Université Pierre et Marie Curie

École doctorale d'Astronomie et d'Astrophysique d'Île-de-France (ED127)

Institut d'Astrophysique de Paris

Bayesian large-scale structure inference and cosmic web analysis

Thèse de doctorat

Discipline : Cosmologie

présentée et soutenue publiquement le 24 septembre 2015

par

Florent Leclercq

devant un jury composé de :

Bertrand Laforge	Professeur, Université Pierre et Marie Curie	*Président du jury*
Benjamin Wandelt	Professeur, Université Pierre et Marie Curie	*Directeur de thèse*
Oliver Hahn	Professeur, Université Nice Sophia Antipolis	*Rapporteur*
Alan Heavens	Professeur, Imperial College London	*Rapporteur*
Ofer Lahav	Professeur, University College London	*Examinateur*
Will Percival	Professeur, University of Portsmouth	*Examinateur*
Rien van de Weygaert	Professeur, University of Groningen	*Invité*
Matías Zaldarriaga	Professeur, Institute for Advanced Study, Princeton	*Invité*

Because all things balance – as on a wheel – and we cannot see nine-tenths of what is real, our claims of self-reliance are pieced together by unpanned gold.

— Franklin D'Olier Reeve (1995), *Coasting*

There's gold, and it's haunting and haunting;
    It's luring me on as of old;
Yet it isn't the gold that I'm wanting
    So much as just finding the gold.
It's the great, big, broad land 'way up yonder,
    It's the forests where silence has lease;
It's the beauty that thrills me with wonder,
    It's the stillness that fills me with peace.

— Robert William Service (1907), *The Spell of the Yukon*

Preface

> "But especially he loved to run in the dim twilight of the summer midnights, listening to the subdued and sleepy murmurs of the forest, reading signs and sounds as a man may read a book, and seeking for the mysterious something that called – called, waking or sleeping, at all times, for him to come."
> — Jack London (1903), *The Call of the Wild*

This PhD thesis is written as completion of my work at the Institut d'Astrophysique de Paris from 2012 to 2015. Three years ago, I started this project with heartfelt enthusiasm, but it turned out to be much more than expected – an incredibly rewarding journey. This thesis is the report of this long process. It expresses my vision – that I have had incredible trouble in organizing linearly – of the final scientific products.

Unfortunately, it cannot describe the thought process and the strange mechanism by which something out in there in the Universe, be it a galaxy cluster, a scientific tool, or a concept – something that a month ago was a stranger – becomes intimate. Neither does it capture my feelings during the long days – and nights – spent in front of a black board, a paper or a computer, alone or in the lab: the sadness and tiredness with failed endeavors, the bittersweet taste of learning I was wrong, the hope for good results, the joy for successes, and the wonder at the elegance of the cosmos.

As a consequence, this preface takes the occasion to describe how this thesis came into being.

Stars and physical sciences

I have always thought that modern physics is fascinating. Not only does it manipulate extraordinary ideas and concepts (quantum mechanics, relativity), it also deals with important societal issues (energy, natural resources). In this respect, I am fascinated by how far the physical sciences, in just a few hundred years, have taken us in our understanding of nature. Given this incredible evolution, it is amazing to realize that looking at the sky – the amusement during the warm summers in ancient Greece – still gathers so much attention. This strange mixture of tradition and modernity may be the reason why, as far as I remember, I have always had a particular attraction for astrophysics.

For most of the history of humanity, cosmology was part of religion or metaphysics. Only recently did it become a science, its peculiarity being the uniqueness of its object of interest – the Universe as a whole. The idea that the entire Universe can be treated as a physical system was one of the most striking revelations of my life as a student. It is commonly predicted that early 21st century cosmology is on the verge of a revolution. In upcoming years, surprising or unexpected results may or may not be found, but I believe that cosmology will stay one of humanity's greatest intellectual endeavors, and certainly the one that has produced the deepest description of the natural world as we find it.

