Non-Gaussian cosmology : theoretical and statistical challenges for modern galaxy surveys

Matteo Rizzato

THÈSE DE DOCTORAT

Sorbonne Université

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

Institut d'Astrophysique de Paris

Non-Gaussian Cosmology: theoretical and statistical challenges for modern galaxy surveys

PRESENTÉE ET SOUTENUE PUBLIQUEMENT LE 26 SEPTEMBRE 2019

PAR

MATTEO RIZZATO

DEVANT UN JURY COMPOSÉ DE :

MICHAEL JOYCE	LPNHE, SORBONNE UNIVERSITÉ	*PRÉSIDENT DU JURY*
FRANCIS BERNARDEAU	IAP, SORBONNE UNIVERSITÉ	*DIRECTEUR DE THÈSE*
KARIM BENABED	IAP, SORBONNE UNIVERSITÉ	*DIRECTEUR DE THÈSE*
BHUVNESH JAIN	UNIVERSITY OF PENNSYLVANIA	*RAPPORTEUR*
MARTIN CROCCE	INSTITUT DE CIÉNCES DE L'ESPAI (ICE-CSIC, IEEC)	*RAPPORTEUR*
ELISABETH KRAUSE	THE UNIVERSITY OF ARIZONA	*EXAMINATRICE*
ERIC JULLO	LAM, AIX-MARSEILLE UNIVERSITÉ	*EXAMINATEUR*
VINCENT DESJACQUES	TECHNION, ISRAEL INSTITUTE OF TECHNOLOGY	*EXAMINATEUR*

Acknowledgments

The journey that started three years ago has finally come to its end. This experience led me through both thrilling and difficult moments, which eventually worked for the best, prompting a personal growth towards a full professional and personal maturity. I was not alone in this period of my life and I would like to thank all the people who have been by my side during these intense three years.

First and foremost, my intellectual debt is to my supervisors, Francis Bernardeau and Karim Benabed, whose professional experience and leadership were a guiding light for me throughout the whole PhD experience. In addition, I would like to render my warmest thanks to my thesis jury, Micheal Joyce as president, Buvanesh Jain and Martin Crocce as reviewers, Elisabeth Krause, Eric Jullo and Vincent Desjacques as examiners. Special thanks to Martin Kilbinger. I am grateful and fortunate to have a recommendation letter from him. I have greatly benefited from the professional and personal exchanges with Sandrine Codis, Guilhem Lavaux, Silvia Galli and Tom Charnock, whom I want to thank for their constant interest in my work, wisdom and golden pieces of advice. I am in particular grateful to Elena Sellentin, who was an endless source of support and encouragement especially during the last year of my PhD and with whom I will pursue further fruitful collaborations. Thank you Elena for giving me the chance to pursue the goal of becoming a researcher. Fabien Lacasa and Paulo Reimberg have been greatly tolerant and supportive during the preparation of our papers together. I am looking forward to more collaborations in the near future. I also would like to express my very great appreciation to Alexandre Barreira, whose crucial suggestions were of great help in identifying an important weakness in my very first scientific manuscript. My sincere thanks also go to Patrick Peter, my thesis Godfather, who was always ready to sacrifice part of his precious time whenever a young researcher like me needed a small piece of advice. Thank you for your impeccable balance between professionalism and deep human understanding. This PhD study would not have been possible without the technical support of Stephane Rouberol, who smoothly runs the cluster Horizon for us all, and the meticulous work of all the people who alleviated the administrative burden entailed by my activity as a researcher. In this sense, I would like to convey my gratitude to Alexandre Dieng, Amandine Guillemois, Chantal Le Vaillant, Jason Payet, Nitaya Singsengsouvanhm, Olivia Leroy and Roselys Rakotomandimby. Furthermore, I acknowledge financial support from the ILP LABEX (under reference ANR-10-LABX-63), financed by French state funds managed by the ANR within the Investissements d'Avenir programme under reference ANR-11-IDEX-0004-02, and from the Centre National d'Etudes Spatiales (CNES) fellowship program

The PhD is not just an academic experience. The three years I spent in IAP, in Paris, gave me the chance to discover and strengthen deep friendship bounds which will be everlasting. Doogesh, thank you for all the support you have been providing me with since the first day of our PhD, a support which is at the same time human and professional. I found a fantastic person in you. A voice echoing from the back of office 035 makes me remind (of course this is a joke!) of the other great friend I made here: the untouchable Big A(rno). Thank you for your lessons on coding performance and on the healthy properties of oranges in the morning! It was a pleasure to share the office also with Martin Pernot-Borras: your peculiar humour will always be remembered along with your never ending-Friday morning loquacity on politics and french social welfare. Thank you Pierre Boldrini and Shweta Dalal for all the light-hearted moments we spent together. I hope one day we will finally have a pasta at Shweta's place! I will also never forget Alexandre Barthélémy (my Euclid colleague and half-brother), Lukas Furtak and all our coffee breaks together: a daily guaranteed moment of joy.

I am grateful to my mother Francesca and my father Stefano, who have provided me moral and emotional support not only during this difficult experience, but also in every single choice I made in my life. Special thoughts for my grandmother Ofelia who has always been caring even in the distance, even in the sickness. My deepest appreciation goes to Veronica who became one of the most important persons in my life. Finally, I will always be indebted to all my friends from Italy, so far and yet so close, for their unconditional affection and for always being ready to warmly welcome me back home. Our friendships will always be unbreakable. Of course, thank you too, small Matisse, for keeping me awake even in the most peaceful nights.

Cosmologie Non-Gaussienne : défis théoriques et statistiques pour les relevés de galaxies modernes

Résumé

Nous détaillons dans cette thèse les différentes étapes nécessaires à l' implémentation numérique optimale du calcul de la vraisemblance des paramètres cosmologiques appliqué aux relevés modernes de lentillage gravitationnel faible des grand structures. En particulier, nous nous concentrons sur la détection conjointe du spectre de puissance et du bispectre du lentillage gravitationnel faible. Pour ce faire, nous avons relevé les défis numériques requis par une analyse complète. Dans un premier temps, nous dressons l'état de l'art nécessaire à la compréhension du formalisme de la sonde cosmologique susmentionnée et nous décrivons comment obtenir une estimation non biaisée de la matrice de covariance pour les observables considérées. En supposant une vraisemblance gaussienne multivariée, nous avons développé un algorithme de haute performance permettant de prédire les observables tomographiques rééchantillonnées avec leur matrice de covariance conjointe en tenant compte des fonctions de corrélations de 2 à 6 points et des corrélations avec les modes hors du relevé. La performance de notre code nous permet de répondre aux exigences scientifiques des relevés de galaxies des dix prochaines années. Nous montrons que le bispectre du lentillage gravitationnel faible améliore le signal sur bruit (S/N) de notre analyse conjointe du spectre et du bispectre d'environ 10% en comparaison avec une analyse du spectre seulement. Par conséquent, le bispectre est une source non négligeable d'information cosmologique pour les relevés futurs. En outre, nous sommes capables de quantifier l'impact des incertitudes théoriques liées à la description de la matière noire dans le "modele des halos" qui est utilisé pour construire nos observables; cet impact se trouve être négligeable pour l'analyse du S/N. Finalement, nous étudions la possibilité de réduire les données pour optimiser les analyses futures du bispectre du lentillage. Nous trouvons qu'en ignorant les erreurs systématiques une analyse ne concernant que 5 échantillons en décalage vers le rouge permet d'obtenir la même quantité d'information que dans le cas d'un relevé semblable à celui d'Euclid, qui utilise 10 échantillons sans améliorer le S/N. Nous explorons également l'analyse en composantes principales et la dépendance de l'information cosmologique en fonction de la forme des triangles comme méthodes permettant de réduire la complexité du problème.

Non-Gaussian Cosmology: theoretical and statistical challenges for modern galaxy surveys

Abstract

In this thesis, we address key points for an efficient implementation of likelihood codes for modern weak lensing large-scale structure surveys. Specifically, we will focus on the joint weak lensing convergence power spectrum-bispectrum probe and we will tackle the numerical challenges required by realistic analyses. In order to clearly convey the importance of our research, we first provide an in-depth review of the background material required for a comprehensive understanding of the final results. The cosmological context of the study is provided, followed by a description of the technical elements inherent to unbiased covariance matrix estimation for the probe considered. Under the assumption of multivariate Gaussian likelihood, we developed a high performance code that allows highly parallelised prediction of the binned tomographic observables and of their joint non-Gaussian covariance matrix accounting for terms up to the 6-point correlation function and super-sample effects. This performance allows us to qualitatively address several interesting scientific questions. We find that the bispectrum provides an improvement in terms of signal-to-noise ratio (S/N) of about 10% on top of the power spectrum alone, making it a non-negligible source of information for future surveys. Furthermore, we are capable to address the impact of theoretical uncertainties in the halo model used to build our observables; with presently allowed variations we conclude that the impact is negligible on the S/N. Finally, we consider data compression possibilities to optimise future analyses of the weak lensing bispectrum. We find that, ignoring systematics, 5 equipopulated redshift bins are enough to recover the information content of a Euclid-like survey, with negligible improvement when increasing to 10 bins. We also explore principal component analysis and dependence on the triangle shapes as ways to reduce the numerical complexity of the problem.

Contents

Introduction

Understanding the present accelerated expansion of the Universe is one of the most important challenges in modern Cosmology. In order to provide a theoretical explanation for this incredible observation, a plethora of theoretical model have been designed over the last few years, some of which introduce corrections to the standard cosmological scenario ΛCDM. This diversity of theoretical ideas shows our current ignorance and defines the need for new observations. Based on our present-day knowledge, existing plausible models will impact observational signatures by tiny amounts that can only be decisively distinguished using high-precision astronomical surveys covering a major fraction of the sky and great cosmological distances. In preparation for future missions like Euclid, Dark Energy Survey (DES, Flaugher [2005]), Large Synoptic Survey Telescope (LSST, Ivezić et al. [2019]) or the Wide Field Infrared Survey Telescope (WFIRST, Green et al. [2011]), we need to understand the performance of the survey observations in terms of cosmological parameter error forecasts for deviations from the ΛCDM model.

Among the different ways that we can look at the light coming from the Universe, weak lensing is an extremely powerful cosmological probe. We exploit the detection of light from distant sources to quantify the emergent cross-correlations between the observed shapes of distant galaxies induced by the underlying total matter distribution. More importantly, weak lensing does not require any knowledge of the relation between the observed distribution of galaxies and the underlying dark matter one. For this reason, this cosmological probe is not just important by itself, but it can actually be combined with the inferred galaxy distribution in redshift and at different scales (galaxy clustering) to improve the scientific performance of our analyses. This is indeed the strategy of the Euclid and WFIRST missions. However, in order to exploit the information encoded in this probe, we have to deal with the non-Gaussian statistics induced by the non-linear clustering of the matter field being investigated. Therefore, it is necessary to go beyond the 2-point statistics (which is sufficient to fully characterise a Gaussian field) to recover the cosmological information left behind. One possible solution is via the estimation of the 3-point correlation function. In configuration space, these statistical tools are respectively named power spectrum and bispectrum. On top of that, when we perform a combined measurement of the two (even if complementary) we are actually probing the same underlying physics. Therefore, these observables are correlated and we need to account for this crucial aspect in order not to double count the information contained.

In terms of exploitation of the observed weak lensing bispectrum, modern cosmological surveys, such as those

listed above, represent a challenge for various reasons. Their capacity to detect the binned redshift position of galaxies and to observe the sky under a wide angular range are key ingredients to have access to a precious time- and scale-dependent information on the matter clustering. On the other hand, its exploitation comes at a cost: an overwhelming number of observables must be measured and along with them their errors and cross-correlations must be taken into account. Covariance matrices encapsulate this information. For this reason, data analyses and forecasts for future galaxy surveys are expected to be extremely complex and computationally costly.

A second problematic aspect of future analyses is related to the underlying statistical approach employed in some circumstances. An exact cosmological forecast requires the exploration of the parameter posterior via sampling techniques like Markov Chains Monte Carlo or nested sampling methods. This allows us to identify best fit values, to visualise confidence regions and correlations, and eventually to compute marginal and conditional probabilities. Over the past decades, analytical approaches have been introduced in order to avoid the full posterior exploration. This can be very useful whenever we have limited computational resources or when forecasting parameter constraints of a future experiment and we want to test different configurations of it. The Fisher approximation is a famous example. With this approach, the posterior is approximated around the peak via a multivariate Gaussian and it allows an analytical determination of the uncertainties on the parameters. However, the level of non-linearity present in the physical process we are interested in may spoil the quality of this assumption. Another weakness of the Fisher approximation is the need for computing the derivatives of the observables with respect to the cosmological parameters of interest. While we are currently capable of reproducing the amplitude of a given signal up to certain scales, we still lack a way to reproduce their derivatives at high enough precision so as not to bias our cosmological analyses. Therefore, we would like to find a trade-off between the rapidity typical of Gaussian-based techniques and the correctness of the sampling techniques in terms of posterior exploration.

In my doctoral work, I tackled these key points which will hinder, unless deeply understood and solved, the scientific success of upcoming galaxy surveys. The aim of this research is to provide definite answers to the following two questions:

- *How can we optimise the computation of the covariance matrices (errors and correlations in the data sets) for future weak lensing measurements beyond the 2-point statistics?*
- *How can we go beyond the standard data analysis approaches in Cosmology?*

The content in this manuscript has been previously presented in the main paper produced during my research at *Institut Astrophysique de Paris* (IAP)

> Rizzato M., Benabed K., Bernardeau F., Lacasa F.,
> *Information content of the weak lensing bispectrum for the next generation of galaxy surveys,*
> 2018, arXiv: 1812.07437
> **Submitted to MNRAS.**

The first two chapters provide a thorough review of the background material required for understanding the results, which are presented in Chapter 3. Part of the introductory sections on the topic of covariance matrix estimation within Chapter 2 has been inspired by the discussions engendered from the following collaborative work

> Sellentin E., Joachimi B., Barreira A., Giocoli c., Heavens A., Kilbinger M.,
> Kitching T., Rizzato M., Schmidt F., Taylor A.
> *Euclid Weak Lensing Covariance Task Force Report,*
> **Internal Euclid paper**.

In the above manuscript, I wrote as main author an important section regarding the performance of the halo model for the analytical derivation of the power spectrum covariance matrix for the weak lensing convergence.

Conclusions and perspectives derived from my doctoral activity are presented in the back-matter of this thesis. They are drawn from the work related to the paper above and from other projects which were the focus of my last months at IAP

> Reimberg P., Bernardeau F., Nischimichi T., Rizzato M.,
> *Failures of Halofit model for computation of Fisher Matrices,*
> 2018, arXiv: 1811.02976.
> **Submitted to MNRAS.**

> Rizzato M., Sellentin E.,
> *Analytic reconstruction of non-Gaussian posteriors for precision cosmology*
> **Upcoming publication.**

While the content of these papers will not be included in the present manuscript, they were an occasion of enrichment for my scientific awareness regarding the problematic points outlined in this introduction. Also, they will be a major source of inspiration for my future activity as a researcher.

I wish you a pleasant journey.

Matteo Rizzato

1

Our description of the observable Universe

Contents

The goal of this chapter is to provide the reader with some useful concepts in order to understand and contextualise the study presented in the following parts of this manuscript.

We would like to underline that the theoretical background employed in this work will not account for Modified Gravity corrections [Clifton et al., 2012; Ishak, 2019]. Therefore, the results and the conclusions derived, both at the level of background material and at the level of original research, are meant to be valid within the context of a standard general relativistic scenario. However, we leave room for future research aiming at the extension of the efforts here presented to more complex and less standard scenarios.

In line with the *cosmological principle* we will assume that the Universe is statistically homogeneous and isotropic at every time on sufficiently large scales. Specifically, 1) different positions in the Universe are physically indistinguishable and 2) observations performed in different directions, from no matter which point in the Universe (point 1), leads to the same physical observations. These are strong conceptual assumptions: however, we have sound evidences for their validity [Hogg et al., 2005; Ntelis et al., 2017] and they strongly bind (help) the mathematical description of the observable Universe around us. This description will be indeed the focus of Sec. 1.1.

Clearly, a perfectly homogeneous Universe is a model too simplistic for the description of the cosmological observations at smaller scales. The simple existence of this higher dense paper page and of a the lower dense air layer in front of it tells us that indeed it is possible to distinguish two points in space $P_1,\ P_2$ (and time) via, for example, a simple measurement of the matter density field, $\rho_{P_1},\ \rho_{P_2}$. These tiny space-dependent fluctuations are indeed present and of paramount importance for cosmological observations. Usual practice in Cosmology is to separate a given field defined on the space-time $f(\mathbf{x}, t)$ in its background-homogeneous component $f(t)$ and in a "perturbation part" $\delta f(\mathbf{x}, t)$ which is defined to be vanishing, at a given time, if an ensable average is performed on it, i.e. $\langle \delta f(\mathbf{x}, t) \rangle \equiv 0$. At the level of notation we will distinguish between the total field and the background component of it via the dependence on the coordinates, being solely on the time for the latter. More often, the dimensionless *contrast density field* is used

$$\delta_f(\mathbf{x}, t) = \frac{f(\mathbf{x}, t) - f(t)}{f(\mathbf{x}, t)}. \tag{1.1}$$

While the observable Universe has been extensively exploited at the level of background expansion [e.g. Kowalski et al., 2008; Amanullah et al., 2010, and references therein], perturbations at smallest scales are actually precious sources of cosmological information for the present and next generations of galaxy surveys (e.g. Collaboration [2005]; LSST Science Collaboration et al. [2009]; Laureijs et al. [2011]; DESI Collaboration [2016]). While the actual matter perturbations δ_m can not be directly accessed, other fields can be employed as proxies for the actual underlying matter distribution, such as the number density of galaxies δ_g or the shape distortion of far sources due to intermediate matter distribution (weak lensing convergence field κ or the weak lensing shear field γ). However, a single measurement of the properties of these *tracers* have a very low *signal-to-noise ratio* and several detections of the same phenomena have to be performed. In addition, the same initial perturbations imprinted in the *cosmic*

microwave background (CMB) [Kamionkowski & Kosowsky, 1999; Hu & Dodelson, 2002; Samtleben et al., 2007] and seeding the evolution of the observed *large-scale structure* (LSS) [Peebles, 1980; Dodelson, 2003], have an intrinsic stochasticity originating from quantum fluctuations of a scalar field. All these facts motivate a statistical approach to the description of the cosmological fields and we will review the main aspects of the related formalism in Sec. 1.2.

As for the matter field, its anisotropies are originally imprinted in the temperature map of the *cosmic microwave background* where $\delta T/T \sim 10^{-5}$ [Planck Collaboration, 2014, 2016, 2018] . Due to a strong coupling between photons and matter particles, they also seed the fluctuations of the matter field $\delta_{\rm m}\left(\mathbf{x}, t_{\rm CMB}\right) \ll 1$ at the time at which the CMB is produced. Starting from this moment, the matter perturbations $\delta_{\rm m}\left(\mathbf{x}, t > t_{\rm CMB}\right)$ evolve in time, the value of which will be determined by the gravitational collapse across the different cosmological epochs. In general, the associated dynamical equations are highly non-linear. However, at early time or if smoothed on large scales the matter perturbations are small enough to allow a linearization of the dynamical equations around the background homogeneous solution. At small scales and at late time, the non-linear nature of the equations of motion can not be hidden and theoretical efforts are requested to produces a description as much precise as possible of the late time matter distribution. In Sec. 1.3 we will review the model for non-linear clustering employed throughout the work presented in this manuscript.

We would like to comment on one last point before letting the reader start his journey. As the title of this chapter does underline, our results and review material are meant to apply only to that part of the (light-cone) of the Universe which can be observed by us, the *Observable Universe*. In particular (and we will give a more formal definition later) it is the region around us within which events have had the time to be in causal contact with us by today, starting from the origin of the Universe.

1.1 The homogeneous Universe

1.1.1 The Einstein equations

The definition of the cosmological observables relies on a robust formalism that has deep roots in the theory of General Relativity, first formalised in 1916 by Albert Einstein. While several attempts have been recently explored aiming at relaxing some of its assumptions to accommodate for more complex scenarios, we will focus on its original formulation which is anyway well tested and compatible with most of the present observations within the achievable instrumental precision [Bull et al., 2016].

Our starting points are the well known Einstein equations

$$R_{\mu\nu}(x) - \frac{1}{2} g_{\mu\nu}(x) R(x) = \frac{8\pi G}{c^4} \sum_i T^{(i)}_{\mu\nu}(x), \quad x = \{\mathbf{x}, t\} \tag{1.2}$$

All the tensors involved are local, evaluated at a point of coordinates $x = \{\mathbf{x}, t\}$ of the space-time, and have Greek indices taking values $0, 1, 2, 3$, the 0^{th} component referring to the time component. In the present section and throughout this manuscript, we will refer to spatial vectors (either three- or two-dimensional) by writing them in bold italic. The modules of these vectors will instead be written in regular italic, i.e. $x = |\mathbf{x}|$. On the left hand side (l.h.s.) we can identify the *Ricci tensor* $R_{\mu\nu}(x)$ and the *Ricci scalar* $R(x)$, obtained by contracting the former with the metric $g_{\mu\nu}(x)$. On the right hand side (r.h.s.) we have the Newton's gravitational constant G, the speed of light c and the sum over the *energy-momentum tensors* of the different species populating the Universe. For sake of completeness, we remind that the Ricci tensor is defined in the terms of the Riemann tensor as

$$R_{\mu\nu}(x) = R^{\alpha}_{\mu\alpha\nu}(x) = R^{\alpha}_{\alpha\mu_{\nu}}(x), \tag{1.3}$$

which is in turns defined as

$$R^{\mu}_{\nu\rho\sigma}(x) = \partial_{\rho}\Gamma^{\mu}_{\sigma\nu}(x) + \Gamma^{\mu}_{\rho\beta}(x)\Gamma^{\beta}_{\sigma\nu}(x) - \partial_{\sigma}\Gamma^{\mu}_{\rho\nu}(x) + \Gamma^{\mu}_{\sigma\beta}(x)\Gamma^{\beta}_{\rho\nu}(x), \tag{1.4}$$

where $\Gamma^{\alpha}_{\beta\rho}(x)$ is the so called *affine connection*.

$$\Gamma^{\alpha}_{\beta\rho}(x) = \frac{1}{2}g^{\alpha\lambda}(x)\left(\partial_{\beta}g_{\lambda\rho}(x) + \partial_{\rho}g_{\lambda\beta}(x) - \partial_{\lambda}g_{\beta\rho}(x)\right), \qquad \partial_{\alpha} = \frac{\partial}{\partial x^{\alpha}} \tag{1.5}$$

We implied, and we will always do, a sum over repeated indices. The request for isotropy and homogeneity, along with the assumption of describing the different components of the Universe via a perfect fluid model, bounds the shape of the energy-momentum tensor to be

$$T^{(i)}_{\mu\nu}(x) = \left(P_i(t) + \rho_i(t)\right)u_{\mu}u_{\nu} - g_{\mu\nu}(x)\,P_i(t). \tag{1.6}$$

$P_i(t)$ and $\rho_i(t)$ are respectively the pressure and the energy density of the given fluid i^{th} as function of time and u_{μ} is the four-dimensional velocity of the particles relative to the fluid at rest.

1.1.2 Locating objects around us: Friedman-Lemaître-Robertson-Walker metric

Cosmological observations require the capacity to measure distances $\{\Delta\mathbf{x},\ \Delta t\}$ in the Universe between different events, and between us, observing, and some event of interest. The symmetries highlighted in the introduction of the present chapter impose a specific formal shape to the metric $g_{\mu\nu}(x)$. The most general one compatible with a four-dimensional space embedding a three-dimensional sub-space invariant under translations and rotations was independently found by Alexander Friedman and Georges Lemaître, and rigorously derived by Howard P. Robertson and Arthur Geoffrey Walker [Lemaître, 1931, 1933; Robertson, 1935, 1936a,b]. Given a reference frame of axes $\hat{x}_1, \hat{x}_2, \hat{x}_3$ and a set of polar coordinates (r, θ, φ)[1] the Friedman-Lemaître-Robertson-Walker metric

[1] Throughout the manuscript we will always define θ and φ to be respectively the *polar angle* and *azimuth angle*. Therefore $(x_1 = r\sin\theta\cos\varphi, x_2 = r\sin\theta\sin\varphi, x_3 = r\cos\theta)$.

(FLRW) is defined via the following line element

$$g_{\mu\nu}^{\mathrm{FLRW}}(x)dx^{\mu}dx^{\nu} \equiv c^{2}\,dt^{2} - a^{2}(t)\left[\frac{dr^{2}}{1-K^{(3)}r^{2}} + r^{2}\left(d\theta^{2} + \sin^{2}\theta d^{2}\varphi\right)\right] \equiv c^{2}\,dt^{2} - a^{2}(t)d\ell^{2}. \tag{1.7}$$

From now, our conclusions will always be derived (unless differently stated) under the assumption of a FLRW metric: for simplicity we will omit this specification. In Eq. (1.7), we introduced the scale factor $a(t)$, the $K^{(3)}$ space curvature and the three-dimensional, *time independent comoving line element* $d\ell$. Depending on the sign of $K^{(3)}$, this three-dimensional sub-space is locally isometric to a sphere ($K^{(3)} > 0$), to a pseudo-sphere ($K^{(3)} < 0$) or to a flat Euclidean space ($K^{(3)}$=0). The scale factor $a(t)$ describes the contribution from the expansion of the Universe to the dynamics of the physical separation between 2 points. If this contribution is not considered, we are specifically talking about the *comoving distance* between the aforementioned objects and it is related to the physical one simply as

$$d_{\mathrm{phy.}} = a\,(t)\,d_{\mathrm{com.}}. \tag{1.8}$$

In the following, we will assume that the scale factor and the radial coordinates are normalised in such a way that the values of $K^{(3)}$ can be either -1, 0 or 1, i.e.

$$a\,(t) \to \frac{a\,(t)}{|K^{(3)}|}, \qquad [a\,(t)] \sim \text{length}, \tag{1.9}$$

$$r \to r|K^{(3)}|, \qquad r \text{ is dimensionless}. \tag{1.10}$$

In these specific cases the comoving three-dimensional line element will be: 1)

$$d\ell^{2} = d^{2}r + r^{2}\left(d\theta^{2} + \sin^{2}\theta d^{2}\varphi\right) = d^{2}x_{1} + d^{2}x_{2} + d^{2}x_{3}, \quad \text{for } K^{(3)} = 0, \tag{1.11}$$

which is metric for a flat Euclidean space, 2)

$$d\ell^{2} = \frac{dr^{2}}{1-r^{2}} + r^{2}\left(d\theta^{2} + \sin^{2}\theta d^{2}\varphi\right) = d^{3}\omega + \sin^{2}\omega\,\left(d\theta^{2} + \sin^{2}\theta d^{2}\varphi\right) \quad \text{for } K^{(3)} = +1, \tag{1.12}$$

where in the second part of the equation we defined $r = \sin\omega$, 3)

$$d\ell^{2} = \frac{dr^{2}}{1+r^{2}} + r^{2}\left(d\theta^{2} + \sin^{2}\theta d^{2}\varphi\right) = d^{2}\omega + \sinh^{2}\omega\left(d\theta^{2} + \sin^{2}\theta d^{2}\varphi\right) \quad \text{for } K^{(3)} = -1, \tag{1.13}$$

where in the second part of the equation we defined $r = \sinh\omega$. Respectively, Eq. (1.12) and Eq. (1.13) define the metric for a three-dimensional space with constant positive and negative curvature. In a synthetic way, we can

summarise the previous cases via a redefinition of the radial coordinate

$$g_{\mu\nu}(x)dx^{\mu}dx^{\nu} = c^2\,dt^2 - a^2(t)\left[d\omega^2 + \mathcal{F}_K^2(\omega)\,d^2\Omega\right],$$

$$d\omega = \frac{dr}{\sqrt{1-K^{(3)}r^3}}, \qquad \mathcal{F}_K(\omega) = \begin{cases} \sin\omega & \text{if } K^{(3)} = +1, \\ \omega & \text{if } K^{(3)} = 0, \\ \sinh\omega & \text{if } K^{(3)} = -1. \end{cases}, \tag{1.14}$$

where $d^2\Omega = (d\theta^2 + \sin^2\theta d^2\varphi)$. This last notation will be useful for defying cosmological distance in a geometry-independent way. Given that the scale factor $a(t)$ is commonly used as a time variable itself, we will often omit its time dependence. This will also help in simplifying the required equations in the following.

The Friedmann equations The Einstein Equations can be written in terms of the FLRW metric. More importantly, Eq. (1.2) can be written as a system of 2 independent equations, known as Friedman equations

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\sum_i \left(\rho_i(t) + 3P_i(t)\right) = -\frac{4\pi G}{3}\sum_i \rho_i(t)\left(1 + 3\omega_i(t)\right), \tag{1.15}$$

$$H^2(t) = \frac{8\pi G}{3c^4}\sum_i \rho_i(t) - \frac{K^{(3)}}{a^2}, \tag{1.16}$$

where $H(t) = \frac{\dot{a}}{a}$ is the so called Hubble parameter, which usually written as

$$H(t) = H_o E(t) \tag{1.17}$$

and $\omega_i = P_i/t_{\rm i}$ is the *equation of state* for a specific component of the Universe. As commonly indicated in the literature, physical quantities with sub-script o refers to the present value of the corresponding quantity. In the specific case of the Hubble parameter, its present value is usually given in terms of the dimensionless parameter h as

$$H_o \equiv 100\,h\,\mathrm{Km\,s^{-1}\,Mpc^{-1}}. \tag{1.18}$$

Along with Eq. (1.15) and Eq. (1.16), we can introduce the energy conservation equations for the different components in an expansing Universe

$$\frac{\mathrm{d}\rho_i(t)}{\mathrm{d}t} = -3\left(\rho_i(t) + P_i(t)\right)\frac{\dot{a}}{a}. \tag{1.19}$$

For a constant equation of state ω_i, it can be easily solved leading to the dynamical evolution for the component of interest

$$\rho_i(t) = \rho_{i,\rm o}\left(\frac{a_{\rm o}}{a}\right)^{3(1+\omega_i)} \tag{1.20}$$

Just to give a couple of examples, pressure-less matter has $\Omega_{\rm m} = 0$ while radiation has $\omega_{\rm r} = 1/3$. Throughout this

manuscript, we will mainly interested in the former.

Density parameters For later convenience, we can define a dimensionless density parameter for each component in the Universe quantifying its fractional density in time

$$\Omega_i(t) \equiv \frac{8\pi G}{3H^2(t)}\rho_i(t)\,. \tag{1.21}$$

Given this last definition, the spatial curvature of the Universe is in general quantified by the following dimensionless density parameter

$$\Omega_k(t) \equiv 1 - \sum_i \Omega_i(t) = -\frac{K^{(3)}}{a^2H^2(t)}. \tag{1.22}$$

Finally, another important concept worth to introduce is the one of *critical density* ρ^c as

$$\rho^c(t) \equiv \frac{3H^2(t)}{8\pi G}, \qquad \Omega_i(t) \equiv \frac{\rho_i(t)}{\rho^c(t)}. \tag{1.23}$$

We would like to conclude this section with a last remark on our conventions. For a spatially-flat Universe, we will use a dimensionless scale factor normalised to its present value, therefore $a_o = 1$. In particular, the FLRW metric expression introduced in Eq. (1.7) is preserved only if $r \to ra_o$ and $K^{(3)} \to K^{(3)}/a_o^2$, the latter not being equal anymore to only $\pm 1, 0$.

1.1.3 Cosmological distances

Defining and measuring astronomical observables, more importantly the weak lensing ones described in Chapter 2, requires the localisation of an object in the observed sky. While we can easily define a polar coordinate system to locate them on the sphere centred in the observer, the definition for a suitable third coordinate to describe its distance from the same point or from other objects is more ambiguous. In the present section, we will review the main definitions we will need for a deep understanding of this work [Weinberg, 1972; Weedman, 1986; Peebles, 1980].

Redshift Commonly labelled as z, the *redshift* has a simple operational definition as the shift of spectral lines due to expansion of the Universe. If a specific photon of a given a physical (non-comoving) wavelength λ_e is a emitted at a time t_e, then the observed physical wavelength λ_o at time t_o will be

$$\lambda_o = \lambda_e \frac{a(t_o)}{a(t_e)}, \tag{1.24}$$

leading to the following definition of redshift

$$z_{t_o}\left(t_e\right)+1\equiv\frac{a\left(t_o\right)}{a\left(t_e\right)}. \tag{1.25}$$

In particular, the redshift is a quantity we can measure from observations. From now on we will drop the subscript t_o assuming that the observation always occur at present time.

Angular diameter distance Given an object at redshift z, of physical linear size Δ and subtended by an angle θ if observed from a point at redshift $z=0$, the *angular diameter distance* is defined as

$$D_A\left(z\right)\equiv\frac{\Delta}{\theta} \tag{1.26}$$

where clearly both Δ and θ have a dependence on the redshift of the object. In a FLRW metric, we can express the angular diameter distance as

$$D_A\left(z\right)=\begin{cases}\frac{c}{H_o(1+z)\sqrt{|\Omega_{K,o}|}}\mathcal{F}_K\left(\sqrt{|\Omega_{k,o}|}\int_o^z\frac{dz'}{E(z')}\right), & \text{if } K^{(3)}\neq 0,\\ \\ \frac{c}{H_o(1+z)}\int_o^z\frac{dz'}{E(z')}, & \text{if } K^{(3)}=0.\end{cases} \tag{1.27}$$

where the functions $\mathcal{F}_K$ and E were respectively defined in Eq. (1.14) and in Eq. (1.17).

Longitudinal comoving distance The *longitudinal comoving distance*, also known as *line-of-sight distance*, between two events assumed aligned on the three-dimensional sphere and occurring at times z_1 and z_2, is defined as the distance covered by the light while travelling between them. Given the metric in Eq. (1.14), it can be easily derived via the integration of the radial comoving coordinate ω along null-geodesics (i.e. $d^2s=0$)

$$\chi\left(z_1,z_2\right)=\frac{c}{H_o}\int_{z_1}^{z_2}\frac{dz'}{E\left(z'\right)}. \tag{1.28}$$

For later convenience, it is worth to notice that in a spatially-flat Universe the *angular distance* to the point at redshift z_2 in the *comoving* coordinates of the point at redshift z_1 coincides with the *longitudinal comoving distance* between the 2 points

$$D_A\left(z_1,z_2\right)\big|_{K^{(3)}=0,\,z_1}=\chi\left(z_1,z_2\right) \tag{1.29}$$

and Eq. (1.29) is symmetric in the two redshifts.

We can also define a physical distance associated to the recently introduced comoving one. The *proper distance* to an event at z_2 measured from an event at z_1, will be given by the comoving distance between the two multiplied

by the scale factor $a(z_1)$, i.e.

$$d_{\mathrm{p}}^{z_1}(z_2) = a(z_1)\chi(z_1, z_2) \tag{1.30}$$

Therefore, we are now capable to define our *Observable Universe*, term which inspired the title of this chapter: it will be simply given by our proper distance up to the Big Bang, at time $z \to \infty$. In particular, the size of our observable Universe is of about 14.4 Gpc.

1.1.4 ΛCDM and Einstein-De-Sitter cosmological models

Given the formalism developed so far, by solving the Friedman equations (1.15) and (1.16) for a given evolution of the density parameter Ω_i, we can derive a specific solution for the scale factor $a(t)$. Two FLRW models of Universe are of particular interest.

ΛCDM model The ΛCDM (Λ Cold Dark Matter) model is a 6-parameter FLRW-model defining an Universe populated by collision-less, pressure-less, non-relativistic (*cold*) particles of matter which can either have only gravitational interactions (*dark matter* Ω_{d}) or all the known ones. In the latter cases, they are usually dubbed as *baryons* (Ω_{b}). Both these types of matter have the same equation of state, namely $\omega_{\mathrm{m}} = 0$ (pressure-less). However, they are present with different density fractions, respectively $h^2\Omega_{\mathrm{d}} = 0.120 \pm 0.001$ and $h^2\Omega_{\mathrm{b},0} = 0.0224 \pm 0.0001$, according to the latest observations[2] [Planck Collaboration, 2018]. The Greek letter Λ stands for an exotic component with a constant negative equation of state $\omega_\Lambda = -1$ and density of $\Omega_{\Lambda,0} = 0.6889 \pm 0.0056$ [Planck Collaboration, 2018].

The model is also defined by further requests on the *primordial Universe* such as the the primordial fluctuations seeding the CMB anisotropies being *Gaussian* distributed, *adiabatic* and with an almost *scale-invariant power spectrum*. While referring to major reviews for all the details concerning the definition of the *ΛCDM* model, its parameters and extensions [Planck Collaboration, 2014, 2016, 2018, and references therein], we reserve the right to come back to these concepts later, when more tools, specifically the concept of *correlation functions*, will be introduced.

Einstein-De-Sitter The Einstein-De-Sitter model (EdS) is a FLRW-model defining an Universe entirely dominated by matter, i.e. ($\Omega_{\mathrm{m}} = 1$), with no curvature ($K^{(3)} = 0$). Even though this model does not give the right description of our observable Universe, however its simplicity allows to derive several analytical results with regards to the growth of the *large-scale structure* and which are proved to be a good approximation to th actual dynamics occurring in the ΛCDM scenario.

[2] The errors reported correspond to the 1σ confidence region.

1.2 The description of cosmological fields

Given the motivations outlined in the introduction, in this section we will provide the tools to describe the statistical properties of a generic random field $\delta_{\rm f}(\mathbf{x})$ (in this section we will omit the time dependence of the generic random field). We will then move to the more specific case of the cosmological fields.

1.2.1 Ensemble average

We are dealing with a random field whenever the values of the field $\delta_{\rm f}(\mathbf{x})$ at a given set of points are random variables.

The *$n-$points joint probability* $P_{\delta_{\rm f}}([a_n], [\mathbf{x}_n])$ is defined as the probability that the field $\delta_{\rm f}$ takes values $\delta_{\rm f}(\mathbf{x}_i) \in [a_i, a_i + \delta a_i]$ at the point set $\{\mathbf{x}_i\}_{i=1,\dots,n}$. A formal definition can be given

$$P_{\delta_{\rm f}}([a_n], [\mathbf{x}_n]) = \left\langle \prod_{i=1}^{n} \delta_{\rm D}\left(\delta_{\rm f}(\mathbf{x}_i) - a_i\right) \right\rangle_{\delta_{\rm f}} \tag{1.31}$$

where the operator $\langle \star \rangle_{\delta_{\rm f}}$ stands for the *ensemble average* associated to the specific field of interest. In analogy with the probability density distribution for random variables, we can define the *probability density functional* $\mathcal{P}$ over the space of the possible realisations $\hat{\delta}_{\rm f}$ of the field $\delta_{\rm f}$. Similarly, it allows to formally define the ensemble average of a functional operator $\mathcal{O}[\delta_{\rm f}]$ of the field as

$$\langle \mathcal{O}[\delta_{\rm f}] \rangle_{\delta_{\rm f}} = \int \mathcal{D}[\hat{\delta}_{\rm f}]\, \mathcal{P}_{\delta_{\rm f}}[\hat{\delta}_{\rm f}]\, \mathcal{O}[\hat{\delta}_{\rm f}]. \tag{1.32}$$

A more operational definition can be given thanks to the concept of $n-$points joint probability introduced in Eq. (1.31). Given one realisation $\hat{\delta}_{\rm f}$ of the field $\delta_{\rm f}$, the value of the probability density functional for the specific realisation can be built as the limit for $N \to \infty$ of the $N-$points joint probability for the value of that realisation once discretized over a grid of N points [Mcclelland & Silk, 1978]. Therefore, the formal definition (1.31) can actually be rephrased in a more operative way as

$$\langle \mathcal{O}[\delta_{\rm f}] \rangle_{\delta_{\rm f}} = \lim_{N\to\infty} \prod_{l,m,n=-N}^{+N} \int_{-\infty}^{+\infty} d\,\hat{\delta}_{\rm f}^{lmn}\, P[\hat{\delta}_{\rm f}^{lmn}] \star [\hat{\delta}_{\rm f}^{lmn}]. \tag{1.33}$$

The quantities $\{\hat{\delta}_{\rm f}^{lmn}\}_{l,m,n=-N,\dots,+N}$ are simply the values of the realisation $\hat{\delta}_{\rm f}$ over the nodes of the grid labelled by the set of indices l, m, n. Therefore, the quantity $P\left[\hat{\delta}_{\rm f}^{lmn}\right]$ has to be understood as the $n-$points joint probability for the field to takes values in the range $[\hat{\delta}_{\rm f}^{lmn}, \hat{\delta}_{\rm f} + \delta\hat{\delta}_{\rm f}^{lmn}]_{l,m,n=-N,\dots,+N}$ over the nodes of the grid.

1.2.2 Correlation functions

Formal definition In order to fully characterise the statistical properties of the matter density field (later simply noted as $\delta_{\rm m}$), we choose to follow the traditional approach of using the n-point correlation functions [Peebles, 1980; Bertschinger, 1992; Matsubara, 1995], which are defined as the joint ensemble average of the field in an arbitrarily number of locations $\{\mathbf{x}_i\}_{i=1,\dots,n}$, formally

$$\varepsilon_{\delta_{\rm f}}\left(\mathbf{x}_1,\dots,\mathbf{x}_n\right) \equiv \left\langle \delta_{\rm f}\left(\mathbf{x}_1\right)\dots\delta_{\rm f}\left(\mathbf{x}_n\right)\right\rangle_{\delta_{\rm f}}. \tag{1.34}$$

In analogy with the *moments* for real variables, we define the *characteristic functional* of a random field $\delta_{\rm f}$ as

$$\Phi_{\delta_{\rm f}}[k] \equiv \int \mathcal{D}\left[\hat{\delta}_{\rm f}\right]\, \exp\left[i\int d^3\mathbf{r}\, k\left(\mathbf{r}\right)\, \hat{\delta}_{\rm f}\left(\mathbf{r}\right)\right] \mathcal{P}_{\delta_{\rm f}}\left[\hat{\delta}_{\rm f}\right] = \left\langle \exp\left[i\int d^3\mathbf{r}\, k\left(\mathbf{r}\right)\, \hat{\delta}_{\rm f}\left(\mathbf{r}\right)\right]\right\rangle_{\delta_{\rm f}}. \tag{1.35}$$

It can be easily proved that the two concepts introduce here are actually deeply connected, as the following relation holds

$$\varepsilon_{\delta_{\rm f}}\left(\mathbf{x}_1,\dots,\mathbf{x}_n\right) = (-i)^n \frac{\delta^n\,\Phi_{\delta_{\rm f}}[k]}{\delta k\left(\mathbf{x}_1\right)\dots\delta k\left(\mathbf{x}_n\right)}\Bigg|_{k=0}. \tag{1.36}$$

In analogy with the concept of *cumulants* for real variables, we can define the *connected n-point correlation function* for the field $\delta_{\rm f}$

$$\varepsilon^c_{\delta_{\rm f}}\left(\mathbf{x}_1,\dots,\mathbf{x}_n\right) \equiv \left\langle \delta_{\rm f}\left(\mathbf{x}_1\right)\dots\delta_{\rm f}\left(\mathbf{x}_n\right)\right\rangle_{\delta_{\rm f},\,c} = (-i)^n \frac{\delta^n\,\Phi^c_{\delta_{\rm f}}[k]}{\delta k\left(\mathbf{x}_1\right)\dots\delta k\left(\mathbf{x}_n\right)}\Bigg|_{k=0}, \qquad \Phi^c_{\delta_{\rm f}}[k] = \log\Phi_{\delta_{\rm f}}[k] \tag{1.37}$$

Similarly to the the real variable case, a *cumulant expansion* for a generic n-point correlation function is given in terms of the lower order connected correlation functions

$$\varepsilon_{\delta_{\rm f}}\left(\mathbf{x}_1,\dots,\mathbf{x}_n\right) = \varepsilon^c_{\delta_{\rm f}}\left(\mathbf{x}_1,\dots,\mathbf{x}_n\right) + \sum_{\mathcal{S}\in P(\{\mathbf{x}_1,\dots,\mathbf{x}_n\})} \prod_{\sigma\in\mathcal{S}} \varepsilon^c_{\delta_{\rm f}}\left(\mathbf{x}_{\sigma(1)},\dots,\mathbf{x}_{\sigma(n)}\right) \tag{1.38}$$

where the sum is made over the proper partitions (any partition except the set itself) of $\{\mathbf{x}_1,\dots,\mathbf{x}_n\}$ and σ is thus a subset of $\{\mathbf{x}_1,\dots,\mathbf{x}_n\}$ contained in the partition $\mathcal{S}$. A very useful diagrammatic representation of this decomposition is given in the literature [Bernardeau et al., 2002] and some practical examples will be given in the following chapters.

1.2.3 (Non-)Gaussian random fields

In this section, we will introduce some of the main features of *Gaussian random fields* and we will highlight their importance in Cosmology in Sec. 1.2.4. By contrasts, we will define as *non-Gaussian random fields* those who do not belong to this particular group of random fields.

Definition A *Gaussian random field* is defined by a probability density functional, introduced in Sec. 1.2.1, whose logarithm is quadratic in the field itself

$$\mathcal{P}_G\left[\delta_f\right] \approx \exp\left[-\frac{1}{2}\int d^3\mathbf{x}d^3\mathbf{y}\delta_f\left(\mathbf{x}\right)\mathcal{K}\left(\mathbf{x},\mathbf{y}\right)\delta_f\left(\mathbf{y}\right)\right] \tag{1.39}$$

for some kernel $\mathcal{K}\left(\mathbf{x},\mathbf{y}\right)$. By applying Eq. (1.31), we can actually have a more physical intuition of the definition of a Gaussian field. Specifically, the n-points joint probability is simply given by a *multivariate Gaussian*

$$\mathcal{P}_{G,\,\delta_f}\left(\left[a_n\right],\left[\mathbf{x}_n\right]\right) = \frac{1}{\sqrt{\left(2\pi\right)^{2n}\det\left[C\right]}}\exp\left[-\frac{1}{2}\sum_{i,j=1}^{n}a_i C_{ij}^{-1}a_j\right] \tag{1.40}$$

where $C_{ij}\equiv\mathcal{K}\left(x_i,x_j\right)$ is the *covariance matrix* of the field.

Properties of the correlation functions Via the relations in Eq. (1.36), the 2-point correlation function of a Gaussian field is related to the kernel $\mathcal{K}\left(\mathbf{x},\mathbf{y}\right)$ as

$$\varepsilon_{\delta_f}\left(\mathbf{x}_1,\mathbf{x}_2\right) = -\frac{\delta^2\,\Phi_{\delta_f}[k]}{\delta k\left(\mathbf{x}_1\right)\delta k\left(\mathbf{x}_2\right)}\Big|_{k=0} = \mathcal{K}\left(\mathbf{x}_1,\mathbf{x}_2\right) = C_{12} \tag{1.41}$$

Thanks to the same formalism, via recursive functional derivations it is straightforward to prove an important theorem for Gaussian random fields, the *Wick theorem*. Specifically it allows to directly relate all the *even-point correlation function*s to the 2-point correlation function in Eq. (1.41) as

$$\varepsilon_{\delta_f}\left(\mathbf{x}_1,\ldots,\mathbf{x}_{2n}\right) = \sum_{\text{pair ass.}}\prod_{\text{pairs }(i,j)}\varepsilon_{\delta_f}\left(\mathbf{x}_i,\mathbf{x}_j\right) \tag{1.42}$$

where the sum runs over all the possible pair associations of the points considered in the correlation, and the product multiplies all the 2-point correlation functions evaluated at the points of each pair for a given possible association. Further, all the *odd-point correlation functions* are vanishing

$$\varepsilon_{\delta_f}\left(\mathbf{x}_1,\ldots,\mathbf{x}_{2n+1}\right) = 0. \tag{1.43}$$

Therefore, by measuring the 2-point correlation function of a Gaussian field, we have fully described all the statistical properties of the field itself, e.g. the probability density functional and its connected and non-connected correlation functions at any desired order.

1.2.4 Random fields in Cosmology

In this section we will see how the concepts introduced so far apply to real fields in Cosmology.

Ensemble average In general, observables in Cosmology are functional of the matter field δ (from now on we will drop the subscript m) and we would like to address the statistical properties of the former (moments, correlations at different points in the sky ...) to infer those of the latter. Clearly, Eq. (1.33) can not be implemented as it is given that we have just one Universe at our disposal. Usually in Cosmology we employ the *ergodic hypothesis*: the *ensemble average* of any quantity coincides with the *sample average* of the same quantity. Most of the time the sample average is actually a volume average, otherwise it has to be properly defined. Therefore, a common realisation of Eq. (1.33) is

$$\langle \mathcal{O}[\delta_{\rm f}] \rangle_{\delta_{\rm f}} \approx \frac{1}{V} \int_V d^3\mathbf{r}\, \mathcal{O}[\delta_{\rm f}(\mathbf{r})]. \tag{1.44}$$

The *ergodic hypothesis* has been proved to hold for one-dimensional stationary random functions [M IAglom & A Silverman, 1964]. For cosmological applications, the ergodic hypothesis is exact in the limit of infinite volume and within standard inflationary scenarios.

Statistical homogeneity and isotropy We anticipated that cosmological observables can be formally seen as functionals $\mathcal{O}[\delta_{\rm f}]$ of a given cosmological field $\delta_{\rm f}$ (matter field, or tracer of it such as density of halos, galaxies, etc ...). Given one particular observable $\mathcal{O}[\delta_{\rm f}]$, its expectation value can be predicted by taking an ensamble average via the previously defined operator $\langle \mathcal{O}[\delta_{\rm f}] \rangle_{\delta_{\rm f}}$, and, from the observational side, we compare this value with the volume average of the given observable once measured on different points of the sky. Clearly the volume average will be just time dependent and naturally forces the final result to be homogeneous and isotropic. However, we want this valuable property to apply also for the expected theoretical value of the observable itself. This translates in assuming that all the fields we will deal with in Cosmology are *statistically isotropic and homogeneous*: the $n-$points joint probability $\mathcal{P}_{\delta_{\rm f}}([a_n], [\mathbf{x}_n])$ is in invariant under *rotation* or *translation*, respectively, of the points $\mathbf{x}_n$.

A crucial consequence of such a request can be seen at the level of correlation functions

$$\varepsilon^{(c)}_{\delta_{\rm f}}(\mathbf{x}_1, \dots, \mathbf{x}_n) \overset{\text{hom.}}{\equiv} \varepsilon^{(c)}_{\delta_{\rm f}}(\mathbf{x}_1 - \mathbf{x}_o, \dots, \mathbf{x}_n - \mathbf{x}_o) \overset{\text{iso.}}{\equiv} \varepsilon^{(c)}_{\delta_{\rm f}}(R\mathbf{x}_1, \dots, R\mathbf{x}_n), \quad \mathbf{x}_o \in \mathbb{R}^3 \tag{1.45}$$

where R simply rotate the system of coordinates. In particular, thanks to homogeneity, the correlation function will depend just on the relative distances between the points

$$\varepsilon^{(c)}_{\delta_{\rm f}}(\mathbf{x}_1, \dots, \mathbf{x}_n) \overset{\text{hom.}}{\to} \varepsilon^{(c)}_{\delta_{\rm f}}(\mathbf{x}_1, \dots, \mathbf{x}_1 + (\mathbf{x}_n - \mathbf{x}_1)) \overset{\text{hom.}}{\to} \varepsilon^{(c)}_{\delta_{\rm f}}(\mathbf{x}_1, \dots, \mathbf{x}_1 + \mathbf{r}_n) \tag{1.46}$$

where here and in the following $\mathbf{r}_i = \mathbf{x}_i - \mathbf{x}_1$. In particular, Eq. (1.46) just forces homogeneity on $\varepsilon^{(c)}_{\delta_{\rm f}}$, while isotropy would further request no dependence on the orientation of the configuration of the points $\{\mathbf{x}_i\}_{i=1,\dots,n}$. Furthermore, the dependence on the base point $\mathbf{x}_1$ is just fictitious and the generic homogeneous n-point correlation function will actually depend on $(n-1)$ vectors.

The request for isotropy is much more complicated. Let us show a couple of simple cases where intuition can help, even though more rigorous proofs are available in the literature [Marcori & Pereira, 2017]. If we consider

the specific case of the 2-point correlation function, isotropy will further bind the correlations to depend just on the module of the relative distance between the 2 points

$$\varepsilon^{(c)}_{\delta_f}(\mathbf{x}_1, \mathbf{x}_2) = \varepsilon^{(c)}_{\delta_f}(\mathbf{x}_1, \mathbf{x}_1 + \mathbf{r}_2) \overset{\text{hom.}}{\equiv} \varepsilon^{(c)}_{\delta_f}(\mathbf{r}_2) \overset{\text{iso.}}{\equiv} \varepsilon^{(c)}_{\delta_f}(r_2) \tag{1.47}$$

As for the 3-point correlation function $\varepsilon^{(c)}_{\delta_f}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \overset{\text{hom.}}{\equiv} \varepsilon^{(c)}_{\delta_f}(\mathbf{r}_2, \mathbf{r}_3)$, isotropy is implemented by forcing it to depend only on the three combinations of $\mathbf{r}_2, \mathbf{r}_3$ that are rotationally invariant: their modules r_2, r_3 and their scalar product $\mathbf{r}_2 \cdot \mathbf{r}_3$. In summary

$$\varepsilon^{(c)}_{\delta_f}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) \overset{\text{hom.}}{\equiv} \varepsilon^{(c)}_{\delta_f}(\mathbf{r}_2, \mathbf{r}_3) \overset{\text{iso.}}{\equiv} \varepsilon^{(c)}_{\delta_f}(r_2, r_3, \mathbf{r}_2 \cdot \mathbf{r}_3) \tag{1.48}$$

At every steps, we are actually reducing the number of degrees of freedom, from 9, to 6 (homogeneity), to 3 (isotropy).

Cumulants and moments In Cosmology we will define *moments* and *cumulants* as the correlation functions in Eq. (1.34) (non-connected) and in Eq. (1.37) (connected) respectively when evaluated at the same location in space. Therefore, the nth order cumulant and the nth order moment for a random field δ_f, will be defined as

$$m_n \equiv \langle \delta^n_f(\mathbf{x}) \rangle_{\delta_f} = (-i)^n \frac{\delta^n \Phi_{\delta_f}[k]}{\delta^n k(\mathbf{x})} \Big|_{k=0}, \tag{1.49}$$

$$\kappa_n \equiv \langle \delta^n_f(\mathbf{x}) \rangle^c_{\delta_f} = (-i)^n \frac{\delta^n \Phi^c_{\delta_f}[k]}{\delta^n k(\mathbf{x})} \Big|_{k=0}. \tag{1.50}$$

Thanks to the request of *statistical homogeneity and isotropy* they will not actually depend on the point of the space considered.

