Higher order statistics for cosmology : likelihood development for future surveys like Euclid

Virginia Ajani

To cite this version:

Virginia Ajani. Higher order statistics for cosmology : likelihood development for future surveys like Euclid. Other [cond-mat.other]. Université Paris Cité, 2021. English. NNT : 2021UNIP7071 . tel-03644314

HAL Id: tel-03644314
https://theses.hal.science/tel-03644314

Submitted on 19 Apr 2022

Université de Paris

Ecole doctorale:
Astronomie et Astrophysique d'Ile de France n. 127
Laboratoire:
Laboratoire de Cosmologie et Statistiques
Commissariat à l'Energie Atomique et aux Energies Alternatives

Thèse de doctorat en Astronomie et Astrophysique

Higher order statistics for cosmology: likelihood development for future surveys like Euclid

---

Par **Virginia Ajani**

Présentée et soutenue publiquement le 8 Octobre 2021

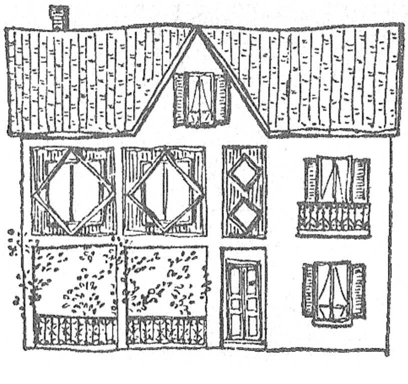
Directrice de thèse: Valeria Pettorino, Directrice de Recherche, CEA
Co-Directeur de thèse: Jean-Luc Starck, Directeur de Recherche, CEA

Composition du Jury:

Françoise Combes	Professeur, Collège de France	Présidente du Jury
David Bacon	Professor, University of Portsmouth	Rapporteur
Alexandre Refregier	Professor, ETH Zurich	Rapporteur
Sophie Maurogordato	Directrice de Recherche, Observatoire de Nice	Examinatrice
James Bartlett	Professeur, Université de Paris	Examinateur
Vincenzo Fabrizio Cardone	Docteur, Astronomical Observatory of Rome	Examinateur

Alla nonna Gisella,

a cui mi sarebbe piaciuto tanto
mostrare questo manoscritto,
sul terrazzo di Viù.



Disegno di Elisabetta Ajani.

SUMMARY

Higher order statistics for cosmology: likelihood development for future surveys like Euclid

Weak gravitational lensing by the large-scale structure is the effect of bending of light emitted by background galaxies due to the presence of foreground matter. It represents a powerful tool for estimating cosmological parameters as it is sensitive to the large-scale structure of the universe. Past, present, and future cosmological surveys, like the upcoming European Space Agency's *Euclid* mission, will use it as one of the main physical probes for investigating unsolved questions in current cosmology, such as: what the properties of the dark components of the universe are, what the origin of its accelerated expansion is and what the sum of neutrino masses is. As weak lensing surveys become deeper, they reveal more non-Gaussian features of the matter density field, requiring statistics beyond the second order to properly extract cosmological information. This has motivated the introduction of several statistics of order higher than the second, such as Minkowski functionals, higher-order moments, the bispectrum, peak counts and, most recently, the scattering transform, wavelet phase harmonic statistics, and machine learning-based methods to account for non-Gaussian information in cosmological analysis. The aim of this thesis is to investigate and develop statistical methods to optimally extract the information encoded in the data in the context of higher order statistics, that can eventually help to improve the constraints on cosmological parameters in current and future cosmological analysis. In a first study, we have compared different summary statistics (from power spectrum to peak counts) and quantified the impact of different multi-scale filtering techniques on cosmological forecasts obtained in an ideal setting. Specifically, we employ a starlet filter, which is an isotropic undecimated wavelet transform that naturally decomposes the original image in several different images of the same size, enabling to extract cosmological information from the maps at different scales simultaneously. The performed study is tomographic, and we compare different summary statistics, assuming *Euclid*-like shape noise. The performance of the starlet is compared with a concatenation of Gaussian filters using peak counts and with the state of the art of summary statistics in weak lensing in an ideal setting without systematics effects. The findings are that in both multi-scale settings, peak counts result to perform better than the state of the art for second order statistics and single scale peak counts. Moreover, when using a multi-scale approach, joining power spectrum and peaks does not add any relevant information over considering just the peaks alone. While the performance and behaviour of both multi-scale filters are similar, for the starlet filter we find that the majority of the information of the data covariance matrix is encoded in the diagonal elements. This can be advantageous when inverting the covariance matrix, speeding up the numerical implementation. Based on the promising performance of the multi-scale approach in the context of peak counts found in the first study, we propose a new high order summary statistics called the starlet ℓ_1-norm. This statistics provides a multi-scale calculation of the full voids and peaks distribution and avoids the problem of the definition of what is a peak and what is a void. Indeed, it is built so that it carries the information encoded in all pixels of the map and not just the one in local maxima and local minima. The outcome is that, in a tomographic ideal setting (without further systematics) with the assumption of *Euclid* shape noise, this new summary statistics outperforms commonly used ones, such as weak lensing peak counts, minimum counts, or combination of the two in terms of constraining power. The starlet ℓ_1-norm represents a new promising unified framework to account for the information encoded in peak counts and voids simultaneously, preserving the advantages of the multi-scale approach. In order to accomplish these two studies, I developed a pipeline written in python that takes as input weak lensing convergence maps and provides as output cosmological parameters constraints. In a second part, the thesis presents preliminary results and procedures of the current work in progress. We aim at extending the previous findings to different survey settings and cosmological probes, by applying high order statistics to galaxy clustering simulations for the Kilo-Degree Survey (KiDS); as the two previous works were carried in an ideal setting (without systematics), we show preliminary results of a first application to real data from the Canada-France Imaging survey (CFIS) to test the robustness of these statistics.

Keywords: Weak gravitational lensing ; higher order statistics ; cosmological parameters ; wavelet transform ; Euclid ; Dark energy ; Cosmology ; multi-scale analysis ; Large-scale structure

RÉSUMÉ DE LA THÈSE EN FRANÇAIS

Statistiques d’ordre supérieur pour la cosmologie : développement de la fonction de vraisemblance pour des missions comme Euclid

L’ effet de lentille gravitationnelle faible lié aux structures à grande échelle est l’effet de la courbure de la lumière émise par les galaxies d’arrière-plan en raison de la présence de matière au premier plan. Il représente un outil puissant pour estimer les paramètres cosmologiques car il est sensible aux structures à grande échelle de l’univers. Les études cosmologiques passées, présentes et futures, comme la prochaine mission *Euclid* de l’Agence spatiale européenne, l’utiliseront comme l’une des principales sondes physiques pour enquêter sur les questions non résolues de la cosmologie actuelle, comme les propriétés des composants sombres de l’univers, l’origine de son expansion accélérée ou de la somme totale des masses des neutrinos. Au fur et à mesure que les relevés par lentilles faibles deviennent plus profonds, ils révèlent davantage les caractéristiques non Gaussiennes du champ de densité de matière, nécessitant ainsi l’utilisation de statistiques au-delà du second ordre pour mieux extraire des données les informations cosmologiques. Ceci a motivé l’introduction de différentes mesures statistiques d’ordre supérieur, telles que les fonctions de Minkowski, le calcul des moments d’ordre supérieurs à 2, le bispectre, les comptages de pics des vides ou, plus récemment, la scattering transform, la decomposition wavelet phase harmonic statistics et des techniques d’intelligence artificielle. Cette thèse étudie le potentiel d’effectuer des mesures statistiques multi-échelles dans le contexte des statistiques d’ordre superieur pour contraindre les paramètres cosmologiques. Le manuscrit commence en fournissant les outils généraux nécessaires pour comprendre les résultats présentés pendant la thèse. En particulier, le modèle cosmologique standard, les principes de formations des structures et certains des défis actuels de la cosmologie moderne sont présentés dans le premier chapitre. Un aperçu des principales sondes cosmologiques qui seront utilisées par *Euclid* est présenté avec les outils statistiques bayésiens qui seront utilisés dans les différents travaux présentés dans la thèse. Le deuxième chapitre traite ensuite plus en détail les lentilles gravitationnelles faibles. La présentation commence avec la dérivation des équations de lentille à partir de la déviation lumineuse. Puis, après avoir introduit la convergence et le cisaillement, le lien entre observations et modèles cosmologiques est traité. Ensuite, des statistiques de second ordre du cisaillement cosmique sont décrites et un bref aperçu des effets systématiques tels que les effets des baryons, l’alignement des ellipticités intrinsèques des galaxies, le biais sur la mesure du cisaillement dans les lentilles faibles et certaines des contraintes actuelles sont présentés. Après, je décrit l’état de l’art des statistiques d’ordre supérieur comme par exemple le bispectre, les fonctions des Minkowski, les moments d’ordre supérieur à 2, le comptage de pics et des vides et enfin des techniques qui utilisent l’apprentissage automatique. L’objectif du troisième chapitre est d’introduire quelques définitions utiles dans le cadre du formalisme en ondelettes. En particulier, je commence par donner une brève introduction de la notion de parcimonie, de transformée en

ondelettes continue et d'ondelette discrète. Grâce à ces outils, la transformée en starlette, qui sera largement utilisée dans les chapitres 4, 5 et 6 est introduite. Ensuite, la procédure suivie pour estimer la bruit dans les images utilisées dans le manuscript et calculer les coefficients d'ondelettes est illustrée. Enfin, un bref aperçu des récentes méthodes de reconstruction des cartes de masse basées sur des représentations parcimonieuses, comme `DeepMass` et `MCALens`, est fourni. Dans le quatrième chapitre je présente la motivation, la méthodologie et les résultats de notre étude présentés dans mon premier papier (Ajani et al. 2020). Nous avons étudié l'impact des approches de filtrage multi-échelles sur les paramètres cosmologiques en employant un filtre starlette (ondelette) et une concaténation de Filtres Gaussiens pour contraindre la somme des masses de neutrinos M_ν, le paramètre de densité de matière Ω_{m} et l'amplitude du spectre de puissance primordial A_{s} en utilisant quatre redshifts tomographiques en simulant un bruit pour une mission comme *Euclid*. Nous avons comparé différentes statistiques: le spectre de puissance, le comptage de pics à une seule échelle et à plusieurs échelles et évalué leurs performances dans un cadre idéal (sans inclure effets systématiques). Nous considérons d'abord deux techniques d'analyse multi-échelles sur des simulations de cartes convergence de lentilles faibles fournies par les manipulations `MassiveNus`. Nous ajoutons du bruit qui simule *Euclid* à ces cartes et nous comparons les deux techniques, afin de quantifier l'impact du choix du filtre sur la capacité à contraindre les paramètres. La première approche est la concaténation de filtres Gaussiens suivie de comptages de pics et la seconde consiste à utiliser une transformée starlette qui est une transformée en ondelettes isotrope non décimée et qui décompose une image en plusieurs bandes de même taille, permettant d'extraire simultanément des informations cosmologiques des cartes à différentes échelles. Cette étude exploite l'information tomographique et considère un bruit d'une amplitude de l'ordre du bruit attendu avec la mission *Euclid*. Nous trouvons que ces deux approches de comptage de pics multi-échelles performent mieux des méthodes couramment utilisées qui sont les statistiques de second ordre comme le spectre de puissance ou le comptage de pics à une seule échelle pour ce qui concerne la puissance de contraindre. De plus, nous observons que dans le cas d'une analyse multi-échelle, combiner les pics multi-échelles avec le spectre de puissance n'ajoute aucune information supplémentaire par rapport à la prise en compte uniquement des pics. L'avantage de la transformée starlette par rapport aux filtres multi-Gaussiens est que la matrice de covariance est quasiment diagonale, ce qui facilite son inversion. Plus spécifiquement, nous trouvons que:

a) Pour les comptages de pics, une approche de filtrage multi-échelle des cartes bruitées conduit à un facteur d'amélioration de plus de deux par rapport à une approche mono-échelle (noyau Gaussien unique pour filtrer la carte de convergence) pour les contraintes conjointes sur $(M_\nu, \Omega_{\mathrm{m}})$, (M_ν, A_{s}) et $(\Omega_{\mathrm{m}}, A_{\mathrm{s}})$ lors de l'utilisation d'un noyau starlette, et un facteur de plus de trois lors de l'utilisation d'un filtre multi-Gaussien. Ceci est encore plus évident dans les contraintes marginalisées, où l'amélioration est respectivement de 43% sur M_ν, 25% sur Ω_{m} et 34% sur A_{s} pour la starlette, alors que pour le *setting* multi-Gaussien c'est 54% sur M_ν, 25% sur Ω_{m} et 45% sur A_{s}. L'utilisation d'un filtre multi-Gaussien au lieu d'un filtre starlette dans le contexte du nombre de pics pourrait améliorer les contraintes de 19% sur M_ν et de 18% sur A_s, alors qu'aucune amélioration est remarquée pour Ω_m. Enfin, les deux comptages de pics multi-échelles fonctionnent mieux que le spectre de puissance sur l'ensemble des paramètres $[M_\nu, \Omega_\mathrm{m}, A_\mathrm{s}]$ respectivement de 63$\%$, 40$\%$ et 72$\%$ lors de l'utilisation d'un filtre starlette et de 70$\%$, 40$\%$ et 77$\%$ lors de l'utilisation d'un filtre Gaussien multi-échelle.

b) Lors de la combinaison des pics multi-échelle avec le spectre de puissance, c'est-à-dire en utilisant une concaténation du nombre de pics et du spectre de puissance comme vecteur de données observé, nous constatons que l'information est principalement contenue dans les seuls pics (pour certains paramètres, tels que Ω_m dans le cas de la starlette, elle est encodé *complètement*). Cela suggère que lors de l'adoption d'une approche multi-échelle, il pourrait être suffisant de travailler avec les pics seuls.

c) L'inclusion de la carte *coarse* dans la décomposition en starlette lors du comptage des pics préserve des informations cruciales. De plus, pour les cartes avec une taille de pixel de 0.4 arcmin, il existe une résolution minimale (c'est-à-dire la plus petite échelle nécessaire) pour les échelles de starlette correspondant à $\theta_\mathrm{ker} = 3.2$ arcmin pour atteindre une puissance de contrainte maximale. Cela nous permet d'exclure les deux premières échelles les plus fines de la décomposition en starlette, qui correspondent aux fréquences les plus élevées et sont les plus sujettes à l'impact du bruit, permettant une analyse plus rapide et plus efficace.

d) Nous remarquons que l'utilisation d'un filtre starlette conduit à une matrice de covariance de données très diagonale, tandis que pour le filtre multi-Gaussien, les termes hors diagonale sont prédominants et les corrélations entre les différentes échelles sont non négligeables. En d'autres termes, la majorité des informations dans le cas du filtre starlette est codée dans les éléments diagonaux de la matrice de covariance. C'est un aspect intéressant du filtre starlette qui pourrait s'avérer utile lorsqu'il s'agit de données de grande dimension où la matrice de covariance peut être difficile à inverser.

En résumé, dans mon premier papier, nous confirmons que le nombre de pics dans le contexte des lentilles faibles est un outil puissant pour dériver contraintes sur les paramètres cosmologiques, en particulier lors de l'étude du régime non linéaire où l'impact de paramètres tels que les masses de neutrinos devient important. Nous soulignons également l'importance d'adopter une approche multi-échelle dans le contexte de comptages de pics de lentilles faibles, qui présentent l'avantage d'analyser les informations encodées à différentes échelles simultanément, conduisant ainsi à des contraintes plus strictes que l'analyse à une seule échelle. Ensuite, dans le cinquième chapitre, je présente l'étude qui à mené à mon deuxième papier (Ajani et al. 2021). Dans cet étude, nous proposons ensuite de remplacer le comptage de pic multi-échelle par un nouveau descripteur statistique, que nous appelons ℓ_1-norm starlette, et qui consiste à sommer la norme ℓ_1 des coefficients starlette de la carte de convergence. Ceci évite de devoir définir la notion de pic ou de vide, et permet de tenir compte de toute l'information contenue dans les coefficients de notre analyse. Outre le fait d'élégamment unifier l'analyse des pics et des vides, cette nouvelle méthodologie est plus efficace non seulement que chacune des deux indépendamment mais aussi de l'analyse jointe pics-vides. Motivé par les avantages des approches multi-échelles pour l'inférence de paramètres cosmologiques comme discuté dans le chapitre 4 et avec les résultats récents de Coulton et al. 2020a en combinant pics et minima, nous proposons la starlette ℓ_1-norm comme nouvelle statistique d'ordre supérieur multi-échelles. La mesure des amplitudes des pics à plusieurs échelles peut être considérée comme une mesure de la norme ℓ_1 d'un sous-ensemble de coefficients d'ondelettes positifs. De même, la mesure des amplitudes des vides peut être considérée comme une mesure de la norme ℓ_1 d'un sous-ensemble de coefficients d'ondelettes négatifs. Les ondelettes fournissent donc un excellent cadre pour une analyse conjointe des pics et des vides, dans laquelle les informations de tous les coefficients d'ondelettes sont incluses. Nous proposons donc d'utiliser une statistique ℓ_1-norm très simple définie comme la somme de la ℓ_1-norm de tous les coefficients dans une case du rapport signal sur bruit S/N (pixel) donnée pour chaque échelle d'ondelettes, définie comme:

$$l_1^{j,i} = \sum_{u=1}^{\#coef(\mathcal{S}_{j,i})} \mid \mathcal{S}_{j,i}[u] \mid = ||\mathcal{S}_{j,i}||_1 \tag{1}$$

si B_i et B_{i+1} sont positifs, cela correspond à la définition de l'ensemble des coefficients $\mathcal{S}_{j,i}$ à l'échelle j et dans le bin i tel que $\mathcal{S}_{j,i} = \{w_{j,k}/B_i < w_{j,k} < B_{i+1}\}$, où k est le pixel index. On peut alors calculer la somme $\sum_{u=1}^{\#coef(\mathcal{S}_{j,i})} \mathcal{S}_{j,i}[u]$ et l'indice u va de 1 au nombre de pixels dans un bac donné i à l'échelle j (i.e. $\#coef(\mathcal{S}_{j,i})$). Nous étudions l'impact de l'utilisation de la starlette norme ℓ_1 en tant que statistique calculée sur des cartes de convergence à lentilles faibles pour estimer les paramètres cosmologiques et nous trouvons qu'elle est potentiellement très efficace en surpasser les statistiques de pointe, le spectre de puissance et la combinaison de pics et de vides monoéchelle, respectivement de 72% et 24% sur M_ν, 60% et 50% sur Ω_m et 75 % et 24 % sur A_s. Nous avons en outre proposé des *starlette extrema* comme statistiques et les avons comparés à la ℓ_1-norm: dans ce cas également, cette dernière est plus performante en termes de puissance de contrainte, dans le cadre idéal actuel, tandis que les extrema de starlette présentent l'avantage d'une matrice de covariance plus diagonale. Pour servir de sonde robuste pour la cosmologie de précision, la puissance statistique seul ne suffit pas; pour leur utilisation, il sera important de tester comment ces statistiques réagissent dans un cadre non idéal, et comment leurs performances sont impactées par les effets systématiques du signal. En fin de chapitre, je propose également une description et visualisation des différentes étapes du pipeline utilisé pour obtenir le résultats des chapitre 4 et 5. Dans la suite, je montre dans le chapitre 6 les résultats préliminaires relatif à l'application de notre nouvelle méthodologie pour le relevé de galaxies KiDS (Kilo-Degree Survey) et également pour une autre sonde, le clustering des galaxies. Plus en détail, je présente la procédure et quelques résultats préliminaires d'un travail en cours (Ajani et al. in prep.) dont le but est d'étendre le pipeline décrit dans le chapitre précédent en ce qui concerne 1) l'espace des paramètres considérés 2) la sonde

cosmologique. La nouvelle sonde envisagée est le *clustering* photométrique des galaxies, qui, avec la lentille faible et le *clustering* spectroscopique des galaxies, sera l'une des principales sondes des relevés de galaxies de prochaine génération comme *Euclid.* Différentes sondes cosmologiques sont affectées de manière différente par les redshifts considérés et par les *ranges* dynamiques de la structure à grande échelle sous-jacente. Ceci, ainsi que les différences dans les instruments employés et les approches de mesure, permettent d'obtenir des informations cosmologiques complémentaires et de traiter facilement les effets systématiques car ils sont généralement distincts et non corrélés dans différentes sondes. La combinaison de sondes peut donc exploiter ces avantages. De plus, comme discuté en détail tout au long de cette thèse, les statistiques d'ordre supérieur au second ordre peuvent fournir des informations non Gaussiennes supplémentaires qui sont généralement perdues lors de l'utilisation de statistiques du second ordre uniquement. En utilisant donc une nouvelle suite de simulations fournies par le premier auteur de Harnois-Déraps, J., Giblin, B., and Joachimi, B. 2019, je décris les paramètres et la procédure suivie jusqu'à présent et que nous envisageons de suivre pour étudier l'impact de la combinaison de sondes dans le cadre des statistiques d'ordre supérieur. Simultanément, j'illustre les questions ouvertes que nous devons encore aborder pour étendre le pipeline pour atteindre cet objectif et l'extension déjà accomplie. Dans le septième chapitre, je fournis la description de la première application de mon pipeline à des données réelles du relevé Canada-France Imagerie (CFIS). Le relevé d'imagerie Canada-France (CFIS), est un relevé en bandes u et r couvrant 5000 deg2 en l'hémisphère Nord. Sa haute résolution et sa profondeur en font un excellent candidat aux sondages pour la science des lentilles faibles à ce jour. Le CFIS est également un élément essentiel des données nécessaires pour redshifts photométriques pour *Euclid.* Nous profitons donc de la disponibilité des données du CFIS et commençons à inclure les systématiques une à une, pour quantifier leurs effets d'abord séparément et éventuellement conjointement. Nous commençons par l'inclusion des effets baryoniques et fournissons une première description qualitative de leur impact dans trois scénarios différents pour une analyse des comptage de pics. Maintenant, dans le but final d'inclure autant de systématiques que possible, nous nous concentrons sur une deuxième systématique, à savoir le biais sur la mésure de cisaillement.

Mots clés: Lentillage gravitationnel faible ; Statistiques d'ordre supérieur ; paramètres cosmologiques; Transformée en ondelettes ; Euclid ; Energie sombre ; Cosmologie ; analyse multi-échelle; Grandes structures de l'Univers

ACKNOWLEDGMENTS

I would like to thank David Bacon, James Bartlett, Vincenzo Fabrizio Cardone, Françoise Combes, Sophie Maurogordato and Alexandre Refregier for accepting to be members of my thesis committee. It is an honour for me to have you as my Jury. A particular thanks to my reviewers David Bacon and Alexandre Refregier, for taking the time to thoroughly read my manuscript and provide comments to improve it.

My deepest thanks go to my two advisors Valeria and Jean-Luc, whose careful and constant supervision during the three years made it possible for me to accomplish this milestone. Thank you for all our scientific discussions and for your precious advice and guidance concerning the PhD projects as well as my professional future. Thank you for being always present and ready to answer any of my questions and for never stop motivating me, both in normal and pandemic times. Thank you for everything you taught me and for the thoughts and care you have dedicated to this PhD. Moreover, I would like to thank Valeria also for involving me in the Euclid IST: Likelihood team: observing the organisation and management of this fantastic team has been inspiring and has taught me so much in terms of team interactions, project organisations and communication. I am profoundly grateful to Jerome Rodriguez and Marc Sauvage for following my work during the three years through the *Comité de suivi*, your point of view has been enriching and truly important for me.

By trying to chronologically retrace the steps of this PhD experience in CosmoStat, my profound gratitude goes also to Austin, Florent, Sam, Santi and Martin for teaching me so much with patience since the very beginning of these three years. A warm and heartfelt thank goes to my *big academic brothers and sister* and friends Morgan, Christophe, Imane and Axel. Thank you for all the time spent together and for always being a point of reference. Thanks to my PhD twin Fadi for the good mood, our jokes and for everything we shared during this experience and thanks to my amazing office mates Tob and Vanshi for making every day in the office very cool. I would like to thank also the magic JCCC organisers Santi, Tob and Isa, I really enjoyed organising the JC with you and learned a lot from you at the same time. Infinite thanks to Sam and François for being always so ready and so fast in kindly answering any question and helping to solve any issue. Thanks again to Santi and Axel, for all the time you dedicated to my questions. I am thankful to all my fellow labmates and all the colleagues I had the luck to meet during these years, Rebecca, Marco, Kostas, Mel, Sandrine, Joana, Remy, Denise, Zac, Benjamin C., Benjamin R., Jiaxin, Hung, André and Sergio, I wish you all the best and thank you for the time spent together. I also had the chance to meet many great interns working at CosmoStat during my stay. Thank you all, especially to Emma, with whom I enjoyed a lot working on peak counts with Martin and Valeria, I wish you all the best for your PhD. Thanks to Vichi, for all the fun time we spent together and for always being there for me and to Christian, Sunayana, Boris for making all coffee breaks and beers at Albert's always interesting and funny.

I would like to use this space to thank the organisers and teachers of the *Euclid* summer school and to the colleagues and friends I was lucky to meet there, especially thinking to Calum, Alex, Sylvain, Marie and Nora. I am deeply grateful to the external collaborators I have worked with during these years, in particular to Jia, Zack, Niall and Joachim for

teaching me so much.
Alongside the amazing people I met at work, coming back home in Paris has always been a joy, thanks to Fede, Ester and all the *gang* of the XIII, Victor, Fra, Luca and Matte; thanks to you the longest winter ever felt like a summer.
Being far from home, the support of all my friends in Italy has been essential, and I want to thank especially Luke, Ludo and Agnes for having shared many moments here. Infinite thanks to my parents and to my family, for their constant love, support and curiosity also during this PhD experience.
Finally, I would like to deeply thank Filippo for making my life better and for always being there, no matter the distance between us (if separated by 6257.59 km or if confined in 19 m^2).

Acknowledgments to fundings, simulations and data

Virginia Ajani acknowledges support by the Centre National d'Etudes Spatiales and the project Initiative d'Excellence (IdEx) of Université de Paris (ANR-18-IDEX-0001). Virginia Ajani thanks the Columbia Lensing group (http://columbialensing.org) for making their simulations available. The creation of these simulations is supported through grants NSF AST-1210877, NSF AST-140041, and NASA ATP-80NSSC18K1093. We thank New Mexico State University (USA) and Instituto de Astrofisica de Andalucia CSIC (Spain) for hosting the Skies & Universes site for cosmological simulation products. We would like to thank Joachim Harnois-Deraps for making public the SLICS mock data which can be found at `http://slics.roe.ac.uk/` and for providing the new set of *cosmo*-SLICS for KiDS-1000. This work is based on data obtained as part of the Canada-France Imaging Survey, a CFHT large program of the National Research Council of Canada and the French Centre National de la Recherche Scientifique. Based on observations obtained with MegaPrime MegaCam, a joint project of CFHT and CEA Saclay, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers (INSU) of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. We would like to thank the authors of Coulton et al. 2020a for providing us with the data arrays of baryonic corrections used in Chapter 7.

CONTENTS

List of Figures

List of Tables

Introduction

A large part of current research in cosmology is driven by two fundamental questions, being what the nature and properties of dark matter and dark energy are. Observations of the Cosmic Microwave Background (CMB) at microwave and infrared frequencies through the European Space Agency (ESA) *Planck* mission enabled to measure several cosmological parameters at the percent level and to establish the standard cosmological model to be the Λ Cold Dark Matter (ΛCDM) model (Planck Collaboration et al. 2020). More precisely, we know that only 5% of the Universe energy-matter content is composed by ordinary visible matter that we are able to describe via the standard model of particle physics. The remaining 95% is composed by two unknown components, referred to as dark matter and dark energy. Also if we take a closer look to the standard model of particle physics, unanswered questions are still present and in search of an answer, such as what the values of the neutrino masses are. To answer these questions, the first step to undertake is to investigate which physical effects and observables are sensitive to dark matter, dark energy and the sum of neutrino masses. We know that dark energy for example, is associated with the accelerated expansion of the universe (Riess et al. 1998; Perlmutter et al. 1999) and that the amount and the behaviour of dark matter and dark energy have an impact on how cosmic structures grow. Combining information coming from the expansion history with the one from the growth of structures will enable to understand whether the current discrepancies between theory and observations lie in how dark energy is modelled (testing the validity of considering it is a constant or if it is actually a fluid evolving in time) or if it is a sign that our current theory of gravity needs to be modified. Next year, in 2022, the European Space Agency Medium Class *Euclid* (Laureijs et al. 2011) satellite will be launched in space to survey the sky for 6 years in the near infrared and visible part of the electromagnetic spectrum. In practice, *Euclid* will collect ~ 2 billions of galaxy shapes and ~ 30 millions of galaxy positions that will make it possible to investigate the expansion history of the universe and the growth of structure in a redshift range $0 < z < 2$, namely covering the period of time during which dark energy played a crucial role in accelerating the expansion of the Universe. This will allow to further test the ΛCDM paradigm at low redshifts, providing complementary information to the CMB, as Planck was designed to test the physics of the very early Universe, which does not bring a lot of information on dark energy. The galaxy shapes and positions that will be collected by *Euclid* are the cornerstone of the primary cosmological probes used by the *Euclid* mission. These are the gravitational lensing by the large-scale structure and galaxy clustering. Gravitational lensing is the phenomenon of bending of light emitted from some background sources due to the presence of massive bodies between the sources and the observer. This bending induces a distortion of the original images of the sources and it represents a powerful tracer of the total foreground matter causing the distortion. Therefore, it can be used to directly probe the distribution of dark matter. In the regime in which these deformations are small, this phenomenon is called *weak lensing* and it is possible to directly map the dark matter distribution and the measured lensing quantities. To make the connection between the physical observable that we will obtain and the theoretical models that need to be tested, statistical techniques to extract the cosmological information present in the data are needed. Generally speaking, this means extracting features that characterise the data with a summary statistics and then compare the summary statistics relative to the observed data with the ones predicted by a given model. A common summary statistics that has been used in the past and current cosmological analysis is the *two-point correlation function* of these data (or its Fourier transform, the power spectrum). Despite being widely employed in current analysis, it is known that second order statistics fail to capture the non-Gaussian information encoded in the data due to non-linear regime effects of structure formation. Indeed, to overcome this information loss, several statistics of order higher than the second have been proposed in the literature, as the bispectrum, Minkowski functionals, higher order moments, peak counts, minima counts and very recently and in parallel to this thesis work, the scattering transform and wavelet phase harmonic statistics. This long list will be discussed more in details in a dedicated chapter. Along with *Euclid*, other wide large-scale imaging and spectroscopic surveys as DESI (DESI Collaboration et al. 2016), the Rubin Observatory's Legacy Survey of Space and Time (LSST) (LSST Science Collaboration et al. 2009) and the Roman Space Telescope (WFIRST) (Spergel et al. 2015) will measure the matter distribution with unprecedented precision, will enable to reach very non-linear scales and will provide an enormous amount of data. Hence, a crucial, parallel challenge of modern cosmology is to master this unprecedented precision, to be able to use and optimise the huge amount of information we will receive to extract useful information, with high accuracy. The path from input data to the final cosmological parameters includes many challenges. Among them, we need to be able to extract the useful features encoded in the data, to optimise the gain in cosmological information. This can be done at different levels of the analysis. For instance, to be able to appreciate the weak lensing distortions, one needs to smooth or filter the shape noise that is dominating over the lensing signal. Some questions arising at this level are then: how does the filtering of the noise impact the information retrieved on cosmological parameters? Are there filters that are more suited than others to extract the information? More in general, which is the statistics that best allows us to retrieve cosmological information and impact our understanding of the universe? Depending on the knowledge that one has on the physical field to be analysed the choice of the summary statistics employed plays a crucial role. For example, if we know that the physical field is very close to a Gaussian field, then the power spectrum can be a good choice to extract features. Or, if it is not far from a Gaussian field one can use n-point correlation functions. At late times, we know that an important contribution of the density maps is composed by dark matter halos and clusters, so peak counts can be an optimal choice. Moreover, it has been proved that combination of certain summary statistics as for example peaks and voids can bring an improvement in the constraining power, so this is also another aspect to keep in mind.

Beyond the combination of summary statistics of the same probe, it is known that combination of different cosmological probes adds additional information to cosmological analysis with respect to a single probe. As explained before, different cosmological probes are sensitive to the underlying large-scale structure field in different ways as well as to systematics effects.

Exploit the advantages coming from both combination of summary statistics of the same probe and of different probes combined, represent an important challenge for next generation galaxy surveys.

The goal of this thesis is to investigate these aspects, to find ways to optimally extract the information encoded in the data in the context of higher order statistics that can eventually help to improve the constraints on cosmological parameters in current and future cosmological analysis.

The content of the thesis is structured as follows: a first part composed by Chapter 1, Chapter 2 and Chapter 3 provides the background tools and the state of the art needed to understand the rest of the thesis. In particular, in Chapter 1 I introduce some basic concepts in General Relativity and cosmology, define cosmological parameters and the current cosmological model and principles of structure formation. Then a brief overview of *Euclid* primary probes is illustrated, followed by some main tools used in Bayesian inference. Chapter 2 outlines the weak lensing formalism, second order summary statistics and current constraints from some of the main current galaxy surveys. Then, the state of the art of statistics of order higher than the second is presented. Chapter 3 describes some concepts in signal processing useful for the purposes of this work, as sparsity, the wavelet starlet transform and some example of usage of wavelet for mass mapping aims. A second part composed by Chapter 4, Chapter 5, Chapter 6 and Chapter 7 illustrates the contributions of my PhD work. In particular, in Chapter 4, I present the analysis and results of a work where we investigate the impact of the multi-scale filtering techniques using weak lensing peak counts. This led to the paper: Ajani et al. 2020. In Chapter 5 I present a new summary statistics that we have introduced motivated from the findings of the first paper along with the search for a unified framework for peaks and voids. This new statistics is called the *starlet* ℓ_1-norm and we presented it for the first time in a second paper: Ajani et al. 2021. Chapter 6 and Chapter 7 describe instead my current ongoing work, so the nature of the results and discussion of these last two chapters is more preliminary and at a earlier stage than Chapter 4 and Chapter 5. Specifically, in Chapter 6 I describe the work I am following to extend the results of this thesis work to photometric galaxy clustering. Chapter 7 presents a first attempt of application on real data. The final discussion and future prospects are then summarised in Chapter 8.

CHAPTER 1

BACKGROUND

The purpose of this first chapter is to provide the general tools needed to understand the rest of this thesis. In particular the standard cosmological model, principles of structure formation and some of the current challenges of modern cosmology are presented. A brief overview of the primary cosmological probes that will be used by *Euclid* is presented along with the Bayesian statistical tools that will be used in the different works presented throughout the thesis.

1.1 Cosmological Principle

Our current understanding of the structure of the universe is based on the Friedmann-Lemaître-Robertson-Walker (FLRW) model. In this model the universe is homogeneous and isotropic on large-scales. Isotropy is rotational invariance around any point. Homogeneity implies that the space is the same in every point and it can be seen as translational invariance. The universe's homogeneity and isotropy constitute the Cosmological Principle, i.e. there is no privileged position nor preferential direction in our universe. The present section summarises the main results of modern cosmology which will be useful for the following chapters.

1.1.1 Homogeneity and isotropy

In order to undertake a description consistent with the Cosmological Principle, it is necessary to work with a metric built assuming homogeneity and isotropy. Friedmann-Lemaître-Robertson-Walker (FLRW) metric is then built using invariants in order to satisfy the Cosmological Principle criteria. Also, it is convenient to work in a comoving reference frame and to choose spherical coordinates which are very well suited for the isotropy requirement. The resulting line-element is:

$$
\begin{aligned}
ds^2 &= dt^2 - a^2(t)\left[\frac{dr}{1-Kr^2} + r^2 d\theta^2 + r^2 sin^2\theta d\phi^2\right] &(1.1)\\
&= dt^2 - a^2(t)\left[d\chi^2 + f_K(\chi)^2\left(d\theta^2 + sin^2\theta d\phi^2\right)\right], &(1.2)
\end{aligned}
$$

where $a(t)$ is the *scale factor*, a parameter that is free to vary as a function of time that physically provides information on how distances between points change as the universe contracts or expands (Hubble 1929); (r, θ, ϕ) the spherical coordinates, χ and $f_K(\chi)$ the comoving radial and transverse distance, defined in Section 1.1.3 and K is a parameter of curvature such that

$$K = \begin{cases} 1 & \text{Closed universe} \quad (\boldsymbol{S}^3) \\ 0 & \text{Flat universe} \quad (\boldsymbol{R}^3) \\ -1 & \text{Open universe} \quad (\boldsymbol{H}^3) \end{cases} \tag{1.3}$$

where $\boldsymbol{S}^3$, $\boldsymbol{R}^3$, $\boldsymbol{H}^3$ are respectively the three-dimensional sphere, hyperboloid and flat Euclidean space embedded in four-dimensional Euclidean space $\boldsymbol{R}^4$. All three possible spatial geometries are homogeneous and isotropic, hence satisfying the Cosmological principle. More in general, the local geometry of space-time can be entirely described by a metric tensor $g_{\mu\nu}$ that defines the infinitesimal distances of a trajectory in a curved space through the line element

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu. \tag{1.4}$$

The kinematics and dynamics of universe's evolution are described by Einstein's Field Equations

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{R} = 8\pi G T_{\mu\nu}. \tag{1.5}$$

The left hand side (l.h.s.) contains the information relative to the geometry and curvature of space-time, as the Ricci curvature tensor $R_{\mu\nu}$ and Ricci scalar $\mathcal{R}$ are expressions of the metric $g_{\mu\nu}$ and its derivatives. The right hand side (r.h.s) provides information about the source of such curvature, encoded in the energy-momentum tensor $T_{\mu\nu}$. Einstein's Equations can be obtained as field equations of General Relativity by applying the principle of least action to

$$\mathcal{S} = \mathcal{S}_{HE} + \mathcal{S}_{\mathcal{M}} = \int d^4x \sqrt{g} \left[-\frac{1}{16\pi G} \mathcal{R} + \mathcal{L}_{\mathcal{M}} \right], \tag{1.6}$$

where $\mathcal{S}_{HE}$ is the Hilbert-Einstein action, that minimised leads to Equation 1.5 while $\mathcal{S}_{\mathcal{M}}$ is the action for matter, which leads exactly to the source of space-time curvature when minimised, i.e.

$$T_{\mu\nu} = -\frac{2}{\sqrt{g}} \frac{\delta \mathcal{S}_{\mathcal{M}}}{\delta g_{\mu\nu}} \tag{1.7}$$

to the energy-momentum tensor.

As mentioned above, through Equation 1.5, it is possible to derive information about the content of the universe and its evolution. In particular, thanks to the cosmological principle, working in a comoving inertial frame, the energy-momentum tensor can be treated as a perfect fluid, completely defined by its rest-frame energy density ρ and momentum p as

$$T^{\mu}{}_{\nu} = T^{\mu\alpha}g_{\alpha\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & p \end{pmatrix} \tag{1.8}$$

where ρ is the energy density of the perfet fluid and p its momentum. This can be formally obtained by considering the four momentum $p^{\mu} = mU^{\mu}$, where U^{μ} is the four velocity $U^{\mu} = \frac{dx^{\mu}}{d\tau}$ with respect to proper time τ. Hence, $T_{\mu\nu}$ represents the flux of the four momentum p^{μ}, and in a non comoving reference frame their relation is given by (Hobson, Efstathiou, and Lasenby 2006):

$$T_{\mu\nu} = (\rho + p)U^{\mu}U^{\nu} - pg_{\mu\nu} \tag{1.9}$$

and its expression of Equation 1.8 is then found by considering a comoving observer in $U^{\mu} = (1, 0, 0, 0)$ and by lowering the index with the metric.

1.1.2 Dynamical evolution of the universe

Thanks to how the metric and the energy-momentum tensor can be written under the Cosmological Principle, it is possible to describe the dynamical evolution of the universe and the different contributions of its components through the Einstein field equations as follows. The resulting equations have been first derived by A. Friedman in 1922 (Friedmann 1922) and are known as *Friedmann Equations*. Practically, this translates in considering the zeroth-component of Equation 1.5 and in substituting for the l.h.s the expressions obtained for the Ricci tensor and the scalar curvature in the FLRW metric, while using $T^{00} = \rho$ for the r.h.s, that leads to

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2} = \frac{8\pi G\rho}{3}, \tag{1.10}$$

which is the *first Friedmann equation*, where $\dot{a}$ indicates the derivative with respect to cosmic time t. It describes the evolution of the scale factor $a(t)$ stating that expansion depends only on the energy density ρ. The first ratio of the l.h.s. represents the *Hubble parameter*

$$H(t) = \frac{\dot{a(t)}}{a(t)} \quad [H] = [t^{-1}] \tag{1.11}$$

whose value today $H_0 = \dot{a}(t_0)$ (given that $a(t_0) = 1$) is given by the Hubble Constant

$$H_0 = 100\,h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}, \tag{1.12}$$

expressed above in terms of the dimensionless reduced Hubble parameter h. Equation 1.10 does not provide information on a possible expansion, but it depicts the velocity of the expansion in terms of the energy density and in dependence on the curvature. The same procedure for the ij^{th}-component of Equation 1.5 always assuming a perfect fluid approximation, leads to

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) = -\frac{4\pi G}{3}\rho(1+3w), \tag{1.13}$$

which is the *second Friedmann equation*, where the parameter $w = p/\rho$ relates the pressure of the fuid to its energy density, representing its *equation of state*. Equation 1.13 encodes the crucial information on whether the expansion is accelerating or decelerating. The first situation arises for $\ddot{a} > 0$ hence requiring the condition $1+3w < 0$, that is $w < -1/3$, while the second implies $w > -1/3$. This consideration provides important information about which component is predominant. In fact, from observations, today's expansion appears to be accelerated (Riess et al. 1998).

Universe energy-matter content

To describe the evolution of ρ with the scale factor one can combine *Friedmann equations*. In particular, by taking the derivative of Equation 1.10 and then using Equation 1.13 one can find

$$\dot{\rho} = -3(1+w)\frac{\dot{a}}{a}. \tag{1.14}$$

Then, considering different i components, each one of them is characterised by its equation of state, given by the solution of Equation 1.14:

$$\rho_i = \rho_{i,0}a^{-3(1+w)}, \tag{1.15}$$

where $\rho_{i,0}$ is the energy density of the i^{th} component at present time t_0. The FLRW model assumes these components to be three non-interacting components. Specifically, considering a system consituted by N particles with mass m, each one characterised by the four-momentum $p^{\mu}_{n}(t) = mu^{\mu}_{n}$, with $n = 1, .., N$:

Non-relativistic components: if the fluid is formed uniquely by non-relativistic matter, the kinetic energy is much smaller than the rest energy. Hence, the energy $E_n = \sqrt{|\boldsymbol{p}_n^2 + m^2|} \sim m + \frac{\boldsymbol{p}_n^2}{2m}$ and consequently its density becomes

$$\rho = \sum_n \Big\{ m + \frac{|\boldsymbol{p}_n|^2}{2m} \Big\} \delta^3(\boldsymbol{x} - \boldsymbol{x}_n(t)) = mn + \sum_n \frac{|\boldsymbol{p}_n|^2}{2m} \delta^3(\boldsymbol{x} - \boldsymbol{x}_n(t)) \tag{1.16}$$

where $n = \sum_n \delta^3(\boldsymbol{x} - \boldsymbol{x}_n(t))$ is the number density of particles. Using then the fact that for a perfect fluid $p = \sum_n \frac{|\boldsymbol{p}_n|^2}{3m} \delta^3(\boldsymbol{x} - \boldsymbol{x}_n(t))$, that leads to $\rho = mn + \frac{3}{2}p$. However, since the thermal energy, which is the source of pressure, is much less than the rest mass, in a non-relativistic fluid the pressure p is considered negligible with respect to the energy density ρ, i.e. $p \ll \rho$. For a *non-relativistic* fluid one can consider $p \sim 0$ and consequent equation of state $w = 0$.

Relativistic components: for a relativistic fluid, rest masses are null or sufficiently small compared to the kinetic energy to be considered negligible. This corresponds to the limit $\boldsymbol{p}_n \sim E_n$, which leads to

$$\rho = \sum_n E_n \delta^3(\boldsymbol{x} - \boldsymbol{x}_n(t)) \tag{1.17}$$

$$p = \sum_n \frac{|\boldsymbol{p}_n|^2}{3m} \delta^3(\boldsymbol{x} - \boldsymbol{x}_n(t)) \sim \frac{1}{3} \sum_n E_n \delta^3(\boldsymbol{x} - \boldsymbol{x}_n(t)) \tag{1.18}$$

and consequently to $p = \frac{1}{3}\rho$, namely to an equation of state given by $w = \frac{1}{3}$.

Cosmological constant Λ: writing Einstein's equations in the case in which the cosmological constant Λ is not absorbed into the energy-momentum tensor yields

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{R} = 8\pi G \Big\{ T_{\mu\nu} + \frac{\Lambda g_{\mu\nu}}{8\pi G} \Big\}, \tag{1.19}$$

where along with the matter-radiation tensor described so far, the energy-momentum tensor of the cosmological constant $T^{(\Lambda)}_{\mu\nu} = \Lambda g_{\mu\nu}$ plays the role of a source term, characterised by an energy density $\rho_\Lambda = T^{00}_{(\Lambda)} = \frac{\Lambda}{8\pi G} g^{00} = \frac{\Lambda}{8\pi G}$ and pressure p such that $T^{ij}_{(\Lambda)} = -p g^{ij}$. This leads to an equation of state for Λ given by $w = -1$.

Now, combining the above discussion with Equation 1.15, it is possible to describe the evolution of each component $i = (m, r, \Lambda)$, where m stands for matter (the non-relativistic component), r indicates radiation (the relativistic component), Λ describes the cosmological constant Λ. Hence, by making the equations of state explicit in Equation 1.15, one finds:

$$\rho_m(t) = \rho_{m,0} a^{-3}(t), \quad \rho_r(t) = \rho_{r,0} a^{-4}(t), \quad \rho_\Lambda = \rho_{\Lambda,0} = \text{const.} \tag{1.20}$$

As already mentioned when describing Equation 1.13, a fluid with an equation of state of $w < -1/3$ would drive an acceleration of the universe expansion. In particular, the cosmological constant Λ can be interpreted as a fluid with $w_\Lambda = -1$, in which case its energy density remains constant with time. However, in order to investigate the nature of dark energy, one can let its equation of state vary with time, to have a more general expression where ρ_{de} is different from ρ_Λ. In particular, its evolution can be expressed as a function of the more general equation of state $w_{de} = f(a)$ that presents several parametrisations. A common parametrisation is given by (Linder and Jenkins 2003):

$$w_{de}(a) = w_0 + (1-a)w_a, \tag{1.21}$$

that includes the case $w_{de} < -1/3$ which is compatible with the expansion of the universe and also recovers the simpler cosmological constant case for $w_0 = -1$ and $w_a = 0$. Now, by manipulating the first Friedmann equation Equation 1.10 to write it as

$$1 + \frac{K}{H^2 a^2} = \frac{8\pi G}{3H^2}\rho, \tag{1.22}$$

and defining the quantity in the. r.h.s as

$$\rho_c \equiv \frac{3H^2}{8\pi G}, \tag{1.23}$$

called the *critical density*, it is possible to introduce for each component $i =$ (matter, radiation, Λ) its corresponding density parameter, defined as

$$\Omega_i = \frac{\rho_i}{\rho_c}. \tag{1.24}$$

This leads to a total density parameter defined as

$$\Omega(a) = \Omega_\mathrm{m} a^{-3} + \Omega_\mathrm{r} a^{-4} + \Omega_\mathrm{de} a^{f(a)} \tag{1.25}$$

obtained by dividing the energy densities for each component by the critical density and by considering Equation 1.20 with a more general equation of state for dark energy: $w = f(a)$.

Then, Equation 1.22 expressed in terms of the density parameter Ω becomes $\frac{K}{H^2a^2} = \Omega - 1$, telling us that for $\Omega > 1$ we expect a universe with positive curvature $K > 0$, for $\Omega < 1$ a universe with positive curvature $K < 0$ and for $\Omega = 1$ a flat universe with $K = 0$.

1.1.3 Kinematics in an expanding universe

To deal with the evolution of the motion of freely-falling particles in a FLRW space-time, geodesic equations in a curved space-time can be obtained. From the General Relativity Equivalence Principle, or by applying the principle of least action to the action

$$\mathcal{S} = \int \sqrt{-g_{\mu\nu} dx^\mu dx^\nu}, \tag{1.26}$$

it is possible to derive the geodesic equations of motion (Weinberg 1972):

$$\frac{d^2x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0, \tag{1.27}$$

representing the evolution of the shortest possible trajectories of a particle in free fall between two points with respect to the affine parameter λ. The study of the radiation emitted by the objects populating the universe provides crucial information about their characteristics, hence representing a fundamental tool in cosmology. Specifically, when investigating the emission and propagation of such radiation, the expansion of the universe needs to be taken into account. Indeed, considering that photons are massless, their four-momentum $p^\mu = (E, p^i)$ follows the equation of motion $p_\mu p^\mu = 0$. This, combined with the Friedmann equations, relates the energy of the photon E the scale factor as $E \propto a^{-1}$ and hence $\lambda \propto a$ (consistently with the solution for the energy density of radiation in Equation 1.14). So, in the presence of expansion, the relation between the light emitted at a given time of emission t_e from a given position r_e, characterised by the wavelength λ_e is related to the wavelength of the same photon observed at time t_0 through

$$\frac{\lambda_e}{\lambda_0} = \frac{a_e}{a_0}. \tag{1.28}$$

In the absence of expansion $a_e = a_0$ and the wavelength of the light stays the same as it propagates through the universe. In the presence of expansion, the scale factor at t_0 is larger than it was when the light was emitted by the distant source, in which case $\lambda_0 > \lambda_e$, namely its spectrum is shifted towards the red. This shift towards the red is called *redshift*, and it is quantified by the parameter

$$z \equiv \frac{\lambda_0 - \lambda_e}{\lambda_e}. \tag{1.29}$$

Along with Equation 1.28, Equation 1.29 sets the relation between the measured redshift z of a distant source to the scale factor at the time of emission t_e:

$$a_e = \frac{1}{1+z}, \tag{1.30}$$

considering that $a(t_0) = 1$. Hence, the presence of a cosmological redshift is due to the expansion of the universe (which is encoded in the scale factor) and from the knowledge of the emission of well-known objects (such as the *standard candles*) at a certain redshift z it is possible to trace the value of the scale factor at the time of emission.

Distances

The redshift of distant sources can be linked to their distance through the cosmological model. Hubble 1929 observed a roughly linear relation between the velocity and the distance of extra-galactic nebulae. Such relation is known as Hubble's law : $v = H_0 r$. It can be recovered from Equation 1.12 and it is only valid for the local universe. On cosmological scales, the evolution of the Hubble parameter as a function of time needs to be taken into account.

Proper distance: The proper distance d_p is the distance seen by a photon. It describes the collection of infinitesimal light paths despite the expansion of the universe. Mathematically, it is defined as $d_p = cdt$. Since the light speed is a constant, the proper distance is simply given by

$$d_p(z_0, z_e) = c[t(z_0) - t(z_e)], \tag{1.31}$$

where z_0 and z_e are respectively the redshift at which the light ay if observed and emitted.

Comoving distance: The comoving radial distance (or simply comoving distance) χ is defined as $d\chi = \frac{cdt}{a}$, namely is the rescaling by the scale factor of the physical distance, such that two particles initially at rest in the evolving FLRW metric stay equidistant in comoving coordinates. As mentioned before, the cosmological model enables to relate the observed redshift of a distant source to its comoving distance with respect to the observer. In particular, by manipulating the Hubble parameter definition as $dt = \frac{da}{aH}$, it is possible to re-write $d\chi = \frac{cdt}{a}$ as

$$\chi(a) = \int_a^1 \frac{c da'}{a'^2 H(a')}, \tag{1.32}$$

where the extreme of the integral indicate that $\chi(a)$ is the comoving distance of an object given that we observe now (at t_0, when holds $a(t_0) = 1$) light that was emitted when the scale factor of the universe was equal to a. Through the scale factor-redshift relation Equation 1.30, this can be explicitly expressed with respect to the redshift as

$$\chi(z) = \int_0^z \frac{c dz'}{H(z')}. \tag{1.33}$$

Comoving transverse distance: The comoving transverse distance f_K is the ratio of the comoving separation between two points at χ to their separation angle. It only depends on the comoving distance χ and it is related with the curvature K as

$$f_K(\chi) = \begin{cases} \frac{1}{\sqrt{K}} \sin\left(\sqrt{K}\chi\right) & \text{for } K > 0 \\ \chi & \text{for } K = 0 \\ \frac{1}{\sqrt{-K}} \sinh\left(\sqrt{-K}\chi\right) & \text{for } K < 0. \end{cases} \tag{1.34}$$

Angular diameter distance: The angular diameter distance D_A is the ratio of a physical separation at the emission epoch (with comoving radial distance χ) to its angular separation from an observer. It can be obtained by scaling the comoving transverse distance to the emission epoch z_e. So, for an observer at $z_0 = 0$ and calling the emission epoch $z_e = z$, the angular diameter distance is defined as

$$D_\mathrm{A}(z) = \frac{f_K[\chi(z)]}{1+z}. \tag{1.35}$$

Luminosity distance: The luminosity distance is based on the scaling of the flux F of a distant source with its distance. It is defined so that the flux of the source at distance with an intrinsic luminosity L corresponds to:

$$D_\mathrm{L} = \sqrt{\frac{L}{4\pi F}}. \tag{1.36}$$

It can be expressed in terms of the transverse comoving distance f_K or in terms of the angular diameter distance as:

$$D_{\mathrm{L}}(z) = (1+z) f_K[\chi(z)] = (1+z)^2 D_{\mathrm{A}}(z) \tag{1.37}$$

for any observer at $z_0 = 0$ and sources at $z_e = z$.

1.2 Cosmological parameters

Given the tools introduced in the previous sections, it is now possible to define some cosmological parameters and their evolution as function of redshift.

1.2.1 Age of the universe

The first Friedmann Equation can be rewritten in terms of today's (at t_0) cosmological parameters and it is known as *Hubble equation*. In particular, it can be rephrased as

$$\dot{a}^2 = H_0^2 \left\{ \sum_i \Omega_i(t_0) a^{-(1+3w_i)} + [1 - \Omega(t_0)] \right\}. \tag{1.38}$$

By integrating the above equation, the evolution for the scale factor $a(t)$ is obtained and it is consequently possible to give an estimation of the expansion rate of the universe today. Considering instead the inverse function of the scale factor $t(a)$, one can derive the age of the universe, given a certain cosmological model. For instance, the calculation of t can be performed by considering the change of variables $dt = da/\dot{a}$ and by integrating it starting from an initial time (imposed to be $t_i = 0$) and supposing the scale factor at that time to be a point-like singularity, i.e. $a(t_i) = 0$. In this way the following expression for t is obtained as

$$t = \frac{1}{H_0} \int_z^\infty \frac{dz(1+z)^{-2}}{\left[\sum_i \Omega_i(t_0)(1+z)^{(1+3w_i)} + 1 - \Omega(t_0)\right]^{1/2}}, \tag{1.39}$$

where the scale factor-redshift relation as $da = -dz(1+z)^{-2}$ has been used. Equation 1.39 returns the age of the universe today when integrated between zero and infinity (corresponding to an integration between 0 and 1 with respect to the scale factor). The estimated (Planck Collaboration et al. 2020) age of the universe today is

$$13.801 \pm 0.024 \,\mathrm{Gyr}. \tag{1.40}$$

1.2.2 Hubble parameter

By simply using that $\dot{a}^2 = H^2 a^2$ in Equation 1.38, the following explicit expression for the *Hubble Parameter* as a function of the redshift is also obtained

$$H^2(z) = H_0^2 \left\{ \sum_i \Omega_i(t_0)(1+z)^{3(1+w_i)} + [1 - \Omega(t_0)](1+z)^2 \right\}, \tag{1.41}$$

so that the expansion rate at a given redshift z can be related to the energy density contribution of each component Ω_i, knowing its value today.

1.2.3 Density Parameter

The connection between the density parameter Ω_i and the redshift leads to the significant consideration that if the universe is flat today will remain so for any value of the redshift and, more generally, if the universe has a certain curvature, this will stay the same for every time. Indeed the density parameter can be formulated as

$$\Omega_i(z) = \frac{\Omega_i(t_0)(1+z)^{3(1+w_i)}}{\left[\sum_i \Omega_i(t_0)(1+z)^{3(1+w_i)+(1-\Omega(t_0))(1+z)^2}\right]}, \tag{1.42}$$

where Equation 1.23 and Equation 1.28 have been used.

1.2.4 Primordial power spectrum

The current theory explaining how structures as galaxies, galaxy clusters or the cosmic web that we observe today formed assumes the existence of small primordial perturbations, namely deviations from the mean density of the universe which slowly increased in amplitude because of the phenomenon of gravitational instability and formed the virialised structures observed at the present time. Among the theories proposed to explain the existence of these density perturbations, *inflation* (Guth 1981; Baumann 2012) predicts that these deviations from the background density field are the product of processes in the very early universe, corresponding to quantum fluctuations which expanded to macroscopic proportions during an extremely fast period of expansion. During this process, these quantum fluctuations reach macroscopic scales and will eventually act as seeds of large-scale structure formation. The scale dependence of such perturbations are described by the *primordial power spectrum*, defined as

$$\mathcal{P}_\zeta(k) = A_\mathrm{s}\left(\frac{k}{k_p}\right)^{n_\mathrm{s}-1} , \tag{1.43}$$

and parametrised as a power law of amplitude A_s and spectral index n_s, that measures the deviation from scale invariance ($n_\mathrm{s} = 1$) and k_p is a pivot scale. In the large-scale structure convention, it is common to normalise the primordial power spectrum of matter fluctuations at present time by requiring that the root mean square (r.m.s.) variance on a sphere of radius $R_8 = 8\,h^{-1}\,\mathrm{Mpc}$ is equal to a normalisation factor, called σ_8.

1.2.5 Sum of neutrino masses

Thanks to the discovery of oscillations among neutrino flavor eigenstates, neutrino masses m_ν are now known to be non-zero. This can be understood intuitively by considering the minimal neutrino mixing scenario where the existence of three species of neutrinos is assumed. The neutrino flavor eigenstates ν_α (with $\alpha = e, \mu, \tau$) can be expressed as a superposition of neutrino mass eigenstates ν_i (with i = 1, 2, 3) through the unitary transformation:

$$\langle\nu_\alpha\rangle = \Sigma_\alpha U_{\alpha i}\,\langle\nu_i\rangle\,, \tag{1.44}$$

and the concept of neutrino mixing is encoded in the unitary matrix $U_{\alpha i}$: if $U_{\alpha i}$ is the identity matrix, the flavor eigenstates coincide with the mass eigenstates and there is no oscillation. However, from solar and atmospheric neutrino experiments, compelling evidences for neutrinos oscillation have been provided (Group, Zyla, Barnett, et al. 2020) as it is observed that the massive states propagate differently between the source (in which neutrinos are produced) and the observer. One of the main quantities considered by these experiments to describe and quantify this phenomenon, is given by the probability that a neutrino of flavor α is later detected as a neutrino with flavor β. Such probability can be derived explicitly and its expression depends on the neutrino mass squared differences $\Delta m^2_\mathrm{ij} \equiv |m^2_\mathrm{i} - m^2_\mathrm{j}|$ (Maki, Nakagawa, and Sakata 1962; Pontecorvo 1968; Kajita 1999). Hence, neutrino oscillation experiments can provide information about the difference of the squared masses while remaining blind with respect to the absolute value of the neutrino mass and to the way in which these masses are ordered. This allows two scenarios in which the masses can be distributed: a *normal hierarchy* where the masses are ordered as $m_1 < m_2 < m_3$ and an *inverted hierarchy* corresponding to $m_3 < m_1 < m_2$. Moreover, from the combination of solar and atmospheric results, a hierarchy among the mass splittings can be derived as well, namely $\Delta m^2_{21} \ll |\Delta m^2_{31}| \sim |\Delta m^2_{32}|$. Assuming then that the lightest neutrino is massless, it is possible to obtain a lower bound from oscillation experiments on the sum of neutrino masses $\sum_\nu m_\nu \equiv M_\nu$ by taking the sum of the mass differences: in a *normal hierarchy* setting this value correspond to $M_\nu > 0.06$ eV and in a *inverted hierarchy* setting to $M_\nu > 0.1$ eV. Particle physics experiments (Wolf 2010; Esfahani, Asner, Böser, et al.

2017) are only sensitive to the mass of the lightest neutrino. Fortunately, massive neutrinos leave an imprint on the expansion history and growth of structure in the universe and this can be used to infer the sum of their masses M_ν. Specifically, the fact that neutrinos with mass less than a few eV remain relativistic and behave like radiation at the time of matter-radiation equality and became then non-relativistic behaving like matter today, delays the beginning of linear growth by moving the matter-radiation equality at a later time. Moreover, neutrinos with small masses have large thermal velocities, and can then stream out of cold dark matter potential wells, causing a suppression of the growth of structure below a certain scale. This scale is approximately equal to the Hubble radius when neutrinos are relativistic while when neutrinos become non-relativistic this scale corresponds to the point in which gravitational instability overcomes the thermal pressure of neutrinos. The strongest upper bound we have today on the sum of neutrino masses comes from cosmology, and it is specifically $M_\nu < 0.12$ eV, from the combination of temperature anisotropies in the cosmic microwave background (CMB), CMB lensing and Baryon Acoustic Oscillation (BAO) (Planck Collaboration et al. 2020).

1.2.6 ΛCDM model

Depending on the cosmological parameters, various cosmological models with different characteristics emerge. The one that best describes current observations is the Λ Cold Dark Matter model (ΛCDM). The name itself gives us information about its two main components: Λ being the cosmological constant appearing in Einstein Fields Equations of General Relativity, related to the simplest parametrisation of a fluid called *dark energy* (DE); and CDM standing for *cold dark matter*. This model is a mathematical parameterisation of Big Bang cosmology, as described by General Relativity and the Friedman-Lemaître-Roberson-Walker (FLRW) equations introduced in Section 1.1.1. It is based on the assumptions that the universe is composed of elementary particles, i.e. ordinary matter (baryons, leptons, photons, neutrinos), and cold (non-relativistic at the time of decoupling) dark matter, which only interacts gravitationally along with dark energy, responsible for the observed acceleration in the Hubble expansion. Dark energy is assumed to take the form of a constant vaccuum energy density, which should coincide with the cosmological constant introduced by Einstein in 1917. It requires only a small set of parameters (illustrated in Table 1.1), to reproduce all cosmological observations and completely specify the cosmological model. The requirement of space flatness comes from these parameters and it is in agreement with the observational data. Indeed, by rewriting the r.h.s. of Equation 1.22 in terms of the density parameter $\Omega_{\mathrm{TOT}} = \sum_i \Omega_i$ it can be seen how for a flat universe $(K = 0)$, Equation 1.22 requires that $\Omega_{\mathrm{TOT}} = 1$. By looking at the values of the energy density parameters for the different components in Table 1.1, one can see how in the ΛCDM paradigm, the vast majority of the total energy-matter density budget is dominated by the cosmological constant Λ, accounting for $\sim 69\%$ of the total energy-matter density budget. The remaining 31% is encoded in the matter density Ω_{m}, whose contributions are respectively given by a 26.2%

of cold dark matter (encoded in the parameter $\Omega_{\rm c}$) and 4.8% of ordinary matter, usually called in cosmology *baryonic* matter (encoded in $\Omega_{\rm b}$). The contribution of Ω_Λ is associated in the ΛCDM framework with the dark energy, responsible for the accelerated expansion of the universe (Riess et al. 1998; Perlmutter et al. 1999), because of its negative pressure, introduced through its equation of state in Section 1.1.2. The necessity for dark matter, the second largest contribution to the total energy content of the universe, was historically invoked to explain the mass deficit of visible matter from observations (Zwicky 1937) and the observations of flat galaxy rotation curves to large distances (Rubin and Ford 1970). Also, the combined weak lensing and X-ray observation of the *Bullet Cluster* provided as well a strong hint of the presence of dark matter through the offset found between the center of the total mass and the center of the baryonic mass peaks of the clusters system (Clowe et al. 2006).

Parameter	Symbol	Value
Hubble parameter	H_0	$(67.66 \pm 0.42)\ \mathrm{km\,s^{-1}Mpc^{-1}}$
Total matter density	$\Omega_{\rm m}$	0.3111 ± 0.0056
Dark matter density	$\Omega_{\rm c}h^2$	0.11933 ± 0.00091
Baryon density	$\Omega_{\rm b}h^2$	0.02242 ± 0.00014
Dark energy density	Ω_Λ	0.6889 ± 0.0056
Power spectrum normalisation	σ_8	0.8102 ± 0.0060
Spectral index	$n_{\rm s}$	0.9665 ± 0.0038
Reionisation optical depth	τ	0.0561 ± 0.0071
Sum of neutrino masses	M_ν	< 0.12 eV

Table 1.1: Some of the main cosmological parameters defining the ΛCDM model from Planck Collaboration et al. 2020 obtained by the combination of CMB TT, TE, EE + lowE + CMB lensing + BAO.

These two dark components, dark matter and dark energy, dominate the current energy-matter content of the universe. However, their nature and properties are still unknown and one of the main goals of the current and future cosmological surveys is exactly to try to determine them, by comparing theory and observations and by performing cosmological parameter inference, as will be described in the next sections, through the imprints these components leave on the cosmological observables.

1.3 Principles of structure formation

The previous section dealt with a homogeneous and isotropic universe, constituted by a perfectly smooth cosmological fluid. However, from both simulations an observations we know that structures as galaxies, galaxy clusters or the cosmic web actually exist. As briefly anticipated in Section 1.2.4, the current theory explaining how those structures formed assumes the existence of small primordial deviations from the mean density which slowly increase in amplitude because of the phenomenon of gravitational instability and form the virialised structures observed at the present time.

1.3.1 Density contrast in the linear regime

The above mentioned deviations from the background can be considered already at the level of the FLRW metric defined in Equation 1.2 by adding to it a perturbative term, in the limit where these perturbations are sufficiently small. This perturbative term propagates in the matter density which becomes the sum of a background density $\bar{\rho}_{\mathrm{m}}(t)$ term and a perturbation term $\delta\rho_{\mathrm{m}}(\boldsymbol{r},t)$. The deviation from the background density is then summarised in the quantity

$$\delta(\boldsymbol{r},t)=\frac{\delta\rho_{\mathrm{m}}(\boldsymbol{r},t)}{\bar{\rho}_{\mathrm{m}}(t)}=\frac{\rho_{\mathrm{m}}(\boldsymbol{r},t)-\bar{\rho}_{\mathrm{m}}(t)}{\bar{\rho}_{\mathrm{m}}(t)}, \tag{1.45}$$

called the *matter density contrast.* At early times and on large scales, this deviation is very small and it is fair to work in the limit $\delta \ll 1$ that allows to derive the evolution of the density field using linear perturbation theory. The dynamical evolution of the density ρ can be described through a *Vlasov-Poisson* system of equations (Peebles 1993) and consequently the *density contrast* δ as well. By combining these equations and working under the assumption of an ideal pressureless fluid, in the linear regime the evolution for the density contrast is given by (Dodelson 2003)

$$\ddot{\delta}+2H\dot{\delta}=4\pi G\bar{\rho}_{\mathrm{m}}\delta, \tag{1.46}$$

which provides a good description of the density fluctuations evolution in the early epochs of the universe, when density fluctuations are still small. The r.h.s. comes from the *Poisson* equation and its the source term governing the evolution of the density contrast, while the second term of the l.h.s. takes into account the universe expansion through the Hubble parameter H. This ordinary second order differential equation admits the general solution of the form (Peebles 1980; Peacock 1999):

$$\delta(t,\boldsymbol{r})=D_{+}(t)\delta_{+}(\boldsymbol{r})+D_{-}(t)\delta_{-}(\boldsymbol{r}), \tag{1.47}$$

where the spatial and temporal coordinates can be decoupled and D_+ and D_- are respectively a *growing* and *decaying* mode. Since the decaying mode will disappear as time goes by, it is not relevant at late times and in this context it can be neglected. Hence, the density contrast can then be expressed using the growing mode whose general solution in terms of the scale factor is given by

$$D_+(a) = \frac{5\Omega_\mathrm{m} H(a)}{2H_0} \int_0^a \frac{da'}{a'^3} \left(\frac{H_0}{H(a')} \right)^3 , \tag{1.48}$$

and it is conventionally normalised such that $D_+(a) = a$ during the matter-dominated era and it is referred to as the *growth factor*. It is convenient to consider its normalised expression as

$$D(a) = \frac{D_+(a)}{D_+(a=1)}, \tag{1.49}$$

to ensure that its value at present time is $D_+(a=1) = D_+(z=0) = 1$ so that the density contrast in the linear regime is given by $\delta(\boldsymbol{r}, z) = D(z)\delta(\boldsymbol{r})$.

