e-prints*.

Holmberg, E. (1941). On the clustering tendencies among the nebulae. ii. a study of encounters between laboratory models of stellar systems by a new integration procedure. *The Astrophysical Journal*, 94:385.

Holmberg, E. (1958). A photographic photometry of extragalactic nebulae. *Meddelanden fran Lunds Astronomiska Observatorium Serie II*, 136:1.

Holmberg, E., Sandage, A., Sandage, M., and Kristian, J. (1975). Galaxies and the universe. *Stars and Stellar Systems*, 9:123.

Hopkins, P. F., Cox, T. J., Kereš, D., and Hernquist, L. (2008). A cosmological framework for the co-evolution of quasars, supermassive black holes, and elliptical galaxies. ii. formation of red ellipticals. *The Astrophysical Journal Supplement Series*, 175(2):390.

Hoskin, T. (2012). Parametric and nonparametric: Demystifying the terms. In *Mayo Clinic*.

Hubble, E. (1926). No. 324. extra-galactic nebulae. *Contributions from the Mount Wilson Observatory/Carnegie Institution of Washington*, 324:1–49.

Hubble, E. (1929). A relation between distance and radial velocity among extra-galactic nebulae. *Proceedings of the National Academy of Sciences*, 15(3):168–173.

Hubble, E. P. (1936). *The realm of the nebulae*, volume 25. Yale University Press.

Huchra, J., Davis, M., Latham, D., and Tonry, J. (1983). A survey of galaxy redshifts. iv-the data. *The Astrophysical Journal Supplement Series*, 52:89–119.

Huchra, J. P., Macri, L. M., Masters, K. L., Jarrett, T. H., Berlind, P., Calkins, M., Crook, A. C., Cutri, R., Erdoğdu, P., Falco, E., et al. (2012). The 2mass redshift survey—description and data release. *The Astrophysical Journal Supplement Series*, 199(2):26.

Hudelot, P., Goranova, Y., Yannick Mellier, Y., McCracken, H., Magnard, F., Monnerville, M., Sémah, G., Cuillandre, J.-C., Withington, K., Regnault, N., Betoule, M., Schultheis, M., and Aussel, H. (2012). *T0007 : The Final CFHTLS Release*.

Huertas-Company, M., Shankar, F., Mei, S., Bernardi, M., Aguerri, J. A. L., Meert, A., and Vikram, V. (2013). No Evidence for a Dependence of the Mass-Size Relation of Early-type Galaxies on Environment in the Local Universe. *The Astrophysical Journal*, 779:29.

Humason, M. L., Mayall, N. U., and Sandage, A. R. (1956). Redshifts and magnitudes of extragalactic nebulae. *The Astronomical Journal*, 61:97–162.

Hyndman, R. J. (1996). Computing and graphing highest density regions. *The American Statistician*, 50(2):120–126.

Hyvärinen, A., Hurri, J., and Hoyer, P. O. (2009). Principal components and whitening. In *Natural Image Statistics*, pages 93–130.

Ilbert, O., Capak, P., Salvato, M., Aussel, H., McCracken, H., Sanders, D., Scoville, N., Kartaltepe, J., Arnouts, S., Le Floc'h, E., et al. (2008). Cosmos photometric redshifts with 30-bands for 2-deg2. *The Astrophysical Journal*, 690(2):1236.

Illingworth, G., Magee, D., Oesch, P., Bouwens, R., Labbé, I., Stiavelli, M., Van Dokkum, P., Franx, M., Trenti, M., Carollo, C., et al. (2013). The hst extreme deep field (xdf): Combining all acs and wfc3/ir data on the hudf region into the deepest field ever. *The Astrophysical Journal Supplement Series*, 209(1):6.

Ishida, E. E. O., Vitenti, S. D. P., Penna-Lima, M., Cisewski, J., de Souza, R. S., Trindade, A. M. M., Cameron, E., and Busti, V. C. (2015). COSMOABC: Likelihood-free inference via Population Monte Carlo Approximate Bayesian Computation. *Astronomy and Computing*, 13:1–11.

Ivezic, Z., Tyson, J., Abel, B., Acosta, E., Allsman, R., AlSayyad, Y., Anderson, S., Andrew, J., Angel, R., Angeli, G., et al. (2008). Lsst: from science drivers to reference design and anticipated data products. *arXiv preprint arXiv:0805.2366*.

Ivison, R., Greve, T., Dunlop, J., Peacock, J., Egami, E., Smail, I., Ibar, E., Van Kampen, E., Aretxaga, I., Babbedge, T., et al. (2007). The scuba half degree extragalactic survey–iii. identification of radio and mid-infrared counterparts to submillimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 380(1):199–228.

Jablonka, P., Martin, P., and Arimoto, N. (1996). The luminosity-metallicity relation for bulges of spiral galaxies. *arXiv preprint astro-ph/9604132*.

Jarrett, T., Chester, T., Cutri, R., Schneider, S., Skrutskie, M., and Huchra, J. (2000). 2mass extended source catalog: overview and algorithms. *The Astronomical Journal*, 119(5):2498.

Jaynes, E. (1985). Highly informative priors.

Jaynes, E. T. (1968). Prior probabilities. *IEEE Transactions on systems science and cybernetics*, 4(3):227–241.

Jaynes, E. T. and Kempthorne, O. (1976). Confidence intervals vs bayesian intervals. In *Foundations of probability theory, statistical inference, and statistical theories of science*, pages 175–257. Springer.

Jeffreys, H. (1938). The correction of frequencies for a known standard error of observation. *Monthly Notices of the Royal Astronomical Society*, 98:190.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. In *Proceedings of the Royal Society of London a: mathematical, physical and engineering sciences*, volume 186, pages 453–461. The Royal Society.

Jeffries, H. (1961). Theory of probability.

Jennings, E. and Madigan, M. (2016). astroABC: An Approximate Bayesian Computation Sequential Monte Carlo sampler for cosmological parameter estimation. *ArXiv e-prints*.

Jeong, W.-S., Mok Lee, H., Pak, S., Nakagawa, T., Minn Kwon, S., Pearson, C. P., and White, G. J. (2005). Far-infrared detection limits–i. sky confusion due to galactic cirrus. *Monthly Notices of the Royal Astronomical Society*, 357(2):535–547.

Jester, S., Schneider, D. P., Richards, G. T., Green, R. F., Schmidt, M., Hall, P. B., Strauss, M. A., Vanden Berk, D. E., Stoughton, C., Gunn, J. E., Brinkmann, J., Kent, S. M., Smith, J. A., Tucker, D. L., and Yanny, B. (2005). The Sloan Digital Sky Survey View of the Palomar-Green Bright Quasar Survey. *Astronomical Journal*, 130:873–895.

Jonsson, P. (2006). Sunrise: polychromatic dust radiative transfer in arbitrary geometries. *Monthly Notices of the Royal Astronomical Society*, 372(1):2–20.

Jonsson, P., Groves, B., and Cox, T. J. (2009). High-Resolution Panchromatic Spectral Models of Galaxies including Photoionisation and Dust. *ArXiv e-prints*.

Joye, W. (2013). Saoimage ds9. vers. 6.1. 2. *Computer software. Future Directions for Astronomical Image Display*.

Justel, A., Peña, D., and Zamar, R. (1997). A multivariate Kolmogorov-Smirnov test of goodness of fit. *Statistics & Probability Letters*, 35(3):251–259.

Kacprzak, T., Herbel, J., Amara, A., and Réfrégier, A. (2017). Accelerating approximate bayesian computation with quantile regression: Application to cosmological redshift distributions. *arXiv preprint arXiv:1707.07498*.

Kangasrääsiö, A., Lintusaari, J., Skytén, K., Järvenpää, M., Vuollekoski, H., Gutmann, M., Vehtari, A., Corander, J., and Kaski, S. (2016). Elfi: Engine for likelihood-free inference. In *NIPS 2016 Workshop on Advances in Approximate Bayesian Inference*.

Kannawadi, A., Mandelbaum, R., and Lackner, C. (2015). The impact of cosmic variance on simulating weak lensing surveys. *Monthly Notices of the Royal Astronomical Society*, 449:3597–3612.

Kant, I. (1755). *Allgemeine Naturgeschichte und Theorie des Himmels*.

Kass, R. E. and Wasserman, L. (1996). The selection of prior distributions by formal rules. *Journal of the American Statistical Association*, 91(435):1343–1370.

Kauffmann, G., Guiderdoni, B., and White, S. (1994). Faint galaxy counts in a hierarchical universe. *Monthly Notices of the Royal Astronomical Society*, 267(4):981–999.

Kauffmann, G., Heckman, T. M., White, S. D., Charlot, S., Tremonti, C., Peng, E. W., Seibert, M., Brinkmann, J., Nichol, R. C., SubbaRao, M., et al. (2003). The dependence of star formation history and internal structure on stellar mass for 105 low-redshift galaxies. *Monthly Notices of the Royal Astronomical Society*, 341(1):54–69.