We would like to underline that the moment and the cumulant of order 1 are identically 0, by definition, i.e. $\langle \delta_f \rangle = 0$. Therefore, in the expansion defined in Eq. (1.38), all the terms involving a singlet will be vanishing. The moment of order two m_2 will be therefore identical to the cumulant of order 2 κ_2 and it is called the variance of the field, usually noted ad σ^2. As for higher order cumulants, they will be in general different from the same-order moments and identically 0 for Gaussian random fields. In line with the tradition, they are usually normalised to the *standard deviation* of the field $\sigma = \sqrt{\sigma^2}$

$$S_n \equiv \frac{\kappa_n}{\sigma^{2n-2}}. \tag{1.51}$$

Polyspectra In Sec. 1.2.2 we defined the correlation functions for a field in real space. Often in cosmology we will employ their counterpart in Fourier space, the *polyspectra*. Given a scalar field in a real d-dimensional space $\delta_f(\mathbf{x}, t)$ (where we made explicit once again the time dependence), we can consider, at every time (dependence

that we locally restore in this paragraph), its Fourier decomposition $\tilde{\delta}_{\rm f}(\mathbf{x},t)$ in configuration space

$$\tilde{\delta}_{\rm f}(\mathbf{k},t) = \int d^{\rm d}\mathbf{x}\,\delta_{\rm f}(\mathbf{x},t)\,e^{+i\mathbf{k}\cdot\mathbf{x}}, \qquad \delta_{\rm f}(\mathbf{x},t) = \int \frac{d^{\rm d}\mathbf{k}}{(2\pi)^{\rm d}}\,\tilde{\delta}_{\rm f}(\mathbf{k},t)\,e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{1.52}$$

which is a field itself on the configuration space. In Eq. (1.52) we show both the direct and inverse Fourier transform for a general dimension to fix our convention on the pre-factors. In this chapter, we will consider $\mathrm{d}=3$. At the level of notation, we will eventually switch to the following much simpler notation for the cosmological field $\tilde{\delta}_{\rm f}(\mathbf{k},t)$: $\delta_{\mathbf{k}}$. This will be the case when deriving the covariance matrix expressions for the observable of interest in Chapter 2.

The decomposition of a scalar cosmological field $f(\mathbf{x},t) = f(t) + \delta f(\mathbf{x},t)$ in its Fourier components $\tilde{f}(\mathbf{k},t)$ has a deeper meaning. As a matter of fact, the assumption of $\tilde{f}(\mathbf{k},t)$ being statistically homogeneous at every time, implies that the background component $\langle f(\mathbf{x},t)\rangle = f(t)$ does not have any space-dependence. Therefore, all the operators that will appear in the equation of motions for the perturbation $\delta f(\mathbf{x},t)$ (Vlasov-Poisson system in Sec. 1.3.1), at linear order, will commute with the Fourier operator on the left of Eq. (1.52), leading to a system of equations for each mode $\mathbf{k}$ of the field. Further, being the plane waves $e^{\pm i\mathbf{k}\cdot\mathbf{x}}$ eigenfunctions of all the space-derivative operators, there will be no couplings between the modes and each of them will simply evolve independently in time. All these features are present as long as the equations for the gravitational dynamics are linear and the statistical homogeneity is assumed, leading to a very simple analysis, in Fourier space, of the linear gravitational collapse.

The *polyspectrum* $P_{\delta_{\rm f}}(\mathbf{k}_2,\ldots,\mathbf{k}_n)$ of a given order n is related to the n-point correlation function of the field $\tilde{\delta}_{\rm f}(\mathbf{k},t)$ as

$$\langle\tilde{\delta}_{\rm f}(\mathbf{k}_1)\ldots\tilde{\delta}_{\rm f}(\mathbf{k}_n)\rangle_{\delta_{\rm f}} = (2\pi)^3\,\delta_{\rm D}(\mathbf{k}_1+\cdots+\mathbf{k}_n)\,P_{\delta_{\rm f}}(\mathbf{k}_2,\ldots,\mathbf{k}_n)\,, \tag{1.53}$$

$$\langle\tilde{\delta}_{\rm f}(\mathbf{k}_1)\ldots\tilde{\delta}_{\rm f}(\mathbf{k}_n)\rangle_{\delta_{\rm f},c} = (2\pi)^3\,\delta_{\rm D}(\mathbf{k}_1+\cdots+\mathbf{k}_n)\,P^c_{\delta_{\rm f}}(\mathbf{k}_2,\ldots,\mathbf{k}_n) \tag{1.54}$$

where the constraint imposed by the *Dirac deltas* $\delta_{\rm D}$, as we will see, comes from the request for homogeneity and bounds the momenta to form closed polygons in configuration space. The relations (1.53) and (1.54) can be proved to be actually consistent with the operational definition of $P^{(\rm c)}_{\delta_{\rm f}}$ as the Fourier transform of the same order correlation function $\varepsilon^{(\rm c)}_{\delta_{\rm f}}$, i.e.

$$P^{(\rm c)}_{\delta_{\rm f}}(\mathbf{k}_2,\ldots,\mathbf{k}_n) \equiv \prod_{i=2}^{n}\left[\int d^3\mathbf{r}_i\,e^{+i\mathbf{k}_i\cdot\mathbf{r}_i}\right]\varepsilon^{(\rm c)}_{\delta_{\rm f}}(\mathbf{r}_2,\ldots,\mathbf{r}_n) \tag{1.55}$$

where we expressed the correlation function in real space in terms of the distances between the points. As a matter

of fact,

$$\left\langle \tilde{\delta}_{\rm f}\left(\mathbf{k}_1\right)\ldots\tilde{\delta}_{\rm f}\left(\mathbf{k}_n\right)\right\rangle_{\delta_{\rm f},\,(c)} = \prod_{i=1}^{n}\left[\int d^3\mathbf{x}_i\, e^{+i\mathbf{k}_i\cdot\mathbf{x}_i}\right]\left\langle \delta_{\rm f}\left(\mathbf{x}_1\right)\ldots\delta_{\rm f}\left(\mathbf{x}_n\right)\right\rangle_{\delta_{\rm f},\,(c)} \tag{1.56}$$

$$= \int d^3\mathbf{x}_1\, e^{+i\mathbf{k}_1\cdot\mathbf{x}_1}\prod_{i=2}^{n}\left[\int d^3\mathbf{r}_i\, e^{+i\mathbf{k}_i\cdot\left(\mathbf{r}_i+\mathbf{x}_1\right)}\right]\left\langle \delta_{\rm f}\left(\mathbf{x}_1\right)\ldots\delta_{\rm f}\left(\mathbf{x}_1+\mathbf{r}_n\right)\right\rangle_{\delta_{\rm f},\,(c)} \tag{1.57}$$

$$= \int d^3\mathbf{x}_1 e^{+i\left(\sum_{i=1}^{n}\mathbf{k}_i\right)\cdot\mathbf{x}_1}\prod_{i=2}^{n}\left[\int d^3\mathbf{r}_i\, e^{+i\mathbf{k}_i\cdot\mathbf{r}_i}\right]\left\langle \delta_{\rm f}\left(\mathbf{x}_1\right)\ldots\delta_{\rm f}\left(\mathbf{x}_1+\mathbf{r}_n\right)\right\rangle_{\delta_{\rm f},\,(c)} \tag{1.58}$$

$$= \left(2\pi\right)^3\delta_{\rm D}\left(\mathbf{k}_1+\cdots+\mathbf{k}_n\right)\prod_{i=2}^{n}\left[\int d^3\mathbf{r}_i\, e^{+i\mathbf{k}_i\cdot\mathbf{r}_i}\right]\left\langle \delta_{\rm f}\left(\mathbf{r}_2\right)\ldots\delta_{\rm f}\left(\mathbf{r}_n\right)\right\rangle_{\delta_{\rm f},\,(c)}. \tag{1.59}$$

In the first line we simply inverse-Fourier-transformed every single mode. In the second line we performed a change of variable $\mathbf{r}_i = \mathbf{x}_n - \mathbf{x}_1$ for $i > 1$. In the fourth line we performed an integration over $\mathbf{x}_1$ to get the Dirac delta, given the fact that $\left\langle \ldots\delta_{\rm f}\left(\mathbf{x}_1\right)\right\rangle_{\delta_{\rm f},\,(c)}$ does not actually depends on the base point $\mathbf{x}_1$, thanks to homogeneity. Finally, we can recognise the relation (1.55), once removed the fictitious dependence on $\mathbf{x}_1$.

Focusing on the 2-point correlation function and on the 3-point correlation function, their polypsectra are respectively dubbed *power spectrum P* and *bispectrum B*. Given the considerations above, we can further simplify their dependencies on the configurations involved. Specifically, for the power spectrum isotropy leads to a dependence on just the module of the only momentum $\mathbf{k}$. As for the bispectrum instead, isotropy leads to a dependence on the modules of the two momenta and on their angle. More often, the dependence of the bispectrum in an homogeneous and isotropic Universe is expressed in terms of the modules of the edges of the triangular configuration involved. Clearly the two definitions have the same physical meaning. In summary

$$P^{(c)}_{\delta_{\rm f}}\left(\mathbf{k},\mathbf{k}'\right) \overset{\text{hom.}}{\equiv} P^{(c)}_{\delta_{\rm f}}\left(\mathbf{k}\right) \overset{\text{iso.}}{\equiv} P^{(c)}_{\delta_{\rm f}}\left(k\right), \tag{1.60}$$

$$B^{(c)}_{\delta_{\rm f}}\left(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3\right) \overset{\text{hom.}}{\equiv} B^{(c)}_{\delta_{\rm f}}\left(\mathbf{k}_2,\mathbf{k}_3\right) \overset{\text{iso.}}{\equiv} B^{(c)}_{\delta_{\rm f}}\left(k_1,k_2,k_3\right). \tag{1.61}$$

Connected vs. non-connected correlation functions In Eq. (1.34) and in Eq. (1.37) we introduced, in a general manner, the concepts of *non-connected* and of *connected correlation functions*. However, in cosmology, we will always employ the latter and from this moment on we will always refer to them in this manuscript, dropping the superscript *c*. A practical reason for this is that, if the field of interest has deviations from the Gaussian statistics, then different modes of the field $\hat{\delta}\left(\mathbf{k}\right)$ will be independent and the connected correlation functions beyond the 2-point one, will be vanishing. If any *higher order correlation function* is detected to be non-varnishing, then this measurement will directly tell us about the level of non-Gaussianity present in the field.

Operative definition For later convenience, let us introduce a more operative definition of the *n*-point correlation function and which is actually closer to the way it is built from catalogues. In Cosmology, may it be in simulations or within actual data sets, the general field δ_f is not continuous: we will always detect the position

and the concentration of particles/pixels (dark matter particles, point sources, galaxies, ...) in the sky. Therefore, given N volumes $\{\Delta V_i\}_{i=1,\dots,N}$ centred in positions $\{\mathbf{x}_i\}_{i=1,\dots,N}$, let us consider the density of the particles in each of them $\{n_i^{\rm f} = N_i^{\rm f}/\Delta V_i\}_{i=1,\dots,N}$, $N_i^{\rm f}$ being the number of particles in the ith volume. Of course we can define a mean background density $\bar{n}^{\rm f} = N_{\rm tot.}^{\rm f}/V$. Therefore, an operational definition of the density contrast $\delta_{\rm f}(\mathbf{x}_i)$ within the i$^{\rm th}$ volume is given by

$$\bar{n}^{\rm f}\left[\delta_{\rm f}(\mathbf{x}_i)+1\right] = n_i^{\rm f} \tag{1.62}$$

The product $\langle N_1^{\rm f}\dots N_N^{\rm f}\rangle$ will then be the average number of N-tuple in the set of volumes $\{\Delta V_i\}_{i=1,\dots,N}$ located at positions $\{\mathbf{x}_i\}_{i=1,\dots,N}$; rephrased, the product of the number of particles in the first volume times the number in the second volume ...times the number of particles in the Nth volume

$$\begin{aligned}\langle N_1\dots N_N\rangle = \bar{n}^N\Big\langle\prod_{i=1}^{N}\Delta V_i\left[\delta_{\rm f}(\mathbf{x}_i)+1\right]\Big\rangle = \\ \bar{n}^N\left[\langle\delta_{\rm f}(\mathbf{x}_1)\dots\delta_{\rm f}(\mathbf{x}_N)\rangle+\dots+1\right](\Delta V)^N = \\ \bar{n}^N\left[\varepsilon^{(c)}_{\delta_{\rm f}}(\mathbf{x}_1,\dots,\mathbf{x}_N)+\dots+1\right](\Delta V)^N\end{aligned} \tag{1.63}$$

where in the last term we are omitting all the $(<N)$-point correlation functions deriving from the product in the second term. We also assumed all the volume elements to be of the same size to simplify our notation, along with removing the superscript *f*. We can invert Eq. (1.63) to get

$$\frac{\langle N_1\dots N_N\rangle}{(\bar{n}\Delta V)^N} = 1+\varepsilon^{(c)}_{\delta_{\rm f}}(\mathbf{x}_1,\dots,\mathbf{x}_N)+\dots. \tag{1.64}$$

Given a single volume element ΔV, the expected number of particles in it is

$$\langle N\rangle = \bar{n}\Delta V. \tag{1.65}$$

If the number of particles $\{N_i\}_{i=1,\dots,N}$ in the volumes $\{\Delta V_i\}_{i=1,\dots,N}$ is not correlated, then the average number of pairs is the product of the average number of particles in the volumes

$$\langle N_1\dots N_N\rangle = (\bar{n}\Delta V)^N. \tag{1.66}$$

Therefore, we can see how the hierarchy of correlation function $\varepsilon^{(c)}_{\delta_{\rm f}}(\mathbf{x}_1,\dots,\mathbf{x}_N)+\dots$ represents the fractional excess of N-tuple at the different positions $\{\mathbf{x}_i\}_{i=1,\dots,N}$, over what would be expected if the distribution were completely random. We refer to the literature for an in depth discussion discussion up to the 4-point correlation function (e.g. Peebles [1980]).

1.3 The large-scale structure of the Universe (LSS)

In this section we will review the main features of the model we will use to describe the growth of structures in the *non-linear regime*, when $|\delta_{\rm m}| \approx 1$: the well known *halo-model* (Cooray & Sheth [2002] for a thorough review). We will start from a brief description of the *Perturbation Theory* (PT) approach to the understanding of the LSS evolution ([Bernardeau et al., 2002] for a thorough review). This will allow us to define several quantities which will be employed in the halo model itself and later in the next chapter.

1.3.1 Perturbation theory, a summary

Gravitational dynamics First of all, let us fix the main variables describing the dynamical evolution of the matter perturbation in an expanding Universe. In this section and in the following ones we will explicit the dependence of the cosmological fields on the time to allow a clearer separation between background and perturbative components. In line with the common practice, we will split the local matter density $\rho(\mathbf{x}, t)$ in a background time-dependent component $\rho(t)$ and in a perturbation term $\delta(\mathbf{x}, t)$

$$\rho(\mathbf{x}, t) = \rho(t)\left(1 + \delta(\mathbf{x}, t)\right). \tag{1.67}$$

When referring to the matter contrast density field δ we will omit the subscript m, being indeed the main cosmological field of interest. While the background evolution is fully determined by the system of Eq. (1.15), Eq. (1.16) and Eq. (1.19), in the following we will focus on the perturbation $\delta(\mathbf{x}, t)$. We start by considering the *phase space density function* $f(\mathbf{x}, \mathbf{p}, t)$. It describes the distribution of the matter particles of mass m, per unit of volume in real (comoving) and in momenta space in terms of its coordinates

$$\mathbf{x} = \frac{\mathbf{r}^{\rm phy.}}{a}, \qquad \mathbf{p} = m\, a\, \mathbf{u}_{\rm p}. \tag{1.68}$$

For simplicity, we will not explicit the time dependence of the scale factor a. We introduced the peculiar velocity $\mathbf{u}_{\rm p}$ of a particle and it is defined as difference between the physical total velocity and the Hubble flow

$$\mathbf{u}_{\rm p}(\mathbf{x}, t) \equiv \frac{d\mathbf{r}^{\rm phy.}}{dt} - \frac{da}{dt}\mathbf{x} = a\frac{d\mathbf{x}}{dt}. \tag{1.69}$$

We can then introduce the *mean velocity flow* $\mathbf{u}$ (which is not to be confused with the peculiar velocity $\mathbf{u}_{\rm p}$) as the first-order momentum of the distribution $f(\mathbf{x}, \mathbf{p}, t)$

$$u_i(\mathbf{x}, t) \equiv \frac{1}{N_{\rm p}} \int d^3\,\mathbf{p}\, \frac{p_i}{ma} f(\mathbf{x}, \mathbf{p}, t), \tag{1.70}$$

$N_{\rm p} = \int d^3\mathbf{p}\, f(\mathbf{x}, \mathbf{p}, t)$ being the mean number density of particles, and the *velocity dispersion tensor* $\sigma_{ij}(\mathbf{x}, t)$ as its second-order momentum

$$\sigma_{ij}(\mathbf{x}, t) \equiv \frac{1}{N_{\rm p}} \int d^3\mathbf{p}\, \frac{p_i}{ma}\, \frac{p_j}{ma} f(\mathbf{x}, \mathbf{p}, t) - \mathbf{u}_i(\mathbf{x}, t)\, \mathbf{u}_j(\mathbf{x}, t)\,. \tag{1.71}$$

In the following, along with the local density contrast δ, the quantity we will solve our equations for will be the *velocity divergence*, defined as

$$\theta(\mathbf{x}, t) \equiv \frac{1}{aH} \nabla_{\mathbf{x}} \cdot \mathbf{u}(\mathbf{x}, t)\,. \tag{1.72}$$

The Vlasov equations According to the ΛCDM model, dark matter is described as a fluid of collision-less particles. In this case, the *Liouville theorem* applies and the phase space density function $f(\mathbf{x}, \mathbf{p}, t)$ is conserved in time

$$\frac{D f(\mathbf{x}, \mathbf{p}, t)}{D t} = \left[\frac{\partial}{\partial t} + \frac{d\mathbf{x}}{dt}\nabla_{\mathbf{x}} + \frac{d\mathbf{p}}{dt}\nabla_{\mathbf{p}}\right] f(\mathbf{x}, \mathbf{p}, t) = 0. \tag{1.73}$$

In terms of the coordinates defined above, the Liouville equation is known as *Vlasov equation*

$$\left[\frac{\partial}{\partial t} + \frac{\mathbf{p}}{m\, a^2}\nabla_{\mathbf{x}} - m\nabla_{\mathbf{x}}\Phi(\mathbf{x}, t) \cdot \nabla_{\mathbf{p}}\right] f(\mathbf{x}, \mathbf{p}, t) = 0, \tag{1.74}$$

and it is coupled with the scalar Newtonian potential $\Phi(\mathbf{x}, t)$

$$\nabla^2\Phi(\mathbf{x}, t) = \frac{4\pi G\, m}{a}\left(\int d^3\mathbf{p}\, (\mathbf{x}, \mathbf{p}, t) - \bar{n}(t)\right) \tag{1.75}$$

with $\bar{n} = \int d^3\mathbf{x}\, d^3\mathbf{p}\, f(\mathbf{x}, \mathbf{p}, t)\, /V$. A typical approach is to solve Eq. (1.74) via a decomposition of it in moments of the phase space function. Therefore, by applying the operators

$$\int d^3\,\mathbf{p}\frac{p_i}{ma}\,\star, \quad \text{1st moment;} \qquad \int d^3\mathbf{p}\,\frac{p_i}{ma}\,\frac{p_j}{ma}\,\star, \quad \text{2nd moment} \tag{1.76}$$

on the l.h.s. and on the r.h.s. of the *Vlasov equation*, we can derive a system of equations for the first-order moment $\mathbf{u}$ and of the second-order moment σ_{ij} of $f(\mathbf{x}, \mathbf{p}, t)$

$$\begin{cases} \frac{\partial\delta(\mathbf{x},t)}{\partial t} + \frac{1}{a}\nabla_{\mathbf{x}} \cdot \left[(1 + \delta(\mathbf{x}, t))\, \mathbf{u}(\mathbf{x}, t)\right] = 0 \\ \frac{\partial\mathbf{u}_i(\mathbf{x},t)}{\partial t} + \frac{\dot{a}}{a}\mathbf{u}_i(\mathbf{x}, t) + \frac{1}{a}\mathbf{u}_j(\mathbf{x}, t)\, \mathbf{u}_{i,j}(\mathbf{x}, t) = -\frac{1}{a}\Phi_{,i}(\mathbf{x}, t) - \frac{\left[\rho(\mathbf{x},t)\sigma_{ij}(\mathbf{x},t)\right]_{,j}}{a\,\rho(\mathbf{x},t)}. \end{cases} \tag{1.77}$$

Finally, we close our system by assuming that the dark matter field can be described by a perfect fluid with an isotropic velocity dispersion which is proportional to the pressure, i.e.

$$\sigma_{ij}(\mathbf{x}, t) = \delta_{ij}\frac{P(t)}{\rho(\mathbf{x}, t)}. \tag{1.78}$$

In the following, we will place ourselves in the commonly used *single flow approximation*:

$|\sigma_{ij}| \ll |\mathbf{u}|$. This approximation is valid at large enough scales where we assume that the flow of the particles induced by gravitational collapse is dominant with respect to the thermal motion of the particles within the different fluids and the pressure is negligible in the dynamics. Within this approximation, the phase space density function can be written as

$$f(\mathbf{x}, \mathbf{p}, t) = \frac{a^3 \rho(\mathbf{x}, t)}{m} \left(\mathbf{p} - m\, a\, \mathbf{u}(\mathbf{x}, t)\right) \tag{1.79}$$

and Eq. (1.75) simplifies to the well known generalised Poisson equation

$$\nabla^2 \Phi\left(\mathbf{x}, t\right) = 4\,\pi\, G\, \rho(t) \delta(\mathbf{x}, t) \tag{1.80}$$

The Vlasov-Poisson system: linear solutions At first place, we will show the solution to the system (1.77) in the linear regime, i.e. when

$$\left[\delta\left(\mathbf{x}, t\right) \mathbf{u}_i\left(\mathbf{x}, t\right)\right]_{,i} \ll \delta\left(\mathbf{x}, t\right), \tag{1.81}$$

$$\left[\mathbf{u}_j\left(\mathbf{x}, t\right) \mathbf{u}_i\left(\mathbf{x}, t\right)\right]_{,j} \ll \delta\left(\mathbf{x}, t\right). \tag{1.82}$$

In this regime of the matter perturbation evolution, the system of equations simplifies to

$$\begin{cases} \left[\frac{\partial^2}{\partial t^2} + 2H(t)\frac{\partial}{\partial t} - \frac{3}{2}H^2(t)\Omega_{\mathrm{m}}\left(t\right)\right] \delta\left(\mathbf{x}, t\right) = 0, \\ \theta\left(\mathbf{x}, t\right) = -\frac{1}{H(t)}\frac{\partial}{\partial t}\delta\left(\mathbf{x}, t\right) \end{cases} \tag{1.83}$$

where we used in first equation the generalised Poisson equation in Eq. (1.80) along with the definition of matter density parameter (1.21). In terms of the *growth factor* $D\left(t\right) \equiv \frac{\delta(\mathbf{x},t)}{\delta(\mathbf{x},t_{\mathrm{i}})}$ that we introduce here for the first time for a convenient choice of the initial time t_{i}, the solution to the system in Eq. (1.83) is

$$\ddot{D}\left(t\right) + 2H(t)\dot{D}\left(t\right) - \frac{3}{2}H^2(t)D\left(t\right) = 0. \tag{1.84}$$

If we assume that our observable Universe can be approximated by the EdS model described in Sec. 1.1.4, then we can find analytical solutions. Specifically, the most general one will be the linear combination of 2 independent modes

$$\delta_+\left(\mathbf{x}, t\right) = D_+\left(t\right) \delta_{\mathrm{m}}\left(\mathbf{x}, t_{\mathrm{i}}\right), \tag{1.85}$$

$$\delta_-\left(\mathbf{x}, t\right) = D_-\left(t\right) \delta_{\mathrm{m}}\left(\mathbf{x}, t_{\mathrm{i}}\right), \tag{1.86}$$

respectively named as *growing* and *decaying* mode. Thanks to the second relation in Eq. (1.83), we can express the linear solution for the field θ

$$\theta_+(\mathbf{x},t) = -\frac{\partial \log D_+(t)}{\partial \log a} D_+(t)\,\delta_{\rm m}(\mathbf{x},t_{\rm i}) \tag{1.87}$$

$$\theta_-(\mathbf{x},t) = -\frac{\partial \log D_+(t)}{\partial \log a} D_-(t)\,\delta_{\rm m}(\mathbf{x},t_{\rm i}) \tag{1.88}$$

For an EdS Universe, we will have $D_+(t) \propto t^{\frac{2}{3}}$ and $D_-(t) \propto t^{-1}$. Even if the above solution is a good approximation also for a ΛCDM Universe, in all our applications we will compute the numerical solution to the exact differential equation (1.84). We will comment in Sec. 1.3.2 about a suitable choice for the initial perturbation $\delta(\mathbf{x},t_{\rm i})$.

The Vlasov-Poisson system: non-linear solutions We now shift to a regime of the matter clustering where the linear approximation defined by Eq. (1.81) and Eq. (1.82) does not hold. We will preserve the single-flow assumption and we will focus on the solution for an Eds model. The system to solve reads

$$\begin{cases} \frac{\partial \delta(\mathbf{x},t)}{\partial t} + \frac{1}{a}\left[(1+\delta(\mathbf{x},t))\,\mathbf{u}_i(\mathbf{x},t)\right]_{,i} = 0, \\ \frac{\partial u_i(\mathbf{x},t)}{\partial t} + \frac{\dot{a}}{a}\mathbf{u}_i(\mathbf{x},t) + \frac{1}{a}\mathbf{u}_j(\mathbf{x},t)\,\mathbf{u}_{i,j}(\mathbf{x},t) = -\frac{1}{a}\Phi_{,i}(\mathbf{x},t)\,. \end{cases} \tag{1.89}$$

Following the standard approach outlined in great details within the review by Bernardeau et al. [2002], let us introduce the following doublet which will allow for a more elegant solution

$$\mathbf{\Psi}(\mathbf{x},t) \equiv (\delta(\mathbf{x},t)\,, -\theta(\mathbf{x},t))\,. \tag{1.90}$$

In Fourier space, we can write the system to solve as

$$\left[\frac{\partial}{\partial\eta}\delta^b_a + \Omega^b_a\right]\tilde{\Psi}_b(\mathbf{k},\eta) = \int \frac{d^3\mathbf{k}_1}{(2\pi)^2}\frac{d^3\mathbf{k}_2}{(2\pi)^2}\,\gamma^{bc}_a(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2)\;\tilde{\Psi}_b(\mathbf{k}_1,\eta)\,\tilde{\Psi}_c(\mathbf{k}_2,\eta) \tag{1.91}$$

where we introduced the time variable $\eta = \log D_+(t)$. For an EdS Universe, the matrix Ω^b_a is independent of time

$$\Omega^b_a = \begin{pmatrix} 0 & -1 \\ -\frac{3}{2} & \frac{1}{2} \end{pmatrix}, \tag{1.92}$$

and the components of the *vertex* $\gamma^{bc}_a(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2)$ are

$$\gamma^{12}_1(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2) = \delta_{\rm D}(\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2)\,\frac{\alpha(\mathbf{k}_1,\mathbf{k}_2)}{2}, \qquad \alpha(\mathbf{k}_1,\mathbf{k}_2) = \frac{(\mathbf{k}_1+\mathbf{k}_2)\cdot\mathbf{k}_1}{k_1^1} \tag{1.93}$$

$$\gamma^{21}_1(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2) = \delta_{\rm D}(\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2)\,\frac{\alpha(\mathbf{k}_2,\mathbf{k}_1)}{2}, \qquad \beta(\mathbf{k}_1,\mathbf{k}_2) = \frac{|\mathbf{k}_1+\mathbf{k}_2|\mathbf{k}_1\cdot\mathbf{k}_2}{2k_1^1k_2^2} \tag{1.94}$$

$$\gamma^{22}_2(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2) = \delta_{\rm D}(\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2)\,\frac{\beta(\mathbf{k}_1,\mathbf{k}_2)}{2}, \tag{1.95}$$

The idea lying behind the philosophy of the PT approach to the emergence of LSS is that we can expand the doublet in Eq. (1.90) at different orders for which we can solve perturbatively

$$\tilde{\boldsymbol{\Psi}}\left(\mathbf{x},\eta\right)=\sum_{p}\tilde{\boldsymbol{\Psi}}^{(p)}\left(\mathbf{x},\eta\right), \tag{1.96}$$

each term being proportional to the p$^{\text{th}}$ power of the linearly evolved over-density $D_{+}(\eta)\delta\left(\mathbf{x},\eta_{\text{i}}\right)$ with $D_{+}(\eta_{\text{i}})=1$ [Goroff et al., 1986; Makino et al., 1992; Jain & Bertschinger, 1996]

$$\tilde{\boldsymbol{\Psi}}^{(p)}\left(\mathbf{x},\eta\right)\equiv\left(\delta^{(p)}\left(\mathbf{x},\eta\right),-\theta^{(p)}\left(\mathbf{x},\eta\right)\right). \tag{1.97}$$

At this point, it is possible to iteratively solve the formal equation Eq. (1.91) at the different order $\Psi_{a}^{(p)}$, starting from lower ones. The final result will be the fields $\delta^{(p)}\left(\mathbf{x},\eta\right),\theta^{(p)}\left(\mathbf{x},\eta\right)$ being function of p-powers of the linear solution of the growing mode $\tilde{\delta}_{+}\left(\mathbf{k},\eta\right)$, convoluted with a kernel. This can eventually be translate in a diagrammatic representation of the modes as well explained in the literature [Bernardeau et al., 2002; Crocce & Scoccimarro, 2006b]. We will not talk about it in this manuscript. In general, we will be capable to write the p-order solution for the vector $\Psi_{a}^{(p)}$ defined in Eq. (1.97) as

$$\tilde{\Psi}_{a}^{(p)}\left(\mathbf{k},\eta\right)=\int\frac{d^{3}\mathbf{k}}{(2\pi)^{3}}\cdots\frac{d^{3}\mathbf{k}_{p}}{(2\pi)^{3}}\delta_{\text{D}}\left(\mathbf{k}-\mathbf{k}_{1\ldots p}\right)\mathcal{F}_{a}^{(p)}\left(\mathbf{k}_{1},\ldots,\mathbf{k}_{p};z\right)\tilde{\delta}_{+}\left(\mathbf{k}_{1},\eta\right)\ldots\tilde{\delta}_{+}\left(\mathbf{k}_{p},\eta\right) \tag{1.98}$$

In the literature, respectively for the local matter field δ (i.e. $a=1$) and the velocity divergence field θ (i.e. $a=2$), the kernel $\mathcal{F}_{a}^{(p)}$ is defined as $F^{(p)}$ and $G^{(p)}$. We will employ this notation later in the chapter.

1.3.2 The linear solution and its power spectrum

Given the formalism introduced in the previous section, we would like to briefly describe the *linear matter power spectrum* $P^{\text{lin.}}(k)$ in cosmology, given that it will play an important role in future definitions. Undoubtedly, the linear matter power spectrum is the power spectrum for linearly-evolved initial matter perturbations

$$P^{\text{lin.}}(k)\delta^{\text{D}}\left(\mathbf{k}'+\mathbf{k}\right)\propto\langle\tilde{\delta}\left(\mathbf{k},t\right)\tilde{\delta}\left(\mathbf{k}',t\right)\rangle_{\text{lin.}}\propto D_{+}^{2}\left(t\right)\langle\tilde{\delta}\left(\mathbf{k},t_{\text{i}}\right)\tilde{\delta}\left(\mathbf{k}',t_{\text{i}}\right)\rangle \tag{1.99}$$

In the last definition we have assumed equal-time correlations, the initial time t_{i} will define a *primordial power spectrum* we will evolve forward in time. The missing term in Eq. (1.99) will eventually account for a geometrical factor (not a random field) describing the change in amplitude for a given mode $\mathbf{k}$ in time.

The *primordial power spectrum* is defined to be the one at the end of the inflation. The scalar potential $\Phi\left(\mathbf{x}\right)$, as set at the end of the inflation, is characterised by a nearly scale-invariant (dimensionless) power spectrum

$$k^{3}P_{\Phi}\left(k\right)=A^{2}\,k^{n_{s}-1}.,\qquad\left[P_{\Phi}\right]\sim\text{length}^{3},\qquad\left[k\right]\sim\text{length}^{-1}. \tag{1.100}$$

In Eq. (1.100), n_s is the *scalar spectral index*, which is surprisingly measured to be $n_s = 0.9649 \pm 0.0042$ (Planck Collaboration [2018] for the latest measurement). If n_s were exactly equal to 1, we would have a perfectly scale invariant primordial power spectrum, case which is well known as *Harrison-Zel'dovich power spectrum*. The potential $\hat{\Phi}(k)$ is related to the density perturbation $\hat{\delta}(k)$ by the Poisson equation

$$\frac{k^2}{a}\tilde{\Phi}(k) = 4\pi G\rho(t)\,\tilde{\delta}(k)\,. \tag{1.101}$$

Therefore, Eq. (1.101) can be used to relate the primordial matter power spectrum after the inflation to the one of the scalar potential Φ as

$$\left\langle \tilde{\delta}(\mathbf{k}, t_i)\,\tilde{\delta}(\mathbf{k}', t_i)\right\rangle_{\text{lin.}} \propto P^{\text{lin.}}(k, t_i) \propto \left[\frac{2\,k^2\,a(t_i)}{3\,\Omega_{m,o}^2\,(H_o)^2}\right]^2 P_\Phi(k) \propto \left[\frac{2\,A\,a(t_i)}{3\,\Omega_{m,o}^2\,H_o^2}\right]^2 k^{n_s} \tag{1.102}$$

The geometrical factor missing in front of Eq. (1.99) is called *transfer function* $\mathcal{T}(k)$ (up to a normalisation factor we will discuss later). It encapsulates the evolution of perturbations through different cosmic epochs up to $a \approx 0.1$ ($z \approx 9$), time domain after which the parametrization in Eq. (1.99) is allowed [Dodelson, 2003]. $\mathcal{T}(k)$ is in general evaluated numerically within well known *Einstein-Boltzmann solvers* like CAMB[3], or CLASS[4]. However, efficient analytical approximations do exist and we will employ for our analysis the one of Eisenstein & Hu [1998]. For completeness, we also refer to Bond & Efstathiou [1984] for a different possible fitting function.

Summarising our considerations, the final expression for the linear matter power spectrum will be

$$P^{\text{lin.}}(k, t) = (A^{\text{Pl.}})^2 \left[\frac{2\,D_+(t)\,\mathcal{T}(k)\;a_i}{3\,\Omega_{m,o}^2\,H_o^2}\right]^2 \left(\frac{k}{k_o}\right)^{n_s} \equiv (A^{\sigma_8})^2\left[\mathcal{T}(k)\,D_+(t)\right]^2 k^{n_s}. \tag{1.103}$$

where we included the previously missing prefactor. In Eq. (1.103) we provide two possible definitions for it. In the first case, $A^{\text{Pl.}}$ is the amplitude of the primordial power spectrum of the scalar perturbations as produced by the Inflation according to the Planck papers' convention [Planck Collaboration, 2014, 2016, 2018] (where also the definition of the pivot scale k_o is provided). It can be predicted theoretically and measured from CMB data to constrain inflationary scenarios. In the second case, the normalisation prefactor A^{σ_8} is measured from the variance σ_8 of the linear matter field at present time (t_o) once smoothed over spheres of (comoving) radius of 8 Mpc/h

$$\sigma_8 = \sigma^2(R)|_{R=8\ \text{Mpc}/h} \equiv \int_o^{+\infty} \left[\Delta^{\text{lin.}}(k, t_o)\right]^2 \tilde{W}_R^2(k)\,d(\log k) \;\Big|_{R=8\ \text{Mpc}/h} \tag{1.104}$$

where $\left[\Delta^{\text{lin.}}(k, t)\right]^2$ is the dimensionless power spectrum defined, according to the literature, as

$$\left[\Delta^{\text{lin.}}(k, t)\right]^2 = \frac{k^3}{2\pi^2} P^{\text{lin.}}(k, t) \tag{1.105}$$

[3] https://camb.info
[4] http://class-code.net

and $\hat{W}_R(k)$ is the Fourier transform of a top hat function with characteristic length R

$$\begin{cases} W_{\mathrm{R}}(\mathbf{x}) = 1, \text{ if } x \leqslant R, \\ W_{\mathrm{R}}(\mathbf{x}) = 0, \text{ if } x > R, \end{cases} \quad , \qquad \tilde{W}_{\mathrm{R}}(q) = \frac{3\left(\sin(qR) - qR\cos(qR)\right)}{(qR)^3}. \tag{1.106}$$

The specific scale of $8\,\mathrm{Mpc}/h$ is interestingly used since it is the scale at which the matter field, once smoothed, has a perturbation $\delta^{R=8\mathrm{Mpc}/h} \approx 1$: at present time this scale can be considered as the transition from the *linear* to the *non-linear* regime (see Reimberg et al. [2018] for an interesting discussion). For our applications, we will normalise the matter power spectrum via the σ_8 convention. We recall that we can smooth the matter field according to the top hat function (1.106) via a convolution in real space

$$\delta^R(\mathbf{x}) \equiv \int d^3\mathbf{y}\, W_{\mathrm{R}}(\mathbf{x} - \mathbf{y})\ \delta(\mathbf{y}). \tag{1.107}$$

Eq. (1.107) also defines our notation for smoothed cosmological fields. In Sec. 1.3.4 we will use this definition extensively.

1.3.3 The matter field bispectrum

As well explained in Sec. 1.2.3, a non-vanishing connected odd-point correlation function would be the signal of a deviation from the Gaussian statistics for the related field. Then, the lowest order correlation function which can be used to address this property is the 3-point one, whose associated polyspectrum is usually dubbed *bispectrum* and denoted with the letter B

$$B_{\delta_{\mathrm{f}}}(\mathbf{k}_1, \mathbf{k}_3, \mathbf{k}_1) \equiv P_{\delta_{\mathrm{f}}}(\mathbf{k}_1, \mathbf{k}_3, \mathbf{k}_1)\,. \tag{1.108}$$

In this section the time dependence of the fields is implied, not being relevant.

Bispectrum at leading order: $\langle\delta\delta\delta\rangle$ term We want to derive its expression at the leading non-vanishing leading order in PT, in terms of doublet expansion (1.96). We start in this paragraph with the matter density perturbation δ. Of course, the bispectrum will not emerge at the linear level (in the following considered as order $p = (1)$): non-Gaussianities are strictly related to non-linearities in the dynamical equations and coupling between different modes are required to have a non-vanishing bispectrum. The leading order at which the bispectrum in PT will appear is when at least one mode in the connected correlation function is in the non-linear regime, then at second order $p = (2)$

$$\begin{aligned} \langle\tilde{\delta}(\mathbf{k}_1)\,\tilde{\delta}(\mathbf{k}_2)\,\tilde{\delta}(\mathbf{k}_3)\rangle_c &= \Big\langle \sum_p \tilde{\delta}^{(p)}(\mathbf{k}_1) \sum_e \tilde{\delta}^{(e)}(\mathbf{k}_1) \sum_{\mathrm{f}} \tilde{\delta}^{(f)}(\mathbf{k}_1) \Big\rangle \\ &\underset{\text{leading}}{=} \langle\tilde{\delta}^{(1)}(\mathbf{k}_1)\,\tilde{\delta}^{(1)}(\mathbf{k}_2)\,\tilde{\delta}^{(2)}(\mathbf{k}_3)\rangle_c + \text{ cycles over } \{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3\} \end{aligned} \tag{1.109}$$

From Eq. (1.98), we have

$$\tilde{\delta}^{(2)}(\mathbf{k}) = \Upsilon_a^{(2)}(\mathbf{k}) = \int \frac{d^3\mathbf{q}}{(2\pi)^3} F^{(2)}(\mathbf{q}, \mathbf{k}-\mathbf{q})\tilde{\delta}^{(1)}(\mathbf{q})\tilde{\delta}^{(1)}(\mathbf{k}-\mathbf{q}). \tag{1.110}$$

For an EdS Universe, the kernel F^2 takes the form

$$F^{(2)}(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{1}{2}\mu(\mathbf{k}_1, \mathbf{k}_2)\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{2}{7}\mu^2(\mathbf{k}_1, \mathbf{k}_2) \tag{1.111}$$

with

$$\mu(\mathbf{k}_1, \mathbf{k}_2) = \frac{\mathbf{k}_1 \cdot \mathbf{k}_1}{k_1 k_2} = \frac{k_3^2 - k_1^2 - k_2^2}{2k_1 k_2}. \tag{1.112}$$

We then plug Eq. (1.110) into Eq. (1.109). For example, the first term reads

$$\left\langle \tilde{\delta}^{(1)}(\mathbf{k}_1)\, \tilde{\delta}^{(1)}(\mathbf{k}_2)\, \tilde{\delta}^{(2)}(\mathbf{k}_3) \right\rangle_c = \int \frac{d^3\mathbf{q}}{(2\pi)^3} F^{(2)}(\mathbf{q}, \mathbf{k}_3 - \mathbf{q}) \left\langle \tilde{\delta}^{(1)}(\mathbf{k}_1)\, \tilde{\delta}^{(1)}(\mathbf{k}_2)\, \tilde{\delta}^{(1)}(\mathbf{q}) \tilde{\delta}^{(1)}(\mathbf{k}_3 - \mathbf{q}) \right\rangle_c. \tag{1.113}$$

Since all the linearly evolved modes $\tilde{\delta}^{(1)}$ are drawn from a Gaussian distribution, we can apply the Wick theorem on the 4-point correlation function appearing in Eq. (1.113), which leads to the following 3 terms

$$\begin{aligned}
&\left\langle \tilde{\delta}^{(1)}(\mathbf{q})\, \tilde{\delta}^{(1)}(\mathbf{k}_3 - \mathbf{q}) \right\rangle_c \left\langle \tilde{\delta}^{(1)}(\mathbf{k}_1)\, \delta^{(1)}(\mathbf{k}_2) \right\rangle_c = \delta_{\mathrm{D}}(\mathbf{q} + \mathbf{k}_3 - \mathbf{q})\, \delta_{\mathrm{D}}(\mathbf{k}_1 + \mathbf{k}_2)\, P^{\mathrm{lin}}(q)\, P^{\mathrm{lin.}}(k_1) \\
&\left\langle \tilde{\delta}^{(1)}(\mathbf{k}_3 - \mathbf{q})\, \tilde{\delta}^{(1)}(\mathbf{k}_2) \right\rangle_c \left\langle \tilde{\delta}^{(1)}(\mathbf{q})\, \tilde{\delta}^{(1)}(\mathbf{k}_1) \right\rangle_c = \delta_{\mathrm{D}}(\mathbf{k}_3 - \mathbf{q} + \mathbf{k}_2)\, \delta_{\mathrm{D}}(\mathbf{q} + \mathbf{k}_1)\, P^{\mathrm{lin}}(\mathbf{k}_3 - \mathbf{q})\, P^{\mathrm{lin.}}(q) \\
&\left\langle \tilde{\delta}^{(1)}(\mathbf{k}_3 - \mathbf{q})\, \tilde{\delta}^{(1)}(\mathbf{k}_1) \right\rangle_c \left\langle \tilde{\delta}^{(1)}(\mathbf{q})\, \tilde{\delta}^{(1)}(\mathbf{k}_2) \right\rangle_c = \delta_{\mathrm{D}}(\mathbf{k}_3 - \mathbf{q} + \mathbf{k}_1)\, \delta_{\mathrm{D}}(\mathbf{q} + \mathbf{k}_2)\, P^{\mathrm{lin}}(\mathbf{k}_3 - \mathbf{q})\, P^{\mathrm{lin.}}(q)\,.
\end{aligned} \tag{1.114}$$

If we perform the integration over $\mathbf{q}$ in Eq. (1.113) and we apply the bispectrum definition as a particular case of Eq. (1.54)

$$\left\langle \tilde{\delta}(\mathbf{k}_1)\, \tilde{\delta}(\mathbf{k}_2)\, \tilde{\delta}(\mathbf{k}_3) \right\rangle_c = (2\pi)^3\, \delta_{\mathrm{D}}(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)\, B^{\mathrm{PT}}_{\delta\delta\delta}(k_1, k_2, k_3) \tag{1.115}$$

we are left with the following expression for the leading order contribution to the bispectrum in PT

$$B^{\mathrm{PT}}_{\delta\delta\delta}(k_1, k_2, k_3) = 2F^{(2)}(\mathbf{k}_1, \mathbf{k}_2) P^{\mathrm{lin}}(k_1)\, P^{\mathrm{lin}}(k_2) + \text{cycles over } \{\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3\} \tag{1.116}$$

Bispectrum at leading order: $\langle \delta\delta\theta \rangle$ term For later convenience, we would like to give one example of bispectrum including the velocity divergence field θ, specifically

$$\begin{aligned}
\left\langle \tilde{\delta}(\mathbf{k}_1)\, \tilde{\delta}(\mathbf{k}_2)\, \theta(\mathbf{k}_3) \right\rangle_c = &\left\langle \sum_p \tilde{\delta}^{(p)}(\mathbf{k}_1) \sum_e \tilde{\delta}^{(e)}(\mathbf{k}_1) \sum_{\mathrm{f}} \theta^{(f)}(\mathbf{k}_1) \right\rangle \underset{\text{leading}}{=} \\
&\left\langle \tilde{\delta}^{(2)}(\mathbf{k}_1)\, \tilde{\delta}^{(1)}(\mathbf{k}_2)\, \theta^{(1)}(\mathbf{k}_3) \right\rangle_c + \left\langle \tilde{\delta}^{(1)}(\mathbf{k}_1)\, \tilde{\delta}^{(2)}(\mathbf{k}_2)\, \theta^{(1)}(\mathbf{k}_3) \right\rangle_c + \left\langle \tilde{\delta}^{(1)}(\mathbf{k}_1)\, \tilde{\delta}^{(1)}(\mathbf{k}_2)\, \theta^{(2)}(\mathbf{k}_3) \right\rangle_c
\end{aligned} \tag{1.117}$$

Shortly, by following the same steps as in the previous paragraph and introducing the divergence velocity kernel (for an EdS Universe)

$$G^{(2)}\left(\mathbf{k}_1, \mathbf{k}_2\right) = \frac{3}{7} + \frac{1}{2}\mu(\mathbf{k}_1, \mathbf{k}_2)\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{4}{7}\mu^2(\mathbf{k}_1, \mathbf{k}_2), \tag{1.118}$$

we can derive the expression for the bispectrum $\langle \delta\delta\theta \rangle$

$$B^{\mathrm{PT}}_{\delta\delta\theta}(k_1, k_2, k_3) = \\ 2G^{(2)}(\mathbf{k}_1, \mathbf{k}_2)P^{\mathrm{lin}}\left(k_1\right)P^{\mathrm{lin}}\left(k_2\right) + 2F^{(2)}(\mathbf{k}_1, \mathbf{k}_3)P^{\mathrm{lin}}\left(k_1\right)P^{\mathrm{lin}}\left(k_3\right) + 2F^{(2)}(\mathbf{k}_2, \mathbf{k}_3)P^{\mathrm{lin}}\left(k_2\right)P^{\mathrm{lin}}\left(k_3\right). \tag{1.119}$$

In the previous derivation, we are assuming that the initial condition for the velocity field θ are actually proportional to those for the matter field density contrast δ (adiabatic initial condition request).

Comparison with simulations N-body simulations are our most precise tool to reproduce cosmological structure formation. We therefore look at them as the best way to inspect the predictive power and the precision of analytical estimations of cosmological observables. In this paragraph we focus on the matter bispectrum as derived from Perturbation Theory. A thorough and recent comparison between perturbation theory predictions and simulations can be found in Lazanu et al. [2016]. It is important to remind that the theoretical framework presented in this section is usually known as Eulerian standard perturbation theory (SPT). While further improvements are available in the literature, the material presented so far is enough for a thorough understanding of the contents of this manuscript. However, Lazanu et al. [2016] do also compare some of these alternative approaches with simulations and they eventually perform better than the tree-level SPT we use here. For simplicity and as not to break the flow of the section, we prefer here to stick with SPT results, even though we introduce a natural improvement to the tree-level results. In fact, we can enrich our estimation by moving to the next non-vanishing order which turns out to be of order 6 in terms of the linearly-evolved field $\delta^{(1)}$. This is usually known as 1-loop correction [Scoccimarro et al., 1998]. Assuming a ΛCDM cosmology, at $z = 0$ the tree-level bispectrum[5] $B(k_1, k_2, k_3)$is in good agreement with the simulations for $k \lesssim 0.1\, h\,\mathrm{Mpc}^{-1}$ on equilateral configurations $[B(k, k, k)]$, and on squeezed configurations $[B(1.4k, 1.4k, 0.14k)]$. The performance improves, at the same redshift, when considering flattened configuration $[B(k, 0.5k, 0.5k)]$ where the results from the simulations are well reproduced up to $k \lesssim 0.3\, h\,\mathrm{Mpc}^{-1}$. As expected, SPT performs better at higher redshifts where the level of non-linearities is lower. At $z = 2$, the tree-level bispectrum on equilateral, squeezed and flattened configurations is well predicted respectively up to $k \lesssim 0.16\, h\,\mathrm{Mpc}^{-1}$, $k \lesssim 0.2\, h\,\mathrm{Mpc}^{-1}$ and $k \lesssim 0.4\, h\,\mathrm{Mpc}^{-1}$. The 1-loop correction helps in converging towards the simulations. Over the scales analysed ($0.07\, h\,\mathrm{Mpc}^{-1} \leqslant k \leqslant 0.4\, h\,\mathrm{Mpc}^{-1}$) the greatest benefit, with respect to the tree-level bispectrum, is achieved at $z = 0$ where the analytical predictions are well within the error bars of the simulated signal up to $k \lesssim 0.11\, h\,\mathrm{Mpc}^{-1}$, $k \lesssim 0.17\, h\,\mathrm{Mpc}^{-1}$ and $k \lesssim 0.2\, h\,\mathrm{Mpc}^{-1}$ for equilateral, squeezed and flattened configurations respectively. In a subsequent paper, Lazanu & Liguori [2018] computed and tested

[5] We will always consider in this paragraph $B(k_1, k_2, k_3) \equiv B^{\mathrm{PT}}_{\delta\delta\delta}(k_1, k_2, k_3)$.

against simulations the 2-loop SPT correction to the matter bispectrum. At $z = 0$ the analytical results match the simulated signal within the error bars up to $k \lesssim 0.2\,h\,\mathrm{Mpc}^{-1}$ for equilateral configurations. Unfortunately, for flattened and squeezed bispectra, the 2-loop correction does not improve the agreement with the simulations. Overall, the best performance is achieved by the renormalised MPTbreeze technique [Scoccimarro, 2001; Crocce & Scoccimarro, 2006a,c, 2008; Bernardeau et al., 2008; Crocce et al., 2012; Bernardeau et al., 2012] which falls outside the SPT formalism we considered. At $z = 0$ and for equilateral configurations, MPTbreeze is competitive with the 2-loop SPT while remarkably improves the estimation for squeezed $[B(k, k, \Delta k = 0.012\,h\,\mathrm{Mpc}^{-1})]$ and flattened $[B(k, 0.5k, 0.5k)]$ configurations. In these last cases, the agreement between simulations and analytical predictions is guaranteed up to scales $k \lesssim 0.25\,h\,\mathrm{Mpc}^{-1}$.

1.3.4 The Halo Model: properties of the halos

The halo model (Cooray & Sheth [2002] for a thorough review) provides an effective and physically motivated ansatz to compute the different polyspectra we will need for our study at much smaller scales than PT allows. This approach relies on the fact, supported by numerical simulations, that we can model the statistical properties of the matter field via halos of dark matter of different masses, redshifts and positions. The reason for going beyond the PT approach described in the previous section will be clearer in the next chapter when we will analyse the weak lensing formalism. Briefly, due to the projection effect, the weak lensing signal is sensitive to very small scales. For a Euclid-like survey, Kitching & Taylor [2011] proved that the power spectrum has to be accurately known to 1% down to $k \approx 50\,h\,\mathrm{Mpc}^{-1}$ to saturate the dark energy figure of merit. This study is in agreement with the previous one from Doré et al. [2009]. Similarly, Huterer & Takada [2005] argued that we typically need the power spectrum within a few percent accuracy up to $k \approx 10\,h\,\mathrm{Mpc}^{-1}$ in order to take advantage of the statistical power of DES- and LSST-like survey. In Eifler [2011]; Hearin et al. [2012] a similar conclusion is obtained. At these scales linear theory for the evolution of the matter perturbations can not be trusted to build our observables. On the other hand, perturbative approaches (2-loop perturbations theory) starts deviating by more than 1% at scales $k \geqslant 0.1\,h\,\mathrm{Mpc}^{-1}$ at $z = 0$ [Taruya et al., 2012], although they perform slightly better at higher redshift. In the above regime cosmological N-body simulations are usually employed to study the nonlinear gravitational evolution and eventually used to tune phenomenological models or fitting formulae for the power spectrum. In Sec.2.2.1 we will extensively describe the inadequacy of N-body simulations for the purpose of our analyses. Therefore, we will shift our attention to the (semi-)analytical models mentioned above. To cite few examples, Peacock & Dodds [1996] provided a fitting formula for the power spectrum based on a scaling ansatz presented in Hamilton et al. [1991]. Later, Smith et al. [2003] proposed a new model of the power spectrum, the so-called halofit model, which is based on a the well known halo model of structure formation (e.g. Ma & Fry [2000] ; Seljak [2000]; Cooray & Hu [2001]). Recently Mead et al. [2015] presented an optimised variant of the halo model, designed to produce accurate matter power spectra well into the non-linear regime for a wide range of cosmological models, including baryonic feedback. However, all these approaches are meant to provide a good fit for the power spectrum alone while providing a poor description of the true underlying physics [Reimberg et al., 2018]. In particular, they can

not be used for building higher order statistics, which are key ingredients of our analyses. We rather build our observables in the standard halo model framework. Recently, Kayo et al. [2013] proved the halo model to have a precision of 20% up to scales $k \approx 10\, h\,\mathrm{Mpc}^{-1}$ at $z = 0$ in terms of weak lensing convergence power spectrum and bisepctrum and of the joint covariance.