Models and beliefs

Contrary to my long-held passion for the Universe, my interest in probability theory came in a rather fortuitous manner. But faced with the immodest, enthralling questions of cosmology, one soon realizes that there is no absolute truth, only beliefs. Colleagues showed me that this viewpoint makes all the problems of modern cosmology appear in a very different light. Then, in a quick succession, reading about probability theory, which truly is the "logic of science", in the words of Jaynes (2003), made me realize that in the much larger and permanent world of plausible reasoning, i.e. rational thinking in the presence of uncertainty, the current problems of physics appear only as details: what matters is the road, not the destination.

Few, if any, of our ordinary-life beliefs are certain to the degree that we cannot imagine them being overthrown by sufficient contradictory evidence. Similarly, typical commonsense inferences rely on applying rules that are general, but not universal. Therefore, deduction does not entirely characterize commonsense reasoning. This thesis exploits theories of inductive and abductive logic, to draw and assess the strength of conclusions from uncertain rules and partial evidence.

Probability theory, when seen as an extension of ordinary logic, incorporates the description of randomness but also statistical inference and becomes a field of logical unity and simplicity. It allows us to solve problems of great complexity, and reproduces many aspects of human cognitive activity, often in disturbing detail. In doing so, it captures something about how our minds operate when we form inductive judgments, of which we may not be consciously aware. This aspect takes a very particular meaning when we deal with the Universe.

Why bother about this thesis?

This thesis focuses on methodology for the analysis of the large-scale structure of the Universe. I should say from the beginning that the methods presented do not have the same degree of maturity as standard techniques for the analysis of galaxy surveys. So why should the reader bother?

1. I believe that new solutions to complex problems involving both data and uncertainty are needed to exploit the full potential of future, but also of existing surveys. However many data sets we record and analyze, if we use the same old models without questioning them, we will always miss the same crucially important feature that the experiment was competent to find, of which we may not be aware. If we want to detect any phenomenon, we must have a data model that at least allows the possibility that it may exist.

2. Innovative methods also allow the possibility to crosscheck the validity of cosmological analyses that are widely accepted. A false premise or a confirmation bias, when built into a model that is never questioned, cannot be removed by any amount of new data. Only a fresh look can get rid of that.

In the hope to make progress in the analysis of the cosmic large-scale structure, this thesis tries to develop a "healthy disrespect for tradition and authority, which have retarded progress throughout the 20th century" (Jaynes, 2003).

How to read this thesis

The thesis is divided into four parts. Part I is a preparatory discussion on the analytical and numerical description of the large-scale structure. The heart of this thesis is part II on Bayesian large-scale structure inference and part IV on cosmic web analysis. The transitional part III focuses on the non-linear regime of structure formation.

It is my hope that at least some chapters are written in a sufficiently engaging textbook style for graduate students. This should be the case in particular for chapter 1 on structure formation, chapter 3 on Bayesian statistics and appendix B on numerical simulations. There, except in section 1.1, I have struggled to avoid the dreadful sentence "It can be shown that..." as much as possible. I have tried to give references to the original literature whenever it is possible, but I certainly did not attempt a true bibliography as can be found in excellent review papers on the large-scale structure and on probability theory. Chapter 4 describes the BORG algorithm, which is the basis of this entire thesis. It gathers information scattered through published journal papers, and hence is also intended for reference use. The rest of the thesis directly draws from the research papers that have been published during my PhD work: chapter 2 from Leclercq *et al.* (2013) and its addendum, Leclercq, Jasche & Wandelt (2015b), chapter 5 from Jasche, Leclercq & Wandelt (2015), chapter 6 from Leclercq *et al.* (2013), chapter 7 from Leclercq *et al.* (2015) and Leclercq, Jasche & Wandelt (2015c), chapter 8 from Leclercq *et al.* (2015), chapter 9 from Leclercq, Jasche & Wandelt (2015c), and chapter 10 from Leclercq, Jasche & Wandelt (2015a). There, the style becomes more succinct and the aim is rather to describe specific projects, give a guide to the literature and report on the results obtained.