1.3.2 Matter power spectrum

A consequence of the cosmological principle is that the density contrast defined in Equation 1.45 is a statistically isotropic and homogeneous field. To be able to perform a quantitative comparison between the theory of structure formation and the actual distribution of the structures that we observe, statistical measures of the density field are needed. A natural way to do it is to invoke the *two-point correlation function* (2PCF) of the field, which is given by

$$\langle \delta(\boldsymbol{r}')\delta(\boldsymbol{r}+\boldsymbol{r}') \rangle_{\boldsymbol{r}'} = \int_{\mathbb{R}^3} \delta(\boldsymbol{r}+\boldsymbol{r}')\delta(\boldsymbol{r}')d\boldsymbol{r}', \tag{1.50}$$

and describes the probability of finding two objects separated by the distance $\boldsymbol{r}$ at the positions $\boldsymbol{r}'$ and $\boldsymbol{r}'+\boldsymbol{r}$ with respect to the probability one would have if those objets were simply scattered independently with uniform probability. Considering the 3D Fourier transform of the density field $\tilde{\delta}(\boldsymbol{k}) = \frac{1}{(2\pi)^3} \int d^3\boldsymbol{r} e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} \delta(\boldsymbol{r})$, it is possible to write the 2PCF of this field in Fourier space as

$$\langle \tilde{\delta}(\boldsymbol{k})\tilde{\delta}^\star(\boldsymbol{k}') \rangle = (2\pi)^3 P_\delta(\boldsymbol{k}) \delta_\mathrm{D}^{(3)}(\boldsymbol{k}-\boldsymbol{k}'), \tag{1.51}$$

where $\delta^\star(\boldsymbol{k})$ is the complex conjugate and $\delta_\mathrm{D}^{(3)}$ is the Dirac delta and $P_\delta(\boldsymbol{k})$ is the 3D matter power spectrum. Given the assumption of homogeneity and isotropy, the dependence on the scale is actually only $P_\delta(\boldsymbol{k}) = P_\delta(|\boldsymbol{k}|)$, with $|\boldsymbol{k}| = k$. Then, it is possible to relate the matter power spectrum $P_\delta(k)$ to the primordial power spectrum of initial fluctuations defined in Equation 1.43, through a *transfer function* $\mathcal{T}(k)$ as for example (Eisenstein and Hu 1998)

$$\mathcal{T}(k) = \frac{\tilde{\delta}(k, a=1)\tilde{\delta}(k=0, a=0)}{\tilde{\delta}(k, a=0)\tilde{\delta}(k=0, a=1)}, \tag{1.52}$$

such that the processed power-spectrum $P_\delta(k)$ is related to its primordial form via $P_\delta(k) \propto \mathcal{T}^2(k)P_\zeta(k)$, accounting for the modulation of each Fourier mode between their original value at $a = 0$ to their current value at $a = 1$. The actual computation of the the power spectrum at sufficiently large scales can be performed by solving linear perturbation theory equations. In particular, it is possible to derive some analytical expressions for the transfer function, either relying on some approximation methods or fitting formulae as in Eisenstein and Hu 1998. Alternatively by employing Boltzmann codes such as CAMB (Lewis, Challinor, and Lasenby 2000a) or CLASS (Lesgourgues 2011). At smaller scales, entering the non-linear regime where $\delta \sim 1$, perturbation theory is no longer applicable and approaches based on N-body simulations are needed. To explore the non-linear gravitational evolution, these N-body simulations are then used to obtain phenomenological halo models or fitting formulae of the non-linear gravitational clustering. For instance, Takahashi et al. 2012, provided a revised prescription of the HALOFIT model of Smith et al. 2003 whose main idea is based on the assumption that all the matter content in the universe is assumed to be bound in dark matter halos. Then the power spectrum can be decomposed into two terms, one accounting for the correlation between two distinct halos, that is dominant at large scales, and another term describing matter correlations within the same dark matter halo, which is dominant at small scales. Boltzmann solvers as CLASS can then be used to compute the linear power spectrum as mentioned before and also to compute the non-linear power spectrum including non-linear corrections computed through the HALOFIT code. It is very important to remind that the functional form for the power spectrum is then based on the halo model, but the model parameters are calibrated from the N-body simulations in consideration. As can be seen by looking at Figure 1.1, the impact of these non linear correction leads to an enhancement of the power at small scales. Moreover, as will be described more in details in Section 2.2.5, for future surveys such as *Euclid* and Rubin Observatory Legacy Survey of Space and Time (LSST) that will have a very high precision and very low statistical uncertainties, it is very important to control and be able to properly model the non-linearities in the matter power spectrum with high accuracy to get accurate and unbiased cosmological constraints from weak gravitational lensing measurements.

1.4 Brief overview of cosmological probes

As described in Section 1.2.6, the nature and properties of the vast majority of the components of our universe are still unknown and from Section 1.2.5 one can see how even in the Standard Model of particle physics there are still some open questions that need

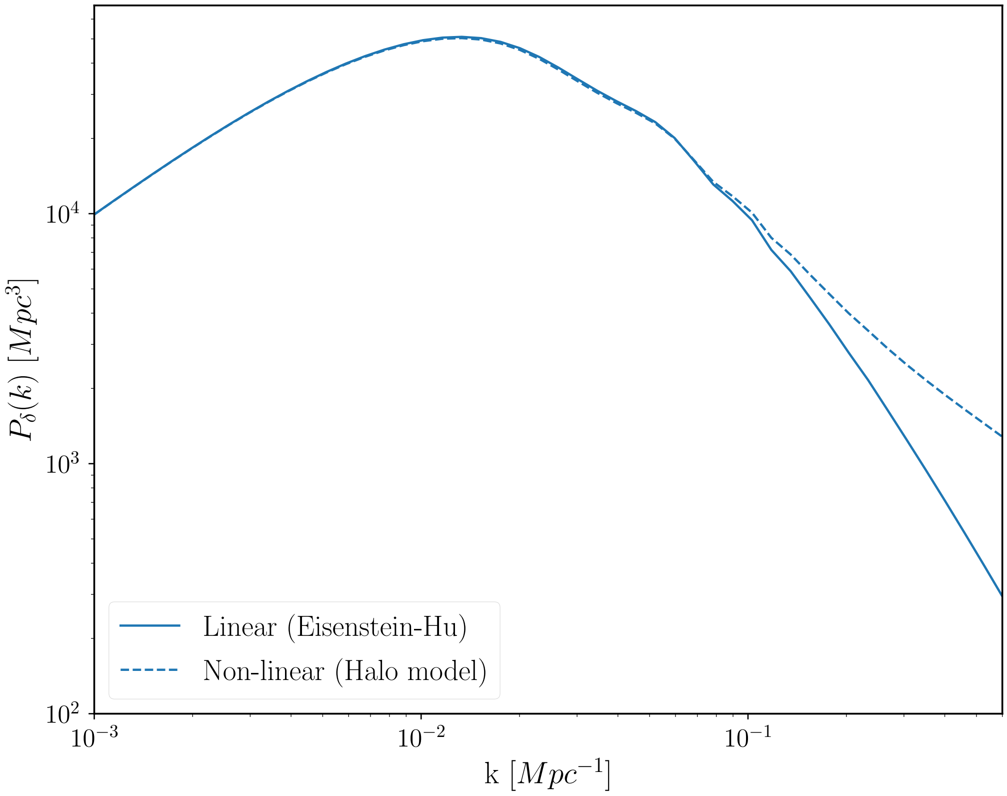


Figure 1.1: Illustration of the linear matter power spectrum (continuous line) and non-linear power spectrum (dashed line) obtained with `pyccl`. The linear one is computed with Eisenstein and Hu 1998 approximation and the non-linear one with the halofit model of Takahashi et al. 2012.

an answer. Fortunately, the behaviour and characteristics of dark energy, dark matter and the sum of neutrino masses, leave an imprint on cosmological observables and this can be used to infer cosmological parameters, with the aim of trying to answering those questions. As an example, Figure 1.2 shows the fractional difference matter power spectrum between power spectra of models including massive neutrinos (the different colours corresponding to different values of the masses) and the power spectrum with massless neutrinos as a function of comoving wavenumber k. The same plot indicates also the sensitivity in the power spectrum for different probes depending on the scale.

For instance, the Cosmic Microwave Background radiation (CMB) (Penzias and Wilson 1965), that is the result of the decoupling of photons from electrons at the epoch of recombination, can be used as a probe to constrain the neutrino mass. Indeed, for low enough

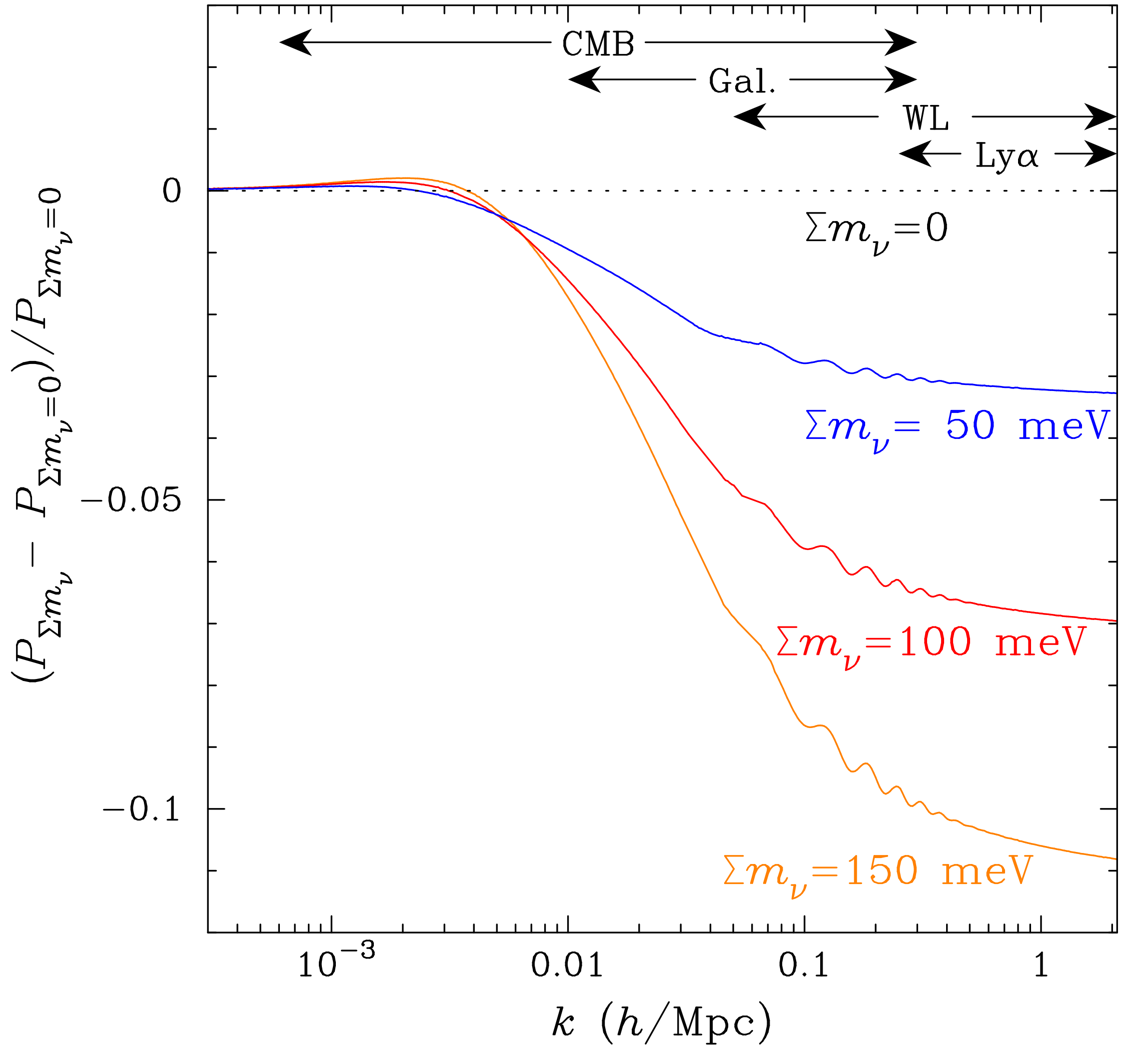


Figure 1.2: Plot from Abazajian et al. 2015, showing thee fractional difference of the matter density power spectrum as a function of comoving scale k for different values of $\sum_\nu m_\nu$.

masses (such that they became non-relativistic after recombination), the angular diameter distance that is degenerate with decreasing H_0 changes, impacting the CMB power spectrum. Small secondary temperature fluctuations induced in the CMB due to the evolution of the gravitational potentials right after recombination can be used to constrain relativistic species, hence also the number of neutrinos and their masses. Moreover, Planck Collaboration et al. 2014 moved the analysis to a regime where the dominant effect comes from the gravitational lensing of the CMB by the large scale structure. Indeed, CMB lensing leads to a direct measurement of the matter power spectrum on scales where the effects of neutrino masses become relevant. More generally, the decrease in temperature due to the expansion of the universe makes the neutrinos transition from being a ultra-relativistic gas behaving as radiation to a cold gas behaving as dark matter. This results in a suppression of power on scales smaller than the neutrinos free-streaming scale. This scale is associated with the scale at which the matter power spectrum peaks, caused by the transition from the radiation to the matter dominated era, when perturbations cross the cosmological horizon. This can be "seen" through the CMB at large scales and by large scale structures probes at small scales. In general, measurements of tracers of the matter clustering in the large-scale structure can be used to constrain cosmological parameters. The combination with absorption lines in the spectra of distant galaxies and quasars arising from the Lyman-α electron transition of the neutral hydrogen due to intervening low density gas, more simply referred to as Ly-α forest (Weinberg et al. 2003) has enabled to provide one of the strongest inferred constraints on neutrino mass for example (Seljak, Slosar, and McDonald 2006). The detection of the 21 cm line from high-redshift neutral hydrogen (HI) represents a powerful probe that with the potential of mapping the distribution of matter until the epoch of reionisation and is the goal of several present and upcoming radio telescopes as for instance the Square Kilometer Array (SKA) (Santos et al. 2015; Combes 2021). The next two sections, are dedicated to two specific probes, weak lensing and galaxy clustering, that I have used in some of the works described in this thesis and that represent the two main cosmological probes that will be employed by *Euclid*.

1.4.1 Galaxy Clustering

Another tracer of the underlying matter distribution is given by the distribution of galaxies. Indeed, galaxies as well as dark matter haloes are sampling the matter density field, however not in a uniform way, in the sense that the spatial clustering of galaxies do not precisely represent the clustering of the totality of the matter in the universe. Galaxies are hence considered as a *biased* tracer of the underlying total matter density field (Kaiser 1984; Bardeen et al. 1986). Following the same philosophy of defining a matter density contrast to investigate the evolution of the overdensity with respect to the background to eventually compute a statistical measure of the field as the 2PCF or the power spectrum, it is possible to define the fluctuations in the galaxy number density field $n(\boldsymbol{x};z)$ in the linear regime as

$$\delta_{\mathrm{g}}(\boldsymbol{x};z) = \frac{n(\boldsymbol{x};z) - \bar{n}(z)}{\bar{n}(z)} = b_{\mathrm{g}}(z)\delta(\boldsymbol{x};z), \tag{1.53}$$

related to the matter density contrast defined in Equation 1.45 through the galaxy bias $b_{\mathrm{g}}(z)$. At large scales, it is fair to assume the bias to be scale independent and to model it uniquely as a function of the redshift. Under these conditions, from Equation 1.53 follows that the galaxies power spectrum can be expressed in terms of the matter power spectrum as

$$P_{\mathrm{g}}(k,z) = b_{\mathrm{g}}^2(z)P_{\delta}(k,z). \tag{1.54}$$

It is important to underline that more complex bias modelling are required, in order to include effects such as primordial non-gaussianities, the presence of massive neutrinos, dark energy, modified gravity and redshift space distortions, as these factors can lead to a more complicated expression for the bias, making it scale-dependent (Desjacques, Jeong, and Schmidt 2018). The discussion of these effects goes beyond the purposes of this work, however their role is non-negligible in current and future cosmological analysis and a huge scientific effort from the community to properly model them is currently ongoing. To map the distributions of the galaxies it is possible to measure their frequency distribution of observed photons per galaxy. Then, through *spectroscopy*, knowing the absorption and emission lines in the rest-frame and obtaining the redshifted ones, it is possible to determine the position of each galaxy in terms of right ascension (RA), declination (DEC) and redshift $z_{\rm spec}$. Thanks to spectroscopic instruments as the Sloan Digital Sky Survey (SDSS), it has been possible to create an extremely detailed 3D map of the universe, using deep multi-color images for one third of the sky, and spectra for more than 3 millions astronomical objects have been retrieved (Ahn et al. 2012) and the future Dark Energy Spectroscopic Instrument (DESI) data releases will obtain spectra and measure the positions and receding velocities of $\sim$ 40 million galaxies (DESI Collaboration et al. 2016). The medium-class ESA mission *Euclid* will enable to measure up to $\sim$ 30 millions spectroscopic redshifts (Pozzetti, L. et al. 2016; Euclid Collaboration et al. 2020b) through the spectroscopic channel of its Near Infrared Spectrometer and Photometer (NISP) instrument. An alternative way to estimate the galaxy redshifts without involving the spectra, consists in measuring the galaxy flux in a number of standard filters that select a relatively broad passband of colours to determine the redshift that is then related to the distance through the Hubble law, defined in Section 1.1.3. The redshifts estimated in this way are referred to as *photometric* redshifts. These measurements lead to the probes for the clustering of galaxies that depending on the technique used are referred to as *Spectroscopic Galaxy Clustering* (GCsp) and *Photometric Galaxy Clustering* (GCph) that along with weak gravitational lensing (see Section 1.4.2) constitute the primary probes of the *Euclid* mission. Section 6.5.2 provides a slightly more detailed description of Photometric Galaxy Clustering, used in one of the projects of this thesis (Chapter 6).

1.4.2 Weak Gravitational Lensing

Another powerful cosmological probe to constrain cosmological parameters is the bending of light emitted by distant background galaxies due to the presence of foreground large-scale structure between the sources emitting the light and the observer. This phenomenon is widely used in current cosmology and is called *Weak Gravitational Lensing* (WL). Indeed, the distortions induced by the presence of the foreground large-scale structure on background galaxy shapes provide a direct probe of the matter power spectrum $P_\delta(k)$ (Kaiser 1992). This is because the distortions, even if small in the weak regime, are induced by the matter distribution along the path of the emitted photons. Hence, it is possible to retrieve information on the properties of this matter distribution, given enough galaxies across the sky. Moreover, since both baryonic and dark matter contributes to this effect, weak lensing traces directly *all* matter, in contrast for example to the distribution of galaxies, that as seen in the previous section, needs to invoke a galaxy bias to retrieve the dark matter distribution using galaxies. The observable signal related to this effect encodes an anisotropic deformation, called the *shear* along with an isotropic magnification, called the *convergence*, of the original source galaxy shape. So what it is actually observed is galaxy shapes, precisely, the observed ellipticity of the galaxy that in the weak regime can be expressed as the sum of its intrinsic ellipticity plus the shear induced by weak lensing. Details of weak lensing formalism, definitions and state of the art results are illustrated in Chapter 2, which is entirely dedicated to weak lensing, as it represents one of the main focus of this thesis in the framework of cosmological parameter inference. As mentioned above, weak lensing constitutes one of the primary probe of future galaxy survey, like *Euclid*. Its visible (VIS) instrument will measure the shapes of 1.5 billions galaxies covering $15000\ \mathrm{deg}^2$ of the sky, up to redshifts of about $z_\mathrm{s} = 2$. The near infrared photometry introduced in the previous section, combined with the VIS data will enable to derive photometric redshifts and rough estimates of these galaxies. Past and current galaxy surveys as the Canada-France-Hawaii Telescope Lensing Survey (CFTHLenS) (Heymans et al. 2012), the Kilo-Degree Survey (KiDS) (Hildebrandt et al. 2016) or the Dark Energy Survey (DES) (Zuntz et al. 2018) have also used and are at present using weak lensing as main cosmological probe and released the associated data products that led to the delivery of some of the most recent results in cosmology (DES Collaboration, Abbott, et al. 2021). In the coming years, along with *Euclid*, also the Vera C. Rubin Observatory Large Synoptic Survey Telescope (LSST) will survey a remarkable fraction of the sky ($\sim 20000\ \mathrm{deg}^2$). Both experiments will deliver source galaxy samples consisting of billions of galaxies and with the combination of deep and wide-field optical photometry that will be provided by LSST along with the wide-field optical photometry, NIR photometry and high resolution of *Euclid* it is crucial to optimise and try to maximise the amount of information they will bring.

1.5 Parameter inference for future galaxy surveys

In order to exploit the richness in information provided by the cosmological probes described in the previous section and be able to measure and estimate the parameters of a certain cosmological model, given some observed data, statistical methods are needed. This section presents an introduction to the main tools in a Bayesian inference framework useful for the purposes of this thesis based on (Heavens 2009; Verde 2010)[1].

---

[1] and on the knowledge about this topic I have acquired during the class on Bayesian Statistics by Professor Benjamin Wandelt.

1.5.1 Bayesian inference

In the Bayesian framework, probabilities are interpreted as degree of belief in a hypothesis. Given a random variable (event) x and denoting the probability density of such event to take a specific value as $\mathcal{P}(x)$, then the probability that the random variable x takes a value between x and $x+dx$ is given by $\mathcal{P}(x)dx$. Being a probability, $\mathcal{P}(x)$ has to be defined within the range $0 \leq \mathcal{P}(x) \leq 1$ and it has to satisfy the condition $\int_{-\infty}^{\infty} \mathcal{P}(x)dx = 1$. Moreover, for mutually exclusive events x_1 and x_2 holds the relation: $\mathcal{P}(x_1 \cup x_2) = \mathcal{P}(x_1) + \mathcal{P}(x_2)$. Considering now the probability of x_1 *and* x_2 to happen, i.e. the *joint* probability density function (PDF) $\mathcal{P}(x_1, x_2)$, this is given by the probability of x_1 times the *conditional* probability of x_2 given that x_1 has already happened $\mathcal{P}(x_2|x_1)$, namely $\mathcal{P}(x_1, x_2) = \mathcal{P}(x_1)\mathcal{P}(x_2|x_1)$. If x_1 and x_2 are independent events, then $\mathcal{P}(x_2|x_1) = \mathcal{P}(x_2)$ and the joint PDF reduces to $\mathcal{P}(x_1, x_2) = \mathcal{P}(x_1)\mathcal{P}(x_2)$. Based on $\mathcal{P}(x_1, x_2) = \mathcal{P}(x_2, x_1)$ the following identity is a direct consequence:

$$\mathcal{P}(x_1, x_2) = \mathcal{P}(x_1)\mathcal{P}(x_2|x_1) = \mathcal{P}(x_2)\mathcal{P}(x_1|x_2). \tag{1.55}$$

For the purposes of this work, it is now interesting to embed the previous discussion in a potential application for cosmology. If we consider a given cosmological model $\mathcal{M}$ characterised by a set of cosmological parameters $\boldsymbol{\Theta} = \{\theta_1, \theta_2, ..., \theta_n\}$, assuming that we have observed a set of data d, we can rely on Equation 1.55 to write the following identity for the conditional probability density function of the cosmological parameters $\boldsymbol{\Theta}$, given the data $\boldsymbol{d}$:

$$\mathcal{P}(\boldsymbol{\Theta}|\boldsymbol{d}, \mathcal{M}) = \frac{\mathcal{P}(\boldsymbol{\Theta}|\mathcal{M})\mathcal{P}(\boldsymbol{d}|\boldsymbol{\Theta}, \mathcal{M})}{\mathcal{P}(\boldsymbol{d}|\mathcal{M})}. \tag{1.56}$$

This identity is known as *Bayes Theorem* and it regulates the relation between the conditional probability density function of the parameters $\boldsymbol{\Theta}$ given the observed data $\boldsymbol{d}$, with the conditional probability density function of observing the data $\boldsymbol{d}$ given a model $\mathcal{M}$ characterised by the parameters $\boldsymbol{\Theta}$ and given a precedent knowledge on the the parameters Θ, that might be given for instance by some previous experiments, and on the data. $\mathcal{P}(\boldsymbol{\Theta}|\boldsymbol{d}, \mathcal{M})$ is called the *posterior* probability distribution, $\mathcal{P}(\boldsymbol{\Theta}|\mathcal{M})$ the *prior*, $\mathcal{P}(\boldsymbol{d}|\boldsymbol{\Theta}, \mathcal{M})$ the *likelihood* and $\mathcal{P}(\boldsymbol{d}|\mathcal{M})$ the *evidence*. With the final goal of determining the posterior distribution for the parameters of a given cosmological model, one needs to define a prior on the parameters, assume a likelihood and define the evidence. The latter, that is the probability of the data before it is known, is fundamental for model comparison but it can be considered as a normalisation constant for parameter inference, in a way that the problem to determine the posterior distribution becomes:

$$\mathcal{P}(\boldsymbol{\Theta}|\boldsymbol{d}) \propto \mathcal{P}(\boldsymbol{\Theta})\mathcal{P}(\boldsymbol{d}|\boldsymbol{\Theta}), \tag{1.57}$$

where to simplify the notation the assumption of the model $\mathcal{M}$ has been omitted. Concerning the *prior*, it can provide a definite information about a certain variable or parameters of a given model or general information about such variable. In the first case it is called *informative prior* and in the second *uninformative prior*. To give an example of the first scenario, in many situations we have knowledge of the parameters from some previous experiment and this information can enter the inference analysis through $\mathcal{P}(\boldsymbol{\Theta})$. For instance, the state-of-the art of cosmological parameters inference from previous experiments such as COBE (Smoot 1999), WMAP (Hinshaw et al. 2013), Planck (Planck Collaboration et al. 2020), SDSS (Tegmark, Strauss, Blanton, et al. 2004), provides a rich and solid theoretical cornerstone for current and future analysis. In the absence of any previous information, the prior can be assumed to be a constant and it is called a *flat prior*. In summary, the prior represents the state of belief on the parameters that will be inferred, prior to any influence by the new observed data $\boldsymbol{d}$.

1.5.2 Likelihood

Once the prior choice is made, the fundamental ingredient that needs to be determined to evaluate the posterior probability distribution for the parameters is the *likelihood* function $\mathcal{P}(\boldsymbol{d}|\boldsymbol{\Theta})$, as one can see from Equation 1.57. The likelihood can have an analytical expression when this is known or it can be estimated through density estimation problems belonging to the class of methods known as *likelihood free inference* (LFI) (Papamakarios and Murray 2018; Alsing et al. 2019), Section 1.5.4 briefly discusses this point. A very common choice for the likelihood function in cosmological analysis is the assumption of a Gaussian likelihood relying on arguments such as the central-limit theorem or the principle of maximum entropy. In these cases, when the likelihood is assumed to be Gaussian, it is convenient to work with its logarithm, namely

$$\log \mathcal{L}(\boldsymbol{d}|\boldsymbol{\Theta}) = \frac{1}{2}\left[(\boldsymbol{d}-\mu(\boldsymbol{\Theta}))^T \boldsymbol{C}^{-1} (\boldsymbol{d}-\mu(\boldsymbol{\Theta}))\right], \tag{1.58}$$

where $\boldsymbol{d}$ indicates the observed data, $\boldsymbol{\Theta}$ the cosmological parameters defining the model, $\boldsymbol{C}$ is the covariance matrix encoding the information about the data errors and the correlation among observables at different scales, and $\mu(\boldsymbol{\Theta})$ is the observable coming from the model as a functions of the parameters $\boldsymbol{\Theta}$. From this expression of the likelihood, one can see how the key ingredients needed to build the likelihood are *a)* the covariance matrix, *b)* the model.

Covariance matrix estimation

Concerning the covariance matrix, this can be computed analytically (following again some assumptions as for example a Gaussian covariance as for example in Euclid Collaboration et al. 2020b) or it can be computed from N-body simulations as for example in Harnois-Déraps et al. 2018, when a Gaussian assumption does not hold anymore for example and the non Gaussian terms become more difficult to estimate. For the purposes of this work, I will focus on the case where the covariance is assumed to be cosmology independent, namely when it is computed for a fixed set of cosmological parameters. Nevertheless, it is important to briefly mention that a way to accurately assess the uncertainty on the measurements relies on the computation of the covariance using a large ensemble of independent N-body simulations at each of the cosmologies that can be sampled in parameters space. However, this requires that the number of simulations per ensemble is significantly larger than the data vector and this scenario needs computing resources far exceeding those currently available. As mentioned before, an alternative approach is to run an ensemble of full N-body simulations at a single cosmology and ignore the variation of the covariance with cosmology. This case also requires a large number of independent simulations with respect to the dimension of the summary statistics and the data array, as in practice it enters the likelihood function with a correction factor, known as the *Hartlap factor* (Hartlap, Simon, and Schneider 2007) that takes into accounts the loss of information due to the finite number of bins $n_{\rm bins}$ and realisations N as

$$C^{-1} = \frac{N - n_{\rm bins} - 2}{N - 1} C_*^{-1}, \tag{1.59}$$

where the element of the covariance matrix C_* are given by

$$C_*^{ij} = \sum_{r=1}^{N} \frac{(x_i^r - \bar{x}_i)(x_j^r - \bar{x}_j)}{N - 1}, \tag{1.60}$$

and x_i^r is the value of the observable in the i^{th} bin for a given realisation r and $\bar{x}_i = \frac{1}{N} \sum_r x_i^r$ is the mean of such observable in a given bin over all the realisations. By looking at Equation 1.59 one can see then how the more the pre-factor is close to one (i.e. $N \gg n_{\rm bins}$) the more the estimated covariance is close to the truth and the uncertainty is smaller. Sellentin and Heavens 2017 have quantified the loss of information associated to this pre-factor for DES, KiDS and *Euclid* settings and in particular they shown how for KiDS-450 and DES SV, the uncertainty due to the covariance matrix is subdominant with respect to parameter degeneracies and the finite data set that are the dominating drivers of the uncertainties of cosmological parameters; for *Euclid* they found a loss of 1% of information for 1900 simulations, as compared to knowing the true the covariance matrix. In the works presented in this manuscript, the condition $N \gg n_{\rm bins}$ is always satisfied, thanks to the large suite of independent N-body simulations employed in the different studies. However, it is important to point out that this is not always the case for weak lensing analysis, depending on the available simulations. When the condition $N \gg n_{\rm bins}$ is not satisfied, the formalism to follow to properly account for all biases to be removed from the parameter inference is the one presented in Sellentin and Heavens 2016 and Sellentin and Heavens 2017. Moreover, as at the late-times tracers of the large-scale structure enter the non-linear regime, deviations from a Gaussian distribution of their PDF at smaller scales increase and it is necessary to properly consider in the analysis the non-Gaussian contributions such as the supersample covariance (SSC) (Hu and Kravtsov 2003). This extra term arises from the non-linear impact caused by density fluctuations that have wavelengths that are greater than the survey size and modify the local observables such that the background density averaged over the survey size is not representative of the averaged density in the universe. This correction is crucial when performing cosmological analysis as it is expected to be the dominant source of statistical error or cosmic variance in the context of weak lensing weak lensing (Barreira et al. 2018). In the works presented in this thesis, this term is not included, but it is important to be aware of this correction for future extensions, considering that ways to take it into account exist or are currently in development, as for example the approximation presented by Lacasa and Grain 2019, that makes fast numerical computation of the SSC possible.

Building the model with Gaussian Processes

The next ingredient needed to build the likelihood defined in Equation 1.58 is the model $\mu(\boldsymbol{\Theta})$, namely a prediction for the observable used to perform the analysis as a function of the cosmological model. When an analytical prescription of the observable is possible to compute, one can use it to directly predict the observable as a function of cosmological parameters. There are several codes that provide an estimate for certain observable and enable to get a theoretical prediction, for example CAMB (Lewis, Challinor, and Lasenby 2000b), `CLASS` (Lesgourgues 2011), `pyccl` (Chisari et al. 2019) for the matter power spectrum or lensing power spectrum, NICAEA (Kilbinger et al. 2017), for calculating second-order correlation functions, and in general pipeline and softwares are available to obtain prediction for cosmological observable as a function of cosmology (e.g. CosmoSIS (Zuntz et al. 2015), nbodykit (Hand et al. 2018), `PyCosmo` (Refregier et al. 2018)). However, there are a number of situations in which an analytical prescription is not always available or presents some drawbacks with respect to effects that can impact the prediction (systematics, masks..). Weak lensing peak counts, that will be introduced and discussed in Section 2.3.5, for example can be used as cosmological observable to derive constraints on parameters. Analytical models to obtain an estimation of their predictions exist, but present some downsides as it will be described in Section 2.3.5. Hence, one of the alternative ways to get the theoretical predictions in these situations relies in employing Gaussian Processes Regression (GPR). Gaussian processes can be seen as an infinite-dimensional multivariate Gaussian distribution or as Gaussian distributions over functions such that they are completely defined by a mean function $m(\theta)$ and a covariance function $K(\theta, \theta')$, that in this context is dubbed the *kernel* and $f(\theta) \sim GP(m(\theta), K(\theta, \theta'))$. Then, the Gaussian Processes act by an assumption of smoothness, i.e. by considering that for a new point in parameter space θ_* which is sufficiently close to a known point θ belonging to the training set, the corresponding observable will be described by a joint normal distribution along with the known observable. This can be summarised by:

$$\begin{bmatrix} f \\ f_* \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m \\ m_* \end{bmatrix}, \begin{bmatrix} K(\theta,\theta) + \sigma_n^2 \mathbb{I} & K(\theta,\theta_*) \\ K(\theta_*,\theta) & K(\theta_*,\theta_*) \end{bmatrix}\right),$$

where $K(\theta, \theta')$ is the *kernel* of the Gaussian processes that assesses the smooth relation among points in parameter space and for the works presented in this thesis we take its form to be an *anisotropic squared exponential* function. σ_n is the standard error of the noise level in the targets. Then, the prediction can be retrieved by making use of the *Multivariate Gaussian Theorem*, so that it is simply defined by its posterior conditional $\mathcal{P}(f_*|\theta_*, \theta, f) = \mathcal{N}(f_*|m_*, \Sigma_*)$, completely defined by its conditional mean and covariance (Rasmussen and Williams 2005):

$$\begin{aligned} m_\star &= m(\theta_*) + K(\theta_*,\theta)[K(\theta,\theta) + \sigma_n^2 \mathbb{I}]^{-1}(f - m(\theta)), \\ \Sigma_* &= K(\theta_*,\theta_*) - K(\theta_*,\theta)[K(\theta,\theta) + \sigma_n^2 \mathbb{I}]^{-1} K(\theta,\theta_*). \end{aligned}$$

The predictions are then exactly given by the conditional means $m_\star$, and variances can be obtained directly taking the diagonal elements of the covariance matrix Σ_*.

1.5.3 Explore parameter space with MCMC

Now that what is needed to build the likelihood is defined, the next step is to explore the parameter space to find the parameter range that provides theoretical predictions $\mu(\boldsymbol{\theta})$ that best fit the data $\boldsymbol{d}$, with the final goal of deriving the posterior probability distribution defined in Equation 1.56. One way to do this relies on Markov Chain Monte Carlo (MCMC). Given a generic probability distribution $p(\theta)$, called the *target* distribution, the MCMC samples the parameter space through a random walk that is initialised at a given point θ_0. A common choice for such point is a value (or a set of n values for n-dimensional parameter space) very close to where it is expected that the likelihood is centred. In the case of forecasts for instance, given a fiducial model, one can initialise the random walk in a Gaussian ball with a tiny radius around the fiducial parameters values. Now, to start the parameter space exploration one needs to determine the next position in parameter space $\tilde{\theta}_{i+1}$. To do this, one can follow the Metropolis-Hastings (MH) algorithm (Metropolis et al. 1953; Hastings 1970). The main idea is based on the fact that a new point in parameter space θ_{i+1}

can be explored by defining a *proposal* probability distribution function $q(\tilde{\theta}_{i+1}, \theta_i)$ that only depends on the previous point θ_i. Then, by defining the following quantity

$$a = \min\left(1, r \equiv \frac{p(\tilde{\theta}_{i+1})q(\theta_i|\tilde{\theta}_{i+1})}{p(\theta_i)q(\tilde{\theta}_{i+1}|\theta_i)}\right), \tag{1.61}$$

if the *acceptance ratio* parameter $r < 1$, the step to the value $\tilde{\theta}_{i+1}$ is *accepted* and $\theta_{i+1} = \tilde{\theta}_{i+1}$ with probability a; if $r > 1$, the proposed value $\tilde{\theta}_{i+1}$ is *rejected* and $\theta_{i+1} = \theta_i$. Then, the procedure is repeated iteratively by doing $i \to i+1$ until the set $\{\theta_i\}_{i=1,..N}$ converges to the a sample of the target $p(\theta)$, in the limit $N \to \infty$. While being a cornerstone of MCMC methods, the MH algorithm presents some limits when dealing with high dimensional parameter space as well as many different models (Akeret et al. 2013). Alternative ways within the MCMC framework have hence been proposed, such as the *Affine invariant ensemble sampler*. First introduced by Goodman and Weare 2010 and then implemented by Foreman-Mackey et al. 2013, this method considers an *ensemble* of walkers spread across the parameter space of the target distribution to explore the parameter space. At each iteration, the walkers are split into groups and are randomly assigned to a partner subgroup chosen from the walkers ensemble and what is proposed as the next step is a random point on a line that links their positions. Then the walkers of one subgroup are updated using the other half of the walkers as reference. Formally, following Goodman and Weare 2010, the proposed position $\tilde{\theta}_{i+1}$ is given by

$$\tilde{\theta}_{i+1} = \theta_i^m + z(\theta_i^k - \theta_i^m) \tag{1.62}$$

where the indexes m, k indicate walkers belonging to the different subgroups and z is a value sampled from a fixed distribution $q(z) = \frac{1}{\sqrt{z}}$ for $z \in [\frac{1}{a}, a]$ and 0 elsewhere, with a being a tuning parameter. Then, the proposed position is accepted if

$$z^{n-1}\frac{p(\tilde{\theta}_{i+1})}{p(\theta_i^m)} \geq r, \tag{1.63}$$

where $r \in [0, 1]$ is a random number and n is the dimension of the target. The implementation of this method by Foreman-Mackey et al. 2013 led to the python package `emcee`. One of the advantages of this affine invariant ensemble sampling of the MCMC is that it can be parallelised since a large number of positions are updated simultaneously at each iteration of the process. In the works presented in this thesis, as some of the simulations used have a vast number of different cosmological models, parameter inference is performed with the package `emcee` to exploit the advantages of parallelisation. A variety of different methods to perform MCMC are however present such as Gibbs sampling, Hamiltonian Monte Carlo and others. A detailed review can be found in Sharma 2017; Robert and Changye 2020.

A corresponding variety of codes and softwares are as well available and widely employed for inference in cosmology as for example `CosmoHammer` (Akeret et al. 2013), `CosmoMC` (Lewis and Bridle 2002), `Cobaya` (Torrado and Lewis 2021)..

1.5.4 Brief mention of likelihood-free inference and ABC methods

Even if the analysis presented in this thesis are all based on a MCMC approach, it is important to mention that exist alternative ways that during the recent years have drawn the attention of the community to perform inference. The main motivation behind the introduction of these methods is to find a solution to situations in which the likelihood function is intractable. Indeed, current cosmological analyses are typically based on the assumption that the likelihood is Gaussian. However, it has been shown that non-Gaussianities are present even for underlying Gaussian fields (e.g. the lensing field) and if neglected by assuming a Gaussian likelihood, they can induce biases on the parameter constraints (Sellentin and Heavens 2017). Moreover, MCMC methods can present some limitations in terms of speed when dealing with very high-dimensional parameter space. In general, to try to overcome the problems arising when the likelihood is unknown and/or too many (in terms of speed) likelihood calls are necessary to obtain well sampled posteriors, the concept of Approximate Bayesian computation (ABC) (Rubin 1984) and density estimation Likelihood Free-Inference (DELFI) (Papamakarios and Murray 2016), have been introduced in cosmological analysis. The main idea behind ABC methods is to draw parameters from some proposal density, then forward simulate data that are compared to the observed data under some distance metric. The proposed parameters are then accepted if this distance is below some threshold value ϵ and the accepted samples provide the samples from an approximate posterior, that recovers the exact posterior one only under the limit $\epsilon \rightarrow 0$. While making it possible to probe the likelihood without a direct evaluation of it, this accept/reject criterion present some disadvantages. As example, as said before these methods recover the exact posterior only in the limit $\epsilon \rightarrow 0$, but they critically slow down in this limit, and what happens in practice is that they draw samples from a distribution that is broader than the posterior (Papamakarios and Murray 2016). Hence, DELFI methods have been introduced to overcome this limitation. In this framework, inference is treated as a density estimation task and the sampling distribution of the data is learned by neural density estimators (NDEs) as a function of the parameters of the model and by adaptively acquiring simulations with active learning. Very schematically, first forward simulations are run to obtain a set of parameters θ and the corresponding data d. Then, following some compression procedure (e.g. Alsing and Wandelt 2018), the D dimensional data are compressed to informative summaries $d \rightarrow t$, where t has the same dimension of the parameter space. From the simulated mocks, the sampling distribution $\mathcal{P}(t|\theta)$ is then learned by the NDE and evaluated at the observed data t_0, providing the likelihood. Some examples of LFI implementations that can be used in cosmological analysis are `PyDelfi`, `BOLFI` and others, that have been proven to potentially be a robust alternative for cosmological inference (Jeffrey, Alsing, and Lanusse 2020; Alsing, Wandelt, and Feeney 2018). Both classes of methods (ABC and LFI) have reached a vast extension in their theoretical development and applications that goes beyond the purposes of this thesis, and a single section won't suffice to provide an extensive overview. However, it is worth to mention these approaches, as even if they are still in a early epoch of their usage (validations and diagnostics of such methods are a very topical object of current research) their potential applications on the enormous amount of data that we will receive from next generation cosmological surveys could bring many advantages.

1.6 Summary

In this Chapter, I defined some background tools that are needed to understand the following Chapters. Starting from the cosmological principle, main cosmological quantities have been derived and our current cosmological model, ΛCDM, has been introduced. Along with its definition, the unsolved questions concerning the two main components of our universe, dark matter and dark energy have been introduced. Then statistical tools as the two point correlation function and the matter power spectrum have been defined, and their connection with cosmological probes as the CMB, weak lensing, galaxy clustering.. has been presented. Lastly, the Bayesian inference framework has been outlined, to show how given a model and some observed data, it is possible to infer the parameters of the model through techniques such as MCMC. The next Chapter will *zoom in* one particular cosmological probe, weak gravitational lensing and will provide more details on how this probe is employed to constraints cosmological parameters.

CHAPTER 2

WEAK GRAVITATIONAL LENSING

This Chapter discusses more in details weak gravitational lensing. The presentation starts with the derivation of lensing equations from light deflection. Then, after introducing the convergence and the shear, the link between observations and cosmological models is illustrated. Second order statistics of cosmic shear are then defined, followed by a brief overview of systematics effects in weak lensing and some of the current constraints. Next, I provide an overview of the state of the art of summary statistics in weak lensing.

2.1 Lensing formalism

The following summarises the most relevant quantities for the purposes of this thesis for weak gravitational lensing based on Seitz, Schneider, and Ehlers 1994; Bartelmann and Schneider 2001; Bacon, Refregier, and Ellis 2000; Refregier 2003; Kilbinger 2015.

2.1.1 Light deflection and the lens equation

There are different ways to derive the equations describing the bending light due to the presence of massive bodies. For example one can derive them directly from the first principles of General Relativity (GR). In particular, we know that in GR the trajectory of light is defined by null geodesics. Hence, given any metric, light deflection from gravitational lensing can be derived from the geodesic equations (introduced in Equation 1.27 and illustrated again for convenience)

$$\frac{d^2 x^\mu}{d\lambda^2} + \Gamma^\mu_{\alpha\beta} \frac{dx^\alpha}{d\lambda} \frac{dx^\beta}{d\lambda} = 0, \tag{2.1}$$

where $\Gamma^\mu_{\alpha\beta}$ are the Christoffel symbols for a given metric and λ the affine parameter. Now, following Seitz, Schneider, and Ehlers 1994, we can consider two arbitrary light rays, where we set $\gamma^\mu_{\text{fid}}(\lambda)$ to be the fiducial and $\gamma^\mu(\lambda, \boldsymbol{\theta})$ to be the one separated by the angular coordinates $\boldsymbol{\theta} = (\theta_1, \theta_2)$ from the fiducial. Given this setting, the quantity of interest to formalise the phenomenon of gravitational lensing is the physical separation between the two rays $\xi(\boldsymbol{\theta}, \lambda) = (x_1, x_2)$ as it depends on the contents of the Universe they pass through, at any arbitrary point $\lambda > 0$. Specifically, the components of $\xi(\boldsymbol{\theta}, \lambda)$ are the transverse components, perpendicular to the tangential direction x_3 of the reference ray in the physical 3D space, as shown in Figure 2.1. The evolution of the physical separation $\xi(\boldsymbol{\theta}, \lambda)$ is then described through the *geodesic deviation equation*, that reads:

$$\frac{d^2 \xi(\boldsymbol{\theta}, \lambda)}{d\lambda^2} = \mathcal{T}(x_1, x_2, x_3)\xi(\boldsymbol{\theta}, \lambda), \tag{2.2}$$

where $\mathcal{T} = \mathcal{T}^{\text{bg}} + \mathcal{T}^{\text{cl}}$ is the optical tidal matrix, accounting for the deviation due to the isotropic background Universe encoded in $\mathcal{T}_{\rm bg}$ and the contribution to the deviation due to local inhomogeneities, or clumps, described in $\mathcal{T}_{\rm cl}$. By assuming a metric, these two last terms can take an explicit form. We work in the weak field limit that implies that the potential is small $\Phi \ll c^2$ as well as the angles through which the light's path is deviated. Under this assumption, in the case of a FLRW metric, Seitz, Schneider, and Ehlers 1994 found the explicit expressions for the two terms constituting the optical tidal matrix to be

$$(T^{\rm bg})_{ij} = \frac{4\pi G}{c^2}\frac{\bar{\rho}_0}{a^5}\delta_{ij}, \tag{2.3}$$

$$(T^{\rm cl})_{ij} = -\frac{1}{c^2 a^2}\left(2\frac{\partial^2}{\partial x_i \partial x_j} + \delta_{ij}\frac{\partial}{\partial x_3}\right)\Phi, \tag{2.4}$$

with $i, j = \{1, 2\}$ being the indeces of the transvere components x_1, x_2 and x_3 being the l.o.s. direction. From now on, we work under the *thin-lens* approximation, namely we consider that inhomogeneities are geometrically thin and the variation along the x_3 direction vanishes, following Seitz, Schneider, and Ehlers 1994. To obtain the *lens equation* one can define the *comoving separation* $\boldsymbol{x} = (x_1, x_2) \equiv \xi/a$, and by explicitly plugging the sum of Equation 2.3 and Equation 2.4 into Equation 2.2 one can obtain the evolution for the *comoving separation* $\boldsymbol{x}$ in comoving coordinates χ as:

$$\frac{d^2\boldsymbol{x}}{d\chi^2} + K\boldsymbol{x}(\chi) = -\frac{2}{c^2}\left[\nabla_\perp \Phi(\boldsymbol{x}, \chi) - \nabla_\perp \Phi(0, (\chi))\right], \tag{2.5}$$

where $\nabla_\perp = (\partial/\partial x_1, \partial/\partial x_2)$ is the comoving transverse gradient, χ the comoving coordinate, and the space-time curvature K arise from the use of Friedmann equation before passing to the comoving reference frame. The solution of Equation 2.5 obtained with boundary conditions $\boldsymbol{x}(\chi = 0) = 0$ and $(\partial \boldsymbol{x}/\partial \chi)|_{\chi=0} = \boldsymbol{\theta}$ is given by

$$\boldsymbol{x}(\chi) = f_K(\chi)\boldsymbol{\theta} - \frac{2}{c^2}\int_0^\chi f_K(\chi - \chi')\left[\nabla_\perp \Phi(\boldsymbol{x}(\chi'), \chi') - \nabla_\perp \Phi(0, \chi')\right] d\chi'. \tag{2.6}$$

Now that we have the expression for the comoving separation $\boldsymbol{x}(\chi)$ we can derive the *lens equation* by considering that in absence of lensing the comoving separation would be seen by the observer under an angle $\beta(\boldsymbol{\theta}, \chi) = \boldsymbol{x}(\chi)/f_K(\chi)$.

Hence, the difference between the angular separations of two light rays in unperturbed and perturbed Universe, at comoving distance χ gives us the *deflection angle* $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha}(\boldsymbol{\theta}, \chi) = \boldsymbol{\theta} - \boldsymbol{\beta}(\boldsymbol{\theta}, \chi) = \frac{2}{c^2}\int_0^\chi \frac{f_K(\chi - \chi')}{f_K(\chi)}\left[\nabla_\perp \Phi(\boldsymbol{x}(\chi'), \chi') - \nabla_\perp \Phi(0, \chi')\right] d\chi', \tag{2.7}$$

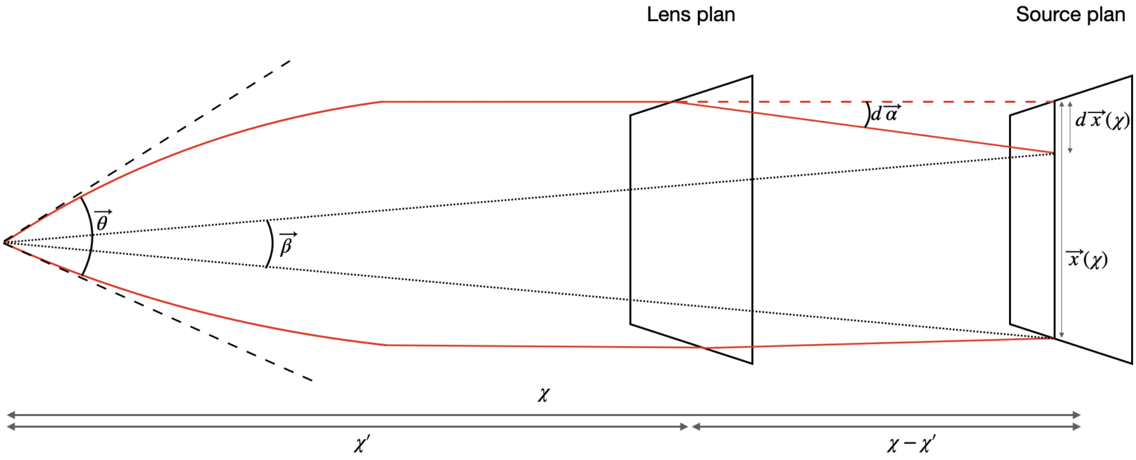


Figure 2.1: Illustration of the propagation of two light beams emitted by a galaxy in the source plan through a cluster located in the lens plan. As described in the text $\boldsymbol{\theta}$ is the angular separation between the observed apparent positions of the source, $\boldsymbol{\beta}$ is the angular separation between the true (unlensed) positions of the source, and $\boldsymbol{\alpha}$ is the deflection angle, representing the difference between the two. χ is the comoving distance between the observer and the source plan and χ' is the comoving distance between the observer and the lens plan. Image credits to Guinot 2020.

that using $\beta(\boldsymbol{\theta}, \chi) = \boldsymbol{x}(\chi)/f_K(\chi)$ basically corresponds to dividing Equation 2.6 by the comoving angular diameter distance $f_K(\chi)$. Assuming that that the potential evaluated on the perturbed light path does not differ substantially from the one on the unperturbed line of sight, i.e. working under the *Born approximation*[1], it is possible to substitute the comoving separation $\boldsymbol{x}(\chi)$ in Equation 2.6 with its 0^{th} order solution $\boldsymbol{x}_0(\chi) = f_K(\chi)\boldsymbol{\theta}$ and Equation 2.6 becomes:

$$\boldsymbol{\alpha}(\boldsymbol{\theta}, \chi) = \boldsymbol{\theta} - \boldsymbol{\beta}(\boldsymbol{\theta}, \chi) = \frac{2}{c^2} \int_0^{\chi} \frac{f_K(\chi - \chi')}{f_K(\chi)} \left[\nabla_{\perp}\Phi(f_K(\chi')\boldsymbol{\theta}, \chi') - \nabla_{\perp}\Phi(0, \chi')\right] d\chi'. \quad (2.8)$$

The above expression is the *lens equation* under the *Born approximation* and it encodes the information on the deviation of the apparent (observed) position of a source at $\boldsymbol{\theta}$ with respect to its true position $\boldsymbol{\beta}(\boldsymbol{\theta}, \chi)$ due to the gravitational potential Φ between the source and the observer. An illustration of the above description can be found in Figure 2.1.

---

[1] Within the *Born approximation*, Shapiro and Cooray 2006 proved using numerical simulations that the corrections up to fourth order in the gravitational potential of large-scale structure on the weak lensing power spectra are unimportant (namely the approximation to be accurate to sub-percent on most scales) for any weak lensing survey, including for a full sky survey limited by cosmic variance.

2.1.2 Shear and convergence fields

Given the *lens equation* derived in the previous section, it is possible to quantify the effect of lensing by mapping the lensed space to the unlensed space. This can be done by defining a linear *distortion matrix* $\mathcal{A}$ that describes the variation of the true position $\boldsymbol{\beta}$ with respect to the observed position $\boldsymbol{\theta}$ trhough the Jacobian $\mathcal{A} \equiv \partial\boldsymbol{\beta}/\partial\boldsymbol{\theta}$, whose elements are explicitly given by

$$\mathcal{A}_{ij}(\boldsymbol{\theta},\chi) = \delta_{ij} - \frac{\partial\alpha_i}{\partial\theta_j} = \delta_{ij} - \frac{2}{c^2}\int_0^{\chi} \frac{f_K(\chi-\chi')f_K(\chi')}{f_K(\chi)} \frac{\partial^2}{\partial x_i \partial x_j}\Phi[f_K(\chi')\boldsymbol{\theta},\chi']d\chi', \quad (2.9)$$

where it has been used that $\nabla_\perp = (\partial/\partial x_1, \partial/\partial x_2)$ and that $\boldsymbol{x}_0(\chi) = f_K(\chi)\boldsymbol{\theta}$ to express the partial derivatives with respect to the comoving transverse components x_1 and x_2. It is useful now to define the *lensing potential* as

$$\psi(\chi,\boldsymbol{\theta}) = \frac{2}{c^2}\int_0^{\chi} \frac{f_K(\chi-\chi')}{f_K(\chi)f_K(\chi')}\Phi[f_K(\chi')\boldsymbol{\theta},\chi']d\chi', \quad (2.10)$$

such that the distortion matrix can be expressed as

$$\mathcal{A}_{ij}(\boldsymbol{\theta},\chi) = \delta_{ij} - \frac{\partial^2\psi(\chi,\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j}. \quad (2.11)$$

We can already see from the equation above how the mathematical formalism starts to give us an intuition of the physics behind the distortion: the impact of lensing on the unlensed scenario that corresponds to the identity encoded in the Krönecker delta δ_{ij} is quantified by the variation of the lensing potential ψ that physically defines how much the gravitational potential Φ arising from a mass distribution changes the direction of the light path.

It is possible now to do a step further and define some fundamental quantities in weak lensing that make even more intuitive the effect of lensing. More specifically, the distortion matrix $\mathcal{A}$ can be parametrised as

$$\mathcal{A} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} = (1-\kappa)\begin{pmatrix} 1-g_1 & -g_2 \\ -g_2 & 1+g_1 \end{pmatrix}, \quad (2.12)$$

where (γ_1, γ_2) are the components of a spin-2 field $\gamma = \gamma_1 + \gamma_2 = |\gamma|e^{2i\phi}$ called *shear*, and κ is a scalar field called *convergence* and the two are related through the *reduced shear* $g_i(\boldsymbol{\theta},\chi) \equiv \gamma_i(\boldsymbol{\theta},\chi)/(1-\kappa)$. The reduced shear is a complex field that can as well be expressed as a function of its modulus and its phase as $g = g_1 + g_2 = |g|e^{2i\phi}$. The *shear* and the

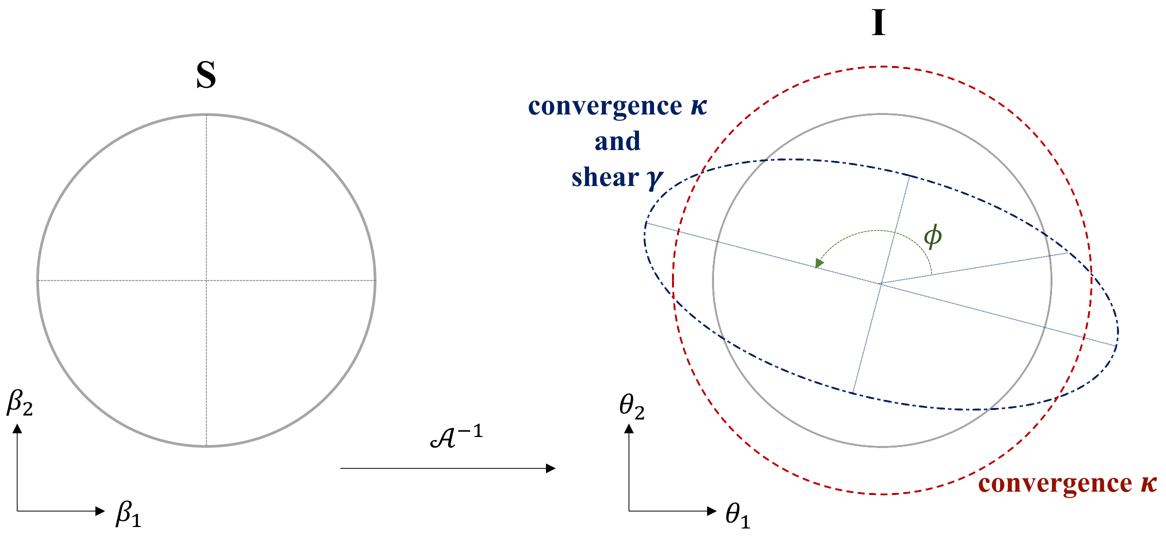


Figure 2.2: Illustration of the mapping between unlensed and lensed space thruogh the distortion matrix $\mathcal{A}$. Given an original circular source image **S** the effect of weak lensing leading to the observed image **I** can be summarised in the action of 1) the shear γ whose effect is to imprint an anisotropic stretching (example shown in blue) where its norm $|\gamma|$ characterizes the flatness of the ellipse and the rotation angle ϕ indicates the distortion direction; 2) the convergence κ that isotropically magnifies or contracts the original shape (example shown in red). This is formalised in the distortion matrix $\mathcal{A}$, being the Jacobian between the original coordinates $\boldsymbol{\beta}$ and the observed coordinates $\boldsymbol{\theta}$.

convergence describe respectively the anisotropic stretching and the isotropic magnification of the galaxy shape due to lensing. Equation 2.11 defines the shear and the convergence fields as second-order derivatives of the lensing potential:

$$\gamma_1 \equiv \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\psi; \quad \gamma_2 \equiv \partial_1\partial_2\psi \tag{2.13}$$

$$\kappa \equiv \frac{1}{2}(\partial_1\partial_1 + \partial_2\partial_2)\psi = \frac{1}{2}\nabla^2\psi. \tag{2.14}$$

With the tools introduced so far, it is now possible to understand the physical meaning of the *shear* and the *convergence*, the relation between the two, and how they are related with the lensing potential. This can be intuitively understood by looking at Figure 2.2. Following Lin 2016 and looking at Figure 2.2, given a circular source **S** of radius $\boldsymbol{\beta}$, the actual observed distorted image **I** can be obtained through the local mapping of the source light distribution to image coordinates given by the inverse of the Jacobian distortion matrix $\mathcal{A}^{-1}$, as

$$\boldsymbol{I} = \mathcal{A}^{-1}\boldsymbol{S} = \frac{\beta}{(1-\kappa)(1-|g|^2)} \begin{pmatrix} \cos t + |g| \cos(2\phi - t) \\ \sin t + |g| \sin(2\phi - t) \end{pmatrix} \tag{2.15}$$

with $t \in \mathbb{R}$ parametrising the circumference $\mathbf{S}$. Applying now a rotation of $-\phi$, the rotated image is given by:

$$\boldsymbol{I}_{\text{rot}} = \begin{pmatrix} \cos(-\phi) & \sin(-\phi) \\ -\sin(-\phi) & \cos(-\phi) \end{pmatrix} \boldsymbol{I} = \begin{pmatrix} \cos(t+\phi)\frac{\beta}{(1-\kappa-|\gamma|)} \\ \sin(t+\phi)\frac{\beta}{(1-\kappa+|\gamma|)} \end{pmatrix}, \tag{2.16}$$

that is the equation of a parametrized ellipse with semi-axes $a = \frac{\beta}{(1-\kappa-|\gamma|)}$ and $b = \frac{\beta}{(1-\kappa+|\gamma|)}$. In the absence of shear ($\gamma = 0$), the equation of a circumference is recovered, telling us that the scalar field κ acts as isotropic transformation. If the shear is non zero, its norm $|\gamma|$ characterizes the flatness of the distortion and the rotation angle ϕ gives the distortion direction.

2.1.3 From observation to cosmology

Now that the formalism to depict the main weak lensing quantities useful for this thesis has been outlined, the next step is to understand the connection between what we actually observe through galaxy surveys with what we can say about cosmology. This can be done by finding the relation between the galaxy observed ellipticities, that is what we actually measure with a survey like *Euclid* for example, with the shear, that allows to quantify the effect of lensing. Then, by linking the shear to the convergence, that as will be illustrated in the following section can be related to the matter density field, it is possible to infer information about the matter distribution between the observer and the source and finally, about the underlying cosmological model.

Estimating shear from galaxy shapes

In the weak regime, the values for the convergence and the shear are of the order of a few percent or less ($|\kappa| \ll 1$ and $|\gamma| \ll 1$). There are no multiple images caused by lensing and each source is mapped uniquely into one image and the matrix $\mathcal{A}$ is invertible. The quantity that is actually measured by galaxy surveys is the observed galaxy ellipticity ϵ^{obs} and it's related to the intrinsic ellipticity of the source galaxy ϵ^{s} for $|g| \leq 1$ through (Seitz and Schneider 1997):

$$\epsilon^{\text{obs}} = \frac{\epsilon^{\text{s}} + g}{1 + g^* \epsilon^{\text{s}}} \quad \text{for} \quad |g| \leq 1, \tag{2.17}$$

where g^* is the complex conjugate of the reduced shear $g = g_1 + ig_2$. In the weak regime this relation can be approximated to $\epsilon^{\rm obs} \approx \epsilon^{\rm s} + \gamma$. The variance of the observed ellipticity is then given by $\sigma^2(\epsilon^{\rm obs}) = \sigma^2(\epsilon^{\rm s}) + \sigma^2(\gamma)$. The variance of the intrinsic ellipticity of the source is usually dominant and it is referred to as *shape noise*. In order to isolate the lensing signal and reduce the shape noise, a very large number of galaxies are required. Under the assumption that the intrinsic ellipticity of galaxies has no preferred orientation (that implies that the expectation value of the intrinsic ellipticities vanishes $\langle\epsilon^{\rm s}\rangle = 0$), the expectation value of the observed ellipticity becomes:

$$\langle\epsilon^{\rm obs}\rangle \approx g \approx \gamma, \tag{2.18}$$

meaning that averaging the ellipticity of many galaxies in a circular area and assuming that the shear is constant over this area, one should obtain an unbiased estimate of the shear itself (provided that there are no systematics in the shape measurement). For the moment, we work under the approximation described above, leading to Equation 2.18. Hence, what is directly measurable is the shear $g \approx \gamma$, while the convergence κ is not, and inversion methods are necessary to retrieve κ from γ.

Projected overdensity and convergence

As mentioned at the beginning of this section, we are walking through the steps to understand the path that starting from the observations lead us to the cosmological information. Having introduced one unbiased estimator of the observed ellipticity through Equation 2.18, the following step is now to provide an interpretation for the other fundamental weak lensing quantity vastly used in weak lensing analysis as well as in this work: the convergence κ. From Equation 2.16 and Figure 2.2 we have seen how the convergence κ is related to an isotropic increase or decrease of the observed size of a source image. Moreover, Equation 2.14 shows how the convergence field can be obtained as the second derivative of the lensing potential. This last point was mentioned *en passant*, but it actually contains all we need to connect the convergence to the cosmological information that we want to extract. Indeed, if we look more carefully into it, we see that the equation relating the lensing potential and the convergence is a *Poisson equation*. We remind here few useful steps to underline why this is important for the aim of this section. By making Equation 2.10 explicit in Equation 2.14 we get:

$$\kappa(\boldsymbol{\theta}, \chi) \equiv \frac{1}{2}\nabla^2\psi(\boldsymbol{\theta}, \chi) = \frac{1}{c^2}\int_0^{\chi} \frac{f_K(\chi - \chi')}{f_K(\chi) f_K(\chi')} \nabla^2\Phi[f_K(\chi')\boldsymbol{\theta}, \chi'] d\chi', \tag{2.19}$$

and using now the fact that the gravitational potential Φ is related to the matter density contrast $\delta = \Delta\rho/\bar{\rho}$ through the Poisson equation $\nabla^2\Phi = 4\pi G a^2 \bar{\rho}_{\rm m}\delta$, the convergence field can be rewritten as

$$\kappa(\boldsymbol{\theta}, \chi) = \frac{3H_0^2\Omega_{\rm m}}{2c^2} \int_0^{\chi} \frac{f_K(\chi - \chi')f_K(\chi')}{f_K(\chi)} \frac{\delta[f_K(\chi')\boldsymbol{\theta}, \chi']}{a(\chi')} d\chi', \tag{2.20}$$

where the mean matter density has been written in terms of the critical density, defined in Equation 1.23, in comoving space.

It is more explicit now how the convergence is connected to the matter distribution: Equation 2.20 depicts the convergence as an integration of the density contrast along the line of sight, weighted by comoving transverse distances and the scale factor. Hence, the convergence κ can be interpreted as the projected overdensity, or the projected mass. Now, given a population of source galaxies with source redshift distribution $n(\chi)$, the mean convergence from such population can be obtained as

$$\kappa(\boldsymbol{\theta}) = \int_0^{\chi_{\rm lim}} n(\chi)\kappa(\boldsymbol{\theta}, \chi) d\chi, \tag{2.21}$$

where $\chi_{\rm lim}$ is the limiting comoving distance of the galaxy sample. Putting everything together:

$$\kappa(\boldsymbol{\theta}) = \frac{3H_0^2\Omega_{\rm m}}{2c^2} \int_0^{\chi_{\rm lim}} g(\chi) f_K(\chi) \frac{\delta[f_K(\chi)\vec{\theta}, \chi]}{a(\chi)} d\chi, \tag{2.22}$$

where the lensing strength at a distance χ of the combined background galaxy distribution at χ' has been summarised in the term

$$g(\chi) = \int_{\chi}^{\chi_{\rm lim}} n(\chi') \frac{f_K(\chi' - \chi)}{f_K(\chi')} d\chi', \tag{2.23}$$

known as *lens efficiency* or *lensing kernel*. Back to the initial intent of this section, we see how the convergence is a measure of the total matter density projected along the line of sight accounting for lensing and the geometry of the universe via the distance ratios, and the source galaxy distribution, given a cosmological model.

Relation between shear and convergence: mass mapping methods

It has been described so far how the shear is related to what is the actual observed quantity (galaxy shapes) and how the convergence provides rich information about the total matter density projected along the line of sight. The last step consists in considering how the shear and the convergence are related, to complete the link between observation and the information of matter distribution along the line of sight. From Equation 2.13 and Equation 2.14 one can see how both the shear and the convergence are related through the lensing potential. Considering these expressions in Fourier space, one gets:

$$\hat{\gamma}_1 = \frac{(k_1^2 - k_2^2)}{2}\hat{\psi}; \quad \hat{\gamma}_2 = k_1 k_2 \hat{\psi} \tag{2.24}$$

$$\hat{\kappa} = \frac{(k_1^2 + k_2^2)}{2}\hat{\psi}. \tag{2.25}$$

Then, one way to obtain the convergence from the shear is to follow Kaiser-Squire (KS) inversion method (Kaiser and Squires 1993) that combines these equations to build a minimum variance estimator of the convergence

$$\hat{\kappa}_\gamma^{\mathrm{est}} = \frac{(k_1^2 - k_2^2)}{k^2}\hat{\gamma}_1 + \frac{2k_1k_2}{k^2}\hat{\gamma}_2 \quad \text{or} \quad \hat{\kappa}_\gamma^{\mathrm{est}} = \frac{k_1^2 + k_2^2}{(k_1 + ik_2)^2}\hat{\gamma}, \tag{2.26}$$

with $k = \sqrt{k_1^2 + k_2^2}$. This is equivalent to write in direct space:

$$\kappa(\boldsymbol{\theta}) = \frac{1}{\pi}\int \mathcal{D}(\boldsymbol{\theta} - \boldsymbol{\theta}')\gamma(\boldsymbol{\theta}')d\boldsymbol{\theta}' + \kappa_0, \tag{2.27}$$

where the convolution kernel $\mathcal{D}$ is the Fourier transform of the inverse of the pre-factor in Equation 2.26. This equation tells us that the convergence can be obtained as a convolution of the shear field, meaning that there exists a linear relation between the shear and the convergence. Nevertheless, Kaiser-Squires inversion is characterised by some complications due to different factors. On one side, the shape noise (defined in Section 2.1.3) is larger than the shear induced by lensing and the random sampling of background galaxies propagates errors through this reconstruction. Therefore, smoothing to reduce the shape noise per pixel and ensure that there are no empty pixels is required. This imprints a loss information at small scales and requires accounting for masks. On the other side, this method suffers from the so called *mass-sheet degeneracy*: a constant convergence does not induce a shear, and this can be seen by noticing that Equation 2.26 is not defined for $k_1 = k_2 = 0$. In real space, this is encoded by the constant term κ_0 in Equation 2.27 and it implies that the global level of the convergence map is undetermined by this method.

Despite some advantageous properties as linearity (in the weak regime), minimum variance and the ability to accommodate reduced shear of this inversion method, the different shortcomings mentioned above led to the introduction in the literature of alternative mass mapping techniques. Some examples are approaches aimed to remove the errors caused by the finite size of the field in weak lensing maps through *inpainting* (Pires et al. 2009) or based on estimating the convergence from the gradient of the convergence (flexion) as proposed in Bacon et al. 2006. Weak-lensing reconstructions of the potential and mass distribution in three dimensions are also possible. In the recent years, powerful methods based on sparsity have also been introduced, such as `GLIMPSE` (Lanusse et al. 2016) where to preserve all available small-scale information, any binning of the irregularly sampled input shear and flexion fields are avoided and the mass mapping problem is treated as a general ill-posed inverse problem, which is regularised using a robust multi-scale wavelet sparsity prior. In Jeffrey et al. 2020 the first reconstruction of dark matter maps from weak lensing observational data using deep learning was presented, with a method called `DeepMass`. Such method estimates the posterior mean of the convergence map from observed weak lensing shear measurements. More specifically, it estimates the convergence maps through a function $\mathcal{F}_{\boldsymbol{\Theta}}$ that maps the pixelised shear to the convergence map such that $\hat{\boldsymbol{\kappa}} = \mathcal{F}_{\boldsymbol{\Theta}}(\boldsymbol{\gamma})$, where what is learned by the CNN are the parameters $\boldsymbol{\Theta}$ by minimising a mean-square-error (MSE) cost function $J(\Theta) = ||\mathcal{F}_{\Theta}(\boldsymbol{\gamma}) - \kappa_{\text{true}}||_2^2$. This cost function is evaluated on a set of training data which consists of pairs of realistic shear and truth convergence maps and the truth maps are drawn from a prior distribution $P(\kappa)$, so that the corresponding noisy shear map is drawn from the likelihood $P(\gamma|\kappa)$. This provides a generalised method where the simulations are trained over a broad prior distribution of cosmological parameters and it has proven to be substantially more accurate than existing mass-mapping methods: compared to Wiener filtering with a fixed power spectrum, the `DeepMass` method improved the mean square error (MSE) by 11%.