SINGLE HALO PROPERTIES: THE SPHERICAL COLLAPSE Interestingly, the spherical collapse is one of the few cases where the dynamics of perturbations can be solved exactly. At some initial time t_i, we smooth the matter field $\delta(\mathbf{x}, t_i) \overset{\text{smooth}}{\rightarrow} \delta^{r_i}(t_i)$ via a spherical top-hat function $W_{r_i}(\mathbf{x})$ (see Eqs. (1.106)-(1.107) for useful definitions). In $\delta^{r_i}(t_i)$ we omitted the spatial coordinate since we will now focus on the dynamics of a single halo. The mass enclosed in a shell of radius r_i, can be related to the background matter density of the Universe $\rho(t_i)$ as

$$m_{<r_i} = \frac{4}{3}\pi r_i^3 \rho(t_i)\left(1+\delta^{r_i}(t_i)\right) = \frac{4}{3}\pi r(t)^3 \rho(t)\left[1+\delta^{r(t)}(t)\right] = m_{<r(t)} \tag{1.120}$$

where the second relation comes from the mass conservation $\rho_i(t_i) r_i^3 = \rho(t) r(t)^3$ along the different cosmological epochs. Once we will have solved the system for the time dependence of the radius $r(t)$, we will know the dynamics of the matter over-density enclosed in the sphere via

$$1+\delta^{r(t)} = \frac{\rho^{r_i}(t)}{\rho(t)} = \frac{1}{\rho(t)}\frac{3m_{<r_i}}{4\pi r^3(t)} \tag{1.121}$$

where we introduce the smoothed mean matter density within the sphere $\rho^{r_i}(t)$. As a consequence of the Birkhoff's theorem, a spherically symmetric matter distribution outside a sphere exerts no force on the sphere itself leading to the following equation of motion

$$\frac{\mathrm{d}^2 r(t)}{\mathrm{d}^2 t} = -G\frac{m_{<r(t)}}{r^2(t)} \approx -G\frac{m_{<r_i}}{r^2(t)}, \tag{1.122}$$

the energy being defined as

$$\frac{1}{2}\left[\frac{\mathrm{d}r(t)}{\mathrm{d}t}\right]^2 - G\frac{m_{<r_i}}{r(t)} = \mathcal{E} \tag{1.123}$$

For a bounded halo, $\mathcal{E} < 0$ and a parametric solution is known

$$r(\theta) = A(1-\cos\theta), \qquad A = \frac{G\, m_{<r_i}}{-2\mathcal{E}}, \tag{1.124}$$

$$t(\theta) = B(\theta - \sin\theta), \qquad B = \frac{G\, m_{<r_i}}{(-2\mathcal{E})^{2/3}} \tag{1.125}$$

where we used the fact that the mass is conserved. The spherical collapse solution $r_{\text{sp.}}(t)$, goes through 3 different phases

shells expand from $r^{\text{sp.}}(0) = 0$ at $\theta = 0$

shells reach a maximum expansion at $r_{\text{max}}^{\text{sp.}}$ at $\theta = \pi$, being $t_{\text{max}} = \pi B$ at this point

shells collapses back to $r^{\text{sp.}} = 0$ at $\theta = 2\pi$, being $t = 2t_{\text{max}}$ at this point.

In particular, this dynamics is typical of every halo, no matter its mass. By using Eq. (1.121), it can be shown that can write the time evolution of the smoothed over-density within the spheres as

$$\delta_{\text{sp.}}^{r(\theta)} = \frac{\rho^{r_i}(\theta)}{\rho(\theta)} = \frac{9(\theta - \sin\theta)^2}{2(1-\cos\theta)^3} - 1 \tag{1.126}$$

with $\rho(\theta) = (6\pi G t^2(\theta))^{-1}$ in EdS models. We specified with the subscript sp. that the solution in Eq. (1.126) is the solution for the over-density of halos undergoing a spherical collapse.

The model considered presents a singularity at $\theta = 2\pi$. However, halos tend not to collapse any further at $r_{\text{v}} = r_{\text{max}}/2$: in line with the virial theorem for a gravitationally bound system of particles, this equilibrium corresponds to the equi-partition between kinetic and potential energy. We define the over-density within a virialised halo δ_{v} as that at the theoretical collapse time

$$\delta_{\text{v}} = \frac{1}{\rho(t_{\text{col.}})} \frac{3m_{<r_i}}{4\pi(2r_{\text{max}})^3} \tag{1.127}$$

Assuming and EdS cosmology, the classical solution is $1 + \delta_{\text{v}} = 18\pi^2$. However, more sophisticated solution can be found for more general cosmologies [Henry, 2000]. For a spatially flat model, with negligible radiation (as it for the redshifts of interest) the following fitting formula holds

$$\delta_{\text{v}}(z) = 18\pi^2 \left(1 + 0.4093\, x^{2.71572}\right), \qquad x = \frac{\left(\frac{1}{\Omega_{\text{m},o}} - 1\right)^{\frac{1}{3}}}{1+z} \tag{1.128}$$

In particular, we will use this latest advanced analytical expression in our implementation.

HALO DISTRIBUTION ON LARGE SCALES: THE MASS FUNCTION The mass function for a given *tracer h* is defined as the number of objects tracers per comoving unit of volume and unit mass interval

$$f_{\text{m}}(z, m) \equiv \frac{\mathrm{d}N_{\text{h}}(z, m)}{\mathrm{d}V\, \mathrm{d}m} \tag{1.129}$$

The computation of the mass function of dark matter halos is a central problem in modern Cosmology. The formation and evolution of dark matter halos is a highly complex dynamical process, and a detailed understanding of it can only come through large-scale N-body simulations. Some analytical understanding is however very desirable, both for obtaining a better physical intuition, and for the flexibility under changes of models or parameters (such as cosmological model, shape of the non-Gaussianities, etc.) which is of paramount importance in the approach we will employ in our work where covariance matrices for the weak lensing observables (see chapter 3) will be computed analytically. In the following part of the section we will review the main efforts towards an

analytical modelling of the mass function for dark matter halos. In the end, we will focus on the model included in our implementation.

THE MASS FUNCTION: PRESS & SCHECHTER APPROACH First, let us define the following quantity

$$\delta_{\rm lin}\left(t\right) \equiv \left[\lim_{\theta\to 0}\delta_{\rm sp.}^{r(\theta)}\right]\frac{D_+\left(t\right)}{D_+\left(\bar{t}\right)} = \frac{3}{20}\left(6\pi\right)^{2/3}\left(\frac{t}{t_{\rm max}}\right)^{\frac{2}{3}} \quad \text{EdS model,} \tag{1.130}$$

This quantity is the value of the smoothed over-density within a halo as if the corresponding initial over-density $\lim_{\theta\to 0}\delta^{r(\theta)}$ were linearly evolved up to a time t. In this case, the time $\bar{t}$ is a reference time we assume to be in regime of linear clustering. In particular, it is the time at which we assume the density field is smoothed in halos which will then collapse. Therefore, we defined it to be $t(\theta = 0) = \bar{t}$. Specifically, at the time of the evolution corresponding to the collapse time of a halo we will have, and define,

$$\delta_{\rm sp.}^{c} \equiv \delta_{\rm lin}\left(t_c = 2t_{\rm max.}\right) = \frac{3}{20}\left(12\pi\right)^{2/3} \quad \text{EdS model.} \tag{1.131}$$

This quantity is usually defined as the *barrier for the spherical collapse*. In Eq. (1.127) we computed the over-density within a formed halo $\delta_{\rm v}$. Once again, this quantity does not depend on the mass, nor the related scale. However, if we perform an observation at a given redshift z, what is the condition for an object to be formed by that observation time? It is important to ask ourselves this question given that we want to account in our model for virialised objects. Eq. (1.130) and Eq. (1.131) are telling us that, if we linearly evolve a smoothed region of the Universe in the initial condition space up to a given time t, if its observed overdensity $\delta^{\rm R}(t)$ is larger than $\delta_{\rm sp.}^{c}$, then we know that the associated halo, undergoing a spherical collapse instead, will have had the time to collapse by t, the observation time. The Press & Schechter [1974] (PS) formalism is based on this concept. Before starting, let us simply change the notation to be in line with the one in the literature. Specifically, instead of having a smoothed density field linearly evolved from a reference time $\bar{t}$ to be compared with the constant barrier $\delta_{\rm sp.}^{c}$, we will consider as static the former (we will omit the time dependence) while we will consider a dynamical barrier as

$$\delta_{\rm sp.}^{c} \to \delta_{\rm sp.}^{c}\left(z\right) \equiv \frac{\delta_{\rm sp.}^{c}}{D_+(z)}, \qquad D_+(\bar{t}) = 1 \tag{1.132}$$

At a given reference time $\bar{t}$, we identify the seeds for halo of different initial radius R via a top-hat smoothing of the (linear) matter field

$$\delta^{\rm R} = \int d^3\mathbf{x}'\delta(\mathbf{x}')W_{\rm R}\left(\mathbf{x}-\mathbf{x}'\right) \tag{1.133}$$

The shape of the filter is given in Eq. (1.106) and we can define an associated mass $M(R) \equiv 4\pi/3\rho\left(t\right)R^3$ which will be easily related (will actually be) to the mass of the associated halo. From now on, we will freely use either M or R to define the size of the filter. The shape of the filter as a sharp top-hat in real space allows for a physically motivated association between the two masses. However, for different filters this will not be the case anymore,

as we will see later. By changing the filter scale, we will be considering halos of different mass. However, as we saw in the previous paragraph, the dynamical evolution of the associated overdensity δ^{R} will not depend on these quantities. This is an extremely valuable feature of the spherical collapse model. The variance of the field δ^{R} is then

$$\sigma^2(R) = \sigma^2(M) = \frac{1}{2\pi^2}\int P^{\mathrm{lin.}}(k)\,\hat{W}_R(k)\,k^2 dk \tag{1.134}$$

If the original field δ is Gaussian, so will be the smoothed field δ^{R}

$$\mathcal{P}\left(\delta^{\mathrm{R}}\right)\mathrm{d}\delta^{\mathrm{R}} = \frac{1}{\sqrt{2\pi}\sigma^2(R)}e^{-\frac{\delta^{\mathrm{R}}}{2\sigma^2(R)}}\mathrm{d}\delta^{\mathrm{R}} \tag{1.135}$$

with the variance given by Eq. (1.134). Following the procedure outlined by PS, we state that *the fraction of mass contained within virialised halos of mass $m_v > M$, $F^{PS}(>M)$, at a certain* observation *time t, is equal to the probability that $\delta^M > \delta^{\mathrm{c}}_{\mathrm{sp.}}(t)$*, i.e.

$$\mathcal{P}\left(\delta^M > \delta^{\mathrm{c}}_{\mathrm{sp.}}(t)\right) = \int_{\delta^{\mathrm{c}}_{\mathrm{sp.}}(t)}^{\infty}\mathcal{P}(\delta_R)\,\mathrm{d}\delta_R = \frac{1}{2}\mathrm{ercf}\left[\frac{\delta^{\mathrm{c}}_{\mathrm{sp.}}(t)}{2\sigma(R)}\right] \equiv F^{\mathrm{PS.}}(>M) \tag{1.136}$$

We recall that the mass function (1.129) is defined as as the number of objects per comoving unit of volume and unit mass interval. To derive its expression, we start by differentiating $\mathrm{d}F^{PS}(>M)/\mathrm{d}m_v$ to find the contribution to the total mass fraction $F^{PS}(>M)$ from halos of a given mass m_v. By multiplying this results for the comoving background density $\rho_{\mathrm{com.}}$ we obtain the total mass (not the fraction) as given by halos of mass m_v. If we finally divide for the mass itself m_v we finally obtain the density we are aiming at. The final expression will then be

$$f^{\mathrm{PS.}}_{\mathrm{m}}(z,m) = \frac{\rho_{\mathrm{com.}}}{m}\frac{\mathrm{d}F^{PS}(>M)}{\mathrm{d}m} = \frac{1}{\sqrt{2\pi}}\frac{\rho_{\mathrm{com.}}}{m^2}\frac{\delta^{\mathrm{c}}_{\mathrm{sp.}}(t)}{\sigma(R)}e^{-\left(\frac{\delta^{\mathrm{c}}_{\mathrm{sp.}}(t)}{2\sigma(R)}\right)^2}\left|\frac{\mathrm{d}\log\,\sigma(R)}{\mathrm{d}\,\log m}\right| \tag{1.137}$$

where we dropped the subscript v for the mass of fully formed halos.

At this point, we can also introduce another very common definition related to the halo concentration: the *multiplicity function* $f_v(v)$ [Jenkins et al., 2001]

$$f_{\mathrm{m}}(z,m)\,dm \equiv \frac{\rho_{\mathrm{com.}}}{m^2}f_v(v)\left|\frac{\mathrm{d}\log v}{\mathrm{d}\log m}\right|dm \tag{1.138}$$

where $v(z,m) = \delta^{\mathrm{c}}_{\mathrm{sp.}}(t)/\sigma(m)$. In particular, the multiplicity function (1.138) has the advantage that to a good accuracy it does not explicitly depend on redshift. For the PS mass function (1.137), the multiplicity function will be

$$f^{\mathrm{PS.}}_v(v) = \frac{1}{\sqrt{2\pi}}v(z,m)\,e^{-\frac{v^2(z,m)}{2}}. \tag{1.139}$$

Even though the PS result represents one of the first attempts to provide analytical insights on the issue of the

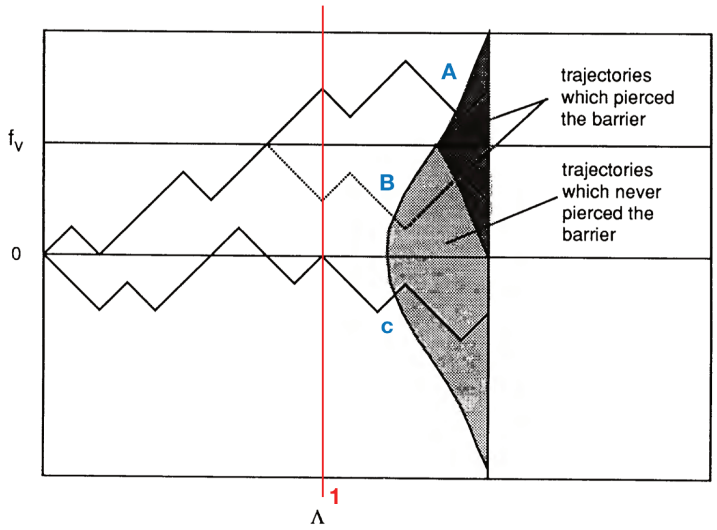

Figure 1.3.1: [Bond et al., 1991]: *This figure shows schematically examples of random walks of* $F(r, \Lambda)$ *with* Λ. *In the limit of infinitesimally fine steps in* Λ, *the distribution of trajectories at fixed* Λ *is Gaussian as indicated by the curve on the right of the figure. The heavily shaded fraction of this Gaussian above the barrier at* $F = f_v$, *represents the fraction of trajectories above the barrier at* Λ. *For such a trajectory, if one reflects in the barrier the portion that lies to the right of where it first pierces the barrier one obtains, for example the dotted trajectory. This new trajectory is just as likely as the one it was formed from. Consequently the heavily shaded region below the barrier, which is simply the reflection of the region above the barrier, represents the probability of a trajectory being below the barrier at A but having pierced the barrier at some lower value of A. Thus the lightly shaded fraction of the area under the Gaussian represents the fraction of trajectories which have not pierced the barrier at any lower value of* Λ.

cosmological mass function for dark matter halos, however it has a pathology. As a matter of fact, given that

$$\lim_{m\to 0} F^{\rm PS.}(>M) \to 0, \qquad \lim_{m\to\infty} F^{\rm PS.}(>M) \to 1/2. \tag{1.140}$$

never more than the 50% of all the matter in the Universe can collapse forming virialised halos. This comes from the assumption that only over-dense regions characterised by $\delta^M > \delta^{\rm c}_{\rm sp.}(t)$ at the level of initial conditions can collapse in halos by the time z. However, under-dense regions, which statistically enclose the 50% of the total mass of the Universe, can be included within the latter. This is called *cloud-in-cloud* problem (e.g. see Sec.2.2 of Peacock & Heavens [1990]). PS corrected for this effect by adding a fudge factor 2 to the derived mass and multiplicity function

$$f^{\rm PS.}_{\rm m}(z, m) \to 2 f^{\rm PS.}_{\rm m}(z, m)\,, \tag{1.141}$$

$$f^{\rm PS.}_{\nu}(\nu)^{\rm PS.} \to 2 f^{\rm PS.}_{\nu}(\nu)\,, \tag{1.142}$$

THE MASS FUNCTION : EXCURSION SET THEORY APPROACH Bond et al. [1991] found a formal solution to the

problem described in the previous section. The new approach is called *Extended Press-Schechter* (EPS) approach. They start their famous paper by focusing on all the possible trajectories of the overdensity of a given point identified in the real space at time $\bar{t}$, as function of the of the filter radius R. They define a 4-dimensional field (once again we will omit the time dependence $\bar{t}$)

$$\mathcal{F}^W(\mathbf{x}, R) = \int d^3\mathbf{x}' W_R(\mathbf{x} - \mathbf{x}')\,\delta(\mathbf{x}') \tag{1.143}$$

where $\mathcal{W}_R(\mathbf{x} - \mathbf{x}')$ is a smoothing function at some resolution R whose shape has to be decided. A one-dimensional example of the topology induced by Eq. (1.143) for a Gaussian field $\delta(\mathbf{x})$ is given in their paper at Fig.1 [Bond et al., 1991] for different choices of smoothing functions. Of great interest is a filter function W_R as a top-hat in Fourier space

$$W_R(\mathbf{r}) = \frac{3}{4\pi R^3}\frac{(\sin x - x\cos x)}{x^3}, \quad x = \frac{r}{R} \tag{1.144}$$

$$\hat{W}_R(\mathbf{k}) = \theta(1 - kR) \tag{1.145}$$

where θ here is the Heaviside function. In particular Eq. (1.144) allows an analytical derivation of the mass function within the EPS formalism. Specifically, for a fixed point in the initial condition space $\bar{\mathbf{x}}$, the trajectories identified by $(\sigma^2(R), \mathcal{F}^W(\bar{\mathbf{x}}, R))$ are Brownian random walks

$$\mathcal{F}^W(\bar{\mathbf{x}}, R) = \int d^3\mathbf{k} W_R(\mathbf{k})\,\tilde{\delta}(\mathbf{k})\,e^{i\mathbf{k}\cdot\bar{\mathbf{x}}} = \int_{k<\frac{1}{R}} d^3\mathbf{k}\tilde{\delta}(\mathbf{k})\,e^{i\mathbf{k}\cdot\bar{\mathbf{x}}} \tag{1.146}$$

In particular, Eq. (1.146) is telling us that our trajectories are characterised by a series of points which are obtained by increasing the resolution, i.e. the maximum mode $k^{\max} = 1/R$ for the convolution in Eq. (1.146). Therefore, a new step differs from the previous one via the inclusion of independent modes of the smoothed matter fields, which is Gaussian. Therefore, any new step is indeed Markovian by definition. This feature is not present with other type of filters. In Fig. 1.3.1, taken from the original paper [Bond et al., 1991] and where we add few graphics to help the explanation, we show the description of the matter collapse in terms of Markovian random walks and how the approach developed by Bond, J.R. et al. can solve the problem of the fudge factor 2. The link between the original terms in the caption with those introduced so far are: $\Lambda = \sigma^2(R)$ while $f_v = \delta^c_{\text{sp.}}(t)$, our usual dynamic barrier for the spherical collapse. More in details, each trajectory corresponds to a mass element identified at the reference time $\bar{t}$. According to the PS ansatz, the fraction of trajectories for which $\mathcal{F}^W(\bar{\mathbf{x}}, R_1) > f_v$, at a certain time t, will contribute to the fraction of mass elements in collapsed objects with mass $m > m(\Lambda_1)$. Specifically, trajectory A will contribute while B will not. However, given that the trajectory B is obtained from the trajectory A just by mirroring it around $f = f_v$, then it will be equally likely as A. Therefore, the actual fraction of mass in halos of $m > m(\Lambda_1)$ will be twice the fraction of trajectories for which $f > f_v$ at Λ_1. This is just a qualitative explanation of the power of the EPS approach and we refer to Bond et al. [1991] for a more rigorous derivation which will also allow to appreciate more the choice for k-sharp filter. In particular the EPS formalism solve the problem of the 2

factor places *ad hoc* by PS.

THE SHETH & TORMEN MASS FUNCTION In this paragraph we will consider one last attempt in improving the analytical expression for the mass function of dark matter halos (1.129). Furthermore, it will be the one actually implemented in our pipeline. Sheth et al. [2001] (ST) were capable to include the dynamics of the ellipsoidal collapse into the PS formalism via a dependence on the mass of the halo in the first-crossing barrier $\delta^{c}_{sp.}(t) \rightarrow \delta^{c}_{el}(t, m)$ of shape

$$\delta^{c}_{el}(z, m) = \delta^{c}_{sp.}(z) \left[1 + \beta \left(\frac{\sigma(m)}{\delta^{c}_{sp.}(z)} \right)^{2\gamma} \right] \tag{1.147}$$

While leaving a detailed description of the ellipsoidal collapse in the Zel'dovich approximation to dedicated papers [Bond & Myers, 1996; Zel'dovich, 1970; Monaco, 1995], we want to underline that the great virtue of the their approach is that, once the barrier shape in Eq. (1.147) is known, then the entire EPS formalism developed on the basis of the spherical collapse dynamics with a constant barrier, can be extended very easily. The final general result is the following expression for the multiplicity function (1.138)

$$\nu f^{ST.}_{\nu}(\nu) = 2A \left(1 + \frac{1}{\nu'^{2q}} \right) \left(\frac{\nu'^{2}}{2\pi} \right)^{\frac{1}{2}} e^{-\frac{\nu'^{2}}{2}} \tag{1.148}$$

with $A \sim 0.322$, $q = 0.3$ and $\nu' = a\nu$. In the original paper from Sheth et al. [2001] $a = 1$. We chose report here the expression with a generic factor a to accommodate for different fitted formulas which proved the analytical expression from ST accurate, up to a tiny dependence on how halos are identified in the simulation (see Sec. 4.1 in ST). For example, a well known parametrization of the mass function comes from Sheth & Tormen [1999], where $a = 0.707$. We will use this last fitted formula in our work.

1.3.5 THE HALO MODEL: MATTER CORRELATION FUNCTIONS

We conclude this chapter with a final section dedicated to the statistical description of the matter field in the Universe, within the halo model approach.

HALO BIAS First of all, the statistical properties of the dark matter halos we introduced so far are not direct proxy of those of the underlying matter field we are interested in. In general, matter halos are *biased tracers* of the underlying matter field. If we define the *halo density contrast* as [Mo & White, 1996; Mo et al., 1997]

$$\delta_{h}(\mathbf{x}, t) \equiv \frac{n_{h}(\mathbf{x}, t)}{\bar{n}_{h}(t)} - 1 \tag{1.149}$$

where $\bar{n}_{\rm h}(t)$ is the mean comoving halo number density and $n_{\rm h}(\mathbf{x}, t)$ is the actual one at position $\mathbf{x}$, then $\delta_{\rm h}(\mathbf{x}, t)$ is not equal to $\delta(t)$ and a more complex relation exists [Desjacques et al., 2018]

$$\delta_{\rm h}(\mathbf{x}, t) = \sum_O \left[b_O(t) + \varepsilon_O(\mathbf{x}, t)\right] \mathcal{O}(\mathbf{x}, t) + \varepsilon(\mathbf{x}, t) \tag{1.150}$$

In Eq. (1.149) and in Eq. (1.150) a dependence on the mass of the halos can eventually be added, then $\delta_{\rm h}(m, \mathbf{x}, t)$ and the biases will refer to halos of a specific mass. In the following we will omit such dependence. In Eq. (1.150), $\mathcal{O}$ are operators, or statistical fields, which describe properties of the halos' environment on which their density can depend. Each operator is multiplied by a corresponding *deterministic bias* parameter b_O, which, at fixed time, is merely a number, and by a *stochastic bias* $\varepsilon_O(\mathbf{x}, t)$ which is relevant for very small scales, along with $\varepsilon(\mathbf{x}, t)$. In our work, we will only account for the deterministic bias expansion and, more specifically, for operators expressed in term of power of the matter density contrast

$$\mathcal{O} = \delta^N, \qquad b_{\mathcal{O}=\delta^N} \equiv \frac{b_N}{N!}. \tag{1.151}$$

The b_N bias parameters are known as *local-in-matter-density* (LIMD) biases ($\propto(\nabla^2\Phi)^N$).

In order to calculate the value for the different terms in the deterministic LIMD expansion, the *peak-background split* (PBS) [Kaiser, 1984; Bardeen et al., 1986; Cole & Kaiser, 1989; Mo & White, 1996] approach is commonly used. According to this approach, the biases $b_N^{(A)}$ for halos included in a given region A (either the observable Universe itself or the volume accessible by our survey) of the Universe are determined via

$$b_N^{(A)}(t) = \frac{1}{\bar{n}_{\rm h}(t, \delta_{\rm b}(t))}\Big|_{\delta_{\rm b}(t)=0} \frac{\partial^N \bar{n}_{\rm h}\left[t, \delta_{\rm b}(t)\right]}{\partial \delta_{\rm b}(t)^N}\Big|_{\delta_{\rm b}(t)=0} \tag{1.152}$$

where $\delta_{\rm b}$ represents the mean matter fractional over-density within the above mentioned region, i.e.

$$\delta_{\rm b}(t) = \frac{1}{V_A}\int_{V_A} d^3\mathbf{x}\, \delta(\mathbf{x}, t) \tag{1.153}$$

and $\bar{n}_{\rm h}(m, z) = a^3 f_{\rm m}(m, z)$ is the physical mean number density of halos, per unit of mass. In particular we will see in Sec. 2.3.2 that also the scale factor a can indeed be shifted by the presence of a local over-density. Precisely, the biases defined in Eq. (1.152) are *Eulerian biases*, noted in the following as $b_N^E(t)$, and they refer to the time-evolved perturbations. However, they can also be defined at the level of initial conditions for the spherical collapse of halos. In this case they are called *Lagrangian biases* $b_N^L(t)$. The two are related, in the case of a spherical collapse in a EdS Cosmology, as

$$b_1^{\rm E}(t) = b_1^{\rm L}(t) + 1, \tag{1.154}$$

$$b_2^{\rm E}(t) = 2(1+a)b_1^{\rm L}(t) + b_2^{\rm L}(t), \qquad a = -\frac{17}{21} \tag{1.155}$$

where we focused on the first 2 orders. The local biases defined in Eq. (1.152), can actually be measured in simula-

tion via the *separate Universe-ansatz* [Dai et al., 2015; Wagner et al., 2015a] given that a specific large-scale mode δ_b can actually be implemented via a change in the cosmological parameters of the simulation [Bertacca et al., 2015]. For our purpose, i.e. an analytical evaluation of the covariance matrix for weak lensing observables, of much more interest is the possibility to analytically derive the LIMD biases from an analytical multiplicity function, as defined in Eq. (1.138). In this case, the biases are usually obtained in Lagrangian space and then mapped in the Eulerian ones via the relations in Eq. (1.154). For the mass function considered in our implementation, i.e. the ST mass function in Eq. (1.148), the first 2 order, mass dependent, Lagrangian biases are [Mo et al., 1997; Scoccimarro et al., 2001]

$$b_1^L(m,z) = \frac{\nu'^2 - 1}{\delta_{sp.}^c(z)} + \frac{2q/\delta_{sp.}^c(z)}{1 + (\nu'^2)^q} \tag{1.156}$$

$$b_2^L(m,z) = \frac{\nu'^2}{\delta_{sp.}^c(z)} \frac{\nu'^2 - 3}{\delta_{sp.}^c(z)} + \frac{1 + 2q}{\delta_{sp.}^c(z)} + 2\frac{\nu'^2 - 1}{\delta_{sp.}^c(z)}. \tag{1.157}$$

where the dependence of ν' over m and z is implied.

Statistical description of halos At this point, we have all the tools to first relate the matter clustering to the halo clustering and, second, to describe the statistical properties of the former via those of the latter. Before tackling this point, let us introduce few more definitions. Given a distribution of N matter halos, their positions $\{\mathbf{x}_i\} i = 1, \dots, N$ and masses $\{m_i\} i = 1, \dots, N$ are random variables and one realisation is characterised by a set of couples $\{\mathbf{x}_i, m_i\}_{i=1,\dots,N}$. In general we can determine the value of an observable for one specific realisation $\hat{O}(\mathbf{x}; \{\mathbf{x}_i, m_i\}_{i=1,\dots,N})$ and then get the expectation value by ensemble averaging over the joint halo position-mass distribution. While the joint theoretical distribution is not known, positions and masses should at least satisfy at every redshift, for the derivations in the present chapter, the following relations involving the halo mass function and the halo-halo correlation function $\varepsilon_{\delta_h}(\mathbf{x}' - \mathbf{x}''; m_1, m_2)$ (Cooray & Hu [2001])

$$\left\langle \sum_i \delta(m - m_i)\, \delta(\mathbf{x} - \mathbf{x}_i) \right\rangle_H \equiv f_m(m,z)\,, \tag{1.158}$$

$$\left\langle \sum_i \delta(\mathbf{x}' - \mathbf{x}_i)\delta(m_1 - m_i) \sum_j \delta(\mathbf{x}'' - \mathbf{x}_j)\delta(m_2 - m_j) \right\rangle_H \equiv f_m(m_1) f_m(m_2) \left[1 + \varepsilon_{\delta_h}(\mathbf{x}' - \mathbf{x}''; m1, m2; z)\right] \tag{1.159}$$

where $\left\langle \sum_i \dots \sum_j \dots \right\rangle_H$ is the average number of pairs of halos of mass m_1 and m_2 at mutual distance $\mathbf{x}' - \mathbf{x}''$. An interesting example of how we can apply the halo model formalism is the evaluation of the *mean background matter density* $\rho(z)$. Its value from a specific realisation is the sum over all the contributions from the different halos of the realisation

$$\hat{\rho}(\mathbf{x}, z) = \sum_i \rho_h(\mathbf{x} - \mathbf{x}_i; m_i; z) = \sum_i m_i u(\mathbf{x} - \mathbf{x}_i; m_i; z) \tag{1.160}$$

where $\rho\left(\mathbf{x}-\mathbf{x}_i;m_i;z\right)$ is the density profile for a single halo of mass m_i as function of the distance from the centre of the halo $\mathbf{x}_i$. In line with the tradition we also defined $u\left(\mathbf{x}-\mathbf{x}_i;m_i;z\right)\equiv\rho_{\rm h}\left(\mathbf{x}-\mathbf{x}_i;m_i;z\right)/m_i$. The final mean matter density can then be estimated via the halo ensemble average

$$\rho(z)=\left\langle\hat{\rho}\left(\mathbf{x},z\right)\right\rangle_{\rm H}=\left\langle\sum_i m_i\,u\left(\mathbf{x}-\mathbf{x}_i;m_i;z\right)\right\rangle_{\rm H}= \\ \int dm\,m\int d^3\mathbf{x}'\,u\left(\mathbf{x}-\mathbf{x}';m;z\right)\left\langle\sum_i\delta\left(\mathbf{x}'-\mathbf{x}_i\right)\delta\left(m-m_i\right)\right\rangle_{\rm H}=\int dm\,m\,f_{\rm m}\left(m,z\right) \quad (1.161)$$

where we used the request for a normalised halo density profile $\int d^3\mathbf{x}'\,u\left(\mathbf{x}-\mathbf{x}';m;z\right)m=m$.

HALO DENSITY PROFILE In the previous paragraph we introduced the *halo density profile* $\rho_{\rm h}\left(\mathbf{r}=\mathbf{x}-\mathbf{x}_i;m_i;z\right)$. For our applications we will use the well known *Navarro-Frank-White* (NFW) profile [Navarro et al., 1996]

$$\rho_{\rm NFW}\left(r,m\right)=\frac{\rho_s(z)}{\frac{r}{r_s(m,z)}\left(1+\frac{r}{r_s(m,z)}\right)^2} \quad (1.162)$$

In Eq. (1.162), $\rho_s(z)$ is actually function of the mass once normalized over the volume of the halo up to the virialization radius $r_v(m,z)$ (the one including the overdensity $\delta_{\rm v}(z)$ defined in Eq. (1.128)), r is the module of the distance from the center of the halo and $r_s(m,z)$ is chosen in order to provide the best fit to simulations. A more convenient way to describe the distribution of halos and fit the 2 left d.o.f. in the model, is the so called concentration parameter $c(m,z)=\frac{r_v(m,z)}{r_s(m,z)}$, proved to be drawn from a log-normal distribution

$$p\left(c,m,z\right)=\frac{1}{c\sqrt{2\pi}\sigma_{\ln c}}\exp\left(-\frac{\left(\ln c-\ln\bar{c}\left(m,z\right)\right)^2}{2\sigma^2_{\ln c}}\right) \quad (1.163)$$

where $\bar{c}\left(m,z\right)$ is the median concentration parameter for every redshift and mass. For our implementation, we will refer to the expressions in Bullock et al. [2001].

MATTER CORRELATION FUNCTIONS We build the $n-$point correlation function for the matter field via the usual halo ensemble average over all the possible realisations of the halo distribution

$$\varepsilon_\delta\left(\mathbf{x}_1,\ldots,\mathbf{x}_n\right)=\left\langle\hat{\varepsilon}_\delta\left(\mathbf{x}_1,\ldots,\mathbf{x}_n\right)\right\rangle_{\rm H}. \quad (1.164)$$

In these last calculations we will imply the time dependencies, being clarified in the previous paragraph. In particular ρ will simply refer to the matter background density $\rho(z)$. Specifically, we can build the product of n instances of the contrast density field for the single realisation

$$\hat{\varepsilon}_\delta\left(\mathbf{x}_1,\ldots,\mathbf{x_n}\right)=\prod_{i=1}^n\delta\left(\mathbf{x}_i\right)=\prod_{i=1}^n\left(\frac{\hat{\rho}(\mathbf{x}_i)}{\rho}-1\right), \quad (1.165)$$

$\hat{\rho}(\mathbf{x}_i)$ being from one realisation. Let us start from the simple case of the 2-point correlation function

$$\hat{\varepsilon}_\delta(\mathbf{x}_1, \mathbf{x}_2) = \delta(\mathbf{x}_1)\delta(\mathbf{x}_2) = \frac{1}{\rho^2}\left[\hat{\rho}(\mathbf{x}_1)\,\hat{\rho}(\mathbf{x}_2))\right] - 1 = \frac{1}{\rho^2}\sum_i m_i\, u(\mathbf{x}_1 - \mathbf{x}_i; m_i)\sum_j m_j\, u\left(\mathbf{x}_2 - \mathbf{x}_j; m_j\right) - 1 \quad (1.166)$$

where we used Eq. (1.160) to relate the value of $\hat{\rho}$ to the halo masses and positions within a particular realisation of the statistical process. Eq. (1.166) can now be split into the sum of two main components. The first one, dubbed 1-halo term, describes the situation where the two points in the correlator are within the same halo and it is dominant at small scales (within the halo virialisation radius). The second instead, the 2-halo term, comes from contributions where the two points are located in two different halos and it is important at larger scales.

$$\hat{\varepsilon}_\delta^{(1)}(\mathbf{x}_1, \mathbf{x}_2) = \left[\delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\right]^{1\text{-halo}} = \frac{1}{\rho^2}\left[\rho(\mathbf{x}_1)\,\rho(\mathbf{x}_2))\right]^{1\text{-halo}} - 1 = \frac{1}{\rho^2}\sum_i m_i\, u(\mathbf{x}_1 - \mathbf{x}_i; m_i)\sum_j m_j\, u\left(\mathbf{x}_2 - \mathbf{x}_j; m_j\right) \times \delta_{ij} - 1, \quad (1.167)$$

$$\hat{\varepsilon}_\delta^{(2)}(\mathbf{x}_1, \mathbf{x}_2) = \left[\delta(\mathbf{x}_1)\delta(\mathbf{x}_2)\right]^{2\text{-halo}} = \frac{1}{\rho^2}\left[\hat{\rho}(\mathbf{x}_1)\,\hat{\rho}(\mathbf{x}_2))\right]^{2\text{-halo}} - 1 = \frac{1}{\rho^2}\sum_i m_i\, u(\mathbf{x}_1 - \mathbf{x}_i; m_i)\sum_{j\neq i} m_j\, u\left(\mathbf{x}_2 - \mathbf{x}_j; m_j\right) - 1 \quad (1.168)$$

where in Eq. (1.167), we force the points $\mathbf{x}_1$ and $\mathbf{x}_2$ to be in the same halo by imposing $i = j$. In Eq. (1.168), on the contrary, we required them to be in 2 separate halos. By taking the the ensamble average of the 2 terms, we will end up with the expressions for the 1-halo and 2-halo term for the 2 point matter correlation function. In details, the 1-halo term is simply

$$\langle\hat{\varepsilon}_\delta^{(1)}(\mathbf{x}_1, \mathbf{x}_2)\rangle_{\mathrm{H}} = \int dm \left(\frac{m}{\rho}\right)^2 f_{\mathrm{m}}(m, z)\int d^3\mathbf{x}\, u(\mathbf{x}_1 - \mathbf{x}; m)\, u(\mathbf{x}_2 - \mathbf{x}; m)\,, \quad (1.169)$$

while the 2-halo term is more complicated, and interesting, to derive

$$\langle\hat{\varepsilon}_\delta^{(2)}(\mathbf{x}_1, \mathbf{x}_2)\rangle_{\mathrm{H}} = \int dm_1 \left(\frac{m_1}{\rho}\right)\int dm_2 \left(\frac{m_2}{\rho_{\mathrm{m}}}\right) \times \int\int u(\mathbf{x}_1 - \mathbf{x}'; m_1)\, u(\mathbf{x}_2 - \mathbf{x}''; m_2) \times \left\langle \sum_i \delta(\mathbf{x}' - \mathbf{x}_i)\delta(m_1 - m_i)\sum_j \delta(\mathbf{x}'' - \mathbf{x}_j)\delta(m_2 - m_j)\right\rangle_{\mathrm{H}} d^3\mathbf{x}'\, d^3\mathbf{x}'' - 1. \quad (1.170)$$

Given the relation (1.159), we can rephrase the ensemble average appearing in Eq. (1.170) in terms of the halo-halo correlation function $\varepsilon_{\delta_{\mathrm{h}}}(\mathbf{x}' - \mathbf{x}''; m_1, m_2)$. At large enough scales (linear), as it is the case where the 2-halo term is dominant, the halo over-density $\delta_{\mathrm{h}}(m)$ is related to the matter over-density via the order 1 LIMD bias $b_1(m)$, i.e. $\delta_{\mathrm{h}}(m) = b_1(m)\delta$. Therefore, we can express the halo-halo correlation function via the linear 2-point

matter one

$$\varepsilon_{\delta_h}\left(\mathbf{x}' - \mathbf{x}''; m1, m2\right) = b_1(m_1) b_1(m_2) \varepsilon^{\text{lin.}}(\mathbf{x}' - \mathbf{x}''). \tag{1.171}$$

Then, the mass integration in Eq. (1.170) can eventually be factorised.

Generalising the formalism to the n-point correlation function, it will always be possible to express it as a sum of terms accounting for all the possible distributions of the points within the halos: from the 1-halo term, where all the points lie within a single halo and which is dominant at the smallest scales, up to the n-halo term where every point lies in a different halo, dominant at the largest scales. Therefore, we will write the correlation function $\varepsilon_\delta\left(\mathbf{x}_1, \ldots, \mathbf{x}_n\right)$ as a sum of its different multi-halo components

$$\varepsilon_\delta\left(\mathbf{x}_1, \ldots, \mathbf{x}_n\right) = \sum_{i=1}^{n} \varepsilon_\delta^{(i)}\left(\mathbf{x}_1, \ldots, \mathbf{x}_n\right). \tag{1.172}$$

However, the number of convolutions appearing within the required n-halo terms increases and the complexity of their numerical computation can easily diverge while moving to higher order correlation function. Since for our analyses we will need the correlations up to the six order, it will be more convenient to move to the Fourier space where the matter polyspectra in the halo model framework take a much simpler expression. We will postpone this discussion to Chapter 3 where we will report the expressions for the polyspectra required for the computation.

2

Weak Lensing

Contents

In the previous chapter we reviewed the basic formalism underlying the description of statistical random fields in Cosmology and we saw how the cosmological model we employ to describe our Universe enters in the game, affecting the different quantities introduced. Now, we want to understand how we can improve our knowledge

about the underlying model. In broad strokes, once a set of dedicated observables has been identified, we measure them from the data, we compare these values against the expected ones for a given theoretical model, and eventually we confirm it or rule it out. While it may sounds very easy, in reality all these steps are far from being immediate.

First, we need to theoretically define a set of observables $\mathcal{O}[\delta]$, as functional of the matter density contrast field δ, sensitive to changes in the underlying model (via its parameters). This is not an easy task. Let us focus on the case of observables employed in late-time Universe observations, as it will be the case for the study carried out in this manuscript. If we could have access to the real matter distribution δ at late time, we would observe a non-linear evolution of its primordial Gaussian distribution. Of course, a desirable property for an observable is to be unmistakably predictable given the values of the parameters under consideration. However, since we do not have an exact model for the matter clustering at all the scales of interest, a level of theoretical uncertainty is introduced in the prediction of the observables. Further, we do not have direct access to the matter field either: we always learn about its distribution via the analysis of *tracers* of it. As a matter of fact, we can study the Universe via the collection of electromagnetic signals emitted by the small fraction of baryonic matter. Therefore, we need to understand the complex relations between those tracers and the underlying matter field. Usually, approximated models are introduced along with *nuisance* parameters in which we are not directly interested in, but which can still spoil our cosmological forecast. In Sec. 2.1, we will build the observables used in our study, i.e. the *binned tomographic weak lensing convergence power spectrum* and *bispectrum* [Bartelmann & Schneider, 2001; Kilbinger, 2015; Bartelmann & Maturi, 2017]. In particular, since the first detection in Bacon et al. [2000]; Kaiser et al. [2000]; Wittman et al. [2000]; Schneider et al. [2002], weak lensing signal has been widely exploited over the last decade to deliver competitive cosmological parameter constraints: the Deep Lens Survey (DLS, Jee et al. [2013, 2016]), the Canada-France-Hawaii Telescope Lensing Survey (CFHTLens, Joudaki et al. [2017]), the Science Verification data from DES (DES-SV, Jarvis et al. [2016]) and the 450 deg^2 release of the Kilo-Degree Survey (KidS-450, Hildebrandt et al. [2017]) are the main examples. Further, weak lensing detection is part of the main scientific program of modern and up-coming surveys such as DES [Flaugher, 2005], the Hyper SuprimeCam Survey (HSC, Aihara et al. [2018]), the Large Synoptic Survey Telescope (LSST, Ivezić et al. [2019]) and Euclid [Laureijs et al., 2011; Amendola et al., 2013]. This trend clearly underlines the importance of studies dedicated to the optimisation of weak lensing detection as main source of cosmological information. The work presented in the following chapter perfectly sits in this line of research.

Secondly, from the data side, we need to identify the quantities which correspond to the observables designed at the previous step. For a given observable $\mathcal{O}[\delta]$ we will measure estimators $\hat{\mathcal{O}}$ of it. The statistical properties of the estimators may however differ form those of the target observables and our goal is to build *unbiased estimators*. Survey specificities like the finite number of tracers (shot noise) or the finite volume of our survey (finite resolution on the measured modes) may affect the statistical properties of the observed quantities and if neglected . *Unbiased estimators* will be function of the data and capable to account for all these effects and correct for them.

Third, given that we usually measure an entire set of observables $\{\mathcal{O}_i\}_{i=1\ldots n}$, along with the design of unbiased

estimators we need to know the errors and eventually the correlations among them. In jargon the set is called *data vector*. In particular, we can correctly infer parameters from a collection of data, or forecast the precision of future survey inference, only if we properly weight our observations (interests in errors) and if we do not double count the information they carry (interest in correlations). All these information are well summarised by the $n \times n$ *covariance matrix* of the data vector, which is formally defined as

$$\mathrm{Cov}\left[\mathcal{O}_i\left[\delta\right], \mathcal{O}_j\left[\delta\right]\right] = \langle \mathcal{O}_i\left[\delta\right] \mathcal{O}_j\left[\delta\right] \rangle_\delta - \langle \mathcal{O}_i\left[\delta\right] \rangle_\delta \langle \mathcal{O}_j\left[\delta\right] \rangle_\delta. \tag{2.1}$$

If we were capable to perform the ensemble average in the above definition, then we would know exactly the covariance matrix for the set. However, this is not the case for late-time Universe observations as we are not capable to analytically map the Gaussian probability density functional for the initial matter field into its final observed one. Then, most of the time, we build approximations for the covariance which are functions of the data themselves, if available, or of the simulated catalogues as expected to be observed by a specific survey. In this sense, several techniques are available in the literature, like *subsampling*, *Jackknife* [Efron, 1982; Bothun et al., 1983] or *bootstrap* [Barrow et al., 1984; Loh, 2008]. However in Sec.2.2 we will argue that these estimators may not be suitable for modern galaxy surveys[Hoffmann et al., 2015; Friedrich et al., 2016; Shirasaki et al., 2017; Lacasa & Kunz, 2017]. Therefore, we will have to reconsider the exact analytical definition given in Eq. (2.1), where the ensemble average will derive from a suitable model for the matter clustering. This approach is also known as *forward modelling* of the covariance. In our case, the model which will provide the tool for the prediction of the covariance matrix will be the halo model, described in Sec. 1.3.4 and in Sec. 1.3.5. Within the present chapter, we will split our discussion about the covariance matrix for the aforementioned weak lensing observables between Sec. 2.2 and Sec. 2.3.In the former we will describe in details the covariance induced by correlations of modes in the matter field which are fully observed by the survey. In the latter instead, we will focus on a less known component which comes from correlations between observed modes and super-survey modes. This last component is usually known as *super-sample covariance* [Hu & Kravtsov, 2003; Rimes & Hamilton, 2006; Takada & Bridle, 2007; Sato et al., 2009; Takada & Jain, 2009; de Putter et al., 2012; Kayo et al., 2013].

2.1 Weak lensing observables, formal definition

In this section we will theoretically define the weak lensing observables employed in our analysis. We will assume a spatially-flat FLRW Universe and we will put ourselves in units defined by $c = 1$ (c being the speed of light). The first assumption is in particular closely verified by the most recent cosmological observations [Hinshaw et al., 2013; Bennett et al., 2013; Planck Collaboration et al., 2016].

Perturbed FLRW metric While moving to small scales, the FLRW metric defined in Eq. (1.7) along with the general shapes of the energy-momentum tensor in Eq. (1.6) are not precise anymore due to the break down of the homogeneity assumption. Specifically, we will assume that the presence of LSS induces small perturbations to the background FLRW metric $\bar{g}$ and to the tensors $\bar{T}$ as

$$g_{\mu\nu}(x) = \bar{g}_{\mu\nu}(x) + h_{\mu\nu}(x), \quad |h_{\mu\nu}(x)| \ll 1, \tag{2.2}$$

$$T_{\mu\nu}(x) = \bar{T}_{\mu\nu}(x) + \Theta_{\mu\nu}(x), \quad |\Theta_{\mu\nu}(x)|/\rho(t) \ll 1. \tag{2.3}$$

The most general form for the perturbed FLRW metric is defined by the following line element (implying the sum over repeated indices) [Bardeen, 1980; Bertschinger, 2000]

$$d^2s = a^2(\eta)\left[(1 + 2\Phi(x))\, d^2\eta - 2w_i(x)\, d\eta\, dx_i - \left((1 - 2\Psi(x))\,\delta_{ij} + 2s_{ij}(x)\right) dx_i dx_j\right] \tag{2.4}$$

where $\Phi = -h_{00}/2$, $w_i = h_{0i}$, $\Psi = -\mathrm{Tr}[h]/6$, $s_{ij} = \left(h_{ij} - \delta^{kl}h_{kl}\delta_{ij}/3\right)/2$ and δ_{ij} is the Kronecker delta defined as $\delta_{ij} = 1$ if $i = j$, $\delta_{ij} = 0$ otherwise. In particular, these last definitions allow to decompose the metric into irreducible components under rotation [Lifshitz, 1946]. We also introduced the *conformal time*

$$d\eta = \frac{dt}{a(t)}, \tag{2.5}$$

which allows to write the metric as proportional to a static one, up to a conformal transformation. This feature will be extremely valuable later when deriving the weak lensing observables. The invariance of the metric under diffeomorfisms, also known as *Gauge freedom*, allows us to remove 4 d.o.f. by binding the tensors in Eq. (2.4) to obey specific constraints. We will work in the *Transverse Gauge*, which is defined by the following relations

$$\partial_i s_{ij}(x) = 0, \tag{2.6}$$

$$\partial_i w_i(x) = 0. \tag{2.7}$$

The Einstein Equations can be written in terms of the perturbed quantities defined in this section. In the chosen Gauge, we have [Carroll, 2004]

$$2\nabla^2\Psi(x) = 8\pi G T_{00}(x), \tag{2.8}$$

$$-\frac{1}{2}\nabla^2 w_i(x) + 2\partial_0\partial_i\Psi(x) = 8\pi G T_{0i}(x), \tag{2.9}$$

$$\left(\delta_{ij}\nabla^2 - \partial_i\partial_j\right)(\Phi(x) - \Psi(x)) - \partial_0\frac{\left(\partial_i\omega_j(x) + \partial_j\omega_i(x)\right)}{2} + 2\delta_{ij}\partial_0^2\Psi(x) - \Box s_{ij}(x) = 8\pi G T_{ij}(x), \tag{2.10}$$

where the component 00, 0i and ij of the Einstein Equations are respectively given.

Newtonian limit We will work out the weak lensing formalism in the *Newtonian limit* which defines a situation where the sources of the metric perturbations are static. At the notation level, we will express the metric perturbations as functions of the three-dimensional vector position $\mathbf{x}$. This correspond to a pressure-less fluid and the matter-dominated LSS perfectly fit in this assumption (we can neglect baryonic effects). In the rest frame for such a fluid, the perturbed Einstein equations turn out to be very simple

$$\nabla^2\Psi(x) = 4\pi G\rho(\mathbf{x}), \tag{2.11}$$

$$\nabla^2 w_i(\mathbf{x}) = 0, \tag{2.12}$$

$$\left(\delta_{ij}\nabla^2 - \partial_i\partial_j\right)(\Phi(\mathbf{x}) - \Psi(\mathbf{x})) - \nabla^2 s_{ij}(\mathbf{x})) = 0. \tag{2.13}$$

The trace of Eq. (2.13) implies that

$$2\nabla^2(\Phi(\mathbf{x}) - \Psi(\mathbf{x})) = 0 \quad \rightarrow \quad \Phi(\mathbf{x}) = \Psi(\mathbf{x}) \tag{2.14}$$

given that the tensor s_{ij} is trace-less, by definition. In general, whenever relativist species are present, the two scalar potential Φ and Ψ will differ. However we assume that we are entirely dominated by non-relativistic dark matter. Therefore, the final solution will be given by

$$\nabla^2\Phi(\mathbf{x}) = 4\pi G\rho(\mathbf{x}), \quad \text{(Poisson equation)} \tag{2.15}$$

$$\mathbf{w}(\mathbf{x}) = 0, \tag{2.16}$$

$$s_{ij}(\mathbf{x}) = 0, \tag{2.17}$$

leading to the following expression for the perturbed metric

$$d^2s = a^2(\eta)\left[(1 + 2\Phi(\mathbf{x}))\, d^2\eta - (1 - 2\Phi(\mathbf{x}))\, dx_i dx_i\right]. \tag{2.18}$$

Fermat principle Tackling the weak lensing problem as an optical one, we can derive the equations for the deviation of light rays via the Fermat principle [Schneider, 1985; Blandford & Narayan, 1992; Bartelmann & Schneider, 2001]. In particular, it states that in a stationary metric a light ray moves along a path for which the variation of the travel time vanishes. The metric given in Eq. (2.18), thanks to the change of variable given in Eq. (2.5), is indeed stationary up to a conformal factor. However, the latter will not affect the formal solution since the null-geodesic condition $ds^2 = 0$ will be always verified. We highlight that it is possible to derive the very same equation of motions for a photon propagating in a perturbed metric via the geodesic equations, assuming a perturbative expansion of the background geodesic. We refer to the literature for this alternative approach [Pyne & Birkinshaw, 1996].

The null-geodesic condition applied to a metric as the one in Eq. (2.18) leads to

$$\left(1+2\Phi\left(\mathbf{x}\right)\right)d^{2}\eta\,\Big|_{\tilde{\gamma}}=\left(1-2\Phi\left(\mathbf{x}\right)\right)d\ell^{2}\left(\mathbf{x}\right)\,\Big|_{\tilde{\gamma}} \tag{2.19}$$

where we can see that the conformal pre-factor $a(\eta)$ disappears and we plugged in the three-dimensional comoving line element $d\ell^{2}$. In addition, we are formally evaluating Eq. (2.19) on the desired solution. We can map our problem into an optical problem by defining an effective refractive index as the ratio between the speed of the light over the path and c ($c=1$ in our units)

$$n\left(x\right)|_{\tilde{\gamma}}\equiv\frac{1}{v_{c}\left(\mathbf{x}\right)}\,\Big|_{\tilde{\gamma}}=\left(\frac{d\ell\left(\mathbf{x}\right)}{d\eta}\right)^{-1}\,\Big|_{\tilde{\gamma}}\approx1-2\Phi\left(\mathbf{x}\right)|_{\tilde{\gamma}} \tag{2.20}$$

where in the last relation we applied the Newtonian limit $\Phi\left(\mathbf{x}\right)\ll1$. In terms of the refractive index, the travel time for a photon is given by

$$\eta=\int_{\eta_{o}\tilde{\gamma}}^{\eta}n\left(x\right)d\ell=\int_{\eta_{o}\tilde{\gamma}}^{\eta}n\left(x\right)|\mathbf{v}| \tag{2.21}$$

and according to the Fermat principle described above the solution $\tilde{\gamma}\left(\eta\right)$ will be given by $\delta\eta|_{\tilde{\gamma}}=0$.