Considering that this thesis is a rare occasion to include whatever I want in a research work, I decided to tackle the difficult task of choosing epigraphs. The various quotes spread throughout the thesis may or may not have something to do with the main text. Some are inspirational, thought-provoking, some are openly provocative, some can just be considered as Easter eggs, and some are just there for free.

The online version of thesis will be revised to correct for any mistakes, typographical and otherwise, found after it goes to press and archiving. I will try to maintain a list of corrections on my website, currently hosted at `http://www2.iap.fr/users/leclercq/`. Please feel free to send me any comments at `florent.leclercq@polytechnique.org`.

Cosmology is a journey

> "Bien lire l'Univers, c'est bien lire la vie."
> — Victor Hugo (1856), *Les Contemplations*

Though physical cosmology is celebrating its first century, it is no relic of the past. We live unique and very exciting times, when we expect to see a qualitative leap in our knowledge of the Universe within a lifetime. I consider myself incredibly fortunate to be part of this adventure.

In my experience, the loneliness felt by some researchers is easily overcome in our field by a simple thought: a cosmologist's quest is the quest of all humanity. This is why, I believe, cosmology resonates with people all around the world well beyond professional scientists, in different places and cultures. It touches everybody intellectually, but also emotionally and spiritually, without prejudice. As probability theory says something about how our mind works, physical cosmology tells us how we can think of ourselves as a species.

Before moving to the traditionally must-read acknowledgement section, I would like to quote Paulo Coelho's prologue to *The Alchemist* (1988). When Narcissus falls into the lake and dies, the lake weeps, and declares: "I weep for Narcissus, but I never noticed that Narcissus was beautiful. I weep because, each time he knelt beside my banks, I could see, in the depths of his eyes, my own beauty reflected." When we look into the deep Universe, the Universe also may be looking deeply into us.

Acknowledgments

> "There are all kinds of love in this world, but never the same love twice."
> — Francis Scott Fitzgerald (1925), *The Great Gatsby*

What a long road this has been. It is difficult to believe that the end has finally arrived. That I could write this thesis, and much more importantly, that I became who I am to write it, is due in no small part to the support of a great many people. It is not a process that started three years ago, but way before that. Along the road, I met thousands of wonderful people. I wish I could mention by name anyone who has helped, in one way or another, but the list is very, very long. I hope that I have been as kind with you as you have been with me, and I apologize for not writing down your name.

I first wish to express my gratitude to my supervisor, Benjamin Wandelt, who was abundantly helpful and offered invaluable assistance, support and guidance. I cannot emphasize enough how delightful it has been to work with you, not only for your knowledge, dedication and modesty, but also for your interest in the human being that lies behind the scientist. Your expertise, understanding, and patience added considerably to my experience during the preparation of this thesis. Our meetings were always rich in ideas, jokes and laughs; I learned so much from you, not only scientifically, but also on how to be successful as a researcher. This is something very precious that I intend to keep forever.

I would like to address very special thanks to Jens Jasche for the opportunity to work within the larger effort surrounding Bayesian large-scale structure inference – such a stimulating endeavor – and for his assistance at all levels during this research project. I am very grateful for the time that you gave selflessly to provide me with direction and technical assistance, especially during my first two years. I greatly enjoyed our many scientific discussions and the way you influenced my scientific approaches. You have been a friendly, creative and supportive advisor.

I am also grateful to several researchers with whom I have had the chance to collaborate in the last three years: Jacopo Chevallard, Héctor Gil-Marín, Nico Hamaus, Guilhem Lavaux, Emilio Romano-Díaz, Paul Sutter, and Alice Pisani. I greatly benefited from the experience of each of you, and it has been a pleasure to work on common projects. To Nico, thanks for the organization of the weekly happy hour and for your unfailing contribution to the good atmosphere. To Paul, thanks for a very nice workshop in Ohio and for a memorable void barbecue at the occasion of the "Breaking of the Fellowship". À Guilhem, merci de m'avoir fait partager ton expérience, tout en me rappelant qu'il ne faut pas trop parler de science comme si c'était un sujet sérieux ! À Alice, merci pour ton sourire, pour nos discussions et tout ce qu'on a partagé.