Another novel approach introducing a new convergence map model based on a mixture of a sparsity-based component to account for the non-Gaussian structure of the field and a Gaussian random field to represent the linear characteristics of the field was presented in Starck et al. 2021 and it is called MCALens. As it uses concepts that need an understanding of sparsity, it will be treated more in detail in Section 3.6. Moreover, very recently, another method, called `KaRMMa` for performing mass map reconstruction based on a Bayesian approach has been introduced (Fiedorowicz et al. 2021) in the literature. It is based on a forward model of the observed shear field as a noisy realsation of an underlying noiseless convergence field, that is assumed to be a non-linear transformation of a Gaussian random field parametrised by a shift parameter, being the minimum possible value that the map can take in a pixel. This method outperforms the Kaiser–Squires reconstruction by exhibiting a narrower distribution of residuals when compared to some true mass maps from simulations. It also appears to be potentially powerful for aspects such as generating void or peak catalogues, as otherwise other methods as KS will also identify prior-sourced structures in peak/void finding algorithms. However, its current implementation does not allow for cosmological inference analysis as a specific cosmology is required to be able to specify the prior; also, it can only perform reconstruction for one single tomographic bin, hence making a joint reconstruction across multiple tomographic bins impossible. Lastly, it is still restricted to relatively low resolutions, therefore, an improved version of it would be needed at this stage in the interests of performing a cosmological inference analysis.

2.2 Cosmic shear

Now that the building blocks to understand the main weak lensing quantities have been described, the next milestone is to understand how to extract cosmological information from weak lensing observables. At this point, we know that the shear is strictly related with what is actually measured, i.e. the galaxy shape, that the convergence can be interpreted as the 2D mass projection along the line of sight providing information about the matter distribution, and that the shear and the convergence are related through the lensing potential and can be obtained with inversion methods such as the ones mentioned in the previous section.

2.2.1 The shear two-point correlation function

In particular, through Equation 2.22, we have seen how the convergence is related to the density contrast and via Equation 2.26 the link between shear and convergence. Since the average density contrast $\langle\delta\rangle$ is zero, following Equation 2.22 the expected value of convergence is also zero. In order to retrieve cosmological information, the first non trivial cosmic shear observable is the real-space shear two-point correlation function (2PCF). It can be built by decomposing the shear in a *tangential* component γ_t and in a *cross* component $\gamma_\times$:

$$\gamma_t = -\Re(\gamma e^{-2i\phi}); \quad \gamma_\times = -\Im(\gamma e^{-2i\phi}) \tag{2.28}$$

where ϕ is the polar angle of the direction $\boldsymbol{\theta}$. Then three two-point correlators can be built from these two shear components: $\langle\gamma_t\gamma_t\rangle$, $\langle\gamma_\times\gamma_\times\rangle$, $\langle\gamma_t\gamma_\times\rangle$ leading to the three components of the 2PCF $\xi(\theta)$:

$$\xi_\pm(\theta) \equiv \langle\gamma_t\gamma_t\rangle(\theta) \pm \langle\gamma_\times\gamma_\times\rangle(\theta); \quad \xi_\times(\theta) \equiv \langle\gamma_t\gamma_\times\rangle(\theta). \tag{2.29}$$

These correlation functions should be invariant under parity transformation, namely: $\gamma_t \to \gamma_t$, $\gamma_\times \to -\gamma_\times$. Then $\xi_\times$ vanishes. Reminding now that in the weak regime Equation 2.18 holds, i.e. the average of the observed ellipticity provides an unbiased estimate of the shear (assuming that the shear is constant over the averaged area and provided there are no systematics in shape measurement), an estimator of the 2PCF is given by (Schneider, P. et al. 2002):

$$\hat{\xi}_\pm(\theta) = \frac{\sum_{ij} w_i w_j [\epsilon_t(\boldsymbol{\theta}_i)\epsilon_t(\boldsymbol{\theta}_j) \pm \epsilon_\times(\boldsymbol{\theta}_i)\epsilon_\times(\boldsymbol{\theta}_j)]}{\sum_{ij} w_i w_j}, \tag{2.30}$$

where ϵ_t and $\epsilon_\times$ are respectively the tangential and cross component of the observed ellipticity ϵ^{obs} (the apex *obs* has been omitted) and each galaxy has a measured ellipticity ϵ^{obs} and an attributed weight w_i to account for the measurement uncertainty. The sum extends over pairs of galaxies (i, j) at corresponding positions on the sky $\boldsymbol{\vartheta}_i$ and $\boldsymbol{\vartheta}_j$ whose separation $|\boldsymbol{\vartheta}_i - \boldsymbol{\vartheta}_j|$ lies in an angular distance bin around θ. More details on this formalism and notation can be found in Schneider, P. et al. 2002.

2.2.2 The convergence power spectrum

Analogously to the previous section, it is possible to define a 2PCF for the convergence field. In particular, working in Fourier space, the two-point correlation function defines the convergence power spectrum $C_\kappa(\ell)$ through:

$$\langle\tilde{\kappa}(\boldsymbol{\ell})\tilde{\kappa}^*(\boldsymbol{\ell}')\rangle = (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{\ell} - \boldsymbol{\ell}')C_\kappa(\ell), \tag{2.31}$$

where δ_{D} is the Dirac delta, and following Schneider 2006, $\tilde{\kappa}$ is the Fourier transform of the convergence field κ:

$$\tilde{\kappa}(\boldsymbol{\ell}) = \int d^2\theta\kappa(\boldsymbol{\theta})e^{i\boldsymbol{\ell}\cdot\boldsymbol{\theta}}. \tag{2.32}$$

Then, using the expression found in Equation 2.22 for the convergence in Equation 2.32 to compute Equation 2.31 and working under the following assumptions:

- *Limber approximation*: correlations along the line of sight are neglected, namely only modes that lie in the plane of the sky are considered (Limber 1953; Kaiser 1992; Simon, P. 2007)
- *Small-angle approximation*: approximating to first order the expansion og trigonometric functions of the angle
- *Flat-sky limit*: replacing the expansion in spherical harmonics by an expansion in Fourier modes (Lemos, Challinor, and Efstathiou 2017),

it is possible to get an explicit expression for the convergence power spectrum as

$$C_\kappa(\ell) = \left(\frac{3H_0^2\Omega_{\mathrm{m}}}{2c^2}\right)^2 \int_0^{\chi_{\mathrm{lim}}} \frac{g^2(\chi)}{a^2(\chi)} P_\delta\left(k = \frac{\ell}{f_K(\chi)}, \chi\right) \mathrm{d}\chi, \tag{2.33}$$

where P_δ is the matter power spectrum, arising from the the density contrast in the right hand side of Equation 2.22 and $g(\chi)$ is the *lensing kernel* introduced in Equation 2.23. Reminding also the relation that holds between the shear and the convergence through Equation 2.26, this implies that $|\tilde{\kappa}| = |\tilde{\gamma}|$. Hence, the shear power spectrum defined through

$\langle\tilde{\gamma}(\boldsymbol{\ell})\tilde{\gamma}^*(\boldsymbol{\ell}')\rangle = (2\pi)^2\delta_{\mathrm{D}}(\boldsymbol{\ell}-\boldsymbol{\ell}')C_\gamma(\ell)$, is actually equivalent to the convergence power spectrum, namely $C_\kappa(\ell) = C_\gamma(\ell)$. Thus, while it is possible to obtain the convergence from shear measurements through inversion methods as described in Section 2.1.3, we can also work directly with the 2-point statistics of shear and still retrieve the cosmological information contained in Equation 2.33 and *viceversa*.

2.2.3 Systematics

There is a range of systematics effects that can mimic a shear signal, namely, multiplicative shear bias, mean redshift bias, baryonic feedback, intrinsic alignment, and boost factor. This section provides a brief overview of these systematics effects, by defining them and by presenting some of the issues that characterise them and some of the current approaches to try to alleviate them.

Baryonic effects

As mentioned in Section 1.3.2 and as will be outlined more in details in Section 2.2.5, predicting the matter distribution in the non-linear regime with high accuracy still represents a challenge. For instance, for Fourier modes as large as $k \sim 10\ h\ \mathrm{Mpc}^{-1}$ baryonic physics can differ from the dynamics of dark matter. In particular, it has been shown that to understand the fact that the star formation rate is higher than expected, the feedback from active galactic nuclei (AGN), i.e. the re-deposition of energy and momentum into the interstellar medium of a galaxy through the outflows and radiation, needs to be taken into account during galaxy formation (Dubois et al. 2013). Hence, as baryonic physics modifies the clustering property of matter, it impacts the non-linear matter power spectrum and, as a direct consequence, it also influences the lensing signal. If the baryonic physics is not properly accounted for, biases are induced on cosmological constraints. In particular, Daalen et al. 2011 have shown that this impacts the matter power spectrum to a level which is higher than the accuracy required for current and future weak lensing surveys, that require an accuracy of a few per cent for a large range of scales.

This becomes more important at more non-linear scales; however, already at the level of second order statistics, it has been shown that ignoring the effects of baryonic physics can induce a bias as large as $\sim 40\%$ on the dark energy equation of state for the AGN scenario (Semboloni et al. 2011). The convergence power spectrum is significantly altered specifically at angular scales corresponding to $\ell \leq 1000$ (or to $\ell \leq 3000$, depending on the assumed statistical uncertainty (Kilbinger 2015)). With this awareness, a lot of effort has been put in the attempt of modelling such complex physics in the ensemble of baryonic processes as radiative cooling, star formation, supernovae, AGN and their associated feedback on the matter density field through hydro-dynamical simulations such as the `Horizon-AGN` (Dubois et al. 2016), the `Illustris` (Vogelsberger et al. 2014), the EAGLE (Schaye et al. 2014), the `Osato15` (Osato, Shirasaki, and Yoshida 2015), the `BAHAMAS` (McCarthy et al. 2016), the `Magneticum path-finder` (Dolag, Komatsu, and Sunyaev 2016), the `Illustris-TNG` (Pillepich et al. 2017). An effort on modelling these effects has been also put in analytical and physically intuitive approaches based on the halo model as for example in Semboloni et al. 2011 or in Fedeli et al. 2014. An alternative method, that does not require hydro-dynamical simulations, consists of modifying the output of N-body simulations by slightly displacing particles around halo centres based on an empirical parametrisation of halo profiles including gas, stars, and dark matter. This approach is called *baryonification*, introduced in Schneider and Teyssier 2015 and Schneider et al. 2019 and it allows to perform many fast realisations of the non-linear cosmic density field with varying baryonic parameters. As already mentioned above, being able to account for baryonic effects becomes more and more important when willing to probe non linear scales. In the next section, dedicated to higher order statistics, some recent studies and results with baryonic effects in non-Gaussian statistics will be discussed. In our first application on real data using CFIS that will be presented in Chapter 7, we consider this effect, alone and in combination with the multiplicative shear bias.

Intrinsic alignment

The interpretation of the lensing signal and convergence maps as introduced in this work so far is based on the hypothesis that the intrinsic galaxy orientation should not have any privileged direction, i.e. on the assumption made in Equation 2.18. In reality, galaxies are continuously exposed to the gravitational interaction with the surrounding matter distribution, during their formation and evolution, and this leads to a coherent alignment of their intrinsic shapes, induced by the underlying tidal field on physically near galaxies and by a complex relation with their host halos. This phenomenon is called *intrinsic alignment* (IA) (Heavens, Refregier, and Heymans 2000). Extensive reviews on this topic are present in Kirk et al. 2015; Kiessling et al. 2015; Joachimi et al. 2015.

At the level of second order statistics, this effect can be included by forgetting the assumption that $\langle \epsilon^{\mathrm{s}} \rangle = 0$ and by considering the 2PCF for the observed galaxy ellipticities $\langle \epsilon^{\mathrm{obs}} \rangle$ (indicating with i and j a pair of galaxies and the average is over all pairs):

$$\langle \epsilon_i^{\mathrm{obs}} \epsilon_j^{\mathrm{obs}} \rangle = \langle \gamma_i \gamma_j \rangle + \langle \epsilon_i^{\mathrm{s}} \epsilon_j^{\mathrm{s}} \rangle + \langle \gamma_i \epsilon_j^{\mathrm{s}} \rangle + \langle \epsilon_i^{\mathrm{s}} \gamma_j \rangle, \tag{2.34}$$

where the $\langle \gamma_i \gamma_j \rangle$ can be be referred to as “GG”, describing the cosmic shear correlation (the only contribution that survives, when working under the assumption that there is no preferred orientation for intrinsic galaxy shapes); the $\langle \epsilon_i^{\mathrm{s}} \epsilon_j^{\mathrm{s}} \rangle$ term that can be referred to as “II”, encodes the correlations between intrinsic ellipticities; and the last two terms $\langle \gamma_i \epsilon_j^{\mathrm{s}} \rangle$, $\langle \epsilon_i^{\mathrm{s}} \gamma_j \rangle$ can be summarised in the notation “GI”, and describe the correlations between the shear for one galaxy and the intrinsic ellipticity of the other. Neglecting the IA effects can introduce important biases on weak lensing signal and consequently on parameters estimation. Indeed, in order to properly model this effect in weak lensing analysis, a wide number of models has been proposed in the literature, such as the *Linear Alignment* (LA) model (Catelan, Kamionkowski, and Blandford 2001), based on the tidal alignment of galaxies taken at linear order or the *Non-Linear Alignment* (NLA), an empirical modification of the LA based on the consideration that it better fits the data, as motivated in Hirata and Seljak 2004; Bridle and King 2007. Kirk et al. 2012 noted that, if IA is imprinted in the primordial phases of galaxy formation, perturbations are still in the linear regime: they then claim that the linear power spectrum should be used for the "II" correlation. Along with this, shear depends on the non-linear power spectrum which must be taken into account in the correlation with IA. Therefore, they proposed to use a *Freeze - in model* (FM) where the overall effect of IA is suppressed at low redshifts and large k because of the use of linear rather than non-linear matter power spectrum.

Many other models that try to account for what is neglected in the LA and NLA models (such as astrophysical effects beyond the lowest order) can be mentioned, such as the *Full tidal alignment* (FTA), obtained by carrying on the computation of the intrinsic-matter power spectrum to the next leading orders (Blazek, Vlah, and Seljak 2015) or the *Tidal alignment and torquing* (TAT) including also terms originating from the tidal torquing expected in rotation supported systems. In Fortuna et al. 2020, a unified framework to account for different IA signatures at large and small scales that also incorporates luminosity and radial dependence has been presented. In this context, the total IA power spectra are given by the sum of the behaviour at large scales (where the majority of the IA signal is due to red central galaxies) and the behaviour at small scales (dominated by the satellite alignment signal). The fact that there is such a vast number of different models is due to the fact that some of these models are not able to represent accurately the real physical effects due to the large number approximations (LA) or they are too hard to compute (as for example FTA and TAT) or not completely understood (FM), or with too many parameters (full NLA).

Given the difficulties to reproduce the physical properties of galaxies coherently with the observations employing DM-only simulations to predict the intrinsic alignment signal, the use of hydro-dynamical simulations has been proven powerful to better predict the complex relation between halo shape and spin and galaxy shape and spin (Codis et al. 2015; Bate et al. 2019). An alternative way to include IA in weak lensing analysis, consists in treating it as an additive component to the cosmological signal, directly obtained from the particle shells of the simulations in a similar fashion as the cosmological convergence signal κ itself, as performed in Zürcher et al. 2021 with the `Ufalcon` pipeline (Sgier et al. 2019).

Multiplicative shear bias

The multiplicative shear bias is a systematics arising at the level of the shear measurement that can be caused by various effects as *noise bias*, a systematics error whose origin is identified in high-order noise terms in the measurement of the shape parameters of galaxies, due to the non-linear relation between image pixels and galaxy shape (Refregier et al. 2012); *model bias* in the context of model fitting methods, if the model is not able to represent realistic galaxy morphologies well enough, then the shape estimator will be biased (Kacprzak et al. 2014); or imperfect PSF corrections (Paulin-Henriksson, S., Refregier, A., and Amara, A. 2009). This systematics can change the overall amplitude of the cosmic shear auto-correlation and its cross-correlation with other probes of the large scale structure and consequently cause biased estimations of cosmological parameters (Liu, Ortiz-Vazquez, and Hill 2016). The simplest way to deal with shear estimation biases consists in estimating the shear errors by processing simulated data with known input shear through the same pipeline as the real data. These calibration methods require simulations that need to be carefully matched to the properties of the given data in order to have correct calibration factors. An alternative way to calibrate the shear measurement is the *metacalibration* (Huff and Mandelbaum 2017). This method does not rely on simulations, and allows to avoid the challenge of creating a large number of time-consuming image simulations and it has proven to be very effective both when applied to the DES Year 1 (Y1) catalogue (Zuntz et al. 2018) and to the Canda-France Imaging Survey (CFIS) r-band data (Guinot et al. in prep.). In our first application on real data using CFIS that will be presented in Chapter 7, we consider this effect, alone and in combination with baryonic effects.

Mean redshift bias

As one can see from Equation 2.21 and Equation 2.23, it is required to have a knowledge of the the source galaxy distribution $n(z)$, in order to be able to interpret weak lensing observables from a cosmological point of view. In particular, for a given desired accuracy to estimate cosmological parameters, the mean redshift $\bar{z}$ has to be determined at least to that accuracy. When performing a tomographic analysis, this applies to each individual redshift bin, as the knowledge of the centers of each bin needs to be better than a per cent to avoid a decrease in the accuracy of dark-energy parameters to 50% as was shown by Huterer et al. 2006. The estimation of galaxy redshifts through photometry (briefly mentioned in Section 1.4.1), or photometric redshifts (photo-zs) is hence needed. The requirements of the accuracy of photometric redshifts depend on the survey. For example, for KiDS, it was found that even a Gaussian 1σ uncertainty on the measured mean redshift of each tomographic bin can degrade the statistical errors on relevant cosmological parameters by $\sim 25\%$ (Hildebrandt et al. 2016). Since the galaxy samples photometric redshift distributions typically present highly non-Gaussian tails, analytic estimates based on Gaussian redshift errors don't represent the optimal way to obtain accurate calibrations. In Hildebrandt et al. 2016, different methods, based on weighted direct calibration obtained by a magnitude-space re-weighting or on angular cross-correlation based calibration, that rely on spectroscopic redshift (spec-z) training samples are compared. Their favoured technique relies on a weighted direct calibration obtained by a magnitude-space re-weighting of spectroscopic redshift catalogues that overlap with KiDS (Lima et al. 2008). They also conclude that for future cosmic shear surveys bringing considerably large data sets, it will be essential to reduce the statistical uncertainty in the redshift calibration. For *Euclid*, Euclid Collaboration et al. 2020a presented a report on the performance of 13 photometric redshift codes for single value redshift estimates and redshift probability distributions on a common set of data, focusing particularly on the redshift range that the *Euclid* will probe. Their finding is that while all methods are able to provide reliable single value estimates, several machine-learning methods do not manage to produce useful redshift probability distributions. Moreover, they find that no machine-learning method provides good results in the regions of galaxy color-space that are sparsely populated by spectroscopic-redshifts, while generally performing better than template-fitting (Bolzonella, Miralles, and Pelló 2000) methods at low redshift, indicating that template-fitting methods do not use all of the information contained in the photometry. They conclude that further work in identifying how to best select between machine-learning and template-fitting approaches for each individual galaxy should be pursued as a priority for future galaxy surveys like *Euclid*.

Boost factor

Another consequence of the uncertainty of photometric redshift estimates is the contamination of the source galaxy catalogue with galaxies associated with the foreground cluster, namely counting galaxies residing at the lens redshift in the $n(z)$ of the sources. As this contamination dilutes the measurement, in order to recover the true signal, the raw signal has to be boosted. Hence, this effect is referred to as the *boost factor*. This systematics is mostly relevant for higher order statistics such as peak counts as the strength of the shear signal varies with the distance of source galaxies to the lens and this causes the detection probability of a peak to depend on the distribution of the redshifts of source galaxies at the position of a peak (Kacprzak et al. 2016). In the context of peak counts, the boost factor can be estimated in different ways. For instance, Kacprzak et al. 2016 corrected the peak function found in their cosmological training set as a function of S/N by evaluating the fractional over-crowding and over-blending rates in peaks of different S/N from a separate catalogue that matches the DES-SV $n(z)$ and by restricting their measurement to $S/N < 4$ where they found that the impact of this effect is minimal (a shift in S_8 of about 0.01) and can then be neglected. For KiDS-450, Shan et al. 2017 used a cluster catalogue that overlaps with the KiDS-450 survey, and evaluated the boost factor from the excess source density around these massive objects. Their finding is that the contamination to the peak function reaches 27% for peaks with $S/N = 5$, being however less than 6% for $S/N < 4$, consistently with Shan et al. 2017.

2.2.4 Constraints with cosmic shear second order statistics

The detection of cosmic shear and its interpretation as direct measure of the mass power spectrum (Bacon, Refregier, and Ellis 2000; Kaiser, Wilson, and Luppino 2000) and the consequent proof of the technical feasibility of using weak lensing surveys to measure dark matter clustering to constrain cosmological parameters (Waerbeke et al. 2000), led to the use of cosmic shear as a primary probe for past, current and future galaxy surveys. The most popular approach adopted by these surveys to perform cosmological parameter inference relies on second order statistics, in the form of correlation functions (or its Fourier equivalent, the power spectrum), or estimated using *pseudo*-C_ℓ, band powers or quadratic estimators. For example, Kilbinger, Fu, Heymans, et al. 2013 used the shear two-point correlation functions computed on CFHTLenS data spanning 154 deg^2 to constrain σ_8 and Ω_m with cosmic shear only and on Ω_K, Ω_DE and w_0 by adding WMAP7 and BOSS data. Again, using the shear two-point correlation function Troxel et al. 2018 constrained σ_8, Ω_m, their combination S_8 and w, for DES Y1, covering 1321 deg^2 of the southern sky and found constraints in good agreement with previous cosmic shear results from KiDS-450 (covering 450 deg^2) (Hildebrandt et al. 2016). For the fourth data release of the KiDS that doubled the survey area (covering 1000 deg^2), Asgari, Marika et al. 2021 compared a COSEBIs (Complete Orthogonal Sets of E/B-Integrals) analysis with complementary analyses of the two-point shear correlation function and band power spectra, finding results to be in great agreement and the sensitivity of the three statistics to systematics to be robust and dominated by statistical errors for the constraints on S_8. Hamana et al. 2020 employed the shear 2PCF from the HSC first-year data, covering 136.9 deg^2 finding constraints for S_8 consistent with the latest Planck CMB result while in tension at a $\sim 2\sigma$ level KiDS + VIKING-450 (Joudaki, S. et al. 2020). To investigate this tension and understand its origin, they plan in future works to analyse the KiDS public shape catalog with their methodology. The most recent results available from cosmic shear to date are from DES-Y3 cosmological results (Amon et al. 2021), with the data spanning 4143 deg^2 for four redshift bins. The analysis provides constraints in the context of the ΛCDM model and find a 3% constraint for S_8, the highest significance measurement of cosmic shear to date, with a signal-to-noise of 40. The two low-redshift measurements are found to be statistically consistent with Planck Cosmic Microwave Background result, however, a $\sim 2.3\sigma$ tension, slightly dependent on the method used, remains for high-redshift predictions (Lemos, DES Collaboration, et al. 2021).

Forthcoming experiments like *Euclid* (Laureijs et al. 2011), the Vera C. Rubin Observatory LSST (LSST Science Collaboration et al. 2009), and the Nancy Grace Roman Space Telescope (Spergel et al. 2015) will also use cosmic shear as a primary probe and will give access to small scales than ever before. Modelling challenges in cosmic shear will become hence more and more important on non-linear scales. This last point is further discussed in the next section, with a focus on the *Euclid* mission.

2.2.5 Non-Linear regime

This paragraph aims to provide a concise description of the impact on the modelling of cosmic shear of the scales at the which matter density field perturbations become larger and linear theory fails to predict the evolution of large-scale structures, mostly based on Martinelli et al. 2020 and Euclid Collaboration et al. 2020b. A wide contribution to the constraining power of the cosmic shear signal lies in non-linear scales, specifically in correspondence of scales up to $k \approx 7$ h Mpc^{-1} (Taylor, Kitching, and McEwen 2019). This requires the capability of modelling the matter power spectrum at high-k. Indeed, if non-linear scales not properly modelled of the shear signal are included in the analysis, this can lead to biases or underestimation of the contours size. For instance, Copeland, Taylor, and Hall 2018 have shown that the impact of baryonic effects on the non-linear power spectrum can lead to a decrease of the FoM by up to 40%, depending on the model and the prior. Hence, it is very important to model the behaviour at high-k and be aware of the choices adopted when modelling the non-linear power spectrum.

A straightforward way to *avoid* the problematic of dealing with non-linear scales that might be difficult to model and interpret is to apply a cut in ℓ-modes by setting a maximum multipole value $\ell_{\rm max}$. This method is the most commonly employed in weak lensing analysis as it yields unbiased results. For example, Euclid Collaboration et al. 2020b identified cuts in ℓ-modes corresponding to an *optimistic* and a *pessimistic* settings in the context of Fisher forecasts. Specifically, for the pessimistic setting the cut is set at $\ell_{\rm max} = 1500$. This choice is motivated by the fact that the information content of the shear power spectrum, defined as the total signal-to-noise ratio, decreases by 30% at $\ell_{\rm max} = 5000$ when in the covariance matrix is considered also the non-Gaussian contribution (Takada and Hu 2013).

The optimistic setting sets $\ell_{\rm max} = 5000$, to quantify the gain that could be obtained extending to the non-linear regime, if one were able to include such high multipoles. As also discussed in Euclid Collaboration et al. 2020b, this approach is practical, but has limited applicability and can lead to imprecise parameter constraints, especially at low redshifts. Indeed, parameter sensitivities are generally scale dependent (Copeland, Taylor, and Hall 2018) and respond differently to different cuts and to non-Gaussian correlations in the covariance. In the analysis presented in this thesis, we follow the optimistic setting we of Euclid Collaboration et al. 2020b when employing the `MassiveNus` simulations, while using a value closer to the pessimistic setting of Euclid Collaboration et al. 2020b (we use $\ell_{\rm max} \sim 2000$) when using the *cosmo*-SLICS simulations.

To overcome these issues, methods to optimally remove sensitivity to poorly modelled scales based on the Bernardeau-Nishimichi-Taruya (BNT) transform (Bernardeau, Nishimichi, and Taruya 2014), where the observed tomographic angular power spectrum can be re-weighted in such a way that each redshift bin retains only the information about lenses within a small redshift range, have been proposed. In particular, employing these methods can reduce the error on S_8 by 32% relative to a correlation function analysis with the same priors but with angular scale cuts and can alleviate baryons modelling uncertainties, as shown when working in configuration space by Taylor, Bernardeau, and Huff 2021. At the same time, Deshpande, Taylor, and Kitching 2020 showed how the *optimistic* scenario of $\ell_{\max} = 5000$ mentioned for Euclid Collaboration et al. 2020b is more achievable when the k-cut cosmic shear technique, that consists in applying a redshift-dependent ℓ-cut after making the BNT transform, is applied.

Besides the complexity of modelling the power spectrum at non-linear scales, the fact that the fluctuations of the cosmic density field are Gaussian in the early Universe, but then grow into a complex structure containing walls, filaments, nodes, and voids, makes second order statistics like the power spectrum *blind* to non-Gaussian features of the large-scale structure field due to non-linear evolution at late times.

Figure 2.3 shows this with the comparison between a convergence map of size $12.25\ \mathrm{deg}^2$ from the `MassiveNus` suite of simulations (left top panel) with a *fake* convergence map of the same size (right top panel) built such that its power spectrum matches the power spectrum of the convergence map. By looking at both maps, it is possible to notice how many structures (such as the big blog at the right top of the left map, or the structures located at its center) are not present in the map of the right and yet, the two maps are described at second order by the same power spectrum. Indeed, by definition, second order statistics can potentially capture all cosmological information encoded in linear, Gaussian density fields. However, as shown in Figure 2.3, a large number of high value convergence are seen in the first convergence map, but are not encoded in the information brought by the power spectrum indicating rich non-Gaussian information and the need to go beyond second order statistics. The next section is dedicated to the state of the art of statistics of order higher than the second, introduced to account for non-Gaussianities.

2.3 State of the art of higher-order statistics

In this section, I will present an overview of the state of the art of current higher-order statistics with the corresponding definitions and main recent results on forecasts and real data constraints. As mentioned at the end of the previous section, in the context of weak lensing, second-order statistics as the two-point correlation function or its Fourier transform (the power spectrum) do not capture the non-Gaussian information encoded in the non-linear features of weak lensing data (Weinberg et al. 2013). This has motivated the introduction of several higher-order statistics, such as Minkowski functionals (Kratochvil et al. 2012; Petri et al. 2015; Vicinanza et al. 2019; Marques et al. 2019; Parroni et al. 2020), higher-order moments (Petri, May, and Haiman 2016; Vicinanza, Cardone, Maoli, et al. 2018; Peel, Pettorino, Giocoli, et al. 2018; Chang et al. 2018; Gatti et al. 2020), the bispectrum (Takada and Jain 2004; Chan and Blot 2017; Coulton et al. 2019), peak counts (Kruse and Schneider 1999; Kratochvil, Haiman, and May 2010; Dietrich and Hartlap 2010; Maturi, Fedeli, and Moscardini 2011; Pires, Leonard, and Starck 2012; Hamana et al. 2012; Hilbert

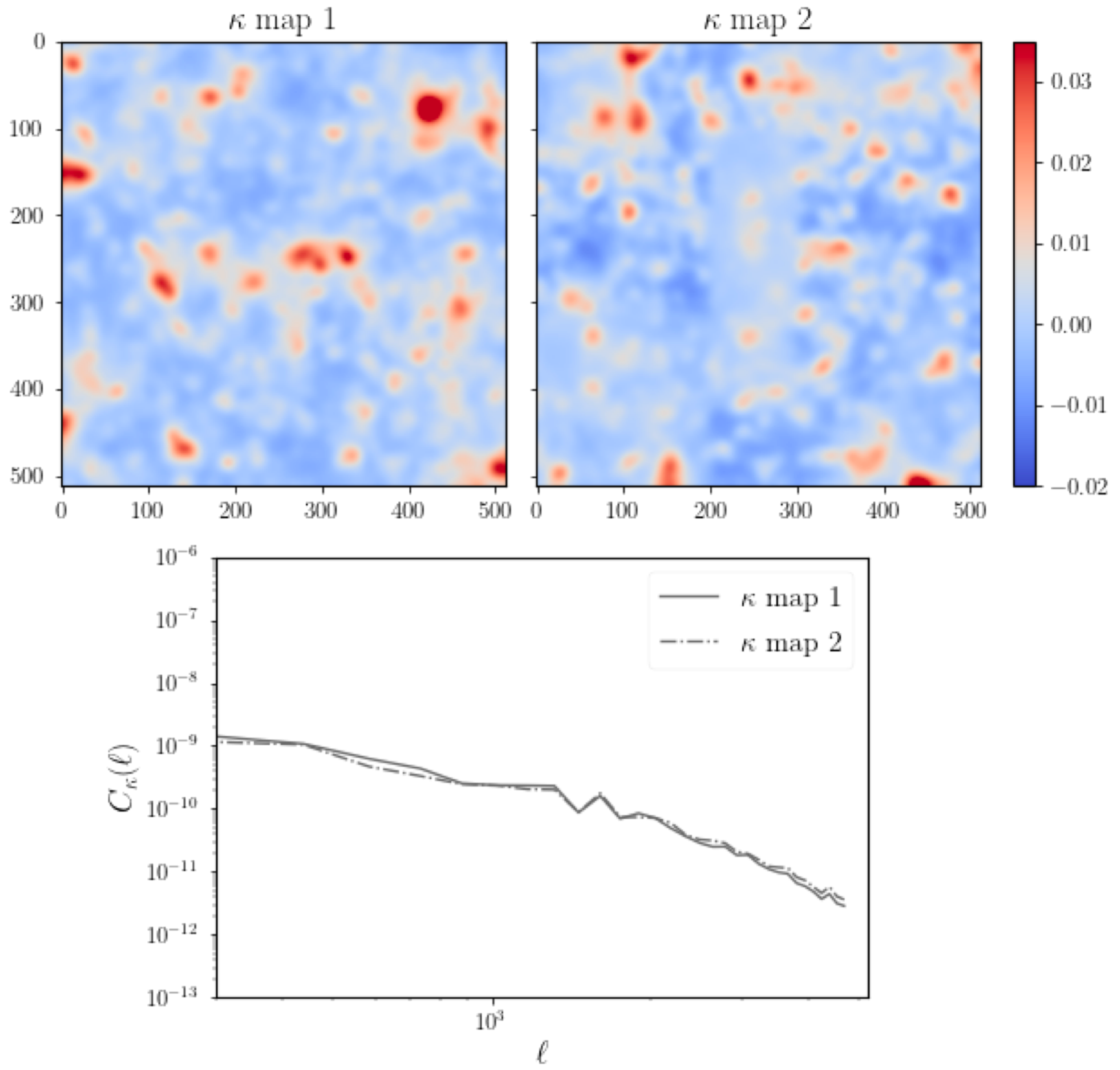


Figure 2.3: Comparison between a convergence map from the `MassiveNus` suite of simulations (left top panel) with a *fake* convergence map (right top panel) built such that its power spectrum matches the power spectrum of the convergence map. The bottom panel shows the power spectrum measured on the convergence map (continuous line) with respect to the power spectrum computed on *fake* convergence map (dashed line). The message of this illustration is to show how a large number of high value peaks (as the big blob at the top right of κ map 1 detectable also by naked eye or the several structures at the center of the same map) are not accounted for when extracting information with the power spectrum, indicating the presence of information beyond Gaussian statistics that is not present in the information brought by the power spectrum.

et al. 2012; Marian et al. 2012; Marian et al. 2013; Martinet et al. 2015; Lin and Kilbinger 2015; Giocoli, Moscardini, Baldi, et al. 2018; Martinet et al. 2018; Peel, Pettorino, Giocoli, et al. 2018; Li, Liu, Zorrilla Matilla, et al. 2019), the scattering transform (Cheng et al. 2020), wavelet phase harmonic statistics (Allys et al. 2020), and machine learning-based methods (Fluri et al. 2018; Peel et al. 2019; Gupta, Matilla, Hsu, et al. 2018; Ribli et al. 2019; Shirasaki et al. 2019), to account for non-Gaussian information in cosmological analysis. Focusing on peak counts, it has been shown that this statistics is particularly powerful in breaking degeneracy between the standard model and fifth forces in the dark sector (Peel, Pettorino, Giocoli, et al. 2018) as well as in constraining cosmological parameters when employed in a multi-scale setting (Liu, Petri, Haiman, et al. 2015; Lin, Kilbinger, and Pires 2016; Fluri, Kacprzak, Sgier, et al. 2018; Ajani et al. 2020; Zürcher et al. 2021). In particular, as it will be discussed in Chapter 4, in Ajani et al. 2020 we have advanced the state of the art by showing that multi-scale peak counts significantly outperform the weak lensing power spectrum, improving the constraints on the sum of neutrino masses $\sum m_\nu \equiv M_\nu$ by 63% when using a starlet filter. In the same study, we have also found that multi-scale peak counts are so constraining that the addition of the power spectrum does not further improve constraints. A very interesting feature that we have found for multi-scale peaks, when they are obtained using the starlet transform (Starck, Fadili, and Murtagh 2007), is the behaviour of the covariance matrix that tends to encode all information in its diagonal elements. Another weak lensing probe of large-scale structure is represented by cosmic voids, namely under-dense regions of the large-scale matter field (Colberg, Pearce, Foster, et al. 2008; Pisani, Massara, Spergel, et al. 2019). Local minima of weak lensing convergence maps, namely pixels with values smaller than their eight neighbouring pixels, have been proposed as tracers of the matter distribution voids to infer cosmological parameters, both in a mono-scale setting (Coulton et al. 2020b; Martinet, Nicolas et al. 2021; Davies, Cautun, Giblin, et al. 2020) and in a multi-scale setting (Zürcher et al. 2021). More specifically, Coulton et al. 2020b found that lensing minima alone are slightly less constraining than the peaks alone and, in agreement with Martinet, Nicolas et al. 2021 and Zürcher et al. 2021, that the combination of the two statistics produces significantly tighter constraints than the power spectrum.

2.3.1 One-point probability distribution function (PDF)

A summary statistics alternative to the power spectrum is the one-point probability distribution function (PDF), which is obtained by taking the histogram of projected weak lensing convergence in pixel values (Thiele, Hill, and Smith 2020). Using the one-point PDF, Liu and Madhavacheril 2019 found that the non-linear growth generates non-Gaussianity in the PDF, demonstrating additional information encoded in the tomographic PDF beyond the power spectrum and that the PDF alone outperforms the power spectrum in constraining cosmology as previously found by Patton et al. 2017. Recently, Boyle et al. 2020 provided a theoretical PDF model to perform forecasts and eliminate the need for relying on expensive N-body simulations, allowing for performing forecasts with the convergence PDF on the full set of ΛCDM parameters and established with Fisher forecasts that the constraining power of the convergence PDF compares favourably to the two-point correlation function for a *Euclid*-like survey area at a single source redshift.

2.3.2 Bispectrum

Another way to go beyond the second-order convergence power spectrum and include the non-Gaussian information encoded at non-linear scales consists in considering the next-leading order statistic defined as the *bispectrum* B_κ. Analogously to how the 2PCF is related to the power spectrum, by measuring the shear three-point correlation function (3PCF) (Takada and Jain 2003), one can relate the 3PCF to the bispectrum, and then to the matter distribution which depends on cosmology (Schneider, P. and Lombardi, M. 2003). The bispectrum measures three-point correlations of the convergence defined on a closed triangle in Fourier space:

$$\langle \tilde{\kappa}(\ell_1)\tilde{\kappa}(\ell_2)\tilde{\kappa}(\ell_3)\rangle = (2\pi)^2 \delta_D(\ell_1 + \ell_2 + \ell_3) \left[B_\kappa(\ell_1, \ell_2) + B_\kappa(\ell_2, \ell_3) + B_\kappa(\ell_1, \ell_3) \right] . \quad (2.35)$$

It is not trivial to define the relation between the 3PCF and the bispectrum and to efficiently evaluate it numerically. One difficulty of directly using the Fourier bispectrum is its large number of terms, which generally must be reduced in some way. Recent works that used the bispectrum to perform parameter inference relied on bispectrum estimators as the binned bispectrum estimator based on Bucher, Racine, and Tent 2016 and adapted to the flat-sky regime. Moreover, under the assumption of homogeneity and isotropy, the bispectrum can be written as a function that only depends on the magnitude of the ℓ modes and the binned bispectrum estimator can be implemented as a simple sum over products of filter maps. Using this assumptions, Coulton et al. 2019 found that the bispectrum and power-spectrum measurements are highly complementary and analysing them jointly produces 30% tighter parameter contours when using the `MassiveNus` simulations in an ideal setting, namely, neglecting systematic effects, baryonic effects, intrinsic alignment, multiplicative bias.

2.3.3 Minkowski functionals

Another powerful tool to account for deviations from Gaussianity is represented by Minkowski functionals (MFs), first introduced by Mecke, Buchert, and Wagner 1994 for galaxy clusters applications and then widely used in astrophysics and cosmology (Saar et al. 2006; Hikage, Komatsu, and Matsubara 2006; Ducout et al. 2012). Minkowski functionals are a complete family of morphological measures invariant under rotations and translations. Their usage allows one to morphologically characterise a given random field in a D dimensional space by a set of $D+1$ functionals V_j, with $j = 0.., D$. For a 2D field as for example when working with a weak lensing convergence map κ, setting a threshold ν and defining σ_0 as the standard deviation of the convergence field, the three Minkowski functionals V_0, V_1, V_2 describe respectively the area, the perimeter and the genus of the excursion set $\Sigma(\nu)$:

$$V_0(\nu) = \frac{1}{A} \int_{\Sigma(\nu)} da, \tag{2.36}$$

$$V_1(\nu) = \frac{1}{4A} \int_{\partial\Sigma(\nu)} dl, \tag{2.37}$$

$$V_2(\nu) = \frac{1}{2\pi A} \int_{\partial\Sigma(\nu)} \mathcal{K} dl, \tag{2.38}$$

where A is the total area of the map, $\Sigma(\nu)$ is the set of pixels with κ such that $\kappa > \kappa_0 = \nu\sigma_0$ and $\partial\Sigma(\nu)$ is the boundary of the excursion set. Given these definitions, the area V_0 is equivalent to the cumulative distribution function, the perimeter V_1 describes the length of its boundary and the genus V_2 is a measure of the total number of connected regions above κ_0 minus number of connected regions below κ_0, measured by the integrated geodesic curvature $\mathcal{K}$ along the boundary. Employing MFs for weak lensing cosmological constraints, Kratochvil et al. 2012 found that a substantial amount of information from non-Gaussian features is encoded by MFs of convergence maps and that this statistic is particularly well suited to break degeneracies and to constrain the dark energy equation of state parameter w. They identified that the non-Gaussian information is brought in part by the one-point function of the convergence through V_0, and in part through combining different smoothing scales for V_0, and through V_1 and V_2. Vicinanza et al. 2019 presented a method to match the theoretical predictions with measured MFs taking into account shape noise, imperfections in the map reconstruction and inaccurate description of the non-linearities in the matter power spectrum and bispectrum and validated their method against simulated maps reconstructed from shear fields generated by the MICE simulations. With a Fisher matrix analysis they forecasted the accuracy on cosmological parameters from a joint MFs and shear tomography analysis and found that MFs are helpful to break the $\Omega_{\mathrm{m}} - \sigma_8$ degeneracy but leading to an overall increase of the Figure of Merit. In a companion study, Parroni, Carolina et al. 2020 discussed the use of MFs in Stage IV lensing surveys updating the procedure used in Vicinanza et al. 2019 finding that MFs can provide a valuable help in increasing the FoM of the lensing survey, provided the nuisance parameters are known with a non negligible precision. Recently, Zürcher et al. 2021 forecasted the performance of the Minkowski Functionals for a stage-3-like weak lensing survey, and found that Minkowski functionals increase the constraining power by a factor of about 2 with respect to the power spectrum, and can help add some robustness against galaxy intrinsic alignment in a non-tomographic setting. Petri et al. 2015 applied Minkowski functionals to perform parameter inference using CFHTLenS data, finding that Minkowski functionals can constrain the non-degenerate direction in parameter space, Σ_8, where the effects of degeneracy are mitigated. However, they also find that constraints on the (Ω_m, σ_8) plane suffer from a strong bias that might be due to uncorrected systematics in the CFHTLenS data when using this statistics.

2.3.4 Higher-order moments

The non-Gaussianity of the convergence field can also be quantified through higher than second order moments. Following Vicinanza et al. 2016, the moment of order n of the convergence κ can be defined as

$$\langle \kappa^n \rangle = \frac{1}{\mathcal{N}_{\text{pix}}} \sum \tilde{\kappa}_i^n, \tag{2.39}$$

where i runs over the number of pixels in the map $\mathcal{N}_{\text{pix}}$ and $\tilde{\kappa}_i^n$ is the value of the convergence in the i-th pixel. Focusing on moments up to 4th order, one can construct the skewness S_3 and the kurtosis S_4 of the convergence field defined as

$$S_3 = \langle \kappa^3 \rangle / (\langle \kappa^2 \rangle)^{3/2}, \quad S_4 = \langle \kappa^4 \rangle / (\langle \kappa^2 \rangle)^2, \tag{2.40}$$

that quantify deviations from a zero centred Gaussian distribution. In particular, using aperture mass computed on weak lensing convergence maps, Peel, Pettorino, Giocoli, et al. 2018 found that these observables can be powerful to discriminate between f(R) models and ΛCDM. Howerver, they also found that peak counts computed in aperture mass maps outperform third and fourth-order moments for this purpose. Using CFHTLenS data Petri et al. 2015, obtained constraints on the parameter set (Ω_m, w, σ_8) using a set of moments up to 4^{th} order and find that high-order moments can break the (Ω_m, σ_8) degeneracy and provide tight constraints with no apparent bias. Nevertheless, high-order moments are very sensitive to outliers and may hence suffer high empirical variance meaning that if the distribution of field intensity has a long tail, the amount of information accessible to n-point functions quickly decrease.

2.3.5 Weak lensing peaks

A high order statistics that has become very popular and more and more adopted in the recent years in weak lensing analysis is peak-count statistics. It can be obtained by counting the local maxima of shear or convergence maps specifically by selecting all pixels that have a higher value than their eight neighbour pixels in 2D. Such local maxima correspond to high signal regions, that are closely associated with massive structures along the line-of-sight. This subsection is dedicated to peak counts in a more detailed way, as peaks are one of the main protagonists of this thesis work, as will be described in Section 4.1.

Models for predicting peak counts

Providing an explicit expression for the peak function, namely a peak number density as a function of the pixels, is a research issue that gave rise to a rich literature in the last decades. At the same time, being able to provide a complete and satisfying theoretical prediction for weak lensing peak counts as a function of the cosmological model, still represents a challenge. Jain and Waerbeke 2000 measured the distribution of peaks from simulated lensing surveys with a peak model based on Bardeen et al. 1986; Bond and Efstathiou 1987; Waerbeke 2000 and showed that the peaks distribution from lensing data contains information on the projected mass function of dark matter halos and represents a probe of non-Gaussian information beyond second order statistics. Later, Bartelmann, M., Perrotta, F., and Baccigalupi, C. 2002 proposed weak-lensing halo counts with the aperture-mass technique as a new method for constraining the nature of dark energy.

Then, Maturi, M. et al. 2010 introduced a new approach based on the peak theory of the Gaussian random field (Bardeen et al. 1986; Bond and Efstathiou 1987) providing a simple analytical recipe to compute the signal-to-noise distribution of detections caused by chance projections of large-scale structures as well as the shape and shot noise contributions by the background galaxies. The same year, Fan, Shan, and Liu 2010 proposed a model to calculate the number of peaks in large-scale reconstructed convergence maps from weak-lensing observations, including both true peaks corresponding to clusters of galaxies and false peaks from intrinsic ellipticities of background galaxies based again on the Gaussian peak theory. However, while in this study peaks are identified as local maxima in the map, in Maturi, M. et al. 2010 a different definition simplifying the analytical expression is given, according to which peaks are determined as contiguous areas with values above a given threshold. Cardone et al. 2013 computed the number density of haloes based on the theoretical mass function to constrain $f(R)$ theories and then the total number of peaks by integrating over the full redshift range, smoothing out the dependence of the mass function on the redshift.

Whilst being the only way to provide explicit expressions for the peak function, analytical models present some downsides. Specifically, when working in realistic conditions, such as in the presence of mask, or the bias from photometric redshifts, as well as the errors from shape measurement, the performance of these models can be strongly limited. Moreover, the link between cosmology and observed lensing peaks is made even more complicated because of some complexity from halo geometry, projections effect and shape noise. To correct for large-scale structure projection effects, Yuan et al. 2018 extended the model of Fan, Shan, and Liu 2010 by incorporating these projection effects by modelling them as a random field, i.e. by adding the stochastic large-scale structure contribution to the convergence map from the halo plus the shape noise. However, being able to model realistic survey settings, including additional cosmological or astrophysical features as for example IA and errors in general still remains challenging for analytical models.

A way to circumvent the difficulties that arise when trying to predict peak counts with analytical models, relies on generating weak lensing simulations densely sampled in cosmological parameter space. For example, Dietrich and Hartlap 2010 generated a set of N-body simulations in the $(\Omega_{\mathrm{m}}, \sigma_8)$ plan for 158 different cosmologies and interpolated using bilinear smoothing splines their shear peak function across parameter space. Liu J. et al. 2015 used a suite of ray-tracing N-body simulations on a grid of 91 cosmological models, varying the three parameters $(\Omega_{\mathrm{m}}, \sigma_8, w)$ and interpolated from the 91 measured peak counts to other cosmological models using two different methods (multi-dimensional Radial Basis Function and Gaussian Processes). Kacprzak et al. 2016 used the same set from Liu J. et al. 2015 and similar interpolation methods to predict the peak counts. With the advent of the `MassiveNus` simulations (Liu et al. 2018), predictions for peak counts based on simulations with a dependence on the sum of neutrino masses M_ν have also become possible, as in Li, Liu, Zorrilla Matilla, et al. 2019 and as we did in Ajani et al. 2020, as will be discussed in Chapter 4, with a set of 100 different cosmological models. Another recent set of cosmological simulations is the *cosmo*-SLICS suite of simulations with 25 cosmologies in the $[\Omega_{\mathrm{m}}, \sigma_8, h, w_0]$ parameter space, that has been used in Harnois-Déraps et al. 2020 to predict weak lensing shear peaks. Somewhere in between analytical models and N-body simulations, a halo-based model that generates fast stochastic simulations of convergence maps called CAMELUS was proposed by Lin, Chieh-An and Kilbinger, Martin 2015. It's a semi-analytic model, requiring only a halo profile and a halo mass function as input to generate a lensing catalogue. It consists in assuming that the main contribution to the lensing signal comes from halos and it discretises the space between the lensed galaxies and the observer in redshift bins, populating them with halos following a Navarro-Frenk-White (NFW) density profile and placing them randomly within its redshift bin. This method was used by Peel et al. 2017a to assess the constraining power of peak counts in a simulated *Euclid*-like survey on the cosmological parameters $(\Omega_{\mathrm{m}}, \sigma_8, w_0)$. They found that using CAMELUS, peak statistics yield a tight but significantly biased constraint in the $(\Omega_{\mathrm{m}}, \sigma_8)$ plane, indicating the need to better understand and control the model's systematics before applying it to a real surveys. Furthermore, Zorrilla Matilla et al. 2016 investigated how CAMELUS predicts the mean abundance of high-S/N peaks compared to N-body simulations, and found reasonable agreement, but that further development is needed to accurately estimate CAMELUS's variance and to probe the low-S/N tail.

Considered the above discussion, despite being costly to run, N-body simulations present several advantages with respect to analytical or semi-analytical models such as the possibility to incorporate the exact survey mask and shape noise in an easy way when one needs to deal with real data. Next section presents a brief overview of the results obtained using peak counts on real data.

State of the art of peaks on real data

The first cosmological constraints from peak analyses were performed approximately at the same time and in an independent way by Liu X. et al. 2015 and Liu J. et al. 2015. Liu X.

et al. 2015 derived constraints on cosmological parameters using weak lensing peak statistics measured on the $\sim 130\ \mathrm{deg}^2$ of the Canada-France-Hawaii Telescope Stripe 82 Survey (CS82) employing the model from Fan, Shan, and Liu 2010 described in the previous paragraph. They found that for flat ΛCDM, the cosmological constraints on $(\Omega_{\mathrm{m}}, \sigma_8)$ from peak analyses are fully consistent with the constraints obtained with both WMAP9 and Planck results and from cosmic shear two-point correlation studies. At the same time, they found that the degeneracy direction of the two parameters is flatter than those from the 2PCF analyses and deduced from this the complementarity of peak counts analysis to 2PCF studies. Liu J. et al. 2015 derived constraints on $(\Omega_{\mathrm{m}}, \sigma_8, w)$ for CFHTLenS $\sim 154\ \mathrm{deg}^2$ data. Their studies are based on interpolations from the suite of simulation templates on a grid of 91 cosmological models introduced in the previous paragraph. From the outcome of this study resulted that constraints from peak counts are comparable to those from the power spectrum, and that when the two statistics are combined, the area of the contours in the $(\Omega_{\mathrm{m}}, \sigma_8)$ plane is reduced by a factor of ~ 2. Moreover, neither observable can constrain w without further external data. Examining the effects of additional smoothing scales, binning of peaks, and the robustness of the results under masking, they demonstrated that using multiple smoothing scales can reduce the size of the area of the error contour by a moderate amount and left to more detailed study the quantification of the impact of these choices on the parameter constraints. With the same methodology adopted in Liu X. et al. 2015, Liu et al. 2016 presented constraints on the Hu-Sawicki $f(R)$ theory with CFHTLenS data. For DES SV data, Kacprzak et al. 2016 measured shear peaks using aperture mass maps adopting the simulation approach to produce WL maps (Dietrich and Hartlap 2010) spanning the $(\Omega_{\mathrm{m}}, \sigma_8)$ plane. They found that, including systematics, peaks perform roughly as good as the power spectrum in a non-tomographic analysis. For the KiDS 450 deg^2 data release, two analysis were performed using peak counts: one obtaining the theoretical prediction using Fan's analytical model (Shan et al. 2017) and one employing N-body simulations (Martinet et al. 2018). Both studies give consistent estimates of the parameter $S_8 \equiv \sigma_8\sqrt{\Omega_m/0.3}$, however, I think it's important to consider that the analytical predictions of the first study, were calibrated on the same simulations employed in the study based on N-body simulations. Considering the combination of shear peaks with non-tomographic measurements of the shear two-point correlation functions brought to an improvement of 20% in the uncertainty on the parameter S_8 in the simulation based analysis, compared to the shear two-point correlation functions alone, highlighting the great potential of peaks as a cosmological probe. The first tomographic peak counts data analysis was recently performed by Harnois-Déraps et al. 2020, with a joint cosmic shear two-point shear correlation function and peak-counts analysis for DES-Y1 data release. To calibrate the cosmological dependence of peak counts they used the *cosmo*-SLICS suite of simulations (details on these simulations are also discussed in Chapter 6) covering the $(\Omega_{\mathrm{m}}, \sigma_8, h, w_0)$ parameter space. They assessed the impact of baryons with the *Magneticum* hydrodynamical simulations, quantified the effect of finite particle mass from high-resolution light-cones, and investigated the impact of intrinsic alignment by infusing intrinsic shapes to mock galaxies following a physical model based on the shape of dark matter halos. They also modelled source-lens clustering by comparing galaxy excess in peaks of different height, and found to have a negligible impact on their results given their peak selection criteria. They found a $\sim 20\%$ gain in precision compared to the 2PCF analysis for the constraints on S_8 when using peak counts. Being the first tomographic analysis with peak counts on real data, they confirm that including cross-redshift bins reinforces the constraints, once systematics such as IA are properly calibrated and conclude that for the peaks statistics to remain competitive with the 2PCF, further development is required in the modelling of baryon feedback.

In summary, it becomes more and more clear that peak counts represent a powerful tool, given their complementarity to second order statistics as the 2PCF or the power spectrum to constrain cosmological parameters, and this has been proven on real data applications. Now, in order for peaks statistics to remain competitive with the 2PCF, one challenge consists in further developing the modelling of baryon feedback. The *baryonification* method described in Section 2.2.3 could be one possible way to further model this, as well as methologies based on hydrodynamical simulations. Another crucial aspect that still has to be addressed concerns intrinsic alignment, as the inclusion in peaks analysis is still critical, and improving the IA modelling is necessary for future analyses. The model for IA adopted in Harnois-Déraps et al. 2020 remains too simple, and as they demonstrated, this effect has an impact on the inferred cosmology at the same level as baryonic effect that can reach 1σ level, limiting in a severe way the analysis. Further exciting work is hence needed to be able to properly model these systematics effects to maximise the exploitation of peak counts potential in cosmological inference.

2.3.6 Weak lensing minima

A complementary probe of non-Gaussian information is given by weak lensing minima that is the number counts of local minima in a lensing convergence map as a function of their depth. In practice, for a 2D map, weak lensing minima are computed as the pixels with values smaller than their eight surrounding pixels. Physically, they are associated to underdense regions (voids) in our universe, being hence complementary to the lensing peaks which are typically associated with massive regions.

Cosmological forecasts with weak lensing minima

The projected mass distribution around voids can be used as a tool for constraining models of dark matter, dark energy, and modified gravity (Hamaus, Sutter, and Wandelt 2014). Deriving forecasts on the parameters $(M_\nu, \Omega_\mathrm{m}, A_\mathrm{s})$ for an LSST-like survey, Coulton et al. 2020a proposed lensing minima as a novel statistic to extract cosmological information in weak lensing data as they probe the emptiest regions (voids) in of the universe and also motivating their choice with the fact that baryonic effects are expected to impact voids differently than overdense regions (peaks). Their results show how in their settings, lensing minima alone are slightly less constraining than lensing peaks but that when the two are combined, they produce significantly tighter constraints with respect to the power spectrum. In the same study, the BAHAMAS hydrodynamical simulations are used to study the effects of baryons on lensing minima and the finding is that baryonic processes tend do decrease the number of very negative minima and high, positive minima, while increasing the number of minima values closer to zero. The bias induced by baryonic effects appears to have a small impact on minima ($\sim 0.5\sigma$), while a more significant one on peak counts ($\sim 4\sigma$). Another recent study that forecasted the cosmological power of weak lensing voids was presentend by Davies et al. 2020. Using the *cosmo*-SLICS source catalogue down-sampled to match LSST specifications, they provide forecasts on the $(\Omega_{\rm m}, S_8, h, w_0)$ parameters and find that these are tighter than the constraints obtained with the 2PCF with the same simulated observational specifications.

These results show how weak lensing minima (voids) represent a promising probe for upcoming cosmological experiments and how when combined with peak counts they represent a very competitive summary statistics to constrain cosmological parameters.

Void definition debate

Since voids may have irregular shapes, the lensing signal depends on the identification scheme considered. A large number of quite different void-finding algorithms is present in the literature, constructed using particles, haloes, and semi-analytical model galaxies extracted from simulations. For example, Bos et al. 2012 used a scheme called Watershed Void Finder based on halos and galaxy distributions to investigate whether void shapes can be used as a probe of the properties of dark energy. They found that the potential of voids as probe for dark energy, when they are identified through halos and galaxies distributions is limited by the discrete, sparse and biased nature of these objects. Another example is the study performed by Davies, Cautun, and Li 2018, where a higher lensing signal for voids is identified using lensing peaks, compared to the one found for voids identified using galaxies. Indeed, since the full mass distribution is not easily observable, observational studies make use of the galaxy distribution to identify voids. This approach can lead to galaxy voids less underdense than one would have using the full matter density field. This was one of the reason of Coulton et al. 2020a for introducing weak lensing minima avoiding void tracers such as halos or lensing peaks. However, this method is based on the assumption that lensing minima are mostly associated with negative lensing signal, but in reality, a small number have imprints from slightly overdense regions as discussed by Coulton et al. 2020a. A detailed comparison of different void finders present in the literature is provided by Colberg, Pearce, Foster, et al. 2008.

2.3.7 Machine learning based methods

In the recent years, with the massive advances in the field of machine learning (ML), machine learning and deep learning (DL) techniques have been proposed as analysis tools for cosmology. In particular Schmelzle et al. 2017 demostrated how a Deep Convolutional Neural Networks (DCNN) can be used to discriminate among cosmologies by learning the relation between five cosmological models and the corresponding convergence maps. By comparing the performance of this approach to commonly used non-Gaussian statistics as the skewness and kurtosis (defined in Section 2.3.4) of the convergence maps, they found that the DCNN implemented in their study outperforms the skewness and kurtosis statistics, especially for high noise levels and can efficiently break the $(\sigma_8, \Omega_\mathrm{m})$ degeneracy with weak lensing convergence maps alone. Another study performed by Gupta, Matilla, Hsu, et al. 2018 applied a convolutional neural network to simulated noiseless lensing maps covering 96 different cosmological models with varying $(\sigma_8, \Omega_\mathrm{m})$ and showed that their neural network could yield five times tighter constraints than the power spectrum, and four times tighter than the lensing peaks. A further step was then achieved by Ribli, Pataki, and Csabai 2018 that improved the convolutional neural network architecture to recover cosmological parameters more accurately from simulated weak lensing maps. Then, Fluri et al. 2018 explored the constraining power of CNNs compared to a standard power spectrum analysis for varying levels of noise and smoothing scales applied to convergence maps for the same set of parameters. They showed how the network is able to significantly improve the constraints with respect to the power spectrum, but also how the gain tends to decrease as the noise level and smoothing scales adopted increase. Peel et al. 2019 explored whether a neural network can discriminate better than peak counts between $\Lambda\mathrm{CDM}$ and modified gravity models with massive neutrinos and found that it can potentially discriminate between such models better than conventional higher-order statistics. Merten et al. 2019 investigated observational degeneracies between nine models of modified gravity and massive neutrinos using three different machine learning techniques and compared the performance with classical summary statistics as power spectrum, peak counts and Minkowski functionals. Their outcome is that the CNN provides the most discriminating representation of the data in a optimal noise-free maps setting. Matilla et al. 2020 analysed the neural network employed in Ribli et al. 2019 and compared its performance with constraints obtained from a combination of three commonly used WL statistics (power spectrum, lensing peaks, and Minkowski functionals), finding that the network can improve the inferred parameter constraints relative to this combination by 20% even in the presence of realistic levels of shape noise. However, the study needs to be verified under more realistic treatments that include realistic galaxy distributions and noise, as well as systematic errors. The first study that used a CNN to measure the cosmological parameters from observed weak lensing data has been performed for KiDS-450 (Fluri et al. 2019). By comparing the results obtained with the CNN to the ones obtained with the power spectrum analysis on the same maps they found an improvement of about 30% on S_8 for the CNN.

In summary, machine learning based methods can be promising tools to retrieve cosmological information beyond second-order statistics efficiently. Moreover, the advancement in parallelised training strategies opens the gates to the possibility of training of larger networks so that the geometry of large survey (as next generation surveys) can be taken into account. Despite this enormous potential, these techniques still need a huge work of investigation to become a robust and accepted tool by the community as when applied to real data for example, systematic errors that are not taken into account in the training process are difficult to get checked and controlled. Moreover, the large number of fitted parameters involved, and the depth of the many layers of non-linearities, on one side enable the learning of complex non-linear relationships between data and the parameters, while on the other side are still very difficult to interpret from a physical point of view.

2.3.8 The scattering transform, wavelet-phase harmonic statistics

Very recently, simultaneously and independently of the work presented in the thesis, two high order summary statistics based on wavelet transform were introduced: the *scattering transform* (Cheng et al. 2020) and Wavelet Phase Harmonics (WPH) statistics (Allys et al. 2020). The first one shares ideas with convolutional neural networks but doesn't require training nor tuning. More specifically, it employs the scattering transform introduced by Mallat 2012 and consists in a wavelet convolution of the weak lensing input image followed by a pixelwise modulus. The scattering operation is then iterated at different scales and finally the spatial average of every scattering image is taken. The constraining power of this statistic for the $(\Omega_\mathrm{m}, \sigma_8)$ plane has been compared to the one of the power spectrum, peak counts and the CNN adopted in Ribli et al. 2019. Their outcome is that on convergence maps with and without galaxy shape noise, the scattering transform outperforms the power spectrum and peak counts, and performs as well as the state-of-the-art CNNs. Then, Cheng and Menard 2021 have tested its performance extending the parameter space also to the dark energy equation of state parameter w_0 using the set of Dark Energy simulations from the Columbia Lensing Group and to the sum of neutrino masses M_ν using the `MassiveNus` simulations and compared their performance with respect to the power spectrum and the bispectrum. They found an improvement for all parameters, and their results on M_ν are consistent with the improvement we have found in both our studies for multi-scale higher order statistics (Ajani et al. 2020; Ajani et al. 2021), as they also state in Cheng and Menard 2021.

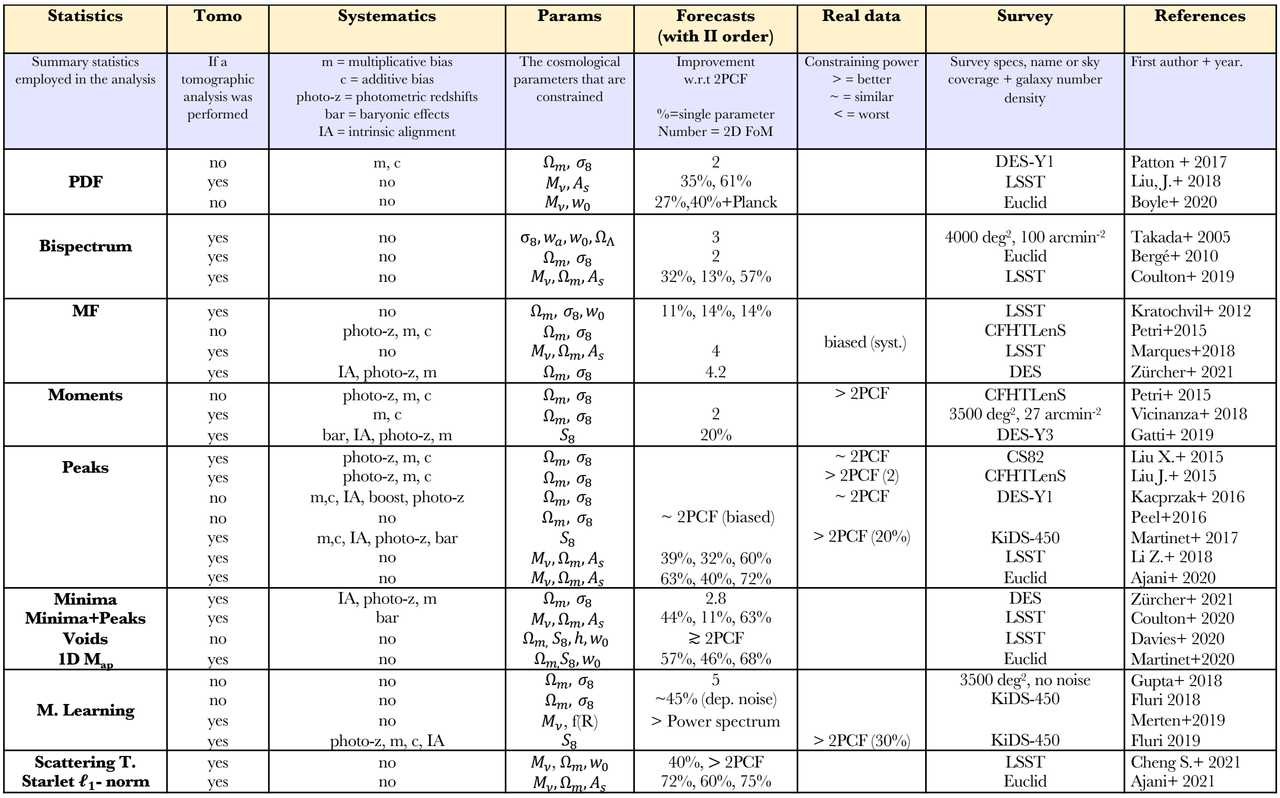
The second study from Allys et al. 2020 applied Wavelet Phase Harmonics (WPH) (Mallat, Zhang, and Rochette 2019) to constrain the parameters $(\Omega_\mathrm{m}, \Omega_\mathrm{b}, h, n_\mathrm{s}, \sigma_8)$ using the Qujote simulations (Villaescusa-Navarro et al. 2020). The computation of this statistics consists in applying the phase harmonic operator to the 2D projection of the large-scale structure matter density field provided from the simulations and convolved with the bump steerable wavelets introduced in Mallat, Zhang, and Rochette 2019. By computing Fisher information matrices, they find that the WPH statistics lead to stringent constraints on all parameters except for σ_8 when compared to statistics based on the combination of the power spectrum and bispectrum.

As in our works (Ajani et al. 2020; Ajani et al. 2021), that are presented in details in Chapter 4 and Chapter 5, these two statistics exploit the advantages of performing multi-scale analysis, thanks to the properties of wavelet functions. From these results, is becoming more and more clear that the possibility of extracting cosmological information from different scales simultaneously can bring great improvement to the cosmological constraints and further effort should be put in testing the robustness of such multi-scale statistics with respect to systematics, to prepare them to be applied to real data for future galaxy surveys.

2.3.9 Summary

As second-order statistics fail to fully capture the non-Gaussian information encoded in the non-linear features of weak lensing data, an increasing number of statistics of order higher than second have been proposed by the weak lensing community. Depending on the survey settings, systematics, noise level and cosmological parameter considered, they can perform in different ways with respect to each other in constraining cosmological parameters. I show here an attempt of summary of the state of the art of studies that have provided clear quantification of the performance of these statistics with respect to second order statistics for forecasts as well as for real data analysis, when available.

- **Statistics**: lists the names of the summary statistics employed to perform cosmological parameter inference.
- **Tomography**: this column indicates whether a tomographic study has been performed or not.
- **Systematics**: indicates if systematic effects have been included in the analysis and if so which ones.
- **Parameters**: indicates which cosmological parameters have been constrained in the analysis. Notice that some analysis might have taken into account certain parameters but couldn't constrain them (e.g. w_0) for CFHTLenS peak counts. In this case the non-constrained parameters are not listed. Only parameters that have been constrained are listed. For clarity and for the purposes of this summary, only references that present a clear quantification of the improvement in constraining power with respect to second order statistics are shown.
- **Forecasts**: specifies the improvement for forecasts analysis (i.e. without the use of real data) for the high order statistics combined with second order statistics w.r.t. the performance of second order statistics alone. The percentages indicates the improvement for the marginalised 1D probability of the corresponding parameter, the



Statistics	Tomo	Systematics	Params	Forecasts (with II order)	Real data	Survey	References
Summary statistics employed in the analysis	If a tomographic analysis was performed	m = multiplicative bias c = additive bias photo-z = photometric redshifts bar = baryonic effects IA = intrinsic alignment	The cosmological parameters that are constrained	Improvement w.r.t 2PCF %=single parameter Number = 2D FoM	Constraining power > = better ~ = similar < = worst	Survey specs, name or sky coverage + galaxy number density	First author + year.
PDF	no	m, c	Ω_m, σ_8	2		DES-Y1	Patton + 2017
	yes	no	M_ν, A_s	35%, 61%		LSST	Liu, J.+ 2018
	no	no	M_ν, w_0	27%,40%+Planck		Euclid	Boyle+ 2020
Bispectrum	yes	no	$\sigma_8, w_a, w_0, \Omega_\Lambda$	3		4000 deg^2, 100 arcmin^{-2}	Takada+ 2005
	yes	no	Ω_m, σ_8	2		Euclid	Bergé+ 2010
	yes	no	M_ν, Ω_m, A_s	32%, 13%, 57%		LSST	Coulton+ 2019
MF	yes	no	Ω_m, σ_8, w_0	11%, 14%, 14%		LSST	Kratochvil+ 2012
	no	photo-z, m, c	Ω_m, σ_8		biased (syst.)	CFHTLenS	Petri+2015
	yes	no	M_ν, Ω_m, A_s	4		LSST	Marques+2018
	yes	IA, photo-z, m	Ω_m, σ_8	4.2		DES	Zürcher+ 2021
Moments	no	photo-z, m, c	Ω_m, σ_8		> 2PCF	CFHTLenS	Petri+ 2015
	yes	m, c	Ω_m, σ_8	2		3500 deg^2, 27 arcmin^{-2}	Vicinanza+ 2018
	yes	bar, IA, photo-z, m	S_8	20%		DES-Y3	Gatti+ 2019
Peaks	yes	photo-z, m, c	Ω_m, σ_8		~ 2PCF	CS82	Liu X.+ 2015
	yes	photo-z, m, c	Ω_m, σ_8		> 2PCF (2)	CFHTLenS	Liu J.+ 2015
	no	m,c, IA, boost, photo-z	Ω_m, σ_8		~ 2PCF	DES-Y1	Kacprzak+ 2016
	no	no	Ω_m, σ_8	~ 2PCF (biased)			Peel+2016
	yes	m,c, IA, photo-z, bar	S_8		> 2PCF (20%)	KiDS-450	Martinet+ 2017
	yes	no	M_ν, Ω_m, A_s	39%, 32%, 60%		LSST	Li Z.+ 2018
	yes	no	M_ν, Ω_m, A_s	63%, 40%, 72%		Euclid	Ajani+ 2020
Minima	yes	IA, photo-z, m	Ω_m, σ_8	2.8		DES	Zürcher+ 2021
Minima+Peaks	yes	bar	M_ν, Ω_m, A_s	44%, 11%, 63%		LSST	Coulton+ 2020
Voids	no	no	Ω_m, S_8, h, w_0	≳ 2PCF		LSST	Davies+ 2020
1D M_{ap}	yes	no	Ω_m, S_8, w_0	57%, 46%, 68%		Euclid	Martinet+2020
M. Learning	no	no	Ω_m, σ_8	5		3500 deg^2, no noise	Gupta+ 2018
	no	no	Ω_m, σ_8	~45% (dep. noise)		KiDS-450	Fluri 2018
	yes	no	M_ν, f(R)	> Power spectrum			Merten+2019
	yes	photo-z, m, c, IA	S_8		> 2PCF (30%)	KiDS-450	Fluri 2019
Scattering T.	yes	no	M_ν, Ω_m, w_0	40%, > 2PCF		LSST	Cheng S.+ 2021
Starlet ℓ_1- norm	yes	no	M_ν, Ω_m, A_s	72%, 60%, 75%		Euclid	Ajani+ 2021

Figure 2.4: Table summarising the current state of the art of high order statistics where an estimate of the constraining power and its performance with respect to second order statistics have been provided by the authors. I would like to stress that this table does not include all studies ever performed with statistics of order higher than the second but the studies where cosmological inference has been performed *and* an explicit comparison with second order statistics is provided in the paper.