Keel, W. C. (2007). *The road to galaxy formation*. Springer Science & Business Media.

Kelkar, K., Aragón-Salamanca, A., Gray, M. E., Maltby, D., Vulcani, B., De Lucia, G., Poggianti, B. M., and Zaritsky, D. (2015). Galaxy sizes as a function of environment at intermediate redshift from the eso distant cluster survey. *Monthly Notices of the Royal Astronomical Society*, 450(2):1246–1255.

Kereš, D., Katz, N., Fardal, M., Davé, R., and Weinberg, D. H. (2009). Galaxies in a simulated λcdm universe–i. cold mode and hot cores. *Monthly Notices of the Royal Astronomical Society*, 395(1):160–179.

Kessy, A., Lewin, A., and Strimmer, K. (2015). Optimal whitening and decorrelation. *ArXiv e-prints*.

Khochfar, S. and Burkert, A. (2006). Orbital parameters of merging dark matter halos. *Astronomy & Astrophysics*, 445(2):403–412.

Khochfar, S. and Silk, J. (2006). A simple model for the size evolution of elliptical galaxies. *The Astrophysical Journal Letters*, 648(1):L21.

King, I. R. (1971). The profile of a star image. *Publications of the Astronomical Society of the Pacific*, 83(492):199.

Kinney, A. L., Calzetti, D., Bohlin, R. C., McQuade, K., Storchi-Bergmann, T., and Schmitt, H. R. (1996). Template ultraviolet to near-infrared spectra of star-forming galaxies and their application to k-corrections. *The Astrophysical Journal*, 467:38.

Kinney, W. H. (2003). Cosmology, inflation and the physics of nothing. *Techniques and Concepts of High-Energy Physics XII*, pages 189–243.

Kirkpatrick, C. D. Gelatt, M. P. V. (1983). Optimization by simulated annealing. *Science*, 220(4598):671–680.

Klypin, A. A., Trujillo-Gomez, S., and Primack, J. (2011). Dark matter halos in the standard cosmological model: Results from the bolshoi simulation. *The Astrophysical Journal*, 740(2):102.

Knuth, K. H. (2006). Optimal Data-Based Binning for Histograms. *ArXiv Physics e-prints*.

Kolb, E. W., Salopek, D. S., and Turner, M. S. (1990). Origin of density fluctuations in extended inflation. *Physical Review D*, 42(12):3925.

Kormendy, J. (1993). Kinematics of extragalactic bulges: evidence that some bulges are really disks. In *Galactic Bulges*, volume 153, page 209.

Kormendy, J. (2013). Secular evolution in disk galaxies. *Secular evolution of galaxies. Cambridge University Press, Cambridge, UK*, pages 1–159.

Kormendy, J. (2016). Elliptical galaxies and bulges of disc galaxies: Summary of progress and outstanding issues. In *Galactic Bulges*, pages 431–477. Springer.

Kormendy, J. and Bender, R. (2011). A revised parallel-sequence morphological classification of galaxies: structure and formation of s0 and spheroidal galaxies. *The Astrophysical Journal Supplement Series*, 198(1):2.

Kormendy, J. and Kennicutt Jr, R. C. (2004). Secular evolution and the formation of pseudobulges in disk galaxies. *Annu. Rev. Astron. Astrophys.*, 42:603–683.

Kriek, M. and Conroy, C. (2013). The Dust Attenuation Law in Distant Galaxies: Evidence for Variation with Spectral Type. *The Astrophysical Journall*, 775:L16.

Kron, R. G. (1980). Photometry of a complete sample of faint galaxies. *Astrophysical Journal Supplement Series*, 43:305–325.

Kullback, S. and Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79–86.

Kurinsky, N. and Sajina, A. (2014). Robust Constraint of Luminosity Function Evolution through MCMC Sampling. In Heavens, A., Starck, J.-L., and Krone-Martins, A., editors, *Statistical Challenges in 21st Century Cosmology*, volume 306 of *IAU Symposium*, pages 295–297.

Lacey, C. G., Baugh, C. M., Frenk, C. S., Benson, A. J., Bower, R. G., Cole, S., Gonzalez-Perez, V., Helly, J. C., Lagos, C. D., and Mitchell, P. D. (2016). A unified multiwavelength model of galaxy formation. *Monthly Notices of the Royal Astronomical Society*, 462(4):3854–3911.

Lang, D., Hogg, D. W., and Mykytyn, D. (2016). The Tractor: Probabilistic astronomical source detection and measurement. Astrophysics Source Code Library.

Lani, C., Almaini, O., Hartley, W. G., Mortlock, A., Häußler, B., Chuter, R. W., Simpson, C., van der Wel, A., Grützbauch, R., Conselice, C. J., et al. (2013). Evidence for a correlation between the sizes of quiescent galaxies and local environment to z 2. *Monthly Notices of the Royal Astronomical Society*, 435(1):207–221.

Laplace, P. (1814). *Théorie Analytique des Probabilités*.

Laureijs, R., Amiaux, J., Arduini, S., Augueres, J.-L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., et al. (2011). Euclid definition study report. *arXiv preprint arXiv:1110.3193*.

Le Borgne, D. and Rocca-Volmerange, B. (2002). Photometric redshifts from evolutionary synthesis with PÉGASE: The code Z-PEG and the z=0 age constraint. *Astronomy and Astrophysics*, 386:446–455.

Le Fèvre, O., Cassata, P., Cucciati, O., Garilli, B., Ilbert, O., Le Brun, V., Maccagni, D., Moreau, C., Scodeggio, M., Tresse, L., et al. (2013). The vimos vlt deep survey final data release: a spectroscopic sample of 35 016 galaxies and agn out to z˜ 6.7 selected with 17.5 ≤ iab ≤ 24.75. *Astronomy & Astrophysics*, 559:A14.

Le Fèvre, O., Vettolani, G., Garilli, B., Tresse, L., Bottini, D., Le Brun, V., Maccagni, D., Picat, J., Scaramella, R., Scodeggio, M., et al. (2005). The vimos vlt deep survey-first epoch vvds-deep survey: 11 564 spectra with $17.5 leq$ i ≤24, and the redshift distribution over 0 ≤ z ≤5. *Astronomy & Astrophysics*, 439(3):845–862.

Léna, P., Rouan, D., Lebrun, F., Mignard, F., and Pelat, D. (2012). *Observational astrophysics*. Springer Science & Business Media.

Lilly, S. J., Tresse, L., Hammer, F., Crampton, D., and Le Fevre, O. (1995). The Canada-France Redshift Survey. VI. Evolution of the Galaxy Luminosity Function to Z approximately 1. *The Astrophysical Journal*, 455:108.

Lin, J. (1991). Divergence Measures Based on the Shannon Entropy. *IEEE TRANSACTIONS ON INFORMATION THEORY*, 37(I).

Lin, L., Koo, D. C., Willmer, C. N., Patton, D. R., Conselice, C. J., Yan, R., Coil, A. L., Cooper, M. C., Davis, M., Faber, S., et al. (2004). The deep2 galaxy redshift survey: Evolution of close galaxy pairs and major-merger rates up to z˜ 1.2. *The Astrophysical Journal Letters*, 617(1):L9.

Lin, L., Patton, D. R., Koo, D. C., Casteels, K., Conselice, C. J., Faber, S., Lotz, J., Willmer, C. N., Hsieh, B., Chiueh, T., et al. (2008). The redshift evolution of wet, dry, and mixed galaxy mergers from close galaxy pairs in the deep2 galaxy redshift survey. *The Astrophysical Journal*, 681(1):232.

Lintott, C., Schawinski, K., Bamford, S., Slosar, A., Land, K., Thomas, D., Edmondson, E., Masters, K., Nichol, R. C., Raddick, M. J., et al. (2010). Galaxy zoo 1: data release of morphological classifications for nearly 900 000 galaxies. *Monthly Notices of the Royal Astronomical Society*, 410(1):166–178.

Loeb, A. (2010). *How did the first stars and galaxies form?* Princeton University Press.

Loeb, A. and Barkana, R. (2001). The reionization of the universe by the first stars and quasars. *Annual review of astronomy and astrophysics*, 39(1):19–66.

Loeb, A. and Zaldarriaga, M. (2004). Measuring the small-scale power spectrum of cosmic density fluctuations through 21 cm tomography prior to the epoch of structure formation. *Physical Review Letters*, 92(21):211301.

Longhetti, M., Saracco, P., Severgnini, P., Della Ceca, R., Mannucci, F., Bender, R., Drory, N., Feulner, G., and Hopp, U. (2006). The kormendy relation of massive elliptical galaxies at z 1.5: evidence for size evolution. *Monthly Notices of the Royal Astronomical Society*, 374(2):614–626.

Lotz, J. M., Primack, J., and Madau, P. (2004). A new nonparametric approach to galaxy morphological classification. *The Astronomical Journal*, 128(1):163.