Single ray deflection In order to find the path that minimises Eq. (2.21), i.e.

$$\delta\eta=\delta\left[\int_{\eta_{o}\tilde{\gamma}}^{\eta}n\left(x\right)|\mathbf{v}|\right]=0, \tag{2.22}$$

we parametrise the general geodesic of the photon via an affine parameter λ

$$\delta\int_{\lambda_{o}}^{\lambda_{\eta}}n\left(\mathbf{x}\left(\lambda'\right)\right)|\mathbf{v}\left(\lambda'\right)|d\lambda'=0 \tag{2.23}$$

where $\mathbf{v}=\mathrm{d}\mathbf{x}/\mathrm{d}\lambda$. The integrand in Eq. (2.23) takes the role of the Lagrangian and we can solve Eq 2.23 via the Euler equations

$$\left[\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\partial}{\partial\mathbf{v}}-\frac{\partial}{\partial\mathbf{x}}\right]\mathcal{L}\left(\mathbf{x},\mathbf{v}\right)=0,\qquad\mathcal{L}\left(\mathbf{x},\dot{\mathbf{x}}\right)=n\left(\mathbf{x}\right)|\mathbf{v}|. \tag{2.24}$$

In details,

$$\frac{\partial}{\partial\mathbf{x}}\mathcal{L}\left(\mathbf{x},\mathbf{v}\right)=\nabla n\left(\mathbf{x}\right),\qquad\frac{\partial}{\partial\mathbf{v}}\mathcal{L}\left(\mathbf{x},\mathbf{v}\right)=n\left(\mathbf{x}\right)\frac{\mathbf{v}}{|\mathbf{v}|}=n\left(\mathbf{x}\right)\hat{\mathbf{v}}, \tag{2.25}$$

then the Euler equations read

$$n\left(\mathbf{x}\right)\dot{\hat{\mathbf{v}}}=\nabla n\left(\mathbf{x}\right)-\dot{\mathbf{v}}\left[\nabla n\left(\mathbf{x}\right)\cdot\mathbf{v}\right]. \tag{2.26}$$

We can see that the second term on the r.h.s. of Eq. (2.26) is the projection of the gradient on the direction $\hat{\mathbf{v}}$, tangential to the path. Therefore, the entire r.h.s. will be the projection of the gradient onto the direction orthogonal

to the path. We will write

$$\dot{\hat{\mathbf{v}}} = \frac{1}{n(\mathbf{x})} \nabla_{\perp} n(\mathbf{x}) \,. \tag{2.27}$$

Eq. (2.27) tells us how the direction of the photon, given by the versor $\hat{\mathbf{v}}$ tangential to the path, changes on the fly. Given an initial direction $\hat{\mathbf{v}}_o = \hat{\mathbf{v}}(\lambda_o)$ and a final one $\hat{\mathbf{v}} = \hat{\mathbf{v}}(\lambda_\eta)$, we integrate Eq. (2.27) to find the total deflection of the path while the photon travelled for a time $\eta - \eta_o$

$$\boldsymbol{\Delta}(\eta, \eta_o) = \int_{\lambda_o, \tilde{\gamma}}^{\lambda_\eta} \frac{\nabla_{\perp} n(\mathbf{x}(\lambda'))}{n(\mathbf{x}(\lambda'))} d\lambda' \approx -2 \int_{\lambda_o, \tilde{\gamma}}^{\lambda_\eta} \nabla_{\perp} \Phi(\mathbf{x}(\lambda')) \, d\lambda' \tag{2.28}$$

where we used the relation in Eq. (2.20) and approximated at first order in Φ. Eq. (2.28) defines the total "deflection" $\boldsymbol{\Delta}(\eta, \eta_o)$ induced by a Newtonian perturbation Φ to the metric for a single photon travelling from time η_o to time η. Please note that the integrals entering the definition are meant to be evaluated on the actual path $\tilde{\gamma}$ followed by the photon and that the quantity $\boldsymbol{\Delta}$ is not an angle in general but a coordinate difference.

2.1.2 Deflection of light bundles

In the previous section we derived the equations which describe the deflection of a photon within a weakly perturbed metric. However, we do not measure single photons in our surveys. What we observe are distorted shapes of far extended sources, which are continuous objects and which are better described by the distortion of bundles of neighbour geodesics. In this section we will promote the formalism explored in the previous part of this chapter to this more realistic scenario.

Jacobian matrix In a bundle of light rays we consider a reference geodesic γ_{b}. We can study the evolution of the comoving separation $\mathcal{E}(\lambda)$ of a general geodesic γ_{a} in the bundle away from the reference one with the help of Eq. (2.27),

$$\frac{\mathrm{d}^2 \mathcal{E}(\lambda)}{\mathrm{d}\lambda^2} = \frac{\mathrm{d}^2 \mathbf{x}_{\mathrm{a}}(\lambda)}{\mathrm{d}\lambda^2} - \frac{\mathrm{d}^2 \mathbf{x}_{\mathrm{b}}(\lambda)}{\mathrm{d}\lambda^2} = \hat{\mathbf{v}}_{\mathrm{a}}(\lambda) - \hat{\mathbf{v}}_{\mathrm{b}}(\lambda) = -2 \left[\nabla_{\perp} \Phi(\mathbf{x}_{\mathrm{a}}(\lambda)) - \nabla_{\perp} \Phi(\mathbf{x}_{\mathrm{b}}(\lambda)) \right]. \tag{2.29}$$

where $\mathbf{x}_a(\lambda)$ and $\mathbf{x}_b(\lambda)$ are points on the support of γ_a and γ_b respectively at a given value of the affine parameter λ. The affine parameter for the second geodesic γ_a is the same as for the reference geodesic and in particular we assume the 2 photons to have the same initial conditions

$$\mathbf{x}_{\mathrm{a}}(\lambda_o) = \mathbf{x}_{\mathrm{b}}(\lambda_o), \quad \mathbf{v}_{\mathrm{a}}(\lambda_o) = \mathbf{v}_{\mathrm{b}}(\lambda_o) \tag{2.30}$$

For later convenience, we recall that the *comoving angular diameter distance* (1.27) allows to relate the comoving transverse distance $\mathbf{r}_{\perp}$ with the corresponding two-dimensional angular size Θ

$$\Theta = \frac{\mathbf{r}_{\perp}}{D_A(z(\omega))}. \tag{2.31}$$

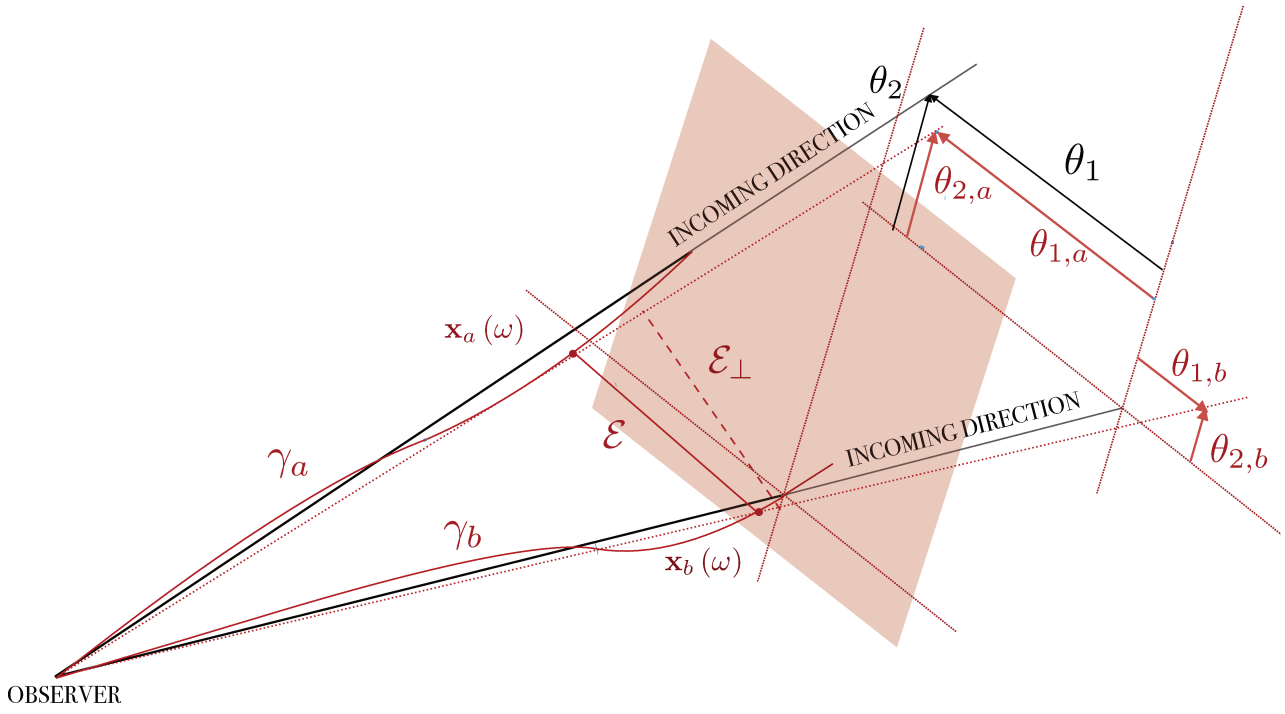

Figure 2.1.1: Graphic depicting the geometrical setup for the mathematical description of the light bundles. The total comoving separation $\mathcal{E}(\omega)$ (we are implying the dependence in the figure) is defined as the comoving distance between the points $\mathbf{x}_a$ and $\mathbf{x}_b$ on the support of the geodesic γ_a and γ_b at the same comoving distance ω from the observer. The transverse comoving separation $\mathcal{E}_\perp(\omega)$ (red dashed line) is the projection of the total separation $\mathcal{E}(\omega)$ onto the red plane defined as orthogonal to the incoming direction of the photon γ_b at comoving distance ω. Angular separations as seen by the observer are measured with respect to the incoming direction of γ_b. Respectively, $\Theta_a = \{\theta_{1,a}, \theta_{2,a}\}$ and $\Theta_b = \{\theta_{1,b}, \theta_{2,b}\}$ are the angular separations between the the reference axe just defined and the extrema of $\mathcal{E}_\perp(\omega)$ (as seen by the observer).

In the following, we will simplify our notation by confusing the functional dependence on $z(\omega)$ with ω for functions previously defined (in Chapter 1) as dependent on redshift. For example, we will simply write $f(\omega)$ implying $f(z(\omega))$ for a general function f. Also, being in a spatially-flat Universe, $D_A(z(\omega)) = \omega$. We will study the deviation of light bundles with the help of the geometrical setup shown in Fig. 2.1.2 and described here below:

- we parametrize geodesics in the bundle via the comoving distance ω of the points on their support from the observer,

- we will solve Eq. (2.29) in terms of the *transverse comoving separation* $\mathcal{E}_\perp(\omega)$ between the two geodesic γ_a and γ_b for the same value of the comoving distance. The separation $\mathcal{E}_\perp(\omega)$ is built as the projection of the three-dimensional separation $\mathcal{E}(\omega)$ (between the points $\mathbf{x}_a$ and $\mathbf{x}_b$ on the support of the geodesic γ_a and γ_b at the same comoving distance ω from the observer) onto the plane orthogonal to the incoming direction of the

reference geodesic γ_b, at angular comoving distance ω from the observer[1]. Then we will consider just two of the three components of the gradient in Eq. (2.29),

- angular deviations $\Theta = \{\theta_1, \theta_2\}$, associated to a specific transverse displacement via Eq. (1.27), are given as relative to the incoming angular direction of the geodesic γ_b, taken as reference.

Given these points, the differential system to solve in terms of $\mathcal{E}_\perp(\omega)$ for an incoming angular direction $(\theta_1, \theta_2) = \Theta$ of γ_a is

$$\frac{\mathrm{d}^2\mathcal{E}_\perp(\omega)}{\mathrm{d}\omega^2} = -2\left[\nabla_\perp\Phi(\omega\Theta_a, \omega) - \nabla_\perp\Phi(\omega\Theta_b, \omega)\right], \qquad \mathcal{E}_\perp(0) = 0, \quad \frac{\mathrm{d}\mathcal{E}_\perp(0)}{\mathrm{d}\omega} = \frac{\mathrm{d}D_A(\omega)}{\mathrm{d}\omega}\Theta = \Theta \tag{2.32}$$

where the initial conditions do actually implement 1) the fact that the 2 incoming photons will reach the observer at the same position (the origin of the reference frame) and the ω-derivative of the transverse displacement at the origin gives the direction of the incoming photon. Respectively, $\Theta_a = \{\theta_{1,a}, \theta_{2,a}\}$ and $\Theta_b = \{\theta_{1,b}, \theta_{2,b}\}$ are the angular separations between the the reference axe defined above and the the projection of the points $\mathbf{x}_a(\omega)$ and $\mathbf{x}_b(\omega)$ on the transverse plane. We highlight that Eq. (2.32) can also be derived within a general relativistic framework as the Sachs-Jacobi equation for a perturbed FLRW metric [e.g. Bartelmann & Schneider, 2001, and references therein]. In such scenario, the source of deviation for the light bundle is the sum of two term: one related to the curvature of the Universe $K^{(3)}$ and another one related to the local matter perturbation. For a spatially-flat Universe the former is vanishing and the system (2.32) is indeed the Sachs-Jacobi equation itself. While the solution of the associated homogeneous system can be easily derived

$$\mathcal{E}_\perp^{\mathrm{ho.}}(\omega) = \omega\,\Theta, \tag{2.33}$$

we will get a particular one via the convolution with the well known Green function $G(\omega, \omega')$ for the operator $\mathrm{d}^2/\mathrm{d}\omega^2$. Under the assumption of a spatially-flat Universe we can compute the difference between 2 comoving distances via Eq. (1.29),

$$G(\omega, \omega') = D_A(\omega, \omega') = \chi(\omega, \omega'). \tag{2.34}$$

in terms of the corresponding redshifts. Therefore, the formal solution to the system (2.32) is given by

$$\mathcal{E}_\perp(\Theta, \omega) = \omega\,\Theta - 2\int_0^\omega d\omega'\chi(\omega, \omega')\left[\nabla_\perp\Phi(\omega\Theta_a, \omega') - \nabla_\perp\Phi(\omega'\Theta_b, \omega')\right] \tag{2.35}$$

and we made explicit the dependence on the angular direction Θ of the incoming photon γ_a. In analogy with standard lens theory we define the *Jacobian matrix* as the derivative of the unperturbed, original angular separation,

[1] A more rigorous definition of the transverse quantities can be derived [Bernardeau et al., 2010]. However, at leading order in the metric perturbations, the differences are negligible.

i.e.

$$\Theta_{\rm np}\left(\Theta,\omega\right)=\frac{\mathcal{E}\left(\Theta,\omega\right)}{\omega},\tag{2.36}$$

with respect to the perturbed one between the incoming photons, i.e.

$$\Theta_{\rm p}\left(\omega\right)=\frac{\omega\Theta}{\omega}=\Theta,\tag{2.37}$$

which is

$$A\left(\Theta,\omega\right)\equiv\frac{\partial\Theta_{\rm np}\left(\Theta,\omega\right)}{\partial\Theta_{\rm p}\left(\omega\right)}=\frac{1}{\omega}\frac{\partial\mathcal{E}^{\theta}\left(\Theta,\omega\right)}{\partial\Theta}.\tag{2.38}$$

In the future, we will omit the dependence on ω in $\Theta_{\rm p}$ being constant. It is worth underlying that the angle associated to the homogeneous solution Θ is actually the direction of the perturbed incoming geodesic γ_a, as specified by the initial conditions in the second line of Eq. (2.32). Instead, the full solution $\mathcal{E}\left(\omega\right)$, which is function of the comoving distance from the observer, is the original reconstructed comoving separation between the 2 geodesics before their fly through the potential Φ, from ω to $\omega=0$ (the observer). In particular, the Jacobian defined in Eq. (2.38) defines the linearization of the *inverse lensing map* $\Theta_{\rm p}\rightarrow\Theta_{\rm np}\left(\omega\right)$, i.e.

$$\Theta_{\rm np}\left(\Theta_{\rm p},\omega\right)|_{\rm lin.}=\left[1_2+A\left(\Theta_{\rm p},\omega\right)\right]\Theta_{\rm p}.\tag{2.39}$$

where 1_2 is the identity matrix of dimension 2. The more natural mapping from an original geodesic bundle to the lensed one is instead realised by the inverse of the Jacobian (at linear level), the *magnification tensor* $M=A^{-1}$

$$\Theta_{\rm p}\left(\Theta_{\rm np}\left(\omega\right)\right)|_{\rm lin.}=\left[1_2+M\left(\Theta_{\rm np},\omega\right)\right]\Theta_{\rm np}\left(\omega\right).\tag{2.40}$$

In the weak lensing regime we will see around Eq. (2.64) that the matrix A is always invertible.

Born approximation The perpendicular gradients of the Newtonian potential in Eq. (2.35) are to be evaluated along the true paths of the photons. The *Born approximation* will help us simplifying the actual computation of the transverse separation $\mathcal{E}_\perp$. In particular it states that the change of the actual transverse comoving separation $\mathcal{E}_\perp$ between the two real geodesics is small compared to the comoving separation between the unperturbed path. Formally

$$\frac{\left|\mathcal{E}_\perp\left(\omega\right)-\omega\,\Theta\right|}{\left|\omega\,\Theta\right|}\ll 1.\tag{2.41}$$

From the computational perspective, the Born approximation entitles us to compute the gradients within Eq. (2.35) over straight paths (the geodesics $\gamma_a^{\rm u}$ and $\gamma_b^{\rm u}$ in Fig. 2.1.2) separated by the angle Θ, then

$$\nabla_\perp\Phi\left(\omega\Theta_a,\omega\right)\approx\nabla_\perp\Phi\left(\omega\Theta,\omega\right),\tag{2.42}$$

$$\nabla_\perp\Phi\left(\omega\Theta_b,\omega\right)\approx\nabla_\perp\Phi\left(0,0,\omega\right).\tag{2.43}$$

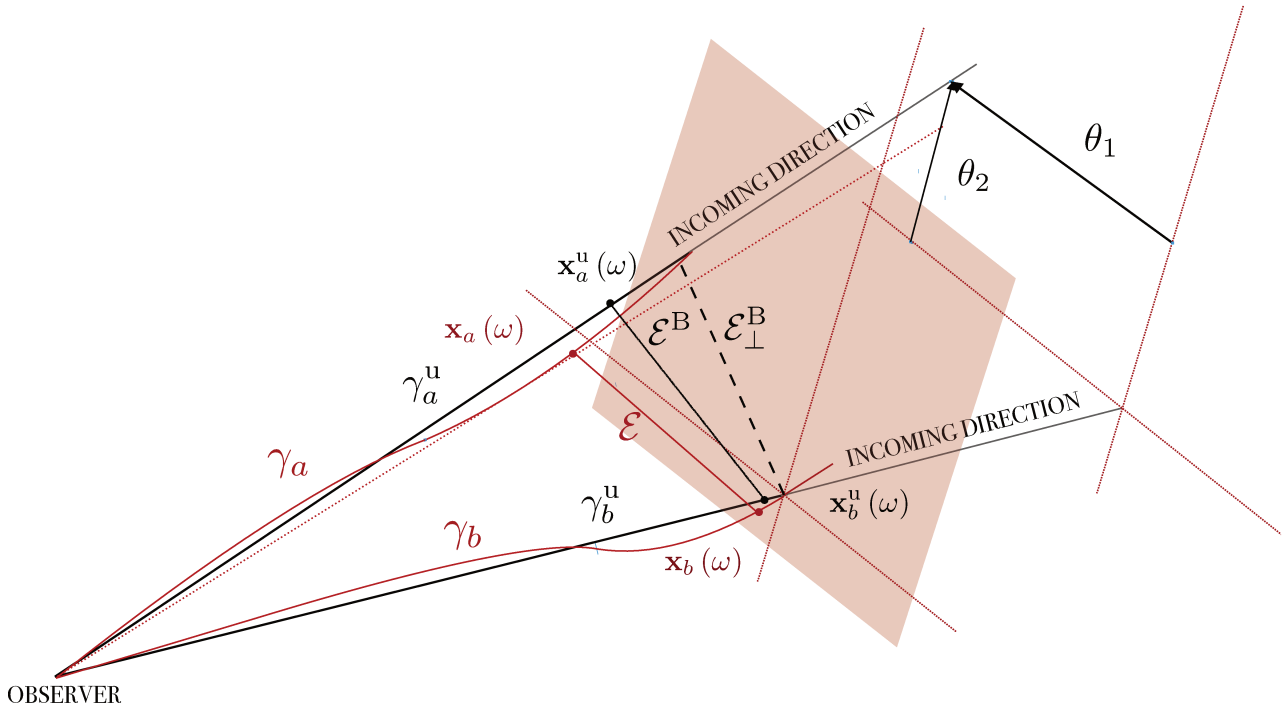

Figure 2.1.2: Graphic depicting the geometrical setup for the mathematical description of the light bundles, under the Born approximation. Different from the analysis in Fig. 2.1.1, the transverse comoving separation $\mathcal{E}_\perp^{\rm B}(\omega)$ (we are implying the dependence in the figure) is now built via the projection of the total separation $\mathcal{E}^{\rm B}(\omega)$ between the unperturbed geodesics $\gamma_a^{\rm u}$ and $\gamma_b^{\rm u}$ at points $\mathbf{x}_a^{\rm u}(\omega)$ and $\mathbf{x}_b^{\rm u}(\omega)$. The red plane is the transverse plane defined in Fig. 2.1.1. The angular separations as seen by the observer are still measured with respect to the incoming direction of γ_b, which now defines the direction of the straight unperturbed geodesic $\gamma_b^{\rm u}$. Also $\Theta = \{\theta_1, \theta_2\}$, which used to be final angular displacement of the photon γ_a, now coincides with the angular separation between the unperturbed geodesics $\gamma_a^{\rm u}$ and $\gamma_b^{\rm u}$.

where in the second approximation (2.43) we made use of the incoming direction of the geodesic $\gamma_{\rm b}$ being defined as the reference axe $(0, 0)$. Fig. 2.1.2 shows the new geometrical setup to further simplify Eq. (2.35). Under this approximation, the comoving separation we will be actually solving for is $\mathcal{E}^{\rm B}(\omega)$, defined as connecting two points $\mathbf{x}_{\rm a}^{\rm u}(\omega)$ and $\mathbf{x}_{\rm b}^{\rm u}(\omega)$ respectively on the support of the unperturbd geodesics $\gamma_{\rm a}^{\rm u}$ and $\gamma_{\rm b}^{\rm u}$ at comoving distance ω from the observer. In terms of Eq. (2.35), and focusing on the transverse component, we will write

$$\mathcal{E}_\perp^{\rm B}(\Theta, \omega) = \omega\,\Theta - 2\int_0^\omega d\omega' \chi(\omega, \omega')\,\nabla_\perp\left[\Phi(\omega'\Theta, \omega') - \Phi(0, 0, \omega')\right] \tag{2.44}$$

while the Jacobian matrix in Eq. (2.38) will simply be

$$\left[A(\Theta, \omega)\right]_{ij} = \frac{1}{\omega}\frac{\partial \mathcal{E}_{\perp,i}^{\rm B}(\Theta, \omega)}{\partial \Theta_j} = \delta_{ij} - 2\int_o^\omega d\omega'\,\mathcal{W}(\omega, \omega')\,\partial_i\partial_j\Phi(\Theta, \omega')\,, \qquad \partial_j = \frac{\partial}{\partial \Theta_j}. \tag{2.45}$$

In Eq. (2.45) we introduced the well known *lensing kernel*

$$\mathcal{W}(\omega,\omega') = \begin{cases} \frac{\chi(\omega,\omega')}{\omega\omega'}, & \text{if } \omega > \omega' \\ 0, & \text{if } \omega \leqslant \omega' \end{cases} \tag{2.46}$$

which will be of great help for introducing survey specificities in the above definitions.

Deflection angle Starting from Eq. (2.44), we can define the *(absolute) deflection angle* as the net displacement of a geodesic bundle away from the initial, reconstructed, angular direction $\mathcal{E}^{\mathrm{B}}_{\perp}(\omega)/\omega$

$$a_i(\Theta,\omega) \equiv \Theta_{\mathrm{p},i} - \Theta_{\mathrm{np},i}(\Theta,\omega) = \frac{\omega\,\Theta_i - \mathcal{E}^{\mathrm{B}}_{\perp,i}(\omega)}{\omega}. \tag{2.47}$$

Absolute deflection angles cannot be measured. All measurable effects of light deflection only depend on the derivatives of the deflection angle a and the choice of the fiducial geodesic $\tilde{\gamma}_b$ is then irrelevant for practical purposes. Regarding this point, it is interesting to notice that the Jacobian in Eq. (2.45) gives exactly this information, being

$$[A(\Theta,\omega)]_{ij} = \delta_{ij} - \partial_j a_i(\Theta,\omega). \tag{2.48}$$

2.1.3 Weak lensing observables

In the present section we will formally introduce the 2 main fields used to exploit weak lensing signal for cosmological purposes, namely the *weak lensing shear* and *weak lensing convergence* fields.

Lensing potential To make our future notation lighter and to get closer to the weak lensing literature [Bartelmann & Schneider, 2001; Kilbinger, 2015; Bartelmann & Maturi, 2017], let us introduce the *lensing potential* as

$$\psi_{\mathrm{L}}(\Theta,\omega) = 2\int_0^{\omega} d\omega'\,\mathcal{W}(\omega,\omega')\;\Phi(\omega'\Theta,\omega). \tag{2.49}$$

In terms of it, the Jacobian can be expressed as

$$[A(\Theta,\omega)]_{ij} = \delta_{ij} - \partial_j\partial_i\psi_{\mathrm{L}}(\Theta,\omega) \tag{2.50}$$

Convergence and shear We define for the first time the scalar field *convergence* $\kappa(\Theta,\omega)$ and the spin-2 field *shear* $\gamma(\Theta,\omega) = \{\gamma_1(\Theta,\omega), \gamma_2(\Theta,\omega)\}$ as

$$A(\Theta,\omega) = \begin{pmatrix} 1-\kappa(\Theta,\omega)-\gamma_1(\Theta,\omega) & -\gamma_2(\Theta,\omega) \\ -\gamma_2(\Theta,\omega) & 1-\kappa(\Theta,\omega)+\gamma_1(\Theta,\omega) \end{pmatrix}. \tag{2.51}$$

Inversely they are related to the lensing potential as

$$\kappa(\Theta,\omega) = \left[\partial_1^2 + \partial_2^2\right]\psi_{\rm L}(\Theta,\omega)\,, \tag{2.52}$$

$$\gamma_1(\Theta,\omega) = \frac{1}{2}\left[\partial_1\partial_1 - \partial_2\partial_2\right]\psi_{\rm L}(\Theta,\omega)\,, \tag{2.53}$$

$$\gamma_2(\Theta,\omega) = \partial_1\partial_2\psi_{\rm L}(\Theta,\omega)\,, \tag{2.54}$$

while we can directly relate them, in real space, via

$$\partial_1\kappa(\Theta,\omega) = \partial_1\gamma_1(\Theta,\omega) + \partial_2\gamma_2(\Theta,\omega)\,, \tag{2.55}$$

$$\partial_2\kappa(\Theta,\omega) = -\partial_2\gamma_1(\Theta,\omega) + \partial_1\gamma_2(\Theta,\omega)\,, \tag{2.56}$$

$$\nabla^2\kappa(\Theta,\omega) = \left(\partial_1^2 - \partial_2^2\right)\gamma_1(\Theta,\omega) + 2\partial_1\partial_2\gamma_2(\Theta,\omega)\,. \tag{2.57}$$

We postpone to Sec. 2.1.4 a comment on the relations among the introduced fields in Fourier space. We remind that the above field expressions have been derived under the assumption of observed light bundles having a small angular size. A full-sky derivation is proposed in Bernardeau et al. [2010]

$$\kappa(\theta,\varphi,\omega) = \frac{1}{2}\bar{\eth}\bar{\eth}\psi_{\rm L}(\theta,\varphi,\omega)\,, \tag{2.58}$$

$$\gamma(\theta,\varphi,\omega) = \frac{1}{2}\eth\eth\psi_{\rm L}(\theta,\varphi,\omega)\,, \tag{2.59}$$

where the operators $\bar{\eth}$, $\eth$ are the spin operators on the sphere

$$\bar{\eth}\,[{}_s{\rm X}] = -\sin^s\theta\left(\partial_\theta + i\frac{\partial_\varphi}{\sin\theta}\right)\frac{1}{\sin^s\theta}\,[{}_s{\rm X}]\,, \tag{2.60}$$

$$\eth = -\frac{1}{\sin^s\theta}\theta\left(\partial_\theta - i\frac{\partial_\varphi}{\sin\theta}\right)\sin^s\theta\,[{}_s{\rm X}]\,, \tag{2.61}$$

given in Cartesian representation with s the spin of the field ${}_s{\rm X}$. Also, the angles θ, φ have been locally introduced to describes the field on the sphere, respectively being the polar and azimuth angle. For the shear field, a complex representation is usually given

$$\gamma(\Theta,\omega) \equiv \gamma_1(\Theta,\omega) + i\gamma_2(\Theta,\omega) \equiv |\gamma(\Theta,\omega)|e^{2i\varphi_\gamma(\Theta,\omega)} \tag{2.62}$$

and we can write the Jacobian matrix in another interesting form

$$A(\Theta,\omega) = (1-\kappa(\Theta,\omega))\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} - \gamma(\Theta,\omega)\begin{pmatrix}\cos\left[2\varphi_\gamma(\Theta,\omega)\right] & \sin\left[2\varphi_\gamma(\Theta,\omega)\right]\\ \sin\left[2\varphi_\gamma(\Theta,\omega)\right] & -\cos\left[2\varphi_\gamma(\Theta,\omega)\right]\end{pmatrix}. \tag{2.63}$$

We can express the determinant of the Jacobian matrix in terms of the fields defined above

$$\det\left[A\left(\Theta,\omega\right)\right]\frac{1}{\left(1-\kappa^{2}\left(\Theta,\omega\right)\right)^{2}-\left|\gamma\left(\Theta,\omega\right)\right|^{2}}\underset{\text{W-L}}{\approx}1+2\kappa\left(\Theta,\omega\right). \tag{2.64}$$

The weak lensing regime, i.e. $\kappa,\gamma\ll 1$, ensures $\det\left[A\right]\neq 0$ which makes it always invertible. The representation of the Jacobian in Eq. (2.63) allows for a more physical interpretation of the fields introduced.

Physical intuition Let's assume to have a source (a galaxy for example) defined by an isotropic profile of radius R_o, at a given comoving distance ω. We parametrise the angular coordinates of a general point P of the profile as $P\rightarrow\left(\theta,\varphi\right)=\left(R_o\cos t,R_o\sin t\right)$, $t\in\left[0,2\pi\right)$. The point P will be mapped into a point P' of angular coordinates (θ',φ'), at the observer position, due to the distortions of light bundles. In terms of the formalism studied in the previous section, $(\theta',\varphi')=\Theta_p$ and $(\theta,\varphi)=\Theta_{\text{np}}\left(\omega\right)$. In this paragraph we desire to inquire about the physical impact of the shear and converge fields. To do that, we will work with the more natural linearised direct lensing map, which is implemented via the magnification tensor $M=A^{-1}$.

If we assume to have just a convergence component $\kappa\left(\theta,\omega\right)$, then the linear modification of the angular coordinates will be, according to Eq. (2.40),

$$\left[\Theta_{\text{p}}-\Theta_{\text{np}}\left(\omega\right)\right]_{\text{lin.}}=M\left(\Theta,\omega\right)\cdot\Theta_{\text{np}}\left(\omega\right)=\begin{pmatrix}\frac{1}{1-\kappa(\Theta,\omega)} & 0\\ 0 & \frac{1}{1-\kappa(\Theta,\omega)}\end{pmatrix}\cdot\begin{pmatrix}R_o\cos t\\ R_o\sin t\end{pmatrix}=\begin{pmatrix}\frac{R_o}{1-\kappa(\Theta,\omega)}\cos t\\ \frac{R_o}{1-\kappa(\Theta,\omega)}\sin t\end{pmatrix} \tag{2.65}$$

and we dropped the subscript np on the angular dependencies within the matrix M since differences between the angle Θ_{p} and the angle Θ_{np} would induce corrections beyond the linear level. Therefore, the original galaxy is mapped into a new circular shape with radius $R=\frac{R_o}{1-\kappa(\Theta,\omega)}$ with a dependence on the distance from the observer. We can then see that the convergence field has the main effect of inducing a magnification, i.e. an isotropic deviation, of the light bundles.

If we also consider a non-vanishing shear component $\gamma\left(\Theta,\omega\right)$, then the original shape will be distorted into an ellipse with major and minor axe respectively given by

$$a=\frac{R_o}{1-\kappa\left(\Theta,\omega\right)-\left|\gamma\left(\Theta,\omega\right)\right|}, \tag{2.66}$$

$$b=\frac{R_o}{1-\kappa\left(\Theta,\omega\right)+\left|\gamma\left(\Theta,\omega\right)\right|}. \tag{2.67}$$

The ratio between the original circular area of the light bundle and the final is

$$\mu=\frac{a\cdot b\cdot\pi}{R_o\pi^{2}}=a\cdot b\underset{\text{W-L}}{\approx}1+2\kappa\left(\Theta,\omega\right) \tag{2.68}$$

and we can see thay in the weak lensing regime, i.e. $\Phi\ll 1\rightarrow\kappa\ll 1$, the convergence field is directly related to

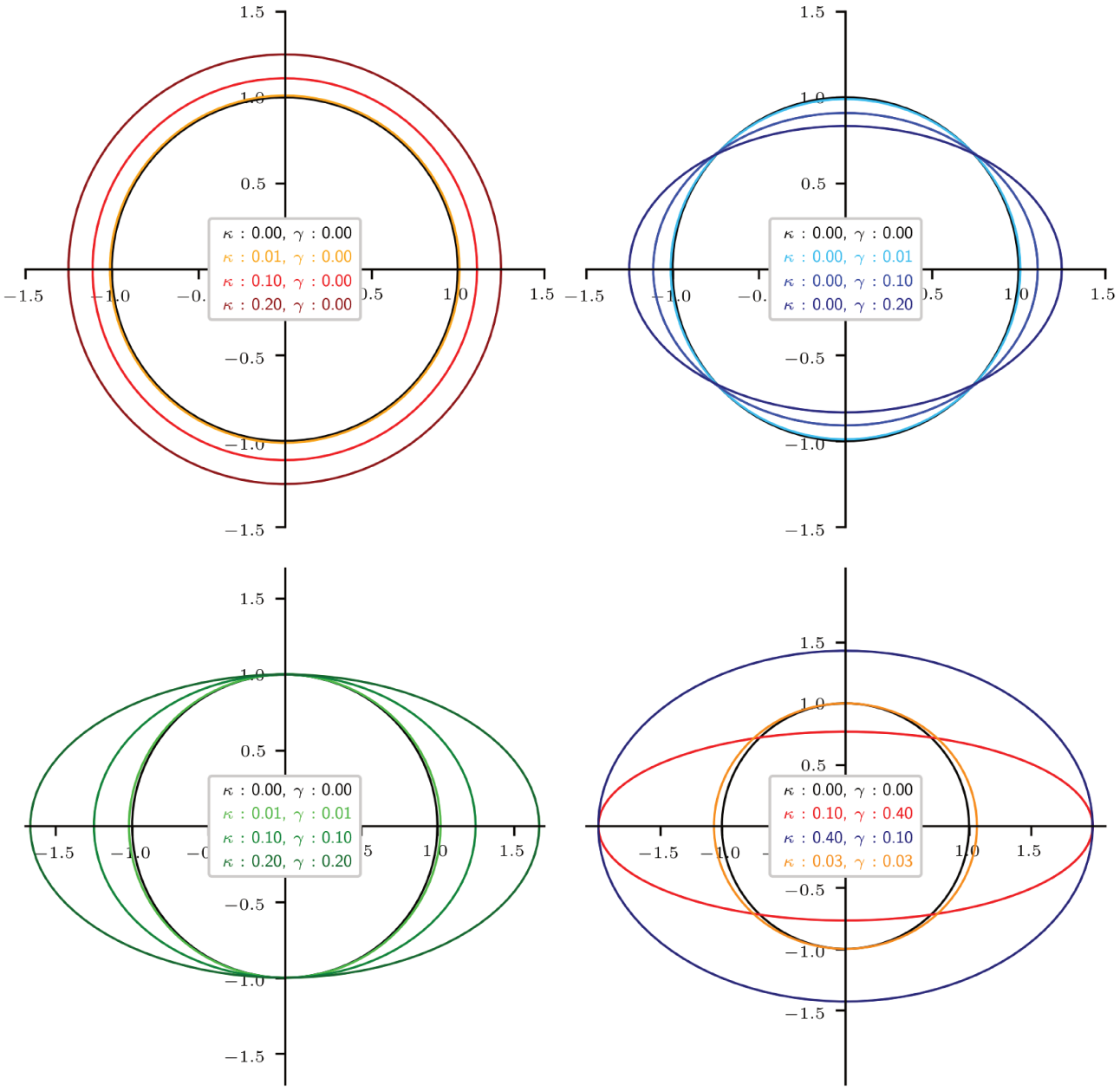

Figure 2.1.3: Starting from the same circular isotropic profile parametrized as $\Theta_{\rm np}(t) = (R_o \cos t, R_o \sin t)$, we propose different observed lensed shapes obtained via the direct linear transformation $\Theta_{\rm p}(t) = \left[1 + A^{-1}\right] \cdot \Theta_{\rm np}(t)$. We implement different values of the convergence field and of the shear field. Specifically, on the *top-left* panel we focus on the magnification effect induced by the convergence field alone. On the *top-right* panel, we look at the distortions induced by the shear component alone. On the *bottom* row we allow for more complex modifications including both the fields. In general, we keep $\varphi_\gamma = 0$ then no rotations of the final ellipses are induced.

the magnification.

We conclude this paragraph by proposing in Fig. 2.1.3 several lensed profiles of the same original isotropic source for different values of the shear and convergence field. We refer to the caption for more details about the implementation.

2.1.4 More on the convergence field

In this section we will focus on the convergence field being the main observable used in our study. We will see how the convergence can provide a direct way to inspect the clustering properties of matter and we will introduce its polyspectra along with the approximations involved.

Convergence field and matter density field Via the Poisson equation in polar coordinates, we can interestingly express the convergence field κ in Eq. (2.52) in terms of the matter density contrast

$$\kappa\left(\Theta,\omega\right)=\frac{3\left(H_{0}\right)^{2}\Omega_{\mathrm{m},0}}{2}\int_{o}^{\omega}\frac{d\omega'}{a\left(\omega'\right)}\mathcal{W}\left(\omega,\omega'\right)\delta\left(\Theta,\omega'\right). \tag{2.69}$$

This relation will be of crucial importance later in this manuscript where we will actually relate the statistical properties of the convergence field to those of the matter field. Indeed, the former are just a "projection" of the latter.

Tomographic convergence field Modern galaxy surveys will have the capacity to detect the binned positions of the sources in redshift. Of course, the comoving number density of galaxies within each redshift bin, along with the number and the width of the bins, is fixed by the specificities of a given survey. While we will apply our analysis to a Euclid-like one, here we will consider a general comoving density distribution $n\left(z\right)$.

The expression reported in Eq. (2.69) gives the actual level of convergence for a single source at comoving distance ω from the observer. If we allow, as it is the case in this section, for a distribution of sources in redshift, then the above expression will have to be convoluted with the function $n\left(z\right)$, properly normalised. Therefore, the total convergence relative to the redshift bin $b_i=[z_i,z_{i+1}]$ will be given by the following more complex expression that will define the *tomographic convergence field*

$$\kappa_{(i)}\left(\Theta\right)=\frac{3\,H_{0}^{2}\Omega_{m,0}}{2}\int_{o}^{z_{i+1}}dz\frac{\mathrm{d}\omega(z)}{\mathrm{d}z}\left(1+z\right)\int_{z}^{\infty}dz_{s}\,\frac{n_{i}(z_{s})}{\bar{n}_{i}}\,\mathcal{W}\left(z_{s},z\right)\delta\left(\Theta,z\right), \tag{2.70}$$

where we employed the redshift as integration variable. We can rephrase Eq. (2.70) in a more compact form

$$\kappa_{(i)}\left(\Theta\right)=\int_{o}^{z_{i+1}}dz\,\frac{\mathrm{d}\omega(z)}{\mathrm{d}z}\,\mathcal{Q}_{(i)}\left(z\right)\delta\left(\Theta,z\right). \tag{2.71}$$

The function $\mathcal{Q}_{(i)}$ is the convolution of a cosmology (only) dependent function $\mathcal{S}\left(z_{s},z\right)$ and a function $\mathcal{F}\left(z_{s},i\right)$

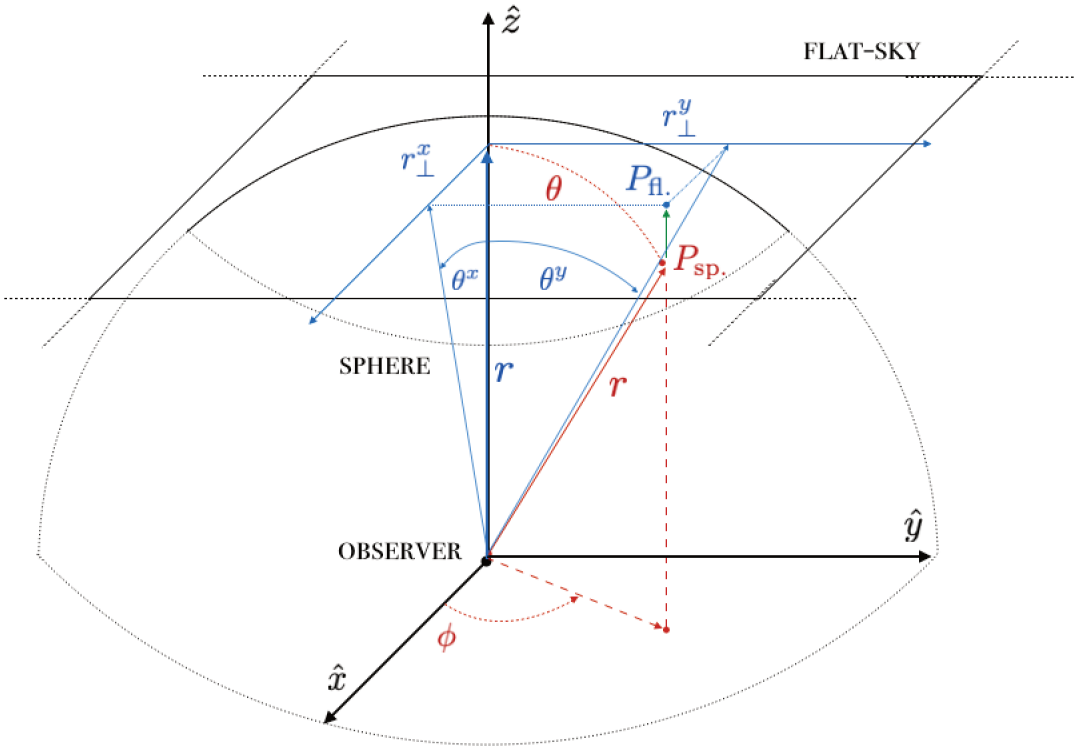

Figure 2.1.4: Graphic representation of how points in the three-dimensional space are identified when the *flat-sky approximation* is assumed. Specifically, the point $P_{\text{sp.}}$ is mapped on the point $P_{\text{fl.}}$ to which an Cartesian system of coordinates is assigned.

which is dependent on the properties of the survey

$$\mathcal{Q}_{(i)}(z) = \int_z^\infty dz_s \, \mathcal{F}(z_s, i) \, \mathcal{S}(z_s, z) \, . \tag{2.72}$$

The function $\mathcal{S}(z_s, z)$, also known as *lensing efficiency*, is

$$\mathcal{S}(z_s, z) = \frac{3 (H_0)^2 \Omega_{m,0}}{2} (1+z) \, \mathcal{W}(z_s, z) = \frac{3 (H_0)^2 \Omega_{m,0}}{2} (1+z) \, \omega(z) \, \frac{\chi(z_s, z)}{\omega(z_s)} \tag{2.73}$$

while $\mathcal{F}(z_s, i)$ will simply be, in absence of any photometric error, the expected (normalised) projected number density of the sources in the i^{th} bin,

$$\mathcal{F}(z_s, i) \equiv \frac{n_i(z_S)}{\bar{n}_i}, \quad n_i(z) = \begin{cases} n(z_s), & \text{for} \quad z_s \in b_i, \\ 0, & \text{for} \quad z_s \notin b_i. \end{cases} \tag{2.74}$$

The quantity $\bar{n}_i$ is the total expected number of sources within the i^{th} bin.

FLAT-SKY APPROXIMATION Following the discussion around Eq. (1.52), we would like to compute the polyspectra for the cosmological field introduced. The important difference is that the tomographic weak lensing convergence field in Eq. (2.71) is defined over a sphere centred on the observer. Therefore, spherical harmonic functions are the suitable basis onto which expand our field. However, it may not be useful for us to explore the largest scales at which the spherical formalism actually applies. As a matter of fact, in Chapter 3 we will see that the weak lensing observables we will use for our cosmological analyses do have a very low signal-to-noise ratio at these scales while being extremely informative at the small ones. When this limit is considered, the modes can be treated as flat leading to an almost negligible impact on the polyspectra. Conventionally, such approximation is called *flat-sky approximation*. In a nutshell, at small enough scales it assumes that the description of our field can be properly done on a flat surface orthogonal to a specific line-of-sight. As long as this approximation holds we will be allowed to move to the configuration space via a much simpler two-dimensional Fourier transform. We will single out the details in the following and we refer to Fig. 2.1.4 for a graphic help.

In general, the value of the matter field in a given point $P_{\rm sp.}$ of the (real, spherical) sky at comoving distance ω from the observer is described by a set of coordinates defined on the sphere as $\delta(P_{\rm sp.}) = \delta\left(\mathbf{r} = \hat{\mathbf{n}}\omega\right)$, $\hat{\mathbf{n}}$ being the unity vector giving the angular position of the point. For sake of completeness $\hat{\mathbf{n}} = \{\sin\theta\,\cos\varphi, \sin\theta\,\cos\varphi, \cos\theta\}$. We also introduce a second set of coordinates by fixing a preferential axe, in Fig. 2.1.4 the $\hat{z}$ one. At each value of the comoving distance ω on it, we can define an orthogonal plane to it. In this new system, a general point $P_{\rm fl.}$ will be described by 1) the comoving distance ω of the plane on which it lies, and 2) by a two-dimensional vector $\mathbf{r}_\perp = \{r^x_\perp, r^y_\perp\}$ defining its position on the plane. For small angular deviation $\Theta = \{\theta^x, \theta^y\}$ of the point $P_{\rm fl.}$ we can approximate $\mathbf{r}_\perp \approx \Theta D_A(\omega)$. We are then ready to define the *flat-sky approximation*. Specifically, for small enough angular deviations Θ we can assume that $\delta(P_{\rm sp.}) \approx \delta(P_{\rm fl.})$ and in terms of coordinates $\delta(\hat{\mathbf{n}}\omega) \approx \delta(\omega, \Theta D_A(\omega))$. We will call the chosen axe, here $\hat{z}$, as *line-of-sight*. Since we are considering spatially-flat Universe, $D_A(\omega) = \omega$.

The main consequence of the flat-sky approximation is the capacity to employ a two-dimensional Fourier transform for each of the planes identified on the line-of-sight

$$\kappa^{(i)}_{\boldsymbol{\ell}} = \int d^2\Theta\, \kappa_{(i)}\left(\Theta\right) e^{+i\boldsymbol{\ell}\cdot\Theta}, \tag{2.75}$$

$$\kappa_{(i)}\left(\Theta\right) = \int_o^{z_{i+1}} dz \left[\frac{\mathrm{d}\omega(z)}{\mathrm{d}z}\right] \mathcal{Q}_{(i)}\left(z\right) \delta\left[\omega(z), \Theta\omega(z)\right] \tag{2.76}$$

In the light of the up-coming calculations, we employ the simpler notation $\kappa_{\boldsymbol{\ell}} \equiv \tilde{\kappa}\left(\boldsymbol{\ell}\right)$. Even if the work presented in this manuscript is mainly focused on the convergence field, realistic survey analyses are performed at the level of the shear field $\gamma = \gamma_1 + i\gamma_2$ defined in Eq. (2.51). However, we can univocally relate them at the level of Fourier modes by inverting Eq. (2.57) (neglecting the tomography)

$$\kappa_{\boldsymbol{\ell}} = \ell^{-2}\left[(\ell_1^2 - \ell_2^2)\,\tilde{\gamma}_1(\boldsymbol{\ell}) + 2\ell_1\ell_2\tilde{\gamma}_1(\boldsymbol{\ell})\right]. \tag{2.77}$$

It is worth highlighting that the explanation given in the previous paragraph holds in real space. We should also

prove the flat-sky approximation in configuration space to build our polyspectra. Luckily, the small scale limit of an harmonic expansion is approximated by a Fourier expansion [White et al., 1999; Hu, 2000; Bernardeau et al., 2011]. We will review in the following the main steps of this peculiar proof.

A general scalar field $\psi(r,\theta,\varphi)$ defined on a sphere can be decomposed in terms of spherical harmonics as

$$\psi(r,\theta,\varphi) = \sum_{\ell} \sum_{m=-\ell}^{+\ell} \hat{\psi}_{\ell,m}(r)\, Y_\ell^m(\theta,\varphi) \tag{2.78}$$

for which a commonly used representation is

$$Y_\ell^m(\theta,\varphi) = \sqrt{\frac{2\ell+1}{4\pi}\frac{(\ell-m)!}{(\ell+m)!}} P_\ell^m(\cos\theta)\, e^{im\varphi}. \tag{2.79}$$

P_ℓ^m is the associated Legendre polynomial. The angle $\theta \in [0,\pi]$ and $\varphi \in [0,2\pi)$ are respectively the polar and azimuth angle on the sphere, as already introduced in the first chapter and identified in Fig. 2.1.4. In configuration space, the flat-sky limit corresponds to small azimuth angular deviation $\theta \ll 1$. Angles are given in radian. We define the two-dimensional Fourier transform of the field $\psi(r,\theta,\varphi)$ as the following function of the the components $\hat{\psi}_{\ell,m}(r)$

$$\hat{\psi}(r,\boldsymbol{\ell}) \equiv \sqrt{\frac{4\pi}{2\ell+1}} \sum_m i^{-m} \hat{\psi}_{\ell,m}(r)\, e^{i\,m\varphi_\ell}, \tag{2.80}$$

and the inverse is clearly

$$\tilde{\psi}_{\ell,m}(r) = \sqrt{\frac{2\ell+1}{4\pi}} i^m \int \frac{2\varphi_\ell}{2\pi} e^{-im\varphi_\ell} \tilde{\psi}(r,\boldsymbol{\ell}) \tag{2.81}$$

with $\ell \equiv |\boldsymbol{\ell}|$ and $\varphi_\ell \equiv \tan^{-1}(\ell_1/\ell_2)$. It is then possible to prove that in the small scale limit Eq. (2.80) does actually provide the coefficient of a Fourier expansion of the field. In details,

$$\begin{aligned}\psi(r,\theta,\varphi) = \sum_{\ell} \sum_{m=-\ell}^{+\ell} \hat{\psi}_{\ell,m}(r)\, Y_\ell^m(\theta,\psi) \approx \\ \sum_{\ell} \frac{\ell}{2\pi} \int\int \frac{2\varphi_\ell}{2\pi}\, \hat{\psi}(r,\boldsymbol{\ell}) \sum_{m=-\ell}^{+\ell} J_m(\ell\,\theta)\, i^m e^{i\,m\,(\varphi-\varphi_\ell)} \approx \\ \int \frac{d^2\ell}{(2\pi)^2} \tilde{\psi}(r,\boldsymbol{\ell})\, e^{i\,\boldsymbol{\ell}\cdot\boldsymbol{\Theta}}\end{aligned} \tag{2.82}$$

where we used the approximation for small angles $\theta \ll 1$ for the spherical harmonic

$$Y_\ell^m(\theta,\varphi) \approx \sqrt{\frac{2\nu}{4\pi}} (-1)^m J_m(\nu\theta) e^{i\,m\,\varphi}, \qquad \nu = \ell + \frac{1}{2}, \tag{2.83}$$

and the Jacobi-Anger expansion for two-dimensional plane waves

$$e^{i\,z\cos\theta} = \sum_{n=-\infty}^{+\infty} J_n(z)\, e^{i\,n\,\theta}, \tag{2.84}$$

with J_n being the n$^{\text{th}}$ order Bessel function of the first kind.