.

integer numbers illustrate the improvement of the figure of merit. If a generic "≥", "∼" is shown, it means that the exact improvement has not been mentioned by the authors.

- **Real data**: report the results in terms of constraining power found in real data analysis.
- **Survey**: indicates the name of the survey if the survey settings used in the analysis are explicitly associated by the authors to a specific survey, otherwise the sky coverage and galaxy number density employed are reported.
- **References**: indicates the corresponding study, identified by first author name + year.

CHAPTER 3

WAVELET FORMALISM FOR ASTROPHYSICAL IMAGES

The aim of this chapter is to introduce some definitions useful for the purposes of this thesis in the context of wavelet formalism. In particular, I start by giving a brief introduction of the concept of sparsity, of the continuous wavelet transform and of the discrete wavelet transform. Thanks to these tools, the starlet transform, that will be widely use in Chapter 4, Chapter 5 and Chapter 6 is introduced. Then, the procedure followed to estimate the noise and compute the wavelet coefficients is illustrated. Finally, a brief overview of recent mass mapping methods based on sparse representations is provided.

3.1 Sparse signal representation

In a context where the goal is to try to maximise the information that we can extract from cosmological data, there are different representations of the data that can be considered. Indeed, any given signal can be represented in a variety of different domains without losing information, but depending on the properties of the signal and on the application, a given representation may be preferred. One very well established representation is the sparsity model, that has proven to be very powerful for astrophysical applications (Starck, Murtagh, and Fadili 2016; Lanusse 2015; Lin, C.-A. et al. 2016). The following is based on Starck, Fadili, and Murtagh 2007; Starck, Murtagh, and Fadili 2016; Starck et al. 2021.

Sparsity data model

A signal X, $X = [x_1, ..., x_N]^T$ is defined *sparse* if most of its entries are equal to zero. More specifically, a $k-$sparse signal is a signal with only k non-zero samples values. It is also possible to give a less stringent definition when considering a signal where only a few of its entries have a large magnitude and the majority of them are close to zero: in this case, one can define the signal as *weakly sparse* or *compressible*. Hence, in a sparse signal, either only a finite number of coefficients is non zero, or the coefficients decrease fast when rank-ordered. If a signal is not sparse, it can be *sparsified* using a given data representation. Intuitively, this can be understood by thinking to periodic signals as for example $sin(x)$ and by looking at Figure 3.1: in real space, sine functions are not sparse, but they are sparse in Fourier space, in particular "1-sparse", as they become the Dirac delta functions.

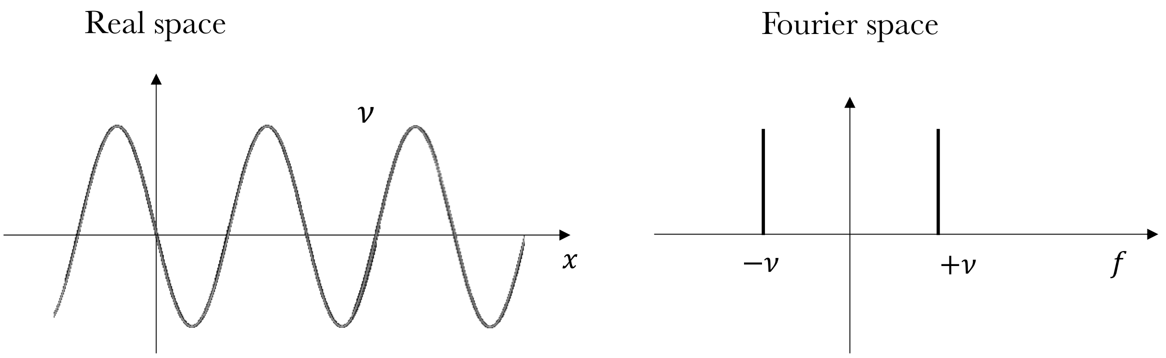


Figure 3.1: The periodic function $sin(x)$ is not sparse in real space (**left image**) but it can be *sparsified* by changing the data representation and work instead in Fourier space (**right image**).

Having this in mind, a signal X is defined *sparse in the Fourier domain* if it has sparse Fourier coefficients, defined as $\hat{X}[u] = \frac{1}{N}\sum_k X[k]e^{2\pi \frac{uk}{N}}$. This concept can be generalised by introducing a dictionary $\mathbf{\Phi} = [\phi_1, ..., \phi_M]$, whose elements are T elementary waveforms that enable to model the signal $X \in \mathbb{R}^{\mathrm{N}}$ as the linear combination:

$$X = \mathbf{\Phi}\alpha = \sum_{i=1}^{M} \alpha[i]\phi_i, \tag{3.1}$$

where $\alpha \in \mathbb{R}^{\mathrm{M}}$ are the decomposition coefficients of the signal X in the dictionary $\mathbf{\Phi}$. This is of interest for this work because, if one can find a good representation that allows to separate the signal from the noise efficiently, this leads to a gain of extracted information. The representation basis treated in this work is the wavelet *starlet* transform, as it is very well suited for astronomical images that often contain isotropic sources (Starck, Fadili, and Murtagh 2007).

3.2 Continuous Wavelet transform

This section provides an introduction to the fundamentals of the continuous and discrete wavelet transforms, in order to have the tools to introduce and understand the starlet transform, that will be used in the works presented in this thesis. A wavelet is a function $\psi(x)$ satisfying the *admissibility condition*:

$$\int_0^\infty |\hat{\psi}(k)|^2 \frac{dk}{k} < \infty, \tag{3.2}$$

where $\hat{\psi}(k)$ is the Fourier transform of $\psi(x)$. In order for Equation 3.2 to be satisfied, ψ has to satisfy the condition $\int \psi(x)dx = 0$. This condition implies that wavelets are highly localized functions with vanishing mean. Wavelet transforms decompose a signal into a family of scaled and translated functions by generating a family of *daughter wavelets*:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right), \tag{3.3}$$

where $\psi^\star$ is the complex conjugate of ψ, a is a scaling parameter and b is a translation (or position) parameter. Hence, $\psi_{a,b}$ are scaled and translated versions of ψ, called the *mother* wavelet. Given a scale parameter, $a > 0$, and the translation parameter $b \in \mathbb{R}$, the continuous wavelet transform (CWT)(Grossmann and Morlet 1984) of a function $f(x)$ can be defined as:

$$\mathcal{W}_\psi(a,b)(f) = \int_{-\infty}^{\infty} f(x)\psi^\star_{a,b}(x)dx = \frac{1}{\sqrt{a}}\int_{-\infty}^{\infty} f(x)\psi^\star\left(\frac{x-b}{a}\right)dx. \tag{3.4}$$

By varying continuously the scale parameter a and the position parameter b one can obtain the wavelet coefficients for a given scale and position. Intuitively, one can think at this transform by considering how Fourier transforms measure the similarity between a signal and an analysing function. In the Fourier transform case, the analysing functions are complex exponential $e^{j\omega t}$, while in the case of the continuous wavelet transform, the analysing function is a wavelet, ψ, that compares the signal to shifted and compressed or dilated versions of the mother wavelet. In other words, as the Fourier transform outputs coefficients as a function of the frequency, the wavelet transform outputs coefficients which are a function of their scale and translation. Another nice, canonical way, commonly used in lectures on wavelets, to understand the properties of the wavelet function, is to consider the information in terms of time and frequency that can be extracted through time series analysis, Fourier transforms, short-time Fourier transforms (STFT) and wavelet transforms, as shown in Figure 3.2. From time series analysis (top left panel of Figure 3.2), namely the analysis of a sequence of data points (signal) that occur in successive order over some period of time, information on the signal is provided for every single point in time, missing however information about its frequency. The Fourier transform (top right panel) converts the time information in the frequency domain, providing great frequency resolution and being well suited for stationary signals, but loosing track of the time information. To overcome the poor time resolution of the Fourier transform, the short-time Fourier transform (STFT) (Gabor 1946) was introduced. The STFT (bottom left panel of Figure 3.2) provides a time-frequency representation of the signal and it is based on the assumption that a portion of the non-stationary signal is stationary. In this way, a small portion of the signal is analysed at a time. Then the Fourier transform of each stationary portion is taken and the results are added up, to recover the entire signal. With this approach, one can recover frequency resolution for a given time resolution. This is done by taking a window function of fixed length that is moved along the signal. The drawback is hence that once the window size is chosen, it remains unchanged during the entire analysis. This leads to a trade-off between the time and frequency resolution that it is possible to derive: by choosing a narrow windowing in time, the time resolution will be good but the frequency resolution will decrease; by choosing instead a wider windowing in the time axis, the time resolution will decrease as the frequency resolution increases. This leads to an inevitable loss of information, as low frequency components usually last a longer time, requiring good resolution in frequency, while high frequency components are often shorten in time, invoking the need of good resolution in time. Wavelets transform improve this context, as they allow to analyse the signal into different frequencies at different resolutions (bottom right panel of Figure 3.2). By looking at the frequency-time plane in Figure 3.2, one can see how in the high frequencies regime, the resolution is good in time and poor in frequency. In other words, where the frequencies are high, the vertical lines are much denser than the horizontal lines. At the same time, at low frequencies, there is better resolution in frequencies and poorer resolution in time. In contrast to the STFT, the width and the central frequency of the wavelet can be changed by changing the scale parameter a, while moving across the signal. In this way, *expanded* wavelets, namely with large values of a, will be good in resolving low frequencies components of the signal with bad time resolution. Similarly, *narrower* wavelets, characterised by a smaller values of a, will be better in resolving high frequency components with better time resolution. So, when the frequency of the wavelet matches part of the signal with a very similar frequency, the output of the wavelet coefficients $\mathcal{W}_{\psi}(a,b)(f)$ is high. In practice, starting from smaller to increasing scales a, the signal is anaysed from high frequencies to low frequencies: high frequency wavelets will trace the high frequencies of the signal, and low frequencies wavelets will detect low frequency of the signal.

3.3 Discrete Wavelet transforms

In the case of CWT, the scale parameter a and the translation parameter b are continuous, implying that a huge amount of wavelet coefficients is generated. To reduce the amount of data generated, one can consider discrete values for the parameters a and b. With this in mind, a discrete wavelet transform (DWT) can be defined through the concept of *multiresolution analysis* (MRA) (Mallat 1989; Meyer 1992). In this framework, the scale and translation parameters are chosen so that they are based on powers of two (dyadic). The main idea of MRA is to build a sequence of approximations f_j of a function $f \in L^2(\mathbb{R})$ by smoothing it with a kernel of increasing width proportional to 2^j. Each of these approximations then belong to a subspace that contains all possible approximations at scale 2^j. For each approximation subspace is then possible to build an orthogonal subspace containing all the details that are lost between two consecutive approximations. Performing this procedure recursively enables to decompose the entire $L^2(\mathbb{R})$ space into a direct sum of the orthogonal subspace containing all the details that were lost between two consecutive approximations. Then for each $j \in \mathbb{Z}$, it is possible to build a wavelet ψ, such that the family of wavelets:

$$\left\{ \psi_{j,n} = \frac{1}{\sqrt{2^j}} \psi \left(\frac{t - 2^j n}{2^j} \right) \right\}_{n \in \mathbb{Z}}, \tag{3.5}$$

is an orthonormal basis of the subspace containing all the details that are lost between two consecutive approximations. The scale parameter is now $a = 2^j$ and the translation (position) parameter $b = 2^j n$. In practice, what happens when analysing a signal f with a DWT, can be understood considering *multilevel decomposition*, reading Figure 3.3, as follows. The input signal f is passed into a low pass filter h and a high pass filter g, so that low pass filters will select the low frequency components while rejecting the high frequency components and *vice-versa*. The low pass portions of the signal are called *approximation* coefficients a_j and they are iteratively filtered by the low pass filter and the high pass filter. The output of the high pass filters are instead the *detail* (or wavelet) coefficients d_j, corresponding to high frequencies information. At each level of this process the output is downsampled in a way that its size is halved. This last step is called *decimation* and takes place when a decimated DWT is used to analyse the signal. The result of the entire process is a set of approxima-

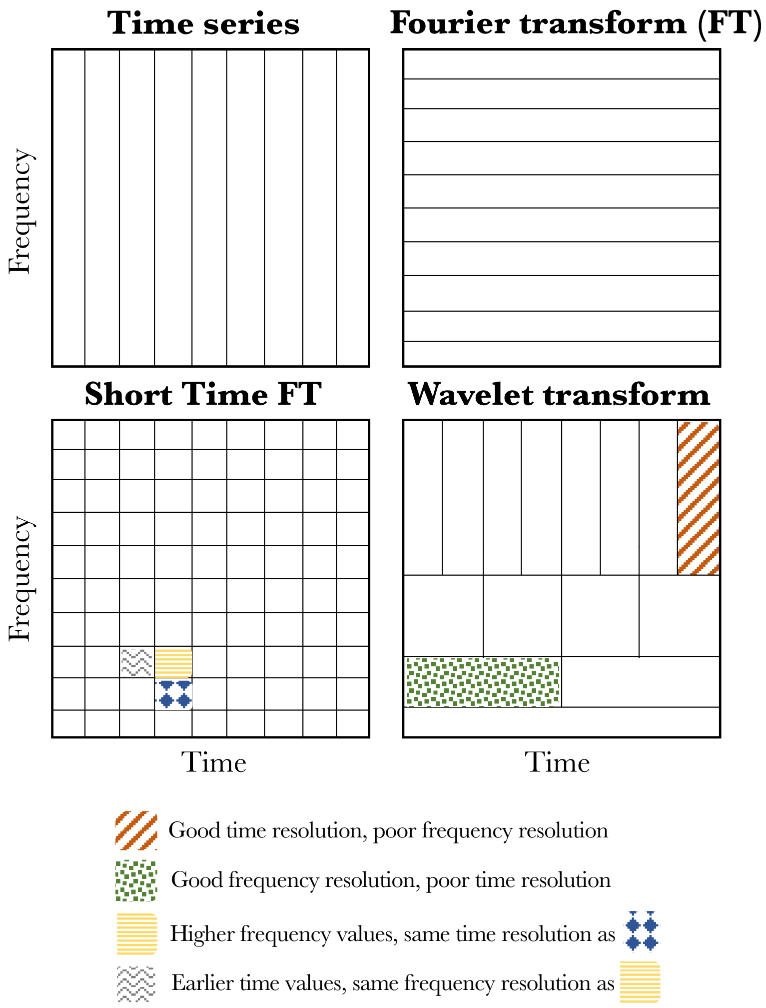


Figure 3.2: Illustration of the time-frequency diagram for Time series analysis (**top left panel**), Fourier transform (**top right panel**), Short Time Fourier transform (STFT, **bottom left panel**), the Wavelet transform (**bottom right panel**). As desribed in the text, time series analysis provide good resolution in time, but none in frequency; Fourier transform, convert the time information in frequency information, loosing all information on time; STFT provide both frequency-time resolution, with the drawback that the window is fixed, and inevitable loss of information is present because of this. The wavelet transform enables a *multiresolution analysis*, i.e. to analyse the signal into different frequency ranges at different time resolutions.

tion coefficients $\{a_1, a_2, a_3, .., a_j, ..a_J\}$ and a set of detail coefficients $\{d_1, d_2, d_3, .., d_j, ..d_J\}$, bringing respectively the low and high frequency components of the input signal.

To give a more formal description, a consequence of the multiresolution approximation

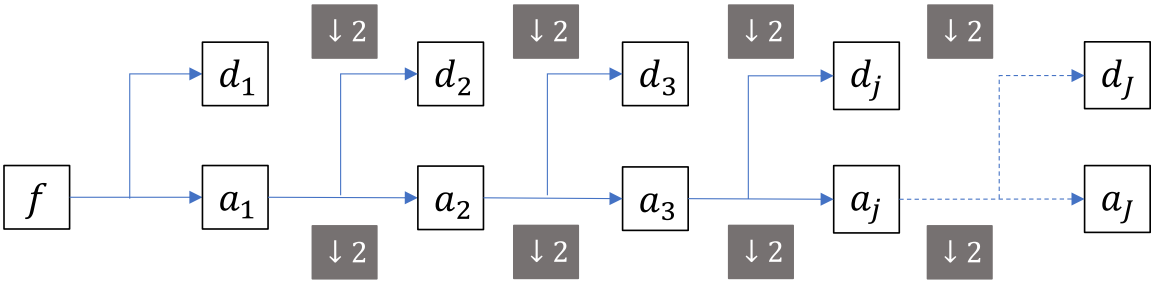


Figure 3.3: Illustration of the multilevel decomposition of the signal f. By passing through a low pass filter, the approximation coefficients $\{ a_1, a_2, a_3, .., a_j, ..a_J \}$ are generated. Each approximation coefficient is iteratively passed through a low pass filter generating the a_js and through a high pass filter that generates the detail coefficients $\{ d_1, d_2, d_3, .., d_j, ..d_J \}$. If a decimated DWT is employed, at each step the size of the coefficients is halved because the output is downsampled by a factor of 2.

(Mallat 1989; Meyer 1992) is that it is possible to build a *scaling* function ϕ that constitutes an orthonormal basis in the approximation subspace, such that

$$\phi_{j,n} = \frac{1}{\sqrt{2^j}} \phi \left(\frac{t-n}{2^j} \right), \tag{3.6}$$

for each $j \in \mathbb{Z}$, that enables to compute the approximations f_j of f at the scale 2^j as the orthogonal projection of f in the approximation subspace as follows:

$$f_j = \sum_{n \in \mathbb{Z}} a_j[n] \phi_{j,n}, \tag{3.7}$$

where $a_j[n]$ are the approximation coefficients given by $a_j[n] = \langle f, \phi_{j,n} \rangle$. Now, looking at the scaling functions defined in Equation 3.6, one can consider the scaled function $\phi_{1,0} = \frac{1}{\sqrt{2}} \phi \left(\frac{t}{2} \right)$, and by exploiting again the implications of the multiresolution approximation (Mallat 1989; Meyer 1992), it is possible to introduce the discrete filter h such that

$$\frac{1}{\sqrt{2}} \phi \left(\frac{t}{2} \right) = \sum_{n \in \mathbb{Z}} h[n] \phi(t-n), \tag{3.8}$$

where $h[n] = \langle \frac{1}{\sqrt{2}} \phi \left(\frac{t}{2} \right), \phi(t-n) \rangle$, namely the $\phi_{1,0}$ has been decomposed on the basis $\{\phi_{0,n}\}_{n \in \mathbb{Z}}$. This can be generalised to any scale 2^{j+1} and basis $\{\phi_{j,p}\}_{p \in \mathbb{Z}}$ in a way that the approximation coefficients at the scale $j+1$ can be obtained recursively from the coefficients at the previous scale j, as

$$a_{j+1}[p] = \sum_{n\in\mathbb{Z}} h[n-2p]a_j[n]. \tag{3.9}$$

Furthermore, it can be shown (Mallat 1989; Meyer 1992) that the projection of f onto the family of wavelets defined in Equation 3.5 forming an orthonormal basis of the subspace containing all the details that were lost between two consecutive approximations, enables to decompose f as

$$f = \sum_{j\in\mathbb{Z}}\sum_{n\in\mathbb{Z}} d_j[n]\psi_{j,n}, \tag{3.10}$$

through the details (wavelet) coefficients $d_j[n] = \langle f, \psi_{j,n}\rangle$, with scale parameter 2^j and position parameter $2^j n$. Then, analogously to Equation 3.9, it is possible to define a filter g to obtain the details coefficients at the scale $j+1$ from the approximation coefficients of the previous scale j:

$$d_{j+1}[p] = \sum_{n\in\mathbb{Z}} g[n-2p]a_j[n] \tag{3.11}$$

as schematically summarised in Figure 3.3. Depending on the properties of the filters g and h, the wavelet employed in the analysis has different properties that can be more or less convenient depending on the goal of the analysis. For the purposes of this thesis, we are interested in a particular type of DWT, called the *starlet* transform, introduced in the next section.

3.4 The starlet transform

The Isotropic Undecimated Wavelet Transform (IUWT) (Starck, Fadili, and Murtagh 2007), also called starlet transform, is an *undecimated* DWT, meaning that the decimation step introduced in the previous section is removed from its implementation. Moreover it is *isotropic*, meaning that

- the filters h and g must be symmetric, namely $\bar{h}[k] = h[-k] = h[k]$ and $\bar{g}[k] = g[-k] = g[k]$,
- in dimensions ≥ 2, the filters h, g, the scaling function ϕ and the wavelet ψ must be nearly isotropic.

À trous algorithm

The starlet decomposition can be implemented through the *à trous algorithm* (Holschneider et al. 1989) as follows. The approximations and details coefficients introduced in Equation 3.9 and Equation 3.11 are recursively computed as

$$a_{j+1}[k] = \sum_{l \in \mathbb{Z}} h[l] a_j[l - 2^j k], \tag{3.12}$$

$$d_{j+1}[k] = \sum_{l \in \mathbb{Z}} g[l] a_j[l - 2^j k], \tag{3.13}$$

where at each scale j different filters $h^{(j)}$ and $g^{(j)}$ are derived from h and g by inserting $2^j - 1$ zeros between each of the values, so that for $j = 1$ there will be one zero between each value of the filter, for $j = 2$ there will be three zeros.. and so on as

$$h^{(1)} = (..., h[-2], 0, h[-1], 0, h[0], 0, h[1], 0, h[2], ...)$$
$$h^{(2)} = (..., h[-1], 0, 0, 0, h[0], 0, 0, 0, h[1], 0, 0, 0, h[2], ...)$$

so that they become sparse. Then, as mentioned before, the scaling function ϕ in $D \geq 2$ must be isotropic and it is chosen to be the spline of order 3 (Daubechies 1992), i.e.

$$\phi_1(x) = \frac{1}{12}(|x-2|^3 - 4|x-1|^3 + 6|x|^3 - 4|x+1|^3 + |x+2|^3) \tag{3.14}$$

with $\phi(x, y) = \phi_1(x)\phi_1(y)$ and the wavelet function is built as the difference between two successive approximations of this isotropic scaling function as

$$\frac{1}{4}\psi\left(\frac{x}{2}, \frac{y}{2}\right) = \phi(x, y) - \frac{1}{4}\phi\left(\frac{x}{2}, \frac{y}{2}\right), \tag{3.15}$$

so that ψ is also isotropic. This is reflected also in the relation between the filters h and g, that is

$$g[k, l] = \delta[k, l] - h[k, l], \tag{3.16}$$

where $\delta[0, 0] = 1$ and $\delta[k, l] = 0$ otherwise and

$$h[k, l] = h^{(1\mathrm{D})}[k] h^{(1\mathrm{D})}[l], \tag{3.17}$$

with

$$h^{(1\mathrm{D})} = \frac{[1, 4, 6, 4, 1]}{16}. \tag{3.18}$$

Hence, when convolved with a 2D starlet, an original image *I* is decomposed by this transform into a coarse version of it a_J plus several images of the same size at different resolution scales *j*, as shown in Figure 3.4:

$$I[k, l] = a_0[k, l] = a_J[k, l] + \sum_{j=1}^{J} d_j[k, l], \tag{3.19}$$

where the wavelet images d_j represent the details of the original image at dyadic (powers of two) scales corresponding to a spatial size of 2^j pixels. The maximum j is determined by the map size as $\log_2 N$ for an $N \times N$ image. In practice, this is done by implementing the above discussion, by first considering the filter bank $h_{1D} = \{\frac{1}{16}, \frac{1}{4}, \frac{3}{8}, \frac{1}{4}, \frac{1}{16}\}$ and initializing $j = 0$. Each coefficient $a_{j+1}[k, l]$ is hence smoothed by convolving the previous coefficient $a_j[k, l]$ with the corresponding *sparsified* weight h_{1D} at the scale j. After this smoothing, the discrete wavelet transform is given by the difference $a_j[k, l] - a_{j+1}[k, l]$. More formally, each wavelet coefficient is obtained as:

$$a_{j+1}[k, l] = \sum_m \sum_n h_{1D}[m] h_{1D}[n] a_j[k + 2^j m, l + 2^j n] \tag{3.20}$$

and

$$d_{j+1}[k, l] = a_j[k, l] - a_{j+1}[k, l] \tag{3.21}$$

so that the ensemble of $\{d_1, d_2, ...a_J\}$ are the wavelet coefficients, i.e. the starlet transform of the original image.

Therefore, the IUWT is entirely defined by the choice of the scaling function ϕ and associated filter h. One advantage of employing a starlet filter is provided by its multi-scale analysis, namely its ability to investigate and extract the information encoded at different scales at the same time (Starck, Murtagh, and Bijaoui 1998). Moreover, compared to the DWT, it presents useful properties such as invariance per rotation and translation. Furthermore, as we are interested in optimising the extraction of information from weak lensing maps, the shape of this filter emphasises round features making it very efficient when dealing with for example with peak counts, which are point-like features. Moreover, peak counts (the signal) are sparse in the starlet domain, as they are described by a few number of non-zero coefficients, while white noise is not sparse in any representation. Hence, as we

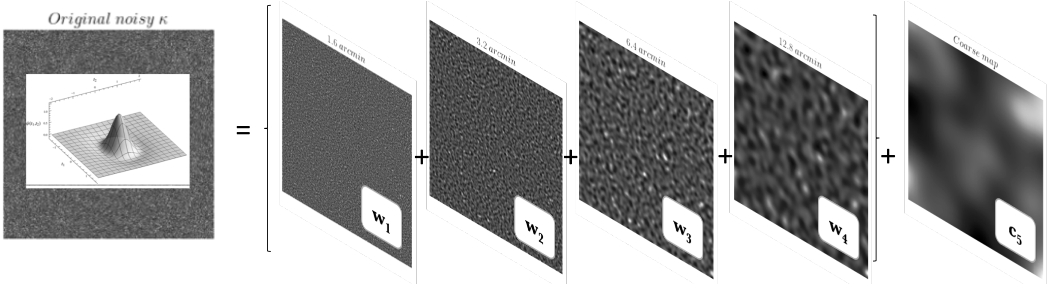


Figure 3.4: Illustration of the starlet decomposition with $j_{\max} = 4$ scales. As the j index increases, larger scales are considered and the resolution decreases.

are employing the starlet to smooth the noise to access the cosmological information from the maps at smaller scales where the noise is more dominant, it represents a powerful tool to separate the noise from the signal. Further details on the advantages of wavelet starlet analysis can be also found in Lin, C.-A. et al. 2016, Peel et al. 2017b and Peel, Pettorino, Giocoli, et al. 2018. Figure 3.5 shows the starlet 1D and 2D profiles.

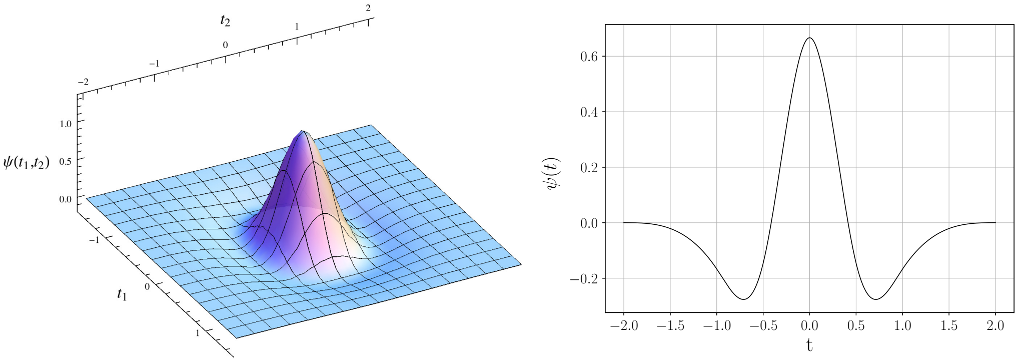


Figure 3.5: We show the 2D starlet function (top panel) as defined in Eq. (3.15) and its 1D profile (bottom panel). Being a wavelet it is a compensated function, i.e. it integrates to zero over its domain. This comes from the *admissibility condition* for the wavelet function ψ: $\int_0^{+\infty} |\hat{\psi(k)}|^2 \, \frac{\mathrm{d}k}{k} < +\infty$ which implies that $\int \psi(x)\mathrm{d}x = 0$ and it has compact support in $[-2,2] \times [-2,2]$. Its shape emphasises round features, making it very efficient when dealing with peak counts.

3.5 Wavelet coefficients and noise estimation

The summary statistics that we compute on the noisy weak lensing maps in Chapter 4 and Chapter 5, except for the power spectrum that is computed as function of the angular multipoles ℓ, are expressed as a function of the signal to noise ratio $\nu \equiv S/N$. Therefore,

Starlet scale	σ_j^e
d_1	0.8891
d_2	0.2017
d_3	0.086
d_4	0.0414
a_5	0.0324

Table 3.1: Values of the noise estimates for each starlet scale j for a map with $(n_{\rm pix}, n_{\rm pix})$ $= (512, 512)$ convolved with a starlet with $j_{\rm max} = 4$ scales.

it is necessary to be able to estimate the noise at each scale. We define the signal to noise field $\nu = S/N$ as the ratio between the noisy convergence κ convolved with the filter $\mathcal{W}$ over the smoothed standard deviation of the noise for each of the map we consider in our analysis:

$$\nu = \frac{(\mathcal{W} * \kappa)(\theta)}{\sigma_n^{\rm filt}(\theta)}. \tag{3.22}$$

If $\mathcal{W}$ is a Gaussian kernel, the denominator in the above equation will be given by the standard deviation of the noise smoothed with a Gaussian kernel of the same size as the one used to smooth the signal. For instance, if the noisy convergence map is smoothed with a Gaussian kernel of size $\theta_{\rm ker} = 1$ arcmin, then $\sigma_n^{\rm filt}(\theta)$ is the standard deviation of the noise smoothed with a Gaussian kernel of size $\theta_{\rm ker} = 1$ arcmin. When $\mathcal{W}$ is a starlet kernel instead, one needs to take into account the fact that the noise level varies at each scale. This can be done using the automatic noise estimation from the multiresolution support presented by Starck and Murtagh 1998. Specifically, the appropriate value of the noise at the scale j in the succession is assessed from the standard deviation of the noise in the original image σ_I and from the behavior of the noise in the wavelet space. The procedure to estimate the noise level σ_j at each starlet scale is the following:

- we define the number of scales for the starlet decomposition,
- we compute the coefficients σ_j^e by taking the starlet transform of a Gaussian distribution of the same size of the original image with standard deviation one,
- to estimate σ_I we take the *median absolute deviation*of the noisy convergence map.

The estimate of the noise level at each wavelet scale, is then given by $\sigma_j = \sigma_j^e \sigma_I$. An example for these values is provided in Table 3.1.

3.6 Wavelets and Dark Matter Mass Map Reconstruction

As described in Section 2.1.3, it is essential to be able to recover the convergence from the shear in a proper way. Mass mapping methods solve an ill-posed problem due to the irregular sampling of the lensing field and the low signal to-noise ratio on small scales (Lanusse et al. 2016), that consists in finding the unknown underlying convergence field $\boldsymbol{\kappa}$ (the signal) from observed observed shear data (noisy measurements), $\boldsymbol{\gamma}$:

$$\boldsymbol{\gamma} = \mathbf{A}\boldsymbol{\kappa} + \boldsymbol{n}, \tag{3.23}$$

where $\boldsymbol{n}$ is the statistical uncertainty vector associated with the data. As mentioned in Section 2.1.3, there exist mass mapping methods based on sparsity. For example, the Gravitational Lensing Inversion and MaPping using Sparse Estimators (GLIMPSE) (Leonard, Lanusse, and Starck 2014; Lanusse et al. 2016) relies on the assumption that the density field can be sparsely represented in an appropriate dictionary in order to regularise the inversion of the lensing operator. Being based on sparse representation, GLIMPSE is well suited for recovering piece-wise smooth features. A different approach, that relies on a Bayesian framework, is the one of Wiener filtering (Wiener 1949). In this approach a Gaussian random field is assumed as a prior for the convergence map, which is responsible for inserting some bias that prevents the estimate from over-fitting (Zaroubi et al. 1995). It has been shown (Lanusse et al. 2016; Price et al. 2019) that the solutions derived using a wavelet regularization significantly outperform standard techniques such as Wiener filtering for peaks recovery, while Wiener filtering performs better on the Gaussian content of the map (Jeffrey et al. 2018). As explained in Section 2.3, since the density field is inhomogeneous, it encompasses Gaussian-type large-scale structures, as well as non-Gaussian features. It is therefore essential for mass mapping methods to preserve both Gaussian and non-Gaussian features during the reconstruction process. Motivated by the above considerations, Starck et al. 2021 proposed a method, called MCALens, to model the matter density field using both linear and non-linear characteristics. More specifically, the idea is to consider that modelling the convergence map as a Gaussian random field leads to a good recovery of the large scales of the convergence map but not the non-Gaussian structures, while assuming the convergence map is compressible in the wavelet domain (i.e. sparse modelling) recovers non-Gaussian structures as peaks extremely well but poorly the Gaussian content. Hence, the sparse recovery is somehow complementary to the Wiener solution. To address these limitations, Starck et al. 2021 introduce a novel approach where the convergence field $\boldsymbol{\kappa}$ is given by a Gaussian and a non-Gaussian part:

$$\boldsymbol{\kappa} = \boldsymbol{\kappa}_{\mathrm{G}} + \boldsymbol{\kappa}_{\mathrm{NG}}. \tag{3.24}$$

The non-Gaussian part of the signal $\boldsymbol{\kappa}_{\mathrm{NG}}$ undergoes a sparse decomposition in a wavelet dictionary, while the component $\boldsymbol{\kappa}_{\mathrm{G}}$ is assumed to be inherently non-sparse and Gaussian. These two components are separated through the Morphological Component Analysis (MCA) proposed in Starck, Elad, and Donoho 2004 that allows to separate two components mixed in a single image when these have different morphological properties. If one can exploit the morphological differences between the components, it was shown that it is possible to extract and separate them.

This requires to define the functions $\boldsymbol{C}_{\mathrm{G}}$ and $\boldsymbol{C}_{\mathrm{NG}}$, such that the following is minimised

$$\min_{\boldsymbol{\kappa}_{\mathrm{G}},\boldsymbol{\kappa}_{\mathrm{NG}}} \left\{ \|\boldsymbol{\gamma} - \boldsymbol{A}(\boldsymbol{\kappa}_{\mathrm{G}} + \boldsymbol{\kappa}_{\mathrm{NG}})\|^2_{\boldsymbol{\Sigma}_n} + \boldsymbol{C}_{\mathrm{G}}(\boldsymbol{\kappa}_{\mathrm{G}}) + \boldsymbol{C}_{\mathrm{NG}}(\boldsymbol{\kappa}_{\mathrm{NG}}) \right\}, \tag{3.25}$$

where $\boldsymbol{\Sigma}_n$ is a pre-defined covariance matrix of the noise, $\boldsymbol{A}$ is the inversion matrix from Equation 3.23 and $\boldsymbol{\gamma}$ is the shear.

Morphological Component Analysis solves this by performing an alternating minimization scheme, as follows:

- estimate $\boldsymbol{\kappa}_{\mathrm{G}}$ assuming $\boldsymbol{\kappa}_{\mathrm{NG}}$ is known:

$$\min_{\boldsymbol{\kappa}_{\mathrm{G}}} \left\{ \|(\boldsymbol{\gamma} - \boldsymbol{A}\boldsymbol{\kappa}_{\mathrm{NG}}) - \boldsymbol{A}\boldsymbol{\kappa}_{\mathrm{G}})\|^2_{\boldsymbol{\Sigma}_n} + \boldsymbol{C}_{\mathrm{G}}(\boldsymbol{\kappa}_{\mathrm{G}}) \right\}, \tag{3.26}$$

- estimate $\boldsymbol{\kappa}_{\mathrm{NG}}$ assuming $\boldsymbol{\kappa}_{\mathrm{G}}$ is known:

$$\min_{\boldsymbol{\kappa}_{\mathrm{NG}}} \left\{ \|(\boldsymbol{\gamma} - \boldsymbol{A}\boldsymbol{\kappa}_{\mathrm{G}}) - \boldsymbol{A}\boldsymbol{\kappa}_{\mathrm{NG}})\|^2_{\boldsymbol{\Sigma}_n} + \boldsymbol{C}_{\mathrm{NG}}(\boldsymbol{\kappa}_{\mathrm{NG}}) \right\}. \tag{3.27}$$

A recent range of MCA applications in astrophysics can be found in Melchior et al. 2018; Joseph et al. 2019 and Wagner-Carena et al. 2020

The Gaussian component

The Gaussian component is then modelled following the standard Wiener modelling where $\boldsymbol{\kappa}_{\mathrm{G}}$ is assumed to be a Gaussian random field:

$$\boldsymbol{C}_G(\boldsymbol{\kappa}_{\mathrm{G}}) = \|\boldsymbol{\kappa}\|^2_{\boldsymbol{\Sigma}_k}, \tag{3.28}$$

and the solution of Equation 3.26 is obtained using the iterative Wiener filtering method described in Jeffrey et al. 2018.

The non-Gaussian component

There are different ways to use a sparse model in the MCA framework. One way consists in using the standard ℓ_1 or ℓ_0-norm regularisation in a wavelet-based sparsity model, as it is done in the GLIMPSE algorithm (Lanusse et al. 2016). This results in:

$$\boldsymbol{C}_{NG}(\boldsymbol{\kappa}_{\mathrm{NG}}) = \lambda \|\boldsymbol{\Phi}^{\star}\boldsymbol{\kappa}_{\mathrm{NG}}\|_p, \tag{3.29}$$

where $p = 0$ or 1, Φ is the wavelet matrix, and λ is a Lagrangian parameter. However, it seems that this approach leads to large wavelet scales and low Fourier frequencies that are relatively close, making it difficult to separate the information.

For this reason, Starck et al. 2021 implement another approach, that involves first estimating the set of scales and positions where wavelet coefficients are above a given threshold, typically between 3 and 5 times the noise standard deviation relative to each wavelet coefficient. This set is composed by some coefficients named Ω, called active coefficients. These coefficients can be interpreted as a mask in the wavelet domain, where $\Omega_{j,x} = 1$ if a wavelet coefficient detected at scale j and position x and 0 otherwise. The noise $\sigma_{j,x}$ at scale j and position x can be determined using noise realizations.

The noise is whitened, since the noise factor considered is Gaussian with standard deviation equal to one and with a uniform power spectrum. Once this wavelet mask Ω is estimated, the non Gaussian component $\boldsymbol{\kappa}_{\mathrm{NG}}$ can be estimated through

$$\min_{\boldsymbol{\kappa}_{\mathrm{NG}}}\{|\Omega \odot \boldsymbol{\Phi}^{\star}\left((\boldsymbol{\gamma} - A\boldsymbol{\kappa}_{\mathrm{G}}) - A\boldsymbol{\kappa}_{\mathrm{NG}})\right)|^2 + \boldsymbol{C}_{\mathrm{NG}}(\boldsymbol{\kappa}_{\mathrm{NG}})\}, \tag{3.30}$$

with $\boldsymbol{C}_{NG}(\boldsymbol{\kappa}_{\mathrm{NG}}) = i_{\mathbb{R}}(\boldsymbol{\kappa}_{NG})$. Once Ω is fixed, the algorithm is almost linear and only a positivity constraint remains. Such positivity constraint is applied on the non-Gaussian component $\boldsymbol{\kappa}_{\mathrm{NG}}$. Peaks in the map can be on top of voids, and therefore have negative pixel values. As peaks are captured by the non-Gaussian component, they are positive by construction, but the convergence map $\boldsymbol{\kappa} = \boldsymbol{\kappa}_{\mathrm{G}} + \boldsymbol{\kappa}_{\mathrm{NG}}$ can still be negative at peaks positions. The larger the non-Gaussianities, the more MCALens is expected to improve over linear methods such as the Wiener one. The experiments run using this method on simulated data show a significant improvement compared to the state of art.

Example

Here is shown an example performed in Starck et al. 2021 of application of MCAlens to reconstruct a convergence map of the 1.64 deg^2 HST/ACS COSMOS survey (Scoville et al. 2007). The bright galaxies shape catalogue produced for Schrabback et al. 2010 is employed.

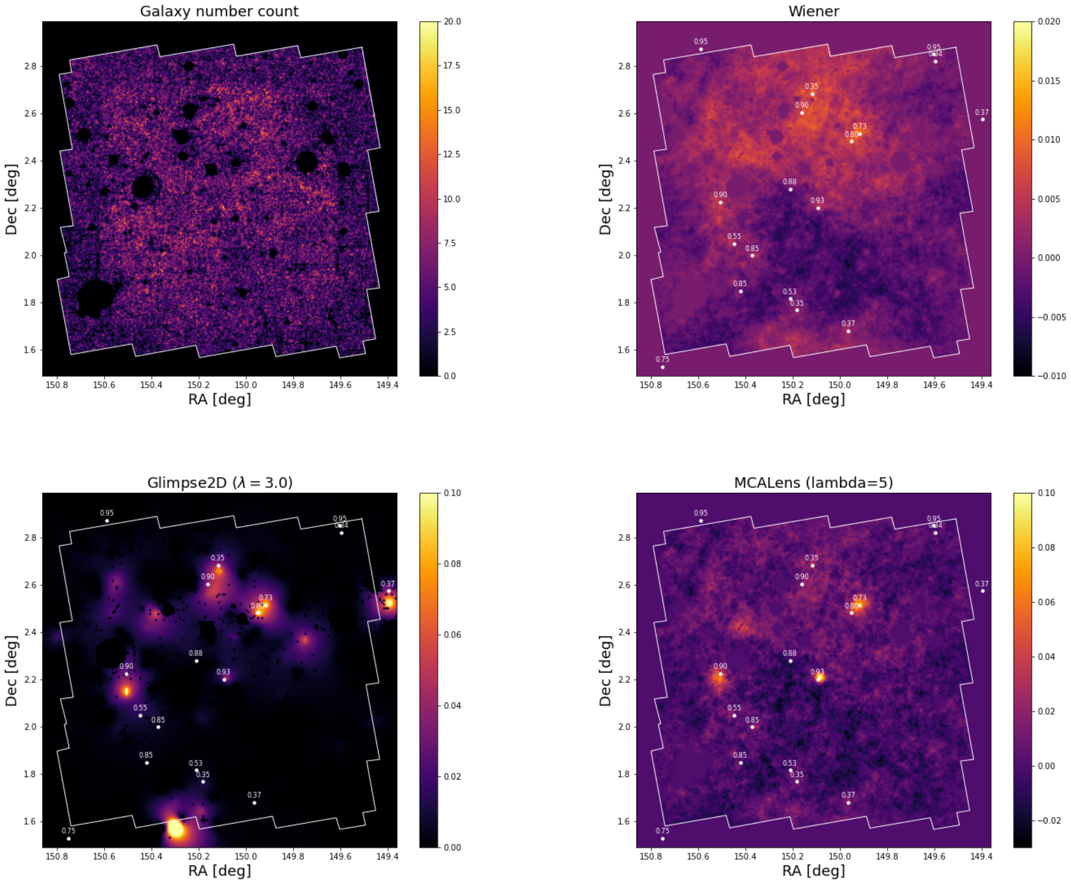


Figure 3.6: COSMOS data: Top, galaxies count map (left) and Wiener map (right) smoothed with a Gaussian having a Full Width at Half Maximum of 2.4 arcmin. Bottom, GLIMPSE(left) and MCAlens(right). Images and caption from Starck et al. 2021.

The results of Starck et al. 2021 after applying MCAlens on COSMOS data are presented in Figure 3.6. The top row shows the galaxy count map (left) and the Wiener map (right) smoothed with a Gaussian at a Full Width at Half Maximum of 2.4 arcmin. The bottom row shows respectively the wavelet solution obtained with GLIMPSE (left) and the MCAlens map (right). The white dots show the locations and redshifts of X-ray selected massive galaxy clusters from the XMM-Newton Wide Field Survey (Finoguenov et al. 2007) with $0.3 < z < 1.0$. From Figure 3.6, one can see how the Wiener filter returns the Gaussian components, GLIMPSE well reconstructs the overdense non-Gaussian parts in correspondence of the massive galaxy clusters and how MCAlens returns both Gaussian and non-Gaussian features.

3.7 Summary

To summarise, wavelets are functions that allow to decompose data into different frequency components, and then study each component with a resolution matched to its scale. The reason for the success of wavelets is due to the fact that wavelet bases represent well a large class of signals, especially astronomical data where many objects present isotropic features. The starlet transform is very well designed for representing these type of data. Sparse wavelet regularization is very efficient for inverse problems in astrophysics such as denoising or deconvolution. Very nice results have been obtained based on weak lensing mass map recovery with the algorithm MCALens (Starck et al. 2021) and Chapter 4.1.8 will show the potential of starlet-based summary statistics as a tool to extract information from cosmological data and estimate cosmological parameters.

CHAPTER 4

RESULTS WITH MULTI-SCALE PEAK COUNTS FOR WEAK LENSING

The goal of this chapter is to illustrate the methodology and results of our study presented in the paper Ajani et al. 2020. We investigated the impact of multi-scale filtering approaches on cosmological parameters by employing a starlet (wavelet) filter and a concatenation of Gaussian filters to constrain the sum of neutrino masses M_ν, the matter density parameter Ω_m and the amplitude of the primordial power spectrum A_s for four tomographic redshifts in a *Euclid*-like setting. We have compared different summary statistics: power spectrum, single-scale peak counts and multi-scale peak counts and assessed their performance in an ideal setting (no systematics).

4.1 Forecasts on neutrino masses using multi-scale peak counts

4.1.1 Overview

As described in Section 2.3, second order statistics such as the two-point correlation function or the power spectrum fail to capture the non-Gaussianities of the large-scale structure field and it is necessary to consider statistics of order higher than the second to distinguish the cosmological information hidden in non-linear modes that can represent new signals beyond the standard ΛCDM model. Moreover, as illustrated in Section 3.1, multi-scale approaches based on wavelets can be used for several applications in astrophysics. In the context of weak lensing peak counts, Liu J. et al. 2015 have shown that a combination of certain smoothing scales can tighten the errors on the cosmological constraints, using the publicly available data from the 154 deg^2 CFHTLenS survey. However, they claim that the combination that they find is only a limited investigation of the benefit of using multiple smoothing scales and leave the open question of identifying optimal filter shapes, sizes, and combinations that can help further tighten constraints using peak counts. An investigation in this direction has then been performed by Lin, Kilbinger, and Pires 2016, where they compared linear filtering with a Gaussian and two compensated filters (the starlet wavelet and the aperture mass), and the non-linear filtering method MRLens when inferring the matter density parameter Ω_{m}, the power spectrum normalisation σ_8, and the dark-energy parameter w_0. While they find that a compensated filter function with counts included separately from different smoothing scales yields the tightest constraints on cosmological parameters from weak lensing peaks, their non-tomographic configuration does not allow to impose constraints on w_0. The goal of the study presented in this chapter is to compare and quantify the performance of different filtering techniques in a multi-scale framework and to compare different summary statistics in this context: as second order statistics we employ the convergence power spectrum, as higher order statistics we employ peak counts of lensing convergence maps. We use as synthetic data the lensing convergence maps from `MassiveNus` simulations (Petri 2016; Liu et al. 2018). Using the same suite of simulations, Liu and Madhavacheril 2019, Marques et al. 2019, and Coulton et al. 2019 have already shown for a LSST-like survey (LSST Science Collaboration et al. 2009) that combining the lensing power spectrum with higher-order statistics can provide tighter constraints on parameters. For this purpose, we perform our analysis using the lensing power spectrum and peak counts as summary statistics following Li, Liu, Zorrilla Matilla, et al. 2019. We extend the study by considering a survey with *Euclid*-like noise, and to smooth the noisy convergence maps we employ a multi-scale approach, investigating a concatenation of Gaussian filters and separately a starlet filter (Starck, Murtagh, and Fadili 2010).

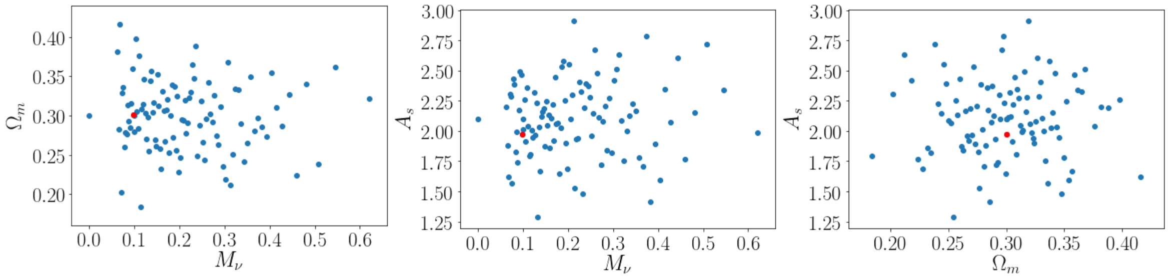


Figure 4.1: Parameter space of the 100 cosmological models available from the `MassiveNus` simulations. The red dot shows the fiducial model corresponding to { M_ν, Ω_m, $10^9 A_\mathrm{s}$ } = { 0.1, 0.3, 2.1 }.

4.1.2 Simulations

We use the Cosmological Massive Neutrino Simulations (`MassiveNus`), a suite of publicly available N-body simulations released by the Columbia Lensing group [1]. It contains 101 different cosmological models obtained by varying the sum of neutrino masses M$_\nu$, the total matter density parameter Ω_m and the primordial power spectrum amplitude A_s at the pivot scale $k_0 = 0.05$ Mpc^{-1}, in the range $M_\nu = [0, 0.62]$ eV, Ω_m =[0.18, 0.42] and $A_\mathrm{s} \cdot 10^9$=$[1.29, 2.91]$. The reduced Hubble constant $h = 0.7$, the spectral index $n_s = 0.97$, the baryon density parameter $\Omega_b = 0.046$ and the dark energy equation of state parameter $w = -1$ are kept fixed under the assumption of a flat universe. The fiducial model is set at $\{M_\nu, \Omega_\mathrm{m}, 10^9 A_\mathrm{s}\}$=$\{0.1, 0.3, 2.1\}$. The presence of massive neutrinos is taken into account assuming normal hierarchy and using a linear response method, where the evolution of neutrinos is described by linear perturbation theory but the clustering occurs in a non-linear cold dark matter potential. The simulations have a 512 Mpc/h box size with 1024^3 CDM particles. They are implemented using a modified version of the public tree-Particle Mesh (tree-PM) code `Gadget2` with a neutrino patch, describing the impact of massive neutrinos on the growth of structures up to $k = 10\ h$ Mpc^{-1}. A complete description of the implementation and the products is presented in Liu et al. 2018. We use the simulated convergence maps as mock data for our analysis. When dealing with real data, the actual observable is the shear field that can be converted into the convergence field as discussed in Section 2.1.3. We bypass this step from γ to κ and work with the convergence maps directly provided as products from `MassiveNus`. The maps are generated using the public ray-tracing package `LensTools` (Petri 2016) for each of the 101 cosmological models at five source redshifts $z_s = \{0.5, 1.0, 1.5, 2.0, 2.5\}$. Each redshift has 10000 different map realisations obtained by rotating and shifting the spatial planes. Each κ map has 512^2 pixels, corresponding to a 12.25 deg^2 total angular size area in the range $\ell \in [100, 37000]$ with a resolution of 0.4 arcmin. Figure 4.1 shows the 2D parameter space planes for the 100 cosmological models available from the simulations.

---

[1] The simulations are publicly available at `http://columbialensing.org`.

4.1.3 Noise and survey specifications

The method described in this work can be applied to any given survey. For illustration purposes, we perform here a tomographic study using the source redshifts $z_{\rm s} = \{\ 0.5, 1.0, 1.5, 2.0\ \}$ and mimicking the noise expected for a survey like *Euclid*[2](Laureijs et al. 2011; Euclid Collaboration et al. 2020b). Specifically, at each source redshift we produce 10000 map realisations of Gaussian noise with mean zero and variance

$$\sigma_{\rm n}^2 = \frac{\langle \sigma_\epsilon^2 \rangle}{n_{\rm gal} A_{\rm pix}}\ , \tag{4.1}$$

where we set the dispersion of the ellipticity distribution to $\sigma_\epsilon = 0.3$, and the pixel area for these simulations is by $A_{\rm pix} \simeq 0.16$ arcmin^2. The redshift dependence that makes a tomographic investigation possible is encoded in the source galaxy redshift distribution, for which we assume the parametric form

$$n(z) = \mathcal{C} \left(\frac{z}{z_0} \right)^\alpha \exp \left[- \left(\frac{z}{z_0} \right)^\beta \right], \tag{4.2}$$

with $\alpha = 2$, $\beta = 3/2$ $z_0 = 0.9/\sqrt{2}$ as in (Laureijs et al. 2011; Euclid Collaboration et al. 2020b), and $\mathcal{C}$ is the normalization constant to guarantee the constraint $\int_{z_{\rm min}}^{z_{\rm max}} n(z)\, {\rm d}z = 30$ gal/arcmin^{-2}. Then, we compute the galaxy number density at each bin as

$$n_{\rm gal}^i = \mathcal{C} \int_{z_i^-}^{z_i^+} n(z) {\rm d}z, \tag{4.3}$$

where z_i^-, z_i^+ are the edges of the i^{th} bin. We adapt the binning choice to the provided simulation settings, assuming that we observe galaxies within a small range around the actual source redshift. This leads to the values for the galaxy number density $n_{\rm gal}$ per source redshift bin $z_{\rm s}$ provided in Table 4.1:

$\mathbf{z_s}$	0.5	1.0	1.5	2.0
$\mathbf{n}_{\rm gal}$	11.02	11.90	5.45	1.45

Table 4.1: Values of $n_{\rm gal}$ for each source redshift z_s. We adapt the binning choice to the provided simulation settings, assuming that we observe galaxies within a small range around the actual redshift. In practice, this means considering as bin edges $\{0.001, 0.75, 1.25, 1.75, 2.25\}$, in order to compute the integral in Equation 4.3.

---

[2] `https://www.Euclid-ec.org`

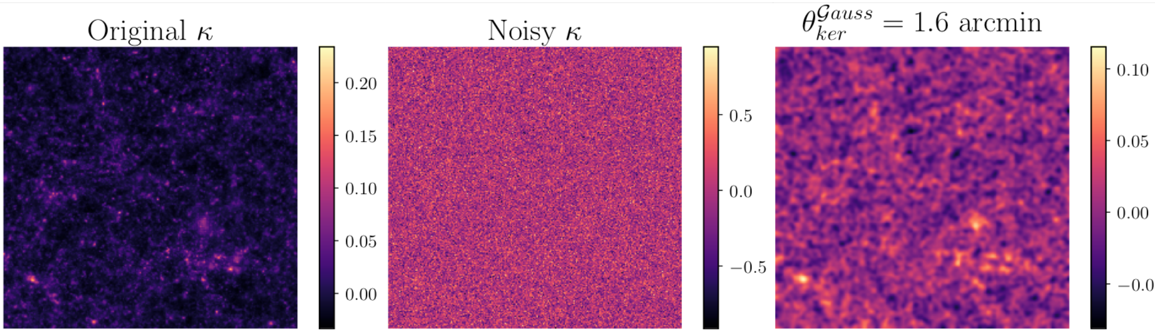


Figure 4.2: Convergence maps κ are noiseless. We apply Gaussian noise and then filter the map using either the Gaussian or starlet filtering. For illustration purposes, we show the Gaussian filtering with $\theta^{\mathcal{G}}_{\text{ker}} = 1.6$ arcmin of one map realisation for the fiducial model $\{M_\nu, \Omega_{\text{m}}, 10^9 A_{\text{s}}\}$=$\{0.1, 0.3, 2.1\}$. The colour bar on the right of each map describes values of the convergence field κ. For each realisation of the 10000 maps provided for each redshift we generate 10000 noise maps as described in Section 4.1.3 corresponding to the different value of n_{gal} respectively for $z_s = [0.5, 1.0, 1.5, 2.0]$.

4.1.4 Filtering strategy

In this section, I will describe the strategy used for filtering. Some definitions introduced in Section 3.1 will be re-established in order to keep reading smooth. In order to access the signal in the convergence maps at small scales, where they are mostly dominated by noise, we filter them, considering a multi-scale analysis compared to a single-scale analysis. First, we use a single Gaussian kernel of size θ_{ker}, defined as

$$\mathcal{G}(\theta; \theta_{\text{ker}}) = \frac{1}{\sqrt{2\pi}\theta_{\text{ker}}} e^{-\theta^2/(2\theta^2_{\text{ker}})} \,, \tag{4.4}$$

which was also used for a single-scale filter in Li, Liu, Zorrilla Matilla, et al. 2019. We then compare the results with those obtained when applying instead a concatenation of Gaussian filters and an *Isotropic Undecimated Wavelet Transform*, also known as a starlet transform (Starck, Fadili, and Murtagh 2007), introduced in Section 3.4. This allows us to decompose our original image (the noisy convergence map) I as a sum of wavelet coefficient images w_j plus a coarse resolution image c_J:

$$I(x, y) = c_J(x, y) + \sum_{j=1}^{j_{\max}} w_j(x, y), \tag{4.5}$$

where the wavelet images w_j represent the details of the original image at dyadic (powers of two) scales corresponding to a spatial size of 2^j pixels and $J = j_{\max} + 1$.

One of the advantages of employing a starlet filter is provided by its multi-scale analysis, namely its ability to investigate and extract the information encoded at different scales at the same time (Starck, Murtagh, and Bijaoui 1998). Hence, the starlet transform presents the properties to compute efficiently J scales with a fast algorithm with a complexity of $\mathcal{O}(N^2 \log N)$ for an image of $N \times N$ pixels and to analyse data with compensated aperture filters with finite support. Further details on the advantages of wavelet starlet analysis have been discussed in Chapter 3 and are also present in Peel et al. 2017; Peel, Pettorino, Giocoli, et al. 2018. The following example illustrates how we can compare results from these two different filtering schemes. Applying a starlet transform with $j_{max} = 4$ to a map with 0.4 arcmin pixel size results in a decomposition of four maps with resolutions [0.8, 1.6, 3.2, 6.4] arcmin plus the coarse-scale map. For our study, we will consider as finest scale $\theta_{ker}^{St} = 1.6$ arcmin, being a more realistic choice in terms of resolution for convergence maps coming from *Euclid*-like survey data. We will therefore focus on the set of scales [1.6, 3.2, 6.4] arcmin plus the coarse map. Concerning the multi-Gaussian filters, to fairly compare them to the starlet, we set the standard deviations of the Gaussian filters such that their maximum matches that of the corresponding single starlet scale profile as shown in Figure 4.3, resulting in a concatenation of Gaussians respectively with $\theta_{ker}^{\mathcal{G}} = [1.2, 2.7, 5.5, 9.5]$ arcmin. Based on the above, in our study we compare cosmological constraints obtained using noisy maps smoothed from a single-Gaussian kernel with the ones obtained from a multi-Gaussian analysis and from a starlet decomposition. We exclude the observables corresponding to 0.8 arcmin in our analysis after having verified that this does not cost any loss of information. The starlet transform can be seen as multi-Gaussian filtering where each Gaussian kernel is replaced by a compensated filter. In Figure 4.2 we show the result of the filtering procedure for a Gaussian kernel: given the original convergence map κ, we add white noise as described in Section 4.1.3 and then we filter the noisy map with the chosen kernel.

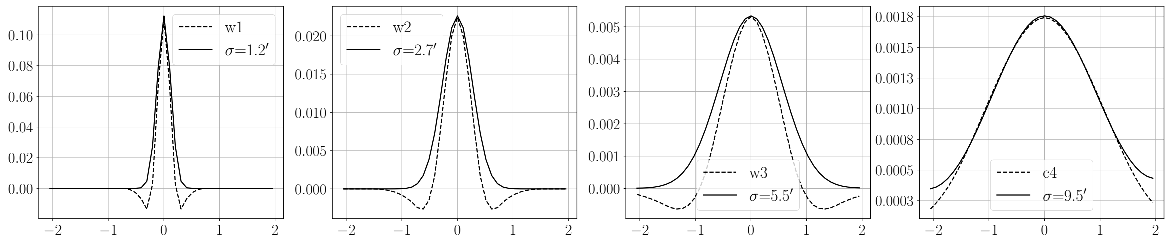


Figure 4.3: We show the matching between the Gaussian filter and the starlet at the different scales. We chose the kernel for the multi-Gaussian concatenation such that the maximum of the two profiles match. From left to right are the finest scale to the smoothest scale where $[\, w_1, w_2, w_3, c_4 \,] = [\, 1.6, 3.2, 6.2, 12.8 \,]$ arcmin.

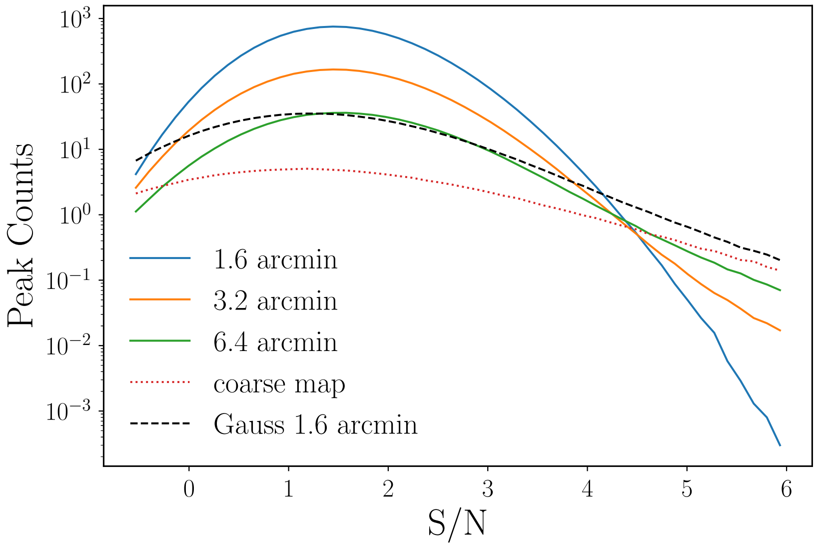


Figure 4.4: Peak Counts distribution in logarithmic scale for each starlet scales resolutions: $[\,1.6, 3.2, 6.4\,]$ arcmin and the coarse maps (red dotted line) and the Gaussian case (black dashed line). Due to the decomposition at different scales, for each map filtered with the starlet there are 4 different distributions. Indeed, the number of counts depends on the resolution: the larger the smoothing size (the lower the frequency) the smaller the number of peaks.

4.1.5 Summary statistics: power spectrum and multi-scale peaks

To extract and investigate the cosmological information encoded in the weak lensing convergence maps, we compute the power spectrum (PS) and peak counts as summary statistics.

Convergence power spectrum

In this study, we compute as second order summary statistics the power spectrum as defined in Section 2.1 of the noisy filtered convergence maps. For a given cosmology we add Gaussian noise to each realisation of κ. For each redshift we generate a different set of noise maps following Equation 4.1. To filter the maps we employ a Gaussian kernel with smoothing size $\theta_{\rm ker}^{\mathcal{G}} = 1$ arcmin and consider angular scales with logarithmically spaced bins in the range $\ell = [300, 5000]$. We compute the power spectra using the python package `LensTools` for each of the 10000 realisations per cosmology and then we take the average over the realisations. We parallelise our code using `joblib`[3] to accelerate processing due to the large number of realisations per cosmology.

---

[3] `https://joblib.readthedocs.io/`

Multi-scale peaks

As higher order statistics, we detect and count weak lensing peaks on the noisy filtered maps using our code [4] as follows. We compute peaks as local maxima of the signal-to-noise field $\nu \equiv S/N$ i.e. as a pixel of larger value than its eight neighbours in the image. We define the signal to noise field ν as the ratio between the noisy convergence κ convolved with the filter $\mathcal{W}(\theta_{\text{ker}})$ over the smoothed standard deviation of the noise for each realisation per redshift:

$$\nu = \frac{(\mathcal{W} * \kappa)(\theta_{\text{ker}})}{\sigma_n^{\text{filt}}}, \tag{4.6}$$

where $\mathcal{W}(\theta_{\text{ker}})$ can be the single-Gaussian, the multi-Gaussian or the starlet filter. Concerning σ_n^{filt}, its definition depends on the employed filter. For a Gaussian kernel it is given by the standard deviation of the smoothed noise maps, while for the starlet case we need to estimate the noise at each wavelet scale for each image per redshift. To estimate the noise level at each starlet scale we follow Starck and Murtagh 1998. As described in Section 3.5, we use the fact that the standard deviation of the noise at the scale j is given by $\sigma_j = \sigma_j^e \sigma_I$, where σ_I is the standard deviation of the noise of the image and σ_j^e are the coefficients obtained by taking the standard deviation of the starlet transform of a Gaussian distribution with standard deviation one at each scale j. To estimate σ_I we take the *median absolute deviation*[5] of the noisy convergence map. We do this for each one of the 10000 realisations for each cosmology and then take the average over the realisations. We consider the peaks distribution for 41 linearly spaced bins within the range $\nu = [-0.6, 6]$, based on the outcomes of Li, Liu, Zorrilla Matilla, et al. 2019 where it is shown that including the low ($S/N < 1$), medium ($1 < S/N < 3$) and high peaks ($S/N > 3$) jointly give the best constraints. Moreover, low and medium peaks, typically formed due to multiple much smaller halos than the single halos that cause the high peaks (Liu and Haiman 2016), contain a similar level of information as the high peaks. In Figure 4.4 we show for illustration purposes the peak counts distribution in logarithmic scale for each starlet scale and for a single-Gaussian filter case. We see that the number of counts depends on the resolution: the larger the smoothing size (the lower the frequency) the smaller the number of peaks. We have investigated the impact of the binning on the results by testing different boundaries, and we have found that choosing 41 bins instead of 50 decreases the condition number of the data matrix, hence facilitating its inversion during the likelihood analysis. We have also considered the minimum and the maximum values of the S/N maps as bin edges, and we have seen that this choice is not very convenient, since it increases the condition number by two orders of magnitude.

---

[4] Code publicly available at `https://github.com/CosmoStat/lenspack.`

[5] For a Gaussian distribution the *median absolute deviation* (MAD) and the standard deviation are directly related as: $\text{MAD}/\sigma = 0.6745$. We choose to use this estimator since it is more robust when dealing with non-normal distributions (being more resistant to outliers in a data set) to have a more general implementation in our pipeline.

4.1.6 Covariance matrices

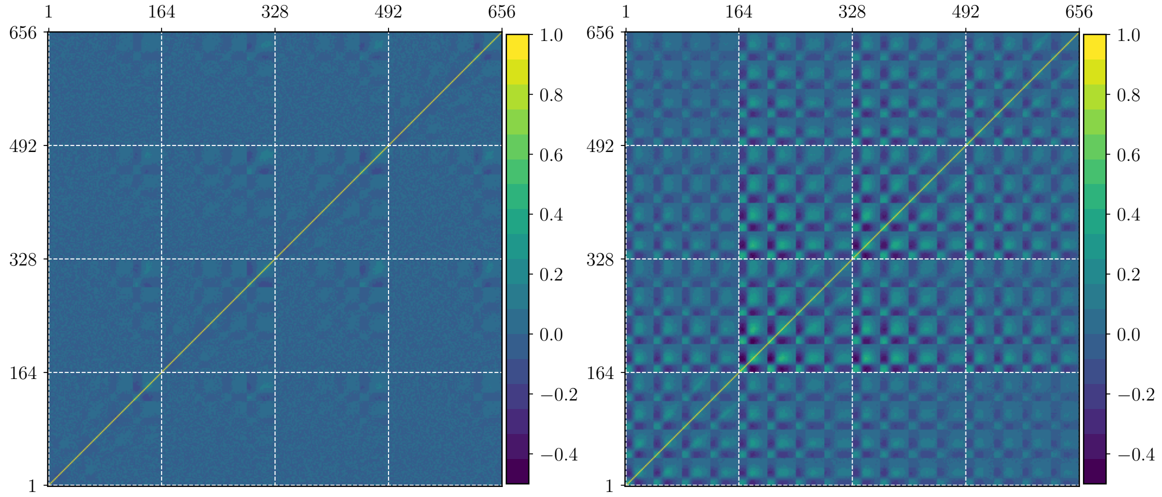


Figure 4.5: We show the correlation coefficients respectively for the starlet (left) and the multi-Gaussian (right) peak counts. The white dashed lines split the contribution of the different redshifts: the bin range [1,164] refers to the correlations at $z_s = 0.5$, the bin range [165, 328] refers to the correlations at $z_s = 1.0$, the bin range [328, 492] refers to the correlations at $z_s = 1.5$ and the bin range [493, 656] refers to the correlations at $z_s = 2.0$. Each redshift contribution is split in the four different scales (the four mini-blocks inside the white boxes) in increasing order, i.e. [1.6, 3.2, 6.4, coarse] arcmin for the starlet and [1.2, 2.7, 5.5, 9.5] arcmin for the multi-Gaussian, where each scale is binned in 41 values of S/N in the range $[-0.6, 6]$. We notice how the starlet decomposition has a tendency to make the observable correlation matrix more diagonal while the off-diagonal terms of the multi-Gaussian matrix show more correlations along the scales and among S/N values in the mini-blocks.