Low, F., Young, E., Beintema, D., Gautier, T., Beichman, C., Aumann, H., Gillett, F., Neugebauer, G., Boggess, N., and Emerson, J. (1984). Infrared cirrus-new components of the extended infrared emission. *The Astrophysical Journal*, 278:L19–L22.

Lu, Y., Kereš, D., Katz, N., Mo, H., Fardal, M., and Weinberg, M. D. (2011). On the algorithms of radiative cooling in semi-analytic models. *Monthly Notices of the Royal Astronomical Society*, 416(1):660–679.

Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. (2012). *The BUGS book: A practical introduction to Bayesian analysis*. CRC press.

Lynden-Bell, D. (1992). Eddington-malmquist bias, streaming motions, and the distribution of galaxies. In *Statistical challenges in modern astronomy*, pages 201–216. Springer.

MacDonald, C. J. and Bernstein, G. (2010). Photometric Redshift Biases from Galaxy Evolution. *Publications of the Astronomical Society of the Pacific*, 122:485–489.

Maceachern, S. N. and Berliner, L. M. (1994). Subsampling the gibbs sampler. *The American Statistician*, 48(3):188–190.

Madau, P. (1995). Radiative transfer in a clumpy universe: The colors of high-redshift galaxies. *The Astrophysical Journal*, 441:18.

Madau, P. and Dickinson, M. (2014). Cosmic star-formation history. *Annual Review of Astronomy and Astrophysics*, 52:415–486.

Madau, P., Ferguson, H. C., Dickinson, M. E., Giavalisco, M., Steidel, C. C., and Fruchter, A. (1996). High-redshift galaxies in the Hubble Deep Field: colour selection and star formation history to z˜4. *Monthly Notices of the Royal Astronomical Society*, 283:1388–1404.

Madau, P. and Rees, M. J. (2001). Massive black holes as population iii remnants. *The Astrophysical Journal Letters*, 551(1):L27.

Madgwick, D. S., Hawkins, E., Lahav, O., Maddox, S., Norberg, P., Peacock, J. A., Baldry, I. K., Baugh, C. M., Bland-Hawthorn, J., Bridges, T., Cannon, R., Cole, S., Colless, M., Collins, C., Couch, W., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Frenk, C. S., Glazebrook, K., Jackson, C., Lewis, I., Lumsden, S., Peterson, B. A., Sutherland, W., and Taylor, K. (2003). The 2dF Galaxy Redshift Survey: galaxy clustering per spectral type. *Monthly Notices of the Royal Astronomical Society*, 344:847–856.

Mahalanobis, P. C. (1936). On the generalised distance in statistics. In *Proceedings National Institute of Science, India*, volume 2, pages 49–55.

Maller, A. H., Berlind, A. A., Blanton, M. R., and Hogg, D. W. (2009). The intrinsic properties of sdss galaxies. *The Astrophysical Journal*, 691(1):394.

Malmquist, G. K. (1920). A study of the stars of spectral type a. *Meddelanden fran Lunds Astronomiska Observatorium Serie II*, 22:3–69.

Malmquist, K. G. (1922). On some relations in stellar statistics. *Meddelanden fran Lunds Astronomiska Observatorium Serie I*, 100:1–52.

Maltby, D. T., Aragón-Salamanca, A., Gray, M. E., Barden, M., Häußler, B., Wolf, C., Peng, C. Y., Jahnke, K., McIntosh, D. H., Böhm, A., et al. (2010). The environmental dependence of the stellar-mass–size relation in stages galaxies. *Monthly Notices of the Royal Astronomical Society*, 402(1):282–294.

Mancini, C., Matute, I., Cimatti, A., Daddi, E., Dickinson, M., Rodighiero, G., Bolzonella, M., and Pozzetti, L. (2009). Searching for massive galaxies at z ≥ 3.5 in goods-north. *Astronomy & Astrophysics*, 500(2):705–723.

Mandelbaum, R., Hirata, C. M., Leauthaud, A., Massey, R. J., and Rhodes, J. (2012). Precision simulation of ground-based lensing data using observations from space. *Monthly Notices of the Royal Astronomical Society*, 420(2):1518–1540.

Marin, J.-M., Pudlo, P., Robert, C. P., and Ryder, R. (2011). Approximate Bayesian Computational methods. *ArXiv e-prints*.

Marjoram, P., Molitor, J., Plagnol, V., and Tavare, S. (2003). Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences of the United States of America*, 100(26):15324–8.

Markov, A. (1913). Example of statistical research on text of "eugene onegin", illustrating interconnection of trials in chain. *Izvestiya Akademii Nauk SPb*, 6:153–162.

Markov, A. (1954). Theory of algorithms (russian), trudy mat. *Istituta imeni VA Steklova*, 42.

Martinez-Manso, J., Guzman, R., Barro, G., Cenarro, J., Perez-Gonzalez, P., Sanchez-Blazquez, P., Trujillo, I., Balcells, M., Cardiel, N., Gallego, J., et al. (2011). Velocity dispersions and stellar populations of the most compact and massive early-type galaxies at redshift ~ 1. *The Astrophysical Journal Letters*, 738(2):L22.

Marzke, R. O. (1998). The Galaxy Luminosity Function at Redshift Zero: Constraints on Galaxy Formation. In *The Evolving Universe*, pages 23–39.

Masters, K. L., Giovanelli, R., and Haynes, M. P. (2003). Internal extinction in spiral galaxies in the near-infrared. *The Astronomical Journal*, 126(1):158.

Masters, K. L., Nichol, R., Bamford, S., Mosleh, M., Lintott, C. J., Andreescu, D., Edmondson, E. M., Keel, W. C., Murray, P., Raddick, M. J., Schawinski, K., Slosar, A., Szalay, A. S., Thomas, D., and Vandenberg, J. (2010). Galaxy zoo: dust in spiral galaxies. *Monthly Notices of the Royal Astronomical Society*, 404(2):792–810.

Mathis, J. S., Rumpl, W., and Nordsieck, K. H. (1977). The size distribution of interstellar grains. *The Astrophysical Journal*, 217:425–433.

McBride, C., Berlind, A., Scoccimarro, R., Wechsler, R., Busha, M., Gardner, J., and van den Bosch, F. (2009). Lasdamas mock galaxy catalogs for sdss. In *Bulletin of the American Astronomical Society*, volume 41, page 253.

Mehrotra, K., Mohan, C. K., and Ranka, S. (1997). *Elements of artificial neural networks*. MIT Press.

Méndez-Abreu, J., Debattista, V. P., Corsini, E., and Aguerri, J. (2014). Secular-and merger-built bulges in barred galaxies. *Astronomy & Astrophysics*, 572:A25.

Meneghetti, M., Melchior, P., Grazian, A., De Lucia, G., Dolag, K., Bartelmann, M., Heymans, C., Moscardini, L., and Radovich, M. (2008). Realistic simulations of gravitational lensing by galaxy clusters: extracting arc parameters from mock dune images. *Astronomy & Astrophysics*, 482(2):403–418.

Messier, C. (1781). Catalogue des nebuleuses & des amas d'etoiles (catalog of nebulae and star clusters). *Connoissance des Temps for 1784, p. 227-267*, pages 227–267.

Metropolis, N. (1987). The beginning of the monte carlo method. special issue 15. *Los Alamos Science*.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, 21(6):1087.

Metropolis, N. and Ulam, S. (1949). The monte carlo method. *Journal of the American Statistical Association*, 44(247):335–341.

Miralda-Escudé, J. (2003). The dark age of the universe. *Science*, 300(5627):1904–1909.

Miyoshi, K. and Kihara, T. (1975). Development of the correlation of galaxies in an expanding universe. *Publications of the Astronomical Society of Japan*, 27:333–346.

Mo, H., Van den Bosch, F., and White, S. (2010). *Galaxy formation and evolution*. Cambridge University Press.

Mo, H. J., Mao, S., and White, S. D. M. (1998). The formation of galactic discs. *Monthly Notices of the Royal Astronomical Society*, 295(2):319–336.

Möllenhoff, C., Popescu, C., and Tuffs, R. (2006). Modelling the spectral energy distribution of galaxies-iv. correcting apparent disk scalelengths and central surface brightnesses for the effect of dust at optical and near-infrared wavelengths. *Astronomy & Astrophysics*, 456(3):941–952.

Monet, D. G., Levine, S. E., Canzian, B., Ables, H. D., Bird, A. R., Dahn, C. C., Guetter, H. H., Harris, H. C., Henden, A. A., Leggett, S. K., et al. (2003). The usno-b catalog. *The Astronomical Journal*, 125(2):984.

Moore, G. E. (2006). Cramming more components onto integrated circuits, reprinted from electronics, volume 38, number 8, april 19, 1965, pp. 114 ff. *IEEE Solid-State Circuits Society Newsletter*, 20(3):33–35.

Morgan, W. (1958). A preliminary classification of the forms of galaxies according to their stellar population. *Publications of the Astronomical Society of the Pacific*, 70(415):364–391.