Convergence polyspectra Given the tomographic convergence field in configuration space in Eq. (2.75), we can define the general polyspectra in the flat-sky approximation as

$$\left\langle \kappa^{(i_1)}_{\boldsymbol{\ell}_1} \dots \kappa^{(i_n)}_{\boldsymbol{\ell}_n} \right\rangle \equiv (2\pi)^2\, P_{i_1\dots i_n}(\boldsymbol{\ell}_2, \dots, \boldsymbol{\ell}_n)\, \delta_{\mathrm{D}}(\boldsymbol{\ell}_1 + \dots + \boldsymbol{\ell}_n)\,. \tag{2.85}$$

and in particular it is proved to be consistent with the small scale limit of the definition for the polyspectra on the sphere [Hu, 2000]. We saw in Eq. (2.69) that it is possible to relate the convergence field to the density matter field. Naturally, we can find a relation between their polyspectra. Indeed we will evaluate those for the former as a projection of those for the latter. In a general fashion,

$$\begin{aligned}
&\left\langle \kappa^{(i_1)}_{\boldsymbol{\ell}_1} \dots \kappa^{(i_n)}_{\boldsymbol{\ell}_n} \right\rangle_c \\
&= \left\langle \prod_{i=1}^{n} \int_0^{\omega_i^{\mathrm{u}}} d\omega_i \left[\mathcal{Q}_{(i)}(\omega_i) \int d^2\Theta_i\, \delta\left[\mathbf{x}_i^{\perp} = \Theta_i \omega_i,\ \omega_i\right] e^{+i\boldsymbol{\ell}_i\cdot\Theta_i} \right] \right\rangle_c \\
&= (2\pi)^{-n} \left\langle \prod_{i=1}^{n} \int_0^{\omega_i^{\mathrm{u}}} d\omega_i \left[\mathcal{Q}_{(i)}(\omega_i) \int dk_i^{\parallel} e^{-i\,k_i^{\parallel}\omega_i} \int \frac{d^2\mathbf{x}_i^{\perp}}{\omega_i^2}\, \delta\left[\mathbf{x}_i^{\perp} = \Theta_i\omega_i,\ k_i^{\parallel}\right] e^{+i\frac{\boldsymbol{\ell}_i}{\omega_i}\cdot\mathbf{x}_i^{\perp}} \right] \right\rangle_c \\
&= (2\pi)^{-n+3} \left[\prod_{i=1}^{n} \int_0^{\omega_i^{\mathrm{u}}} d\omega_i\, \frac{\mathcal{Q}_{(i)}(\omega_i)}{\omega_i^2} \right] \left[\prod_{i=1}^{n} \int dk_i^{\parallel} e^{-i\,k_i^{\parallel}\omega_i} \right] \left(P\left(\mathbf{k}_2^{\perp}, \dots \mathbf{k}_n^{\perp}\right) \delta_D\left(\mathbf{k}_1 + \dots + \mathbf{k}_n\right)|_{k_i^{\perp} = \ell_i/\omega_i} \right) \\
&\underset{\text{Lim.}}{\approx} (2\pi)^{-n+3} \left[\prod_{i=1}^{n} \int_0^{\omega_i^{\mathrm{u}}} d\omega_i\, \frac{\mathcal{Q}_{(i)}(\omega_i)}{\omega_i^2} \right] P(\boldsymbol{\ell}_2/\omega_2, \dots, \boldsymbol{\ell}_n/\omega_n)\,|_{k_i^{\parallel}\approx 0} \delta_D\left(\mathbf{k}_1^{\perp} + \dots + \mathbf{k}_n^{\perp}\right) \times \\
&\qquad\qquad \prod_{i=1}^{n} \int dk_i^{\parallel} e^{-i\,k_i^{\parallel}\omega_i} \delta^{(n)}\left(k_{\parallel}^1 + \dots + k_{\parallel}^n\right)
\end{aligned} \tag{2.86}$$

where, in the third line, we performed the change of variable $\Theta_i \to \mathbf{x}_i^{\perp} = \omega_i\Theta_i$ to allow, in the fourth line, a standard three-dimensional Fourier transformation. We also defined the upper integration limit $\omega_i^{\mathrm{u}} \equiv \omega(z_{i+1})$. At this point the ensemble average applies on the product of the instances of the matter field δ leading to the three-dimensional matter polyspectrum according to its definition in Eq. (1.53). We also assumed that the latter does not depend, in general, on $k_{\parallel}$. This approximation is called *Limber approximation* [Loverde & Afshordi, 2008] and it is valid on small scales, i.e. as long as the *flat-sky approximation* holds (see for example Sec. 9.1 in Dodelson [2003] for an in depth discussion). Shortly, if we are interested in measuring the level of correlation for a configuration involving one multipole $\boldsymbol{\ell}$, then we will have to project over matter polyspectra depending on the three-dimensional mode $\mathbf{k} = (\boldsymbol{\ell}/\omega, k_{\parallel})$, up to a given redshift. However, polyspectra involving $k_{\parallel} \gtrsim 1/\omega$

are suppressed by the oscillating factor $e^{-i\,k^{\parallel}\omega}$. If ℓ is enough large to ensure $\ell/\omega \gg k_{\parallel}$ over the typical length of the projection, then we will simply approximate the full vector as $\mathbf{k} \approx (\boldsymbol{\ell}/\omega, 0)$. Under this assumption, we can re-phrase the last term of the integral as

$$\prod_{i=1}^{n}\int dk_{\parallel}^{i} e^{-i\,k_i^{\parallel}\omega_i}\delta_D\left(k_1^{\parallel}+\cdots+k_n^{\parallel}\right) = \prod_{i=2}^{n}\int dk_i^{\parallel} e^{-i\,k_i^{\parallel}\omega_i} e^{+i\,\omega_1\sum_{j=2}^{n}k_j^{\parallel}} = \prod_{i=2}^{n}\int dk_i^{\parallel} e^{i\,k_i^{\parallel}(\omega_1-\omega_i)} = (2\pi)^{+n-1}\prod_{i=2}^{n}\delta_{\mathrm{D}}\left(\omega_1-\omega_i\right). \quad (2.87)$$

Therefore

$$\left\langle \kappa_{\boldsymbol{\ell}_1}^{(i_1)}\ldots\kappa_{\boldsymbol{\ell}_n}^{(i_n)}\right\rangle_c = (2\pi)^2\int_0^{\infty} d\omega\,\frac{\left[\prod_{\hat{i}=i_1}^{i_n}\mathcal{Q}_{(\hat{i})}(\omega)\right]}{\omega^{2n}}P\left(\boldsymbol{\ell}_2/\omega,\ldots,\boldsymbol{\ell}_n/\omega\right)\big|_{k_i^{\parallel}\approx 0}\delta_D\left(\mathbf{k}_1^{\perp}+\cdots+\mathbf{k}_n^{\perp}\right) = (2\pi)^2\int_0^{\infty} d\omega\,\frac{\left[\prod_{\hat{i}=i_1}^{i_n}\mathcal{Q}_{(\hat{i})}(\omega)\right]}{\omega^{2n-2}}P\left(\boldsymbol{\ell}_2/\omega,\ldots,\boldsymbol{\ell}_n/\omega\right)\big|_{k_i^{\parallel}\approx 0}\delta_D\left(\boldsymbol{\ell}^1+\cdots+\boldsymbol{\ell}^n\right), \quad (2.88)$$

where in the second line we changed variable in the Dirac delta for each of the two components of $\mathbf{k}_i^{\perp}$. The powers of (2π) cancel out according to our conventions (1.52). Finally we recognise the structure of Eq. (2.85), leading to the following relation

$$P_{i_1\ldots i_n}\left(\boldsymbol{\ell}_2,\ldots,\boldsymbol{\ell}_n\right) = \int_0^{\infty} d\omega\,\omega^{2-2n}\left[\prod_{\hat{i}=i_1}^{i_n}\mathcal{Q}_{(\hat{i})}(\omega)\right]P\left(\mathbf{k}\left(\boldsymbol{\ell}_2,\omega\right),\ldots,\mathbf{k}\left(\boldsymbol{\ell}_n,\omega\right)\right) \equiv \int_0^{\infty} d\omega\,\mathcal{T}\left(i_1,\ldots,i_n;\omega\right)P\left(\mathbf{k}\left(\boldsymbol{\ell}_2,\omega\right),\ldots,\mathbf{k}\left(\boldsymbol{\ell}_n,\omega\right)\right). \quad (2.89)$$

the three-dimensional mode given by the Limber dependence $\mathbf{k}(\boldsymbol{\ell},\omega) \approx (\boldsymbol{\ell}/\omega, 0)$. Before moving to the next section, we would like to underline that the fact the three-dimensional polyspectrum is evaluated on contributions coming from the same comoving distance ω is deeply connected to the assumption we made of weak dependence on transverse modes $k^{\parallel}$ (Limber approximation).

2.2 Covariance Matrix, a complex evaluation

We now turn to the evaluation of the covariance. In this section we will discuss about this topic in the context of the weak lensing convergence field. Specifically, our observables will be the tomographic power spectrum and the

tomographic bispectrum of it

$$P_{ij}(\ell) = P_{ij}(\boldsymbol{\ell}), \tag{2.90}$$

$$B_{ijk}(\ell_1, \ell_2, \ell_3) = P_{ijk}(\boldsymbol{\ell}_2, \boldsymbol{\ell}_3), \tag{2.91}$$

where, again, the dependence on just the modules of the momenta involved comes from the request of isotropy and homogeneity. We recall that, according to the notation introduced in Chapter 1, $\ell \equiv |\boldsymbol{\ell}|$.

2.2.1 Why analytical covariances?

Covariance matrix, an introduction Given a data vector $\{\mathcal{O}_i\}_{i=1,\dots,N}$, the measured values are stochastic events and their expectation values can be derived theoretically if the probability density functional of the field their are drawn from is known. The *variance* of each observable $\{\sigma_i^2\}_{i=1,\dots,N}$ and the *covariances* among them $\{\sigma_{ij}^2 = \sigma_{ji}^2\}_{i,j=1,\dots,N}$ are in turn stochastic variables, being function of them

$$\sigma_{ij}^2 \equiv \langle \mathcal{O}_i \mathcal{O}_j \rangle - \langle \mathcal{O}_i \rangle \langle \mathcal{O}_j \rangle, \tag{2.92}$$

$$\sigma_i^2 \equiv \sigma_{ii}^2 \tag{2.93}$$

Variance and correlations for the data vector are given as element of the *Covariance matrix*. The covariance matrix is a square, symmetric and positive defined matrix which has dimension N, N begin the number of observables we want to measure, where the diagonal elements are the variances and the off-diagonal elements are the covariances among them

$$\mathrm{Cov}[\mathcal{O}_i, \mathcal{O}_j] = \sigma_{ij}^2. \tag{2.94}$$

The *correlations* are simply the covariances once normalised over the variances

$$\mathrm{corr}\left[\mathcal{O}_i, \mathcal{O}_j\right] = \mathrm{corr}\left[\mathcal{O}_j, \mathcal{O}_i\right] = \frac{\sigma_{ij}^2}{\sigma_i^2\, \sigma_j^2}, \quad i, j = 1, \dots, N. \tag{2.95}$$

Whenever the above definition cannot be applied due to our ignorance on the probability density functional (PDF) of the underlying field, we have to find a way around to evaluate the elements in the covariance via the single realisation we have at our disposal. Let us briefly clarify some important points of our discussion before moving one with our presentation.

First of all, we will be interested in *forecast* for cosmological parameters. Specifically, we want to address the scientific performance of future missions. In this case, real data are not yet available but still we want to understand how well cosmological parameters will be measured and and how we can optimise future data analyses and computational resources. All the considerations given above still apply. However, when referring to data we will be thinking about *mock* data from simulations which mimic the sky observed by a specific survey accounting for all the systematics real data will be affected by.

Secondly, the observables we will consider are polyspectra measuring the amount of correlations in the weak lensing convergence field at different scales and among different binned source positions (tomography). The correlations in the data set will be directly related to the underlying matter clustering . As we saw in Sec. 1.3 couplings among modes arise while evolving the matter field in time and they will be directly mapped into couplings among angular scales at the level of the weak lensing signal. Therefore, the value of the correlations at a given scale (configuration) will be correlated to that one at an another scale (configuration) and so on leading to a significant impact at the level of cosmological parameter error bars [Sato et al., 2009] . Then, the problem of estimating a covariance matrix for modern galaxy survey is actually indeed an interesting physics problem: covariance matrix estimation is difficult because the physics of non-linear gravitational collapse is not sufficiently well described by up-to-date models of clustering.

Estimating the Covariance matrix In an ideal scenario, the optimal way to compute a covariance matrix would be to run $N_s \to \infty$ large and accurate hydro-simulations of cosmological structure formation. The optimality in this approach arises from the fact that numerical simulations are to date our best tool to describe non-linear structure formation and the properties of the visible structures within. These faithful mock realisations of the Universe would then be used to estimate the true covariance via the sample covariance of the observables $\mathcal{O}_i$ and $\mathcal{O}_j$, built as

$$\hat{\mathrm{Cov}}_{ij} = \frac{1}{N_s - 1} \sum_{I=1}^{N_s} \left(\mathcal{O}_{i,I} - \langle \mathcal{O}_i \rangle\right) \left(\mathcal{O}_{j,I} - \langle \mathcal{O}_j \rangle\right) , \tag{2.96}$$

where $\mathcal{O}_I$ is the I$^{\text{th}}$ measurement of the observable $\mathcal{O}$, and the expectation value of it can be estimated via the arithmetic mean $\langle \mathcal{O} \rangle \equiv \sum_I \mathcal{O}_I / N_s$. However, we do have tight minimum requirements on N_s for a given set of observables. Taylor et al. [2013] derived the minimum number of samples N_s that are used to compute the covariance to obtain a fractional precision better than ε on parameter constraints, assuming a Gaussian weak lensing field

$$n_s > \frac{2}{\varepsilon^2} + N + 4, \tag{2.97}$$

N being the dimension of the data vector. As we will see, we will be dealing with vectors of dimension $\sim 10^4$, which makes a precise evaluation of the covariance matrices from simulations extremely challenging. Furthermore, pure dark-matter simulations produce power spectra to about 1% accuracy up to $k < 1\, h\,\mathrm{Mpc}^{-1}$, or 3% up to $k < 10\, h\,\mathrm{Mpc}^{-1}$ [Schneider et al., 2016]. For a Euclid-like survey for example, Kitching & Taylor [2011] proved that the power spectrum has to be accurately known to 1% down to $k \approx 50\, h\,\mathrm{Mpc}^{-1}$ to saturate the dark energy figure of merit and reach Euclid design goals. This study is in agreement with the previous one from Doré et al. [2009]. As this vastly exceeds the computational power available to the community (and will continue to do so in the next decade), and is out of proportion compared to other central processing tasks in cosmological analysis, such as modelling the cosmological signal or data reduction, compromises have to be sought that find a balance between accuracy, statistical precision, and computational (or human) cost.

Another class of estimators do exist requiring a lower numerical complexity. Within the so called class of *re-*

sampling methods, or *internal covariance methods*, different estimators allow the computation of the covariance matrix from the data themselves or a single realisation in a simulation. The underlying idea is to create a number of $N_{\text{sub.}}$ spatial sub-volumes from the available data and to estimate the statistical properties of their underlying distribution from the variation of the samples. One example is the *Jackknife* (J) method. In this case, the I$^{\text{th}}$ sample is built as the full data set minus the I$^{\text{th}}$ sub-sample. The covariance matrix estimator for this method is

$$\hat{\text{Cov}}_{ij}^{(J)} = \frac{N_{\text{sub.}}}{N_{\text{sub.}}-1}\sum_{I=1}^{N_{\text{sub.}}}\left(\mathcal{O}_{i,I}^{(J)} - \langle\mathcal{O}_{i}^{(J)}\rangle\right)\left(\mathcal{O}_{j,I}^{(J)} - \langle\mathcal{O}_{j}^{(J)}\rangle\right), \tag{2.98}$$

where $\mathcal{O}_{,I}^{(J)}$ is the value of an observable as estimated from the I$^{\text{th}}$ Jackknife sample via arithmetic mean and $\langle\mathcal{O}^{(J)}\rangle$ is the mean of all the all the $\mathcal{O}_{,I}^{(J)}$.

While these methods are correct with respect to the estimation of the covariance between modes which we can measure from within the data set, Lacasa & Kunz [2017] showed, as it can be expected, that they are unable to catch correlations with modes larger or equal to the size of the survey footprint. As we will see in Sec. 2.3 these type of correlations are indeed present and play a non-negligible role in our study: they increase the correlations between the observed scales and they correlate different sub-samples as extracted from the same data set. Specifically *internal covariance methods* fail because they intrinsically assume that the sub-samples are independent.

At this point, it should be pretty clear the answer to the question addressed in the title of this section. From one side, the present performance of N-body simulations does not allow a precise estimation of the covariance given the number of observables modern galaxy surveys will give the chance to measure. On the other hand, internal estimators miss super-survey correlations overestimating the available cosmological information. Therefore, even if it is an extremely challenging path to follow, the exact implementation of Eq. (2.92) for a given theoretical model is the most reliable way to produce unbiased covariance matrices. *This will be indeed the approach we will consider*. In the following section, we will provide a detailed explanation of how to derive the covariance matrix for the two observables of interest from an analytical perspective. For the observed scales of interest, i.e. much smaller the typical length of the survey view field, we will schematically split the covariance (and we will prove to be allowed to do so in this limit) as

$$\text{Cov}\left[P_{ij}(\ell), P_{i'j'}(\ell')\right] = \text{Cov}\left[\dots\right]_{\text{Gauss}} + \text{Cov}\left[\dots\right]_{\text{NGins}} + \text{Cov}\left[\dots\right]_{\text{NGssc}}, \tag{2.99}$$

$$\text{Cov}\left[B_{ijk}(\ell_1,\ell_2,\ell_3), B_{i'j'k'}(\ell_1',\ell_2',\ell_3')\right] = \text{Cov}\left[\dots\right]_{\text{Gauss}} + \text{Cov}\left[\dots\right]_{\text{NGins}} + \text{Cov}\left[\dots\right]_{\text{NGssc}}, \tag{2.100}$$

$$\text{Cov}\left[P_{ij}(\ell), B_{i'j'k'}(\ell_1,\ell_2,\ell_3)\right] = \text{Cov}\left[\dots\right]_{\text{NGins}} + \text{Cov}\left[\dots\right]_{\text{NGssc}}. \tag{2.101}$$

In Eq (2.99) and (2.101), we label as *Gauss* the covariance terms containing only 2-point statistics and which would be the only contribution if the matter density field were Gaussian. The other covariance terms arise due to the non-Gaussian statistics induced by the gravitational collapse and correlate different modes and the different probes, i.e. power spectrum and bispectrum. We distinguish two classes of non-Gaussian covariance, respectively labelled as *NGins* and *NGssc*. The former is sourced by correlations between observed *intra-survey* modes (see later

in the section), while the latter is sourced by correlations between observed modes and background *super-survey* modes and is known in the literature as *super-sample covariance* (see Sec. 2.3). Within the present Chapter, whenever needed, we will label the total non-Gaussian component of a specific covariance matrix with the subscript *NG*.

For completeness, we highlight that analytical covariance matrices were used for cosmological analyses in Krause & Eifler [2017] where cosmic shear, galaxy-galaxy lensing, galaxy clustering, cluster number counts and cluster weak lensing were combined to access the information content for a LSST-like survey. Subsequently, the COSMOLIKE likelihood code developed as result of the above work was officially used within the DES community [Krause et al., 2017] as part of the scientific validation of the the DES-Year1 release [Dark Energy Survey Collaboration et al., 2016]. However, all the works here mentioned were done at the level of 2-point statistics.

2.2.2 Observable estimators

While we have discussed a lot about the covariance estimators, it is necessary to start with a preliminary but crucial step. So far we have discussed about the theoretical values of the polyspectra: we started from the matter field δ, we defined its three-dimensional polyspectra via Eq. (1.53) and then we built those for the convergence field via the *line-of-sight integration* in Eq. (2.89). However, in real data analysis, we will never measure these quantities, nor a realisation of them.

Finite sky effects In real world application we will always measure the values of the polyspectra for a pixelized and domain-limited version of the underlying true field. This may induce differences between the observed properties and the expected ones. The super-sample covariance (we will see) is a nice example of this feature and it is related to the finite sky coverage of our survey. The finite real-space domain has an impact already at the level of observables. Neglecting for the moment the resolution of the instrument, in realistic situations the actual field whose polyspectra we are measuring is a filtered version of the real one

$$\kappa^{\mathrm{w},(i)}(\Theta) = W(\Theta)\,\kappa^{(i)}(\Theta) \tag{2.102}$$

where the window function $W(\Theta)$ describes the footprint of the survey. In our analyses we will assume a simple binary function with $W = 0$ or $W = 1$ for points outside and inside the survey field respectively. The Fourier transform of the masked field will then be the convolution between the Fourier transform of the window function $\hat{W}$ and of the true underlying convergence field

$$\kappa^{\mathrm{w},(i)}_{\ell} = \int d^2\mathbf{q}\tilde{W}(\mathbf{q})\,\kappa^{(i)}_{\ell-\mathbf{q}}. \tag{2.103}$$

We introduce now very useful survey related quantities: $\Theta_{\rm sky}$ (angular coverage of the survey), $\Omega_{\rm sky}$ (solid angle covered by the survey) and $f_{\rm sky}$ (fraction of the sky covered by the survey). These quantities are related as follows

$$\Omega_{\rm sky} = 2\pi\left(1-\cos\Theta_{\rm sky}\right), \quad f_{\rm sky} = \frac{\Omega_{\rm sky}}{4\pi}, \tag{2.104}$$

all the angles being expressed in radian or steradian. If the survey footprint has an angular size of $\Theta_{\rm sky}$ steradian, the convolution in Eq. (2.103) will induce correlations at the level of *true field* at scales $\ell \leqslant \ell_{\rm f}$ where $\ell_{\rm f} = 2\pi/\Theta_{\rm sky}$ is the typical domain length of $\hat{W}$. The frequency $\ell_{\rm f}$ is also known as *fundamental frequency* of the survey. At the level of collected data, the finite real-space domain implies that we cannot measure, in the *observed field*, angular modes $\ell \leqslant \ell_{\rm f}$ and also $\ell_{\rm f}$ provides a minimum resolution for the module of the multipoles we can access

$$\boldsymbol{\ell}^{n_y}_{n_x} = \ell_{\rm f}\{n_x, n_y\}, \quad n_x,\ n_y \in \mathbb{N}^{\times}. \tag{2.105}$$

Therefore, a good approximation to describe the *observed field* within the survey in configuration space is

$$\kappa^{(i)}_{\boldsymbol{\ell}^{n_y}_{n_x}} = \int_{\Omega_{\rm sky}} d^2\Theta\, \kappa^{(i)}(\Theta)\, e^{-i\,\boldsymbol{\ell}^{n_y}_{n_x}\cdot\Theta}, \tag{2.106}$$

where the *Discrete Fourier Transform* has to be used [Jasche et al., 2009]. We can see from Eq. (2.106) that all the modes smaller than the fundamental frequency are excluded while still present in the real field. The question is: *will the statistical properties (expectation value, polyspectra, covariances, ...) of the measured* $\kappa^{(i)}_{\boldsymbol{\ell}^{n_y}_{n_x}}$ *in Eq.* (2.106) *be the same of the true* $\kappa^{{\rm w},(i)}_{\boldsymbol{\ell}}$ *in Eq.* (2.103) *?* The answer is not obvious and indeed we will see that while it is true for the observables (under certain assumptions), the covariance will be deeply affected. From a more general perspective, whenever we want to measure an observable $\mathcal{O}$ whose true value is linked to underlying cosmology (our final target), we build an *estimator* $\hat{\mathcal{O}}$ of it as function of the data. If the expectation value of the estimator is equal to the desired observable, then the estimator is said to be *unbiased*. With the help of this terminology, we can anticipate one of the most important results in this section by saying that the covariance matrix of the discretized (observed) field in Eq. (2.106) does provide a biased estimator of the true covariance of the actual masked field at Eq. (2.103), the difference being the super-sample covariance. Indeed, the curse of this bias lies in neglecting modes $\ell \leqslant \ell_{\rm f}$ in the observed field, which are indeed present, while not observable. We will see this more quantitatively in Sec. 2.3. In the following part of the present section we will focus instead on the correlations between modes entirely observed for which the approximation in Eq. (2.106) applies. Interestingly, we can test this component of the covariance in simulations: unless run as sub-boxes of much larger simulations, the periodicity of the boundary conditions makes Eq. (2.106) exact for the simulated field, dissolving the super-sample correlations (which are indeed entirely absent) (see de Putter et al. [2012] for a thorough discussion). For our study we will consider a specific class of polyspectra estimators which are the *binned estimators*: for optimising of the analysis, to avoid numerical inaccuracy and to increase the signal-to-noise ratio, we will bin the Fourier space according to the module of the momenta considered. These bins can be much larger than the fundamental length $\ell_{\rm f}$ and the value of the estimator associated to them will be constructed via an average over the fundamental modes included

in each.

BINNED ESTIMATORS We define the following estimator for the *binned tomographic convergence power spectrum* (BTP) in the bin ℓ^b and for the *binned tomographic convergence bispectrum* (BTB) on the bins $\{\ell_1^b, \ell_2^b, \ell_3^b\}$

$$\hat{P}_{ij}\left(\ell^b\right) \equiv \frac{1}{\Omega_{\rm sky}\, N_{\rm p}\left(\ell^b\right)} \sum_{\boldsymbol{\ell}'} \kappa^{(i)}_{\boldsymbol{\ell}'} \kappa^{(j)}_{-\boldsymbol{\ell}'} \Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right), \tag{2.107}$$

$$\hat{B}_{ijk}\left(\ell_1^b, \ell_2^b, \ell_3^b\right) \equiv \frac{1}{\Omega_{\rm sky}\, N_{\rm tri.}\left(\ell_1^b, \ell_2^b, \ell_3^b\right)} \sum_{\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3} \kappa^{(i)}_{\boldsymbol{\ell}'_1} \kappa^{(j)}_{\boldsymbol{\ell}'_2} \kappa^{(k)}_{\boldsymbol{\ell}'_3} \Delta^3_{\ell_1^b, \ell_2^b, \ell_3^b}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right). \tag{2.108}$$

In the following, we will refer to two-dimensional multipoles by writing them in bold, e.g. $\boldsymbol{\ell}$, while the corresponding module will simply be ℓ. The bins are defined at the level of modules: $\boldsymbol{\ell}' \in \ell^b$ if $\ell - \Delta\ell^b/2 \leqslant \ell' \leqslant \ell + \Delta\ell^b/2$, ℓ being the central value of the bin ℓ^b of width $\Delta\ell^b$. The selection function $\Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right) = 1$ if and only if the length of the mode $\boldsymbol{\ell}'$ falls into the required bin ℓ^b. The selection function $\Delta^3_{\ell_1^b, \ell_2^b, \ell_3^b}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right) = 1$ applies this constraint separately on the 3 modes, i.e. $\ell_i - \Delta\ell_i^b/2 \leqslant \ell'_i \leqslant \ell_i + \Delta\ell_i^b/2$ and further requires that $\boldsymbol{\ell}'_1 + \boldsymbol{\ell}'_2 + \boldsymbol{\ell}'_3 = 0$. $N_{\rm p}\left(\ell\right)$ and $N_{\rm tri.}\left(\ell_1^b, \ell_2^b, \ell_3^b\right)$ respectively normalise the sums over the total number of independent elements in the bin. We will derive them in the approximation ℓ_i (center of the bin) $\gg \ell_{\rm f}$. Specifically, given a bin ℓ^b of width $\Delta\ell^b$, the number of independent modes discriminated by the fundamental length $\ell_{\rm f}$ is [Joachimi et al., 2009; Kayo et al., 2013]

$$N_{\rm p}\left(\ell^b\right) \equiv \sum_{\boldsymbol{\ell}'} \Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right) \approx \frac{2\pi\ell\Delta\ell^b}{\left(\frac{2\pi}{\Theta_{\rm sky}}\right)} = 2\,\ell\,\Delta\ell\, f_{\rm sky}. \tag{2.109}$$

The number of independent triplets we can combine given a set of bins and related widths $\{\ell_i^b, \Delta\ell_i^b\}_{i=1,2,3}$ can be derived by simple geometrical considerations based on the discretisation of the momenta according to the fundamental frequency $\ell_{\rm f}$ [Joachimi et al., 2009; Kayo et al., 2013]:

$$N_{\rm tri.}\left(\ell_1^b, \ell_2^b, \ell_3^b\right) = \sum_{\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3} \Delta^3_{\ell_1^b, \ell_2^b, \ell_3^b}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right) \approx 2 \times \left(\frac{2\pi\ell_1\Delta\ell_1^b}{\ell_{\rm f}^2}\right) \times \left(\frac{\ell_2\Delta\ell_2^b\Delta\varphi_{12}}{\ell_{\rm f}^2}\right) \tag{2.110}$$

where the term within the first parenthesis gives the number of possible momenta within the annuls of target length ℓ_1 and width $\Delta\ell_1^b$. The term within the second ones is instead the number of allowed momenta within the bin ℓ_2^b. The information on the constraint to form a close configuration with the multiple $\boldsymbol{\ell}'_3 \in \ell_3^b$ is given by the angle $\Delta\varphi_{12}$ which is the maximum allowed angular variation of the angle $\varphi_{12} \in [0, \pi]$ between the multipoles $\boldsymbol{\ell}'_1 \in \ell_1^b$ and $\boldsymbol{\ell}'_2 \in \ell_2^b$. The total number of allowed triplets will simply be the product of these terms. The factor 2 accounts for the parity symmetry around $\boldsymbol{\ell}'_1$. We can find the allowed angular variation $\Delta\varphi_{12}$ by considering a fixed

multipole $\boldsymbol{\ell}'_1 \in \ell^b_1$ and verify the value of its variation while varying $\boldsymbol{\ell}'_3 \in \ell^b_3$

$$\frac{\mathrm{d}\cos\varphi_{12}}{\mathrm{d}\,\boldsymbol{\ell}'_3} = \frac{\mathrm{d}}{\mathrm{d}\,\boldsymbol{\ell}'_3}\left[\frac{\boldsymbol{\ell}'^2_1 + \boldsymbol{\ell}'^2_2 - \boldsymbol{\ell}'^2_3}{2\,\boldsymbol{\ell}'_1\,\boldsymbol{\ell}'_1}\right] = \frac{\boldsymbol{\ell}'_3}{\boldsymbol{\ell}'_2\,\boldsymbol{\ell}'_1},$$
$$\sim \frac{\mathrm{d}\varphi_{12}}{\mathrm{d}\boldsymbol{\ell}'_3} = \frac{\boldsymbol{\ell}'_3}{\boldsymbol{\ell}'_1\,\boldsymbol{\ell}'_2}(\sin\varphi_{12})^{-1},$$
$$\sim \Delta\varphi_{12} \approx \frac{\ell_3\Delta\ell^b_3}{\ell_1\,\ell_2}(\sin\varphi_{12})^{-1} \quad (2.111)$$

where we approximated the derivatives at the central bin values. The angle $\sin\varphi_{12}$ may be expressed in terms of the edges via the Law of sines.

$$\frac{1}{\sin\varphi_{12}} = \frac{2\ell_1\,\ell_2}{\sqrt{2\left(\ell^2_1\,\ell^2_2 + \ell^2_1\,\ell^2_3 + \ell^2_2\,\ell^2_3\right) - \ell^4_1\ell^4_2\,\ell^4_3}}. \quad (2.112)$$

The final desired expression will then be

$$N_{\text{tri.}}\left(\ell^b_1, \ell^b_2, \ell^b_3\right) \approx \frac{2\Omega_{\text{sky}}\ell_1\ell_2\ell_3\Delta\ell^b_1\Delta\ell^b_2\Delta\ell^b_3}{\sqrt{2\ell^2_1\ell^3_2\ell^2_3 - \ell^4_1 - \ell^4_2 - \ell^4_3}}. \quad (2.113)$$

Going back to the estimators (2.107) and (2.108), if the true power spectrum and bispectrum do not changed significantly with the bins, then the two estimators are unbiased

$$\left\langle\hat{P}_{ij}\left(\ell^b\right)\right\rangle = \frac{1}{\Omega_{\text{sky}}\,N_{\text{p}}\left(\ell^b\right)}\sum_{\boldsymbol{\ell}'}\langle\kappa^{(i)}_{-\boldsymbol{\ell}'}\kappa^{(j)}_{\boldsymbol{\ell}'}\rangle\Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right)$$
$$= \frac{1}{\Omega_{\text{sky}}\,N_{\text{p}}\left(\ell^b\right)}\sum_{\boldsymbol{\ell}'}\Omega_{\text{sky}}P_{ij}\left(\ell'\right)\Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right) \approx \frac{P_{ij}\left(\ell\right)}{N_{\text{p}}\left(\ell^b\right)}\sum_{\boldsymbol{\ell}'}\Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right) = P_{ij}\left(\ell_b\right), \quad (2.114)$$

$$\left\langle\hat{B}_{ijk}\left(\ell^b_1, \ell^b_2, \ell^b_3\right)\right\rangle = \frac{1}{\Omega_{\text{sky}}\,N_{\text{tri.}}\left(\ell^b_1, \ell^b_2, \ell^b_3\right)}\sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2,\boldsymbol{\ell}'_3}\langle\kappa^{(i)}_{\boldsymbol{\ell}'_1}\kappa^{(j)}_{\boldsymbol{\ell}'_2}\kappa^{(k)}_{\boldsymbol{\ell}'_3}\rangle\Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right)$$
$$= \frac{1}{\Omega_{\text{sky}}\,N_{\text{tri.}}\left(\ell^b_1, \ell^b_2, \ell^b_3\right)}\sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2,\boldsymbol{\ell}'_3}\Omega_{\text{sky}}B_{ijk}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right)\Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right) \approx$$
$$\approx \frac{B_{ijk}\left(\ell_1, \ell_2, \ell_3\right)}{N_{\text{tri.}}\left(\ell^b_1, \ell^b_2, \ell^b_3\right)}\sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2,\boldsymbol{\ell}'_3}\Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right) = B_{ijk}\left(\ell_1, \ell_2, \ell_3\right) \quad (2.115)$$

and we are in particular free to associate the estimated quantities to the spectra evaluated at the central values of the bins. In Eq. (2.114) and in Eq. (2.115) we employed the definition of polyspectra for a discrete Fourier decomposition of the convergence field [Takada & Bridle, 2007]

$$\left\langle\kappa^{(i_1)}_{\boldsymbol{\ell}_1}\ldots\kappa^{(i_n)}_{\boldsymbol{\ell}_n}\right\rangle \equiv \Omega_{\text{sky}}\,P_{i_1\ldots i_n}\left(\boldsymbol{\ell}_2, \ldots, \boldsymbol{\ell}_n\right)\delta^{\text{K}}_{\boldsymbol{\ell}_1+\cdots+\boldsymbol{\ell}_n}. \quad (2.116)$$

In Eq. (2.116), $\delta^{\mathrm{K}}_{\boldsymbol{\ell}_1+\cdots+\boldsymbol{\ell}_n}$ is the Kronecker delta function enforcing $\boldsymbol{\ell}_1 + \cdots + \boldsymbol{\ell}_n = 0$ for each (discretized) component.

2.2.3 Binned tomographic power spectrum covariance matrix

Once defined suitable estimators for the polyspectra of interest, we need to build the covariance for them. In this section we will derive the covariance for the binned tomographic convergence power spectrum estimator (BTP) by replacing its expression (2.107) in the covariance definition (2.92)

$$\mathrm{Cov}\left[\hat{P}_{ij}\left(\ell^b\right), \hat{P}_{i'j'}(\ell'^b)\right] \equiv \langle \hat{P}_{ij}\left(\ell^b\right) \hat{P}_{i'j'}(\ell'^b)\rangle - \langle \hat{P}_{ij}\left(\ell^b\right)\rangle\langle \hat{P}_{i'j'}(\ell'^b)\rangle = \\ \frac{1}{\Omega_{\mathrm{sky}} N_{\mathrm{p}}\left(\ell\right)} \frac{1}{\Omega_{\mathrm{sky}} N_{\mathrm{p}}\left(\ell'\right)} \sum_{\bar{\boldsymbol{\ell}},\bar{\boldsymbol{\ell}}'} \langle \kappa^{(i)}_{\bar{\boldsymbol{\ell}}} \kappa^{(j)}_{-\bar{\boldsymbol{\ell}}} \kappa^{(i')}_{\bar{\boldsymbol{\ell}}'} \kappa^{(j')}_{-\bar{\boldsymbol{\ell}}'} \rangle \Delta^2_{\ell^b}\left(\bar{\boldsymbol{\ell}}\right) \Delta^2_{\ell'^b}(\bar{\boldsymbol{\ell}}') - P_{ij}\left(\ell\right) P_{i'j'}(\ell'). \quad (2.117)$$

The 4-point correlation function function appearing in Eq. (2.117) can be further split in its non-vanishing [2] connected components as

$$\langle \kappa^{(i)}_{\bar{\boldsymbol{\ell}}} \kappa^{(j)}_{-\bar{\boldsymbol{\ell}}} \kappa^{(i')}_{\bar{\boldsymbol{\ell}}'} \kappa^{(j')}_{-\bar{\boldsymbol{\ell}}'} \rangle = \\ \langle \kappa^{(i)}_{\bar{\boldsymbol{\ell}}} \kappa^{(j)}_{-\bar{\boldsymbol{\ell}}} \rangle_c \langle \kappa^{(i')}_{\bar{\boldsymbol{\ell}}'} \kappa^{(j')}_{-\bar{\boldsymbol{\ell}}'} \rangle_c + \langle \kappa^{(i)}_{\bar{\boldsymbol{\ell}}} \kappa^{(j')}_{-\bar{\boldsymbol{\ell}}'} \rangle_c \langle \kappa^{(i')}_{\bar{\boldsymbol{\ell}}'} \kappa^{(j)}_{-\bar{\boldsymbol{\ell}}} \rangle_c + \langle \kappa^{(i)}_{\bar{\boldsymbol{\ell}}} \kappa^{(i')}_{\bar{\boldsymbol{\ell}}'} \rangle \langle \kappa^{(j)}_{-\bar{\boldsymbol{\ell}}} \kappa^{(j')}_{-\bar{\boldsymbol{\ell}}'} \rangle_c + \\ \langle \kappa^{(i)}_{\bar{\boldsymbol{\ell}}} \kappa^{(j)}_{-\bar{\boldsymbol{\ell}}} \kappa^{(i')}_{\bar{\boldsymbol{\ell}}'} \kappa^{(j')}_{-\bar{\boldsymbol{\ell}}'} \rangle_c \quad (2.118)$$

then, by applying Eq. (2.116),

$$\langle \kappa^{(i)}_{\bar{\boldsymbol{\ell}}} \kappa^{(j)}_{-\bar{\boldsymbol{\ell}}} \kappa^{(i')}_{\bar{\boldsymbol{\ell}}'} \kappa^{(j')}_{-\bar{\boldsymbol{\ell}}'} \rangle = \\ \Omega^2_{\mathrm{sky}} \left[\delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}-\bar{\boldsymbol{\ell}}} P_{ij}\left(\bar{\ell}\right) \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}'-\bar{\boldsymbol{\ell}}'} P_{i'j'}\left(\bar{\ell}'\right) + \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}-\bar{\boldsymbol{\ell}}'} P_{ij'}\left(\bar{\ell}\right) \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}'-\bar{\boldsymbol{\ell}}} P_{i'j}\left(\bar{\ell}'\right) + \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}+\bar{\boldsymbol{\ell}}'} P_{ii'}\left(\bar{\ell}\right) \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}+\bar{\boldsymbol{\ell}}'} P_{jj'}\left(\bar{\ell}\right) \right] + \\ \Omega_{\mathrm{sky}} \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}-\bar{\boldsymbol{\ell}}+\bar{\boldsymbol{\ell}}'-\bar{\boldsymbol{\ell}}'} T_{iji'j'}(\bar{\boldsymbol{\ell}}, -\bar{\boldsymbol{\ell}}, \bar{\boldsymbol{\ell}}' - \bar{\boldsymbol{\ell}}') \quad (2.119)$$

We introduced the trispectrum T as the fourth-order polypsectrum. By replacing this last result in Eq. (2.117)

$$\mathrm{Cov}\left[\hat{P}_{ij}\left(\ell^b\right), \hat{P}_{i'j'}\left(\ell'^b\right)\right] = \frac{1}{N_{\mathrm{p}}\left(\ell^b\right) N_{\mathrm{p}}\left(\ell'^b\right)} \times \\ \left\{ \sum_{\bar{\boldsymbol{\ell}},\bar{\boldsymbol{\ell}}'} \left[P_{ij}\left(\bar{\ell}\right) P_{i'j'}\left(\bar{\ell}'\right) + \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}-\bar{\boldsymbol{\ell}}'} P_{ij'}\left(\bar{\ell}\right) \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}'-\bar{\boldsymbol{\ell}}} P_{i'j}\left(\bar{\ell}'\right) + \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}+\bar{\boldsymbol{\ell}}'} P_{ii'}\left(\bar{\ell}\right) \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}+\bar{\boldsymbol{\ell}}'} P_{jj'}\left(\bar{\ell}\right) \right] \Delta^2_{\ell^b}\left(\bar{\boldsymbol{\ell}}\right) \Delta^2_{\ell'^b}(\bar{\boldsymbol{\ell}}') + \\ \frac{1}{\Omega_{\mathrm{sky}}} \sum_{\bar{\boldsymbol{\ell}},\bar{\boldsymbol{\ell}}'} \left[T_{iji'j'}(\bar{\boldsymbol{\ell}}, -\bar{\boldsymbol{\ell}}, \bar{\boldsymbol{\ell}}', \bar{\boldsymbol{\ell}}') \right] \delta^{\mathrm{K}}_{\bar{\boldsymbol{\ell}}-\bar{\boldsymbol{\ell}}+\bar{\boldsymbol{\ell}}'-\bar{\boldsymbol{\ell}}'} \Delta^2_{\ell^b}\left(\bar{\boldsymbol{\ell}}\right) \Delta^2_{\ell'^b}(\bar{\boldsymbol{\ell}}') \right\} - P_{ij}\left(\ell\right) P_{i'j'}(\ell) \quad (2.120)$$

We start from the first line of the covariance (2.120) which will lead to the Gaussian component of the power spectrum error, i.e. the first term introduced in Eq. (2.99). Let us focus for example on the first (non trivial, then

[2] $\langle \kappa \rangle = 0$.

second) term within the sum. We can remove one of the sum (e.g. $\sum_{\bar{\ell}'}$) via the constraint given by the Kronecker Delta. In general the constraint will be of the kind $\bar{\boldsymbol{\ell}}' = \pm\bar{\boldsymbol{\ell}}$. Given that the power spectrum will depend on the module of the vector, we will simply write $P(\ell') = P(\ell)$. Further, always under the assumption of P not varying too much within the width of the bin, we will factorise the spectra out of the sum and calculate them on the central value of the bin. This implies that the second term in Eq. (2.117) will exactly cancel out the first product within the sum. As for the selection function, the request for $\bar{\boldsymbol{\ell}}' = \pm\bar{\boldsymbol{\ell}}$ will lead to terms like $\sum_{\bar{\boldsymbol{\ell}}} \Delta^2_{\ell^b}\left(\bar{\boldsymbol{\ell}}\right) \Delta^2_{\ell'^b}(\pm\bar{\boldsymbol{\ell}})$ which are not vanishing only if $\bar{\ell}$ lies within the same ℓ^b bin. Therefore, this part of the covariance matrix is not vanishing only if the correlations are between modes both included in the bins ℓ'^b and ℓ^b, i.e. diagonal in the binned momenta space ($\ell = \ell'$ in terms of the central values of the bins). Finally, the Gaussian part of the covariance matrix will be given by

$$\mathrm{Cov}\left[\hat{P}_{ij}\left(\ell^b\right), \hat{P}_{i'j'}(\ell'^b)\right]_{\mathrm{Gauss}} = \frac{\delta_{\ell^b\ell'^b}}{N_{\mathrm{p}}\left(\ell^b\right)}\left[P_{ij'}\left(\ell\right)P_{i'j}\left(\ell\right) + P_{ii'}\left(\ell\right)P_{jj'}\left(\ell\right)\right]. \tag{2.121}$$

where we introduced the synthetic notation $\delta_{\ell^b\ell'^b} = 1$ if $\ell^b = \ell'^b$ or $\delta_{\ell^b\ell'^b} = 0$ if $\ell^b \neq \ell'^b$. As for the component of the covariance associated to the trispectrum, there are no further steps to take and the exact calculation requires to average the trispectrum over the bins ℓ^b and ℓ'^b

$$\mathrm{Cov}\left[P_{ij}\left(\ell\right), P_{i'j'}\left(\ell'\right)\right]_{\mathrm{NGins}} = \frac{1}{N_{\mathrm{p}}\left(\ell^b\right)N_{\mathrm{p}}\left(\ell'^b\right)\Omega_{\mathrm{sky}}}\sum_{\bar{\boldsymbol{\ell}},\bar{\boldsymbol{\ell}}'}\left[T_{iji'j'}\left(\bar{\boldsymbol{\ell}}, -\bar{\boldsymbol{\ell}}, \bar{\boldsymbol{\ell}}', \bar{\boldsymbol{\ell}}'\right)\right]\Delta^2_{\ell^b}\left(\bar{\boldsymbol{\ell}}\right)\Delta^2_{\ell'^b}(\bar{\boldsymbol{\ell}}'). \tag{2.122}$$

The term (2.122) correspond to the second one identified in Eq. (2.99): being sourced by the connected trispectrum, it would disappear in case of a purely Gaussian field. Plus, it simply involves correlations between modes which are entirely observed. In Chapter 3, Sec. 3.1, we will assume that the trispectrum can be we approximated by its 1-halo component at the scales of interest of weak lensing. This will allow to simplify Eq. (2.122) thanks to its dependence just on the modules of the multipoles involved. Further, the assumption of slowly varying polyspectra within the bins will allow to simplify the calculation and compute the trispectrum on the central value of the considered bins.

2.2.4 Binned tomographic bispectrum covariance matrix

Similarly to Eq. (2.117), the covariance for the BTB estimator defined in Eq. (2.108) is

$$\mathrm{Cov}\Big[\hat{B}_{ijk}\big(\ell^b_1, \ell^b_2, \ell^b_3\big), \hat{B}_{i'j'k'}(\ell'^b_1, \ell'^b_2, \ell'^b_3)\Big] \equiv \\ \langle\hat{B}_{ijk}\big(\ell^b_1, \ell^b_2, \ell^b_3\big)\hat{B}_{i'j'k'}(\ell'^b_1, \ell'^b_2, \ell'^b_3)\rangle - \langle\hat{B}_{ijk}\big(\ell^b_1, \ell^b_2, \ell^b_3\big)\rangle\langle\hat{B}_{i'j'k'}(\ell'^b_1, \ell'^b_2, \ell'^b_3)\rangle \tag{2.123}$$

and from the definitions of the estimator (2.108)

$$\mathrm{Cov}\Big[\hat{B}_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big),\hat{B}_{i'j'k'}\big(\ell_1'^b,\ell_2'^b,\ell_3'^b\big)\Big] = \sum_{\bar{\boldsymbol{\ell}}_1,\bar{\boldsymbol{\ell}}_2,\bar{\boldsymbol{\ell}}_3}\sum_{\bar{\boldsymbol{\ell}}_1',\bar{\boldsymbol{\ell}}_2',\bar{\boldsymbol{\ell}}_3'}\big\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\kappa^{(i')}_{\bar{\boldsymbol{\ell}}_1'}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}_2'}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}_3'}\big\rangle\frac{\Delta^3_{\ell_1^b,\ell_2^b,\ell_3^b}\left(\bar{\boldsymbol{\ell}}_1,\bar{\boldsymbol{\ell}}_2,\bar{\boldsymbol{\ell}}_3\right)\Delta^3_{\ell_1'^b,\ell_2'^b,\ell_3'^b}\left(\bar{\boldsymbol{\ell}}_1',\bar{\boldsymbol{\ell}}_2',\bar{\boldsymbol{\ell}}_3'\right)}{\Omega^2_{\rm sky}N_{\rm tri.}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)N_{\rm tri.}\left(\ell_1'^b,\ell_2'^b,\ell_3'^b\right)} - B_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big)B_{i'j'k'}\big(\ell_1^b,\ell_2^b,\ell_3^b\big). \quad (2.124)$$

Once again, we can split the 6-point correlation function appearing in Eq. (2.124) in its non-vanishing connected components. For the bispectrum the computation can become easily challenging. As a matter of fact, the 6-point correlation function can be split into the sum of products like (2-×2-×2-), (3-×3-), (2-×4-)connected correlation functions plus the related 6-point connected component. We will inspect these terms one by one, the first leading to the Gaussian component of the covariance.

Connected 2-point correlation function × Connected 2-point correlations function × Connected 2-point correlation function (Gauss). Within this computation, we will neglect terms involving correlations between scales of the field belonging to the same external binned triangular configuration. As a matter of fact, given Eq. (2.116), these terms would lead to a degenerate triangular configurations. For example if we were to calculate $\big\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_i}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_j}\big\rangle_c$:

$$\big\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_i}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_j}\big\rangle_c \propto \delta^{\rm K}_{\bar{\boldsymbol{\ell}}_i+\bar{\boldsymbol{\ell}}_j}P_{ij'}\left(\bar{\ell}_i\right) \quad \rightarrow \quad \bar{\boldsymbol{\ell}}_i = -\bar{\boldsymbol{\ell}}_j \quad (2.125)$$

$$\Delta^3_{\ell_i^b,\ell_j^b,\ell_k^b}\left(\bar{\boldsymbol{\ell}}_i,\bar{\boldsymbol{\ell}}_j,\bar{\boldsymbol{\ell}}_k\right) \quad \rightarrow \quad \bar{\boldsymbol{\ell}}_i+\bar{\boldsymbol{\ell}}_j+\bar{\boldsymbol{\ell}}_k = 0 \quad (2.126)$$

implying that $\bar{\boldsymbol{\ell}}_k = 0$. Therefore we expect 3! terms, namely

$$\begin{aligned}&\big\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(i')}_{\bar{\boldsymbol{\ell}}_1'}\big\rangle_c\Big[\big\langle\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}_2'}\big\rangle_c\big\langle\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}_3'}\big\rangle_c + \big\langle\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}_3'}\big\rangle_c\big\langle\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}_2'}\big\rangle_c\Big] + \\ &\big\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}_2'}\big\rangle_c\Big[\big\langle\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(i')}_{\bar{\boldsymbol{\ell}}_1'}\big\rangle_c\big\langle\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}_3'}\big\rangle_c + \big\langle\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}_3'}\big\rangle_c\big\langle\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\kappa^{(i')}_{\bar{\boldsymbol{\ell}}_1'}\big\rangle_c\Big] + \\ &\big\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}_3'}\big\rangle_c\Big[\big\langle\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(i')}_{\bar{\boldsymbol{\ell}}_1'}\big\rangle_c\big\langle\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}_2'}\big\rangle_c + \big\langle\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}_2'}\big\rangle_c\big\langle\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\kappa^{(i')}_{\bar{\boldsymbol{\ell}}_1'}\big\rangle_c\Big] \quad (2.127)\end{aligned}$$

In a similar fashion as for Eq. (2.119), we can derive the *PPP* component of the covariance. Specifically, we can apply the definition (2.116) for the power spectrum to the terms above. As an exemplification, the second term will lead to

$$\big\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(i')}_{\bar{\boldsymbol{\ell}}_1'}\big\rangle_c\big\langle\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}_3'}\big\rangle_c\big\langle\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}_2'}\big\rangle_c = \Omega^3_{\rm sky}\delta^{\rm K}_{\boldsymbol{\ell}_1+\boldsymbol{\ell}_1'}P_{ii'}(\boldsymbol{\ell}_1')\delta^{\rm K}_{\boldsymbol{\ell}_2+\boldsymbol{\ell}_3'}P_{jk'}(\boldsymbol{\ell}_2')\delta^{\rm K}_{\boldsymbol{\ell}_3+\boldsymbol{\ell}_2'}P_{kj'}(\boldsymbol{\ell}_3'). \quad (2.128)$$

The sum (e.g.) over $\{\boldsymbol{\ell}_1^{'},\boldsymbol{\ell}_2^{'},\boldsymbol{\ell}_3^{'}\}$ in Eq. (2.124) can be removed thanks to the constraints imposed by Kronecker delta functions. Furthermore, under the assumption of slowing varying polyspectra within the mul-

tipole bins, we factorise the power spectra out from the remaining sum over $\{\ell_1, \ell_2, \ell_3\}$ and we evaluate them on the central value of the bins. Finally, we are left with terms like

$$\mathrm{Cov}\left[\hat{B}_{ijk}\left(\ell_1^b, \ell_2^b, \ell_3^b\right), \hat{B}_{i'j'k'}(\ell_1'^b, \ell_2'^b, \ell_3'^b)\right]_{\mathrm{Gauss}} = \\ \cdots \sim \left[\sum_{\boldsymbol{\ell}_1, \boldsymbol{\ell}_2, \boldsymbol{\ell}_3} \Delta^3_{\ell_1^b, \ell_2^b, \ell_3^b}\left(\bar{\boldsymbol{\ell}}_1, \bar{\boldsymbol{\ell}}_2, \bar{\boldsymbol{\ell}}_3\right) \Delta^3_{\ell_1'^b, \ell_2'^b, \ell_3'^b}\left(-\bar{\boldsymbol{\ell}}_{\sigma(1)}, -\bar{\boldsymbol{\ell}}_{\sigma(2)}, -\bar{\boldsymbol{\ell}}_{\sigma(3)}\right)\right] \ldots \quad (2.129)$$

σ begin one possible permutation of the multipole indices. The terms as the one highlighted in Eq. (2.129) will result in symmetry factors being equal to 1 or 0 according to the configuration request for the specific element of the covariance. Via a similar reasoning as in the power spectrum case, we can easily find that the factor corresponding to the generic case displayed in Eq. (2.129) is

$$\mathrm{Cov}\left[\hat{B}_{ijk}\left(\ell_1^b, \ell_2^b, \ell_3^b\right), \hat{B}_{i'j'k'}(\ell_1'^b, \ell_2'^b, \ell_3'^b)\right]_{\mathrm{Gauss}} = \\ \cdots \sim N_{\mathrm{tri.}}\left(\ell_1^b, \ell_2^b, \ell_3^b\right) \delta_{l_1^b l'^b_{\sigma^{-1}(1)}} \delta_{l_2^b l'^b_{\sigma^{-1}(2)}} \delta_{l_3^b l'^b_{\sigma^{-1}(3)}} \ldots . \quad (2.130)$$

By iterating the same procedure to all the elements in Eq. (2.127), the final expression for this part of the covariance matrix reads

$$\mathrm{Cov}\left[\hat{B}_{ijk}\left(\ell_1^b, \ell_2^b, \ell_3^b\right), \hat{B}_{i'j'k'}\left(\ell_1'^b, \ell_2'^b, \ell_3'^b\right)\right]_{\mathrm{Gauss}} = \quad (2.131)$$

$$\frac{\Omega_{\mathrm{sky}}}{N_{\mathrm{tri.}}\left(\ell_1^b, \ell_2^b, \ell_3^b\right)}\Big\{P_{ii'}\left(\ell_1\right)\delta_{\ell_1\ell_1'}\left[P_{jj'}\left(\ell_2\right)P_{kk'}\left(\ell_3\right)\delta_{\ell_2\ell_2'}\delta_{\ell_3\ell_3'} + P_{jk'}\left(\ell_2\right)P_{kj'}\left(\ell_3\right)\delta_{\ell_2\ell_3'}\delta_{\ell_3\ell_2'}\right] \\ + 2 \text{ terms obtained from } \left(i' \leftrightarrow j', \ell_1' \leftrightarrow \ell_2'\right) \\ + 2 \text{ terms obtained from } \left(i' \leftrightarrow k', \ell_1' \leftrightarrow \ell_3'\right)\Big\}. \quad (2.132)$$

Connected 3-point correlation function × *Connected 3-point correlations function* (NGins,BB term). In this case, the expected number of terms will be given by the un-ordered partitions of 6 elements into 2 classes of 3 elements each[3], specifically 10.

$$\left\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\right\rangle_c\left\langle\kappa^{(i')}_{\bar{\boldsymbol{\ell}}'_1}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}'_2}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}'_3}\right\rangle_c + \left\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(i')}_{\bar{\boldsymbol{\ell}}'_1}\right\rangle_c\left\langle\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}'_2}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}'_3}\right\rangle_c + \\ \left\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}'_2}\right\rangle_c\left\langle\kappa^{(i')}_{\bar{\boldsymbol{\ell}}'_1}\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}'_3}\right\rangle_c + \left\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}'_3}\right\rangle_c\left\langle\kappa^{(i')}_{\bar{\boldsymbol{\ell}}'_1}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}'_2}\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\right\rangle_c + \\ \left\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(i')}_{\bar{\boldsymbol{\ell}}'_1}\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\right\rangle_c\left\langle\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}'_2}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}'_3}\right\rangle_c + \left\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}'_2}\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\right\rangle_c\left\langle\kappa^{(i')}_{\bar{\boldsymbol{\ell}}'_1}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}'_3}\right\rangle_c + \\ \left\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}'_3}\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\right\rangle_c\left\langle\kappa^{(i')}_{\bar{\boldsymbol{\ell}}'_1}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}'_2}\right\rangle_c + \left\langle\kappa^{(k')}_{\bar{\boldsymbol{\ell}}'_3}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\right\rangle_c\left\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(i')}_{\bar{\boldsymbol{\ell}}'_1}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}'_2}\right\rangle_c + \\ \left\langle\kappa^{(j')}_{\bar{\boldsymbol{\ell}}'_2}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\right\rangle_c\left\langle\kappa^{(i')}_{\bar{\boldsymbol{\ell}}'_1}\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}'_3}\right\rangle_c + \left\langle\kappa^{(i')}_{\bar{\boldsymbol{\ell}}'_1}\kappa^{(j)}_{\bar{\boldsymbol{\ell}}_2}\kappa^{(k)}_{\bar{\boldsymbol{\ell}}_3}\right\rangle_c\left\langle\kappa^{(i)}_{\bar{\boldsymbol{\ell}}_1}\kappa^{(j')}_{\bar{\boldsymbol{\ell}}'_2}\kappa^{(k')}_{\bar{\boldsymbol{\ell}}'_3}\right\rangle_c. \quad (2.133)$$

[3] A set of n elements can be partitioned into k un-ordered subsets of r elements each ($kr = n$) in $N = n!/(k!(r!)^k)$ ways.