Two people, with whom I have not worked directly, deserve however a very special mention: Joseph Silk and Matías Zaldarriaga. Joe is one of the scientists I admire most, not only for his immense knowledge in astrophysics and cosmology, but also for his kindness and simplicity. Many thanks for sharing valuable insights into your vision of science and research. Matías followed my work at various stages during my PhD project. I am grateful for your invitations to Princeton, for your encouragements and for many useful discussions and suggestions. This year, applying for postdoctoral positions has been both stimulating and challenging – yet another experience of what it means to grow up. In this respect, I want to thank my recommendation letter writers, Ben Wandelt, Jens Jasche, Joe Silk and Matías Zaldarriaga, who very kindly accepted to support my applications. I am very obliged to Will Percival for his trust and for the perspective of thought-provoking exchanges, stimulating discussions and a fruitful collaboration.

I am indebted to my defense committee, Oliver Hahn, Alan Heavens, Bertrand Laforge, Ofer Lahav, Will Percival, Rien van de Weygaert and Matías Zaldarriaga for the honor to judge my work. I am extremely proud to be able to gather such a prestigious jury. Additional thanks go to the rapporteurs, Oliver Hahn and Alan Heavens, for having spent part of their summer reviewing the manuscript and for their positive judgment of my work. I am very obliged to Alan Heavens for the honor to feature one of the cosmic web maps presented in this thesis as the front cover of the proceedings of IAU Symposium 306 "Statistical challenges in 21st century cosmology".

Participating in various international conferences and summer schools has always been a fruitful and enjoyable experience. I want to thank the organizers of the ICTP summer school on cosmology and workshop on large-scale structure (2012), the Varenna summer school (2013), the Les Houches summer school (2013), the Rencontres de Moriond (2014, Cosmology session), the IAU symposia 306 and 308 in Lisbon and Tallinn (2014), the CCAPP workshop on cosmic voids (2014), COSMO 2014 in Chicago, the MPA-EXC workshop on the dynamic Universe (2014, Garching), the ICTP workshop on cosmological structures (2015), the ESO-MPA-EXC large-scale structure conference (2014, Garching), and the Rencontres du Vietnam in Quy Nhon (2015, Cosmology session). Best greetings, in particular, to the Les Houches' cosmologists group; thanks also to the organizers of the student conferences I attended: Elbereth 2012, 2013, 2014, and the SCGSC 2013. On various occasions, I have had the chance to have friendly and interesting discussions (even if sometimes short) within the cosmology community. In particular, my work benefited from interactions with Niayesh Afshordi, Raul Angulo, Stephen Bailey, Robert Cahn, Olivier Doré, Torsten Enßlin, Luigi Guzzo, Oliver Hahn, Jean-Christophe Hamilton, Alan Heavens, Shirley Ho, Mike Hudson, Eiichiro Komatsu, Ofer Lahav, Mark Neyrinck, Nelson Padilla, Bruce Partridge, Will Percival, David Schlegel, Uroš Seljak, Sergei Shandarin, Ravi Sheth, Svetlin Tassev, and Rien van de Weygaert (among many others). At this point, it also seems needed to acknowledge the decisive contribution of a familiar ~ 20 Mpc/h void (at coordinates $x \approx -100$, $y \approx 200$ in the slice that I usually show), which very nicely makes my point during presentations.