Morgan, W. W. and Mayall, N. (1957). A spectral classification of galaxies. *Publications of the Astronomical Society of the Pacific*, 69(409):291–303.

Morishita, T., Ichikawa, T., and Kajisawa, M. (2014). The evolution of galaxy size and morphology at z 0.5-3.0 in the goods-n region with hubble space telescope/wfc3 data. *The Astrophysical Journal*, 785(1):18.

Moster, B. P., Somerville, R. S., Newman, J. A., and Rix, H.-W. (2011). A cosmic variance cookbook. *The Astrophysical Journal*, 731(2):113.

Nagayama, T., Woudt, P., Nagashima, C., Nakajima, Y., Kato, D., Kurita, M., Nagata, T., Nakaya, H., Tamura, M., Sugitani, K., et al. (2004). A deep near-infrared survey around the giant radio galaxy pks 1343–601. *Monthly Notices of the Royal Astronomical Society*, 354(4):980–990.

Nair, P. B. and Abraham, R. G. (2010). A catalog of detailed visual morphological classifications for 14,034 galaxies in the sloan digital sky survey. *The Astrophysical Journal Supplement Series*, 186(2):427.

Nair, P. B., van den Bergh, S., and Abraham, R. G. (2010). The environmental dependence of the luminosity–size relation for galaxies. *The Astrophysical Journal*, 715(1):606.

Negrello, M., Magliocchetti, M., Moscardini, L., De Zotti, G., Granato, G. L., and Silva, L. (2004). Confusion noise at far-infrared to millimetre wavelengths. *Monthly Notices of the Royal Astronomical Society*, 352(2):493–500.

Newman, A. B., Ellis, R. S., Bundy, K., and Treu, T. (2012). Can minor merging account for the size growth of quiescent galaxies? new results from the candels survey. *The Astrophysical Journal*, 746(2):162.

Newman, J. A., Cooper, M. C., Davis, M., Faber, S., Coil, A. L., Guhathakurta, P., Koo, D. C., Phillips, A. C., Conroy, C., Dutton, A. A., et al. (2013). The deep2 galaxy redshift survey: Design, observations, data reduction, and redshifts. *The Astrophysical Journal Supplement Series*, 208(1):5.

Nguyen, H., Schulz, B., Levenson, L., Amblard, A., Arumugam, V., Aussel, H., Babbedge, T., Blain, A., Bock, J., Boselli, A., et al. (2010). Hermes: The spire confusion limit. *Astronomy & Astrophysics*, 518:L5.

Nipoti, C., Treu, T., Auger, M., and Bolton, A. (2009). Can dry merging explain the size evolution of early-type galaxies? *The Astrophysical Journal Letters*, 706(1):L86.

Noll, S., Burgarella, D., Giovannoli, E., Buat, V., Marcillac, D., and Muñoz-Mateos, J. C. (2009). Analysis of galaxy spectral energy distributions from far-UV to far-IR with CIGALE: studying a SINGS test sample. *Astronomy and Astrophysics*, 507:1793–1813.

Noll, S., Pierini, D., Pannella, M., and Savaglio, S. (2007). Presence of dust with a UV bump in massive, star-forming galaxies at $1 < z < 2.5$. *Astronomy and Astrophysics*, 472:455–469.

Norberg, P., Cole, S., Baugh, C. M., Frenk, C. S., Baldry, I., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Cross, N. J. G., Dalton, G., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Glazebrook, K., Jackson, C., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Madgwick, D., Peacock, J. A., Peterson, B. A., Sutherland, W., Taylor, K., and 2DFGRS Team (2002). The 2dF Galaxy Redshift Survey: the b_J-band galaxy luminosity function and survey selection function. *Monthly Notices of the Royal Astronomical Society*, 336:907–931.

Nowak, N., Thomas, J., Erwin, P., Saglia, R. P., Bender, R., and Davies, R. I. (2010). Do black hole masses scale with classical bulge luminosities only? the case of the two composite pseudo-bulge galaxies ngc 3368 and ngc 3489. *Monthly Notices of the Royal Astronomical Society*, 403(2):646–672.

Okabe, A., Boots, B., Sugihara, K., and Chiu, S. N. (2000). *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. New York: Wiley.

Oke, J. and Gunn, J. (1983). Secondary standard stars for absolute spectrophotometry. *The Astrophysical Journal*, 266:713–717.

Oke, J. B. and Sandage, A. (1968). Energy distributions, k corrections, and the stebbins-whitford effect for giant elliptical galaxies. *The Astrophysical Journal*, 154:21.

Oser, L., Ostriker, J. P., Naab, T., Johansson, P. H., and Burkert, A. (2010). The two phases of galaxy formation. *The Astrophysical Journal*, 725(2):2312.

Overzier, R., Lemson, G., Angulo, R. E., Bertin, E., Blaizot, J., Henriques, B. M. B., Marleau, G.-D., and White, S. D. M. (2013). The Millennium Run Observatory: first light. *Monthly Notices of the Royal Astronomical Society*, 428:778–803.

Overzier, R. A., Heckman, T. M., Kauffmann, G., Seibert, M., Rich, R. M., Basu-Zych, A., Lotz, J., Aloisi, A., Charlot, S., Hoopes, C., et al. (2008). Hubble space telescope morphologies of local lyman break galaxy analogs. i. evidence for starbursts triggered by merging. *The Astrophysical Journal*, 677(1):37.

Palombo, G. (2011). Multivariate Goodness of Fit Procedures for Unbinned Data: An Annotated Bibliography. *ArXiv e-prints*.

Papovich, C., Bassett, R., Lotz, J. M., Van Der Wel, A., Tran, K.-V., Finkelstein, S. L., Bell, E. F., Conselice, C. J., Dekel, A., Dunlop, J., et al. (2012). Candels observations of the structural properties of cluster galaxies at z= 1.62. *The Astrophysical Journal*, 750(2):93.

Papovich, C., Shipley, H., Mehrtens, N., Lanham, C., Lacy, M., Ciardullo, R., Finkelstein, S., Bassett, R., Behroozi, P., Blanc, G., et al. (2016). The spitzer-hetdex exploratory large-area survey. *The Astrophysical Journal Supplement Series*, 224(28):30pp.

Pardo, L. and Menéndez, M. L. (2006). Analysis of Divergence in Loglinear Models When Expected Frequencies are Subject to Linear Constraints. *Metrika*, 64(1):63–76.

Patanchon, G., Ade, P. A., Bock, J. J., Chapin, E. L., Devlin, M. J., Dicker, S. R., Griffin, M., Gundersen, J. O., Halpern, M., Hargrave, P. C., et al. (2009). Submillimeter number counts from statistical analysis of blast maps. *The Astrophysical Journal*, 707(2):1750.

Patil, A., Huard, D., and Fonnesbeck, C. J. (2010). PyMC: Bayesian Stochastic Modelling in Python. *Journal of statistical software*, 35(4):1–81.

Peebles, P. (1973). Statistical analysis of catalogs of extragalactic objects. i. theory. *The Astrophysical Journal*, 185:413–440.

Peng, C. Y., Ho, L. C., Impey, C. D., and Rix, H.-W. (2010). Detailed Decomposition of Galaxy Images. II. Beyond Axisymmetric Models. *Astronomical Journal*, 139:2097–2129.

Penzias, A. A. and Wilson, R. W. (1965). A measurement of excess antenna temperature at 4080 mc/s. *The Astrophysical Journal*, 142:419–421.

Pérez, I. and Sanchez-Blazquez, P. (2011). Study of stellar populations in the bulges of barred galaxies. *Astronomy & Astrophysics*, 529:A64.

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R., Nugent, P., Castro, P., Deustua, S., Fabbro, S., Goobar, A., Groom, D., et al. (1999). Measurements of ω and λ from 42 high-redshift supernovae. *The Astrophysical Journal*, 517(2):565.

Peterson, B. M. (1997). *An introduction to active galactic nuclei*. Cambridge University Press.

Peterson, J. R. and Jernigan, J. G. (2013). phosim: photon simulator. *Astrophysics Source Code Library*.

Petrosian, V. (1976). Surface brightness and evolution of galaxies. *The Astrophysical Journal*, 209:L1–L5.

Pichon, C., Pogosyan, D., Kimm, T., Slyz, A., Devriendt, J., and Dubois, Y. (2011). Rigging dark haloes: why is hierarchical galaxy formation consistent with the inside-out build-up of thin discs? *Monthly Notices of the Royal Astronomical Society*, 418(4):2493–2507.

Poisson, S. (1837). Research on the probability of judgments in criminal and civil matters. *Paris, France: Bachelier*.

Pope, A., Borys, C., Scott, D., Conselice, C., Dickinson, M., and Mobasher, B. (2005). The hubble deep field north scuba super-map-iii. optical and near-infrared properties of submillimetre galaxies. *Monthly Notices of the Royal Astronomical Society*, 358(1):149–167.