To inspect the calculation in details, we plug Eq. (2.133) into Eq. (2.124) and consider the second (first non trivial) element of the sum

$$\sum_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}\sum_{\bar{\ell}'_1,\bar{\ell}'_2,\bar{\ell}'_3}\left\langle\kappa^{(i)}_{\bar{\ell}_1}\kappa^{(j)}_{\bar{\ell}_2}\kappa^{(i')}_{\bar{\ell}'_1}\right\rangle_c\left\langle\kappa^{(k)}_{\bar{\ell}_3}\kappa^{(j')}_{\bar{\ell}'_2}\kappa^{(k')}_{\bar{\ell}'_3}\right\rangle_c = \tag{2.134}$$

$$\Omega^2_{\rm sky}\sum_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}\sum_{\bar{\ell}'_1,\bar{\ell}'_2,\bar{\ell}'_3}B_{ijk}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}'_1\right)B_{i'j'k'}\left(\bar{\ell}_3,\bar{\ell}'_2,\bar{\ell}'_3\right)\delta^{\rm K}_{\bar{\ell}_1+\bar{\ell}_2+\bar{\ell}'_1}\delta^{\rm K}_{\bar{\ell}_3+\bar{\ell}'_2+\bar{\ell}'_3}\times$$

$$\Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\Delta^3_{\ell'^b_1,\ell'^b_2,\ell'^b_3}\left(\bar{\ell}'_1,\bar{\ell}'_2,\bar{\ell}'_3\right)\approx \tag{2.135}$$

$$\Omega^2_{\rm sky}\sum_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}B_{ijk}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\left[\sum_{\bar{\ell}'_2,\bar{\ell}'_3}B_{i'j'k'}\left(\bar{\ell}_3,\bar{\ell}'_2,\bar{\ell}'_3\right)\Delta^3_{\ell'^b_1,\ell'^b_2,\ell'^b_3}\left(\bar{\ell}_3,\bar{\ell}'_2,\bar{\ell}'_3\right)\right] = \tag{2.136}$$

$$\Omega^2_{\rm sky}\delta_{\ell'_1,\ell_3}B_{ijk}\left(\ell_1,\ell_2,\ell_3\right)B_{ijk}\left(\ell'_1,\ell'_2,\ell'_3\right)\sum_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}\Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\sum_{\bar{\ell}'_2,\bar{\ell}'_3}\Delta^3_{\ell'^b_1,\ell'^b_2,\ell'^b_3}\left(\bar{\ell}_3,\bar{\ell}'_2,\bar{\ell}'_3\right) = \tag{2.137}$$

$$\Omega^2_{\rm sky}\delta_{\ell'_1,\ell_3}B_{ijk}\left(\ell_1,\ell_2,\ell_3\right)B_{ijk}\left(\ell'_1,\ell'_2,\ell'_3\right)\sum_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3}\Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)\left[\frac{2\Delta\ell'^b_2\ell'_2\Delta\varphi_{32'}}{\left(2\pi/\Theta_{\rm sky}\right)^2}\right] = \tag{2.138}$$

$$\Omega^2_{\rm sky}\delta_{\ell'_1,\ell_3}N_{\rm tri.}\left(\ell^b_1,\ell^b_2,\ell^b_3\right)N_{\rm tri.}(\ell'^b_1,\ell'^b_2,\ell'^b_3)B_{ijk}\left(\ell_1,\ell_2,\ell_3\right)B_{ijk}\left(\ell'_1,\ell'_2,\ell'_3\right)\frac{2\pi}{\Omega_{\rm sky}\ell'_1\Delta\ell'^b_1}. \tag{2.139}$$

In the second line we removed the sum over $\bar{\ell}'_1$, the several geometrical constrains imposing $\bar{\ell}'_1 = \bar{\ell}_3$. Further, given that $\bar{\ell}_3$ appears both in $\Delta^3_{\ell'^b_1,\ell'^b_2,\ell'^b_3}(\bar{\ell}_3,\bar{\ell}'_2,\bar{\ell}'_3)$ and in $\Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3\right)$, the only non-vanishing contribution to the sum comes from the configuration with $\ell'^b_1 = \ell^b_3$. As usual, we also assumed that the bispectra do not very that much within the width of the bin and we factorised them out. In the third and in the fourth line we counted the number of triangular configurations allowed within the bins' width. Similarly to the derivation of Eq. (2.113), it is now just matter of combining the allowed number of multipoles within the bins. However, the first multipole ℓ'_1 is fixed to the value ℓ_3 then we do not have to account for the associated term (the first parenthesis in Eq. (2.113)) in the calculation. By applying the outlined strategy to all the terms in Eq. (2.133), we find the *BB* component of the non-Gaussian covariance

$$\begin{aligned}
&{\rm Cov}\left[\hat{B}_{ijk}\left(\ell^b_1,\ell^b_2,\ell^b_3\right),\hat{B}_{i'j'k'}\left(\ell'^b_1,\ell'^b_2,\ell'^b_3\right)\right]_{\rm NGins,BB} = \\
&\quad\frac{2\pi}{\Omega_{\rm sky}}\Bigg\{\frac{1}{\ell_1\Delta\ell_1}\Big[B_{i'jk}\left(\ell'_1,\ell_2,\ell_3\right)B_{ij'k'}\left(\ell_1,\ell'_2,\ell'_3\right)\delta_{\ell_1\ell'_1}+ \\
&\qquad\qquad B_{j'jk}\left(\ell'_2,\ell_2,\ell_3\right)B_{i'ik'}\left(\ell'_1,\ell_1,\ell'_3\right)\delta_{\ell_1\ell'_2}+B_{k'jk}\left(\ell'_3,\ell_2,\ell_3\right)B_{i'j'i}\left(\ell'_1,\ell'_2,\ell_1\right)\delta_{\ell_1\ell'_3}\Big] \\
&\qquad+3\text{ terms obtained from perm. of }\left(i\leftrightarrow j,\ell_1\leftrightarrow\ell_2\right) \\
&\qquad+3\text{ terms obtained from perm. of }\left(i\leftrightarrow k,\ell_1\leftrightarrow\ell_3\right)\Bigg\}.
\end{aligned} \tag{2.140}$$

Connected 2-point correlation function × *Connected 4-point correlations function* (NGins,PT term). Similarly to the (2-×2-) case, the conditions given in Eq. (2.125) and in Eq. (2.126) imply that the only non-vanishing contributions to the covariance are given by terms where the connected 2-point correlation function is computed for scales of the field not belonging to the same binned triangle. Therefore we expect $3^2 = 9$ terms, namely

$$\begin{aligned}
&\langle \kappa^{(i)}_{\bar{\ell}_1} \kappa^{(i')}_{\bar{\ell}'_1} \rangle_c \langle \kappa^{(j)}_{\bar{\ell}_2} \kappa^{(k)}_{\bar{\ell}_3} \kappa^{(j')}_{\bar{\ell}'_2} \kappa^{(k')}_{\bar{\ell}'_3} \rangle_c + \langle \kappa^{(i)}_{\bar{\ell}_1} \kappa^{(j')}_{\bar{\ell}'_2} \rangle_c \langle \kappa^{(j)}_{\bar{\ell}_2} \kappa^{(k)}_{\bar{\ell}_3} \kappa^{(i')}_{\bar{\ell}'_1} \kappa^{(k')}_{\bar{\ell}'_3} \rangle_c + \langle \kappa^{(i)}_{\bar{\ell}_1} \kappa^{(k')}_{\bar{\ell}'_3} \rangle_c \langle \kappa^{(j)}_{\bar{\ell}_2} \kappa^{(k)}_{\bar{\ell}_3} \kappa^{(i')}_{\bar{\ell}'_1} \kappa^{(j')}_{\bar{\ell}'_2} \rangle_c + \\
&\langle \kappa^{(j)}_{\bar{\ell}_2} \kappa^{(i')}_{\bar{\ell}'_1} \rangle_c \langle \kappa^{(i)}_{\bar{\ell}_1} \kappa^{(k)}_{\bar{\ell}_3} \kappa^{(j')}_{\bar{\ell}'_2} \kappa^{(k')}_{\bar{\ell}'_3} \rangle_c + \langle \kappa^{(j)}_{\bar{\ell}_2} \kappa^{(j')}_{\bar{\ell}'_2} \rangle_c \langle \kappa^{(i)}_{\bar{\ell}_1} \kappa^{(k)}_{\bar{\ell}_3} \kappa^{(i')}_{\bar{\ell}'_1} \kappa^{(k')}_{\bar{\ell}'_3} \rangle_c + \langle \kappa^{(j)}_{\bar{\ell}_2} \kappa^{(k')}_{\bar{\ell}'_3} \rangle_c \langle \kappa^{(i)}_{\bar{\ell}_1} \kappa^{(k)}_{\bar{\ell}_3} \kappa^{(i')}_{\bar{\ell}'_1} \kappa^{(j')}_{\bar{\ell}'_2} \rangle_c + \\
&\langle \kappa^{(k)}_{\bar{\ell}_3} \kappa^{(i')}_{\bar{\ell}'_1} \rangle_c \langle \kappa^{(i)}_{\bar{\ell}_1} \kappa^{(j)}_{\bar{\ell}_2} \kappa^{(j')}_{\bar{\ell}'_2} \kappa^{(k')}_{\bar{\ell}'_3} \rangle_c + \langle \kappa^{(k)}_{\bar{\ell}_3} \kappa^{(j')}_{\bar{\ell}'_2} \rangle_c \langle \kappa^{(i)}_{\bar{\ell}_1} \kappa^{(j)}_{\bar{\ell}_2} \kappa^{(i')}_{\bar{\ell}'_1} \kappa^{(k')}_{\bar{\ell}'_3} \rangle_c + \langle \kappa^{(k)}_{\bar{\ell}_3} \kappa^{(k')}_{\bar{\ell}'_3} \rangle_c \langle \kappa^{(i)}_{\bar{\ell}_1} \kappa^{(j)}_{\bar{\ell}_2} \kappa^{(i')}_{\bar{\ell}'_1} \kappa^{(j')}_{\bar{\ell}'_2} \rangle_c
\end{aligned} \tag{2.141}$$

As done for the *BB* component, we will detail the calculation for one the above terms (e.g. the first one):

$$\begin{aligned}
&\sum_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3} \sum_{\bar{\ell}'_1,\bar{\ell}'_2,\bar{\ell}'_3} \langle \kappa^{(i)}_{\bar{\ell}_1} \kappa^{(i')}_{\bar{\ell}'_1} \rangle_c \langle \kappa^{(j)}_{\bar{\ell}_2} \kappa^{(k)}_{\bar{\ell}_3} \kappa^{(j')}_{\bar{\ell}'_2} \kappa^{(k')}_{\bar{\ell}'_3} \rangle_c = \Omega^2_{\rm sky} \sum_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3} \sum_{\bar{\ell}'_1,\bar{\ell}'_2,\bar{\ell}'_3} P_{ii'}\left(\bar{\ell}_1\right) T_{jkj'k'}\left(\bar{\ell}_2, \bar{\ell}_3, \bar{\ell}'_2, \bar{\ell}'_3\right) \times \\
&\qquad \delta^{\rm K}_{\bar{\ell}_1+\bar{\ell}'_1} \delta^{\rm K}_{\bar{\ell}_2+\bar{\ell}_3+\bar{\ell}'_2+\bar{\ell}'_3} \Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3\right) \Delta^3_{\ell'^b_1,\ell'^b_2,\ell'^b_3}(\bar{\ell}'_1, \bar{\ell}'_2, \bar{\ell}'_3) =
\end{aligned} \tag{2.142}$$

$$\begin{aligned}
&\Omega^2_{\rm sky} \delta_{\ell'^b_1,\ell^b_1} \sum_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3} \sum_{\bar{\ell}'_2,\bar{\ell}'_3}^{\bar{\ell}'_1=-\bar{\ell}_1} P_{ii'}\left(\bar{\ell}_1\right) T_{jkj'k'}(\bar{\ell}_2, \bar{\ell}_3, \bar{\ell}'_2, \bar{\ell}'_3) \times \\
&\qquad \delta^{\rm K}_{\bar{\ell}_2+\bar{\ell}_3+\bar{\ell}'_2+\bar{\ell}'_3} \Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3\right) \Delta^3_{\ell^b_1,\ell'^b_2,\ell'^b_3}(-\bar{\ell}_1, \bar{\ell}'_2, \bar{\ell}'_3) \approx
\end{aligned} \tag{2.143}$$

$$\begin{aligned}
&\Omega^2_{\rm sky} \delta_{\ell'^b_1,\ell^b_1} P_{ii'}\left(\ell_1\right) T_{jkj'k'}\left(\ell_2, \ell_3, \ell'_2, \ell'_3\right) \times \\
&\qquad \sum_{\bar{\ell}_1,\bar{\ell}_2,\bar{\ell}_3} \sum_{\bar{\ell}'_2,\bar{\ell}'_3}^{\bar{\ell}'_1=-\bar{\ell}_1} \Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\bar{\ell}_1, \bar{\ell}_2, \bar{\ell}_3\right) \Delta^3_{\ell'^b_1,\ell'^b_2,\ell'^b_3}(-\bar{\ell}_1, \bar{\ell}'_2, \bar{\ell}'_3) =
\end{aligned} \tag{2.144}$$

$$\Omega^2_{\rm sky} \delta_{\ell'^b_1,\ell^b_1} N_{\rm tri.}\left(\ell^b_1, \ell^b_2, \ell^b_3\right) N_{\rm tri.}\left(\ell'^b_1, \ell'^b_2, \ell'^b_3\right) P_{ii'}\left(\ell_1\right) T_{jkj'k'}\left(\ell_2, \ell_3, \ell'_2, \ell'_3\right) \frac{2\pi}{\Omega_{\rm sky} \ell'_1 \Delta\ell'^b_1} \tag{2.145}$$

where: in the second line, the relation $\bar{\ell}'_1 = -\bar{\ell}_1$ implies $\ell'^b_1 = \ell^b_1$; in the third line 1) we removed $\delta^{\rm K}_{\bar{\ell}_2+\bar{\ell}_3+\bar{\ell}'_2+\bar{\ell}'_3}$, since its effect is already assured by the left geometrical factors and 2) we assumed the trispectrum to depend only on the modules of the multipole (1-halo term, see Chapter 3 for more details), besides the usual assumption of slow variation within the bins. The final line of the calculation is exactly the same

as in Eq. (2.134). Finally, the *PT* part of the covariance is

$$\begin{aligned}
&\mathrm{Cov}\Big[\hat{B}_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big),\hat{B}_{i'j'k'}\left(\ell_1'^b,\ell_2'^b,\ell_3'^b\right)\Big]_{\mathrm{NGins,PT}} = \\
&\quad\frac{2\pi}{\Omega_{\mathrm{sky}}}\Bigg\{\frac{1}{\ell_1\Delta\ell_1}\Big[P_{ii'}\left(\ell_1\right)T_{jkj'k'}\left(\ell_2,\ell_3,\ell_2',\ell_3'\right)\delta_{\ell_1\ell_1'}+ \\
&\qquad\qquad P_{ij'}\left(\ell_1\right)T_{jki'k'}\left(\ell_2,\ell_3,\ell_1',\ell_3'\right)\delta_{\ell_1\ell_2'}+P_{ik'}\left(\ell_1\right)T_{jki'j'}\left(\ell_2,\ell_3,\ell_1',\ell_2'\right)\delta_{\ell_1\ell_3'}\Big] \\
&\qquad+\ 3\text{ terms obtained from perm. of }\left(i\leftrightarrow j,\ell_1\leftrightarrow\ell_2\right) \\
&\qquad+\ 3\text{ terms obtained from perm. of }\left(i\leftrightarrow k,\ell_1\leftrightarrow\ell_3\right)\Bigg\}
\end{aligned}\tag{2.146}$$

Connected 6-point correlation function (6P term) The derivation of this final component of the covariance follows the same idea as for the trispectrum term in Eq. (2.122). This component is in particular irreducible, being already itself a connected correlation function, and the covariance will just be equal to the average of the correlations over the bins involved in the configuration, i.e.

$$\begin{aligned}
&\mathrm{Cov}\Big[\hat{B}_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big),\hat{B}_{i'j'k'}\big(\ell_1'^b,\ell_2'^b,\ell_3'^b\big)\Big]_{\mathrm{NG,6P}} = \\
&\qquad\sum_{\bar{\boldsymbol{\ell}}_1,\bar{\boldsymbol{\ell}}_2,\bar{\boldsymbol{\ell}}_3}\sum_{\bar{\boldsymbol{\ell}}_1',\bar{\boldsymbol{\ell}}_2',\bar{\boldsymbol{\ell}}_3'}P_{ijki'j'k'}\left(\bar{\boldsymbol{\ell}}_1,\bar{\boldsymbol{\ell}}_2,\bar{\boldsymbol{\ell}}_3,\bar{\boldsymbol{\ell}}_1',\bar{\boldsymbol{\ell}}_2',\bar{\boldsymbol{\ell}}_3'\right)\frac{\Delta^3_{\ell_1^b,\ell_2^b,\ell_3^b}\left(\bar{\boldsymbol{\ell}}_1,\bar{\boldsymbol{\ell}}_2,\bar{\boldsymbol{\ell}}_3\right)\Delta^3_{\ell_1'^b,\ell_2'^b,\ell_3'^b}(\bar{\boldsymbol{\ell}}_1',\bar{\boldsymbol{\ell}}_2',\bar{\boldsymbol{\ell}}_3')}{\Omega_{\mathrm{sky}}\,N_{\mathrm{tri.}}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)N_{\mathrm{tri.}}\left(\ell_1'^b,\ell_2'^b,\ell_3'^b\right)}.
\end{aligned}\tag{2.147}$$

Also in this case, we will later assume that the 6-point correlation function can be we approximated by its 1-halo component at the scales of interest for weak lensing, allowing a further simplification of this expression along with its evaluation on the central values of the involved bins.

Covariance between the BTP and the BTB The two observables are not independent since they probe the same underlying physics. It is then important to calculate the cross-covariance terms between the two and specifically among all the possible configurations available. The definition of their covariance naturally follows from the defintion in Eq. (2.92)

$$\mathrm{Cov}\Big[\hat{P}_{ij}\left(\ell^b\right),\hat{B}_{i'j'k'}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)\Big]=\langle\hat{P}_{ij}\left(\ell^b\right)\hat{B}_{i'j'k'}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)\rangle-\langle\hat{P}_{ij}\left(\ell^b\right)\rangle\langle\hat{B}_{i'j'k'}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)\rangle\tag{2.148}$$

In the same exact way as we proved all the previous contributions, the 5-point correlation function which derives from Eq. (2.148) can be split in the sum of terms involving the product of $(2-\times 3-)$ point connected correlation functions and the 5-point connected one. Without repeating all the calculations (which follows exactly the same

ideas outline in the previous derivations), the final result is

$$\mathrm{Cov}\Big[\hat{P}_{ij}\left(\ell^b\right), \hat{B}_{i'j'k'}\left(\ell_1^b, \ell_2^b, \ell_3^b\right)\Big] \tag{2.149}$$

$$\begin{aligned}
&= \frac{2\pi}{\Omega_{\rm sky}}\Big\{\frac{1}{\ell_1\Delta\ell_1}\left[P_{i'j}\left(\ell\right)B_{ij'k'}\left(\ell,\ell_2,\ell_3\right)\delta_{\ell\ell_1} + P_{ii'}\left(\ell_1\right)B_{jj'k'}\left(\ell,\ell_2,\ell_3\right)\delta_{\ell\ell_1}\right] + \\
&\qquad + 2 \text{ terms obtained from perm. of } \left(i'\leftrightarrow j', \ell_1\leftrightarrow\ell_2\right) + \\
&\qquad + 2 \text{ terms obtained from perm. of } \left(i'\leftrightarrow k', \ell_1\leftrightarrow\ell_3\right)\Big\} + \\
&+ \frac{1}{\Omega_{\rm sky}\, N_{\rm p}\left(\ell^b\right) N_{\rm tri.}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)}\sum_{\boldsymbol{\ell}'}\sum_{\boldsymbol{\ell}_1',\boldsymbol{\ell}_2',\boldsymbol{\ell}_3'} P_{iji'j'k'}\left(\bar{\boldsymbol{\ell}}',\bar{\boldsymbol{\ell}}',\bar{\boldsymbol{\ell}}_1,\bar{\boldsymbol{\ell}}_2,\bar{\boldsymbol{\ell}}_3\right)\Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right)\Delta^3_{\ell_1^b,\ell_2^b,\ell_3^b}\left(\boldsymbol{\ell}_1',\boldsymbol{\ell}_2',\boldsymbol{\ell}_3'\right),
\end{aligned} \tag{2.150}$$

where the usual bin-average appears for the highest order correlation function in the expression. Again, the symmetries of the 1-halo term will allow for further simplifications. We underline that the covariance (2.149) is entirely source by non-Gaussianities in the matter field and there is no Gaussian component. In particular, describing correlation between observed modes, it will simply be the NGins term introduced in Eq. (2.101).

2.3 Super-sample covariance matrix

In the previous section we gave a detailed derivation of the covariance terms between the convergence BTP and the convergence BTB based on the binned estimators (2.107) and (2.108). In this final part of the chapter we will focus on a component of the covariance we have been mentioning since few pages back: the *super-sample covariance* (SSC). This effect can be intuitively understood as an uncertainty on the value of our observables related to the fact that we are observing just a fraction of the sky within our survey view field. For example, if we happen to be in an under-dense or in an over-dense region of the Universe, then our estimators would be biased and the value of the background matter density, as an example, will be respectively under-estimated or over-estimated. The issue is, of course, much more complicated and more complex correlations do indeed play a role. The analytical derivation of the SSC is a very recent subject of research and the level of awareness of the scientific community on it has grown intensively over the last decade. To best of our knowledge, no up-to-date reviews are available on the subject. We would like to invite the reader to see the next section as the seed of an up-coming written production aiming at a comprehensive summary of all the aspects related to it.

2.3.1 Estimators for a masked field

The root of the physical understanding of the SSC dates back to the paper of Hamilton et al. [2006] where it was studied for the three-dimensional matter power spectrum. More recent analyses have been carried out at the level of matter bisperctum by Chan et al. [2018] and Barreira [2019]. As discussed around Eq. (2.106) a finite domain of angular size $\Theta_{\rm sky}$ radian implies a minimum resolution in module $\ell_{\rm f} \approx 2\pi/\Theta_{\rm sky}$ for our momenta to be

measured. However, we also underlined that this is just an approximation while the true field is a convolution with the window function $\tilde{W}$ of the survey. Also, the expansion (2.106) assumed the periodicity of the observed signal. In particular, couplings with modes $\ell \leqslant \ell_{\rm f}$ are allowed. In the previous section, we derived our covariance matrices via the binned estimators defined in Eq. (2.107) and in Eq. (2.108) and the finite-sky condition mainly translated into an overall re-scaling of our covariance. However, no extra-correlations were produced. This is actually the crucial point: a weighted (masked) field induces correlations between small scales and large super-survey modes.

Observable estimators for masked convergence field To start with, we define more sophisticated estimators for the binned convergence polyspectra which account for finite volume effects

$$\hat{P}^{\rm w}_{ij}\left(\ell^b\right) \equiv \frac{1}{\Omega_{\rm sky}\, N_{\rm p}\left(\ell^b\right)} \sum_{\boldsymbol{\ell}'} \kappa^{{\rm w},(i)}_{\boldsymbol{\ell}'} \kappa^{{\rm w},(j)}_{-\boldsymbol{\ell}'} \Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right), \tag{2.151}$$

$$\hat{B}^{\rm w}_{ijk}\left(\ell^b_1, \ell^b_2, \ell^b_3\right) \equiv \frac{1}{\Omega_{\rm sky}\, N_{\rm tri.}\left(\ell^b_1, \ell^b_2, \ell^b_3\right)} \sum_{\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3} \kappa^{{\rm w},(i)}_{\boldsymbol{\ell}'_1} \kappa^{{\rm w},(j)}_{\boldsymbol{\ell}'_2} \kappa^{{\rm w},(k)}_{\boldsymbol{\ell}'_3} \Delta^3_{\ell^b_1, \ell^b_2, \ell^b_3}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right). \tag{2.152}$$

where the selection function $\Delta^{2,(3)}$ are the same as in Eqs.(2.107)-(2.108) and $\kappa^{{\rm w},(i)}_{\boldsymbol{\ell}}$ is the Fourier transform of the masked field $\kappa^{{\rm w},(i)}\left(\Theta\right)$ (2.102). In particular, $\kappa^{{\rm w},(i)}_{\boldsymbol{\ell}}$ will be the convergence field $\kappa^{(i)}$ convoluted with the window function of the survey, in Fourier space,

$$\kappa^{{\rm w},(i)}_{\boldsymbol{\ell}} = \int d^2\mathbf{q} \tilde{W}\left(\mathbf{q}\right) \kappa^{(i)}_{\boldsymbol{\ell}-\mathbf{q}}. \tag{2.153}$$

We can simplify the bispectrum estimator (2.152) via a change of variables

$$\boldsymbol{\ell}'_1 - \mathbf{q}_1 \to \boldsymbol{\ell}'_1, \tag{2.154}$$

$$\boldsymbol{\ell}'_2 - \mathbf{q}_2 \to \boldsymbol{\ell}'_2, \tag{2.155}$$

$$\boldsymbol{\ell}'_3 - \mathbf{q}_1 - \mathbf{q}_2 \to \boldsymbol{\ell}'_3, \tag{2.156}$$

$$\tag{2.157}$$

leading to, including Eq. (2.153),

$$\hat{B}^{\rm w}_{ijk}\left(\ell^b_1, \ell^b_2, \ell^b_3\right) = \frac{1}{\Omega_{\rm sky}\, N_{\rm tri.}\left(\ell^b_1, \ell^b_2, \ell^b_3\right)} \sum_{\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3} \int \left[\prod_{a=1}^{3} \frac{d^2\mathbf{q}_a}{\left(2\pi\right)^2} \tilde{W}\left(\mathbf{q}_a\right)\right] \times \kappa^{(i)}_{\boldsymbol{\ell}'_1} \kappa^{(j)}_{\boldsymbol{\ell}'_2} \kappa^{(k)}_{\boldsymbol{\ell}'_3 - \mathbf{q}_{123}} \Delta^3_{\ell^b_1, \ell^b_2, \ell^b_3}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right). \tag{2.158}$$

While the formal dependence of the integrand is preserved, the variable change (2.154)-(2.156) slighlty modifies the bin limits for the the varialbes $\{\boldsymbol{\ell}'_i\}_{i=1,2,3}$ to be summed over. We notice that the typical domain size of $\tilde{W}$ is $q \sim \ell_{\rm f}$, $\ell_{\rm f}$ being fundamental frequency of the survey. Therefore, the above mentioned changes are of order $\sim \ell_{\rm f}$

and will be neglected

$$|\boldsymbol{\ell}_i - \mathbf{q}_i|^2 = \ell_i^2 + q_i^2 - 2\ell_i\, q_i\, \cos\theta_i \approx \ell_i^2, \quad q_i \sim \ell_{\rm f}. \tag{2.159}$$

We remind that in our notation $q = |\mathbf{q}|$. The above estimators are in general biased. As a matter of fact,

$$\left\langle \hat{P}^{\rm w}_{ij}\left(\ell^b\right)\right\rangle = \frac{1}{\Omega_{\rm sky}\, N_{\rm p}\left(\ell^b\right)} \sum_{\boldsymbol{\ell}'} \int \left[\prod_{a=1}^{2} \frac{d^2\mathbf{q}_a}{(2\pi)^2}\, \tilde{W}(\mathbf{q}_a)\right] \left\langle \kappa^{(i)}_{\boldsymbol{\ell}'-\mathbf{q}_1} \kappa^{(j)}_{-\boldsymbol{\ell}'-\mathbf{q}_2}\right\rangle_c \Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right) \tag{2.160}$$

$$= \frac{1}{\Omega_{\rm sky}\, N_{\rm p}\left(\ell^b\right)} \sum_{\boldsymbol{\ell}'} \int \left[\prod_{a=1}^{2} \frac{d^2\mathbf{q}_a}{(2\pi)^2}\, \tilde{W}(\mathbf{q}_a)\right] (2\pi)^2\, \delta_{\rm D}\left(\mathbf{q}_1 + \mathbf{q}_2\right) P_{ij}\left(|\boldsymbol{\ell}' - \mathbf{q}_1|\right) \Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right) \tag{2.161}$$

$$= \frac{1}{\Omega_{\rm sky}\, N_{\rm p}\left(\ell^b\right)} \sum_{\boldsymbol{\ell}'} \int \frac{d^2\mathbf{q}}{(2\pi)^2}\, |\tilde{W}(\mathbf{q})|^2 P_{ij}\left(|\boldsymbol{\ell}' - \mathbf{q}|\right) \Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right) \neq P_{ij}\left(\ell\right) \tag{2.162}$$

$$\left\langle \hat{B}^{\rm w}_{ijk}\left(\ell^b_1, \ell^b_2, \ell^b_3\right)\right\rangle = \frac{1}{\Omega_{\rm sky}\, N_{\rm tri.}\left(\ell^b_1, \ell^b_2, \ell^b_3\right)} \sum_{\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3} \int \left[\prod_{a=1}^{3} \frac{d^2\mathbf{q}_a}{(2\pi)^2}\, \tilde{W}(\mathbf{q}_a)\right] (2\pi)^2$$
$$\left\langle \kappa^{(i)}_{\boldsymbol{\ell}'_1} \kappa^{(j)}_{\boldsymbol{\ell}'_2} \kappa^{(k)}_{\boldsymbol{\ell}'_3-\mathbf{q}_{123}}\right\rangle_c \Delta^3_{\ell^b_1, \ell^b_2, \ell^b_3}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right) \tag{2.163}$$

$$= \frac{1}{\Omega_{\rm sky}\, N_{\rm tri.}\left(\ell^b_1, \ell^b_2, \ell^b_3\right)} \sum_{\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3} \int \left[\prod_{a=1}^{3} \frac{d^2\mathbf{q}_a}{(2\pi)^2}\, \tilde{W}(\mathbf{q}_a)\right] (2\pi)^2 \times$$
$$B_{ijk}\left(\ell'_1, \ell'_2, |\boldsymbol{\ell}'_3 - \mathbf{q}_{123}|\right) \delta_{\rm D}\left(\boldsymbol{\ell}'_{123} - \mathbf{q}_{123}\right) \Delta^3_{\ell^b_1, \ell^b_2, \ell^b_3}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right) \tag{2.164}$$

$$\neq B_{ijk}\left(\ell_1, \ell_2, \ell_3\right), \tag{2.165}$$

where we used the polyspectra definition for a continuous Fourier transformation and the reality condition for the mask function in the third line of Eq. (2.160). Also, we introduced the useful notation $\boldsymbol{\ell}_{1\ldots n} \equiv \boldsymbol{\ell}_1 + \cdots + \boldsymbol{\ell}_n$. Therefore, if we measure the power spectrum (bispectrum) of the observed convergence field we would end up with a biased estimation of the true one, i.e. in a wrong cosmological analysis. In the general case, one would deconvolve the window function in constructing unbiased estimators (see Tegmark [2004]; Hikage et al. [2011] for a discussion at the level of matter power spectrum). Luckily, for small enough observed scales, i.e. $\ell \gg \ell_{\rm f}$, the above estimators are unbiased. We remind that the window function $\tilde{W}$ has a typical length scale of $\ell_{\rm f}$. Then, the following approximation holds on the entire integration domain

$$P_{ij}\left(\boldsymbol{\ell}' - \mathbf{q}\right) \sim P_{ij}\left(|\boldsymbol{\ell}' - \mathbf{q}|\right)$$
$$\approx P_{ij}\left(\ell'\right) + \left.\frac{{\rm d}P\left(\ell\right)}{{\rm d}\ell}\right|_{\ell'} q \underset{\rm Lead.}{\approx} P_{ij}\left(\ell'\right), \tag{2.166}$$

$$B_{ijk}\left(\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2, \boldsymbol{\ell}'_3\right) \sim B_{ijk}\left(\ell'_1, \ell'_2, |\boldsymbol{\ell}'_3 - \mathbf{q}|\right)$$
$$\approx B_{ijk}\left(\ell'_1, \ell'_2, \ell'_3\right) + \left.\frac{{\rm d}B_{ijk}\left(\ell'_1, \ell'_2, \ell\right)}{{\rm d}\ell}\right|_{\ell'_3} q \underset{\rm Lead.}{\approx} B_{ijk}\left(\ell'_1, \ell'_2, \ell'_3\right). \tag{2.167}$$

In particular, thanks to the assumed homogeneity of the Universe, we were capable to write both the power spectrum and the bispectrum in terms of the modules of the multipoles involved, then expanding around the magnitude $q \sim \ell_{\rm f} \ll \ell'$. Not depending anymore on the long mode $\mathbf{q}$, the polyspectra can be taken out from the integrations

$$\left\langle \hat{P}^{\rm w}_{ij}\left(\ell^b\right)\right\rangle = \frac{1}{\Omega_{\rm sky}\, N_{\rm p}\left(\ell^b\right)} \sum_{\boldsymbol{\ell}'} P_{ij}\left(\ell'\right) \int \frac{d^2\mathbf{q}}{\left(2\pi\right)^2} \left|\tilde{W}\left(\mathbf{q}\right)\right|^2 \Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right) = \frac{1}{N_{\rm p}\left(\ell^b\right)} \sum_{\boldsymbol{\ell}'} P_{ij}\left(\ell'\right) \Delta^2_{\ell^b}\left(\boldsymbol{\ell}'\right) \approx P_{ij}\left(\ell\right), \tag{2.168}$$

$$\begin{aligned}\left\langle \hat{B}^{\rm w}_{ijk}\left(\ell^b_1,\ell^b_2,\ell^b_3\right)\right\rangle &= \frac{1}{\Omega_{\rm sky}\, N_{\rm tri.}\left(\ell^b_1,\ell^b_2,\ell^b_3\right)} \sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2,\boldsymbol{\ell}'_3} \int \left[\prod_{a=1}^{3} \frac{d^2\mathbf{q}_a}{\left(2\pi\right)^2}\, \tilde{W}\left(\mathbf{q}_a\right)\right] \times \\ &\qquad \left\langle \kappa^{(i)}_{\boldsymbol{\ell}'_1-\mathbf{q}_1} \kappa^{(j)}_{\boldsymbol{\ell}'_2-\mathbf{q}_2} \kappa^{(k)}_{\boldsymbol{\ell}'_3-\mathbf{q}_3}\right\rangle_c \Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2,\boldsymbol{\ell}'_3\right) \qquad (2.169)\\ &= \frac{1}{\Omega_{\rm sky}\, N_{\rm tri.}\left(\ell^b_1,\ell^b_2,\ell^b_3\right)} \sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2,\boldsymbol{\ell}'_3} B_{ijk}\left(\ell'_1,\ell'_2,\ell'_3\right) \int \left[\prod_{a=1}^{3} \frac{d^2\mathbf{q}_a}{\left(2\pi\right)^2}\, \tilde{W}\left(\mathbf{q}_a\right)\right] \times \\ &\qquad \left(2\pi\right)^2 \Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2,\boldsymbol{\ell}'_3\right) \delta_{\rm D}\left(\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3\right) \qquad (2.170)\\ &= \frac{B_{ijk}\left(\ell_1,\ell_2,\ell_3\right)}{N_{\rm tri.}\left(\ell^b_1,\ell^b_2,\ell^b_3\right)} \sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2,\boldsymbol{\ell}'_3} \Delta^3_{\ell^b_1,\ell^b_2,\ell^b_3}\left(\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2,\boldsymbol{\ell}'_3\right) \approx B_{ijk}\left(\boldsymbol{\ell}_1,\boldsymbol{\ell}_2,\boldsymbol{\ell}_3\right) \qquad (2.171)\end{aligned}$$

where in both Eq. (2.168) and Eq. (2.169) we assumed slowly varying spectra within the bins and employed the following property for a binary window function

$$\Omega_{\rm sky} = \int d^2\theta\, W\left(\theta\right) = \int d^2\theta\, W^n\left(\theta\right) = \int \left[\prod_{a=1}^{n} \frac{d^2\boldsymbol{\ell}_a}{\left(2\pi\right)^2}\, \tilde{W}\left(\boldsymbol{\ell}_a\right)\right] \left(2\pi\right)^2 \delta_{\rm D}\left(\boldsymbol{\ell}_1+\cdots+\boldsymbol{\ell}_n\right) \tag{2.172}$$

for $n=2$ and $n=3$ respectively.

Covariance estimators for masked convergence field We can compute the covariance of the masked estimators (2.151)-(2.152) with the same techniques employed in Sec. 2.2.3 and in Sec. 2.2.4. Also in this case, the covariance estimation will require to split (non-connected) higher order correlation functions in the corresponding irreducible components. Without repeating all the calculations, we will simply focus on few terms with the simple goal of conveying the most important ideas necessary for the understanding of the super-sample terms of the covariance matrices. At the level of power spectrum, the total covariance is given by

$$\begin{aligned}{\rm Cov}\left[\hat{P}^{\rm w}_{ij}\left(\ell^b_1\right)\hat{P}^{\rm w}_{i'j'}\left(\ell^b_2\right)\right]_{\rm Gauss} &= \frac{1}{\Omega_{\rm sky}\, N_{\rm p}\left(\ell^b_1\right)} \frac{1}{\Omega_{\rm sky}\, N_{\rm p}\left(\ell^b_2\right)} \sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2} \int \left[\prod_{a=1}^{4} \frac{d^2\mathbf{q}_a}{\left(2\pi\right)^2}\, \tilde{W}\left(\mathbf{q}_a\right)\right] \times \\ &\qquad \left\langle \kappa^{(i)}_{\boldsymbol{\ell}'_1-\mathbf{q}_1} \kappa^{(j)}_{-\boldsymbol{\ell}'_1-\mathbf{q}_2} \kappa^{(i)}_{\boldsymbol{\ell}'_2-\mathbf{q}_3} \kappa^{(j)}_{-\boldsymbol{\ell}'_2-\mathbf{q}_4}\right\rangle \Delta^2_{\ell^b_1}(\boldsymbol{\ell}'_1)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2) - P_{ij}\left(\ell_1\right) P_{i'j'}\left(\ell_2\right) \qquad (2.173)\end{aligned}$$

where the last term is assured by the estimator being unbiased. In the same manner as for the covariance of the estimator (2.117), we will first focus on the Gaussian component of the correlator appearing in Eq. (2.173). For example, referring to one possible term (e.g. the second one) of the Wick decomposition (2.118), the calculations follows as

$$\begin{aligned}
&\mathrm{Cov}\left[\hat{P}^{\mathrm{w}}_{ij}\left(\ell^b_1\right)\hat{P}^{\mathrm{w}}_{i'j'}\left(\ell^b_2\right)\right]_{\mathrm{Gauss}} \\
&\quad= \cdots + \sim \sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2}\frac{\Delta^2_{\ell^b_1}(\boldsymbol{\ell}'_1)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2)}{\Omega^2_{\mathrm{sky}}N_{\mathrm{p}}\left(\ell^b_1\right)N_{\mathrm{p}}\left(\ell^b_2\right)}\int\left[\prod_{a=1}^{4}\frac{d^2\mathbf{q}_a}{(2\pi)^2}\,\tilde{W}(\mathbf{q}_a)\right]\left\langle\kappa^{(i)}_{\boldsymbol{\ell}'_1-\mathbf{q}_1}\kappa^{(j)}_{-\boldsymbol{\ell}'_2-\mathbf{q}_2}\right\rangle_c\left\langle\kappa^{(i')}_{\boldsymbol{\ell}'_2-\mathbf{q}_3}\kappa^{(j')}_{-\boldsymbol{\ell}'_1-\mathbf{q}_4}\right\rangle_c+\dots \qquad (2.174)\\
&\quad= \cdots + \sim \sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2}\frac{\Delta^2_{\ell^b_1}(\boldsymbol{\ell}'_1)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2)}{\Omega^2_{\mathrm{sky}}N_{\mathrm{p}}\left(\ell^b_1\right)N_{\mathrm{p}}\left(\ell^b_2\right)}\int\left[\prod_{a=1}^{4}\frac{d^2\mathbf{q}_a}{(2\pi)^2}\,\tilde{W}(\mathbf{q}_a)\right](2\pi)^4\,P_{ij}\left(|\boldsymbol{\ell}'_1-\mathbf{q}_1|\right)\,P_{i'j'}\left(|\boldsymbol{\ell}'_2-\mathbf{q}_3|\right)\times \\
&\qquad\qquad \delta_{\mathrm{D}}\left(\boldsymbol{\ell}'_1-\mathbf{q}_1-\boldsymbol{\ell}'_2-\mathbf{q}_2\right)\delta_{\mathrm{D}}\left(\boldsymbol{\ell}'_2-\mathbf{q}_3-\boldsymbol{\ell}'_1-\mathbf{q}_4\right)+\dots \qquad (2.175)\\
&\quad\approx \cdots + \sim \sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2}\frac{\Delta^2_{\ell^b_1}\left(\boldsymbol{\ell}'_1\right)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2)}{\Omega^2_{\mathrm{sky}}N_{\mathrm{p}}\left(\ell^b_1\right)N_{\mathrm{p}}\left(\ell^b_2\right)}P_{ij}\left(\ell'_1\right)\,P_{i'j'}\left(\ell'_2\right)\int\left[\prod_{a=1,3}\frac{d^2\mathbf{q}_a}{(2\pi)^2}\right](2\pi)^4\times \\
&\qquad\qquad \tilde{W}(\mathbf{q}_1)\,\tilde{W}\left(\boldsymbol{\ell}'_1-\boldsymbol{\ell}'_2-\mathbf{q}_1\right)\tilde{W}(\mathbf{q}_3)\,\tilde{W}\left(\boldsymbol{\ell}'_2-\boldsymbol{\ell}'_1-\mathbf{q}_3\right)+\dots \qquad (2.176)\\
&\quad= \cdots + \sim \sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2}\frac{\Delta^2_{\ell^b_1}\left(\boldsymbol{\ell}'_1\right)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2)}{\Omega^2_{\mathrm{sky}}N_{\mathrm{p}}\left(\ell^b_1\right)N_{\mathrm{p}}\left(\ell^b_2\right)}P_{ij}\left(\ell'_1\right)\,P_{i'j'}\left(\ell'_2\right)\tilde{W}\left(\boldsymbol{\ell}'_1-\boldsymbol{\ell}'_2\right)\tilde{W}^*\left(\boldsymbol{\ell}'_1-\boldsymbol{\ell}'_2\right)+\dots \qquad (2.177)\\
&\quad\approx \cdots + \sim P_{ij}\left(\ell_1\right)\,P_{i'j'}\left(\ell_2\right)\sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2}\frac{\Delta^2_{\ell^b_1}\left(\boldsymbol{\ell}'_1\right)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2)}{\Omega^2_{\mathrm{sky}}N_{\mathrm{p}}\left(\ell^b_1\right)N_{\mathrm{p}}\left(\ell^b_2\right)}|\tilde{W}\left(\boldsymbol{\ell}'_1-\boldsymbol{\ell}'_2\right)|^2+\dots \qquad (2.178)\\
&\qquad\qquad \approx \cdots + \sim \frac{\delta_{\ell_1\ell_2}}{N_{\mathrm{p}}\left(\ell^b_1\right)}P_{ij}\left(\ell_1\right)\,P_{i'j'}\left(\ell_2\right)+\dots \qquad (2.179)
\end{aligned}$$

where in the third line we approximated $P_{ij}\left(|\boldsymbol{\ell}'-\mathbf{q}|\right)\approx P_{ij}\left(\ell'\right)$ in the usual limit of observed multipoles $\ell\gg q\sim\ell_{\mathrm{f}}$. At line four, we employed twice the binary window function relation

$$\tilde{W}(\boldsymbol{\ell})=\int\frac{d^2\boldsymbol{\ell}'}{(2\pi)^2}\tilde{W}\left(\boldsymbol{\ell}'\right)\tilde{W}\left(\boldsymbol{\ell}-\boldsymbol{\ell}'\right)=\int\left[\prod_{a=1}^{n}\frac{d^2\mathbf{q}_a}{(2\pi)^2}\,\tilde{W}(\mathbf{q}_a)\right](2\pi)^2\,\delta_{\mathrm{D}}\left(\boldsymbol{\ell}-\mathbf{q}_{1\dots n}\right). \qquad (2.180)$$

In the fifth line, we assumed that the power spectra do slowly vary within the bin width, taking them out of the bin-average. In the sixth line, since the typical domain size of $\tilde{W}$ is of the order of $\sim\ \ell_{\mathrm{f}}$, $\tilde{W}\left(\pm(\boldsymbol{\ell}'_1-\boldsymbol{\ell}'_2)\right)$ will be non-vanishing only if $\boldsymbol{\ell}'_1$ and $\boldsymbol{\ell}'_2$ differ by a vector of magnitude smaller (or comparable to) than the fundamental frequency. Differently said, the two momenta $\boldsymbol{\ell}'_1$ and $\boldsymbol{\ell}'_1$ has to belong to the same bin, being under the assumption $\Delta\ell^b\gg\ell_{\mathrm{f}}$. We approximated the impact of the window function in the last line via $\tilde{W}\left(\pm(\boldsymbol{\ell}'_1-\boldsymbol{\ell}'_2)\right)\approx\Omega_{\mathrm{sky}}\delta_{\ell_1\ell_2}$.

As main consequence of the above calculation, this component of the covariance matrix is not affected by couplings with the super-survey modes of the order of ℓ_{f} and indeed it is the same as the Gaussian covariance (2.120) obtained from the estimator (2.107).

Moving to the non-Gaussian component as derived from the connected part of the correlator in Eq. (2.173), one has

$$\mathrm{Cov}\left[\hat{P}^{\mathrm{w}}_{ij}\left(\ell^b_1\right)\hat{P}^{\mathrm{w}}_{i'j'}\left(\ell^b_2\right)\right]_{\mathrm{NG}} = \frac{1}{\Omega^2_{\mathrm{sky}}\,N_{\mathrm{p}}\left(\ell^b_1\right)N_{\mathrm{p}}\left(\ell^b_2\right)}\sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2}\Delta^2_{\ell^b_1}\left(\boldsymbol{\ell}'_1\right)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2)\int\left[\prod_{a=1}^{4}\frac{d^2\mathbf{q}_a}{(2\pi)^2}\,\tilde{W}(\mathbf{q}_a)\right]\times$$
$$(2\pi)^2\,\delta_{\mathrm{D}}\left(\mathbf{q}_{1234}\right)T_{ijij'}\left(\boldsymbol{\ell}'_1-\mathbf{q}_1,-\boldsymbol{\ell}'_1-\mathbf{q}_2,\boldsymbol{\ell}'_2-\mathbf{q}_3,-\boldsymbol{\ell}'_2-\mathbf{q}_4\right). \quad (2.181)$$

Similarly to the steps performed at the level of bispectrum estimator, we can rephrase Eq.(2.181) by exploiting the usual limit $\ell \gg q \sim \ell_{\mathrm{f}}$ such that

$$\mathrm{Cov}\left[\hat{P}^{\mathrm{w}}_{ij}\left(\ell^b_1\right)\hat{P}^{\mathrm{w}}_{i'j'}\left(\ell^b_2\right)\right]_{\mathrm{NG}} \quad (2.182)$$
$$= \frac{1}{\Omega^2_{\mathrm{sky}}\,N_{\mathrm{p}}\left(\ell^b_1\right)N_{\mathrm{p}}\left(\ell^b_2\right)}\sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2}\Delta^2_{\ell^b_1}\left(\boldsymbol{\ell}'_1\right)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2)\int\left[\prod_{a=1}^{4}\frac{d^2\mathbf{q}_a}{(2\pi)^2}\,\tilde{W}(\mathbf{q}_a)\right]\times$$
$$(2\pi)^2\,\delta_{\mathrm{D}}\left(\mathbf{q}_{1234}\right)T_{ijij'}\left(\boldsymbol{\ell}'_1,-\boldsymbol{\ell}'_1-\mathbf{q}_{12},\boldsymbol{\ell}'_2,-\boldsymbol{\ell}'_2-\mathbf{q}_{34}\right), \quad (2.183)$$
$$= \frac{1}{\Omega^2_{\mathrm{sky}}\,N_{\mathrm{p}}\left(\ell^b_1\right)N_{\mathrm{p}}\left(\ell^b_2\right)}\sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2}\Delta^2_{\ell^b_1}\left(\boldsymbol{\ell}'_1\right)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2)\int\left[\prod_{a=1}^{3}\frac{d^2\mathbf{q}_a}{(2\pi)^2}\,\tilde{W}(\mathbf{q}_a)\right]\tilde{W}(-\mathbf{q}_{123})\times$$
$$T_{ijij'}\left(\boldsymbol{\ell}'_1,-\boldsymbol{\ell}'_1-\mathbf{q}_{12},\boldsymbol{\ell}'_2,-\boldsymbol{\ell}'_2-\mathbf{q}_{34}\right), \quad (2.184)$$
$$= \frac{1}{\Omega^2_{\mathrm{sky}}\,N_{\mathrm{p}}\left(\ell^b_1\right)N_{\mathrm{p}}\left(\ell^b_2\right)}\sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2}\Delta^2_{\ell^b_1}\left(\boldsymbol{\ell}'_1\right)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2)\int\left[\prod_{a=1,3}\frac{d^2\mathbf{q}_a}{(2\pi)^2}\,\tilde{W}(\mathbf{q}_a)\right]\int\frac{d^2\mathbf{q}}{(2\pi)^2}\tilde{W}(\mathbf{q}-\mathbf{q}_1)\,\tilde{W}(-\mathbf{q}-\mathbf{q}_3)\times$$
$$T_{ijij'}\left(\boldsymbol{\ell}'_1,-\boldsymbol{\ell}'_1-\mathbf{q}_{12},\boldsymbol{\ell}'_2,-\boldsymbol{\ell}'_2-\mathbf{q}_{34}\right), \quad (2.185)$$

where in the first line we performed the change of variables $\boldsymbol{\ell}'_1-\mathbf{q}_1\to\boldsymbol{\ell}'_1, \boldsymbol{\ell}'_2-\mathbf{q}_3\to\boldsymbol{\ell}'_2$ and we integrated over $\mathbf{q}_4$ in the second one. Finally in the last line we defined the new integration variable $\mathbf{q}\equiv\mathbf{q}_{12}$. We can simplify the non-Gaussian term (2.185) with the help of the binary window function relation (2.180)

$$\mathrm{Cov}\left[\hat{P}^{\mathrm{w}}_{ij}\left(\ell^b_1\right)\hat{P}^{\mathrm{w}}_{i'j'}\left(\ell^b_2\right)\right]_{\mathrm{NG}} = \frac{1}{\Omega^2_{\mathrm{sky}}\,N_{\mathrm{p}}\left(\ell^b_1\right)N_{\mathrm{p}}\left(\ell^b_2\right)}\sum_{\boldsymbol{\ell}'_1,\boldsymbol{\ell}'_2}\Delta^2_{\ell^b_1}\left(\boldsymbol{\ell}'_1\right)\Delta^2_{\ell^b_2}(\boldsymbol{\ell}'_2)$$
$$\int\frac{d^2\mathbf{q}}{(2\pi)^2}|\tilde{W}(\mathbf{q})|^2\,T_{ijij'}\left(\boldsymbol{\ell}'_1,-\boldsymbol{\ell}'_1+\mathbf{q},\boldsymbol{\ell}'_2,-\boldsymbol{\ell}'_2-\mathbf{q}\right). \quad (2.186)$$

At this point, Eq. (2.186) cannot be simplified any further: the window function $\tilde{W}(\mathbf{q})$ does not have a trivial approximation over the integration range and the truncation of the expansion in Eq. (2.166) and in Eq. (2.167) are not allowed in the present case. As a matter of fact, we could indeed write the trispectrum as function of the modules involved

$$T_{ijij'}\left(\boldsymbol{\ell}'_1,-\boldsymbol{\ell}'_1-\mathbf{q},\boldsymbol{\ell}'_2,-\boldsymbol{\ell}'_2-\mathbf{q}\right)\sim T_{ijij'}\left(\ell'_1,|-\boldsymbol{\ell}'_1+\mathbf{q}|,\ell'_2,|-\boldsymbol{\ell}'_2-\mathbf{q}|,|\boldsymbol{\ell}'_1-\boldsymbol{\ell}'_2+\mathbf{q}|,q\right). \quad (2.187)$$

The allowed approximation from the usual limit $\ell \gg \ell_{\rm f}$ would be

$$T_{ijij'}\left(\ell_1', |-\boldsymbol{\ell}_1'+\mathbf{q}|, \ell_2', |-\boldsymbol{\ell}_2'-\mathbf{q}|, |\boldsymbol{\ell}_1'-\boldsymbol{\ell}_2'+q|, q\right) \approx T_{ijij'}\left(\ell_1', \ell_1', \ell_2', \ell_2', |\boldsymbol{\ell}_1'-\boldsymbol{\ell}_2'|, q\right) \tag{2.188}$$

and in particular we cannot get rid of the coupling with the super-survey mode $q \leqslant \ell_{\rm f}$.