To compute the covariance matrices of the data, we use the independent fiducial massless neutrino simulation, defined by $[M_\nu, \Omega_\mathrm{m}, 10^9 A_\mathrm{s}] = [0.0, 0.3, 2.1]$ and obtained from initial conditions different from the massive simulations used to perform inference. We consider a parameter-independent covariance to reduce the risk of assigning an excess of information to the observables in the context of a Gaussian likelihood assumption, following the results of Carron, J. 2013. The covariance matrix elements are computed as

$$C_{ij} = \sum_{r=1}^{N} \frac{(x_i^r - \mu_i)(x_j^r - \mu_j)}{N-1}, \tag{4.7}$$

where N is the number of observations (in this case the 10000 realizations), x_i^r is the value of the power spectrum or the peak counts in the i^{th} bin for a given realisation r and

$$\mu_i = \frac{1}{N}\sum_r x_i^r, \tag{4.8}$$

is the mean of the power spectrum or the peak counts in a given bin over all the realisations.

In Figure 4.5 we show the correlation coefficients of the multi-scale tomographic peak counts: the starlet case (in the left panel) and the multi-Gaussian case (in the right panel). The matrices are organised as follows. From left to right, the four main blocks are for the tomographic redshifts respectively in the order $z_s = [0.5, 1.0, 1.5, 2.0]$. Within each of the main blocks, there are four sub-blocks representing the filter scale, i.e. the scales $[1.6', 3.2', 6.4', \mathrm{coarse}]$ for the starlet and $[1.2', 2.7', 5.5', 9.5']$ for the multi-Gaussian. Each scale is binned in 41 values of signal to noise in the range $S/N = [-0.6, 6]$. We see that the starlet decomposition has a tendency to make the matrix more diagonal, while the off-diagonal terms for a multi-Gaussian show more correlations between the scales and for small and high values of S/N. Consistently with this, we notice that the most correlated bins in the starlet case are the ones corresponding to the coarse scale (the last mini-block for each of the main blocks) whose profile indeed closely mimics a Gaussian, as one can see in the last panel of Figure 4.3. Furthermore, we take into account the loss of information due to the finite number of bins and realisations by adopting for the inverse of the covariance matrix the estimator introduced by Hartlap, Simon, and Schneider 2007:

$$C^{-1} = \frac{N - n_{\mathrm{bins}} - 2}{N-1} C_*^{-1}, \tag{4.9}$$

where N is the number of realisations, n_{bins} the number of bins, and C_* the covariance matrix computed for the power spectrum and peak counts, whose elements are given by Equation 4.7. We also scale the covariance for a *Euclid* sky coverage by the factor $f_{\mathrm{map}}/f_{\mathrm{survey}}$, where $f_{\mathrm{map}} = 12.25\ \mathrm{deg}^2$ is the size of the convergence maps and $f_{Euclid} = 15000\ \mathrm{deg}^2$.

In using Equation 4.9, we do not expect all biases to be removed from our parameter inference, as this has already been ruled out in Sellentin and Heavens 2016 and in Sellentin and Heavens 2017. Nevertheless, we rely on the fact that the number of realisations that we are using (10000) is sufficiently large and greater than n_{bins} to consider it a reliable estimator for our purposes[6]. By looking at the correlation matrices of Figure 4.5 it is clear that the starlet (left panel) has the tendency to make the matrix more diagonal while the multi-Gaussian case (right panel) presents non-trivial off-diagonal terms. This finding, namely the fact that the majority of the information seems to be encoded in the diagonal terms of the covariance

---

[6] Indeed, the value of the correction coefficient is close to 1 for each analysis we perform. However, considering that the results of (Sellentin and Heavens 2017) quantify the loss of information also in the case of a *Euclid*-like survey, it would be worth it to reproduce our study applying their restoration technique to generalize our analysis.

matrix in the starlet case, will be further investigated in the next section where I illustrate the results we obtained in this study.

4.1.7 Likelihood

To perform Bayesian inference and get the probability distributions of the cosmological parameters, we use a Gaussian likelihood for a cosmology-independent covariance:

$$\log \mathcal{L}(\theta) = \frac{1}{2}(d - \mu(\theta))^T C^{-1}(d - \mu(\theta)), \tag{4.10}$$

where d is the data array, C is the covariance matrix of the observable, μ the expected theoretical prediction as a function of the cosmological parameters θ. In our case, the data array is the mean over the (simulated) realisations of the power spectrum or peak counts or combination of the two for our fiducial model. Cosmological parameters are the ones for which simulations are available, namely $\{M_\nu, \Omega_\mathrm{m}, A_\mathrm{s}\}$. In order to determine the relation between the observable and the models $\mu(\theta)$, i.e. to be able to have a prediction of the power spectrum and the peak counts given a new set of points in parameter space, we employ an interpolation with Gaussian Processes Regression (GPR, (Rasmussen and Williams 2005)) using the `scikit-learn` python package. As discussed in Section 1.5.2, Gaussian Processes are a generic supervised learning method that, via an assumption of smoothness between parameters with close values, allows one to compute the prediction for an observable at a new given point in parameter space. The cosmological parameters and the corresponding observables (power spectrum and peak counts or the two statistics combined) from the simulations are used as a training set, i.e. as the input for the GPR. Then, the Gaussian Processes act by assuming that for a new point in parameter space θ_* which is sufficiently close to a known point θ belonging to the training set, the corresponding observable will be described by a joint normal distribution along with the known observable. For this interpolation, we employ an *anisotropic squared exponential* function. We estimate the standard error of the noise level in the targets as the noise given by the fact that we take the mean over 10000 realisations for each observable and each bin. More specifically, for each bin we compute the observable mean and its corresponding standard error over the 10000 realisations available from the simulations. The fitting procedure takes then as input the three cosmological parameters and the re-scaled mean of the observable for each bin, while the standard errors are added as dual coefficients of the training data points in kernel space, i.e. as a regularisation term to the diagonal of the kernel matrix to take into account the noise level on the mean. This results in a number of GP corresponding to the number of bins that are then taken in by a prediction function. The latter reads new points in parameter space and returns the corresponding observable predictive distribution (power spectrum, peaks or the two statistics jointly) and its standard deviation. The hyper-parameters are fit with the standard marginal likelihood approach. For the validation we compare the prediction of a given model obtained with the GP emulator excluding the corresponding model from the simulation, for 10 models near the fiducial model following Liu and Madhavacheril 2019.

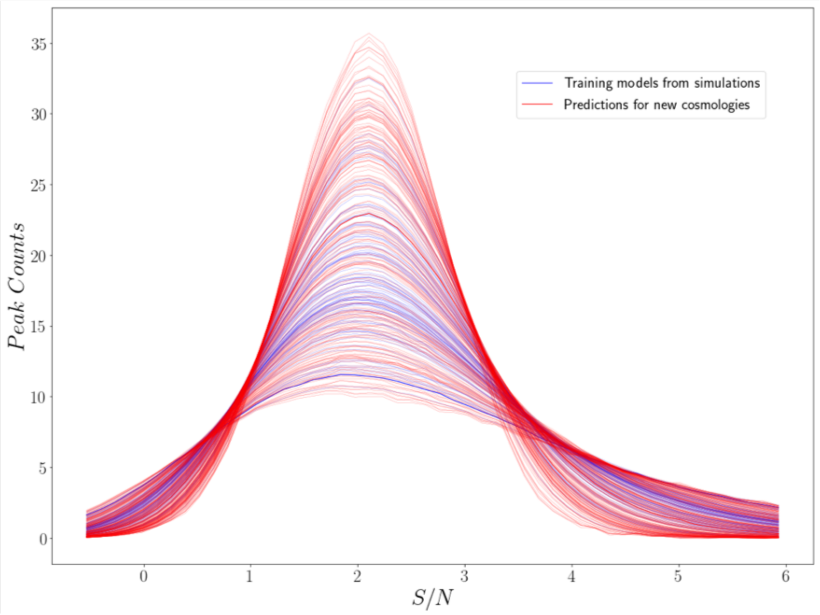


Figure 4.6: Visualisation of Gaussian processes predictor input and output: the training set, constituted by the cosmological parameters and corresponding peaks distributions obtained from the simulations are represented in blue. The output for the prediction is given by the red curves: for 200 new given cosmologies in parameter space fed to the Gaussian processes prediction function we get the corresponding peaks distribution.

We find differences at the sub-percent level that always lie within the statistical error consistent with Li, Liu, Zorrilla Matilla, et al. 2019; Coulton et al. 2019; Marques et al. 2019. Intuitively, this can be understood by looking at Figure 4.6: the training set, constituted by the cosmological parameters and corresponding peaks distributions obtained from the simulations are represented in blue. The output for the prediction is given by the red curves: for 200 new given cosmologies in parameter space fed to the Gaussian processes prediction function we get the corresponding peaks distribution. During the likelihood analysis, parameter space is explored in this way: by computing the summary statistics corresponding to each new set of cosmological parameters explored by the MCMC walkers.

4.1.8 Results

In this section I illustrate forecast results on the sum of neutrino masses M_ν, on the matter density parameter Ω_m and on the primordial power spectrum amplitude A_s for a survey with *Euclid*-like noise in a tomographic setting with four source redshifts $z_\mathrm{s} = [0.5, 1.0, 1.5, 2.0$

], and compare results for different observables (power spectrum and peak counts) and filters (single-Gaussian, starlet and multi-Gaussian) to answer the questions lying at the origin of the study performed in Ajani et al. 2020:

A) What is the performance of multi-scale peak counts in terms of constraining power with respect to the state of the art summary statistics in weak lensing?

B) Does the combination of second order statistics and multi-scale higher-order statistics bring an improvement in tightening the constraints?

C) What is the impact of the different scales on the parameter constraints?

D) What is the impact of the filter choice on the covariance matrix?

E) What is the impact of the different tomographic redshift bins on the parameter constraints?

To answer these questions and compare the different contours corresponding to the different summary statistics and filtering settings we perform a Bayesian likelihood analysis and use the `emcee` package[7] to explore the parameter space with MCMC. We assume a flat prior, specifically following Coulton et al. 2019, a Gaussian likelihood function as defined in Equation 4.10, and a model-independent covariance matrix as discussed in Section 4.1.6. The walkers are initialised in a tiny Gaussian ball of radius 10^{-3} around the fiducial cosmology $[M_\nu, \Omega_\mathrm{m}, 10^9 A_\mathrm{s}] = [0.1, 0.3, 2.1]$ and we estimate the posterior using 120 walkers. Our chains are stable against the length of the chain, and we verify their convergence by employing *Gelman Rubin* diagnostics (Gelman and Rubin 1992). To plot the contours we use the `ChainConsumer` python package (Hinton 2016).

As our results consist in comparing the constraining power for different summary statistics and filtering settings, in order to quantify our results, we use estimators common in the literature, whose definitions are recalled here.

Figure of Merit

To have an approximate quantification of the size of the parameter contours that we use to compare their constraining power, we consider the following Figure of Merit (FoM) as defined in Euclid Collaboration et al. 2020b:

$$\mathrm{FoM} = \left(\det(\tilde{F})\right)^{1/n}, \tag{4.11}$$

---

[7] `emcee` is a python implementation of the affine-invariant ensemble sampler for Markov chain Monte Carlo (MCMC) introduced by (Foreman-Mackey et al. 2013) and described in Section 1.5.3.

FoM	$(M_\nu, \Omega_\mathrm{m})$	(M_ν, A_s)	$(\Omega_\mathrm{m}, A_\mathrm{s})$	$(M_\nu, \Omega_\mathrm{m}, A_\mathrm{s})$
Power spectrum	1585	77	1079	2063
Single-Gaussian Peaks	3559	200	2861	6537
Single-Gaussian Peaks + PS	5839	322	4688	11205
Starlet Peaks	7818	434	6428	12755
Starlet Peaks diagonal	8434	471	6936	16528
Starlet Peaks + PS	9796	540	7966	16166
Multi-Gaussian Peaks	11804	655	9729	18647
Multi-Gaussian Peaks diagonal	13612	770	11231	29780
Multi-Gaussian Peaks + PS	13471	742	10983	22115

Table 4.2: Values of the FoM as defined in Equation 4.11 for the different parameters pairs (α, β) for each observable employed in the likelihood analysis: the power spectrum, the peaks counted on maps smoothed with the kernel in consideration and *Peaks + PS* always refer to the constraints obtained with the peaks relative to some filters and the power spectrum while the term *diagonal* refers to the contours obtained with a likelihood analysis performed by only considering the diagonal elements of the data covariance matrix. We provide in the last column the 3D FoM given as the inverse of the volume in $(M_\nu, A_\mathrm{s}, \Omega_\mathrm{m})$ space.

where $\tilde{F}$ is the marginalised Fisher submatrix that we estimate as the inverse of the covariance matrix among the set of cosmological parameters $\{M_\nu, \Omega_\mathrm{m}, 10^9 A_\mathrm{s}\}$ obtained with the MCMC chains. In the exponent, n is equal to the parameter space dimensionality, e.g. $n = 2$ for FoM in a 2-dimensional plane between two parameters while $n = 3$ if we take the Fisher matrix among the three parameters. We show the values of the FoM four our observables in Table 4.2.

Figure of correlation

To quantify the correlations among the parameters we use the *Figure of Correlation* (Casas et al. 2017; Euclid Collaboration et al. 2020b):

$$\mathrm{FoC} = \sqrt{\det(\mathbf{P^{-1}})}, \tag{4.12}$$

where $\mathbf{P}$ is the correlation matrix, whose elements are defined as $P_{\alpha\beta} = C_{\alpha\beta}/\sqrt{C_{\alpha\alpha}C_{\beta\beta}}$, with $C_{\alpha\beta}$ being the covariance between the cosmological parameters α and β as defined in the previous section. When the parameters are fully uncorrelated FoC $= 1$, while for FoC > 1 the off-diagonal terms are non-zero, indicating an increasing presence of correlations among parameters as the FoC increases. The values of the FoC for our constraints are shown in Table 4.3.

FoC	$(M_\nu, \Omega_\mathrm{m})$	(M_ν, A_s)	$(\Omega_\mathrm{m}, A_\mathrm{s})$
Power spectrum	1.00	1.98	1.10
Single-Gaussian Peaks	1.21	1.42	1.01
Single-Gaussian Peaks + PS	1.19	1.31	1.04
Starlet Peaks	1.14	1.18	1.14
Starlet Peaks + PS	1.13	1.13	1.20
Multi-Gaussian Peaks	1.13	1.16	1.17
Multi-Gaussian Peaks + PS	1.13	1.14	1.18

Table 4.3: Value of the Figure of Correlation for each pair of cosmological parameters corresponding to the different tomographic observables: the power spectrum alone (PS), the Peaks alone for different filters and the two statistics combined (Peaks + PS). As explained in the text, FoC = 1 corresponds to uncorrelated parameters, while the further the FoC is to 1, the more correlations are present. Qualitatively, this can be appreciated by looking at the inclination of the contours: by looking at Figure 4.7 we can see more ‘oblique’ contours for the Gaussian peaks in the plane $(M_\nu, \Omega_\mathrm{m})$ compared to the power spectrum, while for the pair (M_ν, A_s) the power spectrum shows more correlation than peaks.

Condition Number	**Single-Gaussian Peaks**	**Starlet Peaks**	**Multi-Gaussian Peaks**
41 bins	10^5	10^6	10^7
50 bins	10^5	10^{16}	-

Table 4.4: Values of the Condition Number for the data covariance matrices. The smaller the number the easier it is to invert the matrix. In this case we get very large values for this estimator, leading to the conclusion that the data covariance matrices for Starlet Peaks and Gaussian Peaks show very singular behaviours.

Matrix condition number

To estimate how difficult it is to invert our data covariance matrices, we compute the corresponding *condition number*: if the matrix is singular, the associated condition number is infinite, i.e. matrices with large condition numbers are more difficult to invert. We compute the condition number through the 2-norm of the matrix using singular value decomposition (SVD). As shown in Table 4.4, the condition number depends on the binning choice, especially for the multi-scale analysis. Indeed, we find that choosing 41 linearly spaced bins for the peak counts instead of 50 reduces the condition number of the starlet peaks of about 10 orders of magnitude. For this reason we choose 41 bins when performing inference using peak counts.

$\sigma_{\alpha\alpha}$	$\mathbf{M}_\nu$	$\mathbf{\Omega}_\mathrm{m}$	$\mathbf{A}_\mathrm{s}$
Power spectrum	0.127	0.005	0.204
Single-Gaussian Peaks	0.083	0.004	0.086
Single-Gaussian Peaks + PS	0.061	0.003	0.066
Starlet Peaks	0.047	0.003	0.057
Starlet Peaks diagonal	0.049	0.003	0.053
Starlet Peaks + PS	0.040	0.003	0.052
Multi-Gaussian Peaks	0.038	0.003	0.047
Multi-Gaussian Peaks diagonal	0.042	0.002	0.039
Multi-Gaussian Peaks + PS	0.035	0.002	0.044

Table 4.5: Values of 1-σ marginalised error for each cosmological parameter for the different observables.

Marginalised error

To estimate the marginalised 1σ error on a single parameter θ_α (which means having included all the degeneracies with respect to other parameters), we use the quantity:

$$\sigma_{\alpha\alpha} = \sqrt{C_{\alpha\alpha}}, \tag{4.13}$$

where $C_{\alpha\alpha}$ are the diagonal elements of the parameter covariance matrix. We show the values of the $\sigma_{\alpha\alpha}$ for our observables in Table 4.5.

A) Multi-scale peak counts performance vs state of the art summary statistics

In the left panel of Figure 4.7 we compare constraints obtained from the single-scale and the multi-scale peak counts analysis against the power spectrum contours. For the single scale we employ a Gaussian filter with $\theta_{ker}^{\mathcal{G}} = 1.6'$. For the multi-scale analysis we employ a starlet filter and a concatenation of Gaussian filters with smoothing widths chosen such that the profiles match the starlet scales, as described in Section 4.1.4. More specifically, we show the comparison among the power spectrum (blue contours), the single-scale peaks (green contours), the starlet peaks (red contours) and the multi-Gaussian peaks (black contours). We confirm that peak counts outperform power spectrum constraints even in the single-Gaussian case, as found in Li, Liu, Zorrilla Matilla, et al. 2019. In addition, we find that a multi-scale approach leads to a remarkable improvement with respect to a single-scale approach in terms of constraining power, as expected, due to its higher information content concerning structure formation. We quantify these outcomes by considering the Figure of Merit defined in Equation 4.11. As shown in Table 4.2, the FoM for the single-Gaussian peaks in the parameter space plane (M_ν, Ω_m) is more than twice the one given by the power spectrum; the one from starlet peaks is more than twice that obtained with the single-Gaussian peaks, and the multi-Gaussian peaks FoM is more than three times the one corresponding to the single-Gaussian case. Concerning the (M_ν, A_s) and (Ω_m, A_s) planes, the FoM for Gaussian peaks is about three times that for the power spectrum, the one from starlet peaks is again about twice that obtained with the single-Gaussian peaks, and the multi-Gaussian peaks give again a FoM about three times that seen for single-Gaussian peaks contours.

As further investigation, we compute the Figure of Correlation as defined in Equation 4.12 to study the correlation among the parameters. By looking at Table 4.3, one can see how values for the power spectrum for the pairs (M_ν, Ω_m) and (Ω_m, A_s) are close to one, suggesting that correlation among them appears to be very small, while the plane (M_ν, A_s) shows more correlation, as its FoC is nearly twice as large. Qualitatively this can be appreciated by looking at the inclination of the contours. More specifically, concerning the plane (M_ν, Ω_m), the power spectrum contours are horizontal and show also visually that these two parameters are not correlated; constraints obtained via peak counts show a slightly larger correlation, increasing by 21% for Gaussian peaks and by about 13-14% for the multi-scale analysis, with respect to the power spectrum. It is interesting to note that in the (M_ν, A_s) plane, the correlation decreases by 30% when using single-Gaussian peaks and by about 40% when using a multi-scale approach compared to the power spectrum, suggesting that peak counts can play an important role in breaking the degeneracy for this pair of parameters. Independently of the correlation, all constraints obtained with multi-scale filtering are tighter than the ones obtained via single-scale filtering, and both are tighter than the ones for the power spectrum.

B) Performance of second order statistics and multi-scale peaks combined

Based on the previous result, we are now interested in the constraints obtained when considering the two statistics jointly and on the impact of the different filter settings in this context. In particular, by focusing on the right panel of Figure 4.7, where we compare the 95% confidence contours of the power spectrum (blue contours) with the single-Gaussian peaks (green contours) and the two joint statistics (violet contours), we notice how in the single-scale approach the addition of the power spectrum to the peak counts brings a non-negligible improvement in terms of constraining power with respect to the peaks alone. More specifically, reading the values presented in Table 4.2, we see that the FoM of the joint contours is roughly 1.6 times that of the peaks alone, and more than three times that of the power spectrum.

Focusing now on the multi-scale approaches, in Figure 4.8 we show the same comparison in the left panel by comparing the power spectrum with the starlet peaks alone (red contour) and the two joint statistics (orange contours) and in the right panel for the multi-Gaussian case with the multi-Gaussian peaks alone (black contours) and the two joint statis-

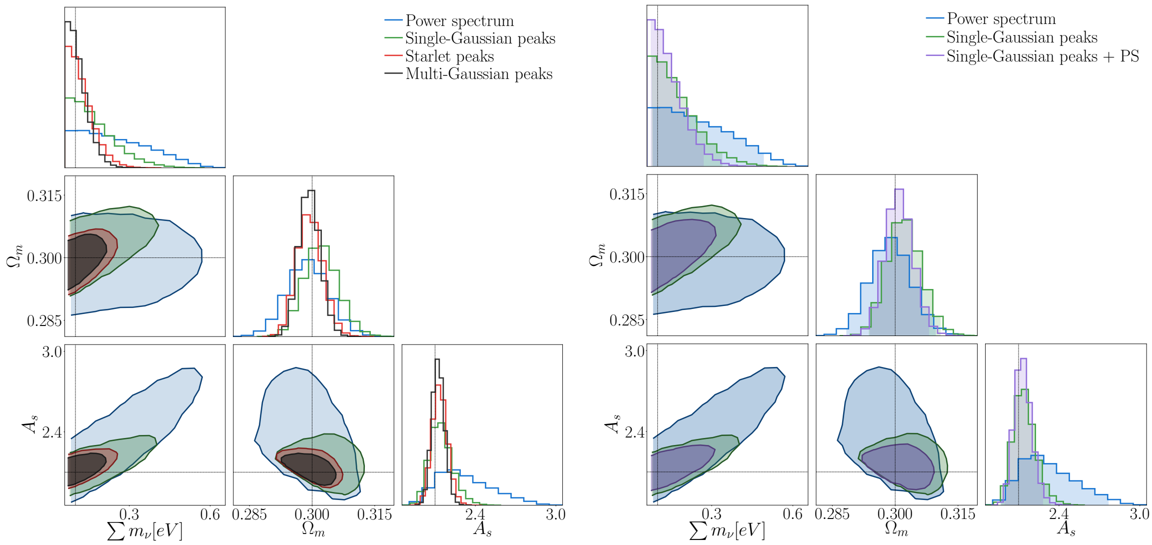


Figure 4.7: 95 % confidence contours tomography with source redshifts $z_{\rm s} = [\ 0.5,1.0,1.5,\ 2.0\ ]$ and corresponding galaxy number density: $n_{\rm gal}$ =[11.02, 11.90, 5.45, 1.45]. The black dotted line is the fiducial model: $[\sum m_\nu,\ \Omega_{\rm m},\ 10^9 A_{\rm s}] = [0.1,\ 0.3,\ 2.1]$. **Left panel**: constraints from power spectrum (blue contours) computed on noisy maps smoothed with a Gaussian kernel $\theta_{\rm ker} = 1$ arcmin, constraints from Gaussian Peak counts (green contours) computed on noisy maps smoothed with a Gaussian kernel $\theta_{\rm ker} = 1.6$ arcmin, constraints from Starlet Peak counts (red contours) computed on noisy maps smoothed with a Starlet kernel with corresponding resolutions [1.6, 3.2, 6.4] arcmin + coarse map, constraints from multi-Gaussian Peak counts computed on noisy maps smoothed with a Starlet kernel with corresponding resolutions [1.2, 2.7, 5.5, 9.5] arcmin. **Right panel**: constraints from power spectrum (blue contours) computed on noisy maps smoothed with a Gaussian kernel $\theta_{\rm ker} = 1$ arcmin, constraints from Gaussian Peak counts (green contours) computed on noisy maps smoothed with a Gaussian kernel $\theta_{\rm ker} = 1.6$ arcmin and the two statistics jointly (violet contours).

tics (turquoise contours). We notice in this case that the FoM of the combined statistics are roughly $1.1 - 1.2$ times the peaks alone in the starlet case and 1.1 times the peaks alone in the multi-Gaussian case, suggesting that the information given by the joint statistics is mostly contained in the multi-scale peak counts alone. Peak counts therefore appear to be competitive and sufficient statistics for parameter inference when dealing with weak lensing convergence maps as input data, in our ideal setting. This further confirms that lensing peaks are a powerful tool in the context of cosmological parameter inference, emphasising as well the importance of the role played by the filtering choice.

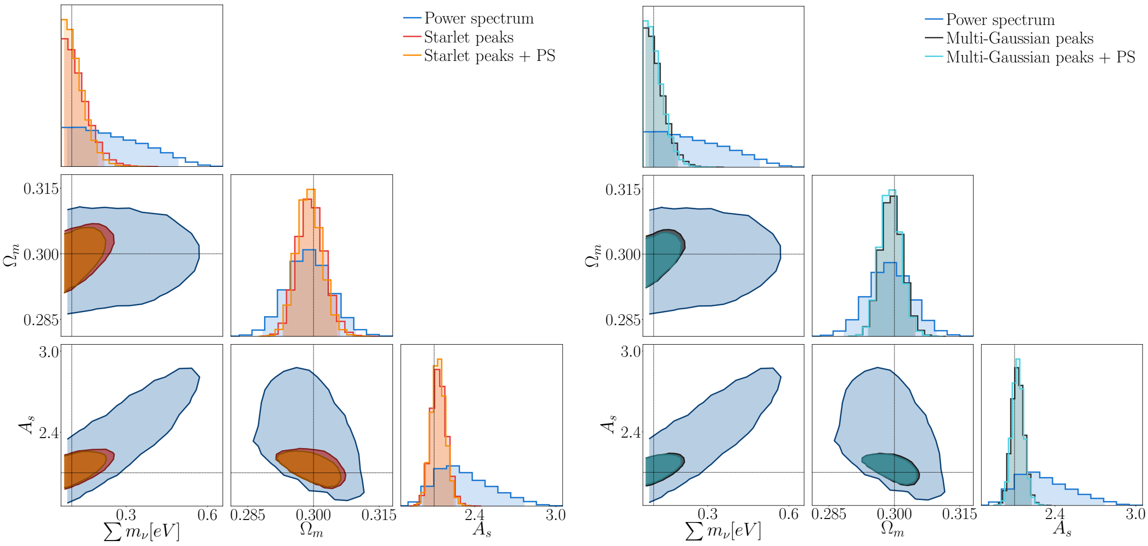


Figure 4.8: 95 % confidence contours tomography with source redshifts z_s = [0.5,1.0,1.5, 2.0] and corresponding galaxy number density: n_gal = [11.02, 11.90, 5.45, 1.45]. The black dotted line is the fiducial model: [$\sum m_\nu$, Ω_m, $10^9 A_\mathrm{s}$] = [0.1, 0.3, 2.1]. **Left panel**: constraints from power spectrum (blue contours) computed on noisy maps smoothed with a Gaussian kernel θ_ker = 1 arcmin, constraints from Starlet Peak counts (red contours) computed on noisy maps smoothed with a Starlet kernel with corresponding resolutions [1.6, 3.2, 6.4] arcmin + coarse map and constraints from the two statistics joint (orange contours) . **Right panel**: constraints from power spectrum (blue contours) computed on noisy maps smoothed with a Gaussian kernel θ_ker = 1 arcmin, constraints from multi-Gaussian Peak counts (black contours) computed on noisy maps smoothed with a multi-Gaussian kernel with corresponding resolutions [1.2, 2.7, 5.5, 9.5] arcmin and the two statistics jointly (light blue contours).

C) Starlet scales impact on constraints

The left panel of Figure 4.9 shows the impact of the different starlet decomposition scales on the constraints. The MCMC chain used for the starlet decomposition results of the analysis has been obtained by considering all starlet scales, i.e. $[1.6, 3.2, 6.4]$ arcmin + coarse scale, shown in red. To check that we are allowed to exclude the finest scale in the entire analysis, namely not to include the resolution corresponding to 0.8 arcmin - which won't satisfy the survey requirements - we compare the constraints relative to $[0.8, 1.6, 3.2, 6.4]$ arcmin + coarse scale with the ones for $[1.6, 3.2, 6.4]$ arcmin + coarse scale and we verify that they overlap. We then investigate the impact of the different starlet scales and we obtain that it is sufficient to consider the setting $[3.2, 6.4]$ + coarse scale to obtain results competitive with the full set of scales, as shown by the dark blue contours in the figure that match the full starlet decomposition contours. Hence, we identify $w_2 = 3.2$ arcmin as the smallest scale needed to obtain the maximal constraints with convergence maps of resolution 0.4 arcmin. We also perform the inference by adding one scale at a time in the observable array to show how the contours shrink as a function of the number of starlet scales. We find that the only setting that recovers almost the full information is given by [3.2, 6.4, coarse]. We also notice from the contours relative to $[w_1, w_2, w_3]$ that excluding the coarse scale leads to a loss of information (precisely 28% on M_ν, 33% on Ω_m and 19% on A_s).

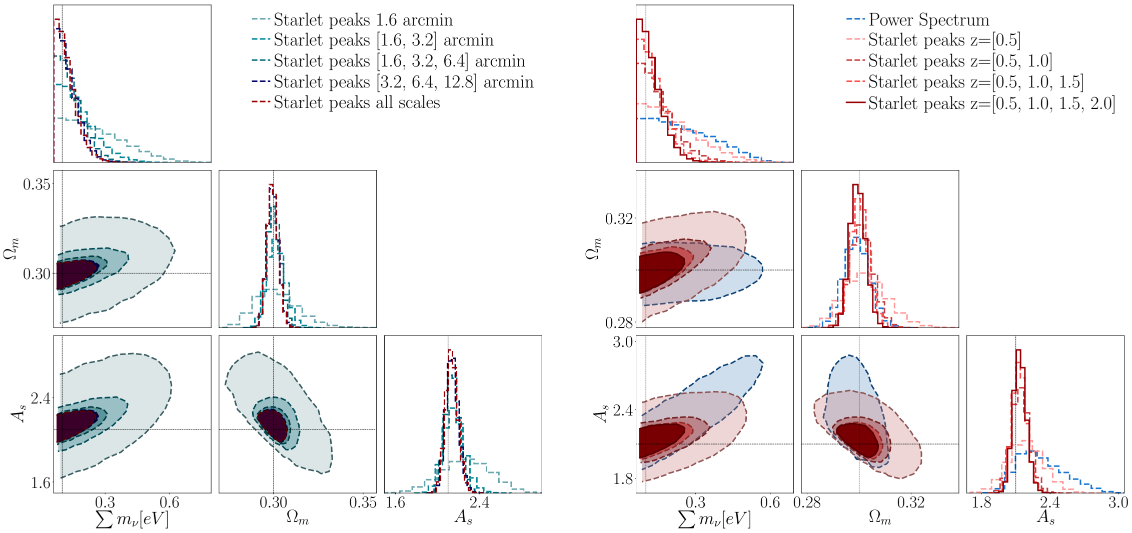


Figure 4.9: 95 % confidence contours using tomography with redshifts $z_\mathrm{s} = [\ 0.5, 1.0, 1.5, 2.0\ ]$ and corresponding galaxy number densities $n_\mathrm{gal} = [\ 11.02,\ 11.90,\ 5.45,\ 1.45\ ]$. The black dotted line is the fiducial model: $[\ \sum m_\nu,\ \Omega_\mathrm{m},\ 10^9 A_\mathrm{s}\ ] = [\ 0.1,\ 0.3,\ 2.1\ ]$. **Left panel**: we show the impact of the different starlet scales and we prove there exists a minimum resolution $\theta_\mathrm{ker} = w_2 = 3.2$ arcmin that allows us to obtain constraints comparable to what is achieved with the full wavelet decomposition and that the information contained in the coarse map cannot be neglected. Starting from the first starlet scale alone $\theta_\mathrm{ker} = w_1 = 1.6$ arcmin (dashed big contours in light blue) we add scale by scale in light blue until [w1, w2, w3]. In dark blue we show the constraints corresponding to $[w_2, w_3, c_4\ ]$ that almost match with the constraints provided by the full starlet decomposition (in red). **Right panel**: we show here the constraints obtained by adding each tomographic redshift at the time: the dashed pink corresponds to the contours relative to $z_\mathrm{s} = 0.5$, the lighter red to $z_s = [\ 0.5, 1.0\ ]$, the darker pink to $z_\mathrm{s} = [\ 0.5,\ 1.0,\ 1.5\ ]$ and the dark red to the full set of redshifts in the starlet case. We plot this against the power spectrum (blue contours) to show how each source redshift contribution to shrinks the contours and helps break the degeneracy with respect to the power spectrum.

D) Impact on covariance matrix: Gaussian multi-scale and Starlet comparison

In this section we investigate the impact of the choice of the filter on the data covariance matrix. Indeed, by looking at the correlation matrices of Figure 4.5 it is clear that the starlet (left panel) has the tendency to *diagonalise* the matrix while the multi-Gaussian case (right panel) presents non-trivial off-diagonal terms as introduced in Section 4.1.6. To further explore this aspect we have run the likelihood analysis considering just the diagonal elements of the covariance matrices in order to compare results with full covariance case. We find the constraints illustrated in Figure 4.10: in the left panel we plot the starlet peak counts contours obtained with the full covariance matrix (red) against the diagonal-only version of the data covariance matrix (dashed dark red). On the right panel we show the same comparison for the multi-Gaussian peaks case with the full covariance case (black) against diagonal-only contours (dashed gray). We see that for the starlet filter the majority of the information is indeed encoded in the diagonal elements, while for the multi-Gaussian case the presence of non-trivial correlations among the scales makes the contours slightly larger for $\Omega_{\rm m}$ and $A_{\rm s}$, while it adds some information on M_ν with respect to the diagonal case. We can quantify this by taking the ratio between the FoM relative to the full covariance and the diagonal elements cases: for the starlet we find a ratio of 1.07 and 1.15 for the multi-Gaussian. The same interpretation arises when looking at the 1-σ marginalised error of Table 4.5: excluding the off-diagonal terms in the starlet data covariance matrix implies a loss of information of 4% on M_ν, no loss on $\Omega_{\rm m}$ and a gain of 7% on $A_{\rm s}$. For the multi-Gaussian case, the same procedure leads to a loss of 11% on M_ν, a gain of 33% on $\Omega_{\rm m}$ and a gain of 17% on $A_{\rm s}$. This is an interesting aspect of the starlet filter that could prove to be useful when dealing with high dimensional data and the covariance matrix can be difficult to invert.

E) Impact of different tomographic redshifts

Here we investigate how parameter constraint contours shrink by adding the tomographic redshift bins by one by one. In particular, in the right panel of Figure 4.9 we show the information gain resulting from the addition of each tomographic redshift bin. In blue we show the power spectrum, and in darkening shades of red we plot contours for the starlet peaks as follows. Contours for source redshift $z_{\rm s} = 0.5$ are dashed pink, they are dashed darker pink for source redshift $z_{\rm s} = [0.5, 1.0]$, and so on until reaching the dark red contours which are obtained by concatenating all source redshifts. As expected, the peaks contours show a different degeneracy direction from that of the power spectrum due to the higher-order information they contain. Indeed, the contours at $z_{\rm s} = 0.5$ already show different degeneracy compared to the power spectrum contours between M_ν and $\Omega_{\rm m}$ with a FoC=1.09, giving though larger marginalised constraints on $\Omega_{\rm m}$. Adding $z_{\rm s} = 1.0$ provides FoM that are more three times those for $z_{\rm s} = 0.5$ alone for the planes $(M_\nu, \Omega_{\rm m})$ and $(M_\nu, A_{\rm s})$ but increases the correlation between M_ν and $\Omega_{\rm m}$ to FoC=1.17. Finally, the concatenation of the

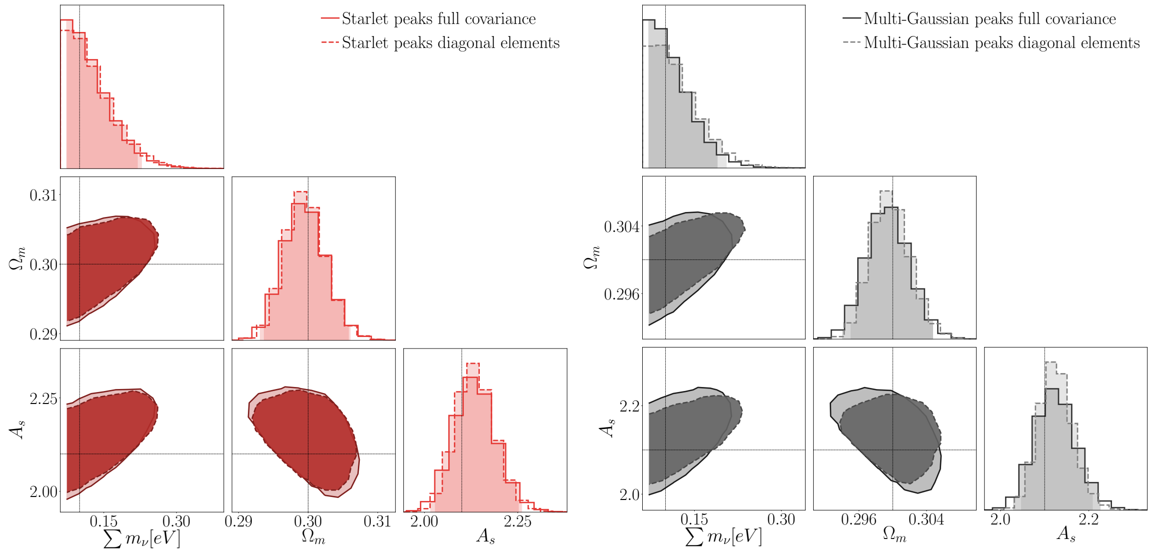


Figure 4.10: 95 % confidence contours tomography with redshifts $z_s = [0.5, 1.0, 1.5, 2.0]$ and corresponding galaxy number density: $n_{\mathrm{gal}} = [11.02, 11.90, 5.45, 1.45]$. The black dotted line is the fiducial model: [$\sum m_\nu$, Ω_{m}, $10^9 A_{\mathrm{s}}$] = [0.1, 0.3, 2.1]. **Left panel**: constraints from starlet peak counts (continuous red contours) obtained employing the full covariance matrix against constraints from starlet peak counts (dashed red contours) obtained employing the diagonal elements only of the covariance matrix in the likelihood analysis. **Right panel**: constraints from multi-Gaussian peak counts (continuous black contours) obtained employing the full covariance matrix against constraints from multi-Gaussian peak counts (dashed gray contours) obtained employing the diagonal elements only of the covariance matrix in the likelihood analysis

source redshifts $z_{\mathrm{s}} = [0.5, 1.0, 1.5]$ further adds correlations between the two parameters (FoC=1.19) but shrinks the contours to almost reach those of the complete set of source redshifts.

Marginalised errors

In Figure 4.11 we show the marginalised constraints on each cosmological parameter corresponding to the different observables. To compare the improvement obtained by employing the different statistics we compute the 1σ marginalised error for each parameter, summarised in Table 4.5. In particular, we find an improvement of 35%, 20% and 58% respectively on M_ν, Ω_{m} and A_{s} when employing the single-Gaussian peaks instead of the power spectrum, an improvement of 63%, 40% and 72% when employing the starlet peaks instead of the power spectrum alone, and an improvement of 70%, 40% and 77% when employing the multi-Gaussian peaks instead of the power spectrum alone. Namely, the starlet peaks outperform the single-Gaussian peaks by 43% on M_ν, 25% on $\Omega_{\rm m}$ and 34% on $A_{\rm s}$, and the multi-Gaussian peaks outperform the single-Gaussian peaks by 54% on M_ν, 25% on $\Omega_{\rm m}$ and 45% on $A_{\rm s}$. Finally, employing a multi-Gaussian instead of a starlet filter in the context of peak counts might improve the constraints by 19% on M_ν, and 18% on $A_{\rm s}$, while no improvement is noticed for $\Omega_{\rm m}$.

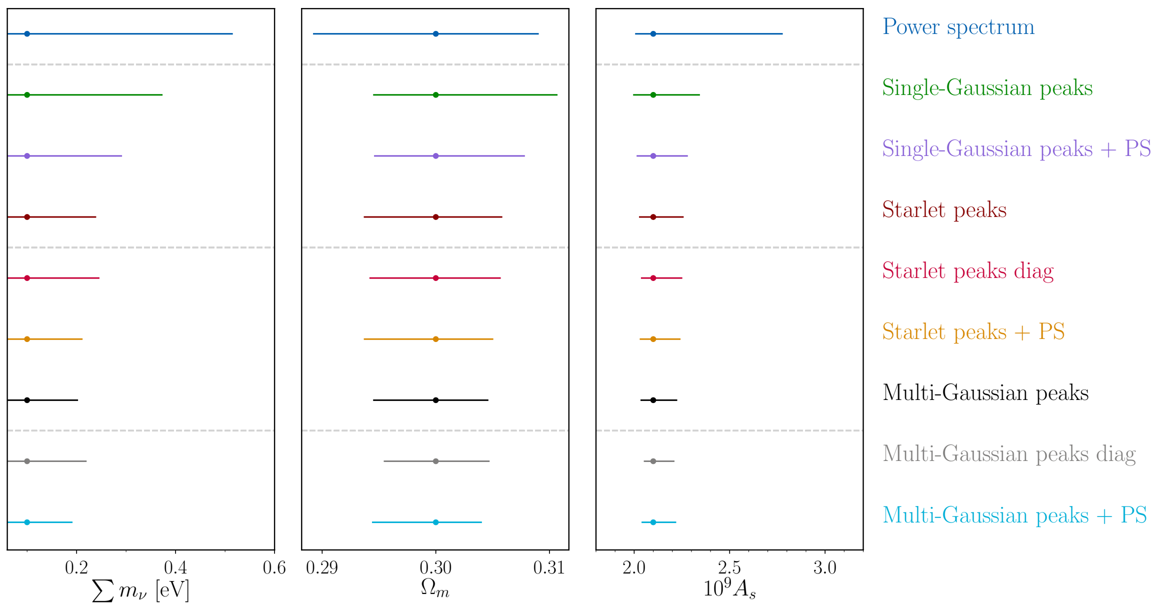


Figure 4.11: Marginalized constraints on each parameter for forecasts showing the 2.5 and 97.5 percentiles with respect to the fiducial model. These marginalised constraints refer to a tomographic setting with $z = [0.5, 1.0, 1.5, 2.0]$ with the fiducial model set at $[M_\nu, \Omega_{\rm m}, 10^9 A_{\rm s}] = [0.1, 0.3, 2.1]$ corresponding to the different observables employed within the likelihood analysis. The values are listed in Table 4.6.

4.1.9 Conclusions and future prospects

In the presented study, we infer the sum of neutrino masses M_ν, the matter density parameter $\Omega_{\rm m}$ and the amplitude of the primordial power spectrum $A_{\rm s}$ for a survey with *Euclid*-like noise using tomographic weak lensing. Our goal is to compare the constraining power of multi-scale filtering approaches, namely the starlet filter and a concatenation of Gaussian filters, with respect to a single-Gaussian one in the context of peak counts. We also compute the constraints with standard second-order statistics, in particular using the lensing power spectrum as a benchmark for the comparison. We compare the outcomes obtained from filtering the lensing convergence maps, which have a resolution of 0.4 arcmin, with a

Observable	$\mathbf{M_\nu+}$	$\mathbf{\Omega_m-}$	$\mathbf{\Omega_m+}$	$\mathbf{A_s-}$	$\mathbf{A_s+}$
Power spectrum	0.514	0.289	0.308	2.008	2.774
Single-Gaussian Peaks	0.372	0.295	0.311	1.998	2.340
Single-Gaussian Peaks + PS	0.290	0.294	0.307	2.016	2.277
Starlet Peaks	0.238	0.293	0.305	2.029	2.255
Starlet Peaks diagonal	0.244	0.294	0.305	2.039	2.248
Starlet Peaks + PS	0.210	0.293	0.305	2.032	2.239
Multi-Gaussian Peaks	0.201	0.294	0.305	2.037	2.222
Multi-Gaussian Peaks diagonal	0.218	0.295	0.305	2.054	2.207
Multi-Gaussian Peaks + PS	0.190	0.294	0.304	2.042	2.216

Table 4.6: Values of the 2.5 and 97.5 percentiles for each cosmological parameter as illustrated in Figure 4.11. In this table we also show the values corresponding to the marginalised constraints obtained using only the diagonal elements of the covariance matrices. They are very similar to the ones obtained by employing the full covariance. We further investigate this aspect in Section 4.1.8.

Gaussian kernel of smoothing size 1.6 arcmin, a starlet kernel, and a concatenation of Gaussians. More specifically, the starlet filter is an isotropic undecimated wavelet transform that allows us to extract the information encoded in different spatial scales simultaneously. Setting the number of scales in the transform to four, the starlet kernel sizes for our maps correspond to [0.8, 1.6, 3.2, 6.4] arcmin + the coarse map, since the starlet transform returns maps filtered at dyadic scales. In deriving parameter constraints we exclude the first scale and work with [1.6, 3.2, 6.4] arcmin + coarse map. To fairly compare it with a multi-Gaussian, we set the standard deviations of the Gaussian kernels at each scale such that their profile peaks match the corresponding starlet scale peaks, resulting in a concatenation of Gaussians with standard deviations of [1.2, 2.7, 5.5, 9.5] arcmin respectively.

We find the following results:

a) For peak counts, a multi-scale filtering approach of the noisy maps leads to an improvement factor of more than two over a single-scale approach (single-Gaussian kernel) for the joint constraints on $(M_\nu, \Omega_\mathrm{m})$, (M_ν, A_s) and $(\Omega_\mathrm{m}, A_\mathrm{s})$ when using a starlet kernel, and a factor of more than three when using a multi-Gaussian filter. This is even more evident in the marginalised constraints, where the improvement is respectively 43% on M_ν, 25% on Ω_m and 34% on A_s for the starlet, while for the multi-Gaussian it is 54% on M_ν, 25% on Ω_m and 45% on A_s. Employing a multi-Gaussian instead of a starlet filter in the context of peak counts might improve the constraints by 19% on M_ν, and 18% on A_s, while no improvement is noticed for Ω_m. Finally, both multi-scale peak counts perform better than the power spectrum on the set of parameters $\{M_\nu, \Omega_\mathrm{m}, A_\mathrm{s}\}$ respectively by 63%, 40% and 72% when using a starlet filter and by 70%, 40% and 77% when using a multi-scale Gaussian filter.

b) When combining multi-scale peaks and the power spectrum, i.e. using a concatenation of peak counts and the power spectrum as the observed data vector, we find that the information is mostly encoded in the peaks alone (for certain parameters, such as Ω_{m} in the starlet case, it is *completely* encoded). This suggests that when adopting a multi-scale approach, it might be sufficient to work with the peaks alone.

c) The inclusion of the coarse map when counting peaks preserves crucial information. Moreover, for maps with a pixel size of 0.4 arcmin, there exists a minimum resolution (i.e. smallest scale needed) for the starlet scales corresponding to $\theta_{\mathrm{ker}} = 3.2$ arcmin to achieve maximal constraining power. This enables us to exclude the first two finest scales of the starlet decomposition, which correspond to the highest frequencies and are the most prone to the impact of noise, allowing for a faster and more efficient analysis.

d) We notice that employing a starlet filter leads to a highly diagonal data covariance matrix, while for the multi-Gaussian filter the off-diagonal terms are prominent, and correlations among the different scales are non-negligible. In other words, the majority of the information in the starlet filter case is encoded in the diagonal elements of the covariance matrix. This is an interesting aspect of the starlet filter that could prove useful when dealing with high dimensional data where the covariance matrix can be difficult to invert.

In summary, we confirm that weak-lensing peak counts are a powerful tool to infer cosmological parameters, especially when investigating the non-linear regime where the impact of parameters such as the neutrino masses becomes relevant. We also point out the importance of adopting a multi-scale approach in the context of weak-lensing peak counts, which bring the advantage of analysing the information encoded at different scales simultaneously, thereby leading to tighter constraints than single-scale analysis. As we have shown in Figure 4.11, the multi-scale filters considered (the multi-Gaussian filter and the starlet filter) lead to similar constraints. This is expected, as we choose the Gaussian kernels such that each profile peak matches with a starlet scale. Minimal residual differences between the two filters may be related to the binning: while this is the same for both, it might be that the two filters are optimal with different choices of the binning. We leave the investigation of the optimal binning for both multiscale Gaussian and starlet peaks to future work. There is however an advantage, in using the starlet filter over a multi-Gaussian filter: the starlet has the tendency to remove the off-diagonal terms in the covariance matrix, hence making the matrix more diagonal, easy and faster to invert. Moreover, Leonard, Pires, and Starck 2012 have proved that it offers a clear and significant time advantage over standard aperture mass algorithms for all scales of interest. We implemented a pipeline that allows us to go from simulated lensing convergence maps as input data to constraints on cosmological parameters as final output, employing different filtering techniques with second-order (the power spectrum) and higher-order statistics (peak-counts). Hence, a future project will be to generalise the pipeline in terms of flexibility of the input data, including systematics effects and modelling of the noise. In particular, being able to control systematics errors and baryonic effects is as important as the statistical power to guarantee a robust analysis. In the context of weak-lensing peak counts, baryons can change the shape of the distribution of peaks by increasing the low S/N end and decreasing the high S/N values by a few percent, as quantified by Fong et al. 2019. It has been shown as well by Coulton et al. 2020b that ignoring baryonic effects can lead to strong biases in inferences from peak counts and that in principle these biases can be mitigated without significantly degrading cosmological constraints when baryonic effects are modelled and marginalised. In Chapter 7 I will present a first application to real data where we include baryonic effects. Concerning intrinsic alignment and noise uncertainty, further investigations is needed in future studies with the aim of including such modelling in our pipeline and to ultimately apply our pipeline to real data coming from future galaxy surveys.

CHAPTER 5

A NEW STATISTICS THAT COMBINES MULTI-SCALE PEAKS AND VOIDS

In this chapter I illustrate the methodology and results we presented in Ajani et al. 2021. Motivated by the advantages of multi-scale approaches for cosmological parameter inference as discussed in Chapter 4 along with recent results of Coulton et al. 2020a in combining peaks an minima, we propose a new multi-scale higher order summary statistics that we called the *starlet* ℓ_1-norm that allows for a multi-scale analysis that naturally encodes the information brought by peaks and minima. At the end of the chapter, I also provide a description and visualisation of the different steps of the pipeline employed to obtain the results of this chapter and Chapter 4.

5.1 Starlet ℓ_1-norm for weak lensing cosmology

5.1.1 Overview

The previous section outlined the potential of peak statistics in a multi-scale framework for cosmological parameter inference. As previously mentioned, besides peak statistics, several alternative high order statistics have been proven to be very powerful in many studies to account for the non-Gaussian information encoded in weak lensing data. In particular, Coulton et al. 2020a have introduced weak lensing minima as a new probe of non-Gaussian information related to underdense regions (voids) in our universe and have shown how the combination of lensing peaks and minima leads to an improvement of respectively 44%, 11% and 63% on the sum of neutrino masses M_ν, the matter density parameter $\Omega_{\rm m}$ and on the amplitude of the primordial power spectrum $A_{\rm s}$ with respect to the power spectrum contours. In a study to compare the constraining power of high order statistics in the ($\Omega_{\rm m}$, σ_8) plane, Zürcher et al. 2021 found that peak counts, minimum counts and Minkowski functionals yield stronger constraints when compared to the angular power spectrum analysis in a non-tomographic, as well as a in tomographic setup. From the results of these studies, it becomes clearer and clearer that the information encoded in lensing minima is complementary to what we can extract through peak counts and that their combination, namely the full distribution, is more powerful than considering the two statistics alone in terms of constraining power. In this direction, Martinet et al. 2021 have performed forecasts with aperture mass ($\rm M_{ap}$) map statistics for future cosmic shear surveys, including peaks, voids, and the full distribution of pixels (1D $\rm M_{ap}$) and have shown that for the the structure growth parameter S_8, the dark energy equation of state parameter w_0 and the matter density parameter $\Omega_{\rm m}$, the full distribution of pixels with aperture mass has a superior constraining power than peaks and voids because it includes both peaks and voids. Motivated by the general context just described and by our previous findings on the potential of multi-scale wavelet based framework for high order statistics, we have then proposed in Ajani et al. 2021 a new summary statistics called *starlet* ℓ_1 *norm* that allows to perform very naturally a joint multi-scale peaks and voids analysis, accounting for the information present in all pixels in the weak lensing convergence map. The aim of this chapter is to present the methodology and results of the study Ajani et al. 2021.

5.1.2 Toward the starlet ℓ_1 norm

Starlet peaks

Multi-scale peaks counts can be derived either using a set of Gaussian kernels of different sizes or through a wavelet decomposition such as the starlet transform as done in Ajani et al. 2020. The starlet transform decomposes a convergence map κ of size $N \times N$ into a set $\mathcal{W} = \{w_1, ..., w_{j_{max}}, c_J\}$ of $J = j_{max} + 1$ bands of same size, where j_{max} is the number of wavelet scales considered, c_J is the coarse scale, namely a very smoothed version of the original image κ, and w_j are the wavelet bands at scale 2^j pixels. An example to illustrate this is shown in Figure 5.1 and details of this have been outlined in Section 3.4 and in Section 4.1.4.

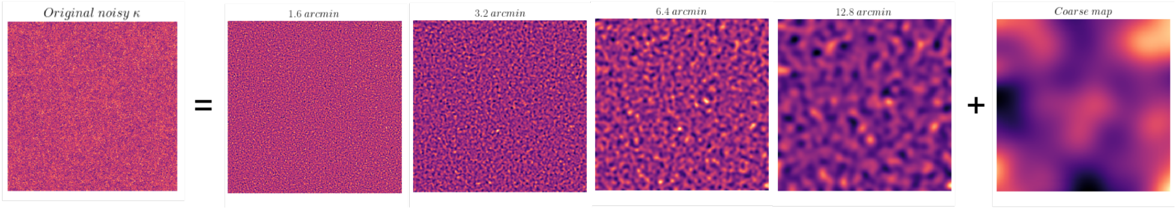


Figure 5.1: Visualisation of the starlet decomposition of an original noisy convergence map. The original noisy convergence map is obtained by applying shape noise to a noiseless realisation of the `MassiveNus` convergence maps products. Then by convolution with a 2D starlet, the original image is decomposed in several different images corresponding to the scale 2^j plus a smoother (coarse) version of it. Here we show the decomposition for a map of the fiducial cosmology. The resolution for the `MassiveNus` simulations convergence maps is ~ 0.4 arcmin.

It was also shown in Leonard et al. 2012 that the starlet transform can be interpreted as a fast multiscale aperture mass decomposition where the aperture mass kernels have a compact support and are compensated (namely, the kernel integral is null). Starlet peaks are then derived by considering n bins with bin edges given by the minimum and the maximum values of each band in $\mathcal{W}$. An interesting advantage of such an approach is that each wavelet band covers a different frequencies range which leads to an almost diagonal peaks counts covariance matrix (Ajani et al. 2020). This is not the case when a standard multiscale Gaussian analysis is applied on the convergence map.

Starlet extrema

In this section we introduce the notation *starlet extrema* to refer to peaks computed on positive signal and minima computed on negative signal that we will use as a statistics to compare against the starlet ℓ_1-norm. We employ starlet peaks computed on maps with only $S/N > 0$ contribution combined with starlet minima computed on maps with only $S/N < 0$ contribution using a 4 scales corresponding to [$1.6'$, $3.2'$, $6.4'$, coarse] based on the resolution of the convergence maps that we use, with 29 linearly spaced bins for each scales between the minimum and maximum values of each S/N map.

As mentioned in the introduction, cosmic voids analysis is an alternative to peaks to study convergence maps, and the combination of both improves the constraints on cosmological parameters. It is interesting to notice how a starlet decomposition can naturally include a multiscale voids analysis. Instead of extracting only maxima (peaks) in each band, we can also extract minima (pixels with values smaller than their eight neighbours), and a joint peaks-voids multiscale is therefore obtained extracting wavelet coefficients extrema (minima + maxima). The starlet decomposition provides therefore a very natural framework for a joint multiscale peaks and voids analysis.

5.1.3 Definition of starlet ℓ_1-norm

A particularity of peak and void statistics is that only a few pixels are considered, while other high-order statistics, such as bispectrum or Minkowski functionals, use *all* the pixels. In a starlet framework, we should emphasise that starlet peaks have mainly positive values and starlet voids negative values due to the property of the wavelet function. So instead of counting the number of peaks or voids in a given bin i defined by two values, B_i and B_{i+1}, we could take the sum of all wavelet coefficients with an amplitude between B_i and B_{i+1}. If B_i and B_{i+1} are positive, this corresponds to the definition of the set of coefficients $\mathcal{S}_{j,i}$ at scale j and in bin i such that $\mathcal{S}_{j,i} = \{w_{j,k} / B_i < w_{j,k} < B_{i+1}\}$, where k is the pixel index. We can then compute the sum $\sum_{u=1}^{\#coef(\mathcal{S}_{j,i})} \mathcal{S}_{j,i}[u]$. This can be generalised to positive and negative bins using:

$$l_1^{j,i} = \sum_{u=1}^{\#coef(\mathcal{S}_{j,i})} \mid \mathcal{S}_{j,i}[u] \mid = ||\mathcal{S}_{j,i}||_1 \tag{5.1}$$

where $||.||_1$ is the standard ℓ_1-norm (i.e. $||x||_1 = \sum_k |x_k|$) and the index u runs from 1 to the number of pixels in a given bin i at scale j (i.e. $\#coef(\mathcal{S}_{j,i})$). The quantity $l_1^{j,i}$ defined in Equation (5.1) is nothing more than the ℓ_1-norm of the binned pixel values of the starlet coefficients of the original image κ map. In the following, we will name S_{ℓ_1}, the starlet ℓ_1-norm, as the set S_{ℓ_1} of all $l_1^{j,i}$ numbers obtained from the different scales j and bins i. This approach enables us to extract the information encoded in the absolute value of all pixels in the map instead of characterising it only by selecting local minima or maxima. An interesting advantage is that it avoids the open issue of how to define a void (Colberg, Pearce, Foster, et al. 2008), briefly discussed in Section 2.3.6. It is interesting to notice that this S_{ℓ_1} statistic is also directly related to the density probability function of the starlet coefficients at different scales.

5.1.4 Methodology

We provide constraints on the sum of neutrino masses M_ν, on the matter density parameter Ω_m and on the power spectrum amplitude A_s by employing five different summary statistics applied to the `MassiveNus` in the same setting as the one described in Section 4.1.2 and

Section 4.1.3. We compute the summary statistics on maps of the signal to noise field where we define the signal to noise as the ratio between the noisy convergence κ convolved with the filter $\mathcal{W}(\theta_{\mathrm{ker}})$ over the smoothed standard deviation of the noise for each realisation per redshift. We estimate the noise level at each starlet scale following Starck and Murtagh 1998 as described in Section 3.5. For the purposes of this study, we mimic the shape noise expected for a survey like *Euclid*[1](Laureijs et al. 2011; Euclid Collaboration et al. 2020b) as defined in Section 4.1.3. We perform a tomographic analysis using four source redshifts $z_s = \{0.5, 1.0, 1.5, 2.0\}$ with corresponding values for the galaxy number density n_{gal} per source redshift bin $n_{gal} = \{11.02, 11.90, 5.45, 1.45\}$. The covariance matrix is computed in the same way as described in Section 4.1.6 and the model to build the likelihood is obtained with the Gaussian Process Regressor as in Section 4.1.7 for the different summary statistics. Analogously to the multi-scale peaks study detailed in the previous chapter, we explore and constrain the parameter space with the `emcee` package, which is a python implementation of the affine invariant ensemble sampler for Markov chain Monte Carlo (MCMC) introduced by (Foreman-Mackey et al. 2013). We assume a flat prior, a Gaussian likelihood function as defined in Section 4.1.7, and a model-independent covariance matrix as discussed in Section 4.1.6. The walkers are initialised in a tiny Gaussian ball of radius 10^{-3} around the fiducial cosmology $[M_\nu, \Omega_{\mathrm{m}}, 10^9 A_{\mathrm{s}}] = [0.1, 0.3, 2.1]$ and we estimate the posterior using 120 walkers. Our chains are stable against the length of the chain, and we verify their convergence by employing *Gelman Rubin* diagnostics (Gelman and Rubin 1992). To plot the contours we use the `ChainConsumer` python package (Hinton 2016). To quantify our results, we employ the same result estimators as the ones introduced in Section 4.1.8.

5.1.5 Summary statistics

In this section I illustrate the different summary statistics used for the analysis.

a) We compute the power spectrum on noisy convergence maps filtered with a Gaussian kernel with smoothing size $\theta_{\mathrm{ker}}^{\mathcal{G}}$ = 1 arcmin. We consider angular scales with 24 logarithmically spaced bins in the range $\ell = [300, 5000]$.

b) Monoscale peaks + voids: we compute peaks (as pixels with values larger than their eight neighbours) and voids (as pixels with values smaller than their eight neighbours) on noisy convergence maps filtered with a single-Gaussian kernel with smoothing size $\theta_{\mathrm{ker}}^{\mathcal{G}}$ = 2 arcmin with 29 linearly spaced bins for peaks between the minimum and maximum of the map in S/N and 29 linearly spaced bins for voids between the negative maximum and positive minimum of the maps in S/N.

c) Starlet peak counts are computed as pixels with values larger than their eight neighbours, computed on noisy convergence maps filtered with a starlet kernel with four

---

[1] `https://www.Euclid-ec.org`

scales corresponding to [$1.6'$, $3.2'$, $6.4'$, coarse], with 29 linearly spaced bins for each scales between the minimum and maximum values of each S/N maps.

d) Starlet extrema are obtained combining peaks computed on maps with only $S/N > 0$ contribution combined with starlet voids computed on maps with only $S/N < 0$ contribution, using a starlet as in c), with 58 (29 for peaks + 29 for voids) linearly spaced bins.

e) the starlet ℓ_1-norm S_{ℓ_1} is computed following Equation 5.1 on noisy convergence maps filtered with a starlet kernel with four scales as in c) and d).

Statistics (d) and (e) are our new proposals for this study. In all statistics where we employ the starlet decomposition, the finest frequency we consider is $\theta_{\mathrm{ker}}^{St} = 1.6$ arcmin, corresponding to the maximum angular scale $\ell_{\mathrm{max}} = 2149$.

5.1.6 Results

In this section, I illustrate forecast results on the sum of neutrino masses M_ν, on the matter density parameter Ω_{m} and on the primordial power spectrum amplitude A_{s} for a survey with *Euclid*-like shape noise in a tomographic setting with four source redshifts z_{s}= [0.5,1.0,1.5,2.0]. We compare results for the different observables defined in Section 5.1.5 and investigate the impact of the choice of the filter on the covariance matrix.

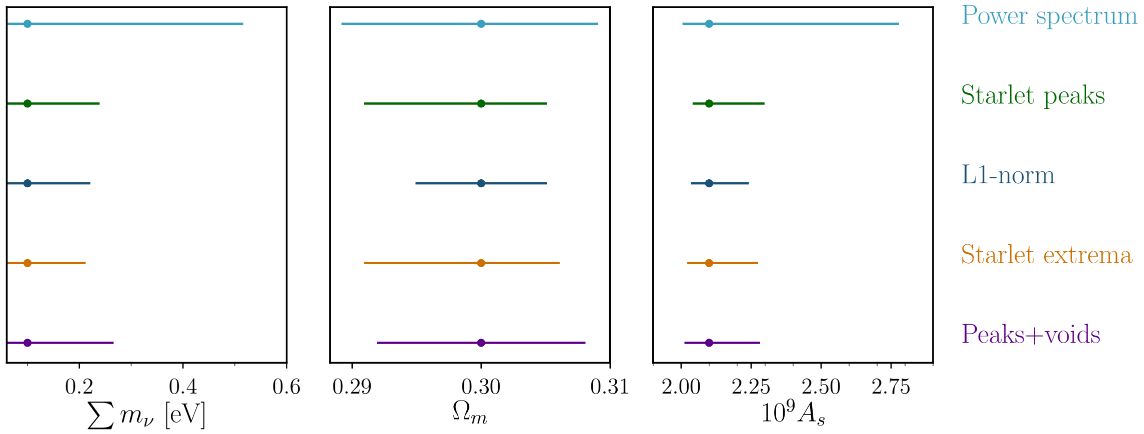


Figure 5.2: Marginalised errors for the observables described in Section 5.1.5, on each parameter showing the 2.5 and 97.5 percentiles with respect to the fiducial model. They refer to a tomographic setting with $z = [0.5, 1.0, 1.5, 2.0]$ with the fiducial model set to $[M_\nu, \Omega_{\mathrm{m}}, 10^9 A_{\mathrm{s}}] = [0.1, 0.3, 2.1]$. The last observable refers to monoscale peaks and voids, as described in the text.

In Figure 5.3 we show the comparison between the constraints obtained using different summary statistics, as described in Section 5.1.5. As expected, we see that all higher-order statistics are more constraining than the power spectrum. The new result of this study is represented by the starlet ℓ_1-norm: the inclusion of all pixels enables to retrieve tighter constraints than the combination of local minima and maxima (voids + peaks). Specifically, for all parameter space planes, the starlet ℓ_1-norm FoM values, illustrated in Table 5.1, are about twice the ones for the local minima and local maxima combined (and more than seven times larger than the power spectrum FoM values).

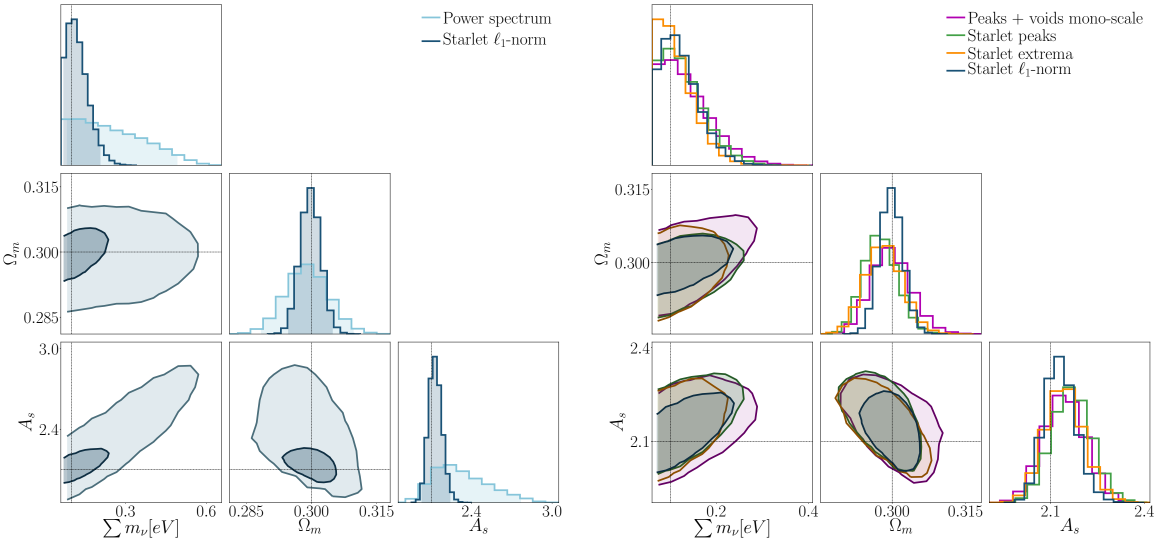


Figure 5.3: 95 % confidence contours tomography with source redshifts $z_\mathrm{s} = [\,0.5, 1.0, 1.5, 2.0\,]$ and corresponding galaxy number density: $n_\mathrm{gal} = [\,11.02, 11.90, 5.45, 1.45\,]$. The black dotted line is the fiducial model: $[\sum m_\nu, \Omega_\mathrm{m}, 10^9 A_\mathrm{s}] = [0.1, 0.3, 2.1]$. **Left panel**: constraints from the power spectrum (light blue contours) computed on noisy maps smoothed with a Gaussian filter with θ_ker =1 arcmin compared to constraints from starlet ℓ_1-norm (dark blue contours) computed on noisy maps filtered with a four scales starlet kernel. **Right panel**: constraints from the combination of peaks and voids (magenta contours) computed on noisy maps smoothed with a Gaussian filter with θ_ker =2 arcmin against constraints from starlet peak counts (green contours), *starlet extrema* (orange contours) and ℓ_1-norm (dark blue contours) computed on noisy maps filtered with a four scales starlet kernel.

In this work, we have also introduced *starlet extrema* as new summary statistics to constrain the parameters. Similarly to the ℓ_1-norm, starlet extrema are computed between the minimum and maximum S/N value of the map, but they are defined as the combination of local maxima computed on $S/N > 0$ and local minima computed on $S/N < 0$ (namely, they do not encode the information present in all pixels). We see that starlet extrema FoM values

Table 5.1: Values of the Figure of Merit (FoM) as defined in Equation 4.11 for different parameters pairs for each observable employed in the likelihood analysis. We provide in the last column the 3D FoM given as the inverse of the volume in $(M_\nu, A_\mathrm{s}, \Omega_\mathrm{m})$ space.