Cette thèse n'aurait pas été la même si elle n'avait pas été préparée dans l'excellent environnement scientifique de l'IAP. Pour cette raison, je suis reconnaissant à tous ses chercheurs, et en particuliar, pour d'intéressantes conversations, à Francis Bernardeau, Luc Blanchet, François Bouchet, Jean-François Cardoso, Stéphane Charlot, Yohan Dubois, Florence Durret, Silvia Galli, Valérie de Lapparent, Matt Lehnert, Gary Mamon, Henry McCracken, Jean-Philippe Uzan, Sébastien Peirani, Patrick Peter, Cyril Pitrou, et Sébastien Renaux-Petel. Merci à Valérie de Lapparent pour son aide, utile pour la rédaction de l'introduction de cette thèse, concernant l'historique des relevés de galaxies et plusieurs questions observationnelles. Thanks to Matt Lehnert for sharing his experience of the politics of research. Une pensée amicale aux étudiants en thèse que j'ai eu l'opportunité de côtoyer pendant trois ans: Sylvain et Guillaume (docteurs en 2013); Vincent, Vincent, Maxime, Jean, Hayley, Guillaume, Manuel et Flavien (docteurs en 2014); Nicolas, Hélène, Sandrine, Charlotte, Vivien et Clément (vous me devancez de quelques jours, félicitations !); et ceux pour qui ce n'est pas encore fini (bon courage !): Pierre, Thomas, Alba, Clotilde, Jean-Baptiste, Julia, Laura, Mélanie, Rebekka, Caterina, Erwan, Federico, Nicolas, Sébastien, Tilman. Amitiés également à quelques uns des labos voisins: Fabien, Julian, Benjamin, Julien, Mathilde, Tico, Anaïs, François, Linc, Agnès, Marta. Pour la gestion sans faille des serveurs Horizon, dont mon travail a beaucoup bénéficié, merci à Stéphane Rouberol et Christophe Pichon. Enfin, merci aux services informatique et administratif de l'IAP, en particulier à Isabelle Guillerme et Olivia Leroy pour la gestion de mes missions.

Je voudrais exprimer ma reconnaissance à l'École polytechnique pour une formation hors pair. Tout particulièrement, je dois à certains professeurs de m'avoir ouvert les yeux sur la relativité, l'astrophysique et la cosmologie, et d'avoir partagé avec moi leur vision de la recherche : Francis Bernardeau, David Langlois et Martin Lemoine. Merci à Fabio Iocco d'avoir guidé mes premiers pas dans le monde de la recherche, et à tous

les professeurs de la première promotion du M2 Physique des Hautes Énergies à l'X, en particulier aux directeurs Ignatios Antoniadis et Jean-Claude Brient, de m'avoir engagé sur la bonne voie pour commencer cette thèse. Pour leur suivi pendant ces trois ans, merci à l'ED 127 et à la commission des thèses de physique, en particulier, respectivement, à Florence Durret et Benoît Semelin.

Bisous à mes amis de longue date (je ne suis pas sûr que vous pensiez trouver votre nom ici !): les BIP Antoine (comment vont les caribous ?), Sébastien, Nico, Vincent, Arnaud et Nicopathe; Alex et Élise – plus un p'tit gars dont je ne connais pas encore le prénom au moment d'écrire ces lignes; Thomas et Agathe; Amandine, Arnaud et Aymeric; Anaëlle, Justin, Kéan et Aaron; Manuela, Thibaut, Marine, Clotilde, Benji, Julien, Samuel, Stéphane, et tous les autres. Merci pour les soirées, les vacances, les jeux, et pour me rappeler que s'évader n'est pas uniquement faire de la cosmologie.

« *Vivre d'orgies est ma seule espérance, le seul bonheur que j'aie pu conquérir. C'est sur les flots qu'j'ai passé mon enfance, c'est sur les flots qu'un forban doit mourir.* » Une pensée très particulière pour les corsaires de la Ciotat et les bretteurs de l'X: Cap'taine Daniel; mes différents partenaires Kornichon, Sasha, Seva, et Daniel le Savant; Emmanuel le Gaucher, Guillaume la Montagne, Charles Ventre-Sec, Dino l'Italien, Lauriane Joli-cœur, Céline Fleur d'épine, Benoît la Grenouille; et tous les corsaires ciotadens. Merci pour les combats, l'ambiance et les chants. May your swords stay sharp!

Enfin, il conviendrait sûrement d'exprimer ma reconnaissance à ma famille, mais ce que j'aurais à leur dire va bien au-delà de ce qu'il est possible d'écrire ici.

My final thanks go to you, the reader. If you are reading this thesis linearly, you have already read 9 pages and only have 228 to go. I hope the galaxies shine brightly over you through the rest of your life!

Florent Leclercq
Paris, September 2015