Since we are working in the framework provided by the Limber approximation, we can evaluate the trispecturm in (2.186) as the projected matter trispectrum along the line-of-sight defined in Eq. (2.89)

$$\begin{aligned} T_{ijij'}\left(\boldsymbol{\ell}_1', -\boldsymbol{\ell}_1'-\mathbf{q}, \boldsymbol{\ell}_2', -\boldsymbol{\ell}_2'-\mathbf{q}\right) &= \int_0^\infty d\omega\, \mathcal{T}\left(iji'j';\omega\right) T\left(\frac{\boldsymbol{\ell}_1'}{\omega}, \frac{-\boldsymbol{\ell}_1'+\mathbf{q}}{\omega}, \frac{\boldsymbol{\ell}_2'}{\omega}, \frac{-\boldsymbol{\ell}_2'-\mathbf{q}}{\omega}\right) \\ &\sim T\left[\mathbf{k}\left(\boldsymbol{\ell}_1',\omega\right), \mathbf{k}\left(-\boldsymbol{\ell}_1'+\mathbf{q},\omega\right), \mathbf{k}\left(\boldsymbol{\ell}_2',\omega\right), \mathbf{k}\left(-\boldsymbol{\ell}_2'-\mathbf{q},\omega\right)\right]. \end{aligned} \tag{2.189}$$

In Eq. (2.189) we implied the explicit redshift dependence of the matter trispectrum and we expressed the three-dimensional modes via the Limber dependence $\mathbf{k}(\boldsymbol{\ell},\omega) \approx (\boldsymbol{\ell}/\omega, 0)$. The three dimensional matter trispectrum can be evaluated at this point with the preferred model of non-linear clustering. The crucial point is that the trispectrum is now coupling the 2 observed modes, $\mathbf{k}_1 = \mathbf{k}\left(\boldsymbol{\ell}_1',\omega\right)$ and $\mathbf{k}_2 = \mathbf{k}\left(\boldsymbol{\ell}_2',\omega\right)$, with the "soft" mode $\varepsilon = \mathbf{k}\left(\mathbf{q},\omega\right) \lesssim 2\pi/\sqrt[3]{V}$. In our case, V is the volume slice at comoving distance ω of the projection light cone. This can be rephrased as: the mode ε induces an extra correlation between the observed modes $\mathbf{k}_1$ and $\mathbf{k}_2$. This mode is an *un-observed mode*, much larger than, or close to the size of the survey. For this reason, we dub the induced extra covariance as *super-survey covariance*.

Over the past decade, several attempts have been done in order to understand in details the role of these correlations. Hamilton et al. [2006]; Takada & Jain [2009]; Barreira et al. [2018] used Perturbation Theory (PT) to calculate the trispectrum corresponding to these peculiar configurations resulting in an additional term which sums to the standard intra-survey non-Gaussian covariance. Since PT is valid up to scales in the mildly non-linear regime, they were not capable to describe entirely the effect of super-sample modes and their calculation lead to the usually dubbed *beat-coupling effect*. For didactic reason, we would like to derive the beat-coupling covariance at leading order (l.o.) in the PT framework.

$$\begin{aligned} T\left(\mathbf{k}_1, -\mathbf{k}_1-\mathbf{p}, \mathbf{k}_2, -\mathbf{k}_2-\mathbf{p}\right) \underset{\text{l.o.}}{\approx} & T\left(\mathbf{k}_1, -\mathbf{k}_1, \mathbf{k}_2, -\mathbf{k}_2\right) + \\ & 16\, P^{\rm lin.}\left(k_1\right)\, P^{\rm lin.}\left(k_2\right)\, P^{\rm lin.}\left(p\right)\, F^{(2)}\left(\mathbf{k}_1, -\mathbf{p}\right) F^{(2)}\left(\mathbf{k}_2, \mathbf{p}\right). \end{aligned} \tag{2.190}$$

By replacing it within Eq. (2.189) and Eq. (2.186), we find indeed a distinct separation between the standard intra-survey component $\mathrm{Cov}\left[\ldots\right]_{\rm NGins}$ of the covariance (2.122) (from the first term in Eq. (2.190)) and the

super-survey term $\mathrm{Cov}\left[\dots\right]_{\mathrm{NGssc}}$

$$\mathrm{Cov}\left[\hat{P}^{\mathrm{w}}_{ij}\left(\ell^b_1\right)\hat{P}^{\mathrm{w}}_{i'j'}\left(\ell^b_2\right)\right]_{\mathrm{NGssc}} \tag{2.191}$$

$$= \frac{16}{\Omega_{\mathrm{sky}} N_{\mathrm{p}}\left(\ell^b_1\right) N_{\mathrm{p}}\left(\ell^b_2\right)} \int_0^\infty d\omega\, \mathcal{T}\left(iji'j';\omega\right) \sum_{\bar{\ell}'_1,\bar{\ell}'_2} \Delta^2_{\ell^b_1}\left(\bar{\ell}_1\right) \Delta^2_{\ell^b_2}(\bar{\ell}'_2) \int \frac{d^2\mathbf{q}}{(2\pi)^2} |\tilde{W}(\mathbf{q})|^2 \times$$
$$P^{\mathrm{lin.}}\left(\frac{\ell'_1}{\omega}\right) P^{\mathrm{lin.}}\left(\frac{\ell'_2}{\omega}\right) P^{\mathrm{lin.}}\left(\frac{q}{\omega}\right) F^{(2)}\left(\mathbf{k}\left(\ell'_1,\omega\right),-\frac{\mathbf{q}}{\omega}\right) F^{(2)}\left(\mathbf{k}\left(\ell'_2,\omega\right),\frac{\mathbf{q}}{\omega}\right) \tag{2.192}$$

$$= \frac{16}{\Omega^2_{\mathrm{sky}}} \int_0^\infty d\omega\, \mathcal{T}\left(iji'j';\omega\right) \int \frac{d^2\ell'_1}{A_{\ell^b_1}} \int \frac{d^2\ell'_2}{A_{\ell^b_2}} \int \frac{d^2\mathbf{q}}{(2\pi)^2} |\tilde{W}(\mathbf{q})|^2 \times$$
$$P^{\mathrm{lin.}}\left(\frac{\ell'_1}{\omega}\right) P^{\mathrm{lin.}}\left(\frac{\ell'_2}{\omega}\right) P^{\mathrm{lin.}}\left(\frac{q}{\omega}\right) F^{(2)}\left(\mathbf{k}(\boldsymbol{\ell}'_1),-\mathbf{q}\right) F^{(2)}\left(\mathbf{k}(\boldsymbol{\ell}'_2),\mathbf{q}\right) \tag{2.193}$$

$$= \frac{16}{\Omega^2_{\mathrm{sky}}}\left(\frac{36}{49}\right) \int_0^\infty d\omega\, \mathcal{T}\left(iji'j';\omega\right) \int \frac{d\ell'_1}{\ell_1 \Delta\ell^b_1} P^{\mathrm{lin.}}\left(\frac{\ell'_1}{\omega}\right) \int \frac{d\ell'_2}{\ell_2 \Delta\ell^b_2} P^{\mathrm{lin.}}\left(\frac{\ell'_2}{\omega}\right) \times$$
$$\int \frac{d^2\mathbf{q}}{(2\pi)^2} |\tilde{W}(\mathbf{q})|^2 P^{\mathrm{lin.}}\left(\frac{q}{\omega}\right) \tag{2.194}$$

$$\propto \frac{1}{\Omega_{\mathrm{sky}}} \int_0^\infty d\omega\, \mathcal{T}\left(iji'j';\omega\right) \int P^{\mathrm{lin.}}\left(\frac{\ell'_1}{\omega}\right) P^{\mathrm{lin.}}\left(\frac{\ell'_2}{\omega}\right) \sigma^2_{\mathrm{v}}(\omega)\,. \tag{2.195}$$

Let us give few details about the intermediate calculations, the PT second-order kernels $F^{(2)}$ (see Eq. (1.111)) should depend on three-dimensional momenta: we employed the notation introduced around Eq. (2.89) to refer the the three-dimensional momenta under the Limber approximation. In the second line, given the shape of the kernel $F^{(2)}$ we removed the overall scalar factor $1/\omega$. Also, to have a simpler notation, we moved to the integral form of the bin-average. In the third line, we averaged the $F^{(2)}$ kernel of the polar angle of $\boldsymbol{\ell}'_1$ and $\boldsymbol{\ell}'_2$, leading to the factor $36/49 = (6/7)^2$. In the fourth line 1) we assumed that the linear matter power spectra do not vary that much within the multiple bins, 2) we replaced the expression for the annulus area related to the bins ℓ^b_1 and ℓ^b_2 and we identified the last term as the variance of the linear matter field over the survey footprint at comoving distance ω

$$\sigma^2_{\mathrm{v}}(\omega) = \int \frac{d^2\mathbf{q}}{(2\pi)^2} |\tilde{W}(\mathbf{q})|^2 P^{\mathrm{lin.}}\left(\frac{q}{\omega}\right). \tag{2.196}$$

At the level of bispectrum and cross power spectrum-bisepctrum covariance for the masked estimators defined in Eq. (2.151) and (2.152), similar results can be derived. It is possible to analytically prove (under the usual limit of observed modes much larger than the fundamental frequency) that all the terms identified in Sec. 2.2 are indeed not biased except for the 6- and 5-point correlation function related ones (see Barreira [2019] for a detailed

calculation at the level of matter bispectrum)

$$\mathrm{Cov}\Big[\hat{B}_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big),\hat{B}_{i'j'k'}(\ell_1'^b,\ell_2'^b,\ell_3'^b)\Big]_{\mathrm{Gauss}} \underset{\ell\gg\ell_{\mathrm{f}}}{\approx} \mathrm{Cov}\Big[\hat{B}^{\mathrm{w}}_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big),\hat{B}^{\mathrm{w}}_{i'j'k'}(\ell_1'^b,\ell_2'^b,\ell_3'^b)\Big]_{\mathrm{Gauss}}, \tag{2.197}$$

$$\mathrm{Cov}\Big[\hat{B}_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big),\hat{B}_{i'j'k'}(\ell_1'^b,\ell_2'^b,\ell_3'^b)\Big]_{\mathrm{NGins,BB}} \underset{\ell\gg\ell_{\mathrm{f}}}{\approx} \mathrm{Cov}\Big[\hat{B}^{\mathrm{w}}_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big),\hat{B}^{\mathrm{w}}_{i'j'k'}(\ell_1'^b,\ell_2'^b,\ell_3'^b)\Big]_{\mathrm{NGins,BB}}, \tag{2.198}$$

$$\mathrm{Cov}\Big[\hat{B}_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big),\hat{B}_{i'j'k'}(\ell_1'^b,\ell_2'^b,\ell_3'^b)\Big]_{\mathrm{NGins,PT}} \underset{\ell\gg\ell_{\mathrm{f}}}{\approx} \mathrm{Cov}\Big[\hat{B}^{\mathrm{w}}_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big),\hat{B}^{\mathrm{w}}_{i'j'k'}(\ell_1'^b,\ell_2'^b,\ell_3'^b)\Big]_{\mathrm{NGins,PT}}, \tag{2.199}$$

$$\mathrm{Cov}\Big[\hat{P}_{ij}\left(\ell^b\right),\hat{B}_{i'j'k'}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)\Big]_{\mathrm{NGins,PB}} \underset{\ell\gg\ell_{\mathrm{f}}}{\approx} \mathrm{Cov}\Big[\hat{P}_{ij}\left(\ell^b\right),\hat{B}_{i'j'k'}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)\Big]_{\mathrm{NGins,PB}}, \tag{2.200}$$

while

$$\mathrm{Cov}\Big[\hat{B}^{\mathrm{w}}_{ijk}\big(\ell_1^b,\ell_2^b,\ell_3^b\big),\hat{B}^{\mathrm{w}}_{i'j'k'}(\ell_1'^b,\ell_2'^b,\ell_3'^b)\Big]_{\mathrm{NG,6P}} = \frac{\sum_{\boldsymbol{\ell}_1',\boldsymbol{\ell}_2',\boldsymbol{\ell}_3'}\sum_{\boldsymbol{\ell}_1'',\boldsymbol{\ell}_2'',\boldsymbol{\ell}_3''}\Delta^3_{\ell_1'^b,\ell_2'^b,\ell_3'^b}\left(\boldsymbol{\ell}_1',\boldsymbol{\ell}_2',\boldsymbol{\ell}_3'\right)\Delta^3_{\ell_1^b,\ell_2^b,\ell_3^b}\left(\boldsymbol{\ell}_1',\boldsymbol{\ell}_2',\boldsymbol{\ell}_3'\right)}{\Omega_{\mathrm{sky}}\,N_{\mathrm{tri.}}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)N_{\mathrm{tri.}}\left(\ell_1'^b,\ell_2'^b,\ell_3'^b\right)}$$
$$\int\frac{d^2\mathbf{q}}{(2\pi)^2}|\tilde{W}(\mathbf{q})|^2P_{ijki'j'k'}\left(\boldsymbol{\ell}_1',\boldsymbol{\ell}_2',\boldsymbol{\ell}_3'+\mathbf{q},\boldsymbol{\ell}_1'',\boldsymbol{\ell}_2'',\boldsymbol{\ell}_3''-\mathbf{q}\right), \tag{2.201}$$

$$\mathrm{Cov}\Big[\hat{P}^{\mathrm{w}}_{ij}\left(\ell^b\right),\hat{B}^{\mathrm{w}}_{i'j'k'}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)\Big]_{\mathrm{NG,5P}} = \frac{\sum_{\boldsymbol{\ell}'}\sum_{\boldsymbol{\ell}_1',\boldsymbol{\ell}_2',\boldsymbol{\ell}_3'}\Delta^2_{\ell_1^b}\left(\boldsymbol{\ell}_1'\right)\Delta^3_{\ell_1^b,\ell_2^b,\ell_3^b}\left(\boldsymbol{\ell}_1',\boldsymbol{\ell}_2',\boldsymbol{\ell}_3'\right)}{\Omega_{\mathrm{sky}}\,N_{\mathrm{tri.}}\left(\ell_1^b,\ell_2^b,\ell_3^b\right)N_{\mathrm{p}}\left(\ell^b\right)}$$
$$\int\frac{d^2\mathbf{q}}{(2\pi)^2}|\tilde{W}(\mathbf{q})|^2P_{iji'j'k'}\left(\boldsymbol{\ell}_1',\boldsymbol{\ell}_2',\boldsymbol{\ell}_3'+\mathbf{q},\boldsymbol{\ell}',-\boldsymbol{\ell}'-\mathbf{q}\right). \tag{2.202}$$

In Eqs. (2.197)-(2.197) we use ℓ below the relational signs to refer to the general observed mode. Also in this case PT allows to split the components of the non-Gaussian covariance matrices in two terms: one involving only correlations between modes within the survey footprint (leading to the NGins-type terms in Eqs. (2.99)-(2.101)), and one involving correlation between the 6, 5 (respectively) observed modes and a background mode $\mathbf{q}$ (the NGssc-type terms in Eqs. (2.99)-(2.101)). Schematically then,

$$\mathrm{Cov}\left[PP\right]^{PT}_{\mathrm{NG}} = \mathrm{Cov}\left[PP\right]_{\mathrm{NGins}} + \mathrm{Cov}\left[PP\right]_{\mathrm{NGssc}}, \tag{2.203}$$

$$\mathrm{Cov}\left[BB\right]^{PT}_{\mathrm{NG}} = \mathrm{Cov}\left[PP\right]_{\mathrm{NGins.}} + \mathrm{Cov}\left[PP\right]_{\mathrm{NGssc}}, \tag{2.204}$$

$$\mathrm{Cov}\left[PB\right]^{PT}_{\mathrm{NG}} = \mathrm{Cov}\left[PB\right]_{\mathrm{NGins}} + \mathrm{Cov}\left[PB\right]_{\mathrm{NGssc}}. \tag{2.205}$$

We have justified the expressions anticipated in Eq. (2.99)- 2.101 within the PT framework. Actually, we will see in the following that this structure is granted independently of the model for non-linear clustering employed.

2.3.2 SSC via the response approach

Starting from the studies summarised in the previous section, a very interesting formalism was recently introduced capable to include the leading-order PT results (beat coupling) and also provide a better fit to simulations,

including missing terms in the non-linear regime of matter clustering.

MATTER POLYSPECTRUM CONSISTENCY RELATIONS The *response approach* (in this context) was first introduced by Takada & Hu [2013]. In analogy with similar effects for primordial non-Gaussianity [Maldacena, 2003], the polyspectra on the squeezed configurations required in Eqs. (2.189), (2.202) and (2.201) should be determined by the reposnse of the power spectrum and bispectrum to the long mode $\delta_{\rm b}$ defined as

$$\delta_b \equiv \frac{1}{V}\int_V d^3\mathbf{x}\,\delta\left(\mathbf{x}\right), \qquad \tilde{\delta}_b\left(\mathbf{q}\right) = \left(2\pi\right)^3\delta_{\rm D}\left(\mathbf{q}\right)\delta_{\rm b}. \tag{2.206}$$

where, we recall, δ_D is the Dirac delta. Takada & Hu [2013] defined the following *trispectrum consistency relation*.

$$T\left(\mathbf{k}_1,-\mathbf{k}_1+\mathbf{p},\mathbf{k}_2,-\mathbf{k}_2-\mathbf{p}\right) \approx T\left(\mathbf{k}_1,-\mathbf{k}_1,\mathbf{k}_2,-\mathbf{k}_2\right) + \frac{{\rm d}P(k_1)}{{\rm d}\delta_{\rm b}}\frac{\partial P(k_2)}{\partial\delta_{\rm b}}P^{\rm lin}\left(p\right). \tag{2.207}$$

with $p = |\mathbf{p}|$ and $P^{\rm lin}$ is the linear matter power spectrum. This results generalises the leading-order PT-based results of the previous section: it is always possible to split the covariance for windowed estimators in 1) correlations between observed modes, and in 2) a super-sample component involving correlations between the observed and the un-observed modes, the latter entering via the variance of the linear matter field over the survey volume

$$\begin{aligned}{\rm Cov}\left[\hat{P}^{\rm w}_{ij}\left(\ell^b_1\right)\hat{P}^{\rm w}_{i'j'}\left(\ell^b_2\right)\right]_{\rm NG} = {\rm Cov}\left[\hat{P}_{ij}\left(\ell^b_1\right)\hat{P}_{i'j'}\left(\ell^b_2\right)\right]_{\rm NGins} + \\ \frac{1}{\Omega_{\rm sky}}\int_{\rm o}^{\infty}d\omega\,\mathcal{T}\left(iji'j';\omega\right)\frac{{\rm d}P}{{\rm d}\delta_{\rm b}}\left(\mathbf{k}(\ell,\omega)\right)\frac{{\rm d}P}{{\rm d}\delta_{\rm b}}\left(\mathbf{k}(\ell',\omega)\right)\sigma^2_{\rm v}\left(\omega\right).\end{aligned} \tag{2.208}$$

Via the same argument, Chan et al. [2018] found an analogous relation for the bispectrum covariance and the power spectrum-bispectrum cross-covariance

$${\rm Cov}\left[\hat{B}^{\rm w}_{ijk}\left(\ell^b_1,\ell^b_2,\ell^b_3\right),\hat{B}^{\rm w}_{i'j'k'}\left(\ell'^b_1,\ell'^b_2,\ell'^b_3\right)\right]_{\rm NG,6P} = {\rm Cov}\left[\ldots\right]_{\rm NGins} + {\rm Cov}\left[\ldots\right]_{\rm NGssc} \tag{2.209}$$

$${\rm Cov}\left[\hat{P}_{ij}\left(\ell^b\right),\hat{B}_{i'j'k'}\left(\ell^b_1,\ell^b_2,\ell^b_3\right)\right]_{\rm NG,5P} = {\rm Cov}\left[\ldots\right]_{\rm NGins} + {\rm Cov}\left[\ldots\right]_{\rm NGssc} \tag{2.210}$$

with

$${\rm Cov}\left[\hat{B}^{\rm w}_{ijk}\left(\ell^b_1,\ell^b_2,\ell^b_3\right),\hat{B}^{\rm w}_{i'j'k'}\left(\ell'^b_1,\ell'^b_2,\ell'^b_3\right)\right]_{\rm NGins} = {\rm Cov}\left[\hat{B}_{ijk}\left(\ell^b_1,\ell^b_2,\ell^b_3\right),\hat{B}_{i'j'k'}\left(\ell'^b_1,\ell'^b_2,\ell'^b_3\right)\right]_{\rm NGins}, \tag{2.211}$$

$${\rm Cov}\left[\hat{P}^{\rm w}_{ij}\left(\ell^b\right),\hat{B}^{\rm w}_{i'j'k'}\left(\ell^b_1,\ell^b_2,\ell^b_3\right)\right]_{\rm NGins} = {\rm Cov}\left[\hat{P}_{ij}\left(\ell^b\right),\hat{B}_{i'j'k'}\left(\ell^b_1,\ell^b_2,\ell^b_3\right)\right]_{\rm NGins}, \tag{2.212}$$

and

$$\mathrm{Cov}\left[\hat{B}^{\mathrm{w}}_{ijk}\left(\ell^b_1,\ell^b_2,\ell^b_3\right),\hat{B}^{\mathrm{w}}_{i'j'k'}(\ell'^b_1,\ell'^b_2,\ell'^b_3)\right]_{\mathrm{NGssc}}=$$
$$\frac{1}{\Omega_{\mathrm{sky}}}\int_0^\infty d\omega\,\mathcal{T}\left(ijki'j'k';\omega\right)\frac{\mathrm{d}B}{\mathrm{d}\delta_{\mathrm{b}}}\left(\mathbf{k}(\ell_1,\omega),\mathbf{k}(\ell_2,\omega),\mathbf{k}(\ell_3,\omega)\right)\frac{\mathrm{d}B}{\mathrm{d}\delta_{\mathrm{b}}}\left(\mathbf{k}(\ell'_1,\omega),\mathbf{k}(\ell'_2,\omega),\mathbf{k}(\ell'_3,\omega)\right)\sigma^2_{\mathrm{v}}\left(\omega\right),\quad(2.213)$$

$$\mathrm{Cov}\left[\hat{P}^{\mathrm{w}}_{ij}\left(\ell^b\right),\hat{B}^{\mathrm{w}}_{i'j'k'}\left(\ell^b_1,\ell^b_2,\ell^b_3\right)\right]_{\mathrm{NGssc}}=$$
$$\frac{1}{\Omega_{\mathrm{sky}}}\int_0^\infty d\omega\,\mathcal{T}\left(iji'j'k';\omega\right)\frac{\mathrm{d}P}{\mathrm{d}\delta_{\mathrm{b}}}\left(\mathbf{k}(\ell,\omega)\right)\frac{\mathrm{d}B}{\mathrm{d}\delta_{\mathrm{b}}}\left(\mathbf{k}(\ell'_1,\omega),\mathbf{k}(\ell'_2,\omega),\mathbf{k}(\ell'_3,\omega)\right)\sigma^2_{\mathrm{v}}\left(\omega\right).\quad(2.214)$$

The quantities $\mathrm{d}\star/\mathrm{d}\delta_{\mathrm{b}}$ are called in the literature *responses* : their derivation will be the subject of the next paragraph.

Polyspectra responses For the evaluation of the responses one usually considers three main components according to the separate Universe picture [Sirko, 2005; Baldauf et al., 2011; Sherwin & Zaldarriaga, 2012; Li et al., 2014; Wagner et al., 2015a; Baldauf et al., 2016] and as interestingly outlined in Lazeyras et al. [2016] at the level the power spectrum and in Chan et al. [2018] for the bispectrum: the observables built in a local over-dense region of the Universe (P^{L}) as altered in presence of a long mode δ_{b} are related to the quantities in the global one (P^{G}). We can easily extend their considerations to a general matter polyspectrum. The responses can be seen as the coefficients of an expansion of the true globally derived polyspectrum around the background mode δ_{b}, which is assumed to be in the linear regime

$$P^{\mathrm{G}}\left(\mathbf{k}_1,\ldots,\mathbf{k}_n|\delta_{\mathrm{b}}\right)=P^{\mathrm{G}}\left(\mathbf{k}_1,\ldots,\mathbf{k}_n;0\right)+\sum_{n>0}\frac{R_n\left(\mathbf{k}_1,\ldots,\mathbf{k}_n\right)}{n!}\delta^n_{\mathrm{b}},$$
$$R_m\left(\mathbf{k}_1,\ldots,\mathbf{k}_n\right)=\frac{\mathrm{d}^m P\left(\mathbf{k}_1,\ldots,\mathbf{k}_n|\delta_{\mathrm{b}}\right)}{\mathrm{d}\delta^m_{\mathrm{b}}}\bigg|_{\delta_{\mathrm{b}}=0}\quad(2.215)$$

In this work we will consider *linear responses* $R_1\left(\mathbf{k}_1,\ldots,\mathbf{k}_n\right)$.

Local average effect this effect was already identified within the calculation of the non-linear power spectrum covariance matrix by de Putter et al. [2012]. The background mode δ_{b} can induce a *local average effect* on small scales via a shift on the mean background matter density

$$\rho^{\mathrm{L}}=\left(1+\delta_{\mathrm{b}}\right)\rho,\quad(2.216)$$

and at the level of locally observed contrast density field

$$\delta^{\mathrm{L}}\left(\mathbf{x}\right)=\frac{\delta\left(\mathbf{x}\right)-\delta_{\mathrm{b}}}{1+\delta_{\mathrm{b}}}\quad\rightarrow\quad\delta^{\mathrm{L}}\left(\mathbf{k}\right)=\frac{\delta\left(\mathbf{k}\right)-\delta_{\mathrm{b}}\delta_{\mathrm{D}}\left(\mathbf{k}\right)}{1+\delta_{\mathrm{b}}}.\quad(2.217)$$

At the level of polyspectra, the local $P^{\rm L}$ ones relate to the global ones P(from now on we will omit the superscript G), in Fourier space, via

$$P\left(\mathbf{k}_1,\ldots,\mathbf{k}_n|\delta_{\rm b}\right) = \left(1+\delta_{\rm b}\right)^n\, P^{\rm L}\left(\mathbf{k}_1,\ldots,\mathbf{k}_n\right) \tag{2.218}$$

We can see that an intuitive behaviour is recovered: given a fixed local polyspectrum, the corresponding one related to the true underlying matter density is enhanced. Clearly, for a fixed background mode, the lower is the local power, the lower will be the global one.

Dilation effect A positive background over-density implies a slower expansion, which can be quantified by the relative difference in the scale factor between the local and global cosmology, $a^{\rm L} = a/\sqrt[3]{1+\delta_{\rm b}}$. Therefore, for the same physical scale $r_{\rm phy.}$ the corresponding comoving one $r_{\rm com.}$ differs as $r_{\rm com.}a = r^{V}_{\rm com.}a_L$, and the momenta $k \approx 1/r_{\rm com.}$ are therefore differently evaluated between the local and global Universe as

$$\mathbf{k}^{\rm L} = \frac{\mathbf{k}}{\sqrt[3]{1+\delta_{\rm b}}}. \tag{2.219}$$

The impact of such a coordinate scaling at the level of polyspectra can be derived via a generalisation of the calculation performed in Pajer et al. [2013] and in Chan et al. [2018]. Under a change of coordinated alone (excluding local-average effect and the growth-only response (see later in the section)), a general n-point correlation function transform as a scalar field, specifically in this case (with $\mathbf{x} \to \mathbf{x}^{\rm L} = \mathbf{x}\sqrt[3]{1+\delta_{\rm b}}$)

$$\varepsilon^{\rm L}\left(\mathbf{r}^{\rm L}_2,\ldots,\mathbf{r}^{\rm L}_n\right) = \varepsilon\left(\mathbf{r}^{\rm L}_2/\sqrt[3]{1+\delta_{\rm b}},\ldots,\mathbf{r}^{\rm L}_n/\sqrt[3]{1+\delta_{\rm b}}\right). \tag{2.220}$$

We can relate them to the corresponding n-order polyspectra

$$\varepsilon\left(\mathbf{r}^{\rm L}_2/\sqrt[3]{1+\delta_{\rm b}},\ldots,\mathbf{r}^{\rm L}_n/\sqrt[3]{1+\delta_{\rm b}}\right) = \prod_{i=2}^{n}\left[\int \frac{d^3\mathbf{k}_i}{(2\pi)^3}\, e^{-i\frac{\mathbf{k}_i\cdot\mathbf{r}^{\rm L}_i}{\sqrt[3]{1+\delta_{\rm b}}}}\right] P\left(\mathbf{k}_2,\ldots,\mathbf{k}_n\right), \tag{2.221}$$

$$\varepsilon^{\rm L}\left(\mathbf{r}^{\rm L}_2,\ldots,\mathbf{r}^{\rm L}_n\right) = \prod_{i=2}^{n}\left[\int \frac{d^3\mathbf{k}^{\rm L}_i}{(2\pi)^3}\, e^{-i\mathbf{k}^{\rm L}_i\cdot\mathbf{r}^{\rm L}_i}\right] P^{\rm L}\left(\mathbf{k}^{\rm L}_2,\ldots,\mathbf{k}^{\rm L}_n\right), \tag{2.222}$$

We then translate the equivalence in Eq. (2.220) into a relation between locally and globally evaluated polyspectra

$$P^{\rm L}\left(\mathbf{k}^{\rm L}_2,\ldots,\mathbf{k}^{\rm L}_n\right) = \left(1+\delta_{\rm b}\right)^{n-1} P\left(\mathbf{k}^{\rm L}_2\sqrt[3]{1+\delta_{\rm b}},\ldots,\mathbf{k}^{\rm L}_n\sqrt[3]{1+\delta_{\rm b}}\right) \sim$$
$$P\left(\mathbf{k}_2,\ldots,\mathbf{k}_n\right) = \left(1+\delta_{\rm b}\right)^{-n+1} P^{\rm L}\left(\mathbf{k}_2/\sqrt[3]{1+\delta_{\rm b}},\ldots,\mathbf{k}_n/\sqrt[3]{1+\delta_{\rm b}}\right). \tag{2.223}$$

Combining the local average effect (2.218) and the dilation effect (2.223), we have

$$P\left(\mathbf{k}_2,\ldots,\mathbf{k}_n|\delta_{\rm b}\right) = \left(1+\delta_{\rm b}\right) P^{\rm L}\left(\mathbf{k}_2/\sqrt[3]{1+\delta_{\rm b}},\ldots,\mathbf{k}_n/\sqrt[3]{1+\delta_{\rm b}}\right). \tag{2.224}$$

Growth-only response We are then left with the growth-only response which stands for a modification of the intrinsic growth of structures, at fixed scales. If the background perturbation is positive, then gravity is stronger in the local Universe and the intrinsic growth is enhanced. Unfortunately this last quantity cannot be predicted via the above shifts but it has to be either tuned in simulations or predicted according to a given theoretical prescription. In general we will refer to it as the response to the locally observed polyspectrum to the long mode

$$G_n(\mathbf{k}_2, \dots, \mathbf{k}_n) = \frac{\mathrm{d}^n P^{\mathrm{L}}\left(\mathbf{k}, \dots, \mathbf{k}|\delta_{\mathrm{b}}\right)}{\mathrm{d}\delta_{\mathrm{b}}^n}\bigg|_{\delta_{\mathrm{b}}=0} \frac{1}{P\left(\mathbf{k}, \dots, \mathbf{k}\right)}. \tag{2.225}$$

and we added a dependence on the background mode due to the alteration of the growth of structures

Given all the above effects, the true globally evaluated polyspectrum will depend on the long background mode δ_{b} as

$$P\left(\mathbf{k}_1, \dots, \mathbf{k}_n|\delta_{\mathrm{b}}\right) = (1+\delta_{\mathrm{b}}) \left[\left(1 + \sum_{m>0} \frac{1}{m!} G_m(\mathbf{k}_1^{\mathrm{L}}, \dots, \mathbf{k}_n^{\mathrm{L}})\, \delta_b^m\right) P\left(\mathbf{k}_1^{\mathrm{L}}, \dots, \mathbf{k}_n^{\mathrm{L}}\right)\right]_{\mathbf{k}_*^{\mathrm{L}} = \frac{\mathbf{k}_*}{\sqrt[3]{1+\delta_{\mathrm{b}}}}} \tag{2.226}$$

where $P^{\mathrm{L}}\left(\mathbf{k}_2, \dots, \mathbf{k}_n|0\right) = P\left(\mathbf{k}_2, \dots, \mathbf{k}_n\right)$ since we are not considering the effect of the long mode on the growth of structure and we are evaluating them at the same configurations. Finally, the linear response will be

$$R_1\left(\mathbf{k}_1, \dots, \mathbf{k}_n\right) = 1 + G_1(\mathbf{k}_1, \dots, \mathbf{k}_n) - \frac{1}{3}\sum_{i=1}^{n} \frac{\mathrm{d}P\left(\mathbf{k}_1, \dots, \mathbf{k}_n\right)}{\mathrm{d}\log \mathbf{k}_i} \tag{2.227}$$

In the next chapter, we will review all the polyspectra included in our analysis. Along with them, we will describe the expressions used for their responses in the light of Eq. (2.227)

Weak lensing SSC, few comments In the present chapter, given the scales of interest for our analyses and the simplicity in the computation, we built our observables in the flat-sky and Limber approximations. We will see in the next chapter that these assumptions are indeed well justified for our work. These approximations were also employed for deriving the super-sample covariance terms in the present section, the soft mode, generally identified with $\mathbf{q}$ being defined on a plane. In the most general case modes should be treated as curved and even if the small scales limit may apply for the observed ones, the background long modes do lie on a sphere and can potentially correlate different small "flat" patches of the observed sky. In the interesting paper from Barreira et al. [2018], the authors do actually provide a rigorous proof for different expressions of the SSC for the weak lensing convergence field relaxing different approximations on the geometry of the modes: in Sec. 4.2 they derive the SSC with curved super-survey modes only, while in Appendix D they derive the SSC with no flat-sky approximation, nor Limber, for both the background and observed modes. While these expressions can actually be derived analytically, they can be computationally costly. As a matter of fact, in the second case (the most realistic), the non-Gaussian lensing covariance is given by a high-dimensional integral of rapidly oscillating functions with unequal-time matter trispectrum in general configurations. Furthermore, this full expression does not even permit a clean separation

at the level of trispectrum between super-sample contributions and the rest of the connected non-Gaussian covariance (as it was the case for the flat-sky analyses performed so far). For the bispectrum covariance, all these considerations lead to an even more complex analysis. In the less complex case with flat-sky for the observed modes, the analysis still involves 2 integrations over the redshift for the evaluation of the projected variance σ_v^2, for every configuration of interest. This complexity was also found and successfully tackled by Lacasa et al. [2016]. For all these reasons, we will employ the flat-sky approximation for both the modes observed and beyond the survey. By generalising to higher order the consistency relations (2.207) the following expression is given by simply projecting the three-dimensional polyspectra responses

$$\mathrm{Cov}\left[\hat{P}^{\mathrm{w}}_{i_1,\ldots,i_n}\left(\boldsymbol{\ell}_1,\ldots,\boldsymbol{\ell}_n\right),\hat{P}^{\mathrm{w}}_{i'_1,\ldots,i'_{n'}}\left(\boldsymbol{\ell}'_1,\ldots,\boldsymbol{\ell}'_{n'}\right)\right]_{\mathrm{NGssc}} \tag{2.228}$$

$$= \int_0^\infty d\omega\, \mathcal{T}\left(i_1,\ldots,i_n;\omega\right)\, \mathcal{T}\left(i'_1,\ldots,i'_{n'};\omega\right) \times \tag{2.229}$$

$$\frac{\mathrm{d}\, P\left(\mathbf{k}\left(\boldsymbol{\ell}_1,\omega\right),\ldots,\mathbf{k}\left(\boldsymbol{\ell}_n,\omega\right)|\delta_{\mathrm{b}}\right)}{\mathrm{d}\,\delta_{\mathrm{b}}} \frac{\mathrm{d}\, P\left(\mathbf{k}\left(\boldsymbol{\ell}'_1,\omega\right),\ldots,\mathbf{k}\left(\boldsymbol{\ell}'_{n'},\omega\right)|\delta_{\mathrm{b}}\right)}{\mathrm{d}\,\delta_{\mathrm{b}}} \sigma_v^2\left(\omega\right) \tag{2.230}$$

where we employ the Limber approximation for the three-dimensional modes within the responses. The super-sample covariance expression (2.228) is also consistent with the work of Lacasa & Rosenfeld [2016] (see also Lacasa et al. [2016] for an application with partial-sky coverage). They derived a general expression for the super-sample covariance between two observables $\mathcal{O}_1$ and $\mathcal{O}_2$ for a projected cosmological field

$$\mathrm{Cov}\left[\mathcal{O}_1,\mathcal{O}_2\right]_{\mathrm{NGssc}} = \int \mathrm{d}\omega_1\, \mathrm{d}\omega_2\, \omega_1^2\, \omega_2^2 \frac{\mathrm{d}o_1\left(\omega_1\right)}{\mathrm{d}\delta_{\mathrm{b}}} \frac{o_2\left(\omega_2\right)}{\mathrm{d}\delta_{\mathrm{b}}} \sigma_{\mathrm{v}}^2\left(\omega_1,\omega_2\right). \tag{2.231}$$

where

$$o(\omega) = \frac{1}{\omega^2}\frac{\mathrm{d}\mathcal{O}(\omega)}{\mathrm{d}\omega}, \quad \sigma_{\mathrm{v}}^2\left(\omega_1,\omega_2\right) = \int \frac{d^3\mathbf{k}}{\left(2\pi\right)^3}\, \tilde{W}(\mathbf{k},\omega_1)\tilde{W}^*(\mathbf{k},\omega_2)\, P^{\mathrm{lin.}}(k;\omega_1,\omega_2). \tag{2.232}$$

In particular we recover our expression under the limit $\sigma_{\mathrm{v}}^2\left(\omega_1,\omega_2\right)$ being peaked around $\omega_1 = \omega_2$ (Limber approximation regime).

We leave for Chapter 3 more details regarding the shape of the window function used for the evaluation of the variance σ_{v}^2. We just underline here that weak lensing surveys dealing with projected fields are characterised by an angular window function $W\left(\omega,\theta,\varphi\right)$. In the most general case, the depth of the survey (related to the distance ω) is coupled with the direction in the sky, (related to the angular coordinates). This increases the numerical complexity highlighted in the paragraph above. In our analysis, we will allow for a factorisation in a radial and in an angular component

$$W\left(r,\theta,\varphi\right) = W_{\mathrm{R}}\left(r\right) \times W_{\mathrm{an}}\left(\theta,\varphi\right). \tag{2.233}$$

While the radial component will be absorbed as extrema in the line-of-sight integration in Eq. (2.228) (and in particular will be equal to the size of the tomographic bins employed), the angular component will contribute to three-dimensional variance in Eq. (2.196) once projected onto plane waves. We will show this calculation in

details for the chosen window function in the next chapter.

3

Weak lensing bispectrum information content

Contents

The purpose of this third chapter is to finally apply the formalism introduced so far to the optimisation of future data analyses for modern galaxy surveys. While the concepts we will introduce are absolutely general, we will derive our conclusions for a Euclid-like survey whose detection specificities are given in Sec. 3.1.2.

All the ingredients introduced in the first two chapters do play an important role both at the level of *forecast* and at the level of actual *data analysis*, where the parameter inference process takes place. The purpose of the forecast analysis is to predict how well cosmological parameters will be constrained by a set of observables as measured with a specific survey. Always at this level, we want to identify non informative observables to discard and build summary statistics which maximise the cosmological content in the observables while minimising the computational burden in the future data pipelines. Finally, we want to test the robustness of the upcoming parameter inference against the uncertainties on *nuisance parameters*. These type of parameters are necessary to model the expected signal but the uncertainty on their values, or correlations, may spoil the precision of the future inference. It is important to underline that the forecast process has major consequences for the actual data analysis: the latter is optimised based on the outcome of the former. For example, summary statistics identified during the forecast can be estimated at the level of data along with removing non-informative combinations of observables, leading to faster parameter posterior exploration and probe combination. Moreover, given the importance of an analytical approach as described in Sec. 2.2.1, the covariances and the observables computed for a fiducial cosmology will enter directly the *likelihood* codes used for inferring the cosmological parameters as part of the heritage of a mission as whole.

In this final chapter we will address key points for an efficient implementation of likelihood codes for modern weak lensing large-scale structure surveys. Specifically, we will focus on the joint weak lensing convergence power spectrum-bispectrum probe and we will tackle the numerical challenges required by a thorough analysis. Under the assumption of a multivariate Gaussian likelihood, we developed a high performance code that allows highly parallelised prediction of the binned tomographic observables and of their joint non-Gaussian covariance matrix accounting for terms up to the 6-point correlation function and super-sample effects.

So far, we have been mentioning the *cosmological information content* as the quantity to maximise, providing a guiding criterion for our research. Qualitatively, we can associate it with the overall precision expected for the cosmological parameters to be estimated from the available data, given their uncertainties. In big-data context, the constraining power of a probe can be addressed either via Fisher forecast [Tegmark, 1997], DALI forecast [Sellentin et al., 2014; Sellentin, 2015] or Markov Chain Monte Carlo (MCMC) analyses (just to give few examples). However, given the novelty and the complexity of the study here carried out, all the above techniques are not computationally feasible. This difficulty is exacerbated by the fact that we aim at exploring different configurations of the target experiment. Instead, we propose a signal-to-noise ratio (S/N) analysis as suggested by Tegmark et al. [1997] and already exploited in previous literature within the cosmological context [Rimes & Hamilton, 2005; Sato et al., 2009; Takada & Jain, 2009; Kayo et al., 2013]. Following Kayo & Takada [2013], we define the information content of an observable as the expected inverse variance of its amplitude A as estimated from a

set of measured values $\mathbf{x}$, assuming a fixed shape. In other terms, the second power of the signal-to-noise ratio is the inverse of the unique element of the Fisher information matrix [Tegmark, 1997] in such a single parameter experiment. Specifically, we can write

$$\left(\frac{S}{N}\right)^2 \equiv \frac{\partial\left(\mathbf{x}-A\mathbf{D}\right)^t}{\partial A}\cdot \mathrm{C}^{-1}\cdot\frac{\partial\left(\mathbf{x}-A\mathbf{D}\right)}{\partial A} = \mathbf{D}^t\cdot\mathrm{C}^{-1}\cdot\mathbf{D}, \tag{3.1}$$

where $A\mathbf{D} = \langle\mathbf{x}\rangle$ and $\mathrm{C}_{ij} = \mathrm{Cov}[x_i, x_j]$. We will use this tool extensively in what follows. Even if we are aware that the S/N is not suitable for a precise forecast, however its simplicity allows a fast analysis of the level of correlations within our data vector. Nonetheless, the computation of the covariance matrix is a key step towards the forecasting techniques listed above and this work already finalises this calculation.

In this chapter we will present in Sec. 3.2 a thorough analysis of the most important forecast related aspects for a Euclid-like survey considering as observables the binned tomographic convergence power spectrum and bispectrum. The matter presented here below has been previously proposed in Rizzato et al. [2018].

3.1 Implementation specificities

The general goal of the present section is to specialise the theoretical discussions performed in Chapter 1 and in Chapter 2 to the actual implementation for our analysis. In Sec. 3.1.1, we will start by giving the structure of the chosen data vector. In Sec. 3.1.2, we will list the survey specificities of a Euclid-like mission. Finally, in Sec. 3.1.3, we will provide in-depth information regarding the construction of the halo model polyspectra required for the forward modelling of the data vector and of the covariance, responses included.

3.1.1 Vector of observables

We will be dealing with the tomographic convergence power spectrum and bispectrum, i.e.

$$P_{ij}\left(\ell\right), \qquad B_{ijk}\left(\ell_1,\ell_2,\ell_3\right). \tag{3.2}$$

To be precise, the quantities entering the data vector $\mathbf{D}$ in Eq. (3.1) are the binned estimators (2.151) and (2.152). We already proved them unbiased and, under the assumption of negligible variations within a given ℓ-bin, they are approximately equal to respectively the power spectrum and to the bispectrum as evaluated at the central value of the bin, or of the bins, involved in a given configuration. Then, we will populate our data vector with the observables (3.2) computed on the central value of the bins in ℓ. We build the vector with the overall structure

$$\mathbf{D} = \{\mathbf{P},\ \mathbf{B}\}. \tag{3.3}$$

In each of the two sub-vectors **P** and **B**, the observables (power spectrum and bisepctrum respectively) are stored for decreasing (increasing) value of the correlation scales (multipole) used for the configurations. The triangular configurations for the bispectrum, labelled via the triangle edges $\{\ell_i, \ell_j, \ell_k\}$, are ordered within the vector **B** such as $\ell_i \leqslant \ell_j \leqslant \ell_k$ where the index j is the fastest and the index k is the slowest varying one while moving along the vector towards higher indices. The angular multipole range we will investigate is $\ell \in [10, 5000]$[1], binned in 14 regularly spaced intervals in $\log \ell$. The choice for such a binning is arbitrary and it is mainly motivated by computational resource optimisation. In future analyses, the binning will be instead optimised to increase the signal-to-noise ratio of the observables. As for the different tomographic contributions, they are contiguous within the vector for the same spatial configuration. For the power spectrum, the tomographic indexes (i, j) are ordered such that $i \leqslant j$, j being the faster varying index through the vector. For the bispectrum, the observables associated to the tomographic bins (i, j, k) are ordered such that k is the fastest index and i is the slowest one. Symmetries in the spatial triangles help in removing redundant tomographic redshift contributions from the bispectrum vector **B**: bispectra on isosceles or equilateral configurations lead to the following symmetries within the corresponding row (column) in the covariance (e.g.)

$$\mathrm{Cov}\left[B_{ijk}\left(\ell,\ell,\ell'\right),\star\right] = \mathrm{Cov}\left[B_{jik}\left(\ell,\ell,\ell'\right),\star\right],$$
$$\mathrm{Cov}\left[B_{ijk}\left(\ell,\ell,\ell\right),\star\right] = \mathrm{Cov}\left[B_{[\sigma(i)\sigma(j)\sigma(k))]}\left(\ell,\ell,\ell\right),\star\right] \quad (3.4)$$

σ being a generic permutation of the tomographic indices and $\star$ being another observable. In the example (3.4) the tomographic indices will respectively obey the relation $i \leqslant j$, $i \leqslant k$, $j \leqslant k$ for $\ell_i = \ell_j$, $\ell_i = \ell_k$, $\ell_j = \ell_k$ respectively (isosceles configurations) and $i \leqslant j \leqslant k$ for $\ell_i = \ell_j = \ell_k$(equilateral configurations).

3.1.2 Survey specificities

Euclid-like tomography While all the analytical results in this manuscript do apply to any weak lensing survey, we now switch to the specificties of the Euclid mission. To this end, we make use of the requirements presented in Laureijs et al. [2011]. For the sky coverage, we use $\Theta_{\rm sky} = 1.29$ rad, $\Omega_{\rm sky} = 15.000$ deg$^2 \approx 4.57$ sterad, $f_{\rm sky} = 0.36$. For the photometric properties of the survey, we use a total comoving number of observed sources $n_{\rm tot} = 30$ gal arcmin^{-2} from $z_{\rm min} = 0.001$ up to $z_{\rm max} = 2.5$. The expected distribution of the sources is

$$n\left(z\right) \propto \left(\frac{z}{\bar{z}}\right)^2 e^{-\left(\frac{z}{\bar{z}}\right)^{1.5}} \quad (3.5)$$

with $\bar{z} = 0.9/\sqrt{2}$ and it will be normalised over the observed range. The sources will be split up to a maximum of 10 equipopulated redshift bins with extrema: $\{0.001, 0.418, 0.560, 0.678, 0.789, 0.900, 1.019, 1.155, 1.324, 1.576, 2.500\}$. It is worth mentioning at this stage that the power spectrum measurements for an actual survey are also affected by intrinsic shape noise due to the finite number of sources and the intrinsic variability of galaxy shapes. Assuming

[1] This range is in line with present Euclid forecasts [Laureijs et al., 2011] and in particular it has been chosen such that the effects of baryonic feedback on the lensing power spectrum are minimised [Semboloni et al., 2011; Kitching & Taylor, 2011]

that the orientation of intrinsic galaxy shapes is random and the shapes of different galaxies are uncorrelated, this component is Gaussian[2] and we account for this effect in the following way

$$P_{ij}(\ell) \to P_{ij}(\ell) + \frac{\sigma_\varepsilon^2}{\bar{n}_{(i)}} \delta_{ij}^K. \tag{3.6}$$

In particular, we will account for this contribution only at the level of covariance since it is possible to remove it from the data. We will use the value $\sigma_\varepsilon = 0.3$ which is representative of the expected Euclid sample [Amendola et al., 2013] and we will ignore weak lensing systematics like intrinsic alignments [Schneider & Bridle, 2010; Joachimi et al., 2013; Troxel & Ishak, 2015; Kiessling et al., 2015; Blazek et al., 2017; Schmitz et al., 2018], photometric redshift errors [Ma et al., 2006], blending [Hartlap et al., 2011] and point spread function (PSF) [Jarvis & Jain, 2008; Hamana et al., 2013; Chang et al., 2013]. In Eq. (3.6), $\bar{n}_{(i)}$ is the expected projected number of sources per unit of solid angle within the i^{th} redshift bin.

Window function In order to compute the super-sample covariance (2.228) we also need to specify the shape of the mask function $W(r, \theta, \varphi)$. This function binds the total volume of the Universe accessible for the data collection, thus defining the variance of the long mode δ_b. We will assume that it can be factorised into an angular component $W_{\text{an.}}(\theta, \varphi)$ and into a radial component $W_r(r)$. Both components can take values either 1 or 0 respectively within and outside the detected volume (binary window function). The Fourier transform of such a cylindrical mask, which is a top-hat in both redshift and angular coordinates, is given in flat sky by [Lima & Hu, 2007; Hu & Kravtsov, 2003; Aguena & Lima, 2018]

$$W_{\text{th}}(\mathbf{k}, \delta_\omega, \hat{\omega}) = 2 \exp\left(i\, k_\parallel \hat{\omega}\right) j_0\left(\frac{1}{2} k_\parallel \delta_\omega\right) \frac{J_1\left(k_\perp \hat{\omega}\, \Theta_{\text{sky}}\right)}{k_\perp \hat{\omega}\, \Theta_{\text{sky}}} V \tag{3.7}$$

where the vector $\mathbf{k}$ is split in $\{\mathbf{k}_\perp, k_\parallel\}$. These components are respectively associated to the coordinates on the plane orthogonal to the line of sight ($\mathbf{k}_\perp$) and to the comoving distance to it ($k_\parallel$). More precisely, Eq. (3.7) is the Fourier transform of the selection function for a comoving cylindrical volume of depth δ_ω, centred in $\hat{\omega}$ and derived under the assumption of a slowly varying ω and Hubble factor $H(\omega)$ within δ_ω. The angle Θ_{sky} (in radian) stands for the linear angular coverage of the survey. For sake of completeness, we underline that in Eq. (3.7) we are omitting the component related to the photometric error. The special function j_0 and J_1 are respectively the 0^{th} order spherical Bessel function and the 1^{st} order Bessel function of the first kind. The super-sample covariance requires to compute the variance $\sigma_v^2(\omega)$ of the linear matter field at a specific redshift. Then, we will compute $\sigma_v^2(\omega)$ over a disk-like volume, obtained under the limit $\delta_\omega \to 0$, $\hat{\omega} \to \omega$ of Eq. (3.7)

$$\sigma_v^2(\omega) = 4 \int \frac{d^2\varepsilon_\perp}{(2\pi)^2} \left[\frac{J_1\left(\varepsilon_\perp \omega\, \Theta_{\text{sky}}\right)}{\varepsilon_\perp \omega\, \Theta_{\text{sky}}}\right]^2 P^{\text{lin.}}(\varepsilon_\perp, \omega). \tag{3.8}$$

[2] A more realistic case would have a binomial shot noise, leading to a non vanishing bispectrum component. We leave this consideration to future works.

3.1.3 Halo model matter polyspectra

In Secs. 1.3.4-1.3.5 we described the general features of the chosen matter clustering model: the halo model. We gave a thorough review about the precision requirements as not to degrade the information content in modern galaxy surveys. We concluded Sec. 1.3.5 introducing the general framework for building the matter correlation functions at a desired order and we postponed the derivation of the detailed expressions for our specific case. We want now to tackle this point along with the approximations considered to reduce the computational burden.

The evaluation of the data vector $\mathbf{D}$ (3.1) and of the associated covariance (2.99)-(2.101) requires the computation of the correlation functions (2.85) up to the 6-point one, plus the responses of the observables to the background modes for the super-sample terms. In the next three paragraphs we will briefly derive the halo model expressions necessary for this evaluation. We anticipate that for the observables in the vector $\mathbf{D}$, we will use all the required n-halo terms, both for the power spectrum (1- and 2-halo) and the bispectrum (1-, 2- and 3-halo). Instead, for the covariance, we will approximate the polyspectra of order larger than 2 with the 1-halo component. We will give motivations for this choice in the related paragraph.

The data vector In Sec. 1.3.5 we saw that within the halo model framework we can compute the matter polyspectra as sum of different terms corresponding to different point allocations within one or more halos. While the value of the matter density at a specific point in space is related to the halo properties via the halo density profile (1.162), its statistical properties do not reflect directly those of the halo number density. We formally expressed this concept with the bias expansion (1.150) and we will use it to relate the correlations between halos to those of the underlying matter field. We will describe the power spectrum and the bispectrum at leading order in the linearly-evolved matter density field. In terms of the aforementioned bias expansion, this requires terms up to the second power in δ. In this scenario, two local biases $b_1(m,t), b_2(m,t)$ [Fry & Gaztanaga, 1993] and a non-local bias term $b_{s_2}(m,t)$ [Chan et al., 2012; Baldauf et al., 2012] are required

$$\delta_{\rm h}(\mathbf{k},t,m) = b_1(m,t)\,\delta(\mathbf{k},t) + \frac{b_2(m,t)}{2}\int\frac{d^3\mathbf{q}}{(2\pi)^3}\delta(\mathbf{q},t)\,\delta(\mathbf{k}-\mathbf{q},t) + \frac{b_{s_2}(m,t)}{2}\int\frac{d^3\mathbf{q}}{(2\pi)^3}\delta(\mathbf{q},t)\,\delta(\mathbf{k}-\mathbf{q},t)\,S_2(\mathbf{q},\mathbf{k}-\mathbf{q})\,, \quad (3.9)$$

with [Baldauf et al., 2012]

$$S_2(\mathbf{k}_1,\mathbf{k}_2) = \frac{(\mathbf{k}_1\cdot\mathbf{k}_2)^2}{k_1^2k_2^2} - \frac{1}{3}, \qquad b_{s_2}(m,t) = -\frac{2}{7}(b_1(m,t)-1)\,. \quad (3.10)$$

The local biases $b_1(m,t)$ and $b_2(m,t)$ can be predicted via the peak-background split (1.152) from the mass function $f_{\rm m}(m)$ (1.129). [Kaiser, 1984; Bardeen et al., 1986; Cole & Kaiser, 1989; Mo & White, 1996]. In the following, we will imply the time dependence.