FoM	$(M_\nu, \Omega_\mathrm{m})$	(M_ν, A_s)	$(\Omega_\mathrm{m}, A_\mathrm{s})$	$(M_\nu, \Omega_\mathrm{m}, A_\mathrm{s})$
Power spectrum	1585	77	1079	2063
Starlet Peaks	5904	331	4856	10147
ℓ_1-norm	11408	619	9126	16688
Starlet extrema	6967	442	5956	6740
Peaks + voids monoscale	5321	307	4286	6114

Table 5.2: Values of 1-σ marginalised error as defined in Equation 4.13 for each cosmological parameter for the different observables.

$\sigma_{\alpha\alpha}$	$\mathbf{M}_\nu$	$\mathbf{\Omega}_\mathrm{m}$	$\mathbf{A}_\mathrm{s}$
Power spectrum	0.147	0.005	0.204
Starlet peaks	0.057	0.003	0.064
ℓ_1-norm	0.041	0.002	0.051
Starlet extrema	0.040	0.004	0.062
Peaks + voids monoscale	0.054	0.004	0.067

are larger than starlet peaks and peaks + voids monoscale suggesting that starlet extrema can be a good candidate as well as multiscale higher order statistics. However, the ℓ_1-norm remains the statistics that performs the best with respect to all the summary statistics we have considered in terms of constraining power. In Figure 5.2 we show the marginalised constraints on each cosmological parameter corresponding to the different observables. To compare the improvement obtained by employing the different statistics we compute the 1σ marginalised error for each parameter, summarised in Table 5.2. We find that the starlet ℓ_1-norm outperforms the power spectrum by 72% on M_ν, 60% on Ω_m and 75% on A_s, and the state of the art peaks + voids for a single smoothing scale respectively by 24% on M_ν, 50% on Ω_m and 24% on A_s. Starlet extrema outperform the power spectrum by 72% on M_ν, 20% on Ω_m and 70% on A_s . We also quantify the improvement provided by the ℓ_1-norm with respect of our previous study (Ajani et al. 2020), finding that the ℓ_1-norm starlet outperforms starlet peaks by 28% on M_ν, 33% on Ω_m and 20% on A_s. We also compare the covariance matrices obtained when using starlet extrema and the ℓ_1-norm and we have found that starlet extrema present a more diagonal covariance for the observable.

Diagonal Covariances

As recalled in the introduction, in Ajani et al. 2020 we found that starlet peak counts have the interesting feature that the corresponding covariance matrix is nearly diagonal. Motivated by this, we have tested in this work whether this feature is also maintained for starlet extrema and for the ℓ_1-norm. Interestingly, we find that starlet extrema keep this characteristic while the ℓ_1-norm show more correlations in the off-diagonal term. This can be seen by looking at Figure 5.4: in the left panel we show the constraints for the starlet extrema when using the full covariance matrix (continuous contours) and compare them with the starlet extrema when using only the diagonal elements of the covariance matrix (dashed contours). Analogously, we show the same comparison for the ℓ_1-norm in the right panel. When using starlet extrema as summary statistics, if we employ in the likelihood analysis a covariance matrix with only its diagonal elements, there is no loss of information on M_ν, and a slight loss of information of 25% on Ω_m and 22% A_s. Concerning the ℓ_1-norm, it is sufficient to look at the contours to notice how the contours with only diagonal terms are considerably different with respect to the contours obtained with the full covariance matrix: they appear shifted and they present a different degree of degeneracy for $(\Omega_\mathrm{m}, A_\mathrm{s})$. Hence, we conclude that the ℓ_1-norm outperforms starlet extrema in terms of constraining power when considering the full covariance in the analysis, but presents a less diagonal matrix than starlet extrema. We conclude that depending on the context the ℓ_1-norm could be a convenient choice when the priority is the constraining power while starlet extrema might be more useful when interested in speeding up the analysis or when the covariance matrix can be difficult to invert.

5.1.7 Conclusions and future prospects

In this work we propose to use starlet ℓ_1-norm statistics on weak lensing converge maps to constrain cosmological parameters. The measure of multiscale peak amplitudes can be seen as a measure of the ℓ_1-norm of a subset of positive wavelet coefficients. Similarly, the measure of voids amplitudes can be seen as a measure of the ℓ_1-norm of a subset of negative wavelet coefficients. Wavelets provide therefore a great framework for a joint peaks and voids analysis, in which information from all wavelet coefficients is included. We propose therefore to use a very simple ℓ_1-norm statistics defined as the sum of the ℓ_1-norm of all coefficients in a given S/N (pixel) bin for each wavelet scale, as defined in Equation 5.1. We investigate the impact of employing the starlet ℓ_1-norm as summary statistics computed on weak lensing convergence maps to estimate cosmological parameters and we find that the ℓ_1-norm outperforms the state of the art summary statistics, the power spectrum and the combination of monoscale peaks and voids, respectively of 72% and 24% on M_ν, 60% and 50% on Ω_m and 75% and 24% on A_s. We have further proposed starlet extrema and compared them to ℓ_1-norm: also in this case the latter performs better in terms of constraining power, within the current ideal setting, while starlet extrema present the

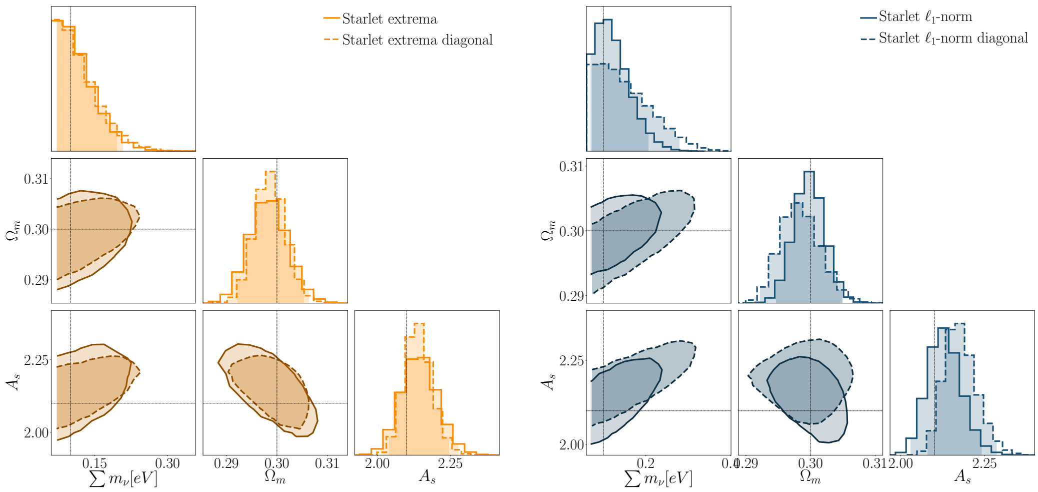


Figure 5.4: 95 % confidence contours tomography with source redshifts $z_{\mathrm{s}} = [\ 0.5,\ 1.0,\ 1.5,\ 2.0\ ]$ and corresponding galaxy number density: $n_{\mathrm{gal}} = [\ 11.02,\ 11.90,\ 5.45,\ 1.45\ ]$. The black dotted line is the fiducial model: $[\ \sum m_\nu,\ \Omega_{\mathrm{m}},\ 10^9 A_{\mathrm{s}}] = [\ 0.1,\ 0.3,\ 2.1\ ]$. **Left panel**: constraints from *starlet extrema* with the full covariance matrix (continuous contours) computed on noisy maps filtered with a four scales starlet kernel against constraints from *starlet extrema* with the only diagonal elements of the covariance matrix (dashed contours). **Right panel**: constraints from ℓ_1-norm with the full covariance matrix (continuous contours) computed on noisy maps filtered with a four scales starlet kernel against constraints from ℓ_1-norm with the only diagonal elements of the covariance matrix (dashed contours).

advantage of a more diagonal covariance matrix. We are aware that to serve as a robust probe for precision cosmology, the statistical power alone is not sufficient; for their usage, it will be important to test how these statistics react in a non-ideal setting, and how their performance is impacted by systematics in the signal. We will dedicate a future study to this aspect.

5.2 Pipeline description and visualisation

This section illustrates a description of the pipeline used in the work presented in this chapter and the one presented in Chapter 4. The input of the pipeline is given by weak lensing convergence maps. These convergence maps are then filtered, with a Gaussian or a starlet filter if they are already noisy, otherwise shape noise mimicking a given survey is added to the maps. Then the summary statistics are computed on the noisy smoothed maps for both simulations at the different cosmologies available and for the data (or fiducial model). With the same settings as the models and data summary statistics, the covariance matrix is computed by concatenating the different observables if tomography or probe combination is taken into account. Then, the observables computed from the simulations (model) are fed to the Gaussian Processes (GP) regressor that learns the relation between the summary statistics and the cosmological parameters and is then able to predict the observable to a new given point in parameter space. Afterwards, the summary statistics for the data, for the model and the covariance, are given as input to the likelihood function and to the MCMC module. The output of this last module provides the chains corresponding to a given summary statistics constraining the parameters that where given as input to the GP training. A visualisation of the procedure is presented in Figure 5.5. Here's a detailed description of the different steps numbered in correspondence of the figure below:

1a) **Convergence map**: The convergence map are provided as initial input in this first version of the pipeline.

1b) **Survey noise**: If the convergence map is built directly from observed ellipticity maps, the noise is intrinsically taken into account following the survey specifications present in the simulations that supply the catalogue. If the convergence map is built directly from γ shear maps, we add the noise to the map by taking the standard deviation of the observed ellipticity of each map for each tomographic bin and follow Equation 4.1.

2) **Noise filtering**: In order to access the cosmological information, the noisy map if convolved with a 2D Gaussian filter or with a 2D *starlet* filter. The *starlet* decomposition with J scales generates $J+1$ different maps, at the scale 2^J, where the $J+1$ map is the *coarse* map. The Gaussian kernel is computed with the python package `scipy`. The starlet decomposition is performed using the CosmoStat python package `lenspack`.

3), 4) **Summary statistics**: On the smoothed/filtered maps (output of 2)), different summary statistics can be computed.

  - Power spectrum
  - Peak counts: single-scale peak counts, multi-scale peak counts, starlet peak counts
  - Void counts: single-scale minima counts, multi-scale minima counts
  - Starlet *extrema*
  - Starlet ℓ_1-norm

  The computation of all summary statistics is implemented in parallel using the python package `joblib`, so that their computation is performed simultaneously for each cosmological model, redshift and realisation.

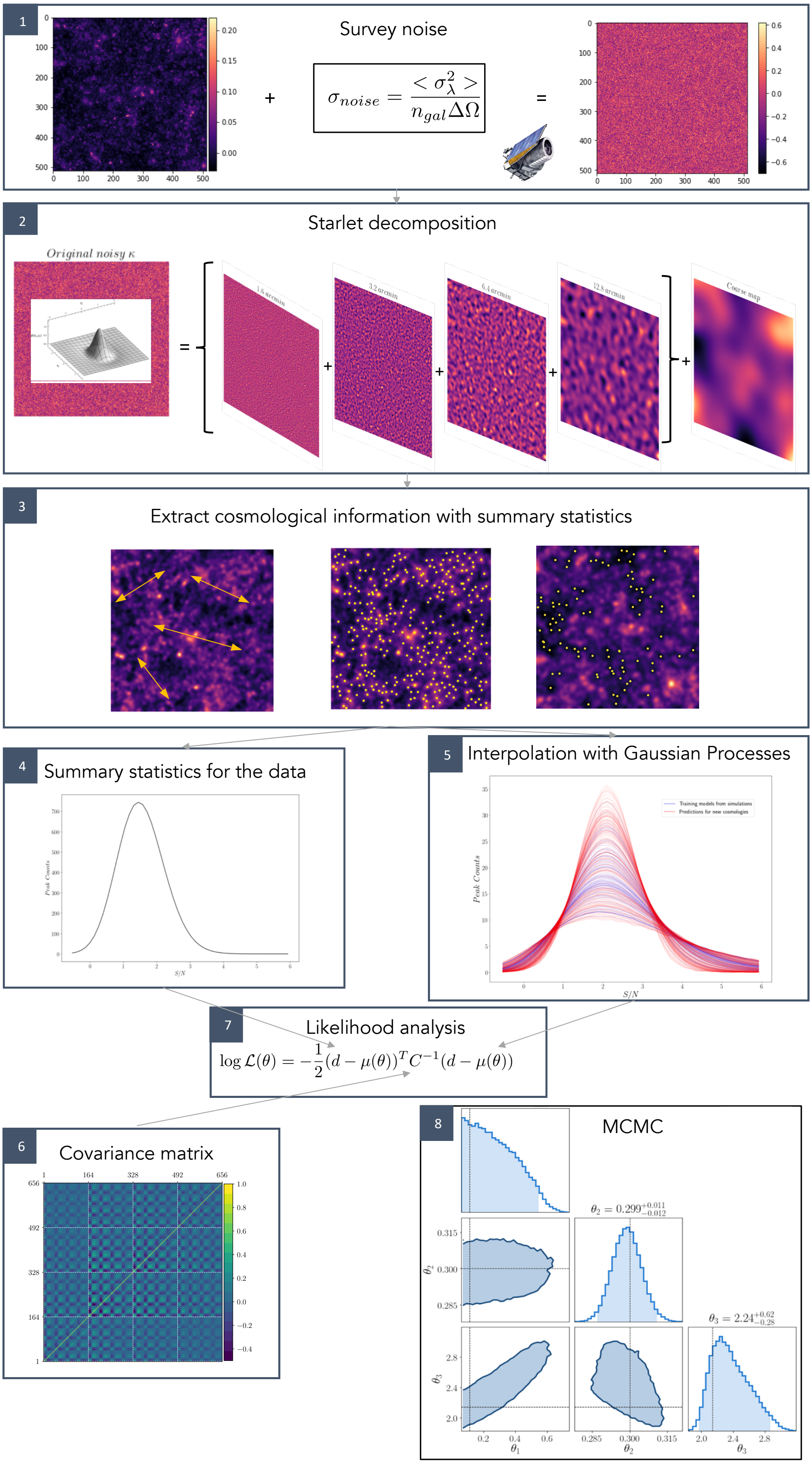
5) **Interpolation with Gaussian Processes**: The interpolation with GP to get the predictions of the observable is implemented for each bin in ℓ or S/N and for each redshift when a tomographic analysis is performed. This means that for an observable with 20 bins and 4 tomographic redshift, 80 different Gaussian Processes will be generated to get the prediction. The Gaussian Process Regressor, the GP kernel and the prediction function are implemented using the `scikit-learn` package.

6) **Covariance**: The covariance matrix takes as input the summary statistic for a given fiducial model and it's computed following Equation 4.9. In practice it is implemented using `numpy`. Depending on the survey sky coverage and the size of the maps, a pre-factor to scale the covariance has to be provided.

7), 8) **Likelihood analysis and MCMC**: The likelihood, prior and posterior functions implemented take as input the cosmological parameters and the GP prediction function. The MCMC is performed using the python packages `emcee` and multiprocessing packages so that parameter space is explored in parallel. The chains are plotted using `ChainConsumer`, that is also employed to compute the mean acceptance fraction and the mean autocorrelation time.

As will be described in Section 6.3.1 a first extension of the input of the pipeline is now in place, allowing to take as initial input directly a galaxy catalogue from which shear maps or convergence maps can be built. This is one of the focus of Chapter 6.

Figure 5.5: Illustration of the different steps of the pipeline.



CHAPTER 6

EXTENSION TO PHOTOMETRIC GALAXY CLUSTERING

This chapter presents the procedure and some preliminary results of an ongoing work (Ajani et al. in prep.[a]) whose aim is to extend the pipeline described in the previous with respect to 1) the considered parameter space 2) the cosmological probe. The new probe in consideration is photometric galaxy clustering, that along with weak lensing and spectroscopic galaxy clustering will be one of the primary probes of next generation galaxy surveys like *Euclid*. Different cosmological probes are affected in different ways by the considered redshifts and by the dynamical ranges of the underlying large-scale structure (Harnois-Déraps, J., Giblin, B., and Joachimi, B. 2019). This, along with the differences in the instruments employed and measurement approaches make it possible to obtain complementary cosmological information and to easily deal with systematic effects as they are typically distinct and uncorrelated in different probes (Harnois-Déraps et al. 2018). Probe combination can therefore exploit these advantages. Moreover, as discussed in details throughout this thesis, statistics of order higher than the second can provide additional non-Gaussian information that is typically lost when using second order statistics only. Using hence a new suite of simulations provided by the first author of Harnois-Déraps, J., Giblin, B., and Joachimi, B. 2019, I describe the settings and procedure followed so far and that we envisage to follow to investigate the impact of probe combination in a higher order statistics framework. Simultaneously, I illustrate the open questions that we still need to address to extend the pipeline to achieve this goal and the extension already accomplished.

6.1 Background definitions

We employ as second order statistics for photometric galaxy clustering, the angular power spectra in harmonic space, $C^{\rm GG}(\ell)$ measured on the galaxy maps, built as described in the next section. Its theoretical description is briefly provided here. Given the galaxy number density $n_i^{\rm G}(z)$ in each tomographic bin i of the photometric survey, the angular galaxy density in such redshift bin is given by

$$\bar{n}_i^{\rm G} = \int_{z_{\rm min}}^{z_{\rm max}} {\rm d}z \, n_i^{\rm G}(z) \, . \tag{6.1}$$

Then, the radial weight function for a given tomographic bin i for galaxy clustering can be defined as

$$W_i^{\rm G}(z) = \frac{n_i^{\rm G}(z)}{\bar{n}_i^{\rm G}} \frac{H(z)}{c} \, . \tag{6.2}$$

The observable spectrum is theoretically given by the following expression

$$C_{ij}^{\rm GG}(\ell) = \int {\rm d}z \, \frac{W_i^{\rm G}(z) W_j^{\rm G}(z)}{H(z) r^2(z)} P_{\rm gg}^{\rm photo} \left[\frac{\ell + 1/2}{r(z)}, z \right] \tag{6.3}$$

where $P_{gg}(\ell, z)$ is the galaxy-galaxy power spectrum, linked to the matter power spectrum $P_{\delta\delta}(\ell, z)$ through the bias $b_{\rm g}$

$$P_{\rm gg}^{\rm photo}(\ell, z) = [b_{\rm g}^{\rm photo}(z)]^2 P_{\delta\delta}(\ell, z). \tag{6.4}$$

In this preliminary part of the work, we consider for our power spectrum measurements only the auto-spectra in the same redshift bin and the galaxy bias $b_{\rm g}$ is fixed in the simulations.

6.2 SLICS and *cosmo*-SLICS simulations

The analysis presented in this section is based on two suites of N-body simulations, the Scinet LIght Cones Simulations (SLICS) (Harnois-Déraps et al. 2018) and the *cosmo*-SLICS (Harnois-Déraps, J., Giblin, B., and Joachimi, B. 2019). The SLICS set is used to get the observables employed in the computation of the covariance matrix, while the *cosmo*-SLICS suite to compute the observables for the model and mock data (fiducial cosmology). Both are based on a series of 100 $\deg^2$ light-cones extracted from N-body simulations with the multiple plane technique. The underlying gravity-only calculations are obtained as in (Harnois-Déraps et al. 2013) and evolve 1536^3 particles in a box of comoving side of 505 h^{-1} Mpc.

The SLICS are specifically designed for the estimation of covariance matrices: they provide of 800 fully independent ΛCDM runs in which the cosmological parameters are fixed to the cosmology $[\Omega_\mathrm{m}, \sigma_8, h, n_s, \Omega_\mathrm{b}] = [0.2905, 0.826, 0.6898, 0.969, 0.0473]$ and the random seeds in the initial conditions are varied. The *cosmo*-SLICS are run with a similar but complementary N-body configuration: the random seeds are fixed, while the cosmological parameters $[\Omega_\mathrm{m}, S_8, h, w_0]$ are sampled with a Latin hypercube at 25 points in parameter space, while keeping the parameters n_s, and Ω_b fixed to the value used in the SLICS. For each cosmology, a pair of simulations is provided, obtained with two different seeds (seed_a and seed_f) to suppress the sampling variance, such that any measurement averaged over the two seeds nearly follows the ensemble mean. For each seed for each cosmology and each bin, 25 different realisations are provided. The parameter space is shown in Figure 6.1. In total, given the for 25 light-cones for the two seeds, for each realisation covering 100 deg^2, the simulations provide 5000 deg^2 per cosmology. Further details on these simulations can be also found in Burger et al. 2021.

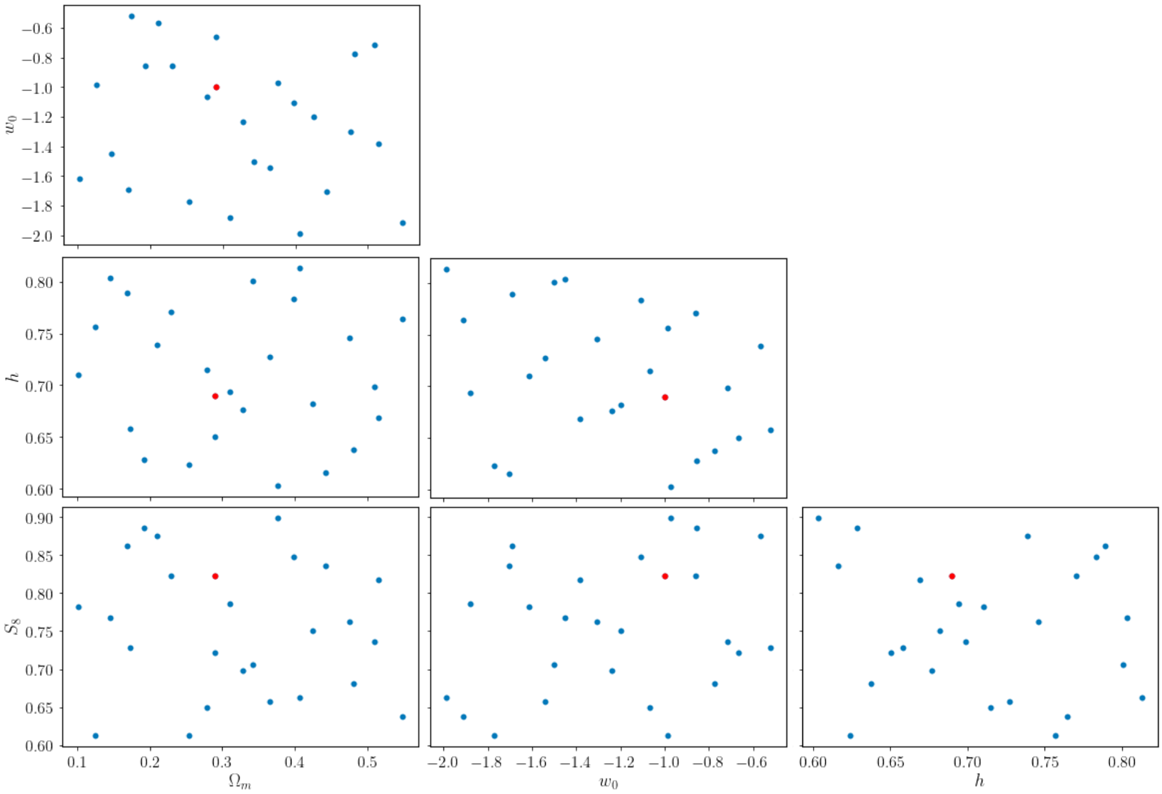


Figure 6.1: Illustration of the different parameter space planes for the cosmo-SLICS suite of simulations. The red dots indicate the values for the fiducial cosmology used to mock the observational data.

Main differences with the `MassiveNus` simulations used in the works described in chapter 4 and chapter 5 mainly consist in:

- the number of cosmologies, being 25 for the *cosmo*-SLICS and 100 for the `MassiveNus`
- the number of realisations per cosmologies, being 25 for the *cosmo*-SLICS and 10000 for the `MassiveNus`
- cosmological parameters: the *cosmo*-SLICS parameter space is 4-dimensional, precisely the varying parameters are $[\Omega_\mathrm{m}, \sigma_8, h, w_0]$. The `MassiveNus` parameter space is 3-dimensional, with varying parameters $[M_\nu, \Omega_\mathrm{m}, A_\mathrm{s}]$.
- the size of the maps, that is $100\ \mathrm{deg}^2$ for the *cosmo*-SLICS and $12.25\ \mathrm{deg}^2$ for the `MassiveNus`.

I want to underline these differences, since one of the aimed extensions for the pipeline described in Section 5.2 is to test its performance for a different parameter space than the `MassiveNus`. Indeed, the `MassiveNus` were the first simulations I have employed as first application of the pipeline, so it is crucial to verify that the pipeline works with different simulations, to be sure that none of the steps is dependent on the simulations and that it works regardless of this choice. We have therefore chosen to use the *cosmo*-SLICS, since they also provide a catalogue for the *lenses* that we will aim to use for combining probes.

6.3 Weak lensing

6.3.1 Building the convergence maps from the catalogue

A first extension that needs to be done concerns the input of the pipeline. Indeed, as the `MassiveNus` directly provide weak lensing convergence maps, the pipeline used to get the results of Chapter 4 and Chapter 5 is built in such a way that it takes directly as input the convergence maps. One of the products of the *cosmo*-SLICS instead, is the galaxy catalogue mimicking KiDS-1000 requirements (Asgari, Marika et al. 2021). This section shows the procedure I followed to build the convergence maps that are then used to compute the summary statistics from the galaxy catalogue provided by the simulations. The final pipeline will hence take as input directly the catalogues and give as output the cosmological constraints.

Galaxy catalogue

The galaxy catalogue provides:

- the positions (x, y) for each galaxy in arcmin

- values for the redshifts z that match to best fit estimates from the data, obtained by a self-organising map method
- values for two components of the shear γ_1, γ_2 assigned to each galaxy
- values for two components of the observed ellipticities $\epsilon_1^{\rm obs}$, $\epsilon_2^{\rm obs}$ assigned to each galaxy.

To get the convergence map, we use Kaiser Squires (KS) inversion: first, we project the shear onto a Cartesian grid and then we smooth it with a Gaussian filter with width 0.7 arcmin following Giblin et al. 2018 to reduce the impact of potential empty pixels. We chose $n_{\rm pix} = 600$ pixels per side. Since the size of each catalogue covers 100 $\rm deg^2$, the resolution is 1 arcmin. Then, we employ Equation 2.26 to get $\hat{\kappa}_\gamma^{\rm est}$. Finally, we perform an inverse Fourier transform to recover $\kappa(\boldsymbol{\theta})$. This is implemented using our CosmoStat python package `lenspack`. In Figure 6.2 we show an example of our reconstructed κ maps.

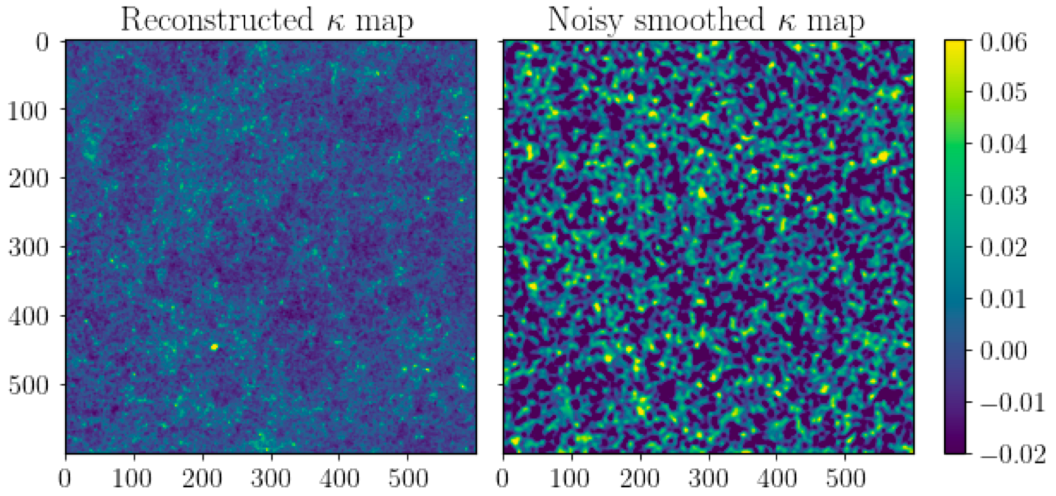


Figure 6.2: κ map reconstructed from the catalogue with KS inversion. The catalogue used to reconstruct this corresponds to the cosmology '00' (the list of cosmologies is available in Table 2 of Harnois-Déraps, J., Giblin, B., and Joachimi, B. 2019) of the *cosmo*-SLICS, for l.o.s.=74. **Left panel**: noiseless κ. **Right panel**: noisy κ, smoothed with a Gaussian kernel of size 3 arcmin.

Noise and filtering

Once the convergence map is obtained, the next step is to add a noise component to take into account shape noise. As the catalogues provide the observed ellipticity for each galaxy, the average shape noise $\langle \sigma_\lambda \rangle$ can be obtained by simply taking the standard deviation of the observed ellipticities for each cosmology and for each tomographic redshift. Then, to get the standard deviation of the noise as in Equation 4.1, we take the total number of galaxies N for each catalogue and we divide it by the area of the map in pix^2 to get the galaxy number density per catalogue. For each tomographic bin, it is possible to build the $n(z)$ with the redshifts values provided from the catalogue. Figure 6.3 shows the $n(z)$ for the different redshifts and the corresponding number of galaxies N that we employ to compute the noise to add to the convergence maps.

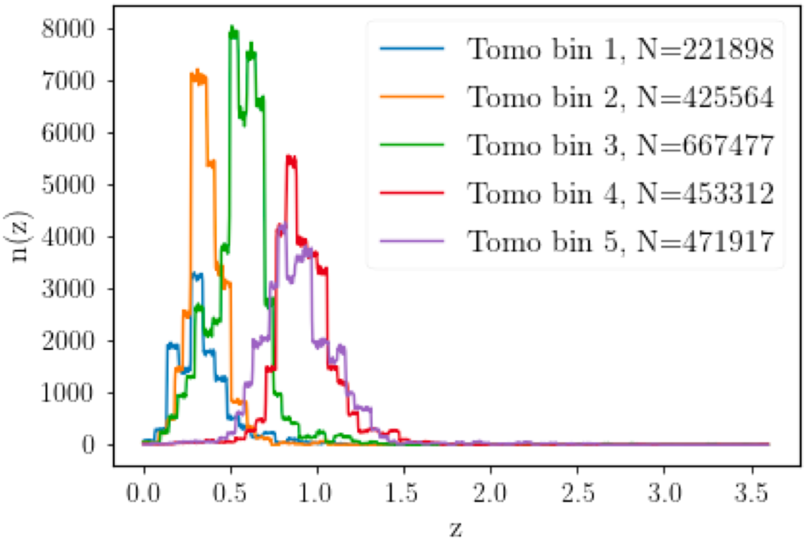


Figure 6.3: Figure showing the different redshift distributions $n(z)$ for each tomographic bin provided by the simulations. The legend shows the number of galaxies N associated to each bin, that is employed to compute the noise that we add to the convergence maps.

As the noisy maps are ready, the next step is to filter them in order to access the cosmological information. We then smooth the noisy convergence maps with a Gaussian kernel of size $\theta_{\mathrm{ker}} = 3$ arcmin.

6.3.2 Map validation with power spectrum

Once the maps are filtered, we compute the power spectrum for each map of each cosmology and each redshift, to build the training set for the Gaussian Processes interpolation. We use the power spectrum to also validate our maps. We do this by comparing our measurement for each cosmology, taking the average over the realisations and over the two different seeds for each tomographic bin to the theoretical predictions, as shown in the left panel of Figure 6.4. We find that the power spectra we measure are in good agreement from $\ell_{\mathrm{min}} = 180$ up to $\ell_{\mathrm{max}} = 4700$ with the theoretical predictions.

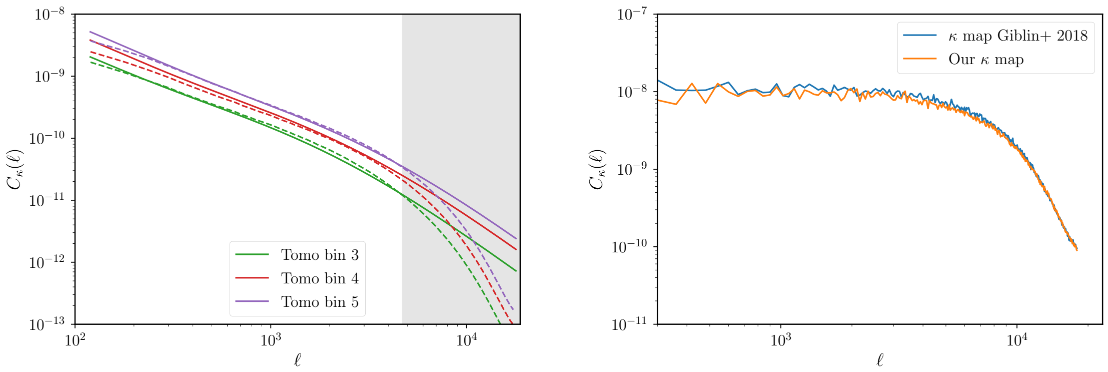


Figure 6.4: Validation of our reconstructed convergence maps from the catalogue with the power spectrum. **Left panel**: comparison between the theoretical predictions (continuous lines) provided by the author of the *cosmo*-SLICS simulations (Harnois-Déraps et al. 2018) and the power spectra measured on our noiseless κ maps reconstructed from the catalogue shear components (γ_1, γ_2). **Right panel**: comparison between the power spectrum measured on the map provided by the authors of Giblin et al. 2018 and the power spectrum measured on our map. Specifically, both maps are built from the observed ellipticities ($\epsilon_1^{\rm obs}$, $\epsilon_2^{\rm obs}$) provided by the catalogue for the first tomographic bin of the SLICS simulations for KiDS-1000, for l.o.s. 74.

We also perform an extra validation, by comparing the power spectrum measured on our reconstructed map with one map provided by the authors of Giblin et al. 2018, as we follow a similar method to reconstruct the κ map [1], and we find very good agreement between the two, as shown in the right panel of Figure 6.4. The power spectra used to validate against the theoretical predictions are measured on noiseless κ maps, built from the two components of the shear (γ_1, γ_2) from the catalogue. The power spectrum used to validate against the map from Giblin+ 2018, is measured on a noisy smoothed map, built from the two components of the observed ellipticities $(\epsilon_1^{\rm obs}, \epsilon_2^{\rm obs})$ from the catalogue. The power spectra employed as summary statistics to perform the parameter estimation analysis are computed as the latter, namely measured on κ maps that are built from the two components of the observed ellipticities $(\epsilon_1^{\rm obs}, \epsilon_2^{\rm obs})$, hence intrinsically accounting for the shape noise of each map. We use 33 bins in the range $\ell = [180, 2040]$, that lies within the region where we are confident that the power spectra are in agreement with the theoretical predictions (white region in the left panel of Figure 6.4).

---

[1] The map they provided is built from the same catalogue (for KiDS-1000) we are using.

6.4 Photometric Galaxy Clustering

6.4.1 Building the maps from the catalogue

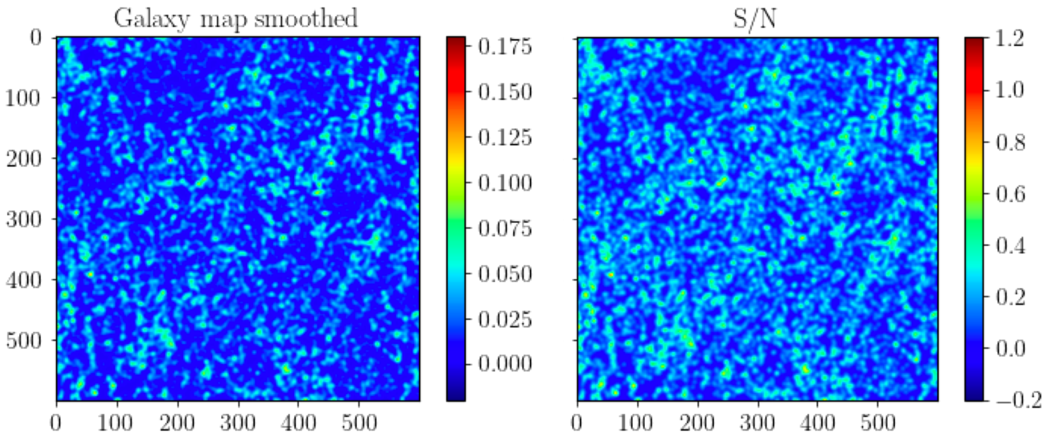


Figure 6.5: **Left panel**: example of galaxy map built from the LRGs cosmo-SLICS catalogue, smoothed with a Gaussian filter of size 3 arcmin. **Right panel**: S/N map built from the LRGs cosmo-SLICS catalogue, where we provide as estimate of the noise, the mean of the map.

In order to obtain the GCph observables we build galaxy density maps from the d KiDS-1000-like lens catalogues of Luminous Red Galaxies (LRGs) provided by the *cosmo*-SLICS simulations. This set of photometric redshift galaxy foreground samples of LRGs mock catalogues is available for each of the *cosmo*-SLICS cosmologies illustrated in Figure 6.1 and for 4 tomographic bins. A recent study employing the same catalogue was performed by Burger et al. 2021. Each catalogue is constituted by the positions of each galaxies in RA and DEC and redshift. The redshifts of the LRGs match exactly the redshifts of projected mass sheets, from which they originate. So they appear to be very discrete, separated by about $\Delta\chi$ = 257 Mpc/h. The catalogues are generated assuming a constant linear galaxy bias $b_{\rm g} = 1.72$, with a galaxy number density of $n_0 = 0.028$ gal/arcmin^2. Starting from this catalogue we build galaxy density maps for each cosmology and each realisation as follows:

- we bin the provided positions in RA and DEC into a 2D grid of 600×600 pixels
- we smooth it with a Gaussian filter of size 3 arcmin

an example of the resulting map is shown in the left panel of Figure 6.5. As for the computation of the high order statistics, we are interested in the S/N, we estimate the noise as the mean of our map. The S/N is then defined as the galaxy map smoothed over the standard deviation of our noise estimate, for each map. One example is shown in the right panel of Figure 6.5.

6.4.2 Map validation with power spectrum

We validate our maps by comparing our measured power spectra against the theoretical prediction obtained with the python library `pyccl`, corresponding to Equation 6.3.

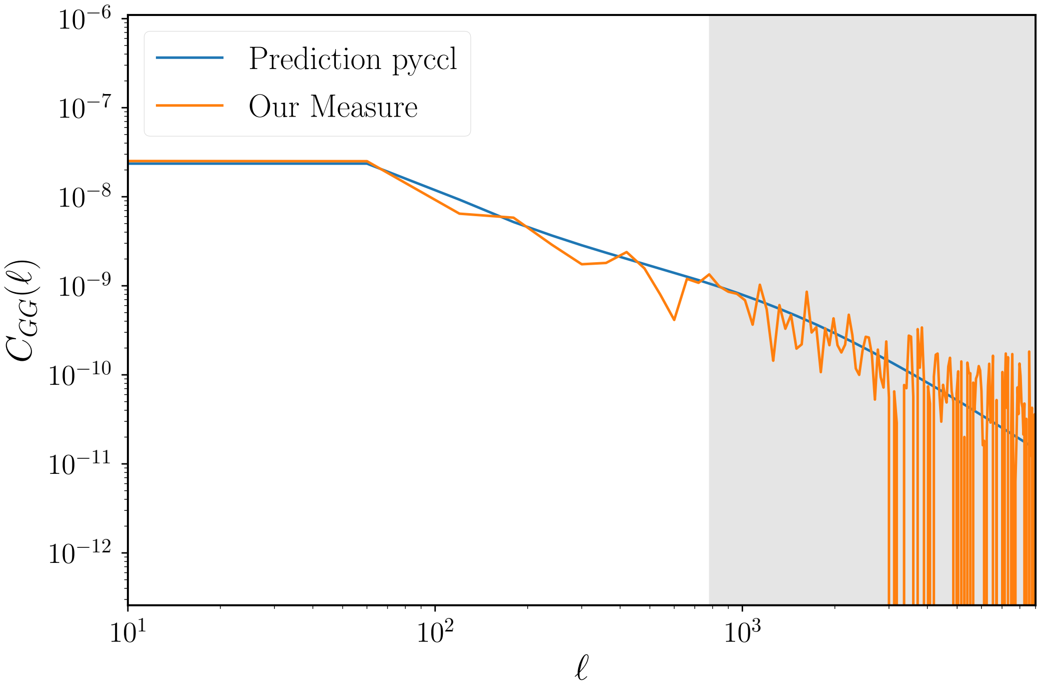


Figure 6.6: Comparison between the power spectrum measured on a galaxy map we built from the LRGs lens catalogue (orange line) and the theoretical prediction obtained with `pyccl` (blue line). The gray region indicates values for the multipoles $\ell > \ell_{\mathrm{max}}$ where $\ell_{\mathrm{max}} = 780$ is the maximum value used for the analysis.

In order to do so we measure the power spectrum on the galaxy density map built as described in the previous section. As the map is intrinsically noisy (due to the presence of Poisson noise), while the theoretical prediction is obtained for a noiseless case, we compute the power spectrum of the noise and then subtract it from the power spectrum of the map. As first approximation, we estimate the noise of the map as its mean. An example of this comparison for one realisation is shown in Figure 6.6. At high ℓ a noisy behaviour can be noticed. This noise probably is intrinsic in the data as the shot noise has already been subtracted as described before. The scales where this happens are in any case excluded from the analysis, as we include multipoles up to $\ell_{\text{max}} = 780$ as explained in Section 6.5.2. Moreover, when averaging over the different l.o.s. for each cosmology (which is what we do when performing the analysis) the noisy behaviour disappears.

6.5 Preliminary results

This section presents some preliminary results of this project. The final goal is to obtain cosmological forecasts for combined weak lensing and photometric galaxy clustering for both second and higher order statistics. Here some preliminary outputs are shown, some of them being useful to validate some already implemented extensions of the pipeline presented in Section 5.2 and others to understand what are the next steps to follow to extend it to reach the aim of this project. I would like to stress that these outputs are still preliminary and do not represent final quantitative results. At this stage of the project, they should be read as first experiments to understand if some approximations we make have to be modified and how. To get the parameter constraints for all summary statistics we follow the same procedure as in Chapter 4 and Chapter 5. To get the model predictions and build the likelihood we employ Gaussian Processes interpolation, in the same fashion we did in Section 4.1.7. We assume flat priors, in particular spanning the following the ranges for each parameter: $\Omega_{\text{m}} = [0.001, 0.6]$, $h = [0.5, 0.9]$, $w_0 = [-2, -0.3]$, $\sigma_8 = [0.5, 1.3]$ and when we consider S_8 instead, $S_8 = [0.1, 1.3]$. To compute the covariance matrix we employ the SLICS suite of simulations, that is obtained at a fixed cosmology ($\Omega_{\text{m}} = 0.2905$, $h = 0.6898$, $w_0 = -1$, $\sigma_8 = 0.826$, $\Omega_{\text{c}} = 0.2432$, $\Omega_\Lambda = 0.7095$). The final SLICS suites for KiDS-1000 will provide 800 independent realisations. However, at the moment these first tests were run, we found some missing l.o.s., therefore we used the 715 realisations that are available at the moment. We scale the covariance matrix with the pre-factor $f_{\text{map}}/f_{\text{sky}} = 100/1000$, since the size of the maps corresponds to $10 \times 10\ \text{deg}^2$ and the catalogue is built for KiDS-1000 covering $1000\ \text{deg}^2$. We then perform MCMC with thee python package `emcee`, described in Section 1.5.3.

6.5.1 Weak lensing (WL) power spectrum

We perform a cosmological parameter inference analysis using the power spectrum computed as described in Section 6.3.2. We use 33 bins in the range $\ell = [180, 2040]$. We provide forecasts for the matter density parameter Ω_{m}, the clustering amplitude σ_8, their combination $S_8 \equiv \sigma_8\sqrt{\Omega_{\text{m}}/0.3}$, the dark energy equation of state parameter w_0 and the reduced Hubble parameter h. We find that the results obtained with our pipeline are consistent with the cosmic shear analysis of the KiDS fourth data release (KiDS-1000) (Asgari, Marika et al. 2021). In particular, in Figure 6.7, we compare our result with the marginal posterior from mock data for the second order statistics employed by Asgari, Marika et al. 2021 for the parameter space plane $(\Omega_\mathrm{m}, \sigma_8)$. We consider this step as an important validation of the first extension of the pipeline. In particular, the step that goes from the catalogue to the convergence maps was not present in the first version of the pipeline as it was used for its first applications (Chapter 4, Chapter 5). After this extension its input is directly the galaxy catalogue, and it provides the possibility to work with shear maps as well as with convergence maps obtained with a KS inversion.

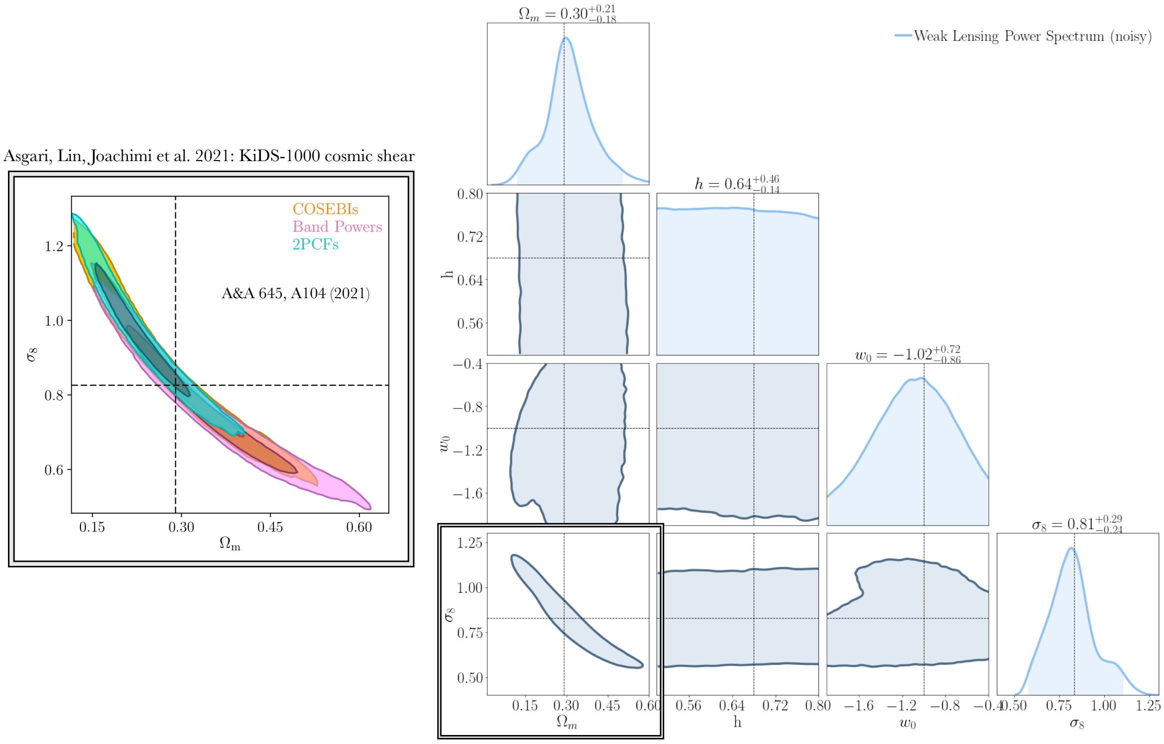


Figure 6.7: Preliminary forecasts derived employing the power spectrum as summary statistics. The analysis is tomographic, performed with the 5 redshift bins shown in Figure 6.3. We compare our results on the right (blue contours) with the results obtained by (Asgari, Marika et al. 2021) for different second order statistics for KiDS-1000 mock data (smaller panel on the left).

6.5.2 Galaxy clustering photometric (GCph) power spectrum

As a second test, we want to provide forecasts for the matter density parameter Ω_m, the clustering amplitude σ_8, their combination $S_8 \equiv \sigma_8\sqrt{\Omega_\mathrm{m}/0.3}$, the dark energy equation of state parameter w_0 and the reduced Hubble parameter h. We perform a cosmological parameter inference analysis measuring the power spectrum on our galaxy maps. The value for the galaxy bias is fixed in the simulations at $b_\mathrm{g} = 1.72$. For the angular scales, we employ a conservative cut, by removing the contributions of multipoles $\ell > 780$ from the analysis, following the *pessimistic* setting of Euclid Collaboration et al. 2020b. Specifically, We use 14 bin in the range $\ell = [120, 780]$. We show the result in the left panel of Figure 6.8. Employing photometric galaxy clustering instead of weak lensing brings very tight constraints in this setting: the constraining power of GCph seems to be actually more constraining than expected. Whether the reason for this lies in some limitations of the simulations or in some approximations we have made, as we estimate the noise for instance, is currently under investigation. What we know for sure is that the clustering varies with the galaxy bias which is fixed in the simulations, but largely unknown in the data. That would inflate the error contours significantly in more realistic conditions. On the right panel of Figure 6.8 we also show the constraints on the $(\Omega_{\rm m}, S_8)$ plane for both probes and their combination. As expected, the combination brings an improvement on the constraining power for both probes, and the GCph contribution is dominant. However, it is important to stress that this is a first test of the pipeline on a new probe, as well as the fact that the bias is actually largely unknown in the data, so we expect that once applied to real data, the error contours for GCph will inflate significantly. At this stage, we do not treat these outcomes as final quantitative results but more as first experiments to understand how to properly include this new probe.

6.5.3 Photometric Galaxy Clustering starlet peaks

The next extension involves the inclusion of higher order statistics computed on the galaxy maps. We therefore attempt a first experiment by applying the starlet peak counts framework described in Section 4.1. We express the maps in terms of their signal to noise ratio S/N and built them as depicted in Section 6.4.1. The size of the filters for the high order statistics are chosen such that when comparing the results for second order statistics (GCph power spectrum) the scales considered roughly match to have a fair comparison. For example, since in the GCph power spectrum analysis the considered range in scale is $\ell = [120, 780]$, for maps with resolution 1 arcmin, we need to consider that for $j_{\rm max} = 6$ scales, the larger scale gives $\theta = 64$ arcmin, corresponding to $\ell = 54$ and the finest scale $\theta = 2$ arcmin, corresponds to $\ell = 1719$.

Hence, we only include scales corresponding to $\theta = [4, 6, 16, 32]$ arcmin leading to $\ell = [859, 430, 215, 107]$. The scales considered in the analysis and corresponding galaxy maps are shown in Figure 6.9. We then compute the peak counts for the different starlet scales and perform the MCMC analysis. We show the corresponding 95% confidence contours in Figure 6.10. Qualitatively we see that for the majority of the parameters, except h and $\Omega_{\rm m}$ the starlet peaks for GCph tend to provide tighter contours. However, we also notice that the constraints on w_0 tend to induce a slight shift. The next step is then to investigate whether this is due to some approximation in the analysis such as the way we estimate the noise or in some limits of the simulations as the fixed galaxy bias or to other origins.

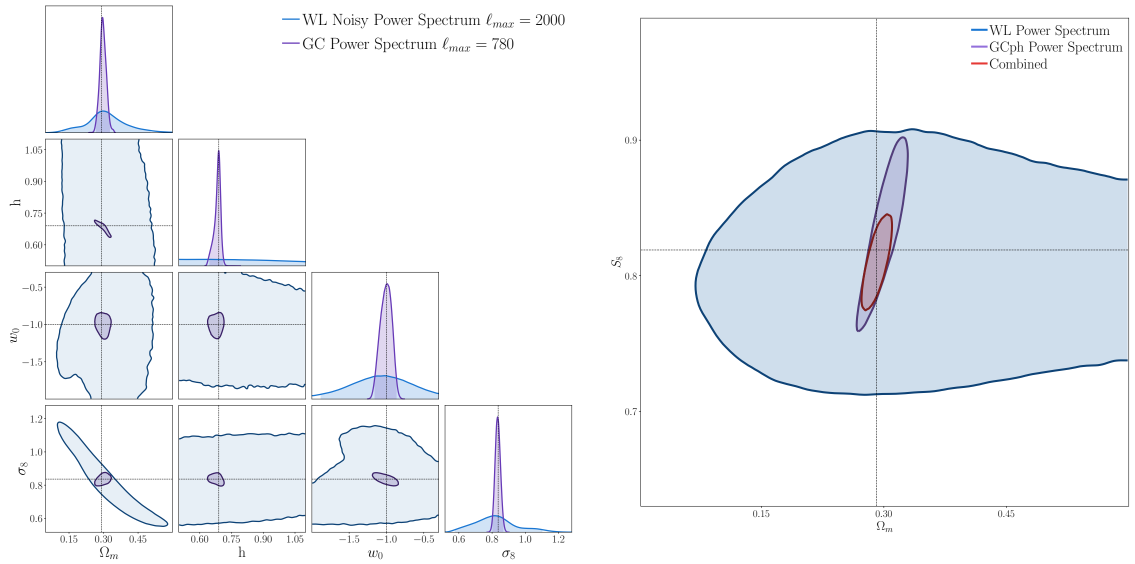


Figure 6.8: The plots shown in figure are very preliminary and do not represent yet quantitative results. They are illustrated to show some first tests we run on a new probe, to understand the next step to undertake to properly incorporate the GCph analysis. **Left panel**: Cosmological forecasts for Ω_{m}, σ_8, w_0 and h for five redshift bins with weak lensing power spectrum (blue contours) and four redshift bins for photometric galaxy clustering power spectrum (violet contours). As mentioned above, the constraining power of GCph seems to be actually more constraining than expected and we expect the error contours to significantly inflate in more realistic conditions, starting with a varying galaxy bias. **Right panel**: joint forecast for Ω_{m} and $S_8 \equiv \sigma_8\sqrt{\Omega_{\mathrm{m}}/0.3}$ for weak lensing power spectrum (blue contours) and galaxy clustering photometric power spectrum (violet contours) and their combination (red contours). We stress also here that the constraining power of GCph seems to more constraining than expected and that the reason for this (if it is due to the simulations or to some approximations we have made) is currently under investigation. What we know for sure is that the value for the galaxy bias is fixed at $b_{\mathrm{g}} = 1.72$ and as the bias is actually largely unknown in the data, we expect that once applied to real data, the error contours for GCph will inflate significantly.

6.6 Ongoing work and future prospects

The results shown in this section are very preliminary as the work on this project is currently ongoing. The goal of this first part of the project mainly consists in understanding in which aspects the pipeline presented in Section 5.2 needs to be extended to be able to deal with a different parameter space and a different cosmological probe. Concerning the first point, using a different set of simulations (the *cosmo*-SLICS instead of the `MassiveNus`)

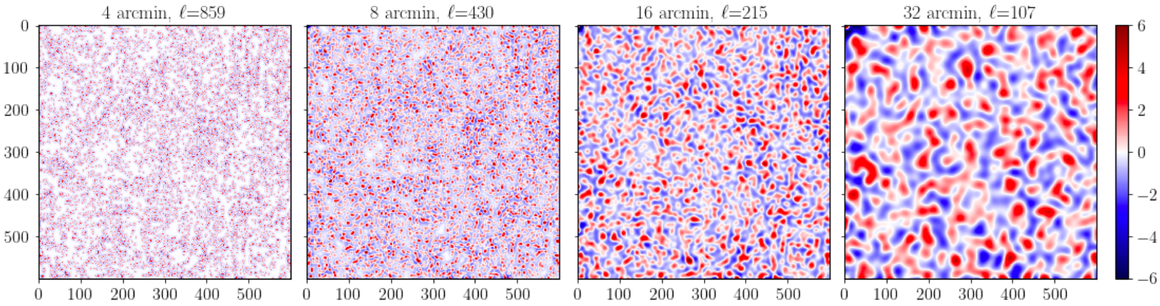


Figure 6.9: Starlet decomposition for a galaxy map. The scales shown in the figure are the ones included in the analysis. The value in arcmin corresponds to the resolution of the corresponding starlet scale, considering that the resolution of the original map is 1 arcmin. The corresponding values in ℓ are also shown.

allowed us to extend the analysis to a $4-$dimensional parameter space, and most importantly to add the initial step of deriving the convergence maps from the catalogue so that now, the initial input is the galaxy catalogue. Concerning the second point, by employing the mock lens catalogue for the KiDS-1000 LRGs set, allowed us to start including in the analysis photometric galaxy clustering as a new probe. It is indeed known that probe combination is crucial to increase the gain in information as different probes are sensitive to different dynamical ranges of the underlying structure formation and to different redshifts (Harnois-Déraps et al. 2018). For example, weak lensing and galaxy clustering depend on two independent combinations of the gravitational potentials (the growth of structure and deflection of light) and can therefore help in testing gravity (Amendola et al. 2018; Harnois-Déraps et al. 2018; Euclid Collaboration et al. 2020b; Heymans et al. 2020). Furthermore, as detailed throughout this thesis, statistics of order higher than the second enable to include the non-Gaussian information of physical fields present at non-linear scales and as a consequence to improve the constraining power. Therefore, we are willing to investigate what is the impact of combining probes for statistics of order higher than the second on the constraints, with a particular interest in testing our starlet ℓ_1-norm that has been proven potentially powerful in the context of weak lensing (Section 5.1.2). As one can see from these first preliminary results, there are many aspects that need to be clarified and are the focus of my current work. In practice, the next steps involve to better understand why the contours for the GCph power spectrum are more constraining than expected and then to understand the behaviour of starlet peaks on the galaxy maps by quantifying the impact of our approximations (concerning the noise estimation and the fixed bias). Once these steps are completed, the goal is to compute the starlet ℓ_1-norm for GCph, WL and study the impact of their combination on cosmological parameter as well as the combination of each single probe with the corresponding second order statistics (the angular power spectra in this case). Depending on the outcome of this analysis in an ideal setting, the natural follow-up will be then to apply this on real data. I would like to underline that, despite the exploratory nature of the results presented so far, the use of a new suite of simulations enabled a first extension of the pipeline in terms of initial input. Indeed, the steps described in Section 5.2 can be extended at the beginning with the following items

Sources Galaxy catalogue: The galaxy catalogue is provided usually as `fits` or `dat` file. It contains the galaxy position (x, y), the redshift z, the observed ellipticities $(\epsilon_1^{\rm obs}, \epsilon_2^{\rm obs})$ and the values for the two components of shear (γ_1, γ_2). Public python packages as `astropy` can be used to read it.

Pixelated shear map: Taking as input the observed ellipticity or the shear values provided by the catalogue, a shear map is built by binning the shear values according to position. This is performed using the CosmoStat python package `lenspack`.

Convergence map: The convergence map is then computed by Kaiser Squires (KS) inversion. This is performed using the CosmoStat python package `lenspack`.

Lenses Galaxy catalogue: The galaxy catalogue is provided usually as `fits` or `dat` file. It contains the galaxy position in RA and DEC and the redshift z. Public python packages as `astropy` can be used to read it.

Galaxy map: The galaxy map is obtained by binning the positions in RA and DEC into a cartesian grid of $n_{\rm pix} \times n_{\rm pix}$.

that basically summarise the procedure described in Section 6.3.1 and Section 6.4.1. In this way, the pipeline illustration of Figure 5.5 can be updated by adding the initial step illustrated in Figure 6.11. Starting from a lenses catalogue instead of a sources one, this is automatically extended to photometric galaxy clustering. The procedure make it now possible to obtain constraints on cosmological parameters as output starting from a mock catalogue of lenses or sources as input. As briefly mentioned also in this Chapter, the natural extension of the pipeline is actually to apply it to more realistic settings, to be ready to apply it to real data. Along with the extension for new probes which is the focus of the project introduced in this Chapter, a parallel effort to extend it to more realistic conditions is described in the next chapter.

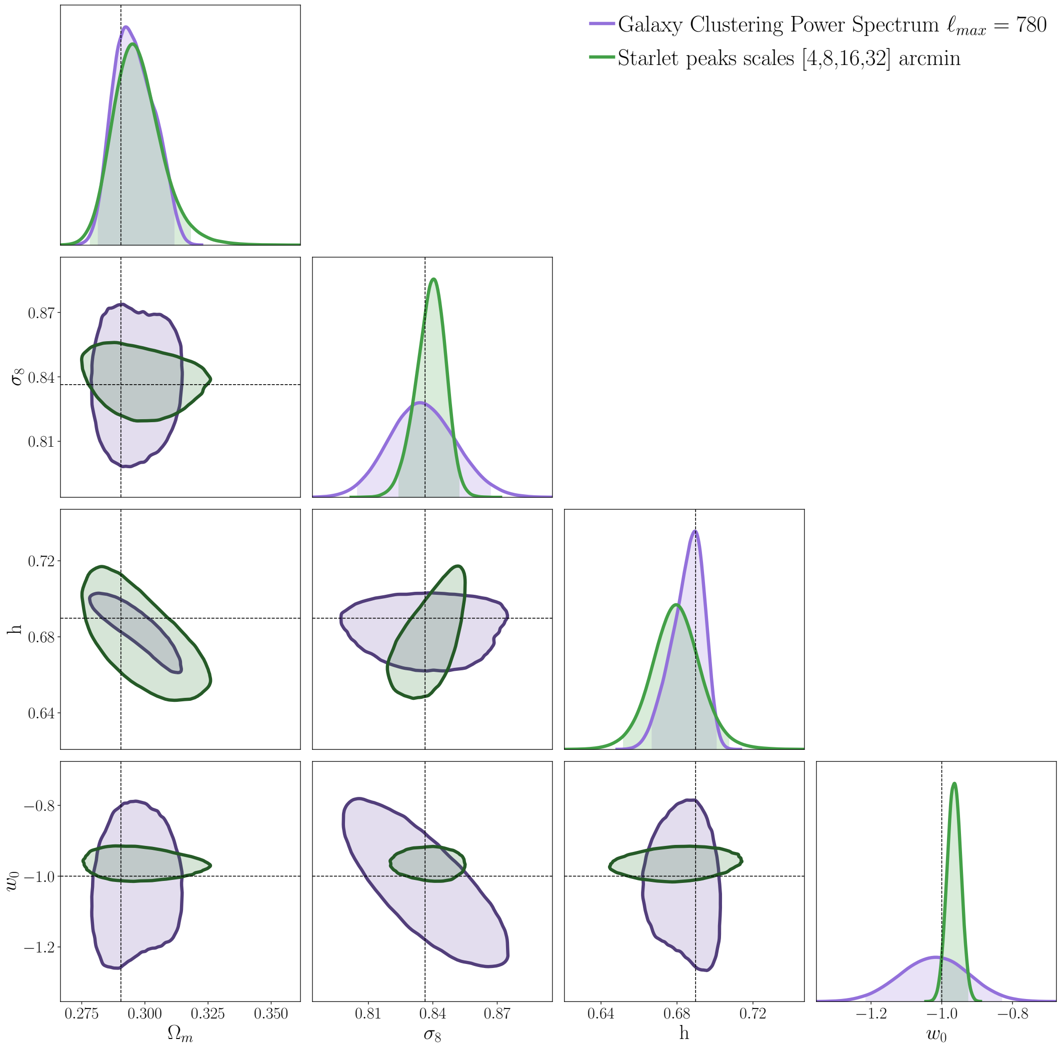
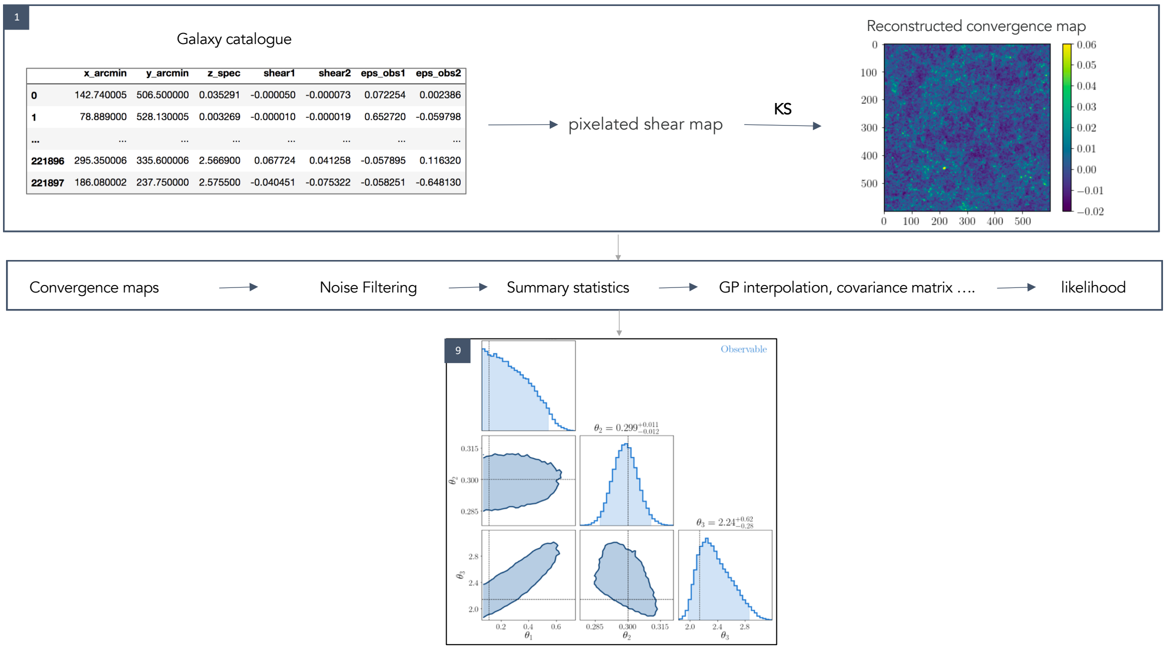


Figure 6.10: This plot represents a preliminary test for the extension of our method to GCph second order statistics and starlet peaks. As discussed above, the contours for the power spectrum seem to be more constraining than expected and we are currently investigating if this might be due to some limitations of the employed simulations or to some approximations we made throughout the analysis. I illustrate here a first attempt to get cosmological forecasts for Ω_m, σ_8, w_0 and h for four redshift bins with galaxy clustering photometric power spectrum (violet contours) and starlet peaks (green contours). The scales considered for the starlet decomposition are chosen to roughly match the range in ℓ for the power spectrum. We notice that the constraints for w_0 tend to slightly shift the contours. It is to be investigated weather this is caused by some assumption made in the analysis as noise estimation or the fixed galaxy bias.



	x_arcmin	y_arcmin	z_spec	shear1	shear2	eps_obs1	eps_obs2
0	142.740005	506.500000	0.035291	-0.000050	-0.000073	0.072254	0.002386
1	78.889000	528.130005	0.003269	-0.000010	-0.000019	0.652720	-0.059798
...	...	...	...	...	...	...	...
221896	295.350006	335.600006	2.566900	0.067724	0.041258	-0.057895	0.116320
221897	186.080002	237.750000	2.575500	-0.040451	-0.075322	-0.058251	-0.648130

Figure 6.11: Extension of the initial input of the pipeline illustrated in Figure 5.5.

CHAPTER 7

FIRST APPLICATION TO REAL DATA

The Canada-France Imaging Survey (CFIS), is a u- and r-band survey covering $5000\,\mathrm{deg}^2$ in the Northern hemisphere. Its high resolution and depth make it a great survey candidates for weak lensing science to date. CFIS is also a critical component of the data needed for photometric redshifts for *Euclid*. A first attempt of applying the pipeline to real data from the CFIS survey and preliminary results are presented. This Chapter presents the procedure and some preliminary results of an ongoing work (Ajani et al. in prep.[b]).

7.1 CFIS survey

The Canada-France Imaging Survey (CFIS), is a u- and r-band survey covering $5000\ \mathrm{deg}^2$ in the Northern hemisphere. Since 2017 it has covered $3000\ \mathrm{deg}^2$, and will cover $4800\ \mathrm{deg}^2$ in 2025. Its high resolution and depth make it a great survey candidates for weak lensing science to date. CFIS is also a critical component of the data needed for photometric redshifts for *Euclid*. In particular, the survey parameters for CFIS are also chosen to provide essential, complementary, sky coverage to *Euclid*. The optimal area for CFIS-r to best complement *Euclid* is given by the the northern $5000\ \mathrm{deg}^2$ of the extragalactic sky. Figure 7.1 illustrates the sky coverage for CFIS relative to some complementary surveys. CFIS-u requires extensive sky coverage away from the Galactic plane, whereas the primary science drivers for CFIS-r require SDSS spectroscopy. An area as large the one of the SDSS region as possible is hence covered, leading to a $n_{\mathrm{gal}} = 7\ \mathrm{gal}/\mathrm{arcmin}^2$ for lensing. This value for the galaxy density is similar to the ones of RCS2, KiDS and DES, but over a much larger area than the first two, and a better delivered image quality than the latter (CFIS coll. 2017).

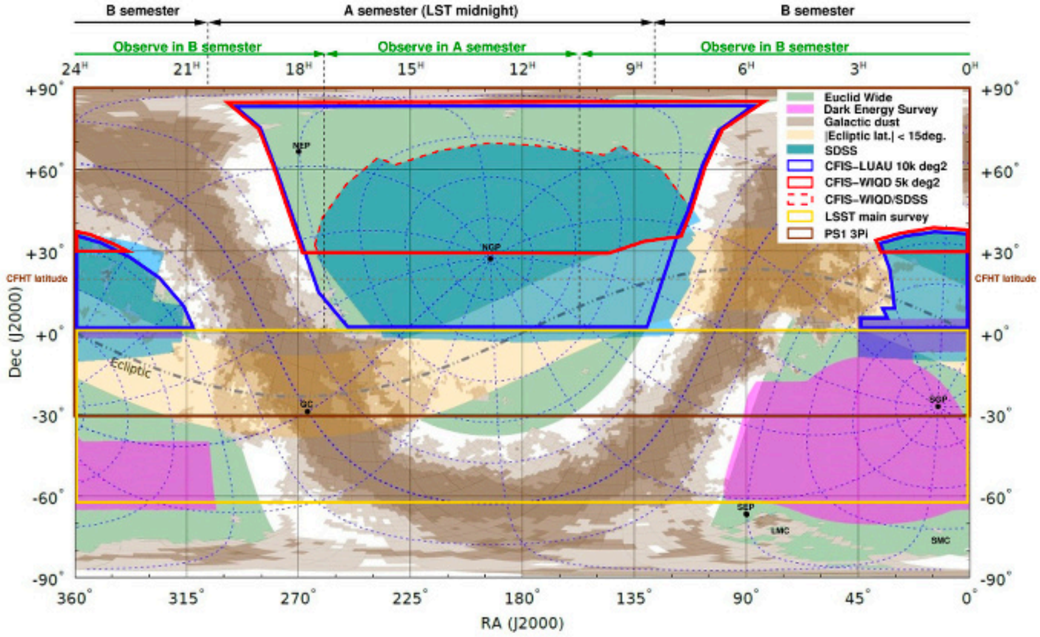


Figure 7.1: Image from `https://www.cfht.hawaii.edu/Science/CFIS/cfissurvey.html`, showing the equatorial projection of the entire sky for CFIS-r (red outline) and CFIS-u (blue outline) with respect to other surveys (listed in the top right caption). The areas that will be observed from the A and B semesters are indicated at the top. Image credit: T. Dwelly.