Popescu, C. C. and Tuffs, R. J. (2005). The effect of dust on the appearance of the outer disks of spiral galaxies. In *American Astronomical Society Meeting Abstracts*, volume 37 of *Bulletin of the American Astronomical Society*, page 1482.

Prevot, M. L., Lequeux, J., Prevot, L., Maurice, E., and Rocca-Volmerange, B. (1984). The typical interstellar extinction in the Small Magellanic Cloud. *Astronomy and Astrophysics*, 132:389–392.

Primack, J. R. (2005). Precision cosmology. *New Astronomy Reviews*, 49(2):25–34.

Pritchard, J. K., Seielstad, M. T., and Perez-Lezaun, A. (1999). Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular biology and ...*.

Racine, R. (1996). The telescope point spread function. *Publications of the Astronomical Society of the Pacific*, 108(726):699.

Read, J. and Trentham, N. (2005). The baryonic mass function of galaxies. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 363(1837):2693–2710.

Reeves, R. and Pettitt, A. (2005). A theoretical framework for approximate bayesian computation. In *Proc. 20th Int. Works. Stat. Mod. Australia.*, pages 393–396. Francis, A.R., Matawie, K.M., Oshlack, A., Smyth, G.K. (Eds.).

Refregier, A. (2003). Shapelets—i. a method for image analysis. *Monthly Notices of the Royal Astronomical Society*, 338(1):35–47.

Refregier, A. and Amara, A. (2014). A way forward for cosmic shear: Monte-carlo control loops. *Physics of the Dark Universe*, 3:1–3.

Rettura, A., Rosati, P., Nonino, M., Fosbury, R., Gobat, R., Menci, N., Strazzullo, V., Mei, S., Demarco, R., and Ford, H. (2010). Formation epochs, star formation histories, and sizes of massive early-type galaxies in cluster and field environments at z= 1.2: insights from the rest-frame ultraviolet. *The Astrophysical Journal*, 709(1):512.

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., et al. (1998). Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009.

Roberts, G. O., Gelman, A., and Gilks, W. R. (1997). Weak convergence and optimal scaling of random walk Metropolis algorithms. *The Annals of Applied Probability*, 7(1):110–120.

Roberts, G. O. and Rosenthal, J. S. (2007). Coupling and ergodicity of adaptive markov chain monte carlo algorithms. *Journal of applied probability*, 44(2):458–475.

Roberts, G. O. and Rosenthal, J. S. (2009). Examples of Adaptive MCMC. *Journal of Computational and Graphical Statistics*, 18(2):349–367.

Roberts, G. O., Rosenthal, J. S., et al. (2001). Optimal scaling for various metropolis-hastings algorithms. *Statistical science*, 16(4):351–367.

Robertson, B., Bullock, J. S., Cox, T. J., Di Matteo, T., Hernquist, L., Springel, V., and Yoshida, N. (2006). A merger-driven scenario for cosmological disk galaxy formation. *The Astrophysical Journal*, 645(2):986.

Robin, A. C., Luri, X., Reylé, C., Isasi, Y., Grux, E., Blanco-Cuaresma, S., Arenou, F., Babusiaux, C., Belcheva, M., Drimmel, R., Jordi, C., Krone-Martins, A., Masana, E., Mauduit, J. C., Mignard, F., Mowlavi, N., Rocca-Volmerange, B., Sartoretti, P., Slezak, E., and Sozzetti, A. (2012). Gaia Universe model snapshot. A statistical analysis of the expected contents of the Gaia catalogue. *Astronomy and Astrophysics*, 543:A100.

Robin, A. C., Reylé, C., Fliri, J., Czekaj, M., Robert, C. P., and Martins, A. M. M. (2014). Constraining the thick disc formation scenario of the Milky Way. *Astronomy & Astrophysics*, 569:A13–.

Rogstad, D. and Shostak, G. (1971). Aperture synthesis study of neutral hydrogen in the galaxy m101: I. observations. *Astronomy and Astrophysics*, 13:99–107.

Roseboom, I., Oliver, S., Kunz, M., Altieri, B., Amblard, A., Arumugam, V., Auld, R., Aussel, H., Babbedge, T., Béthermin, M., et al. (2010). The herschel multi-tiered extragalactic survey: source extraction and cross-identifications in confusion-dominated spire images. *Monthly Notices of the Royal Astronomical Society*, 409(1):48–65.

Roškar, R., Debattista, V. P., Brooks, A. M., Quinn, T. R., Brook, C. B., Governato, F., Dalcanton, J. J., and Wadsley, J. (2010). Misaligned angular momentum in hydrodynamic cosmological simulations: warps, outer discs and thick discs. *Monthly Notices of the Royal Astronomical Society*, 408(2):783–796.

Rowe, B., Hirata, C., and Rhodes, J. (2011). Optimal Linear Image Combination. *The Astrophysical Journal*, 741:46.

Rowe, B., Jarvis, M., Mandelbaum, R., Bernstein, G. M., Bosch, J., Simet, M., Meyers, J. E., Kacprzak, T., Nakajima, R., Zuntz, J., et al. (2015). Galsim: The modular galaxy image simulation toolkit. *Astronomy and Computing*, 10:121–150.

Rubin, V. C. and Ford Jr, W. K. (1970). Rotation of the andromeda nebula from a spectroscopic survey of emission regions. *The Astrophysical Journal*, 159:379.

Rubner, Y., Tomasi, C., and Guibas, L. (1998). A metric for distributions with applications to image databases. In *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pages 59–66. Narosa Publishing House.

Rybizki, J. and Just, A. (2015). Towards a fully consistent Milky Way disc model - III. Constraining the initial mass function. *Monthly Notices of the Royal Astronomical Society*, 447:3880–3891.

Saglia, R., Sánchez-Blázquez, P., Bender, R., Simard, L., Desai, V., Aragón-Salamanca, A., Milvang-Jensen, B., Halliday, C., Jablonka, P., Noll, S., et al. (2010). The fundamental plane of ediscs galaxies-the effect of size evolution. *Astronomy & Astrophysics*, 524:A6.

Salmon, B., Papovich, C., Long, J., Willner, S. P., Finkelstein, S. L., Ferguson, H. C., Dickinson, M., Duncan, K., Faber, S. M., Hathi, N., Koekemoer, A., Kurczynski, P., Newman, J., Pacifici, C., Pérez-González, P. G., and Pforr, J. (2016). Breaking the Curve with CANDELS: A Bayesian Approach to Reveal the Non-Universality of the Dust-Attenuation Law at High Redshift. *The Astrophysical Journal*, 827:20.

Samushia, L., Reid, B. A., White, M., Percival, W. J., Cuesta, A. J., Zhao, G.-B., Ross, A. J., Manera, M., Aubourg, É., Beutler, F., et al. (2014). The clustering of galaxies in the sdss-iii baryon oscillation spectroscopic survey: measuring growth rate and geometry with anisotropic clustering. *Monthly Notices of the Royal Astronomical Society*, 439(4):3504–3519.

Sancisi, R. (1976). Warped hi disks in galaxies. *Astronomy and Astrophysics*, 53:159.

Sandage, A. (1961). *The Hubble atlas of galaxies*, volume 618. Carnegie Institution of Washington Washington, DC.

Sandage, A., Freeman, K. C., and Stokes, N. R. (1970). The Intrinsic Flattening of e, so, and Spiral Galaxies as Related to Galaxy Formation and Evolution. *The Astrophysical Journal*, 160:831.

Sandage, A. and Tammann, G. (1975). Steps toward the hubble constant. v-the hubble constant from nearby galaxies and the regularity of the local velocity field. *The Astrophysical Journal*, 196:313–328.

Saracco, P., Gargiulo, A., Ciocca, F., and Marchesini, D. (2017). Cluster and field elliptical galaxies at z˜ 1.3-the marginal role of the environment and the relevance of the galaxy central regions. *Astronomy & Astrophysics*, 597:A122.

Saracco, P., Longhetti, M., and Gargiulo, A. (2010). The number density of superdense early-type galaxies at $1 < z < 2$ and the local cluster galaxies. *Monthly Notices of the Royal Astronomical Society*, 408:L21–L25.

Saulder, C., Van Kampen, E., Chilingarian, I. V., Mieske, S., and Zeilinger, W. W. (2016). The matter distribution in the local universe as derived from galaxy groups in sdss dr12 and 2mrs. *Astronomy & Astrophysics*, 596:A14.

Scannapieco, C. e. a., Wadepuhl, M., Parry, O., Navarro, J., Jenkins, A., Springel, V., Teyssier, R., Carlson, E., Couchman, H., Crain, R., et al. (2012). The aquila comparison project: the effects of feedback and numerical methods on simulations of galaxy formation. *Monthly Notices of the Royal Astronomical Society*, 423(2):1726–1749.

Scargle, J. D. (1998). Studies in Astronomical Time Series Analysis. V. Bayesian Blocks, a New Method to Analyze Structure in Photon Counting Data. *The Astrophysical Journal*, 504:405–418.