7.2 Analysis setting

This Section describes the setting chosen for the analysis. We employ as data convergence maps built from CFIS catalogue as described in the following and as simulations we use the `MassiveNus` simulations introduced in Section 4.1.2 that we correct including baryonic effects following Coulton et al. 2020a. We compute peak counts as summary statistics, we assume a Gaussian likelihood, a cosmology independent covariance, and perform MCMC using the `emcee` python package.

7.2.1 Data

As starting point of our analysis we take a convergence map built by Guinot 2020 through the pipeline `ShapePipe` (Guinot et al. in prep.). In particular, we consider the patch P3 of the four patches presented in Guinot 2020 of 806, 282, 249 and $358\ \mathrm{deg}^2$, corresponding to P1, P2, P3, and P4, respectively, shown in Figure 7.2.

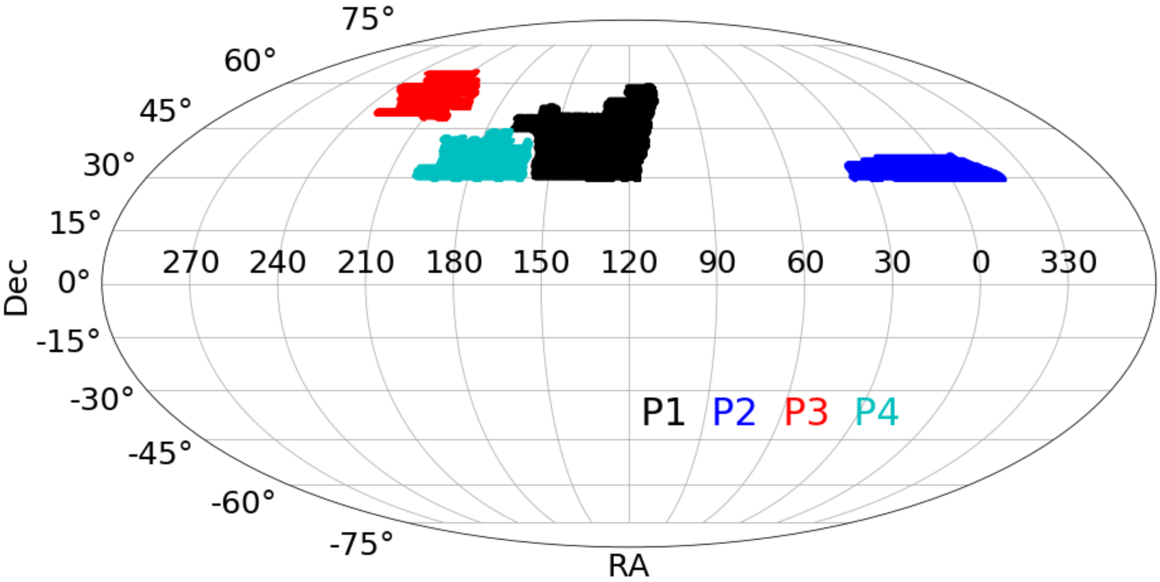


Figure 7.2: Process patches of the CFIS dataset provided by Guinot 2020. In this analysis, we employ the P3 (red patch) corresponding to a size of 249 deg^2.

The patches represent around 5200 single exposures and 6800 tiles, with a total area of $1695\ \mathrm{deg}^2$. From the catalogue provided by `ShapePipe`, the κ map for the patch P3 shown in Figure 7.3 is built, through the KS inversion implemented in `lenspack` in a way that the resolution in pixel corresponds to ~ 0.4 arcmin. The reason behind the choice of this specific value is that for the purposes of this analysis, we employ the `MassiveNus` simulations, (described in Section 4.1.2) and we choose to match the resolution of the simulations. Given the noiseless κ maps from the simulations, we need to have an estimate of the noise to mimic the CFIS survey. For this, we need a redshift estimate for the source galaxies, or at least the galaxy number density distribution as a function of redshift, $n(z)$, for the lensing sample.

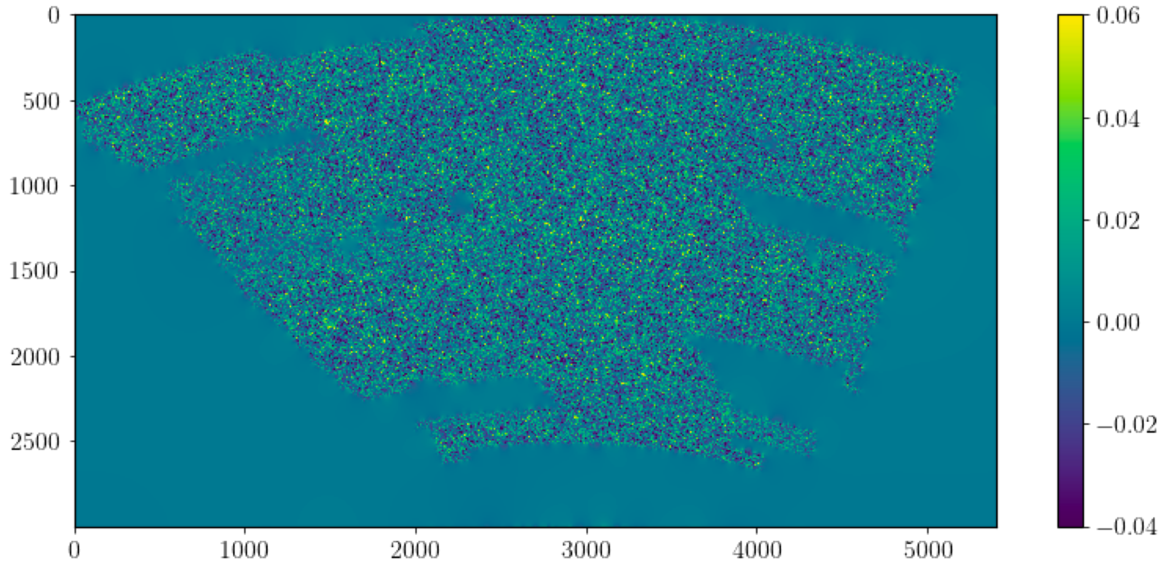


Figure 7.3: Illustration of the E-mode convergence map constructed for the CFIS P3 patch, corresponding to 249 deg^2. The data shown in this image have been provided by A. Guinot (Guinot 2020). The map is built through KS inversion, setting a pixel size of ~ 0.4 arcmin, to match the resolution of the κ maps from the `MassiveNus` simulations.

We employ the $n(z)$ built by Guinot 2020, shown in Figure 7.4. As of today, the galaxy number density distribution is not available based on CFIS data only, since only data on the r-band are available right now [1]. For this reason, the $n(z)$ is derived by matching CFIS galaxies with the ones detected in the W3 region of CFHTLens (Erben et al. 2013), that corresponds to P3. Around 700,000 galaxies have been matched with CFHTLenS by Guinot 2020. Then the $n(z)$ is reconstructed from the photometric redshift measurements in CFHTLenS, using the best-fit value z_{B}. This method provides only an approximation, but it gives a first estimate of the galaxy distribution.

The corresponding mean galaxy number density used in the analysis is $n_{\mathrm{gal}} = 7\,\mathrm{gal/arcmin}^2$. The shape noise is $\sigma_\epsilon = 0.44$.

---

[1] An ongoing effort to obtain redshift distributions with the so-called cluster-z technique as presented in Ménard et al. 2014 is currently pursued by the collaboration.

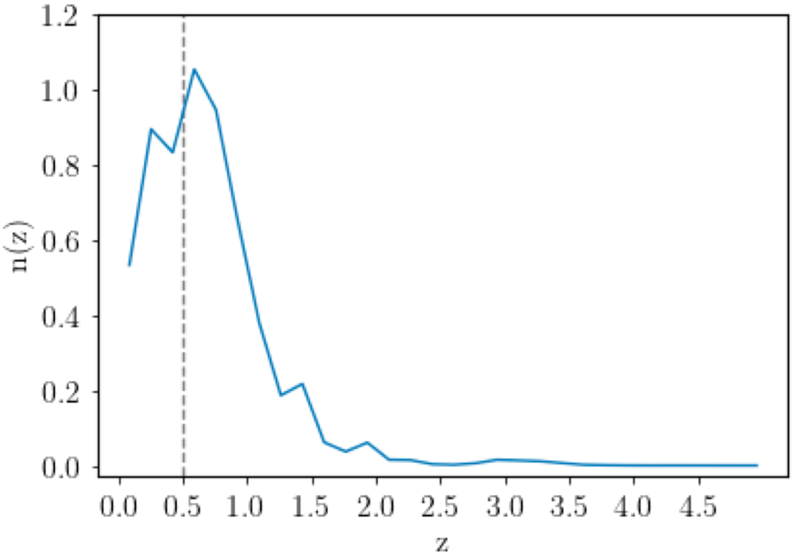


Figure 7.4: Galaxy number density distribution (blue line) obtained by matching the galaxies with the CFHTLenS-W3 patch by Guinot 2020.

7.2.2 Methodology

We perform a likelihood analysis by applying the pipeline I developed, introduced in Section 5.2. In this first attempt, we adapt the data to the available `MassiveNus` simulations, in terms of resolution and size of the maps. Ideally, one would do the contrary, in order to exploit the full information contained in the data. However, in the interest of performing a first experiment, and given the huge amount of maps from the simulations (100 cosmologies $\times$ 10000 realisations = 10^6 maps), we start by mainly adapting the data to the simulations, to get an idea of what some preliminary output would be, in this first application on real data. In order to do so, we cut 13 squared patches of size 512 $\times$ 512 from the original P3 convergence map of Figure 7.3. In this way, the maps from which we compute the peak counts that are use to build the theoretical predictions, and the maps from which we compute the peak counts for the data, have the same resolution (~ 0.4 arcmin) and the same size (512 $\times$ 512 pixels^2). Concerning the redshift for the sources, in chronological order we have acted as follows. As the redshift planes provided by the `MassiveNus` simulations are fixed at $z_\mathrm{s} = [0.5, 1.0, 1.5, 2.0, 2.5]$, we looked at the $n(z)$ for the data, shown in Figure 7.4. We noticed that the value $z = 0.5$ (dashed gray line in Figure 7.4) is actually very close to where the $n(z)$ peaks and decided to perform the analysis employing the convergence maps from the simulations at $z_\mathrm{s} = 0.5$. We have in a first moment performed the entire analysis in this setting, adding Gaussian noise with standard deviation

$$\sigma_{\rm n} = \frac{\langle \sigma_\epsilon \rangle}{\sqrt{2 n_{\rm gal} A_{\rm pix}}}, \tag{7.1}$$

to the noiseless simulated maps, with σ_ϵ and $n_{\rm gal}$ being respectively the shape noise and the mean galaxy number density defined in the previous section and with $A_{\rm pix} = 0.16\ {\rm arcmin}^2$ being the pixel area. At a later time, we realised we could try to adapt the source redshift of the simulations to the actual value at which the $n(z)$ peaks, which is $z = 0.65$. Therefore we have used the presence of convergence maps at $z_{\rm s} = 0.5$ and $z_{\rm s} = 1.0$ from the simulations and performed an interpolation for all 100 cosmologies and 10000 realisations and derived effective convergence maps at $z_{\rm s} = 0.65$. This is done by considering that, as shown in Equation 2.21 the convergence field as a function of θ is given by

$$\kappa(\boldsymbol{\theta}) = \int_0^{z_{\rm lim}} n(z) \kappa(\boldsymbol{\theta}, z) dz, \tag{7.2}$$

and that given that the available redshifts in our case are basically two Dirac delta functions, we can approximate Equation 7.2 for a discrete case to find our effective convergence map at $z_{\rm s} = 0.65$ as

$$\kappa^{\rm eff}_{z_{\rm s}=0.65} = \kappa_{z_{\rm s}=0.5} \lambda + \kappa_{z_{\rm s}=1.0}(1-\lambda), \tag{7.3}$$

where λ is a parameter that weights the convergence maps at a given redshift so that their combination leads to the effective convergence map at the desired redshift. The parameter λ is then obtained by asking that the mean redshift $\bar{z}$ is equal to 0.65, namely by asking that

$$\begin{aligned} \bar{z} &= \int n(z) z dz \qquad (7.4) \\ &= \int [\delta(z-0.5)\lambda + \delta(z-1)(1-\lambda)] z dz \qquad (7.5) \\ &= 0.5\lambda + 1(1-\lambda) = 0.65, \qquad (7.6) \end{aligned}$$

that leads to a value of $\lambda = 0.7$. By performing this interpolation, we obtain the set of effective κ maps for the simulation at $z_{\rm s} = 0.65$. We show in Figure 7.5 one realisation example of the output of the interpolation. The effective convergence map $\kappa^{\rm eff}_{z_{\rm s}=0.65}$ is in the central panel. Qualitatively, one can recognise by eye the same structures in the same positions as $z_{\rm s} = 0.5$ and $z_{\rm s} = 1.0$. In the bottom panel we compare the angular power spectra for the κ maps at $z_{\rm s} = 0.5$ and $z_{\rm s} = 1.0$ with our effective $\kappa^{\rm eff}_{z_{\rm s}=0.65}$ at $z_{\rm s} = 0.65$ as well as the peak counts. Given the difference between the summary statistics (in particular for the peak counts) at $z_{\rm s} = 0.5$ and $\kappa^{\rm eff}_{z_{\rm s}=0.65}$, we expect this difference to be propagated in the results of the inference analysis in the two settings. To further improve this interpolation, by trying for example to match the entire redshift distribution, we would need simulated convergence maps at more intermediate redshift steps, which are not provided by the employed simulations. As a more detailed follow-up study, it could be interesting to see how the constraints behave with simulations at $z_\mathrm{s} = 0.65 \pm \Delta z$. For this study however, we would again need to use other simulations.

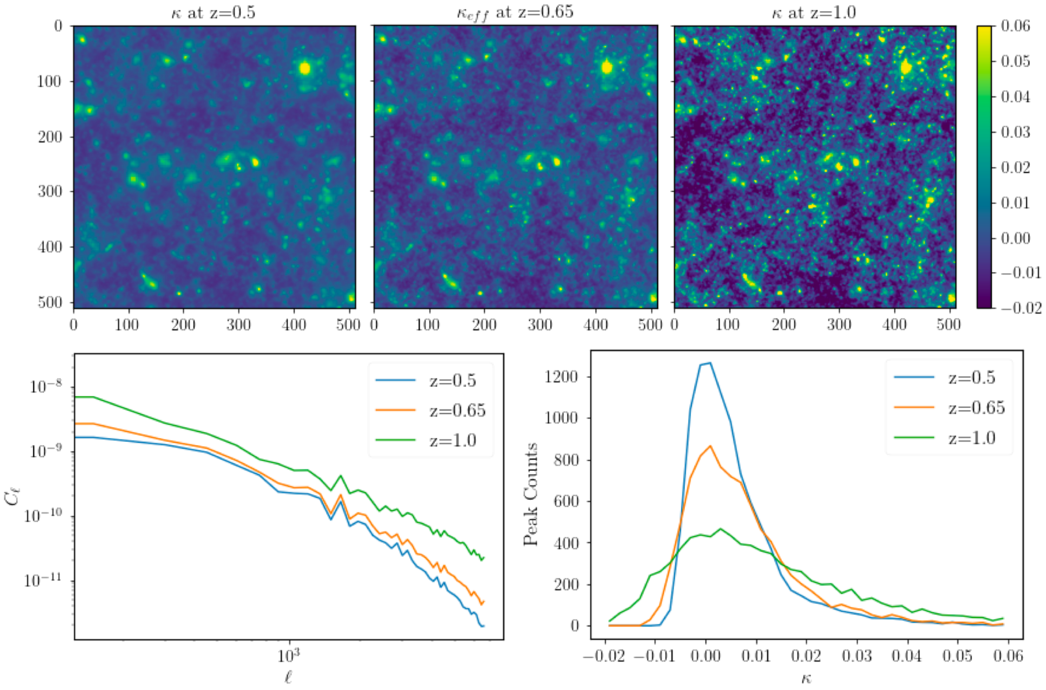


Figure 7.5: Illustration of one realisation example result of our interpolation to get the effective convergence map $\kappa^{\mathrm{eff}}_{z_\mathrm{s}=0.65}$ at $z_\mathrm{s} = 0.65$. **Top panel**: at the edges the maps from the simulations employed for the interpolation are shown, $z = 0.5$ (left) and $z = 1.0$ (right). The central panel shows the result of Equation 7.3. **Bottom panel**: angular power spectra (left), peak counts (right) for the different redshifts. Given the difference between the summary statistics (in particular for the peak counts) at $z_\mathrm{s} = 0.5$ and $\kappa^{\mathrm{eff}}_{z_\mathrm{s}=0.65}$, we expect this difference to be propagated in the results of the inference analysis in the two settings.

7.2.3 Peak counts

As a first result for the constraints on the parameters Ω_m and σ_8 has already been obtained by Guinot 2020 for the 3 × 2-point analysis, we focus on peak counts as summary statistics for our analysis as a first experiment to apply the pipeline. For both settings at $z_\mathrm{s} = 0.5$ and $z_\mathrm{s} = 0.65$, we consider 30 linearly spaced bins in S/N in the range $[-2, 6]$. The signal to noise ratio is defined as in the Gaussian filter case of Section 4.1, namely as the noisy κ map smoothed with Gaussian kernel, over the standard deviation of the noise. The size of the Gaussian kernel is 2 arcmin. To build the theoretical predictions we compute the peak counts for each realisation and each of the 100 cosmologies from the `MassiveNus` simulations and then interpolate through the Gaussian Process Regression (GPR) presented in Section 5.2. For the data, we compute the peak counts for the 13 patches extracted from the P3 patch and the average over them before inserting them in the likelihood. We then rescale the covariance matrix to take into account for the fact that we average over 13 patches. The peak counts are computed on a HPC cluster in parallel with the package `joblib` for each of the 10000 realisations and each of the 100 cosmologies.

7.3 Correction for baryonic effects

As we are dealing with real data but modelling the theoretical predictions with N-body simulations, we should take into account a variety of systematics to consider our analysis robust. For this first experiment we start by including in the analysis a correction for baryonic effects. As described in Section 2.2.3, baryonic effects (which are not described by gravity-only N-body simulations) have a significant impact on current and future weak lensing observations. Weiss et al. 2019 have shown that the strength of the baryon suppression on peak counts is strongly dependent on the scale of the Gaussian smoothing applied to the convergences maps. Specifically, they found that for a DES-like configuration, a smoothing scale of $\theta_{\rm ker} \sim 8$ arcmin or larger is sufficient to wash-out the baryonic effects so that the suppression signal becomes smaller than the statistical error and for a *Euclid*-like setting that a smoothing of at least $\theta_{\rm ker} \sim 16$ arcmin is required to reduce the baryon suppression signal enough so that it falls below the statistical error. If one considers smaller smoothing scales, a proper modelling of baryonic effects is necessary. We apply a smoothing scale of $\theta_{\rm ker} \sim 2$ arcmin to our convergence maps. Hence, to model the baryonic effects impact, as we get the theoretical prediction with the `MassiveNus` simulations, we apply corrections to the peak counts in our analysis following the results of Coulton et al. 2020a. To include baryonic effects, they consider a parameter called $A_{\rm baryon}$, which linearly interpolates the fractional effect of baryons between no baryonic effects ($A_{\rm baryon} = 0$) to the BAHAMAS high-AGN model $A_{\rm baryon} = 3$, with $A_{\rm baryon} = 2$ for the fiducial model, and $A_{\rm baryon} = 1$ for the low-AGN model. We use the peak counts for the three models (low-AGN, fiducial, high-AGN), kindly provided by the authors of Coulton et al. 2020a and compute the ratio between each baryon feedback model and the peak counts obtained by the DM-only simulations. The peak counts sets they provided are computed for 99 linearly spaced bins, in the range $S/N = [-6, 16]$. Since for our likelihood analysis we want to use 30 bins in the range $S/N = [-2, 6]$, we perform a 1-D interpolation with the python package `scipy`, to extract the peak counts in our S/N range of interest for the three baryon feedback models. We show in Figure 7.6 the fractional difference (left panel) and the ratio with respect to the dark matter only (right panel) peak counts due to different feedback strengths, from the arrays provided by the authors of Coulton et al. 2020a. The correction input that we use to get our peak counts accounting for the different baryonic feedback corresponds to the right panel. All three baryon feedback models imprint a deviation from the

DM only distribution, that increases with the strength of the baryonic effects. The counts for $S/N \simeq [-1, 3]$ are boosted, while being suppressed for $S/N \geq 3$.

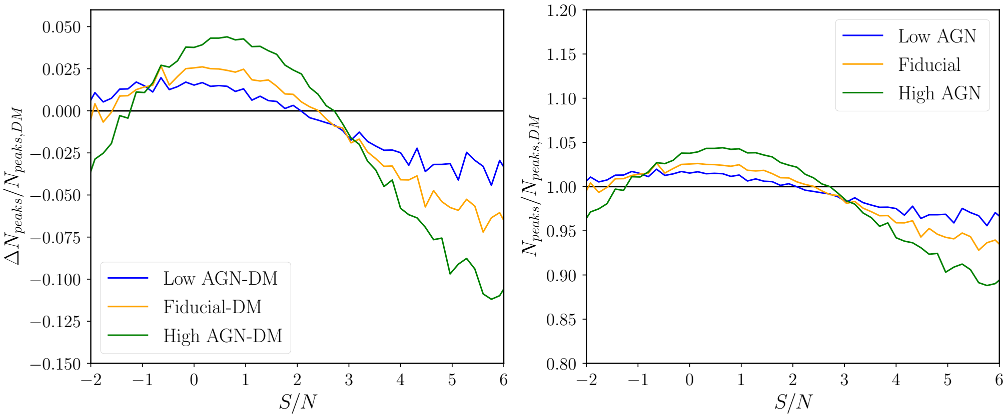


Figure 7.6: **Left panel**:Fractional difference between peak counts from Low-AGN (blue), Fiducial (yellow), High-AGN (green) models with respect to the peak counts computed on DM ony simulations. **Right panel**: Ratio between peak counts from Low-AGN (blue), Fiducial (yellow), High-AGN (green) models with respect to the peak counts computed on DM ony simulations. The arrays plotted in this figure have been kindly provided by the authors of Coulton et al. 2020a.

We run the analysis in the following settings:

1) for source redshift $z_{\rm s} = 0.5$: we run the likelihood analysis using as summary statistics a) DM-only peak counts, b) Low-AGN peak counts, c) Fiducial ($A_{\rm baryon} = 2$) peak counts, d) High-AGN peak counts using as data the peak counts computed on the 13 squared patches extracted from CFIS P3 patch, shown in Figure 7.3.

2) For source redshift $z_{\rm s} = 0.65$: we run the likelihood analysis using as summary statistics a) DM-only peak counts, b) Low-AGN peak counts, c) Fiducial ($A_{\rm baryon} = 2$) peak counts, d) High-AGN peak counts using as data the peak counts computed on the 13 squared patches extracted from CFIS P3 patch, shown in Figure 7.3

7.4 First experiment and discussion

We run a first experiment to get constraints on the sum of neutrino masses M_ν, the matter density parameter $\Omega_{\rm m}$ and the amplitude of the primordial power spectrum $A_{\rm s}$. We take into account the baryonic feedback for three different models, labeled as Low-AGN, fiducial and High-AGN. We assume a Gaussian likelihood

$$\log \mathcal{L}(\theta) = \frac{1}{2}(d - \mu(\theta))^T C^{-1}(d - \mu(\theta)), \tag{7.7}$$

where for the model $\mu(\theta)$ we employ the Gaussian Process Regressor trained on the 100 cosmological models from the `MassiveNus` simulations, for the data d the peak counts computed on the 13 patches of the CFIS P3 patch, and the covariance is computed employing the massless cosmology from the simulations and re-scaled with respect to the area of the 13 patches. We explore parameter space with MCMC using the python package `emcee`, with 120 walkers and 6500 steps. We verify the convergence of our chains with Gelman-Rubin (Gelman and Rubin 1992) and Geweke (Geweke 1991) statistics. We start by looking at the difference induced when employing maps at $z_\mathrm{s} = 0.65$ and maps at $z_\mathrm{s} = 0.5$, namely the constraints shown in Figure 7.7. In particular, for both the dark matter only simulations (left panel) and the fiducial baryonic effects model (right panel), when considering $z_\mathrm{s} = 0.65$ (dashed lines) there is a shift for the matter density parameter Ω_m and the amplitude of the primordial power spectrum A_s, towards smaller values. As already mention, the reason behind the change in source redshift was due to the fact that the `MassiveNus` simulations exist at fixed redshift $z_\mathrm{s} = 0.5$ and $z_\mathrm{s} = 1.0$ and not at $z_\mathrm{s} = 0.65$, for which we have derived some effective convergence maps to try to better match the peak in $n(z)$, shown in Figure 7.4. After ascertaining the difference between the two redshift settings, we consider the setting at $z_\mathrm{s} = 0.65$ the most reliable for our analysis, given the above discussion on the distribution of the sources. We then explore the impact of the different baryonic correction on the constraints. Specifically, by looking at Figure 7.8, we notice that the presence of baryonic effects induces a shift in the $(\Omega_\mathrm{m}, A_\mathrm{s})$ plane towards higher values of both Ω_m and A_s and this effect increases as the parameter A_baryon increases. The effect on the sum of neutrino masses M_ν is harder to detect, from the joint contours, while from the marginalised constraints in Figure 7.9, the Low-AGN and Fiducial correction seem to increase its value and the High-AGN to slightly decrease it, with respect to the case without correction. We show in Table 7.1 the maximum likelihood values for the case without baryonic correction (DM only), and the three baryonic models for the setting at $z_\mathrm{s} = 0.65$.

Table 7.1: Maximum likelihood values for the three parameters for the different settings (no correction, Low-AGN, Fiducial, High-AGN).

Model	M_ν	Ω_m	$10^9 A_\mathrm{s}$
DM only	$0.105^{+0.313}_{-0.043}$	$0.358^{+0.023}_{-0.028}$	$1.60^{+0.43}_{-0.30}$
Low-AGN	$0.21^{+0.20}_{-0.15}$	$0.375^{+0.036}_{-0.021}$	$2.05^{+0.57}_{-0.20}$
Hihg-AGN	$0.091^{+0.229}_{-0.029}$	$0.418^{+0.0091}_{-0.0449}$	$2.09^{+0.41}_{-0.22}$
Fiducial	$0.21^{+0.21}_{-0.15}$	$0.393^{+0.027}_{-0.024}$	$2.08^{+0.46}_{-0.21}$

In summary, we see that in all cases, the impact of taking into account or not the baryonic feedback is pretty important. We notice that Ω_m and A_s are shifted to higher value and M_ν becomes constrained narrower than the prior. The three baryonic corrections yield

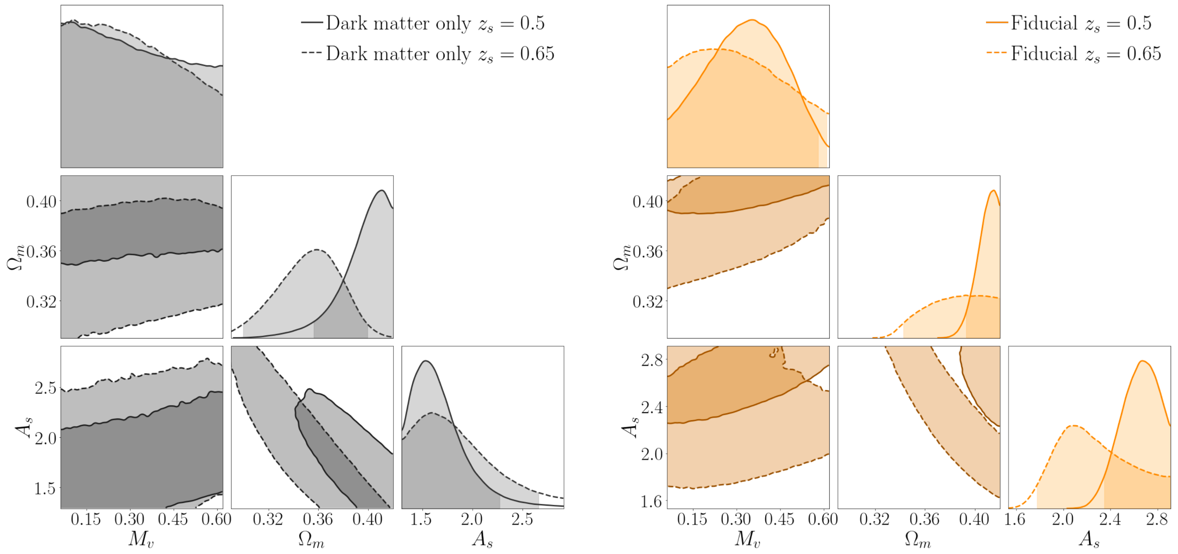


Figure 7.7: Constraints on M_ν, Ω_m and A_s using as data the CFIS P3 patch. **Left panel**: constraints obtained with DM-only peak counts for source redshift $z_\mathrm{s} = 0.5$ (continuous line) and for source redshift $z_\mathrm{s} = 0.65$ (dashed line). **Right panel**: same as left panel but for the baryonic effects fiducial model.

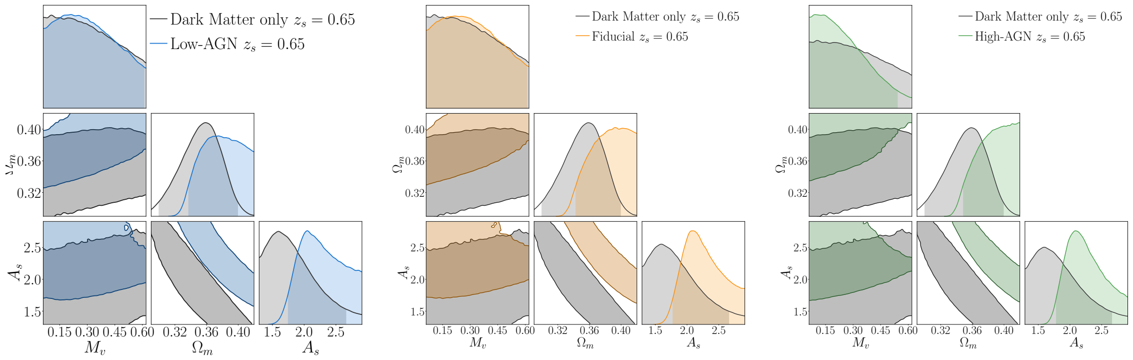


Figure 7.8: Constraints on M_ν, Ω_m and A_s using as data the CFIS P3 patch. **Left panel**: comparison between constraints with no baryonic effects (black contours) and including baryonic effects with the Low AGN model (blue contours, $A_\mathrm{baryon} = 1$). **Central panel**: comparison between constraints with no baryonic effects (black contours) and including baryonic effects with the Fiducial model (orange contours, $A_\mathrm{baryon} = 2$). **Right panel**: comparison between constraints with no baryonic effects (black contours) and including baryonic effects with the High AGN model (green contours, $A_\mathrm{baryon} = 3$).

to constraints which are pretty close to each other. Thus, even if we do not know how important is the baryonic feedback on the data, we can conclude that the specific flavour of

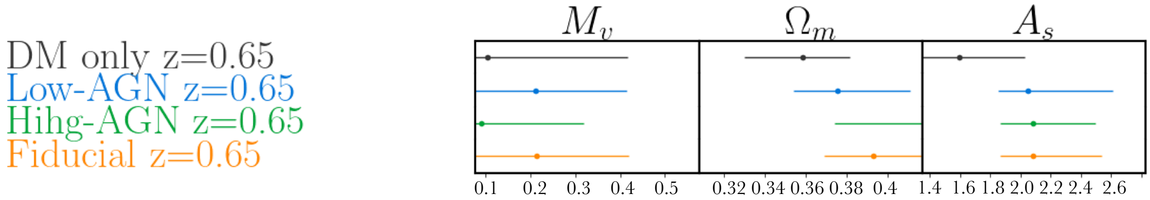


Figure 7.9: Maximum likelihood values and marginalised errors for the 3 parameters.

baryonic feedback (if Low-AGN, Fiducial or High-AGN) is less important than the difference between baryonic feedback and pure dark-matter models. With the baryonic feedback there is less peaks in the model. Hence, this has to be compensated with a higher Ω_{m} because the data does not change.

7.5 Ongoing work and future prospects

The aim of this chapter was to show a first experiment of application of the pipeline presented in Chapter 4 and Chapter 5 using peak counts as summary statistics on real data. We are aware that the preliminary results shown so far are very limited by the fact that the only systematics that we are including are baryonic effects in the simulations, while other crucial systematics such as intrinsic alignment, multiplicative shear bias, mean redshift bias, boost factor (previously discussed in Section 2.2.3) have been neglected. Indeed, an ongoing effort with the collaborators of this project is ongoing, in trying to model each one of these effects first independently and then combined with the baryonic effects. For instance, a further step has been undertaken to extend this work by including multiplicative shear bias using *metacalibration* (Huff and Mandelbaum 2017) with a residual bias Δm and explore its impact on the constraints, independently from and combined with baryonic effects. A first result of this second step is shown in Figure 7.10: first, the two systematics effects (multiplicative shear bias and baryonic effects) are considered separately to have an idea of how the single effects act with respect to the constraints where no correction is applied to the simulations. This is shown in the left panel of Figure 7.10. We can see how the baryonic correction tends to shift Ω_{m} and A_{s} to higher values whereas the redshift and metacalilbration bias tends to shift them to lower values. Then the two are combined, to see how they act together with respect to the constraints where no correction is applied. On the right panel of Figure 7.10, we can see the constraints obtained for data from the local calibration at 1 deg^2 with $\Delta m = 0.007$ (Guinot et al. in prep.) for the simulations at $z_{\mathrm{s}} = 0.65$ corrected by the baryonic feedback with the fiducial model $A_{\mathrm{baryon}} = 2$. We would like to stress that as of today the exact value for Δm is still unknown for CFIS. Therefore we have used a preliminary value for it (Guinot et al. in prep.) as our best guess. Next steps will be then to include in the analysis further systematics that have been neglected, with the final goal of quantifying the impact of each of them and of all of them combined on the final constraints and quantify the robustness and constraining power of peak counts in this context. This work is ongoing, while a straightforward future work that will be interesting to carry out is to test the starlet ℓ_1-norm, introduced in Section 5.1.5, that has been proven powerful in an ideal setting, but still has not been tested on real data.

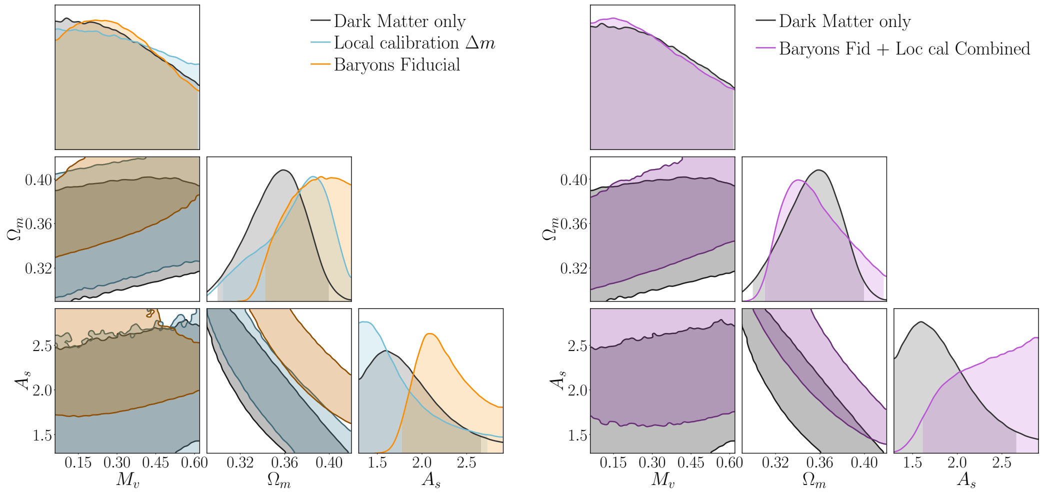


Figure 7.10: **Left panel**: constraints without corrections (black contours), constraints with local calibration and Δm uncertainty (light blue contours), constraints with local calibration and Δm uncertainty (light blue contours), constraints with fiducial baryonic correction (orange contours). **Right panel**: constraints without corrections (black contours), constraints with local calibration and Δm uncertainty combined with fiducial baryonic correction (violet contours).

CHAPTER 8

CONCLUSIONS AND PERSPECTIVES

The main goal of this thesis has been the investigation and development of statistical methods to extract information from cosmological data coming from next generation galaxy surveys like *Euclid*. To address this topic, I first focused on a study where we investigated the impact on the parameter constraints of multi-scale filtering techniques at the level of the noisy data. Comparing the performance of a wavelet starlet filter and a concatenation of Gaussian filters with the state of the art techniques when using weak lensing peak counts as summary statistics, we found that in an ideal setting, both multi-scale techniques significantly outperform the state of the art. For instance, for the set of parameters $M_\nu, \Omega_{\mathrm{m}}$ and A_{s}, we found that starlet peak counts provide an improvement of respectively 63%, 40% and 72% with respect to the standard power spectrum. As second result of this first study, we found that the power spectrum does not seem to bring any additional information to the constraints when multi-scale peak counts are combined with the power spectrum. This is interesting and quite new outcome, as usually when second order statistics and higher order statistics are combined, an improvement on the constraining power is typically found for single-scale statistics, as each statistics bring complementary information to the other. Lastly, when exploring the effect of the filter choice on the different steps of the analysis, we noticed that the covariance matrix for peak counts when using a starlet filter tends to be more diagonal than the one obtained when using multi-Gaussian filter. Intrigued by this feature, we run the same analysis using only the diagonal elements of the covariance matrix, and basically recovered the same constraints as when the full covariance is used. This can be convenient for situations (actually not rare in cosmology) where the covariance matrix is difficult to invert, or there is interest in speeding up the analysis. In a broader sense, this analysis proved some advantages of the starlet transform to optimise the information gain in cosmological analysis, confirmed the potential of higher order statistics in improving the constraining power and showed how the combination of high order statistics in a multi-scale framework can bring a significant improvement in the constraining power of future surveys like *Euclid*. The results of this first work also led to the construction of a pipeline skeleton for multi-scale peak counts, having as initial input weak lensing convergence maps and as output constraints on cosmological parameters in an ideal simulations-based setting. Then, trying to preserve the advantages of a multi-scale approach in a higher order statistics framework, we wondered if our choice of peak counts as higher order statistics was limiting in some sense. Indeed, as discussed in Section 2.3, a vast literature of different type of higher order statistics is present and it was proven that some combination of these statistics can improve even more the constraining power. For example, using the same sets of simulations that we used, Coulton et al. 2020b found that the combination of peak counts (local maxima in convergence maps) and local minima counts, further tighten the constraints. Inspired by this, we wondered: what happens if, instead of selecting only the local minima and local maxima in the map, we consider the information encoded in all pixels, while preserving the starlet multi-scale approach that has been proven advantageous? This question gave origin to my second study, where we have introduced a new multi-scale higher order summary statistics: the *starlet* ℓ_1-norm. This statistics is built in such a way that starting from a starlet decomposition of the input data, it accounts for the information present at different wavelet scales for *all* pixels in the map. Hence, it automatically contains the information encoded in both peaks and minima. This is confirmed by our outcome on its performance in constraining parameters: the *starlet* ℓ_1-norm outperforms the state of the art combination of peaks and voids and it avoids at the same time the problem of defining what is a void (topic on which there is still no unanimous agreement). This second work also led to an extension of the public CosmoStat python package `lenspack`, that now contains the implementation of this new statistics to perform cosmological analysis.

The practical contributions of these first two studies can be summarised has follows:

1) a first paper (Ajani et al. 2020) quantifying the impact of different multi-scale filtering on weak lensing peak counts and between different statistics

2) the construction of a pipeline skeleton for multi-scale peak counts, having as initial input weak lensing convergence maps and as output constraints on cosmological parameters in an ideal simulations-based setting

3) the introduction of a new multi-scale higher order statistics, defined in a second paper (Ajani et al. 2021)

4) an extension of the public CosmoStat python package `lenspack`

5) tomographic MCMC forecasts mimicking a *Euclid*-like shape noise for the parameters $M_\nu, \Omega_\mathrm{m}, A_\mathrm{s}$ obtained with many different statistics such as the power spectrum, single-scale peak counts, multi-scale Gaussian and starlet peaks, voids and starlet ℓ_1-norm.

These results show that the multi-scale higher order statistics approach we developed represents a very promising tool to reach high levels of precision in constraining cosmological parameters. However, to serve as a robust probe for precision cosmology, the statistical power alone is not sufficient and it is very important to guarantee robustness with respect to systematic errors. Furthermore, keeping in mind that the final aim is to maximise the information coming from surveys like *Euclid*, we also want to apply these findings not only to weak lensing, but also to galaxy clustering to fully exploit the information gain deriving from their combination. These reflections led to the ongoing work and preliminary results of Chapter 6 and Chapter 7 where we aim to extend our method to photometric galaxy clustering alone and combined with weak lensing and to real data from the Canada France Imaging Survey (CFIS). This, from a practical point of view, translates in understanding in which parts and in which ways my multi-scale higher order statistics pipeline needs to be extended. First of all, for the extension to photometric galaxy clustering, I had to use different simulations than the `MassiveNus` on which the pipeline was initially built. This change of simulations, naturally led to two extensions simultaneously. The first concerns the different parameter space, both in terms of specific parameters (w_0 and h were actually fixed in the `MassiveNus`) and in terms of dimensionality (the parameters space for the *cosmo*-SLICS is 4-dimensional). The second extension is relative to the input, that is now given by the galaxy catalogues directly instead of the convergence maps. What one deals with in real data analysis is actually a galaxy catalogue, and now the pipeline directly takes as input the catalogue, builds a pixelised shear maps and with Kaiser-Squires inversion gets the convergence maps. Lastly, given a lens catalogue instead of a source one, it is now possible to build galaxy density maps with the idea of performing a photometric galaxy clustering analysis. Running exploratory experiments, we obtained some preliminary results for combined second order statistics (the power spectrum) for both weak lensing and galaxy clustering. As expected, we find that the combination of the two improves the constraints. However, for photometric galaxy clustering, we are applying some approximations in how we estimate the noise and for the galaxy bias. The investigation on these points is ongoing and we expect the contours for GCph to inflate once using real data, where the galaxy bias is largely unknown. We also apply starlet peaks on the galaxy maps built from the lenses catalogue and find improvement for the majority of the parameters in the constraining power. However, a slight shift in the w_0 constraints arises and we are currently trying to understand whether this is caused by approximations made in the procedure or if its origin has a different nature. Lastly, a natural test for the robustness of the pipeline is to adapt our procedure to real data. Fortunately, about $1600\ \mathrm{deg}^2$ of the CFIS data have been already pre-processed and this has enabled a first test of the pipeline on a part of these avaiable data. In Chapter 7 some preliminary work in this sense is outlined. In its first implementation, the pipeline didn't include any systematics effects. We therefore profit of the availability of CFIS data and start including the systematics one by one, to quantify their effects first separately and eventually jointly. We start with the inclusion of baryonic effects, and provide a first qualitative description of their impact in three different scenarios for a peak counts analysis. Now, with the final aim of including as many systematics as possible, we are focusing on a second systematics, namely the multiplicative shear bias. To calibrate the shear, we use *metacalibration* and test the impact of local calibration. Indeed, as the multiplicative shear bias depends on the local density (Hoekstra, Viola, and Herbonnet 2017), a local calibration can be more precise and consequently lead to a smaller multiplicative bias. However, as metacalibration is not perfect, there is a residual bias Δm that needs to be considered and whose value is a current object of research within the CFIS collaboration. The results obtained so far on CFIS data are still very preliminary. However, we start having an idea of some qualitative behaviour, such as that with these data and simulations, the more constrained of the parameters seems to be $\Omega_{\rm m}$. We also find that local calibration with a $\Delta m < 0$ shifts this parameter to higher values compared to baryonic corrections. On the other hand, we can conclude that the specific flavour of baryonic feedback is not so relevant, with respect to the difference between baryonic feedback and pure dark-matter models. Concerning the local calibration, it is sure that a local one is preferred over a global one, as long as the error and standard deviation are small, because it will account for local effects of the catalogue. Finally, we need a more precise value of Δm as changes in this value can lead to large differences on the parameters. Some questions presented above still need an answer, as for example, what the impact of combining multi-scale high order statistics on combined probe is, or what the impact of many systematics effects that were neglected is. However, these unsolved points are already structured in ongoing projects whose outcomes hopefully will provide us with the answers that we are looking for. Besides the above discussed projects, many other interesting questions that are still pending can be addressed in this context. For example, by paying attention to some assumptions made throughout this work, we could focus on the fact that we derive our input convergence maps with a standard Kaiser-Squires inversion. However, as shown in Section 2.1.3 and Section 3.6, new powerful approaches as `DeepMass` (Jeffrey et al. 2018) or `MCALens` (Starck et al. 2021) that have overtaken the drawbacks of standard inversion techniques are available to be used. It is then worth exploring how the combination of these mass mapping approaches with multi-scale higher order statistics can improve the analysis. On a different level, many other questions concerning the considered parameters can arise. For instance, using the `MassiveNus` simulations, we provided forecasts and constraints on the sum of neutrino masses M_ν. However, strong observational degeneracies between the effects of modified $f(R)$ gravity theories and massive neutrinos have been highlighted on structure formation processes over a wide range of scales and redshifts, both in the linear and in the non-linear regime (Baldi et al. 2014). These degeneracies can be broken with the use of high order statistics as was proved by Peel, Pettorino, Giocoli, et al. 2018. However, what the optimal statistics to do so is, still remains an open question, and testing whether the *starlet* ℓ_1-norm could be a good candidate for this task could be interesting and in practice doable using the DUSTGRAIN-*pathfinder* simulations (Giocoli, Baldi, and Moscardini 2018) that include $f(R)$ models. This work is also currently in progress.

A large number of open questions are of course still present. In the context of this work, a summary of what needs to be addressed can be summed up in the fact that next generation cosmological surveys such as *Euclid*, DESI, the Vera Rubin Observatory Legacy Survey of

Space and Time, the Nancy Grace Roman Space Telescope, will provide us with a huge volume of data, characterised by a high resolution up to the very small scales. The main goal is then to fully exploit these small scales up to $k \sim 10h\,\mathrm{Mpc}^{-1}$, as they contain the cosmological information needed to try to answers the big questions about the dark components of the universe outlined at the beginning. A huge collective effort in this context is ongoing. For example, the current work carried on by the Interscience Taskforce for Likelihood development (IST:L) in *Euclid* to which I had the chance to contribute, will provide the tools that we will need to perform a full cosmological analysis including weak lensing, galaxy clustering and their cross-correlation. Concerning the higher order statistics, it is clear that given the large amount of different statistics already available now, one crucial challenge at this stage is also to compare what it is already available and find the optimal statistics in terms of constraining power and robustness to systematics. In this sense, a collective effort is as well ongoing within the higher order weak lensing statistics (HOWLS) work package within the *Euclid* collaboration. Besides the preparation for future surveys, this is already a very exciting time to pursue these investigations, as data from current surveys as KiDS and DES are already available and have very recently provided the most precise constraints on the ΛCDM and wCDM models to date from weak lensing probes (DES Collaboration, Abbott, et al. 2021).

Bibliography

Abazajian, K.N. et al. (2015). "Neutrino physics from the cosmic microwave background and large scale structure". In: *Astroparticle Physics* 63. Dark Energy and CMB, pp. 66–80. ISSN: 0927-6505. DOI: https://doi.org/10.1016/j.astropartphys.2014.05.014. URL: https://www.sciencedirect.com/science/article/pii/S092765051400084X.

Ahn, Christopher P. et al. (2012). "THE NINTH DATA RELEASE OF THE SLOAN DIGITAL SKY SURVEY: FIRST SPECTROSCOPIC DATA FROM THE SDSS-III BARYON OSCILLATION SPECTROSCOPIC SURVEY". In: *The Astrophysical Journal Supplement Series* 203.2, p. 21. DOI: 10.1088/0067-0049/203/2/21. URL: https://doi.org/10.1088/0067-0049/203/2/21.

Ajani, V. et al. (2020). "Constraining neutrino masses with weak-lensing multiscale peak counts". In: *Phys. Rev. D* 102 (10), p. 103531. DOI: 10.1103/PhysRevD.102.103531. URL: https://link.aps.org/doi/10.1103/PhysRevD.102.103531.

Ajani, V. et al. (2021). "Starlet l1-norm for weak lensing cosmology". In: *A&A* 645 (10), p. L11. DOI: 10.1051/0004-6361/202039988. URL: https://doi.org/10.1051/0004-6361/202039988.

Ajani, V. et al. (in prep.[a]). "Combined multi-scale higher order statistics for weak lensing and photometric galaxy clustering".

Ajani, V. et al. (in prep.[b]). "Cosmology with Peak counts for CFIS".

Akeret, Joël et al. (2013). "CosmoHammer: Cosmological parameter estimation with the MCMC Hammer". In: *Astronomy and Computing* 2, pp. 27–39. ISSN: 2213-1337. DOI: https://doi.org/10.1016/j.ascom.2013.06.003. URL: https://www.sciencedirect.com/science/article/pii/S221313371300022X.

Allys, E. et al. (2020). "New interpretable statistics for large-scale structure analysis and generation". In: *Phys. Rev. D* 102 (10), p. 103506. DOI: 10.1103/PhysRevD.102.103506. URL: https://link.aps.org/doi/10.1103/PhysRevD.102.103506.

Alsing, Justin and Benjamin Wandelt (2018). "Generalized massive optimal data compression". In: *Mon. Not. Roy. Astron. Soc.* 476.1, pp. L60–L64. DOI: 10.1093/mnrasl/sly029. arXiv: 1712.00012 [astro-ph.CO].

Alsing, Justin, Benjamin Wandelt, and Stephen Feeney (Mar. 2018). "Massive optimal data compression and density estimation for scalable, likelihood-free inference in cosmology". In: *Monthly Notices of the Royal Astronomical Society* 477.3, pp. 2874–2885. ISSN: 0035-8711. DOI: 10.1093/mnras/sty819. eprint: https://academic.

oup.com/mnras/article-pdf/477/3/2874/24802996/sty819.pdf. URL: https://doi.org/10.1093/mnras/sty819.

Alsing, Justin et al. (July 2019). "Fast likelihood-free cosmology with neural density estimators and active learning". In: *monthly Notices of the Royal Astronomical Society* 488.3, pp. 4440–4458. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1960. eprint: https://academic.oup.com/mnras/article-pdf/488/3/4440/29113037/stz1960.pdf. URL: https://doi.org/10.1093/mnras/stz1960.

Amendola, Luca et al. (2018). "Cosmology and fundamental physics with the Euclid satellite". In: *Living Reviews in Relativity* 21.1. ISSN: 1433-8351. DOI: 10.1007/s41114-017-0010-3. URL: http://dx.doi.org/10.1007/s41114-017-0010-3.

Amon, A. et al. (2021). *Dark Energy Survey Year 3 Results: Cosmology from Cosmic Shear and Robustness to Data Calibration.* arXiv: 2105.13543 [astro-ph.CO].

Asgari, Marika et al. (2021). "KiDS-1000 cosmology: Cosmic shear constraints and comparison between two point statistics". In: *A&A* 645, A104. DOI: 10.1051/0004-6361/202039070. URL: https://doi.org/10.1051/0004-6361/202039070.

Bacon, D. J. et al. (Jan. 2006). "Weak gravitational flexion". In: *monthly Notices of the Royal Astronomical Society* 365.2, pp. 414–428. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2005.09624.x. eprint: https://academic.oup.com/mnras/article-pdf/365/2/414/2978831/365-2-414.pdf. URL: https://doi.org/10.1111/j.1365-2966.2005.09624.x.

Bacon, David J., Alexandre R. Refregier, and Richard S. Ellis (Oct. 2000). "Detection of weak gravitational lensing by large-scale structure". In: *monthly Notices of the Royal Astronomical Society* 318.2, pp. 625–640. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2000.03851.x. eprint: https://academic.oup.com/mnras/article-pdf/318/2/625/18641910/318-2-625.pdf. URL: https://doi.org/10.1046/j.1365-8711.2000.03851.x.

Baldi, Marco et al. (Mar. 2014). "Cosmic degeneracies – I. Joint N-body simulations of modified gravity and massive neutrinos". In: *Monthly Notices of the Royal Astronomical Society* 440.1, pp. 75–88. ISSN: 0035-8711. DOI: 10.1093/mnras/stu259. eprint: https://academic.oup.com/mnras/article-pdf/440/1/75/9380024/stu259.pdf. URL: https://doi.org/10.1093/mnras/stu259.

Bardeen, J. M. et al. (1986). "The Statistics of Peaks of Gaussian Random Fields". In: *ApJ* 304, p. 15. DOI: 10.1086/164143.

Barreira, Alexandre et al. (2018). “Accurate cosmic shear errors: do we need ensembles of simulations?” In: *Journal of Cosmology and Astroparticle Physics* 2018.10, pp. 053–053. DOI: `10.1088/1475-7516/2018/10/053`. URL: `https://doi.org/10.1088/1475-7516/2018/10/053`.

Bartelmann, Matthias and Peter Schneider (2001). “Weak gravitational lensing”. In: *Physics Reports* 340.4, pp. 291–472. ISSN: 0370-1573. DOI: `https://doi.org/10.1016/S0370-1573(00)00082-X`. URL: `https://www.sciencedirect.com/science/article/pii/S037015730000082X`.

Bartelmann, M., Perrotta, F., and Baccigalupi, C. (2002). “Halo concentrations and weak-lensing number counts in dark energy cosmologies”. In: *A&A* 396.1, pp. 21–30. DOI: `10.1051/0004-6361:20021417`. URL: `https://doi.org/10.1051/0004-6361:20021417`.

Bate, James et al. (Nov. 2019). “When galaxies align: intrinsic alignments of the progenitors of elliptical galaxies in the Horizon-AGN simulation”. In: *monthly Notices of the Royal Astronomical Society* 491.3, pp. 4057–4068. ISSN: 0035-8711. DOI: `10.1093/mnras/stz3166`. eprint: `https://academic.oup.com/mnras/article-pdf/491/3/4057/31553821/stz3166.pdf`. URL: `https://doi.org/10.1093/mnras/stz3166`.

Baumann, Daniel (2012). *TASI Lectures on Inflation*. arXiv: `0907.5424 [hep-th]`.

Bernardeau, F., L. van Waerbeke, and Y. Mellier (1997). “Weak lensing statistics as a probe of {OMEGA} and power spectrum.” In: *A&A* 322, pp. 1–18. arXiv: `astro-ph/9609122 [astro-ph]`.

Bernardeau, F. et al. (2002). “Large-scale structure of the Universe and cosmological perturbation theory”. In: *Physics Reports* 367.1-3, 1–248. ISSN: 0370-1573. DOI: `10.1016/s0370-1573(02)00135-7`. URL: `http://dx.doi.org/10.1016/S0370-1573(02)00135-7`.

Bernardeau, Francis, Takahiro Nishimichi, and Atsushi Taruya (Oct. 2014). “Cosmic shear full nulling: sorting out dynamics, geometry and systematics”. In: *monthly Notices of the Royal Astronomical Society* 445.2, pp. 1526–1537. ISSN: 0035-8711. DOI: `10.1093/mnras/stu1861`. eprint: `https://academic.oup.com/mnras/article-pdf/445/2/1526/18198076/stu1861.pdf`. URL: `https://doi.org/10.1093/mnras/stu1861`.

Blazek, Jonathan, Zvonimir Vlah, and Uroš Seljak (2015). “Tidal alignment of galaxies”. In: *Journal of Cosmology and Astroparticle Physics* 2015.08, pp. 015–015. DOI: `10.1088/1475-7516/2015/08/015`. URL: `https://doi.org/10.1088/1475-7516/2015/08/015`.

Bolzonella, M., J. M. Miralles, and R. Pelló (Nov. 2000). "Photometric redshifts based on standard SED fitting procedures". In: *A&A* 363, pp. 476–492. arXiv: astro-ph/0003380 [astro-ph].

Bond, J. R. and G. Efstathiou (1987). "The statistics of cosmic background radiation fluctuations". In: *MNRAS* 226, pp. 655–687. DOI: 10.1093/mnras/226.3.655.

Bos, E. G. Patrick et al. (Oct. 2012). "The darkness that shaped the void: dark energy and cosmic voids". In: *Monthly Notices of the Royal Astronomical Society* 426.1, pp. 440–461. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2012.21478.x. eprint: https://academic.oup.com/mnras/article-pdf/426/1/440/3273501/426-1-440.pdf. URL: https://doi.org/10.1111/j.1365-2966.2012.21478.x.

Boyle, Aoife et al. (2020). *Nuw CDM cosmology from the weak lensing convergence PDF.* arXiv: 2012.07771 [astro-ph.CO].

Bridle, Sarah and Lindsay King (2007). "Dark energy constraints from cosmic shear power spectra: impact of intrinsic alignments on photometric redshift requirements". In: *New Journal of Physics* 9.12, pp. 444–444. DOI: 10.1088/1367-2630/9/12/444. URL: https://doi.org/10.1088/1367-2630/9/12/444.

Bucher, Martin, Benjamin Racine, and Bartjan van Tent (2016). "The binned bispectrum estimator: template-based and non-parametric CMB non-Gaussianity searches". In: *Journal of Cosmology and Astroparticle Physics* 2016.05, pp. 055–055. DOI: 10.1088/1475-7516/2016/05/055. URL: https://doi.org/10.1088/1475-7516/2016/05/055.

Burger, Pierre et al. (2021). *A revised density split statistic model for general filters.* arXiv: 2106.13214 [astro-ph.CO].

Cardone, V. F. et al. (Feb. 2013). "Weak lensing peak count as a probe of f(R) theories". In: *monthly Notices of the Royal Astronomical Society* 430.4, pp. 2896–2909. ISSN: 0035-8711. DOI: 10.1093/mnras/stt084. eprint: https://academic.oup.com/mnras/article-pdf/430/4/2896/3845870/stt084.pdf. URL: https://doi.org/10.1093/mnras/stt084.

Carron, J. (2013). "On the assumption of Gaussianity for cosmological two-point statistics and parameter dependent covariance matrices". In: *A&A* 551, A88. DOI: 10.1051/0004-6361/201220538. URL: https://doi.org/10.1051/0004-6361/201220538.

Casas, Santiago et al. (2017). "Linear and non-linear Modified Gravity forecasts with future surveys". In: *Physics of the Dark Universe* 18, pp. 73–104. DOI: 10.1016/j.dark.2017.09.009. arXiv: 1703.01271 [astro-ph.CO].

Catelan, Paolo, Marc Kamionkowski, and Roger D. Blandford (2001). "Intrinsic and extrinsic galaxy alignment". In: *monthly Notices of the Royal Astronomical Society* 320.1, pp. L7–

L13. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2001.04105.x. eprint: https://academic.oup.com/mnras/article-pdf/320/1/L7/3878727/320-1-L7.pdf. URL: https://doi.org/10.1046/j.1365-8711.2001.04105.x.

CFIS coll. (2017). *Canada-France Imaging Survey*. URL: https://www.cfht.hawaii.edu/Science/CFIS/.

Chan, Kwan Chuen and Linda Blot (2017). "Assessment of the information content of the power spectrum and bispectrum". In: *Phys. Rev. D* 96.2, 023528, p. 023528. DOI: 10.1103/PhysRevD.96.023528. arXiv: 1610.06585 [astro-ph.CO].

Chang et al. (2018). "Dark Energy Survey Year 1 results: curved-sky weak lensing mass map". In: *MNRAS* 475.3, pp. 3165–3190. DOI: 10.1093/mnras/stx3363. arXiv: 1708.01535 [astro-ph.CO].

Cheng, Sihao and Brice Menard (2021). *Weak lensing scattering transform: dark energy and neutrino mass sensitivity*. arXiv: 2103.09247 [astro-ph.CO].

Cheng, Sihao et al. (2020). "A new approach to observational cosmology using the scattering transform". In: *MNRAS*. DOI: 10.1093/mnras/staa3165.

Chisari, Nora Elisa et al. (2019). "Core Cosmology Library: Precision Cosmological Predictions for LSST". In: *The Astrophysical Journal Supplement Series* 242.1, p. 2. DOI: 10.3847/1538-4365/ab1658. URL: https://doi.org/10.3847/1538-4365/ab1658.

Clowe, Douglas et al. (2006). "A Direct Empirical Proof of the Existence of Dark Matter". In: *The Astrophysical Journal* 648.2, pp. L109–L113. DOI: 10.1086/508162. URL: https://doi.org/10.1086/508162.

Codis, S. et al. (Mar. 2015). "Intrinsic alignment of simulated galaxies in the cosmic web: implications for weak lensing surveys". In: *monthly Notices of the Royal Astronomical Society* 448.4, pp. 3391–3404. ISSN: 0035-8711. DOI: 10.1093/mnras/stv231. eprint: https://academic.oup.com/mnras/article-pdf/448/4/3391/2957600/stv231.pdf. URL: https://doi.org/10.1093/mnras/stv231.

Colberg, Jörg M., Frazer Pearce, Caroline Foster, et al. (2008). In: *MNRAS* 387.2, pp. 933–944. DOI: 10.1111/j.1365-2966.2008.13307.x. arXiv: 0803.0918 [astro-ph].

Combes, Francoise (2021). *Science with SKA*. arXiv: 2107.03915 [astro-ph.CO].

Copeland, David, Andy Taylor, and Alex Hall (2018). English. In: *monthly Notices of the Royal Astronomical Society* 480.2, pp. 2247–2265. ISSN: 0035-8711. DOI: 10.1093/MNRAS/STY2001.

Coulton, William R. et al. (2019). In: *Journal of Cosmology and Astroparticle Physics* 2019.05, pp. 043–043. DOI: `10.1088/1475-7516/2019/05/043`. URL: `https://doi.org/10.1088`.

Coulton, William R et al. (Apr. 2020a). In: *monthly Notices of the Royal Astronomical Society* 495.3, pp. 2531–2542. ISSN: 0035-8711. DOI: `10.1093/mnras/staa1098`. URL: `https://doi.org/10.1093/mnras/staa1098`.

— (Apr. 2020b). "Weak lensing minima and peaks: Cosmological constraints and the impact of baryons". In: *monthly Notices of the Royal Astronomical Society* 495.3, pp. 2531–2542. ISSN: 0035-8711. DOI: `10.1093/mnras/staa1098`. URL: `https://doi.org/10.1093/mnras/staa1098`.

Daalen, Marcel P. van et al. (Aug. 2011). "The effects of galaxy formation on the matter power spectrum: a challenge for precision cosmology". In: *monthly Notices of the Royal Astronomical Society* 415.4, pp. 3649–3665. ISSN: 0035-8711. DOI: `10.1111/j.1365-2966.2011.18981.x`. eprint: `https://academic.oup.com/mnras/article-pdf/415/4/3649/17328731/mnras0415-3649.pdf`. URL: `https://doi.org/10.1111/j.1365-2966.2011.18981.x`.

Daubechies, Ingrid (1992). *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics. DOI: `10.1137/1.9781611970104`. eprint: `https://epubs.siam.org/doi/pdf/10.1137/1.9781611970104`. URL: `https://epubs.siam.org/doi/abs/10.1137/1.9781611970104`.

Davies, Christopher T., Marius Cautun, Benjamin Giblin, et al. (2020). "Constraining cosmology with weak lensing voids". In: *arXiv e-prints*, arXiv:2010.11954, arXiv:2010.11954. arXiv: `2010.11954 [astro-ph.CO]`.

Davies, Christopher T, Marius Cautun, and Baojiu Li (July 2018). "Weak lensing by voids in weak lensing maps". In: *monthly Notices of the Royal Astronomical Society: Letters* 480.1, pp. L101–L105. ISSN: 1745-3925. DOI: `10.1093/mnrasl/sly135`. eprint: `https://academic.oup.com/mnrasl/article-pdf/480/1/L101/25440614/sly135.pdf`. URL: `https://doi.org/10.1093/mnrasl/sly135`.

Davies, Christopher T. et al. (2020). *Constraining cosmology with weak lensing voids*. arXiv: `2010.11954 [astro-ph.CO]`.

DES Collaboration, T. M. C. Abbott, et al. (2021). *Dark Energy Survey Year 3 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing*. arXiv: `2105.13549 [astro-ph.CO]`.

Deshpande, Anurag C., Peter L. Taylor, and Thomas D. Kitching (2020). "Accessing the high-ℓ frontier under the reduced shear approximation with k-cut cosmic shear". In:

Phys. Rev. D 102 (8), p. 083535. DOI: 10.1103/PhysRevD.102.083535. URL: https://link.aps.org/doi/10.1103/PhysRevD.102.083535.

DESI Collaboration et al. (2016). *The DESI Experiment Part I: Science,Targeting, and Survey Design.* arXiv: 1611.00036 [astro-ph.IM].

Desjacques, Vincent, Donghui Jeong, and Fabian Schmidt (2018). “Large-scale galaxy bias”. In: *Physics Reports* 733. Large-scale galaxy bias, pp. 1–193. ISSN: 0370-1573. DOI: https://doi.org/10.1016/j.physrep.2017.12.002. URL: https://www.sciencedirect.com/science/article/pii/S0370157317304192.

Dietrich, J. P. and J. Hartlap (2010). “Cosmology with the shear-peak statistics”. In: *MNRAS* 402.2, pp. 1049–1058. DOI: 10.1111/j.1365-2966.2009.15948.x. arXiv: 0906.3512 [astro-ph.CO].

Dietrich, J. P. and J. Hartlap (Feb. 2010). “Cosmology with the shear-peak statistics”. In: *monthly Notices of the Royal Astronomical Society* 402.2, pp. 1049–1058. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2009.15948.x. eprint: https://academic.oup.com/mnras/article-pdf/402/2/1049/3855286/mnras0402-1049.pdf. URL: https://doi.org/10.1111/j.1365-2966.2009.15948.x.

Dodelson, Scott (2003). *Modern cosmology.*

Dolag, K., E. Komatsu, and R. Sunyaev (Aug. 2016). “SZ effects in the Magneticum Pathfinder simulation: comparison with the Planck, SPT, and ACT results”. In: *monthly Notices of the Royal Astronomical Society* 463.2, pp. 1797–1811. ISSN: 0035-8711. DOI: 10.1093/mnras/stw2035. eprint: https://academic.oup.com/mnras/article-pdf/463/2/1797/9685287/stw2035.pdf. URL: https://doi.org/10.1093/mnras/stw2035.

Dubois, Yohan et al. (June 2013). “AGN-driven quenching of star formation: morphological and dynamical implications for early-type galaxies”. In: *monthly Notices of the Royal Astronomical Society* 433.4, pp. 3297–3313. ISSN: 0035-8711. DOI: 10.1093/mnras/stt997. eprint: https://academic.oup.com/mnras/article-pdf/433/4/3297/4933412/stt997.pdf. URL: https://doi.org/10.1093/mnras/stt997.

Dubois, Yohan et al. (Sept. 2016). “The Horizon-AGN simulation: morphological diversity of galaxies promoted by AGN feedback”. In: *monthly Notices of the Royal Astronomical Society* 463.4, pp. 3948–3964. ISSN: 0035-8711. DOI: 10.1093/mnras/stw2265. eprint: https://academic.oup.com/mnras/article-pdf/463/4/3948/18515487/stw2265.pdf. URL: https://doi.org/10.1093/mnras/stw2265.

Ducout, A. et al. (Dec. 2012). “Non-Gaussianity and Minkowski functionals: forecasts for Planck”. In: *monthly Notices of the Royal Astronomical Society* 429.3, pp. 2104–2126.

ISSN: 0035-8711. DOI: 10.1093/mnras/sts483. eprint: https://academic.oup.com/mnras/article-pdf/429/3/2104/3214532/sts483.pdf. URL: https://doi.org/10.1093/mnras/sts483.

Eisenstein, Daniel J. and Wayne Hu (1998). "Baryonic Features in the Matter Transfer Function". In: *The Astrophysical Journal* 496.2, pp. 605–614. DOI: 10.1086/305424. URL: https://doi.org/10.1086/305424.

Erben, T. et al. (June 2013). "CFHTLenS: the Canada–France–Hawaii Telescope Lensing Survey – imaging data and catalogue products". In: *Monthly Notices of the Royal Astronomical Society* 433.3, pp. 2545–2563. ISSN: 0035-8711. DOI: 10.1093/mnras/stt928. eprint: https://academic.oup.com/mnras/article-pdf/433/3/2545/13761746/stt928.pdf. URL: https://doi.org/10.1093/mnras/stt928.

Esfahani, Ali Ashtari, David M Asner, Sebastian Böser, et al. (2017). "Determining the neutrino mass with cyclotron radiation emission spectroscopy—Project 8". In: *Journal of Physics G: Nuclear and Particle Physics* 44.5, p. 054004. DOI: 10.1088/1361-6471/aa5b4f. URL: https://doi.org/10.1088/1361-6471/aa5b4f.

Euclid Collaboration et al. (2020a). In: *A&A* 644, A31. DOI: 10.1051/0004-6361/202039403. URL: https://doi.org/10.1051/0004-6361/202039403.

— (2020b). "Euclid preparation - VII. Forecast validation for Euclid cosmological probes". In: *A&A* 642, A191. DOI: 10.1051/0004-6361/202038071. URL: https://doi.org/10.1051/0004-6361/202038071.

Fan, Zuhui, Huanyuan Shan, and Jiayi Liu (2010). "NOISY WEAK-LENSING CONVERGENCE PEAK STATISTICS NEAR CLUSTERS OF GALAXIES AND BEYOND". In: *The Astrophysical Journal* 719.2, pp. 1408–1420. DOI: 10.1088/0004-637x/719/2/1408. URL: https://doi.org/10.1088/0004-637x/719/2/1408.