Scargle, J. D. (2002). Bayesian blocks in two or more dimensions: Image segmentation and cluster analysis. In Fry, R. L., editor, *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, volume 617 of *American Institute of Physics Conference Series*, pages 163–173.

Scargle, J. D., Norris, J. P., Jackson, B., and Chiang, J. (2013). Studies in Astronomical Time Series Analysis. VI. Bayesian Block Representations. *The Astrophysical Journal*, 764:167.

Schafer, C. M. and Freeman, P. E. (2012). Likelihood-Free Inference in Cosmology: Potential for the Estimation of Luminosity Functions. In *Statistical Challenges in Modern Astronomy V*, pages 3–19.

Schechter, P. (1976). An analytic expression for the luminosity function for galaxies. *The Astrophysical Journal*, 203:297.

Scheuer, P. A. (1957). A statistical method for analysing observations of faint radio stars. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 53, pages 764–773. Cambridge University Press.

Schlafly, E. F. and Finkbeiner, D. P. (2011). Measuring reddening with sloan digital sky survey stellar spectra and recalibrating sfd. *The Astrophysical Journal*, 737(2):103.

Schlaifer, R. and Raiffa, H. (1961). *Applied statistical decision theory*.

Schlegel, D. J., Finkbeiner, D. P., and Davis, M. (1998). Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds. *The Astrophysical Journal*, 500:525–553.

Schreiber, C., Elbaz, D., Pannella, M., Merlin, E., Castellano, M., Fontana, A., Bourne, N., Boutsia, K., Cullen, F., Dunlop, J., et al. (2017). egg: hatching a mock universe from empirical prescriptions. *Astronomy & Astrophysics*, 602:A96.

Schultz, G. V. and Wiemer, W. (1975). Interstellar reddening and IR-excess of O and B stars. *Astronomy and Astrophysics*, 43:133–139.

Scott, D. W. (1979). On optimal and data-based histograms. *Biometrika*, 66(3):605–610.

Scoville, N., Aussel, H., Brusa, M., Capak, P., Carollo, C., Elvis, M., Giavalisco, M., Guzzo, L., Hasinger, G., Impey, C., et al. (2007). The cosmic evolution survey (cosmos): overview. *The Astrophysical Journal Supplement Series*, 172(1):1.

Scoville, N., Faisst, A., Capak, P., Kakazu, Y., Li, G., and Steinhardt, C. (2015). Dust Attenuation in High Redshift Galaxies: “Diamonds in the Sky”. *The Astrophysical Journal*, 800:108.

Seaton, M. J. (1979). Interstellar extinction in the UV. *Monthly Notices of the Royal Astronomical Society*, 187:73P–76P.

Sersic, J. L. (1963). Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy. *Boletin de la Asociacion Argentina de Astronomia La Plata Argentina*, 6:41.

Shankar, F., Marulli, F., Bernardi, M., Mei, S., Meert, A., and Vikram, V. (2012). Size evolution of spheroids in a hierarchical universe. *Monthly Notices of the Royal Astronomical Society*, 428(1):109–128.

Sharma, S. (2017). Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy. *ArXiv e-prints*.

Shen, S., Mo, H., White, S. D., Blanton, M. R., Kauffmann, G., Voges, W., Brinkmann, J., and Csabai, I. (2003). The size distribution of galaxies in the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 343(3):978–994.

Sheth, R. K. (2007). On estimating redshift and luminosity distributions in photometric redshift surveys. *Monthly Notices of the Royal Astronomical Society*, 378(2):709–715.

Shimasaku, K., Fukugita, M., Doi, M., Hamabe, M., Ichikawa, T., Okamura, S., Sekiguchi, M., Yasuda, N., Brinkmann, J., Csabai, I., et al. (2001). Statistical properties of bright galaxies in the sloan digital sky survey photometric system. *The Astronomical Journal*, 122(3):1238.

Shlens, J. (2014). A Tutorial on Principal Component Analysis.

Silk, J. and Mamon, G. A. (2012). The current status of galaxy formation. *Research in Astronomy and Astrophysics*, 12(8):917.

Silva, L., Granato, G. L., Bressan, A., and Danese, L. (1998). Modeling the Effects of Dust on Galactic Spectral Energy Distributions from the Ultraviolet to the Millimeter Band. *The Astrophysical Journal*, 509:103–117.

Simard, L. and Somerville, R. (2000). Testing hierarchical galaxy evolution models with the luminosity-size relation of distant field galaxies. In *Dynamics of Galaxies: from the Early Universe to the Present*, volume 197, page 439.

Simpson, D. P., Rue, H., Martins, T. G., Riebler, A., and Sørbye, S. H. (2014). Penalising model component complexity: A principled, practical approach to constructing priors. *ArXiv e-prints*.

Sinharay, S. (2003). Assessing convergence of the markov chain monte carlo algorithms: A review. *ETS Research Report Series*, 2003(1).

Snyder, G. F., Torrey, P., Lotz, J. M., Genel, S., McBride, C. K., Vogelsberger, M., Pillepich, A., Nelson, D., Sales, L. V., Sijacki, D., et al. (2015). Galaxy morphology and star formation in the illustris simulation at z= 0. *Monthly Notices of the Royal Astronomical Society*, 454(2):1886–1908.

Somerville, R. S. and Davé, R. (2015). Physical Models of Galaxy Formation in a Cosmological Framework. *Annual Review of Astronomy and Astrophysics*, 53:51–113.

Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M., Bennett, C., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., et al. (2003). First-year wilkinson microwave anisotropy probe (wmap) observations: determination of cosmological parameters. *The Astrophysical Journal Supplement Series*, 148(1):175.

Springel, V., White, S. D., Jenkins, A., Frenk, C. S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., et al. (2005). Simulations of the formation, evolution and clustering of galaxies and quasars.

Steinmetz, M. and Bartelmann, M. (1995). On the spin parameter of dark-matter haloes. *Monthly Notices of the Royal Astronomical Society*, 272:570–578.

Stephens, M. A. (1974). Edf statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association*, 69:p.730–737.

Stetson, P. B. (1987). Daophot: A computer program for crowded-field stellar photometry. *Publications of the Astronomical Society of the Pacific*, 99(613):191.

Strazzullo, V., Gobat, R., Daddi, E., Onodera, M., Carollo, M., Dickinson, M., Renzini, A., Arimoto, N., Cimatti, A., Finoguenov, A., et al. (2013). Galaxy evolution in overdense environments at high redshift: Passive early-type galaxies in a cluster at z 2. *The Astrophysical Journal*, 772(2):118.

Stringer, M., Brooks, A., Benson, A., and Governato, F. (2010). Analytic and numerical realizations of a disc galaxy. *Monthly Notices of the Royal Astronomical Society*, 407(1):632–644.

Sunnaker, M., Busetto, A. G., Numminen, E., Corander, J., Foll, M., and Dessimoz, C. (2013). Approximate bayesian computation. *PLoS Comput Biol*, 9(1):1–10.

Suto, Y. (2005). Simulations of large-scale structure in the new millennium. In *Symposium-International Astronomical Union*, volume 216, pages 105–119. Cambridge University Press.

Taghizadeh-Popp, M., Fall, S. M., White, R. L., and Szalay, A. S. (2015). Simulating Deep Hubble Images with Semi-empirical Models of Galaxy Formation. *The Astrophysical Journal*, 801:14.

Takeuchi, T. T. and Ishii, T. T. (2004). A general formulation of the source confusion statistics and application to infrared galaxy surveys. *The Astrophysical Journal*, 604(1):40.

Tammann, G. A. (1987). The cosmic distance scale. In Hewitt, A., Burbidge, G., and Fang, L. Z., editors, *Observational Cosmology*, volume 124 of *IAU Symposium*, pages 151–185.

Teerikorpi, P. (1997). Observational selection bias affecting the determination of the extragalactic distance scale. *Annual Review of Astronomy and Astrophysics*, 35(1):101–136.

Teerikorpi, P. (2004). Influence of a generalized eddington bias on galaxy counts. *Astronomy & Astrophysics*, 424(1):73–78.

Teerikorpi, P. (2015). Eddington-Malmquist bias in a cosmological context. *Astronomy and Astrophysics*, 576:A75.

Tegmark, M., Blanton, M. R., Strauss, M. A., Hoyle, F., Schlegel, D., Scoccimarro, R., Vogeley, M. S., Weinberg, D. H., Zehavi, I., Berlind, A., et al. (2004). The three-dimensional power spectrum of galaxies from the sloan digital sky survey. *The Astrophysical Journal*, 606(2):702.

Thain, D., Tannenbaum, T., and Livny, M. (2005). Distributed computing in practice: the condor experience. *Concurrency - Practice and Experience*, 17(2-4):323–356.

Tolman, R. C. (1930). On the estimation of distances in a curved universe with a non-static line element. *Proceedings of the National Academy of Sciences*, 16(7):511–520.

Tolman, R. C. (1934). Relativity. *Thermodynamics and Cosmology*, 235.