Fedeli, C. et al. (2014). "The clustering of baryonic matter. II: halo model and hydrodynamic simulations". In: *Journal of Cosmology and Astroparticle Physics* 2014.08, pp. 028–028. DOI: 10.1088/1475-7516/2014/08/028. URL: https://doi.org/10.1088/1475-7516/2014/08/028.

Fiedorowicz, Pier et al. (2021). *KaRMMa – Kappa Reconstruction for Mass Mapping*. arXiv: 2105.14699 [astro-ph.CO].

Finoguenov, A. et al. (Sept. 2007). "The XMM-Newton Wide-Field Survey in the COSMOS Field: Statistical Properties of Clusters of Galaxies". In: *ApJS* 172.1, pp. 182–195. DOI: 10.1086/516577. arXiv: astro-ph/0612360 [astro-ph].

Fluri, J., T. Kacprzak, R. Sgier, et al. (2018). "Weak lensing peak statistics in the era of large scale cosmological surveys". In: *JCAP* 2018.10, 051, p. 051. DOI: 10.1088/1475-7516/2018/10/051. arXiv: 1803.08461 [astro-ph.CO].

Fluri, Janis et al. (2018). "Cosmological constraints from noisy convergence maps through deep learning". In: *Phys. Rev. D* 98.12, 123518, p. 123518. DOI: 10.1103/PhysRevD.98.123518. arXiv: 1807.08732 [astro-ph.CO].

Fluri, Janis et al. (2018). "Cosmological constraints from noisy convergence maps through deep learning". In: *Phys. Rev. D* 98 (12), p. 123518. DOI: 10.1103/PhysRevD.98.123518. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.123518.

Fluri, Janis et al. (2019). "Cosmological constraints with deep learning from KiDS-450 weak lensing maps". In: *Phys. Rev. D* 100 (6), p. 063514. DOI: 10.1103/PhysRevD.100.063514. URL: https://link.aps.org/doi/10.1103/PhysRevD.100.063514.

Fong, Matthew et al. (July 2019). "The impact of baryonic physics and massive neutrinos on weak lensing peak statistics". In: *monthly Notices of the Royal Astronomical Society* 488.3, pp. 3340–3357. ISSN: 0035-8711. DOI: 10.1093/mnras/stz1882. URL: https://doi.org/10.1093/mnras/stz1882.

Foreman-Mackey, Daniel et al. (2013). "emcee: The MCMC Hammer". In: *PASP* 125.925, p. 306. DOI: 10.1086/670067. arXiv: 1202.3665 [astro-ph.IM].

Fortuna, Maria Cristina et al. (Dec. 2020). "The halo model as a versatile tool to predict intrinsic alignments". In: *monthly Notices of the Royal Astronomical Society* 501.2, pp. 2983–3002. ISSN: 0035-8711. DOI: 10.1093/mnras/staa3802. eprint: https://academic.oup.com/mnras/article-pdf/501/2/2983/35559288/staa3802.pdf. URL: https://doi.org/10.1093/mnras/staa3802.

Friedmann, A. (Jan. 1922). "Über die Krümmung des Raumes". In: *Zeitschrift fur Physik* 10, pp. 377–386. DOI: 10.1007/BF01332580.

Gabor, D. (1946). "Theory of communication. Part 1: The analysis of information". English. In: *Journal of the Institution of Electrical Engineers - Part III: Radio and Communication Engineering* 93 (26), 429–441(12). ISSN: 0367-7540. URL: https://digital-library.theiet.org/content/journals/10.1049/ji-3-2.1946.0074.

Gatti, M. et al. (2020). "Dark Energy Survey Year 3 results: cosmology with moments of weak lensing mass maps - validation on simulations". In: *MNRAS* 498.3, pp. 4060–4087. DOI: 10.1093/mnras/staa2680. arXiv: 1911.05568 [astro-ph.CO].

Gelman, Andrew and Donald B. Rubin (1992). "Inference from Iterative Simulation Using Multiple Sequences". In: *Statist. Sci.* 7, pp. 457–472. DOI: 10.1214/ss/1177011136.

Gelman, Andrew and Donald B. Rubin (Jan. 1992). "Inference from Iterative Simulation Using Multiple Sequences". In: *Statistical Science* 7, pp. 457–472. DOI: 10.1214/ss/1177011136.

Geweke, John F. (1991). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments.* Staff Report 148. Federal Reserve Bank of Minneapolis. URL: `https://ideas.repec.org/p/fip/fedmsr/148.html`.

Giblin, Benjamin et al. (Aug. 2018). "KiDS-450: enhancing cosmic shear with clipping transformations". In: *Monthly Notices of the Royal Astronomical Society* 480.4, pp. 5529–5549. ISSN: 0035-8711. DOI: `10.1093/mnras/sty2271`. eprint: `https://academic.oup.com/mnras/article-pdf/480/4/5529/25636792/sty2271.pdf`. URL: `https://doi.org/10.1093/mnras/sty2271`.

Giocoli, Carlo, Marco Baldi, and Lauro Moscardini (Sept. 2018). "Weak lensing light-cones in modified gravity simulations with and without massive neutrinos". In: *Monthly Notices of the Royal Astronomical Society* 481.2, pp. 2813–2828. ISSN: 0035-8711. DOI: `10.1093/mnras/sty2465`. eprint: `https://academic.oup.com/mnras/article-pdf/481/2/2813/25813034/sty2465.pdf`. URL: `https://doi.org/10.1093/mnras/sty2465`.

Giocoli, Carlo, Lauro Moscardini, Marco Baldi, et al. (2018). "Weak-lensing peaks in simulated light cones: investigating the coupling between dark matter and dark energy". In: *MNRAS* 478.4, pp. 5436–5448. DOI: `10.1093/mnras/sty1312`. arXiv: `1801.01886 [astro-ph.CO]`.

Goodman, Jonathan and Jonathan Weare (Jan. 2010). "Ensemble samplers with affine invariance". In: *Communications in Applied Mathematics and Computational Science* 5.1, pp. 65–80. DOI: `10.2140/camcos.2010.5.65`.

Grossmann, A. and J. Morlet (1984). "Decomposition of Hardy Functions into Square Integrable Wavelets of Constant Shape". In: *SIAM Journal on Mathematical Analysis* 15.4, pp. 723–736. DOI: `10.1137/0515056`. eprint: `https://doi.org/10.1137/0515056`. URL: `https://doi.org/10.1137/0515056`.

Group, Particle Data, P A Zyla, R M Barnett, et al. (Aug. 2020). "Review of Particle Physics". In: *Progress of Theoretical and Experimental Physics* 2020.8. 083C01. ISSN: 2050-3911. DOI: `10.1093/ptep/ptaa104`. eprint: `https://academic.oup.com/ptep/article-pdf/2020/8/083C01/34673722/ptaa104.pdf`. URL: `https://doi.org/10.1093/ptep/ptaa104`.

Guinot et al. (in prep.). "ShapePipe: a new shape measurement pipeline, weak-lensing application to UNIONS/CFIS data". In:

Guinot, Axel (2020). "Weak lensing analysis of the Canada-France Imaging Survey : from pixels to cosmology, preparation for the Euclid mission". Thèse de doctorat dirigée par Kilbinger, Martin et Mellier, Yannick Physique. Astronomie et astrophysique Université de Paris (2019-....) 2020. PhD thesis. URL: `http://www.theses.fr/2020UNIP7106`.

Gupta, Arushi, José Manuel Zorrilla Matilla, Daniel Hsu, et al. (2018). "Non-Gaussian information from weak lensing data via deep learning". In: *Phys. Rev. D* 97.10, 103515, p. 103515. DOI: 10.1103/PhysRevD.97.103515. arXiv: 1802.01212 [astro-ph.CO].

Guth, Alan H. (1981). "Inflationary universe: A possible solution to the horizon and flatness problems". In: *Phys. Rev. D* 23 (2), pp. 347–356. DOI: 10.1103/PhysRevD.23.347. URL: https://link.aps.org/doi/10.1103/PhysRevD.23.347.

Hamana, Takashi et al. (2012). "Scatter and bias in weak lensing selected clusters". In: *MNRAS* 425.3, pp. 2287–2298. DOI: 10.1111/j.1365-2966.2012.21582.x. arXiv: 1204.6117 [astro-ph.CO].

Hamana, Takashi et al. (Feb. 2020). "Cosmological constraints from cosmic shear two-point correlation functions with HSC survey first-year data". In: *Publications of the Astronomical Society of Japan* 72.1. 16. ISSN: 0004-6264. DOI: 10.1093/pasj/psz138. eprint: https://academic.oup.com/pasj/article-pdf/72/1/16/32408879/psz138.pdf. URL: https://doi.org/10.1093/pasj/psz138.

Hamaus, Nico, P. M. Sutter, and Benjamin D. Wandelt (2014). "Universal Density Profile for Cosmic Voids". In: *Phys. Rev. Lett.* 112 (25), p. 251302. DOI: 10.1103/PhysRevLett.112.251302. URL: https://link.aps.org/doi/10.1103/PhysRevLett.112.251302.

Hand, Nick et al. (2018). "nbodykit: An Open-source, Massively Parallel Toolkit for Large-scale Structure". In: *The Astronomical Journal* 156.4, p. 160. DOI: 10.3847/1538-3881/aadae0. URL: https://doi.org/10.3847/1538-3881/aadae0.

Harnois-Déraps, Joachim et al. (Dec. 2020). "Cosmic Shear Cosmology Beyond 2-Point Statistics: A Combined Peak Count and Correlation Function Analysis of DES-Y1". In: *arXiv e-prints*, arXiv:2012.02777, arXiv:2012.02777. arXiv: 2012.02777 [astro-ph.CO].

Harnois-Déraps, J., Giblin, B., and Joachimi, B. (2019). "Cosmic shear covariance matrix in wCDM: Cosmology matters". In: *A&A* 631, A160. DOI: 10.1051/0004-6361/201935912. URL: https://doi.org/10.1051/0004-6361/201935912.

Harnois-Déraps, J et al. (Aug. 2018). "Cosmological simulations for combined-probe analyses: covariance and neighbour-exclusion bias". In: *monthly Notices of the Royal Astronomical Society* 481.1, pp. 1337–1367. ISSN: 0035-8711. DOI: 10.1093/mnras/sty2319. eprint: https://academic.oup.com/mnras/article-pdf/481/1/1337/25720036/sty2319.pdf. URL: https://doi.org/10.1093/mnras/sty2319.

Harnois-Déraps, Joachim et al. (Sept. 2013). “High-performance P3M N-body code: cubep3m”. In: *Monthly Notices of the Royal Astronomical Society* 436.1, pp. 540–559. ISSN: 0035-8711. DOI: `10.1093/mnras/stt1591`. eprint: `https://academic.oup.com/mnras/article-pdf/436/1/540/2965832/stt1591.pdf`. URL: `https://doi.org/10.1093/mnras/stt1591`.

Hartlap, J., P. Simon, and P. Schneider (2007). “Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix”. In: *A&A* 464.1, pp. 399–404. DOI: `10.1051/0004-6361:20066170`. arXiv: `astro-ph/0608064 [astro-ph]`.

Hastings, W. K. (Apr. 1970). “Monte Carlo sampling methods using Markov chains and their applications”. In: *Biometrika* 57.1, pp. 97–109. ISSN: 0006-3444. DOI: `10.1093/biomet/57.1.97`. eprint: `https://academic.oup.com/biomet/article-pdf/57/1/97/23940249/57-1-97.pdf`. URL: `https://doi.org/10.1093/biomet/57.1.97`.

Heavens, Alan (June 2009). “Statistical techniques in cosmology”. In: *arXiv e-prints*, arXiv:0906.0664, arXiv:0906.0664. arXiv: `0906.0664 [astro-ph.CO]`.

Heavens, Alan, Alexandre Refregier, and Catherine Heymans (Dec. 2000). “Intrinsic correlation of galaxy shapes: implications for weak lensing measurements”. In: *monthly Notices of the Royal Astronomical Society* 319.2, pp. 649–656. ISSN: 0035-8711. DOI: `10.1111/j.1365-8711.2000.03907.x`. eprint: `https://academic.oup.com/mnras/article-pdf/319/2/649/4125368/319-2-649.pdf`. URL: `https://doi.org/10.1111/j.1365-8711.2000.03907.x`.

Heymans, Catherine et al. (2012). “CFHTLenS: the Canada–France–Hawaii Telescope Lensing Survey”. In: *monthly Notices of the Royal Astronomical Society* 427.1, 146–166. ISSN: 1365-2966. DOI: `10.1111/j.1365-2966.2012.21952.x`. URL: `http://dx.doi.org/10.1111/j.1365-2966.2012.21952.x`.

Heymans, Catherine et al. (2020). “KiDS-1000 Cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints”. In: *A&A, Forthcoming article*. DOI: `10.1051/0004-6361/202039063`. URL: `https://doi.org/10.1051/0004-6361/202039063`.

Hikage, Chiaki, Eiichiro Komatsu, and Takahiko Matsubara (2006). “Primordial Non-Gaussianity and Analytical Formula for Minkowski Functionals of the Cosmic Microwave Background and Large-Scale Structure”. In: *The Astrophysical Journal* 653.1, pp. 11–26. DOI: `10.1086/508653`. URL: `https://doi.org/10.1086/508653`.

Hilbert, Stefan et al. (2012). “Measuring primordial non-Gaussianity with weak lensing surveys”. In: *MNRAS* 426.4, pp. 2870–2888. DOI: `10.1111/j.1365-2966.2012.21841.x`. arXiv: `1204.4530 [astro-ph.CO]`.

Hildebrandt, H et al. (Oct. 2016). “KiDS-450: cosmological parameter constraints from tomographic weak gravitational lensing”. In: *monthly Notices of the Royal Astronomical Society* 465.2, pp. 1454–1498. ISSN: 0035-8711. DOI: `10.1093/mnras/stw2805`. eprint: `https://academic.oup.com/mnras/article-pdf/465/2/1454/24243465/stw2805.pdf`. URL: `https://doi.org/10.1093/mnras/stw2805`.

Hinshaw, G. et al. (2013). “NINE-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE (WMAP) OBSERVATIONS: COSMOLOGICAL PARAMETER RESULTS”. In: *The Astrophysical Journal Supplement Series* 208.2, p. 19. DOI: `10.1088/0067-0049/208/2/19`. URL: `https://doi.org/10.1088/0067-0049/208/2/19`.

Hinton, Samuel R. (Aug. 2016). “ChainConsumer”. In: *The Journal of Open Source Software* 1.4, 00045, p. 00045. DOI: `10.21105/joss.00045`.

Hirata, Christopher M. and Uro š Seljak (2004). “Intrinsic alignment-lensing interference as a contaminant of cosmic shear”. In: *Phys. Rev. D* 70 (6), p. 063526. DOI: `10.1103/PhysRevD.70.063526`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.70.063526`.

Hobson, M. P., G. P. Efstathiou, and A. N. Lasenby (2006). *General Relativity: An Introduction for Physicists*. Cambridge University Press. DOI: `10.1017/CBO9780511790904`.

Hoekstra, Henk, Massimo Viola, and Ricardo Herbonnet (Mar. 2017). “A study of the sensitivity of shape measurements to the input parameters of weak-lensing image simulations”. In: *Monthly Notices of the Royal Astronomical Society* 468.3, pp. 3295–3311. ISSN: 0035-8711. DOI: `10.1093/mnras/stx724`. eprint: `https://academic.oup.com/mnras/article-pdf/468/3/3295/13651776/stx724.pdf`. URL: `https://doi.org/10.1093/mnras/stx724`.

Holschneider, M. et al. (1989). “A Real-Time Algorithm for Signal Analysis with the Help of the Wavelet Transform”. In: *Wavelets*. Ed. by Jean-Michel Combes, Alexander Grossmann, and Philippe Tchamitchian. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 286–297. ISBN: 978-3-642-97177-8.

Hu, Wayne and Andrey V. Kravtsov (2003). “Sample Variance Considerations for Cluster Surveys”. In: *The Astrophysical Journal* 584.2, pp. 702–715. DOI: `10.1086/345846`. URL: `https://doi.org/10.1086/345846`.

Hubble, Edwin (1929). “A relation between distance and radial velocity among extra-galactic nebulae”. In: *Proceedings of the National Academy of Sciences* 15.3, pp. 168–173. ISSN: 0027-8424. DOI: `10.1073/pnas.15.3.168`. eprint: `https://www.pnas.org/content/15/3/168.full.pdf`. URL: `https://www.pnas.org/content/15/3/168`.

Huff, Eric and Rachel Mandelbaum (Feb. 2017). “Metacalibration: Direct Self-Calibration of Biases in Shear Measurement”. In: *arXiv e-prints*, arXiv:1702.02600, arXiv:1702.02600. arXiv: `1702.02600 [astro-ph.CO]`.

Huterer, Dragan et al. (Feb. 2006). “Systematic errors in future weak-lensing surveys: requirements and prospects for self-calibration”. In: *monthly Notices of the Royal Astronomical Society* 366.1, pp. 101–114. ISSN: 0035-8711. DOI: `10.1111/j.1365-2966.2005.09782.x`. eprint: `https://academic.oup.com/mnras/article-pdf/366/1/101/6248427/366-1-101.pdf`. URL: `https://doi.org/10.1111/j.1365-2966.2005.09782.x`.

Jain, Bhuvnesh and Ludovic Van Waerbeke (2000). “Statistics of Dark Matter Halos from Gravitational Lensing”. In: *The Astrophysical Journal* 530.1, pp. L1–L4. DOI: `10.1086/312480`. URL: `https://doi.org/10.1086/312480`.

Jaynes, E. T. (2003). *Probability Theory: The Logic of Science.* Ed. by G. LarryEditor Bretthorst. Cambridge University Press. DOI: `10.1017/CBO9780511790423`.

Jeffrey, N. et al. (2018). *Improving Weak Lensing Mass Map Reconstructions using Gaussian and Sparsity Priors: Application to DES SV.* eprint: `arXiv:1801.08945`.

Jeffrey, Niall, Justin Alsing, and François Lanusse (Nov. 2020). “Likelihood-free inference with neural compression of DES SV weak lensing map statistics”. In: *Monthly Notices of the Royal Astronomical Society* 501.1, pp. 954–969. ISSN: 0035-8711. DOI: `10.1093/mnras/staa3594`. eprint: `https://academic.oup.com/mnras/article-pdf/501/1/954/35102475/staa3594.pdf`. URL: `https://doi.org/10.1093/mnras/staa3594`.

Jeffrey, Niall et al. (Jan. 2020). “Deep learning dark matter map reconstructions from DES SV weak lensing data”. In: *monthly Notices of the Royal Astronomical Society* 492.4, pp. 5023–5029. ISSN: 0035-8711. DOI: `10.1093/mnras/staa127`. eprint: `https://academic.oup.com/mnras/article-pdf/492/4/5023/32358675/staa127.pdf`. URL: `https://doi.org/10.1093/mnras/staa127`.

Joachimi, Benjamin et al. (2015). “Galaxy Alignments: An Overview”. In: *Space Science Reviews* 193.1-4, pp. 1–65. DOI: `10.1007/s11214-015-0177-4`. arXiv: `1504.05456 [astro-ph.GA]`.

Joseph, R. et al. (Mar. 2019). “Sparse Lens Inversion Technique (SLIT): lens and source separability from linear inversion of the source reconstruction problem”. In: *A&A* 623, A14, A14. DOI: `10.1051/0004-6361/201731042`. arXiv: `1809.09121 [astro-ph.IM]`.

Joudaki, S. et al. (2020). “KiDS+VIKING-450 and DES-Y1 combined: Cosmology with cosmic shear”. In: *A&A* 638, p. L1. DOI: `10.1051/0004-6361/201936154`. URL: `https://doi.org/10.1051/0004-6361/201936154`.

Kacprzak, T. et al. (Aug. 2016). “Cosmology constraints from shear peak statistics in Dark Energy Survey Science Verification data”. In: *monthly Notices of the Royal Astronomical Society* 463.4, pp. 3653–3673. ISSN: 0035-8711. DOI: `10.1093/mnras/stw2070`. eprint: `https://academic.oup.com/mnras/article-pdf/463/4/3653/18511865/stw2070.pdf`. URL: `https://doi.org/10.1093/mnras/stw2070`.

Kacprzak, Tomasz et al. (May 2014). “Sérsic galaxy models in weak lensing shape measurement: model bias, noise bias and their interaction”. In: *monthly Notices of the Royal Astronomical Society* 441.3, pp. 2528–2538. ISSN: 0035-8711. DOI: `10.1093/mnras/stu588`. eprint: `https://academic.oup.com/mnras/article-pdf/441/3/2528/3867913/stu588.pdf`. URL: `https://doi.org/10.1093/mnras/stu588`.

Kaiser, Nick (1984). “On the Spatial correlations of Abell clusters”. In: *Astrophys. J. Lett.* 284, pp. L9–L12. DOI: `10.1086/184341`.

Kaiser, Nick (Apr. 1992). “Weak Gravitational Lensing of Distant Galaxies”. In: *ApJ* 388, p. 272. DOI: `10.1086/171151`.

Kaiser, Nick and Gordon Squires (Feb. 1993). “Mapping the Dark Matter with Weak Gravitational Lensing”. In: *ApJ* 404, p. 441. DOI: `10.1086/172297`.

Kaiser, Nick, Gillian Wilson, and Gerard A. Luppino (2000). *Large-Scale Cosmic Shear Measurements.* arXiv: `astro-ph/0003338 [astro-ph]`.

Kajita, Takaaki (1999). “Atmospheric neutrino results from Super-Kamiokande and Kamiokande — Evidence for ν_μ oscillations”. In: *Nuclear Physics B - Proceedings Supplements* 77.1, pp. 123 –132. ISSN: 0920-5632. DOI: `https://doi.org/10.1016/S0920-5632(99)00407-7`. URL: `http://www.sciencedirect.com/science/article/pii/S0920563299004077`.

Kiessling, Alina et al. (2015). “Galaxy Alignments: Theory, Modelling & Simulations”. In: *Space Science Reviews* 193.1-4, pp. 67–136. DOI: `10.1007/s11214-015-0203-6`. arXiv: `1504.05546 [astro-ph.GA]`.

Kilbinger, Martin (2015). In: *Reports on Progress in Physics* 78.8, p. 086901. DOI: `10.1088/0034-4885/78/8/086901`. URL: `https://doi.org/10.1088`.

Kilbinger, Martin, Liping Fu, Catherine Heymans, et al. (Feb. 2013). “CFHTLenS: combined probe cosmological model comparison using 2D weak gravitational lensing”. In: *monthly Notices of the Royal Astronomical Society* 430.3, pp. 2200–2220. ISSN: 0035-8711. DOI: `10.1093/mnras/stt041`. eprint: `https://academic.oup.com/mnras/article-pdf/430/3/2200/4928328/stt041.pdf`. URL: `https://doi.org/10.1093/mnras/stt041`.

Kilbinger, Martin et al. (Aug. 2017). “Precision calculations of the cosmic shear power spectrum projection”. In: *Monthly Notices of the Royal Astronomical Society* 472.2, pp. 2126–

2141. ISSN: 0035-8711. DOI: 10.1093/mnras/stx2082. eprint: https://academic.oup.com/mnras/article-pdf/472/2/2126/19943054/stx2082.pdf. URL: https://doi.org/10.1093/mnras/stx2082.

Kirk, Donnacha et al. (Aug. 2012). "The cosmological impact of intrinsic alignment model choice for cosmic shear". In: *monthly Notices of the Royal Astronomical Society* 424.3, pp. 1647–1657. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2012.21099.x. eprint: https://academic.oup.com/mnras/article-pdf/424/3/1647/2976995/424-3-1647.pdf. URL: https://doi.org/10.1111/j.1365-2966.2012.21099.x.

Kirk, Donnacha et al. (2015). "Galaxy Alignments: Observations and Impact on Cosmology". In: *Space Science Reviews* 193.1-4, 139–211. ISSN: 1572-9672. DOI: 10.1007/s11214-015-0213-4. URL: http://dx.doi.org/10.1007/s11214-015-0213-4.

Kratochvil, Jan M., Zoltán Haiman, and Morgan May (2010). "Probing cosmology with weak lensing peak counts". In: *Phys. Rev. D* 81.4, 043519, p. 043519. DOI: 10.1103/PhysRevD.81.043519. arXiv: 0907.0486 [astro-ph.CO].

Kratochvil, Jan M. et al. (2012). "Probing cosmology with weak lensing Minkowski functionals". In: *Phys. Rev. D* 85.10, 103513, p. 103513. DOI: 10.1103/PhysRevD.85.103513. arXiv: 1109.6334 [astro-ph.CO].

Kruse, Guido and Peter Schneider (1999). "Statistics of dark matter haloes expected from weak lensing surveys". In: *MNRAS* 302.4, pp. 821–829. DOI: 10.1046/j.1365-8711.1999.02195.x. arXiv: astro-ph/9806071 [astro-ph].

Lacasa, Fabien and Julien Grain (2019). "Fast and easy super-sample covariance of large-scale structure observables". In: *A&A* 624, A61. DOI: 10.1051/0004-6361/201834343. URL: https://doi.org/10.1051/0004-6361/201834343.

Lanusse, F. et al. (2016). "High resolution weak lensing mass mapping combining shear and flexion". In: *A&A* 591, A2.

Lanusse, Francois (Nov. 2015). "Sparse reconstruction of the dark matter mass map from weak gravitational lensing". Theses. Université Paris-Saclay. URL: https://tel.archives-ouvertes.fr/tel-01281927.

Laureijs, R. et al. (2011). "Euclid Definition Study Report". In: *arXiv e-prints*, arXiv:1110.3193, arXiv:1110.3193. arXiv: 1110.3193 [astro-ph.CO].

Lemos, P., DES Collaboration, et al. (June 2021). "Assessing tension metrics with dark energy survey and Planck data". In: *Monthly Notices of the Royal Astronomical Society* 505.4, pp. 6179–6194. ISSN: 0035-8711. DOI: 10.1093/mnras/stab1670. eprint: https://academic.oup.com/mnras/article-pdf/505/4/6179/38875790/stab1670.pdf. URL: https://doi.org/10.1093/mnras/stab1670.

Lemos, Pablo, Anthony Challinor, and George Efstathiou (2017). "The effect of Limber and flat-sky approximations on galaxy weak lensing". In: *Journal of Cosmology and Astroparticle Physics* 2017.05, pp. 014–014. DOI: `10.1088/1475-7516/2017/05/014`. URL: `https://doi.org/10.1088/1475-7516/2017/05/014`.

Leonard, A. et al. (July 2012). "Fast calculation of the weak lensing aperture mass statistic". In: *MNRAS* 423, pp. 3405–3412. DOI: `10.1111/j.1365-2966.2012.21133.x`. arXiv: `1204.4293`.

Leonard, Adrienne, François Lanusse, and Jean-Luc Starck (Mar. 2014). "GLIMPSE: accurate 3D weak lensing reconstructions using sparsity". In: *Monthly Notices of the Royal Astronomical Society* 440.2, pp. 1281–1294. ISSN: 0035-8711. DOI: `10.1093/mnras/stu273`. eprint: `https://academic.oup.com/mnras/article-pdf/440/2/1281/18498127/stu273.pdf`. URL: `https://doi.org/10.1093/mnras/stu273`.

Leonard, Adrienne, Sandrine Pires, and Jean-Luc Starck (July 2012). "Fast calculation of the weak lensing aperture mass statistic". In: *monthly Notices of the Royal Astronomical Society* 423.4, pp. 3405–3412. ISSN: 0035-8711. DOI: `10.1111/j.1365-2966.2012.21133.x`. URL: `https://doi.org/10.1111/j.1365-2966.2012.21133.x`.

Lesgourgues, Julien (2011). *The Cosmic Linear Anisotropy Solving System (CLASS) I: Overview.* arXiv: `1104.2932 [astro-ph.IM]`.

Lewis, Antony and Sarah Bridle (2002). "Cosmological parameters from CMB and other data: A Monte Carlo approach". In: *Phys. Rev. D* 66 (10), p. 103511. DOI: `10.1103/PhysRevD.66.103511`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.66.103511`.

Lewis, Antony, Anthony Challinor, and Anthony Lasenby (2000a). "Efficient computation of CMB anisotropies in closed FRW models". In: *ApJ* 538, pp. 473–476. DOI: `10.1086/309179`. arXiv: `astro-ph/9911177 [astro-ph]`.

— (2000b). "Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models". In: *The Astrophysical Journal* 538.2, pp. 473–476. DOI: `10.1086/309179`. URL: `https://doi.org/10.1086/309179`.

Li, Zack, Jia Liu, José Manuel Zorrilla Matilla, et al. (2019). "Constraining neutrino mass with tomographic weak lensing peak counts". In: *Phys. Rev. D* 99.6, 063527, p. 063527. DOI: `10.1103/PhysRevD.99.063527`. arXiv: `1810.01781 [astro-ph.CO]`.

Lima, Marcos et al. (Oct. 2008). "Estimating the redshift distribution of photometric galaxy samples". In: *monthly Notices of the Royal Astronomical Society* 390.1, pp. 118–130. ISSN: 0035-8711. DOI: `10.1111/j.1365-2966.2008.13510.x`. eprint: `https://academic.oup.com/mnras/article-pdf/390/1/118/`

2955170/mnras0390-0118.pdf. URL: https://doi.org/10.1111/j.1365-2966.2008.13510.x.

Limber, D. Nelson (Jan. 1953). "The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field." In: *ApJ* 117, p. 134. DOI: 10.1086/145672.

Lin, Chieh-An (2016). *Cosmology with weak-lensing peak counts.* arXiv: 1612.04041 [astro-ph.CO].

Lin, Chieh-An and Martin Kilbinger (2015). "A new model to predict weak-lensing peak counts. II. Parameter constraint strategies". In: *A&A* 583, A70, A70. DOI: 10.1051/0004-6361/201526659. arXiv: 1506.01076 [astro-ph.CO].

Lin, Chieh-An, Martin Kilbinger, and Sandrine Pires (2016). "A new model to predict weak-lensing peak counts. III. Filtering technique comparisons". In: *A&A* 593, A88, A88. DOI: 10.1051/0004-6361/201628565. arXiv: 1603.06773 [astro-ph.CO].

Lin, C.-A. et al. (2016). "A new model to predict weak-lensing peak counts - III. Filtering technique comparisons". In: *A&A* 593, A88. DOI: 10.1051/0004-6361/201628565. URL: https://doi.org/10.1051/0004-6361/201628565.

Lin, Chieh-An and Kilbinger, Martin (2015). "A new model to predict weak-lensing peak counts - I. Comparison with N-body simulations". In: *A&A* 576, A24. DOI: 10.1051/0004-6361/201425188. URL: https://doi.org/10.1051/0004-6361/201425188.

Linder, E. V. and A. Jenkins (2003). "Cosmic structure growth and dark energy". In: *Monthly Notices of the Royal Astronomical Society* 346.2, pp. 573–583. DOI: https://doi.org/10.1046/j.1365-2966.2003.07112.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1365-2966.2003.07112.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1365-2966.2003.07112.x.

Liu, Jia and Zoltán Haiman (2016). "Origin of weak lensing convergence peaks". In: *Phys. Rev. D* 94 (4), p. 043533. DOI: 10.1103/PhysRevD.94.043533. URL: https://link.aps.org/doi/10.1103/PhysRevD.94.043533.

Liu, Jia and Mathew S. Madhavacheril (2019). "Constraining neutrino mass with the tomographic weak lensing one-point probability distribution function and power spectrum". In: *Phys. Rev. D* 99 (8), p. 083508. DOI: 10.1103/PhysRevD.99.083508. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.083508.

Liu, Jia, Alvaro Ortiz-Vazquez, and J. Colin Hill (2016). "Constraining multiplicative bias in CFHTLenS weak lensing shear data". In: *Phys. Rev. D* 93 (10), p. 103508. DOI: 10.1103/PhysRevD.93.103508. URL: https://link.aps.org/doi/10.1103/PhysRevD.93.103508.

Liu, Jia, Andrea Petri, Zoltán Haiman, et al. (2015). "Cosmology constraints from the weak lensing peak counts and the power spectrum in CFHTLenS data". In: *Phys. Rev. D* 91.6, 063507, p. 063507. DOI: 10.1103/PhysRevD.91.063507. arXiv: 1412.0757 [astro-ph.CO].

Liu, Jia et al. (2018). "MassiveNuS: cosmological massive neutrino simulations". In: *JCAP* 2018.3, 049, p. 049. DOI: 10.1088/1475-7516/2018/03/049. arXiv: 1711.10524 [astro-ph.CO].

Liu, Xiangkun et al. (2016). "Constraining $f(R)$ Gravity Theory Using Weak Lensing Peak Statistics from the Canada-France-Hawaii-Telescope Lensing Survey". In: *Phys. Rev. Lett.* 117 (5), p. 051101. DOI: 10.1103/PhysRevLett.117.051101. URL: https://link.aps.org/doi/10.1103/PhysRevLett.117.051101.

Liu J., Jia et al. (2015). "Cosmology constraints from the weak lensing peak counts and the power spectrum in CFHTLenS data". In: *Phys. Rev. D* 91 (6), p. 063507. DOI: 10.1103/PhysRevD.91.063507. URL: https://link.aps.org/doi/10.1103/PhysRevD.91.063507.

Liu X., Xiangkun et al. (May 2015). "Cosmological constraints from weak lensing peak statistics with Canada–France–Hawaii Telescope Stripe 82 Survey". In: *monthly Notices of the Royal Astronomical Society* 450.3, pp. 2888–2902. ISSN: 0035-8711. DOI: 10.1093/mnras/stv784. eprint: https://academic.oup.com/mnras/article-pdf/450/3/2888/18508804/stv784.pdf. URL: https://doi.org/10.1093/mnras/stv784.

Longair, M.S. (1998). *Galaxy Formation*. Astronomy and Astrophysics Library. Springer Berlin Heidelberg. ISBN: 9783540637851. URL: https://books.google.dj/books?id=2ARuLT-tk5EC.

LSST Science Collaboration et al. (2009). "LSST Science Book, Version 2.0". In: *arXiv e-prints*, arXiv:0912.0201, arXiv:0912.0201. arXiv: 0912.0201 [astro-ph.IM].

Maki, Ziro, Masami Nakagawa, and Shoichi Sakata (1962). "Remarks on the unified model of elementary particles". In: *Prog. Theor. Phys.* 28, pp. 870–880. DOI: 10.1143/PTP.28.870.

Mallat, S.G. (1989). "A theory for multiresolution signal decomposition: the wavelet representation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11.7, pp. 674–693.

Mallat, Stéphane (2012). "Group Invariant Scattering". In: *Communications on Pure and Applied Mathematics* 65.10, pp. 1331–1398. DOI: https://doi.org/10.1002/cpa.21413. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/cpa.21413. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/cpa.21413.

Mallat, Stéphane, Sixin Zhang, and Gaspar Rochette (Nov. 2019). "Phase harmonic correlations and convolutional neural networks". In: *Information and Inference: A Journal* *of the IMA* 9.3, pp. 721–747. ISSN: 2049-8772. DOI: 10.1093/imaiai/iaz019. eprint: https://academic.oup.com/imaiai/article-pdf/9/3/721/33720733/iaz019.pdf. URL: https://doi.org/10.1093/imaiai/iaz019.

Marian, Laura et al. (2012). "Optimized detection of shear peaks in weak lensing maps". In: *MNRAS* 423.2, pp. 1711–1725. DOI: 10.1111/j.1365-2966.2012.20992.x. arXiv: 1110.4635 [astro-ph.CO].

— (2013). "The cosmological information of shear peaks: beyond the abundance". In: *MNRAS* 432.2, pp. 1338–1350. DOI: 10.1093/mnras/stt552. arXiv: 1301.5001 [astro-ph.CO].

Marques, Gabriela A. et al. (2019). "Constraining neutrino mass with weak lensing Minkowski Functionals". In: *JCAP* 2019.6, 019, p. 019. DOI: 10.1088/1475-7516/2019/06/019. arXiv: 1812.08206 [astro-ph.CO].

Martinelli, M. et al. (2020). "Euclid: impact of nonlinear prescriptions on cosmological parameter estimation from weak lensing cosmic shear". In: *arXiv e-prints*, arXiv:2010.12382, arXiv:2010.12382. arXiv: 2010.12382 [astro-ph.CO].

Martinet, N. et al. (2021). "Probing dark energy with tomographic weak-lensing aperture mass statistics". In: *A&A* 646, A62. DOI: 10.1051/0004-6361/202039679. URL: https://doi.org/10.1051/0004-6361/202039679.

Martinet, Nicolas et al. (2015). "Constraining cosmology with shear peak statistics: tomographic analysis". In: *A&A* 581, A101, A101. DOI: 10.1051/0004-6361/201425164. arXiv: 1506.02192 [astro-ph.CO].

Martinet, Nicolas et al. (2018). "KiDS-450: cosmological constraints from weak-lensing peak statistics - II: Inference from shear peaks using N-body simulations". In: *MNRAS* 474.1, pp. 712–730. DOI: 10.1093/mnras/stx2793. arXiv: 1709.07678 [astro-ph.CO].

Martinet, Nicolas et al. (2021). "Probing dark energy with tomographic weak-lensing aperture mass statistics". In: *A&A* 646, A62. DOI: 10.1051/0004-6361/202039679. URL: https://doi.org/10.1051/0004-6361/202039679.

Matilla, José Manuel Zorrilla et al. (2020). "Interpreting deep learning models for weak lensing". In: *Phys. Rev. D* 102 (12), p. 123506. DOI: 10.1103/PhysRevD.102.123506. URL: https://link.aps.org/doi/10.1103/PhysRevD.102.123506.

Maturi, M., C. Fedeli, and L. Moscardini (2011). "Imprints of primordial non-Gaussianity on the number counts of cosmic shear peaks". In: *MNRAS* 416.4, pp. 2527–2538. DOI: 10.1111/j.1365-2966.2011.18958.x. arXiv: 1101.4175 [astro-ph.CO].

Maturi, M. et al. (2010). "An analytic approach to number counts of weak-lensing peak detections". In: *A&A* 519, A23. DOI: 10.1051/0004-6361/200912866. URL: https://doi.org/10.1051/0004-6361/200912866.

McCarthy, Ian G. et al. (Oct. 2016). “The bahamas project: calibrated hydrodynamical simulations for large-scale structure cosmology”. In: *monthly Notices of the Royal Astronomical Society* 465.3, pp. 2936–2965. ISSN: 0035-8711. DOI: 10.1093/mnras/stw2792. eprint: https://academic.oup.com/mnras/article-pdf/465/3/2936/8485530/stw2792.pdf. URL: https://doi.org/10.1093/mnras/stw2792.

Mecke, K. R., T. Buchert, and H. Wagner (1994). “Robust morphological measures for large-scale structure in the Universe”. In: *A&A* 288, pp. 697–704. arXiv: astro-ph/9312028 [astro-ph].

Melchior, P. et al. (July 2018). “SCARLET: Source separation in multi-band images by Constrained Matrix Factorization”. In: *Astronomy and Computing* 24, 129, p. 129. DOI: 10.1016/j.ascom.2018.07.001. arXiv: 1802.10157 [astro-ph.IM].

Merten, Julian et al. (Apr. 2019). “On the dissection of degenerate cosmologies with machine learning”. In: *Monthly Notices of the Royal Astronomical Society* 487.1, pp. 104–122. ISSN: 0035-8711. DOI: 10.1093/mnras/stz972. eprint: https://academic.oup.com/mnras/article-pdf/487/1/104/28693917/stz972.pdf. URL: https://doi.org/10.1093/mnras/stz972.

Metropolis, Nicholas et al. (1953). “Equation of State Calculations by Fast Computing Machines”. In: *The Journal of Chemical Physics* 21.6, pp. 1087–1092. DOI: 10.1063/1.1699114. eprint: https://doi.org/10.1063/1.1699114. URL: https://doi.org/10.1063/1.1699114.

Meyer, Y. (1992). “Wavelets and Operators”. In: *Cambridge University Press* 22.

Ménard, Brice et al. (2014). *Clustering-based redshift estimation: method and application to data*. arXiv: 1303.4722 [astro-ph.CO].

Osato, Ken, Masato Shirasaki, and Naoki Yoshida (2015). “IMPACT OF BARYONIC PROCESSES ON WEAK-LENSING COSMOLOGY: POWER SPECTRUM, NONLOCAL STATISTICS, AND PARAMETER BIAS”. In: *The Astrophysical Journal* 806.2, p. 186. DOI: 10.1088/0004-637x/806/2/186. URL: https://doi.org/10.1088/0004-637x/806/2/186.

Papamakarios, George and Iain Murray (Dec. 2016). “Fast e-free Inference of Simulation Models with Bayesian Conditional Density Estimation”. English. In: Advances in Neural Information Processing Systems. 30th Annual Conference on Neural Information Processing Systems, NIPS 2016 ; Conference date: 05-12-2016 Through 10-12-2016, pp. 1028–1036. URL: https://nips.cc/Conferences/2016.

— (2018). *Fast ϵ-free Inference of Simulation Models with Bayesian Conditional Density Estimation*. arXiv: 1605.06376 [stat.ML].

Parroni, Carolina et al. (2020). “Going deep with Minkowski functionals of convergence maps”. In: *A&A* 633, A71, A71. DOI: 10.1051/0004-6361/201935988. arXiv: 1911.06243 [astro-ph.CO].

Parroni, Carolina et al. (2020). “Going deep with Minkowski functionals of convergence maps”. In: *A&A* 633, A71. DOI: 10.1051/0004-6361/201935988. URL: https://doi.org/10.1051/0004-6361/201935988.

Patton, Kenneth et al. (June 2017). “Cosmological constraints from the convergence 1-point probability distribution”. In: *monthly Notices of the Royal Astronomical Society* 472.1, pp. 439–446. ISSN: 0035-8711. DOI: 10.1093/mnras/stx1626. eprint: https://academic.oup.com/mnras/article-pdf/472/1/439/19692830/stx1626.pdf. URL: https://doi.org/10.1093/mnras/stx1626.

Paulin-Henriksson, S., Refregier, A., and Amara, A. (2009). “Optimal point spread function modeling for weak lensing: complexity and sparsity”. In: *A&A* 500.2, pp. 647–655. DOI: 10.1051/0004-6361/200811061. URL: https://doi.org/10.1051/0004-6361/200811061.

Peacock, John A. (1999). *Cosmological Physics.*

Peebles, P. J. E. (1980). *The large-scale structure of the universe.*

— (1993). *Principles of Physical Cosmology.*

Peel, A. et al. (2017a). “Cosmological constraints with weak-lensing peak counts and second-order statistics in a large-field survey”. In: *A&A* 599, A79. DOI: 10.1051/0004-6361/201629928. URL: https://doi.org/10.1051/0004-6361/201629928.

— (2017b). “Cosmological constraints with weak-lensing peak counts and second-order statistics in a large-field survey”. In: *A&A* 599, A79. DOI: 10.1051/0004-6361/201629928. URL: https://doi.org/10.1051/0004-6361/201629928.

Peel, A. et al. (2017). “Cosmological constraints with weak-lensing peak counts and second-order statistics in a large-field survey”. In: *A&A* 599, A79. DOI: 10.1051/0004-6361/201629928. URL: https://doi.org/10.1051/0004-6361/201629928.

Peel, A. et al. (2019). “Distinguishing standard and modified gravity cosmologies with machine learning”. In: *Phys. Rev. D* 100.2, 023508, p. 023508. DOI: 10.1103/PhysRevD.100.023508. arXiv: 1810.11030 [astro-ph.CO].

Peel, Austin, Valeria Pettorino, Carlo Giocoli, et al. (2018). “Breaking degeneracies in modified gravity with higher (than 2nd) order weak-lensing statistics”. In: *A&A* 619, A38, A38. DOI: 10.1051/0004-6361/201833481. arXiv: 1805.05146 [astro-ph.CO].

Penzias, A. A. and R. W. Wilson (July 1965). “A Measurement of Excess Antenna Temperature at 4080 Mc/s.” In: *ApJ* 142, pp. 419–421. DOI: 10.1086/148307.

Perlmutter, S. et al. (1999). “Measurements of Ω and Λ from 42 High-Redshift Supernovae”. In: *ApJ* 517.2, pp. 565–586. DOI: 10.1086/307221. arXiv: astro-ph/9812133 [astro-ph].

Petri, A. (Oct. 2016). “Mocking the weak lensing universe: The LensTools Python computing package”. In: *Astronomy and Computing* 17, pp. 73–79. DOI: 10.1016/j.ascom.2016.06.001. arXiv: 1606.01903 [astro-ph.CO].

Petri, Andrea, Morgan May, and Zoltán Haiman (2016). “Cosmology with photometric weak lensing surveys: Constraints with redshift tomography of convergence peaks and moments”. In: *Phys. Rev. D* 94.6, 063534, p. 063534. DOI: 10.1103/PhysRevD.94.063534. arXiv: 1605.01100 [astro-ph.CO].

Petri, Andrea et al. (2015). “Emulating the CFHTLenS weak lensing data: Cosmological constraints from moments and Minkowski functionals”. In: *Phys. Rev. D* 91.10, 103511, p. 103511. DOI: 10.1103/PhysRevD.91.103511. arXiv: 1503.06214 [astro-ph.CO].

Pillepich, Annalisa et al. (Oct. 2017). “Simulating galaxy formation with the IllustrisTNG model”. In: *monthly Notices of the Royal Astronomical Society* 473.3, pp. 4077–4106. ISSN: 0035-8711. DOI: 10.1093/mnras/stx2656. eprint: https://academic.oup.com/mnras/article-pdf/473/3/4077/21841785/stx2656.pdf. URL: https://doi.org/10.1093/mnras/stx2656.

Pires, S. et al. (May 2009). “FAst STatistics for weak Lensing (FASTLens): fast method for weak lensing statistics and map making”. In: *monthly Notices of the Royal Astronomical Society* 395.3, pp. 1265–1279. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2009.14625.x. eprint: https://academic.oup.com/mnras/article-pdf/395/3/1265/18232501/mnras0395-1265.pdf. URL: https://doi.org/10.1111/j.1365-2966.2009.14625.x.

Pires, Sandrine, Adrienne Leonard, and Jean-Luc Starck (2012). “Cosmological constraints from the capture of non-Gaussianity in weak lensing data”. In: *MNRAS* 423.1, pp. 983–992. DOI: 10.1111/j.1365-2966.2012.20940.x. arXiv: 1203.2877 [astro-ph.CO].

Pisani, Alice, Elena Massara, David N. Spergel, et al. (2019). “Cosmic voids: a novel probe to shed light on our Universe”. In: *Bulletin of the AAS* 51.3, 40, p. 40. arXiv: 1903.05161 [astro-ph.CO].

Planck Collaboration et al. (2014). “Planck 2013 results. XVI. Cosmological parameters”. In: *A&A* 571, A16. DOI: 10.1051/0004-6361/201321591. URL: https://doi.org/10.1051/0004-6361/201321591.

Planck Collaboration et al. (2020). “Planck 2018 results - VI. Cosmological parameters”. In: *A&A* 641, A6. DOI: 10.1051/0004-6361/201833910. URL: https://doi.org/10.1051/0004-6361/201833910.

Pontecorvo, B. (1968). “Neutrino Experiments and the Problem of Conservation of Leptonic Charge”. In: *Sov. Phys. JETP* 26, pp. 984–988.

Pozzetti, L. et al. (2016). “Modelling the number density of Htters for future spectroscopic near-IR space missions”. In: *A&A* 590, A3. doi: 10.1051/0004-6361/201527081. url: https://doi.org/10.1051/0004-6361/201527081.

Price, M. A. et al. (Nov. 2019). “Sparse Bayesian mass mapping with uncertainties: peak statistics and feature locations”. In: *MNRAS* 489.3, pp. 3236–3250. doi: 10.1093/mnras/stz2373. arXiv: 1812.04018 [astro-ph.CO].

Rasmussen, Carl Edward and Christopher K. I. Williams (2005). *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press. isbn: 026218253X.

Refregier, A. et al. (2018). “PyCosmo: An integrated cosmological Boltzmann solver”. In: *Astronomy and Computing* 25, pp. 38–43. issn: 2213-1337. doi: https://doi.org/10.1016/j.ascom.2018.08.001. url: https://www.sciencedirect.com/science/article/pii/S2213133717300963.

Refregier, Alexandre (2003). “Weak Gravitational Lensing by Large-Scale Structure”. In: *Annual Review of Astronomy and Astrophysics* 41.1, pp. 645–668. doi: 10.1146/annurev.astro.41.111302.102207. eprint: https://doi.org/10.1146/annurev.astro.41.111302.102207. url: https://doi.org/10.1146/annurev.astro.41.111302.102207.

Refregier, Alexandre et al. (Sept. 2012). “Noise bias in weak lensing shape measurements”. In: *monthly Notices of the Royal Astronomical Society* 425.3, pp. 1951–1957. issn: 0035-8711. doi: 10.1111/j.1365-2966.2012.21483.x. eprint: https://academic.oup.com/mnras/article-pdf/425/3/1951/3035730/425-3-1951.pdf. url: https://doi.org/10.1111/j.1365-2966.2012.21483.x.

Ribli, DezsHo et al. (2019). “Weak lensing cosmology with convolutional neural networks on noisy data”. In: *MNRAS* 490.2, pp. 1843–1860. doi: 10.1093/mnras/stz2610. arXiv: 1902.03663 [astro-ph.CO].

Ribli, Dezső, Bálint Ármin Pataki, and István Csabai (2018). *An improved cosmological parameter inference scheme motivated by deep learning*. arXiv: 1806.05995 [astro-ph.CO].

Riess, Adam G. et al. (1998). “Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant”. In: *The Astronomical Journal* 116.3, pp. 1009–1038. doi: 10.1086/300499. url: https://doi.org/10.1086/300499.

Robert, Christian P. and Wu Changye (2020). *Markov Chain Monte Carlo Methods, a survey with some frequent misunderstandings*. arXiv: 2001.06249 [stat.CO].

Rubin, Donald B. (1984). “Bayesianly Justifiable and Relevant Frequency Calculations for the Applied Statistician”. In: *The Annals of Statistics* 12.4, pp. 1151 –1172. doi: 10.

1214/aos/1176346785. URL: https://doi.org/10.1214/aos/1176346785.

Rubin, Vera C. and Jr. Ford W. Kent (Feb. 1970). "Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions". In: *ApJ* 159, p. 379. DOI: 10.1086/150317.

Saar, Enn et al. (Dec. 2006). "Multiscale morphology of the galaxy distribution". In: *monthly Notices of the Royal Astronomical Society* 374.3, pp. 1030–1044. ISSN: 0035-8711. DOI: 10.1111/j.1365-2966.2006.11237.x. eprint: https://academic.oup.com/mnras/article-pdf/374/3/1030/3538223/mnras0374-1030.pdf. URL: https://doi.org/10.1111/j.1365-2966.2006.11237.x.

Santos, Mario G. et al. (2015). *Cosmology with a SKA HI intensity mapping survey*. arXiv: 1501.03989 [astro-ph.CO].

Schaye, Joop et al. (Nov. 2014). "The EAGLE project: simulating the evolution and assembly of galaxies and their environments". In: *monthly Notices of the Royal Astronomical Society* 446.1, pp. 521–554. ISSN: 0035-8711. DOI: 10.1093/mnras/stu2058. eprint: https://academic.oup.com/mnras/article-pdf/446/1/521/4139718/stu2058.pdf. URL: https://doi.org/10.1093/mnras/stu2058.

Schmelzle, Jorit et al. (2017). *Cosmological model discrimination with Deep Learning*. arXiv: 1707.05167 [astro-ph.CO].

Schneider, Aurel and Romain Teyssier (2015). "A new method to quantify the effects of baryons on the matter power spectrum". In: *Journal of Cosmology and Astroparticle Physics* 2015.12, pp. 049–049. DOI: 10.1088/1475-7516/2015/12/049. URL: https://doi.org/10.1088/1475-7516/2015/12/049.

Schneider, Aurel et al. (2019). "Quantifying baryon effects on the matter power spectrum and the weak lensing shear correlation". In: *Journal of Cosmology and Astroparticle Physics* 2019.03, pp. 020–020. DOI: 10.1088/1475-7516/2019/03/020. URL: https://doi.org/10.1088/1475-7516/2019/03/020.

Schneider, P. (2006). "Weak Gravitational Lensing". In: *Gravitational Lensing: Strong, Weak and Micro*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 269–451. ISBN: 978-3-540-30310-7. DOI: 10.1007/978-3-540-30310-7_3. URL: https://doi.org/10.1007/978-3-540-30310-7_3.

Schneider, P. and Lombardi, M. (2003). "The three-point correlation function of cosmic shear - I. The natural components". In: *A&A* 397.3, pp. 809–818. DOI: 10.1051/0004-6361:20021541. URL: https://doi.org/10.1051/0004-6361:20021541.

Schneider, P. et al. (2002). “Analysis of two-point statistics of cosmic shear - I. Estimators and covariances”. In: *A&A* 396.1, pp. 1–19. DOI: `10.1051/0004-6361:20021341`. URL: `https://doi.org/10.1051/0004-6361:20021341`.

Schrabback, T. et al. (2010). “Evidence of the accelerated expansion of the Universe from weak lensing tomography with COSMOS”. In: *A&A* 516, A63, A63. DOI: `10.1051/0004-6361/200913577`. arXiv: `0911.0053 [astro-ph.CO]`.

Scoville, N. et al. (2007). “COSMOS: Hubble Space Telescope Observations”. In: *ApJS* 172.1, pp. 38–45. DOI: `10.1086/516580`. arXiv: `astro-ph/0612306 [astro-ph]`.

Seitz, C. and P. Schneider (Feb. 1997). “Steps towards nonlinear cluster inversion through gravitational distortions. III. Including a redshift distribution of the sources.” In: *A&A* 318, pp. 687–699. arXiv: `astro-ph/9601079 [astro-ph]`.

Seitz, Stella, Peter Schneider, and Jürgen Ehlers (1994). “Light propagation in arbitrary spacetimes and the gravitational lens approximation”. In: *Classical and Quantum Gravity* 11.9, pp. 2345–2373. DOI: `10.1088/0264-9381/11/9/016`. arXiv: `astro-ph/9403056 [astro-ph]`.

Seljak, Uroš, Anže Slosar, and Patrick McDonald (2006). “Cosmological parameters from combining the Lyman-alpha forest with CMB, galaxy clustering and SN constraints”. In: *Journal of Cosmology and Astroparticle Physics* 2006.10, pp. 014–014. DOI: `10.1088/1475-7516/2006/10/014`. URL: `https://doi.org/10.1088/1475-7516/2006/10/014`.

Sellentin, Elena and Alan F. Heavens (2016). “Parameter inference with estimated covariance matrices”. In: *MNRAS* 456.1, pp. L132–L136. DOI: `10.1093/mnrasl/slv190`. arXiv: `1511.05969 [astro-ph.CO]`.

Sellentin, Elena and Alan F. Heavens (Sept. 2017). “On the insufficiency of arbitrarily precise covariance matrices: non-Gaussian weak-lensing likelihoods”. In: *Monthly Notices of the Royal Astronomical Society* 473.2, pp. 2355–2363. ISSN: 0035-8711. DOI: `10.1093/mnras/stx2491`. eprint: `https://academic.oup.com/mnras/article-pdf/473/2/2355/21564684/stx2491.pdf`. URL: `https://doi.org/10.1093/mnras/stx2491`.

Sellentin, Elena and Alan F. Heavens (2017). “Quantifying lost information due to covariance matrix estimation in parameter inference”. In: *MNRAS* 464.4, pp. 4658–4665. DOI: `10.1093/mnras/stw2697`. arXiv: `1609.00504 [astro-ph.CO]`.

Semboloni, Elisabetta et al. (Oct. 2011). “Quantifying the effect of baryon physics on weak lensing tomography”. In: *monthly Notices of the Royal Astronomical Society* 417.3, pp. 2020–2035. ISSN: 0035-8711. DOI: `10.1111/j.1365-2966.2011.19385.x`. eprint: `https://academic.oup.com/mnras/article-pdf/417/3/2020/3811548/mnras0417-2020.pdf`. URL: `https://doi.org/10.1111/j.1365-2966.2011.19385.x`.

Sgier, R.J. et al. (2019). “Fast generation of covariance matrices for weak lensing”. In: *Journal of Cosmology and Astroparticle Physics* 2019.01, pp. 044–044. DOI: `10.1088/1475-7516/2019/01/044`. URL: `https://doi.org/10.1088/1475-7516/2019/01/044`.

Shan, HuanYuan et al. (2017). “KiDS-450: cosmological constraints from weak lensing peak statistics – I. Inference from analytical prediction of high signal-to-noise ratio convergence peaks”. In: *monthly Notices of the Royal Astronomical Society* 474.1, pp. 1116–1134. ISSN: 0035-8711. DOI: `10.1093/mnras/stx2837`. eprint: `https://academic.oup.com/mnras/article-pdf/474/1/1116/22370928/stx2837.pdf`. URL: `https://doi.org/10.1093/mnras/stx2837`.

Shapiro, Charles and Asantha Cooray (2006). “The Born and lens–lens corrections to weak gravitational lensing angular power spectra”. In: *Journal of Cosmology and Astroparticle Physics* 2006.03, pp. 007–007. DOI: `10.1088/1475-7516/2006/03/007`. URL: `https://doi.org/10.1088/1475-7516/2006/03/007`.

Sharma, Sanjib (2017). “Markov Chain Monte Carlo Methods for Bayesian Data Analysis in Astronomy”. In: *Annual Review of Astronomy and Astrophysics* 55.1, pp. 213–259. DOI: `10.1146/annurev-astro-082214-122339`. eprint: `https://doi.org/10.1146/annurev-astro-082214-122339`. URL: `https://doi.org/10.1146/annurev-astro-082214-122339`.

Shirasaki, Masato et al. (2019). “Decoding Cosmological Information in Weak-Lensing Mass Maps with Generative Adversarial Networks”. In: *arXiv e-prints*, arXiv:1911.12890, arXiv:1911.12890. arXiv: `1911.12890 [astro-ph.CO]`.

Simon, P. (2007). “How accurate is Limber’s equation?” In: *A&A* 473.3, pp. 711–714. DOI: `10.1051/0004-6361:20066352`. URL: `https://doi.org/10.1051/0004-6361:20066352`.

Smith, R. E. et al. (June 2003). “Stable clustering, the halo model and non-linear cosmological power spectra”. In: *Monthly Notices of the Royal Astronomical Society* 341.4, pp. 1311–1332. ISSN: 0035-8711. DOI: `10.1046/j.1365-8711.2003.06503.x`. eprint: `https://academic.oup.com/mnras/article-pdf/341/4/1311/3500043/341-4-1311.pdf`. URL: `https://doi.org/10.1046/j.1365-8711.2003.06503.x`.

Smoot, G F (1999). *Summary of Results from COBE.* Tech. rep. astro-ph/9902027. URL: `https://cds.cern.ch/record/378075`.

Spergel, D. et al. (2015). *Wide-Field InfrarRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA 2015 Report.* arXiv: `1503.03757 [astro-ph.IM]`.

Starck, J., J. Fadili, and F. Murtagh (2007). “The Undecimated Wavelet Decomposition and its Reconstruction”. In: *IEEE Transactions on Image Processing* 16.2, pp. 297–309. DOI: `10.1109/TIP.2006.887733`.

Starck, J.-L., M. Elad, and D.L. Donoho (2004). “Redundant Multiscale Transforms and their Application for Morphological Component Analysis”. In: *Advances in Imaging and Electron Physics* 132.

Starck, J.-L. et al. (2021). “Weak-lensing mass reconstruction using sparsity and a Gaussian random field”. In: *A&A* 649, A99. DOI: `10.1051/0004-6361/202039451`. URL: `https://doi.org/10.1051/0004-6361/202039451`.

Starck, Jean-Luc, Jalal Fadili, and Fionn Murtagh (Mar. 2007). “The Undecimated Wavelet Decomposition and its Reconstruction”. In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 16, pp. 297–309. DOI: `10.1109/TIP.2006.887733`.

Starck, Jean-Luc and Fionn Murtagh (1998). “Automatic Noise Estimation from the Multiresolution Support”. In: *Publications of the Astronomical Society of the Pacific* 110.744, pp. 193–199. DOI: `10.1086/316124`. URL: `https://doi.org/10.1086`.

Starck, Jean-Luc, Fionn Murtagh, and Jalal Fadili (2010). *Sparse Image and Signal Processing: Wavelets, Curvelets, Morphological Diversity*. New York, NY, USA: Cambridge University Press. ISBN: 0521119138, 9780521119139.

— (2016). *Sparse Image and Signal Processing: Wavelets and Related Geometric Multiscale Analysis*. Cambridge University Press.

Starck, Jean-Luc, Fionn D. Murtagh, and Albert Bijaoui (1998). “The wavelet transform”. In: *Image Processing and Data Analysis: The Multiscale Approach*. Cambridge University Press, 1–45. DOI: `10.1017/CBO9780511564352.002`.

Takada, Masahiro and Wayne Hu (2013). “Power spectrum super-sample covariance”. In: *Phys. Rev. D* 87 (12), p. 123504. DOI: `10.1103/PhysRevD.87.123504`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.87.123504`.

Takada, Masahiro and Bhuvnesh Jain (Sept. 2003). “Three-point correlations in weak lensing surveys: model predictions and applications”. In: *monthly Notices of the Royal Astronomical Society* 344.3, pp. 857–886. ISSN: 0035-8711. DOI: `10.1046/j.1365-8711.2003.06868.x`. eprint: `https://academic.oup.com/mnras/article-pdf/344/3/857/3307546/344-3-857.pdf`. URL: `https://doi.org/10.1046/j.1365-8711.2003.06868.x`.

Takada, Masahiro and Bhuvnesh Jain (2004). “Cosmological parameters from lensing power spectrum and bispectrum tomography”. In: *MNRAS* 348.3, pp. 897–915. DOI: `10.1111/j.1365-2966.2004.07410.x`. arXiv: `astro-ph/0310125 [astro-ph]`.

Takahashi, Ryuichi et al. (2012). “REVISING THE HALOFIT MODEL FOR THE NONLINEAR MATTER POWER SPECTRUM”. In: *The Astrophysical Journal* 761.2, p. 152. DOI: `10.1088/0004-637x/761/2/152`. URL: `https://doi.org/10.1088/0004-637x/761/2/152`.

Taylor, Peter L., Francis Bernardeau, and Eric Huff (2021). "x-cut Cosmic shear: Optimally removing sensitivity to baryonic and nonlinear physics with an application to the Dark Energy Survey year 1 shear data". In: *Phys. Rev. D* 103 (4), p. 043531. DOI: 10.1103/PhysRevD.103.043531. URL: https://link.aps.org/doi/10.1103/PhysRevD.103.043531.

Taylor, Peter L., Thomas D. Kitching, and Jason D. McEwen (2019). "Nonparametric cosmology with cosmic shear". In: *Phys. Rev. D* 99 (4), p. 043532. DOI: 10.1103/PhysRevD.99.043532. URL: https://link.aps.org/doi/10.1103/PhysRevD.99.043532.

Tegmark, Max, Michael A. Strauss, Michael R. Blanton, et al. (2004). "Cosmological parameters from SDSS and WMAP". In: *Phys. Rev. D* 69 (10), p. 103501. DOI: 10.1103/PhysRevD.69.103501. URL: https://link.aps.org/doi/10.1103/PhysRevD.69.103501.

Thiele, Leander, J. Colin Hill, and Kendrick M. Smith (2020). "Accurate analytic model for the weak lensing convergence one-point probability distribution function and its autocovariance". In: *Phys. Rev. D* 102 (12), p. 123545. DOI: 10.1103/PhysRevD.102.123545. URL: https://link.aps.org/doi/10.1103/PhysRevD.102.123545.

Torrado, Jesús and Antony Lewis (2021). "Cobaya: code for Bayesian analysis of hierarchical physical models". In: *Journal of Cosmology and Astroparticle Physics* 2021.05, p. 057. DOI: 10.1088/1475-7516/2021/05/057. URL: https://doi.org/10.1088/1475-7516/2021/05/057.

Troxel, M. A. et al. (2018). "Dark Energy Survey Year 1 results: Cosmological constraints from cosmic shear". In: *Phys. Rev. D* 98 (4), p. 043528. DOI: 10.1103/PhysRevD.98.043528. URL: https://link.aps.org/doi/10.1103/PhysRevD.98.043528.

Verde, L. (2010). "Statistical Methods in Cosmology". In: *Lectures on Cosmology: Accelerated Expansion of the Universe.* Ed. by Georg Wolschin. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 147–177. ISBN: 978-3-642-10598-2. DOI: 10.1007/978-3-642-10598-2_4. URL: https://doi.org/10.1007/978-3-642-10598-2_4.

Vicinanza, M. et al. (2016). "Higher order moments of lensing convergence - I. Estimate from simulations". In: *arXiv e-prints*, arXiv:1606.03892, arXiv:1606.03892. arXiv: 1606.03892 [astro-ph.CO].

Vicinanza, Martina, Vincenzo F. Cardone, Roberto Maoli, et al. (2018). "Increasing the lensing figure of merit through higher order convergence moments". In: *Phys. Rev. D* 97.2, 023519, p. 023519. DOI: 10.1103/PhysRevD.97.023519. arXiv: 1802.02963 [astro-ph.CO].

Vicinanza, Martina et al. (2019). “Minkowski functionals of convergence maps and the lensing figure of merit”. In: *Phys. Rev. D* 99.4, 043534, p. 043534. DOI: 10.1103/PhysRevD.99.043534. arXiv: 1905.00410 [astro-ph.CO].

Villaescusa-Navarro, Francisco et al. (2020). “The Quijote Simulations”. In: *The Astrophysical Journal Supplement Series* 250.1, p. 2. DOI: 10.3847/1538-4365/ab9d82. URL: https://doi.org/10.3847/1538-4365/ab9d82.

Vogelsberger, Mark et al. (Aug. 2014). “Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe”. In: *monthly Notices of the Royal Astronomical Society* 444.2, pp. 1518–1547. ISSN: 0035-8711. DOI: 10.1093/mnras/stu1536. eprint: https://academic.oup.com/mnras/article-pdf/444/2/1518/24040116/stu1536.pdf. URL: https://doi.org/10.1093/mnras/stu1536.

Waerbeke, L. Van et al. (2000). *Detection of correlated galaxy ellipticities on CFHT data: first evidence for gravitational lensing by large-scale structures.* arXiv: astro-ph/0002500 [astro-ph].

Waerbeke, Ludovic Van (Apr. 2000). “Noise properties of gravitational lens mass reconstruction”. In: *monthly Notices of the Royal Astronomical Society* 313.3, pp. 524–532. ISSN: 0035-8711. DOI: 10.1046/j.1365-8711.2000.03259.x. eprint: https://academic.oup.com/mnras/article-pdf/313/3/524/3684463/313-3-524.pdf. URL: https://doi.org/10.1046/j.1365-8711.2000.03259.x.

Wagner-Carena, S. et al. (Mar. 2020). “A novel CMB component separation method: hierarchical generalized morphological component analysis”. In: *MNRAS* 494.1, pp. 1507–1529. DOI: 10.1093/mnras/staa744. arXiv: 1910.08077 [astro-ph.CO].

Weinberg, David H. et al. (2003). “The Lyman-alpha Forest as a Cosmological Tool”. In: *AIP Conference Proceedings* 666.1, pp. 157–169. DOI: 10.1063/1.1581786. eprint: https://aip.scitation.org/doi/pdf/10.1063/1.1581786. URL: https://aip.scitation.org/doi/abs/10.1063/1.1581786.

Weinberg, David H. et al. (2013). “Observational probes of cosmic acceleration”. In: *Phys. Rep.* 530.2, pp. 87–255. DOI: 10.1016/j.physrep.2013.05.001. arXiv: 1201.2434 [astro-ph.CO].

Weinberg, Steven (1972). *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity.*

Weiss, Andreas J. et al. (2019). “Effects of baryons on weak lensing peak statistics”. In: *Journal of Cosmology and Astroparticle Physics* 2019.10, pp. 011–011. DOI: 10.1088/1475-7516/2019/10/011. URL: https://doi.org/10.1088/1475-7516/2019/10/011.

Wiener, Norbert (1949). *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications*. Vol. 7. MIT press Cambridge.

Wolf, Joachim (2010). "The KATRIN neutrino mass experiment". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 623.1. 1st International Conference on Technology and Instrumentation in Particle Physics, pp. 442 –444. ISSN: 0168-9002. DOI: `https://doi.org/10.1016/j.nima.2010.03.030`. URL: `http://www.sciencedirect.com/science/article/pii/S0168900210005942`.

Yuan, Shuo et al. (2018). "Projection Effects of Large-scale Structures on Weak-lensing Peak Abundances". In: *The Astrophysical Journal* 857.2, p. 112. DOI: `10.3847/1538-4357/aab900`. URL: `https://doi.org/10.3847/1538-4357/aab900`.

Zaroubi, S. et al. (1995). "Wiener Reconstruction of the Large-Scale Structure". In: *Astrophysical Journal* 449, p. 446. eprint: `astro-ph/9410080`.

Zorrilla Matilla, José Manuel et al. (2016). "Do dark matter halos explain lensing peaks?" In: *Phys. Rev. D* 94 (8), p. 083506. DOI: `10.1103/PhysRevD.94.083506`. URL: `https://link.aps.org/doi/10.1103/PhysRevD.94.083506`.

Zuntz, J. et al. (2015). "CosmoSIS: Modular cosmological parameter estimation". In: *Astronomy and Computing* 12, pp. 45–59. ISSN: 2213-1337. DOI: `https://doi.org/10.1016/j.ascom.2015.05.005`. URL: `https://www.sciencedirect.com/science/article/pii/S2213133715000591`.

Zuntz, J et al. (Aug. 2018). "Dark Energy Survey Year 1 results: weak lensing shape catalogues". In: *monthly Notices of the Royal Astronomical Society* 481.1, pp. 1149–1182. ISSN: 0035-8711. DOI: `10.1093/mnras/sty2219`. eprint: `https://academic.oup.com/mnras/article-pdf/481/1/1149/25719997/sty2219.pdf`. URL: `https://doi.org/10.1093/mnras/sty2219`.

Zwicky, F. (1937). "On the Masses of Nebulae and of Clusters of Nebulae". In: *ApJ* 86, p. 217. DOI: `10.1086/143864`.

Zürcher, Dominik et al. (2021). "Cosmological forecast for non-Gaussian statistics in large-scale weak lensing surveys". In: *Journal of Cosmology and Astroparticle Physics* 2021.01, pp. 028–028. DOI: `10.1088/1475-7516/2021/01/028`. URL: `https://doi.org/10.1088/1475-7516/2021/01/028`.