Toni, T., Welch, D., Strelkowa, N., Ipsen, A., and Stumpf, M. P. (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface*, 6(31).

Toomre, A. and Toomre, J. (1972). Galactic bridges and tails. *The Astrophysical Journal*, 178:623–666.

Trujillo, I., Cenarro, A. J., de Lorenzo-Cáceres, A., Vazdekis, A., Ignacio, G., and Cava, A. (2009). Superdense massive galaxies in the nearby universe. *The Astrophysical Journal Letters*, 692(2):L118.

Trujillo, I., Conselice, C. J., Bundy, K., Cooper, M., Eisenhardt, P., and Ellis, R. S. (2007). Strong size evolution of the most massive galaxies since z 2. *Monthly Notices of the Royal Astronomical Society*, 382(1):109–120.

Trujillo, I., Förster Schreiber, N. M., Rudnick, G., Barden, M., Franx, M., Rix, H.-W., Caldwell, J. A. R., McIntosh, D. H., Toft, S., Häussler, B., Zirm, A., van Dokkum, P. G., Labbé, I., Moorwood, A., Röttgering, H., van der Wel, A., van der Werf, P., and van Starkenburg, L. (2006). The Size Evolution of Galaxies since z˜3: Combining SDSS, GEMS, and FIRES. *The Astrophysical Journal*, 650:18–41.

Trujillo, I., Graham, A. W., and Caon, N. (2001). On the estimation of galaxy structural parameters: the Sérsic model. *Monthly Notices of the Royal Astronomical Society*, 326:869–876.

Trumpler, R. J. (1930). Absorption of Light in the Galactic System. *Publications of the Astronomical Society of the Pacific*, 42:214.

Tuffs, R. J., Möllenhoff, C., and Popescu, C. C. (2005). Correcting apparent disk scalelengths for the effect of dust. In *American Astronomical Society Meeting Abstracts*, volume 37 of *Bulletin of the American Astronomical Society*, page 1482.

Tully, R. B. and Fisher, J. R. (1977). A new method of determining distances to galaxies. *Astronomy and Astrophysics*, 54:661–673.

Tully, R. B., Pierce, M. J., Huang, J.-S., Saunders, W., Verheijen, M. A., and Witchalls, P. L. (1998). Global extinction in spiral galaxies. *The Astronomical Journal*, 115(6):2264.

Turkkan, N. and Pham-Gia, T. (1997). Algorithm as 308: Highest posterior density credible region and minimum area confidence region: the bivariate case. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 46(1):131–140.

Turkowski, K. (2013). *Filters for Common Resampling Tasks*. Elsevier.

Turner, B. M. and Van Zandt, T. (2012). A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, 56(2):69–85.

Unterborn, C. T. and Ryden, B. S. (2008). Inclination-dependent extinction effects in disk galaxies in the sloan digital sky survey. *The Astrophysical Journal*, 687(2):976.

Valentijn, E. A. (1990). Opaque spiral galaxies. *Nature*, 346(6280):153.

Valentinuzzi, T., Fritz, J., Poggianti, B., Cava, A., Bettoni, D., Fasano, G., D'Onofrio, M., Couch, W., Dressler, A., Moles, M., et al. (2010). Superdense massive galaxies in wings local clusters. *The Astrophysical Journal*, 712(1):226.

Van de Sande, J., Kriek, M., Franx, M., van Dokkum, P. G., Bezanson, R., Whitaker, K. E., Brammer, G., Labbé, I., Groot, P. J., and Kaper, L. (2011). The stellar velocity dispersion of a compact massive galaxy at z= 1.80 using x-shooter: confirmation of the evolution in the mass-size and mass-dispersion relations. *The Astrophysical Journal Letters*, 736(1):L9.

van der Burg, R. F. J., Hildebrandt, H., and Erben, T. (2010). The UV galaxy luminosity function at z = 3-5 from the CFHT Legacy Survey Deep fields. *Astronomy and Astrophysics*, 523:A74.

Van der Kruit, P. (1981). The thickness of the hydrogen layer and the three-dimensional mass distribution in ngc 891. *Astronomy and Astrophysics*, 99:298–304.

Van Der Kruit, P. and Freeman, K. (2011). Galaxy disks. *Annual Review of Astronomy and Astrophysics*, 49:301–371.

Van Der Wel, A., Holden, B. P., Zirm, A. W., Franx, M., Rettura, A., Illingworth, G. D., and Ford, H. C. (2008). Recent structural evolution of early-type galaxies: size growth from z= 1 to z= 0. *The Astrophysical Journal*, 688(1):48.

van Dokkum, P. G., Franx, M., Fabricant, D., Kelson, D. D., and Illingworth, G. D. (1999). A high merger fraction in the rich cluster ms 1054–03 at z= 0.83: Direct evidence for hierarchical formation of massive galaxies. *The Astrophysical Journal Letters*, 520(2):L95.

Van Dokkum, P. G., Franx, M., Kriek, M., Holden, B., Illingworth, G. D., Magee, D., Bouwens, R., Marchesini, D., Quadri, R., Rudnick, G., et al. (2008). Confirmation of the remarkable compactness of massive quiescent galaxies at z˜ 2.3: early-type galaxies did not form in a simple monolithic collapse. *The Astrophysical Journal Letters*, 677(1):L5.

VanderPlas, J. (2014). Frequentism and Bayesianism: A Python-driven Primer. *ArXiv e-prints*.

Vanderplas, J., Connolly, A., Ivezić, Ž., and Gray, A. (2012). Introduction to astroml: Machine learning for astrophysics. In *Conference on Intelligent Data Understanding (CIDU)*, pages 47 –54.

Verter, F. (1993). Effect of Malmquist bias on correlation studies with IRAS data base. *The Astrophysical Journal*, 402:141–159.

Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Nelson, D., and Hernquist, L. (2014). Introducing the illustris project: simulating the coevolution of dark and visible matter in the universe. *Monthly Notices of the Royal Astronomical Society*, pages 1518–1547.

Walcher, J., Groves, B., Budavári, T., and Dale, D. (2011). Fitting the integrated spectral energy distributions of galaxies. *Astrophysics and Space Science*, 331(1):1–51.

Warren, M. S., Quinn, P. J., Salmon, J. K., and Zurek, W. H. (1992). Dark halos formed via dissipationless collapse. I - Shapes and alignment of angular momentum. *The Astrophysical Journal*, 399:405.

Waskom, M., Botvinnik, O., drewokane, Hobson, P., David, Halchenko, Y., Lukauskas, S., Cole, J. B., Warmenhoven, J., de Ruiter, J., Hoyer, S., Vanderplas, J., Villalba, S., Kunter, G., Quintero, E., Martin, M., Miles, A., Meyer, K., Augspurger, T., Yarkoni, T., Bachant, P., Williams, M., Evans, C., Fitzgerald, C., Brian, Wehner, D., Hitz, G., Ziegler, E., Qalieh, A., and Lee, A. (2016). seaborn: v0.7.1 (june 2016).

Wasserman, L. (2013). *All of statistics: a concise course in statistical inference*. Springer Science & Business Media.

Weber, R., Schek, H.-J., and Blott, S. (1998). A quantitative analysis and performance study for similarity-search methods in high-dimensional spaces. In *VLDB*, volume 98, pages 194–205.

Weinberg, S. (2008). *Cosmology*. Oxford University Press.

Weingartner, J. C. and Draine, B. T. (2001). Dust Grain-Size Distributions and Extinction in the Milky Way, Large Magellanic Cloud, and Small Magellanic Cloud. *The Astrophysical Journal*, 548:296–309.

Westra, E., Geller, M. J., Kurtz, M. J., Fabricant, D. G., and Dell'Antonio, I. (2010). Empirical Optical k-Corrections for Redshifts. *Publications of the Astronomical Society of the Pacific*, 122:1258.

Weyant, A., Schafer, C., and Wood-Vasey, W. M. (2013). Likelihood-free Cosmological Inference with Type Ia Supernovae: Approximate Bayesian Computation for a Complete Treatment of Uncertainty. *The Astrophysical Journal*, 764:116.

White, S. D. and Frenk, C. S. (1991). Galaxy formation through hierarchical clustering. *The Astrophysical Journal*, 379:52–79.

White, S. D. and Rees, M. J. (1978). Core condensation in heavy halos: a two-stage theory for galaxy formation and clustering. *Monthly Notices of the Royal Astronomical Society*, 183(3):341–358.

Williams, R. E., Blacker, B., Dickinson, M., Dixon, W. V. D., Ferguson, H. C., Fruchter, A. S., Giavalisco, M., Gilliland, R. L., Heyer, I., Katsanis, R., Levay, Z., Lucas, R. A., McElroy, D. B., Petro, L., Postman, M., Adorf, H.-M., and Hook, R. (1996). The Hubble Deep Field: Observations, Data Reduction, and Galaxy Photometry. *The Astronomical Journal*, 112:1335.

Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., Toft, S., Kriek, M., and Labbé, I. (2010). The Evolving Relations Between Size, Mass, Surface Density, and Star Formation in 3 x 10^4 Galaxies Since z = 2. *The Astrophysical Journal*, 713:738–750.

Willman, B. and Strader, J. (2012). "galaxy," defined. *The Astronomical Journal*, 144(3):76.

Willmer, C. N. A. (1997). Estimating galaxy luminosity functions. *Astronomical Journal*, 114:898–912.

Wolf, C., Dye, S., Kleinheinrich, M., Meisenheimer, K., Rix, H.-W., and Wisotzki, L. (2001). Deep BVR photometry of the Chandra Deep Field South from the COMBO-17 survey. *Astronomy and Astrophysics*, 377:442–449.

Wolf, C., Meisenheimer, K., Rix, H.-W., Borch, A., Dye, S., and Kleinheinrich, M. (2003). The COMBO-17 survey: Evolution of the galaxy luminosity function from 25 000 galaxies with 0.2< z <1.2. *Astronomy and Astrophysics*, 401:73–98.

Wright, E. L., Eisenhardt, P. R., Mainzer, A. K., Ressler, M. E., Cutri, R. M., Jarrett, T., Kirkpatrick, J. D., Padgett, D., McMillan, R. S., Skrutskie, M., et al. (2010). The wide-field infrared survey explorer (wise): mission description and initial on-orbit performance. *The Astronomical Journal*, 140(6):1868.

Yan, R., White, M., and Coil, A. L. (2004). Mock catalogs for the deep2 redshift survey. *The Astrophysical Journal*, 607(2):739.

Yang, R. and Berger, J. O. (1996). *A catalog of noninformative priors*. Institute of Statistics and Decision Sciences, Duke University.

Yasuda, N., Fukugita, M., Narayanan, V. K., Lupton, R. H., Strateva, I., Strauss, M. A., Ivezić, Ž., Kim, R. S., Hogg, D. W., Weinberg, D. H., et al. (2001). Galaxy number counts from the sloan digital sky survey commissioning data. *The Astronomical Journal*, 122(3):1104.

Yip, C.-W., Szalay, A. S., Carliles, S., and Budavári, T. (2011). Effect of Inclination of Galaxies on Photometric Redshift. *The Astrophysical Journal*, 730:54.

Yoachim, P. and Dalcanton, J. J. (2006). Structural parameters of thin and thick disks in edge-on disk galaxies. *The Astronomical Journal*, 131(1):226.

Yoon, Y., Im, M., and Kim, J.-W. (2017). Massive galaxies are larger in dense environments: Environmental dependence of mass–size relation of early-type galaxies. *The Astrophysical Journal*, 834(1):73.

York, D. G., Adelman, J., Anderson, Jr., J. E., Anderson, S. F., Annis, J., Bahcall, N. A., Bakken, J. A., Barkhouser, R., Bastian, S., Berman, E., Boroski, W. N., Bracker, S., Briegel, C., Briggs, J. W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Carr, M. A., Castander, F. J., Chen, B., Colestock, P. L., Connolly, A. J., Crocker, J. H., Csabai, I., Czarapata, P. C., Davis, J. E., Doi, M., Dombeck, T., Eisenstein, D., Ellman, N., Elms, B. R., Evans, M. L., Fan, X., Federwitz, G. R., Fiscelli, L., Friedman, S., Frieman, J. A., Fukugita, M., Gillespie, B., Gunn, J. E., Gurbani, V. K., de Haas, E., Haldeman, M., Harris, F. H., Hayes, J., Heckman, T. M., Hennessy, G. S., Hindsley, R. B., Holm, S., Holmgren, D. J., Huang, C.-h., Hull, C., Husby, D., Ichikawa, S.-I., Ichikawa, T., Ivezić, Ž., Kent, S., Kim, R. S. J., Kinney, E., Klaene, M., Kleinman, A. N., Kleinman, S., Knapp, G. R., Korienek, J., Kron, R. G., Kunszt, P. Z., Lamb, D. Q., Lee, B., Leger, R. F., Limmongkol, S., Lindenmeyer, C., Long, D. C., Loomis, C., Loveday, J., Lucinio, R., Lupton, R. H., MacKinnon, B., Mannery, E. J., Mantsch, P. M., Margon, B., McGehee, P., McKay, T. A., Meiksin, A., Merelli, A., Monet, D. G., Munn, J. A., Narayanan, V. K., Nash, T., Neilsen, E., Neswold, R., Newberg, H. J., Nichol, R. C., Nicinski, T., Nonino, M., Okada, N., Okamura, S., Ostriker, J. P., Owen, R., Pauls, A. G., Peoples, J., Peterson, R. L., Petravick, D., Pier, J. R., Pope, A., Pordes, R., Prosapio, A., Rechenmacher, R., Quinn, T. R., Richards, G. T., Richmond, M. W., Rivetta, C. H., Rockosi, C. M., Ruthmansdorfer, K., Sandford, D., Schlegel, D. J., Schneider, D. P., Sekiguchi, M., Sergey, G., Shimasaku, K., Siegmund, W. A., Smee, S., Smith, J. A., Snedden, S., Stone, R., Stoughton, C., Strauss, M. A., Stubbs, C., SubbaRao, M., Szalay, A. S., Szapudi, I., Szokoly, G. P., Thakar, A. R., Tremonti, C., Tucker, D. L., Uomoto, A., Vanden Berk, D., Vogeley, M. S., Waddell, P., Wang, S.-i., Watanabe, M., Weinberg, D. H., Yanny, B., Yasuda, N., and SDSS Collaboration (2000). The Sloan Digital Sky Survey: Technical Summary. *Astronomical Journal*, 120:1579–1587.

Younger, J. D., Hopkins, P. F., Cox, T., and Hernquist, L. (2008). The self-regulated growth of supermassive black holes. *The Astrophysical Journal*, 686(2):815.

Zehavi, I., Blanton, M. R., Frieman, J. A., Weinberg, D. H., Mo, H. J., Strauss, M. A., Anderson, S. F., Annis, J., Bahcall, N. A., Bernardi, M., Briggs, J. W., Brinkmann, J., Burles, S., Carey, L., Castander, F. J., Connolly, A. J., Csabai, I., Dalcanton, J. J., Dodelson, S., Doi, M., Eisenstein, D., Evans, M. L., Finkbeiner, D. P., Friedman, S., Fukugita, M., Gunn, J. E., Hennessy, G. S., Hindsley, R. B., Ivezić, Ž., Kent, S., Knapp, G. R., Kron, R., Kunszt, P., Lamb, D. Q., Leger, R. F., Long, D. C., Loveday, J., Lupton, R. H., McKay, T., Meiksin, A., Merrelli, A., Munn, J. A., Narayanan, V., Newcomb, M., Nichol, R. C., Owen, R., Peoples, J., Pope, A., Rockosi, C. M., Schlegel, D., Schneider, D. P., Scoccimarro, R., Sheth, R. K., Siegmund, W., Smee, S., Snir, Y., Stebbins, A., Stoughton, C., SubbaRao, M., Szalay, A. S., Szapudi, I., Tegmark, M., Tucker, D. L., Uomoto, A., Vanden Berk, D., Vogeley, M. S., Waddell, P., Yanny, B., and York, D. G. (2002). Galaxy Clustering in Early Sloan Digital Sky Survey Redshift Data. *The Astrophysical Journal*, 571:172–190.

Zhang, Y. and Yang, X. (2017). Size distribution of galaxies in sdss dr7. *arXiv preprint arXiv:1707.04979*.

Zhukovska, S. and Henning, T. (2014). Life Cycle of Dust in the Magellanic Clouds and the Milky Way. *ArXiv e-prints*.

Zucca, E., Ilbert, O., Bardelli, S., Tresse, L., Zamorani, G., Arnouts, S., Pozzetti, L., Bolzonella, M., McCracken, H. J., Bottini, D., Garilli, B., Le Brun, V., Le Fèvre, O., Maccagni, D., Picat, J. P., Scaramella, R., Scodeggio, M., Vettolani, G., Zanichelli, A., Adami, C., Arnaboldi, M., Cappi, A., Charlot, S., Ciliegi, P., Contini, T., Foucaud, S., Franzetti, P., Gavignaud, I., Guzzo, L., Iovino, A., Marano, B., Marinoni, C., Mazure, A., Meneux, B., Merighi, R., Paltani, S., Pellò, R., Pollo, A., Radovich, M., Bondi, M., Bongiorno, A., Busarello, G., Cucciati, O., Gregorini, L., Lamareille, F., Mathez, G., Mellier, Y., Merluzzi, P., Ripepi, V., and Rizzo, D. (2006). The VIMOS VLT Deep Survey. Evolution of the luminosity functions by galaxy type up to z = 1.5 from first epoch data. *Astronomy and Astrophysics*, 455:879–890.

Zwicky, F. (1933). The redshift of extragalactic nebulae. *Helv. Phys. Acta*, 6:110.