Let us start from the power spectrum. We saw in Chapter 1 that it is well approximated at every redshift (a de-

pendence that we will omit in the following) as the sum of the 1-halo term (1.169) and of the 2-halo term (1.170) which respectively captures the contribution given by the two points being in the same halo and in two separate ones. If we focus on the 1-halo term in Fourier space, the convolution between the two halo profiles (1.169) (more precisely of the two normalised profiles $u(\mathbf{x}-\mathbf{x}_i;m)$) will simply transform into a product of the two Fourier-transformed profiles $\tilde{u}(m,\mathbf{k})$. Then, the 1-halo term can be derived with no particular effort

$$P^{1h}(k) = \int_{m^{Min}}^{m^{Max}} dm \left(\frac{m}{\rho_{com.}}\right)^2 f_m(m) \int_{c_v^{Min}}^{c_v^{Max}} dc_v\, p(c_v,m)\, \tilde{u}^2(m,c_v,k)\,. \tag{3.11}$$

In Eq. (3.11) we also accounted for the uncertainty on the halo concentration parameter via the convolution with the probability density distribution $p(c_v,m)$ given in Eq. (1.163). Also, the dependence on solely the magnitude k of $\mathbf{k}$ derives from the symmetries of the chosen NFW profile (1.162). The 2-halo term describes the correlation between 2 points hosted in two different halos of mass (e.g.) m_1 and m_2. Then it will depend on the halo-halo correlation function. Qualitatively we write

$$\langle \delta_h(\mathbf{k}_1,m_1)\,\delta_h(\mathbf{k}_2,m_2)\rangle \approx b_1(m_1)\,b_1(m_2)\langle\delta(\mathbf{k}_1)\,\delta(\mathbf{k}_2)\rangle \approx b_1(m_1)\,b_1(m_2)\,P^{lin.}(k_1)\,\delta(\mathbf{k}_1+\mathbf{k}_2)\,. \tag{3.12}$$

In Eq. (3.12) we stopped the bias expansion (1.150) at the linear level, i.e. $\sim (b_1(m_2)\,\delta)$, since it is a common assumption to consider the scales here involved (beyond the varialization halo radius) in the linear regime. For this reason, we will neglect quadratic corrections and the power spectra are the linear ones. A more quantitative evaluation leads to

$$P^{2h}(k) = \left[\int_{m^{Min}}^{m^{Max}} dm\, b_1(m)\,\frac{m}{\rho_{com.}}\,f_m(m)\left(\int_{c_v^{Min}}^{c_v^{Max}} dc_v\, p(c_v,m)\,\tilde{u}(m,c_v,k)\right)\right]^2 P^{lin.}(k) \tag{3.13}$$

and the total matter power spectrum will simply be the sum of the two terms derived above

$$P^{HM}(k) = P^{1h}(k) + P^{2h}(k)\,. \tag{3.14}$$

We refer to the dedicated section in Chapter 1 for the meaning of the different quantities appearing in the above relations. To simplify the expression for the general matter polyspectrum, we can introduce the following quantity

$$\mathrm{I}^{\beta}_{\mu}(k_1,\ldots,k_\mu) = \int_{m^{Min}}^{m^{Max}} dm\, b_\beta(m)\left(\frac{m}{\rho_{com.}}\right)^{\mu} f_m(m)\left(\int_{c_v^{Min}}^{c_v^{Max}} dc_v\, p(c_v,m)\left[\prod_{i=1}^{\mu}\tilde{u}(m,c_v,k_i)\right]\right) \tag{3.15}$$

where $b_o \equiv 1$. The, the matter power spectrum (3.14) can be written in a more synthetic way as

$$P(k) = \mathrm{I}^0_2(k,k) + [\mathrm{I}^1_1(k)]^2\,P^{lin.}(k)\,. \tag{3.16}$$

Thanks to this lighter notation, we can introduce the three-dimensional matter bispectrum as the sum of 3 multi-halos terms [Cooray & Hu, 2001]: 1) all the points are within the same halo (1-halo term, B^{1h}), 2) 2 out of

3 points are in the same halo while the third is not (2-halo term, B^{2h}) and 3) each point is hosted in a different halo (3-halo term, B^{3h}). In analogy to the calculations performed at the level of power spectrum, the 1-halo term for the bispectrum will simply be the convolution of the halo mass function with the third power of the halo profile

$$B^{1h}\left(k_1,k_2,k_3\right)=\int_{m^{\mathrm{Min}}}^{m^{\mathrm{Max}}}dm\,\left(\frac{m}{\rho_{\mathrm{com.}}}\right)^3 f_{\mathrm{m}}\left(m\right)\left(\int_{c_{\mathrm{v}}^{\mathrm{Min}}}^{c_{\mathrm{v}}^{\mathrm{Max}}}dc_{\mathrm{v}}\,p\left(c_{\mathrm{v}},m\right)\left[\prod_{i=1}^{3}\tilde{u}\left(m,c_{\mathrm{v}},k_i\right)\right]\right)=\mathrm{I}_3^{\mathrm{o}}\left(k_1,k_2,k_3\right). \tag{3.17}$$

The 2-halo term will qualitatively depend on the matter density field as in Eq. (3.12) where a quantitative derivation leads to

$$B^{2h}\left(k_1,k_2,k_3\right)=\mathrm{I}_1^1\left(k_1\right)\mathrm{I}_2^1\left(k_2,k_3\right)P^{\mathrm{lin.}}\left(k_1\right)+\mathrm{I}_1^1\left(k_3\right)\mathrm{I}_2^1\left(k_1,k_2\right)P^{\mathrm{lin.}}\left(k_3\right)+\mathrm{I}_1^1\left(k_2\right)\mathrm{I}_2^1\left(k_3,k_1\right)P^{\mathrm{lin.}}\left(k_2\right). \tag{3.18}$$

Finally, the 3-halo term will have the following dependence

$$\begin{aligned}&\left\langle\delta_{\mathrm{h}}\left(\mathbf{k}_1,m_1\right)\delta_{\mathrm{h}}\left(\mathbf{k}_2,m_2\right)\delta_{\mathrm{h}}\left(\mathbf{k}_3,m_3\right)\right\rangle\approx\\&\quad\left\langle\left(b_1\left(m_1\right)\delta\left(\mathbf{k}_1\right)+\frac{b_2\left(m_1\right)}{2}\delta\left(\mathbf{q}\right)\cdot\delta\left(\mathbf{k}_1-\mathbf{q}\right)+\frac{b_{s_2}\left(m_1\right)}{2}\delta\left(\mathbf{q}\right)\cdot_s\delta\left(\mathbf{k}_1-\mathbf{q}\right)\right)\right.\\&\quad\left(b_1\left(m_2\right)\delta\left(\mathbf{k}_1\right)+\frac{b_2\left(m_2\right)}{2}\delta\left(\mathbf{q}\right)\cdot\delta\left(\mathbf{k}_2-\mathbf{q}\right)+\frac{b_{s_2}\left(m_2\right)}{2}\delta\left(\mathbf{q}\right)\cdot_s\delta\left(\mathbf{k}_2-\mathbf{q}\right)\right)\\&\quad\left.\left(b_1\left(m_3\right)\delta\left(\mathbf{k}_1\right)+\frac{b_2\left(m_3\right)}{2}\delta\left(\mathbf{q}\right)\cdot\delta\left(\mathbf{k}_3-\mathbf{q}\right)+\frac{b_{s_2}\left(m_3\right)}{2}\delta\left(\mathbf{q}\right)\cdot_s\delta\left(\mathbf{k}_3-\mathbf{q}\right)\right)\right\rangle\end{aligned} \tag{3.19}$$

where we wrote in a synthetic way the convolutions (3.9) via the operators $\cdot$ and $\cdot_s$ and $m_{1,2,3}$ are the masses of the three halos hosting the points used for computing the correlation. According to the Perturbation Theory (PT) paradigm (see Bernardeau et al. [2002] for a thorough review), we can expand the field $\delta\left(\mathbf{k}\right)$ at different orders

$$\delta\left(\mathbf{k}\right)=\sum_p\delta^{(p)}\left(\mathbf{k}\right), \tag{3.20}$$

each term of the expansion being proportional to the p^{th} power of the linearly-evolved initial over-density [Goroff et al., 1986; Makino et al., 1992; Jain & Bertschinger, 1996]. The leading non-vanishing term in Eq. (3.19) includes at least one mode at second order in PT $\delta^{(2)}$. Overall, the 3-halo term at leading order will be proportional to the 4^{th} power of the initial linearly-evolved contrast density field, i.e.

$$\begin{aligned}&\left\langle\delta_{\mathrm{h}}\left(\mathbf{k}_1,m_1\right)\delta_{\mathrm{h}}\left(\mathbf{k}_2,m_2\right)\delta_{\mathrm{h}}\left(\mathbf{k}_3,m_3\right)\right\rangle\big|_{4^{\mathrm{th}}}\approx\\&\qquad b_1\left(m_1\right)b_1\left(m_2\right)b_1\left(m_3\right)\left\langle\delta^{(1)}\left(\mathbf{k}_1\right)\delta^{(1)}\left(\mathbf{k}_2\right)\delta^{(2)}\left(\mathbf{k}_3\right)\right\rangle+\text{cycles over }\left\{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3\right\}.\end{aligned} \tag{3.21}$$

The correlation (qualitatively) $\left\langle\delta^{(1)}\delta^{(1)}\delta^{(2)}\right\rangle$ can be related to the tree-level PT matter bispectrum as computed in Sec. 1.3.3. By looking at Eq. (3.19), we can see that at the same order more components are present, specifically those including $b_2\delta^2$ and $b_{s_2}\delta^2$ with δ in the linear regime, i.e. $\delta=\delta^{(1)}$. Similarly to the computation of the tree-level

PT bispectrum in Sec. 1.3.3, these correlations of linear modes can be reduced via the Wick theorem in products of 2-point correlation functions as

$$\begin{aligned}
&\left\langle \delta_{\rm h}\left(\mathbf{k}_1, m_1\right)\delta_{\rm h}\left(\mathbf{k}_2, m_2\right)\delta_{\rm h}\left(\mathbf{k}_3, m_3\right)\right\rangle\big|_{4^{\rm th}} \approx \\
&\qquad b_1\left(m_1\right) b_1\left(m_2\right) b_{2,(s_2)}\left(m_3\right)\left\langle \delta^{(1)}\left(\mathbf{k}_1\right)\delta^{(1)}\left(\mathbf{k}_2\right)\delta^{(1)}\left(\mathbf{q}\right)\delta^{(1)}\left(\mathbf{k}_3-\mathbf{q}\right)\right\rangle + \text{cycles over } \{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3\} = \\
&\qquad b_1\left(m_1\right) b_1\left(m_2\right)\frac{b_{2,(s_2)}\left(m_3\right)}{2}\Big[\left\langle \delta^{(1)}\left(\mathbf{k}_1\right)\delta^{(1)}\left(\mathbf{q}\right)\right\rangle\left\langle \delta^{(1)}\left(\mathbf{k}_2\right)\delta^{(1)}\left(\mathbf{k}_3-\mathbf{q}\right)\right\rangle + \\
&\qquad\qquad \left\langle \delta^{(1)}\left(\mathbf{k}_1\right)\delta^{(1)}\left(\mathbf{k}_3-\mathbf{q}\right)\right\rangle\left\langle \delta^{(1)}\left(\mathbf{k}_2\right)\delta^{(1)}\left(\mathbf{q}\right)\right\rangle\Big] + \text{cycles over } \{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3\} \approx \\
&\qquad b_1\left(m_1\right) b_1\left(m_2\right) b_{2,(s_2)}\left(m_3\right) P^{\rm lin.}\left(k_1\right) P^{\rm lin.}\left(k_2\right) + \text{cycles over } \{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3\}
\end{aligned} \tag{3.22}$$

where we did not consider the contributions forcing $\mathbf{k}_i = \mathbf{k}_j,\ (i,j=1,2,3)$ since they are associated to degenerate triangular configurations in the original bispectrum. A detailed evaluation of the mass-integration over the halo profiles leads to

$$\begin{aligned}
B^{\rm 3h}\left(k_1,k_2,k_3\right) = &\prod_{i=1}^{3} I_1^1\left(k_i\right) B^{\rm PT}\left(k_1,k_2,k_3\right) + \\
&+ I_1^1\left(k_1\right) I_1^1\left(k_2\right) I_1^2\left(k_3\right) P^{\rm lin.}\left(k_1\right) P^{\rm lin.}\left(k_2\right) + 2 \text{ terms from cycles over } \{k_1,k_2,k_3\} + \\
&+ \frac{4}{7}\left[I_1^1\left(k_1\right) I_1^1\left(k_2\right) \mathrm{S}_2\left(\mathbf{k}_1,\mathbf{k}_2\right) P^{\rm lin.}\left(k_1\right) P^{\rm lin.}\left(k_2\right)\left(1 - I_1^1\left(k_3\right)\right) + 2 \text{ terms from cycles over } \{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3\}\right].
\end{aligned} \tag{3.23}$$

and we can recognise the tree-level PT power spectra and bispectrum induced by the correlations obtained in Eq. (3.22) and in Eq. (3.21) respectively. The total bispectrum will then be the sum of the terms derived above

$$B^{\rm HM}\left(k_1,k_2,k_3\right) = B^{\rm 1h}\left(k_1,k_2,k_3\right) + B^{\rm 2h}\left(k_1,k_2,k_3\right) + B^{\rm 3h}\left(k_1,k_2,k_3\right). \tag{3.24}$$

Comparison with simulations In line with the presentation of Sec. 1.3.3, we propose a brief summary of the most up-to-date comparisons between the bispectrum halo model predictions and simulations. In particular, we refer to two thorough works, Fosalba et al. [2005] and Lazanu et al. [2016], where two different halo model implementations are considered. We remind that several attempts to improve the performance of the model are present in the literature [Takahashi et al., 2012; Valageas et al., 2013; Valageas & Nishimichi, 2011a,b; Mead et al., 2015; Mohammed & Seljak, 2014; Seljak & Vlah, 2015]. However, they do rely either on relaxing one or more assumptions of the halo model itself or on sets of parameters which are to be fitted against simulations. We will not focus on them here, being interested on the halo model standard formulation. In Fosalba et al. [2005] the dependence on the upper limit $m^{\rm Max}$ within the mass integration (3.15) is analysed along with the impact of a differently chosen halo boundary $r_{\rm v}'$. These variations are considered on top of a classical setup with an NFW profile [Navarro et al., 1996], the mass function from Sheth et al. [2001] and the concentration parameter from Bullock et al. [2001]. Their best fiducial model is given by $m^{\rm Max} = 10^5\, M_\odot/h$ and $r_{\rm v}' = 1.3 r_{\rm v}$ and it matches the simulated

signal within the error bars for all the scales analysed ($0.05\,h\,\mathrm{Mpc}^{-1} \leqslant k \leqslant 10\,h\,\mathrm{Mpc}^{-1}$). Lazanu et al. [2016] tested their halo model implementation on a similar range of scales as in Fosalba et al. [2005] therefore a comparison between the two works is meaningful. The halo model setup from Lazanu et al. [2016] is slightly different from the previous one. More specifically, the mass function is the one from Tinker et al. [2010] and the concentration parameter is obtained from the Boloshi simulation [Klypin et al., 2011]. The results from Lazanu et al. [2016] show an overall good performance of the halo model in predicting the bispectrum from the simulations for all the scales considered ($0.04h\mathrm{Mpc}^{-1} \leqslant k \leqslant 8h\mathrm{Mpc}^{-1}$) at $z = 0$, regardless of the triangular configurations. A deviation of $\gtrsim 20\%$ can be seen at the transition scales, i.e. ($0.3\,h\,\mathrm{Mpc}^{-1} \leqslant k \leqslant 1\,h\,\mathrm{Mpc}^{-1}$). In particular this is a well known shortcoming of the halo model formalism [Cooray & Hu, 2001; Smith et al., 2008; Figueroa et al., 2012]. This lack of performance is exacerbated for higher redshifts where the halo model predictions underestimate the signal from the simulations by a factor of ~ 2 and ~ 2.5 at $z = 1$ and $z = 2$ respectively, around the same scales. We think that these discrepancies are of minor importance for the results of our work. As a matter of fact, the convergence observables required for our analyses are mainly projections of matter polyspectra at $z \lesssim 1$.

THE COVARIANCE MATRIX In the introduction to this section we anticipated that we will approximate the polyspectra of order higher than 2 within the covariance with the respective 1-halo term. Here we describe the reasons behind this choice and we will also provide the general expression for the 1-halo term. At the level of 4-,5- and 6-order polyspectra this is mainly motivated by the effort in reducing the computational cost of the implementation. As for the bispectra instead, which appear in the NGins,BB term (2.140), the impact of the 2- and of the 3-halo term in the covariance is complex. We will dig into this issue by considering the joint covariance matrix for a 1 bin tomography, i.e. assuming that all the sources are placed within a unique broad bin between redshift 0.001 and 2.5. The other photometric parameters and the angular binning given in Sec. 3.1.2 are kept unchanged. The bottom panels in Fig. 3.1.1 show the fractional impact on the correlations when the 2-halo term (central panel) and the 2-+3-halo term (right panel) are added on top of the 1-halo term alone. We conclude that these multi-halo terms mainly enhance the correlations 1) between power spectra and bispectra at linear scales and 2) between bispectra on squeezed configuration (i.e. $B(\ell_i, \ell_j, \ell_k)$ with $\ell_k \sim \ell_j \gg \ell_i$) and bispectra evaluated on large scales. In Fig. 3.1.1 we report few binned configurations to help following the discussion. The single number, i.e. (14), refers to the respective power spectrum configuration, i.e. the power spectrum evaluated on the largest (the 14^{th}) ℓ-bin. We now move to the top panels in Fig. 3.1.1. Here we show the correlation matrices when including in the NGins,BB term of the joint covariance the 1-halo term only (left panel), the 1-+2-halo term (central panel) and the 1-+2-+3-halo term (right panel). We can see that, when adding the 2-halo term and the 2-+3-halo term on top of the 1-halo term alone, some correlations are dramatically pushed towards 1. Specifically the problematic correlations are between bispectra on squeezed configurations (see definition above) and power spectra at linear scales. At numerical level this translates into extremely ill-conditioned matrices with determinant close to 0 along with the possibility of numerically evaluated negative eigenvalues. This is indeed the situation when the joint power spectrum-bispectrum covariance matrix is analysed. It is worth underlying that, even if we were capable to invert exactly the matrix, the presence of highly correlated observables would lead to a poor improvement of the

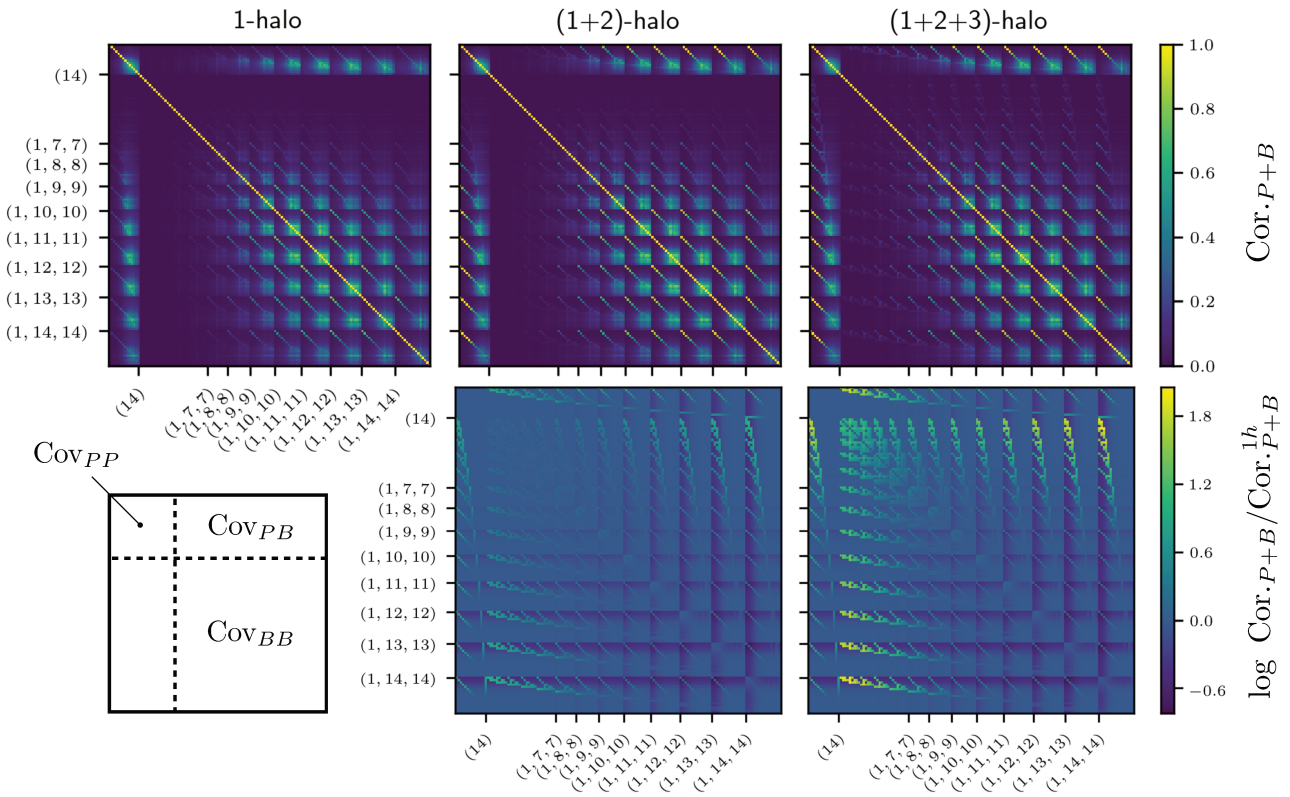

Figure 3.1.1: *Top-row*: joint correlation matrix for the weak lensing convergence power spectrum and bispectrum when approximating the bispectra required for the their computation via the 1- (left panel), 1-+2- (central panel) and 1-+2-+3-halo term (right panel). *Bottom-row*: log-fractional impact one the joint correlation matrix when adding respectively the 2- (central panel) and the 2-+3-halo (right panel) terms on top of the 1-halo alone. In both cases we employ the usual binning in $\log \ell$ (14 regularly spaced bins from $\ell = 10$ to $\ell = 5000$) and all the survey related parameter as chosen in Sec. 3.1.2. One tomographic bin is assumed, between $(0.001, 2.500)$. On the axes, we label few binned configurations to facilitate the analysis. The single number, i.e. (14), refers to the respective power spectrum configuration, i.e. the power spectrum evaluated on the largest (the 14^{th}) ℓ-bin. The left panel in the bottom row represents the overall structure of the covariances used in this work.

information content. In the simpler case of the bispectrum-only covariance, we were capable to perform a reliable analysis proving that the 2- and the 3-halo terms only affect the overall signal-to-noise ratio by a few percent. The physical intuition for such a tiny impact can be found in the poor signal-to-noise ratio for the configurations which are mainly empowered by these new terms. Specifically, these are all configurations where one multipole of the triangle is in the linear regime where the statistics of the matter field is mainly Gaussian. This is expected since the 2- and 3-halo term quantifies the correlations at scales larger than the virialization radius of halos. Then the bispectrum is either dominated by its own variance or highly correlated with the power spectra at those scales. Given this discussion, we will safely assume that the 1-halo approximation is sufficient for our purpose. High correlations and the subsequent numerical inaccuracies may eventually be healed via a different binning in ℓ or via the inclusion of higher-halo terms into the higher-order polyspectra involved. We leave these analyses for future works.

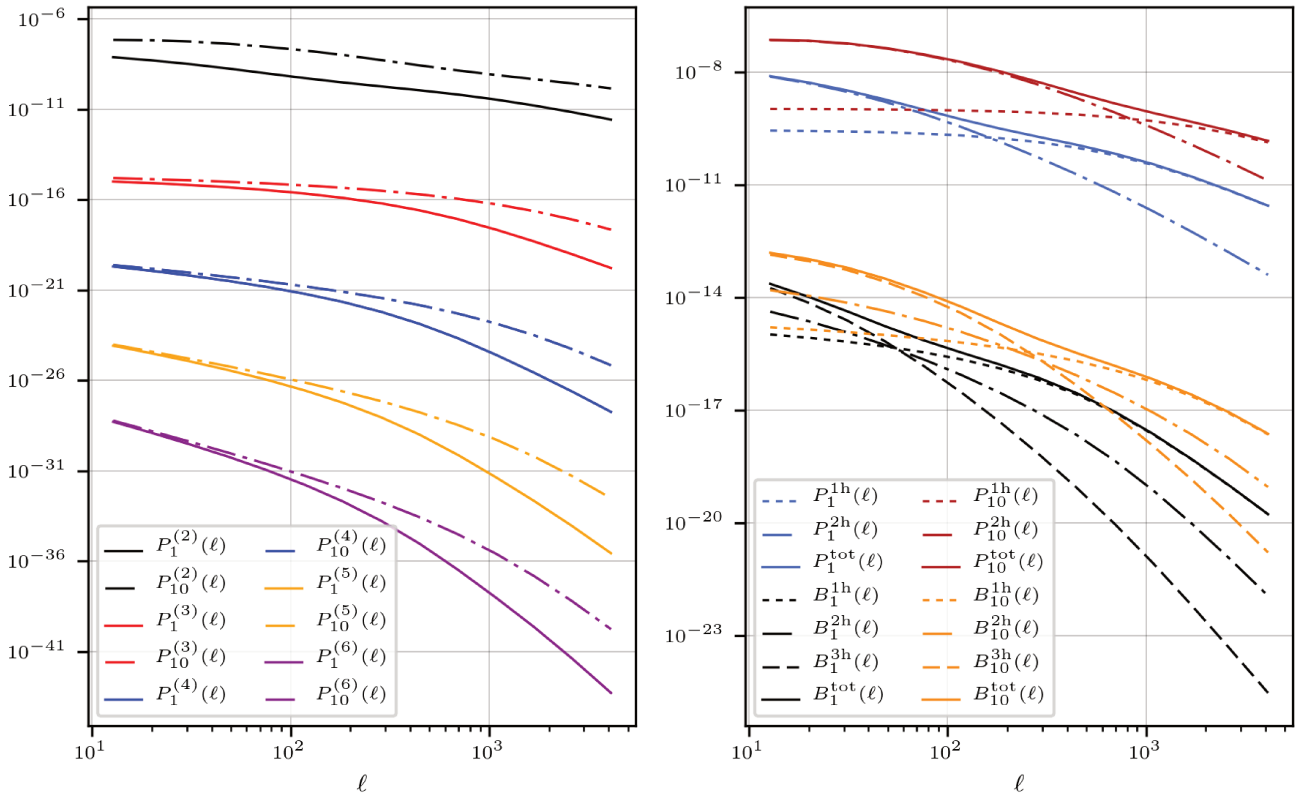

Figure 3.1.2: *Left:* auto-tomographic polyspectra of different orders (from 2 to 6) evaluated on equilateral configurations. We are considering auto-correlations for the 1st and the 10th tomographic bin, in a Euclid-like photometry (Sec. 3.1.2). *Left:* polyspectra employed in the computation of the covariance, i.e. power spectrum including the 1-+2-halo term and higher order polyspectra approximated to the 1-halo component. In the legend we are employing the notation $P_i^{(n)}(\ell) \equiv P_{i_1\dots i_n}(\ell,\dots,\ell)$ with $i_1 = \cdots = i_n$ and $\ell_i = \dots \ell_n$, the second member of the equivalence being introduced in Eq. (2.85). *Right*: polyspectra employed in the computation of the data vector: power spectrum including the 1-+2-halo term (we also depict the separate behaviour) and bispectrum including 1-+2-+3-halo term (we also depict the separate behaviour). In this panel we swtiched to the standard notation for the bispectrum, i.e. $B = P^{(3)}$.

Concluding, we give the expression for the 1-halo term of the general n-order polyspectrum. While in real space this type of correlation would require an n-dimensional convolution of halo profiles, the Fourier analysis makes it much simpler turning it into a one-dimensional (mass) integration of the product of n Fourier transformed halo profiles, i.e.

$$P^{1h}(\mathbf{k}_1,\dots,\mathbf{k}_n) = \mathrm{I}_n^0(k_1,\dots,k_n)\,, \tag{3.25}$$

where we employed the synthetic notation introduced in Eq. (3.15). Also, the dependence on the modules of the vectors derives from the fact that chosen single halo profile $u_{\rm NFW}(r,m)$ (1.162) is spherical in real space. As a summary, in Fig. 3.1.2 we show the behaviour of the polyspectra required for our analysis, both at the level of covariance (right panel) and at the level of data vector (left panel).

THE RESPONSES The halo model also provides a powerful recipe for the computation of the responses in Eqs.(2.208)-(2.214) required for the forward modelling of the super-sample covariance. We can obtain them by taking the derivative of the power spectrum (3.14) and of the bispectrum (3.24) with respect to the long mode δ_b (2.206). We will assume that the single halo profile is not affected by the long mode while it will impact the distribution of halos on larger scales via the bias and the mass function. Following Chan et al. [2018], we will consider the responses up to the linear bias (the second order ones were proved to be sub-leading in Chan et al. [2018] and we will neglect for consistency the tidal component of the response being also quadratic in the perturbation δ). We can formally express the impact of the long mode with the help of Eq. (3.15)

$$\begin{aligned}\frac{\mathrm{d}\mathrm{I}^{\beta}_{\mu}\left(k_1,\ldots,k_\mu|\delta_b\right)}{\mathrm{d}\delta_b}\Big|_{\delta_b=0} &= \int_{m^{\mathrm{Min}}}^{m^{\mathrm{Max}}} dm\left(\frac{m}{\rho_{\mathrm{com.}}}\right)^{\mu}\frac{\mathrm{d}}{\mathrm{d}\delta_b}\left[b_\beta\left(m\right)f_{\mathrm{m}}\left(m\right)\right]\left(\int_{c_{\mathrm{v}}^{\mathrm{Min}}}^{c_{\mathrm{v}}^{\mathrm{Max}}} dc_{\mathrm{v}}\,p\left(c_{\mathrm{v}},m\right)\left[\prod_{i=1}^{\mu}\tilde{u}\left(m,c_{\mathrm{v}},k_i\right)\right]\right)\\ &= \mathrm{I}^{\beta}_{\mu+1}\left(k_1,\ldots,k_\mu\right)\end{aligned} \tag{3.26}$$

where the derivative appearing within the expression above can be computed in the peak-background split approach as

$$b_\beta\left(m\right) = \frac{1}{f_{\mathrm{m}}\left(m\right)}\frac{\mathrm{d}f_{\mathrm{m}}\left(m\right)}{\mathrm{d}\delta_b}. \tag{3.27}$$

Therefore, the halo model power spectrum response will be related to the linear matter power spectrum one as

$$\frac{\mathrm{d}\,P^{\mathrm{HM}}\left(k|\delta_b\right)}{\mathrm{d}\delta_b}\Big|_{\delta_b=0} \approx \left[\mathrm{I}^1_1\left(k\right)\right]^2\frac{\mathrm{d}P^{\mathrm{lin.}}\left(k|\delta_b\right)}{\mathrm{d}\delta_b}\Big|_{\delta_b=0} + \mathrm{I}^1_2\left(k\right). \tag{3.28}$$

In Eq. (3.28) we neglected the term proportional to I^2_2 begin smaller than those already included [Chiang et al., 2014; Wagner et al., 2015b; Chan et al., 2018]. Under the same approximations, the bispectrum response can be expressed as

$$\begin{aligned}\frac{\mathrm{d}B^{\mathrm{HM}}\left(k_1,k_2,k_3|\,\delta_b\right)}{\mathrm{d}\delta_b}\Big|_{\delta_b=0} \approx\ & \mathrm{I}^1_1\left(k_1\right)\mathrm{I}^1_1\left(k_2\right)\mathrm{I}^1_1\left(k_3\right)\frac{\mathrm{d}B^{\mathrm{PT}}\left(k_1,k_2,k_3|\,\delta_b\right)}{\mathrm{d}\delta_b}\Big|_{\delta_b=0} +\\ &+\left[\mathrm{I}^1_1\left(k_1\right)\mathrm{I}^2_2\left(k_2,k_3\right)P^{\mathrm{lin.}}\left(k\right)+\mathrm{I}^1_1\left(k_1\right)\mathrm{I}^1_2\left(k_2,k_3\right)\frac{\mathrm{d}P^{\mathrm{lin.}}\left(k_1|\delta_b\right)}{\mathrm{d}\delta_b}\Big|_{\delta_b=0}\right] + \text{2 cycles} +\\ &+\mathrm{I}^1_3\left(k_1,k_2,k_3\right).\end{aligned} \tag{3.29}$$

Eq. (3.28) and Eq. (3.29) require the evaluation at tree-level in perturbation theory (PT) of the linear response for the power spectrum and for the bispectrum. The derivation is quite technical and we refer to Chan et al. [2018] for a detailed step-by-step explanation: it involves non-linear interactions between short (s) observed modes $\delta(\mathbf{k})$ and the long (l) mode $\tilde{\delta}_b\left(\mathbf{q}\right)$ (2.206). In the following we would like instead to qualitatively convey the main ideas behind this derivation, as not to break the flow of this chapter. To start with, we recall that in PT we describe these non-linear couplings between modes via the convolution kernels $\mathcal{F}^{(p)}_a$ defined in Eq. (1.98).

The tree-level power spectrum is simply given by the linear matter power spectrum, the two modes involved being in the linear regime. Therefore, to compute the tree-level linear response for the power spectrum, we will need to take its derivative with respect to the long mode δ_b when correlating modes described as follows (qualitatively)

$$\tilde{\delta}^{(1)}_{ls} = \tilde{\delta}^{(1)} + \left(\tilde{\delta}^{(1)} \cdot_{F^{(2)}} \tilde{\delta}_b\right). \tag{3.30}$$

Eq. (3.30) describes an observed short mode δ_{ls} obtained as a coupling between one long (l) un-observed mode (we are interested in the linear response) and the linear solution for the short (s) mode within the survey (we are interested in the tree-level solution). In particular the superscript in $\tilde{\delta}^{(1)}_{ls}$ refers to the order of the perturbative expansion in terms of short mode while the subscript refer to the type of coupling involved (long-short). By computing the power spectrum of the modes in Eq. (3.30) and by subsequently taking the derivative of it with respect to δ_b, it is possible to prove the following expression for the desired response

$$\frac{dP^{\text{lin.}}(k|\delta_b)}{d\delta_b}\Big|_{\delta_b=0} \approx \frac{47}{21}P^{\text{lin.}}(k) - \frac{1}{3}\frac{dP^{\text{lin.}}(k)}{d\ln k},. \tag{3.31}$$

Moving to the bispectrum, in Sec. 1.3.3 we derived the PT tree-level expression for it and we noticed that the non-vanishing leading order is given when the modes are, at least, at second order. Then, the tree-level solution for the linear response can be derived from the bispectrum of modes coupled as follows

$$\tilde{\delta}^{(2)}_{lls} = \tilde{\delta}^{(1)} + \left(\tilde{\delta}^{(1)} \cdot_{F^{(2)}} \tilde{\delta}_b\right) + \left(\tilde{\delta}^{(1)} \cdot_{F^{(3)}} \tilde{\delta}^{(1)} \cdot_{F^{(3)}} \tilde{\delta}_b\right). \tag{3.32}$$

In Eq. (3.32) we ordered the terms according to the power of the short (s) mode and once again the superscript and the subscript in $\tilde{\delta}^{(2)}_{lls}$ respectively refers to the order of the expansion in terms of the observed short mode and to the type of couplings involved (short-short-long). In the last term of Eq. (3.32) we introduced a symbolic notation defining the matter density contrast at third order, in perturbation theory,

$$\tilde{\delta}^{(3)}(\mathbf{k}) = 3\int \frac{d^3\mathbf{p}}{(2\pi)^3}\int \frac{d^3\mathbf{p}'}{(2\pi)^3}F^{(3)}(\mathbf{p},\mathbf{p}',\mathbf{k}-\mathbf{p}-\mathbf{p}')\tilde{\delta}^{(1)}(\mathbf{p})\tilde{\delta}^{(1)}(\mathbf{p}')\tilde{\delta}^{(1)}(\mathbf{k}-\mathbf{p}-\mathbf{p}')$$
$$\equiv \tilde{\delta}^{(1)} \cdot_{F^{(3)}} \tilde{\delta}^{(1)} \cdot_{F^{(3)}} \tilde{\delta}_b. \tag{3.33}$$

The third order kernel $F^{(3)}$ is given in the literature [Goroff et al., 1986]. We will not report its expression here so as not to break the flow of this qualitative explanation. A detailed calculation based on the derivative with respect to δ_b of the bispectrum for modes with couplings as those depicted in Eq. (3.32), leads to [Chan et al., 2018]

$$\frac{dB^{\text{PT}}(k_1,k_2,k_3|\delta_b)}{d\delta_b}\Big|_{\delta_b=0} \approx \frac{433}{126}B^{\text{PT}}(k_1,k_2,k_3) + \frac{5}{126}B_{G_2}(k_1,k_2,k_3) - \frac{1}{3}\sum_{i=1}^{3}\frac{dB^{\text{PT}}(k_1,k_2,k_3)}{d\ln k_i}. \tag{3.34}$$

By replacing the responses (3.31)-(3.34) into Eqs. (3.28)- (3.29), we obtain the desired expressions.

HALO MODEL NUMERICAL IMPLEMENTATION We would like to conclude this section by discussing more on a particular point: the numerical integration within Eq. (3.15). The main consequence of having a numerical integration over a finite domain is the exclusion of halos whose masses fall outside the range $(m_v^{\rm Min}, m_v^{\rm Max})$. This cut-off is intrinsically present in all known halo model implementations given that they are calibrated against simulations over a finite mass range. However, we do not expect this feature to have any impacts: very light halos, extending over extremely small scales, should not contribute to our cosmological observables while heavy halos are exponentially suppressed due to the shape of the mass function (1.129). The sensitivity to the mass cut-off is in particular problematic for the lower bound. We can easily see this point by looking at the integrand $\mathcal{I}_\beta(m)$ in the following consistency relations [Cooray & Sheth, 2002; Takada & Jain, 2003]

$$\int_{m^{\rm Min}}^{m^{\rm Max}} \frac{m}{\rho_{\rm com.}} dm\, b_\beta(m)\, f_{\rm m}(m) \equiv \int_{m^{\rm Min}}^{m^{\rm Max}} \mathcal{I}_\beta(m)\, dm = \begin{cases} 1 & \text{if } \beta \leqslant 1, \\ 0 & \text{if } \beta \geqslant 2, \end{cases} \tag{3.35}$$

where the case for $\beta = 0$ is actually the consistency relation for the mass function. In physical terms, the relations Eq. (3.35) requires that the mass of the Universe is entirely enclosed in halos ($\beta = 0$) of mass $m \in (m_v^{\rm Min}, m_v^{\rm Max})$ and that the overall distribution of halos is not biased compared to the total matter distribution ($\beta > 0$), at each order in the bias expansion (1.151). In the top-left panel of Fig. 3.1.3, we plot the integrand $\mathcal{I}_\beta(m)$ defined in Eq. (3.35) for $\beta = 0, 1, 2$ as function of the variable

$$\nu(m, z) \equiv \left(\delta_{\rm sp.}^c(z) / \sigma(m)\right)^2 \tag{3.36}$$

$\delta_{\rm sp.}^c$ and $\sigma^2(m)$ being respectively introduced in Eq. (1.131) and in Eq. (1.134). We see that the integrals we are trying to evaluate are slowly convergent for $\nu \to 0$ ($m \to 0$ implies $\sigma(m) \to \infty$). Therefore we would expect every halo model implementation to be very sensitive to extremely light halos, which contradicts the physical intuition described above and consequently undermines halo models themselves, not being testable at these small masses. The problem comes from the extrapolation of the model beyond its regime of validity assuming that arbitrarily small mass halos are present in our simulation accounting for the whole matter content and biases. Clearly Eq. (3.35) would not be satisfied since we are excluding these halos from the integration. However, once a minimum mass is set, lighter halos are excluded while the total matter budget may be filled by non-virialized dust. This contribution cannot be caught by the mass function $f_{\rm m}(m)$, by definition. In order to fulfil the normalisation constraints (3.35) we will simply assume that the non-virialised matter content provides an effective contribution via the following regularisation of the mass function and of the biases [Schmidt, 2016]

$$f_{\rm m}(m) \to f_{\rm m}(m) + a_0\, \delta_{\rm D}\left(m - m^{\rm Min}\right), \tag{3.37}$$

$$b_\beta(m) \to \begin{cases} b_\beta(m) & \text{if } m > m^{\rm Min}, \\ a_\beta & \text{if } m = m^{\rm Min}, \end{cases} . \tag{3.38}$$

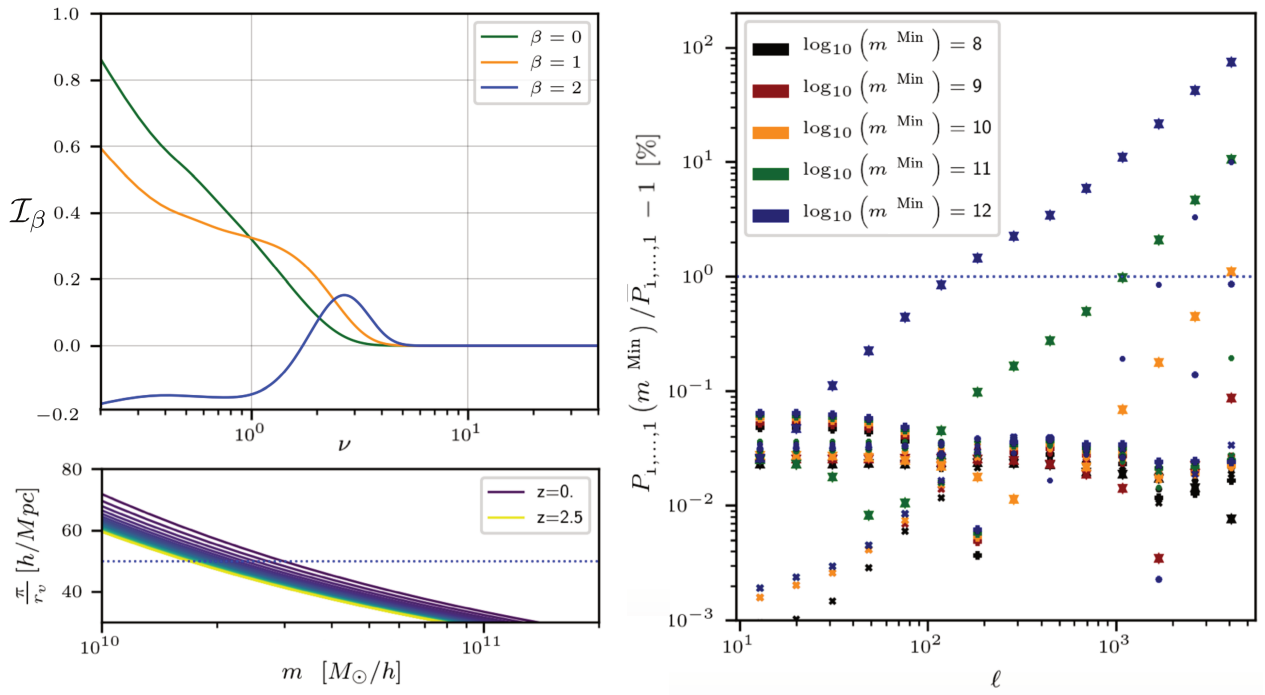

Figure 3.1.3: *Top-left:* integrand $\mathcal{I}_\beta$ in Eq. (3.15) as function of ν (see main text for definition) for $\beta = 0, 1, 2$. *Bottom-left:* inverse halo comoving radius as function of the virialised mass at different redshift. *Right:* impact on the polyspectra, from the order 2 to 6, from excluding light halos from the mass integrations required in their computation. The minimum mass considered in the different cases is given in the legend (and included as further dependence on the spectra) and we are not differentiating polyspectra of different orders since interested in the overall behaviour. We are considering auto-tomographic polyspectra for the first tomographic bin evaluated on equilateral configuration $P_{1,\dots,1}(\ell, \dots, \ell)$ (on the y-axis label we are omitting the dependence on the multipoles). We investigate the sensitivness of the polyspectra to the low-mass cut-off by computing the percentage fractional difference with a conservative case $\bar{P}_{1,\dots,1}(\ell, \dots, \ell)$, derived by integrating over halos of mass $m \geqslant 10^2 M_\odot/h$.

The parameters a_β (a_0) then are fixed in order to satisfy the consistency relations (3.35). We do not consider any corrections related to the upper bound m^{Max} because heavy halos are strongly suppressed by the mass function itself, as we can see from the top-left panel in Fig. 3.1.3. Our numerical integrals converge for $m^{\text{Max}} \approx 10^{16} M_\odot/\text{h}$. In the computation of the weak lensing observables, we pushed our line-of-sight integration up to $k_{\text{Max}} = 50h\,\text{Mpc}^{-1}$ [Kitching & Taylor, 2011]. Therefore, we do not expect halos whose comoving radius is smaller than this scale to be significant for our analyses. In the bottom-left panel in Fig. 3.1.3 we show the (inverse) comoving halo virialization radius as function of their virialised mass. The radius corresponding to the above scale of $k_{\text{Max}} = 50h\,\text{Mpc}^{-1}$ (indicated by the horizontal blue dotted line) encloses a mass $m^{k_{\text{Max}}} \approx 10^{10} M_\odot/h$ (with a small dependence on the redshift). Therefore, if our set-up (once applied the regularisation (3.37)) is consistent, our polyspectra should not be sensitive to halos lighter than $m^{k_{\text{Max}}}$. This is indeed what we prove with the right panel in Fig. 3.1.3. Here, we look at the impact on the polyspectra, from the order 2 to 6, from excluding light halos from the mass integrations required in their computation. The minimum mass considered in the different cases is given in the legend and we are not differentiating polyspectra of different orders since interested in the overall behaviour. Also, we are presenting this analysis with single points in order to have a cleaner figure. We are considering auto-tomographic

polyspectra for the first tomographic bin (being the ones more sensitive to light halos) evaluated on equilateral configuration (for simplicity), i.e. $P_{1,\dots,1}(\ell,\dots,\ell)$ (we are omitting the dependence on the multipoles on the y-axis label) according to the notation given in Eq. (2.85). We investigate the sensitivness of the polyspectra to the low-mass cut-off by comparing them with a conservative case $\bar{P}_{1,\dots,1}(\ell,\dots,\ell)$ which is derived by integrating over halos of mass $m \geqslant 10^2 M_\odot/h$. Our implementation is solid: we observe a first deviation by more than 1% when integrating over halos of mass $m \gtrsim 10^{10} = m^{k_{\rm Max}}$. We chose for the mass integrals $m^{\rm Min} = 10^9 M_\odot/h$ to avoid numerical inaccuracies with a negligible extra computational price. This is extremely interesting also because this resolution in mass can be easily achieved by present *smoothed-particle hydrodynamics* simulations [Hopkins et al., 2014; Genel et al., 2014; Khandai et al., 2015] allowing precise tests of the statistical properties of the halo model at an accuracy which is sufficient for the joint power spectrum-bispectrum analysis. In our specific case, we employed a halo model implementation based on the work of Bullock et al. [2001]. In their work, the distribution of halos was tested at the level of mass function with a simulation covering a mass range $10^{11} M_\odot/h - 10^{14} M_\odot/h$, from redshift $z = 40$ to the present.

3.1.4 Computational effort

Given the specificities outlined at the beginning of this section, we will compute covariance matrices with, at worst, $\gtrsim 10^4 \times 10^4$ elements. Such computational effort can be unsustainable for a single core computer, especially for a binning in ℓ finer than the one actually considered. Moreover, given the equations defining the covariances in Sec.s 2.2.3- 2.2.4, the number of required polyspectra for the intermediate calculations has to be multiplied by factors of order 1-10 if a naive approach is employed. Another important point to keep in mind is the numerical accuracy of our final products. We leave a more complex propagation of the error for future works and we arbitrary assume 1% numerical accuracy to be satisfactory for all our observables (vectors and covariances). However, this level of precision might not be enough to guarantee the numerical stability of our analysis for some extreme configurations. As a matter of fact, the covariance matrices we are dealing with tend to be extremely ill conditioned having eigenvalues that span over the same range in orders of magnitude as the observables in Fig. 3.1.2. On top of that, they are far from being diagonal dominated. For example, especially when the super-sample terms are included, we can reach cross-correlations > 0.9. The numerical inversion of these matrices is challenging and sometimes even impossible given that numerical round off errors might turn our covariances from positive defined to negative defined. For example, we found that 1% accuracy in the evaluation of the covariance matrix for the power spectrum is not enough for a reliable inversion and a target of 0.1% must be set when not including the shot noise. Luckily, if one is interested in analysing the information content in the most realistic cases where the shot noise term is accounted for, the inversion of the matrix becomes less an issue given that this component acts as an additional source of Gaussian noise (under our approximations) leading to more diagonal dominated covariances. This will be the regime in which we will work when dealing with the bispectrum analysis. Accuracy, high number of evaluations and fast production of the observables are achieved via a original and parallelised numerical implementation of the problem. In the following we will briefly go through the strategy used in our code divid-

ing it in different steps. Please note that the best performance of the algorithm is obtained on medium-large size clusters: for this reason we will assume a parallelised use in the following description. However, even if strongly discouraged, a serial run on a single core is allowed.

Covariance matrix evaluation

1. In a first serial part, each processor builds own tables carrying the information related to 1) the points in redshift $\{z_i\}_{i=1,\dots,n_z}$ for the evaluation of the matter polyspectra to be used within the line-of-sight integration (2.89) and to 2) the multipole binning $\{\ell_i^b\}_{i=1,\dots,n_\ell}$ to construct the configurations in the data vector. Since the computation time required for this part is negligible compared to the whole run, we prefer having each processor computing its own tables. We also build in each core two vectors of structures $\mathcal{J}_P$ and $\mathcal{J}_B$ which respectively represents the data vector **P** and **B**. The i^{th} structure within these vectors stores the tomographic positions of the sources and the central values of the bins for the computation of the corresponding observable. On top of that, we store at every redshift $\{z_i\}_{i=1,\dots,n_z}$ the required information on the distribution of halos: this is needed for building matter polyspectra. The tables are global and every core can have access to them. For this task, we rely on the routines written by Mead et al. [2015]. We want to underline that we are not exploiting the improved halo model implementation suggested in their work since it does not provide a recipe for higher order correlation functions. On the other hand, their fitting formulas are extremely advised for power spectrum analyses.

2. Given the number of cores n_{cores} available, the elements in the upper diagonal part of the covariance matrix (which is symmetric) are equally split among the processors: if the number of entries is not a multiple of n_{cores}, then the remaining is equally redistributed starting from the 0^{th} core.

3. At this point, we enter the parallelised part of our pipeline. Each core analyses the assigned part of the covariance (from step [2.]) performing the following operations on the different polyspectra required according to the expressions given in Secs. 2.2.3-2.2.4. More precisely, we have the cores working on the polyspectra order n by order n, from the power spectrum ($n = 2$) up to the 6-order polyspectrum ($n = 6$).

3.1. For every order n, the minimum number of non-equivalent spectra required for the computation of the covariance elements is identified. By looking at Eq. (2.89), our spectra are invariant under permutations of the tomographic indices. Also, the assumed homogeneity of the Universe ensures that the value of a given polyspectrum does not depend on the order of the momenta it depends on. Then, by non-equivalent polyspectra we do mean polyspectra up-to permutations of the indices referring to both the position of the sources and the bins in ℓ used to build the configurations. Also, the minimum number of non-equivalent spectra is identified considering that a specific polyspectrum is needed only if the associated configuration in Fourier space meets symmetry requirements: if these are not met, the core skips that spectrum.

3.2. In a second moment, the polyspectra to be computed are equally shared among the cores: they evaluate the matter polyspectra from the halo model ingredients computed in step [1.] and then they perform the line-of-sight integration (2.89). The resulting projected observables are stored in global tables, available to all the cores.

3.3. Every core goes again through the elements assigned at step [2.], computing them from the polyspectra available in the global tables from step [3.2.]. The value of every element of the covariance computed at this order is finally summed to the previous one, which had gone through the same procedure. If spectra of different orders need to be multiplied between each other (namely the power spectrum and the trisepctrum in the NGins,PT terms in Eq. (2.146)), we store them in global arrays whose life-time covers the computation of the polyspectra of order 2 and 4.

Given that our code analyses the covariance matrix order of correlation by order of correlation, it is very easy or us to analyse the impact of the different orders of correlations. Once the parallelised computation explained so far is finished, the vector **D**, produced as byproduct, and the covariance matrix are ready to be used for our studies.

Covariance analysis Once the covariance matrix has been computed via the steps outlined in the above paragraph, we desire to perform some operations on it. To start with, we will need to invert it in order to compute the signal-to-noise ratio defined in Eq. (3.1). Further, bearing in mind the idea of identifying possible summary statistics, we are also interested in a *principal component analysis* (PCA) of the covariance. Given the ill-conditioned nature of the joint covariance and its remarkable dimension, we will use the high performance library *Math Kernel Library* (MKL)[3] from *Intel* which does also allow for an automatic parallelization whenever several threads are available. While the first task is achieved with satisfactory precision via the function `dpotrs`[4] based on a Cholesky decomposition of the matrix, the PCA is more sensitive to round-off errors occurring in the computation of the smallest eigenvalues. We will see in the next section that these are actually the most important ones for the definition of summary statistics. At this second level we made use of Intel MKL Extended Eigensolver `dfeast_syev`[5] which guarantees a higher computational precision with the price of searching for eigenvalues in a specific range. Then, we will simply solve the eigen-problem over ranges of 1 order of magnitude iteratively. We found this technique being extremely efficient: all the eigenvalues identified are positive as theoretically expected and as numerically proved by the function `dpotrs`.

3.2 Analysis of the Information content in the weak lensing observables

The greatest achievement of the work presented in this manuscript is the forward modelling of remarkably big covariance matrices for the cosmological exploitation of the tomographic weak lensing convergence power spectrum-

[3] https://software.intel.com/en-us/mkl.
[4] https://software.intel.com/en-us/mkl-developer-reference-fortran-potrs.
[5] https://software.intel.com/en-us/mkl-developer-reference-fortran-extended-eigensolver-routines.

bispectrum joint probe. We accomplished this goal via design of an algorithm which will be a key element for future galaxy survey heritage. In order to deeply understand the actual benefit in terms of cosmological parameters, Fisher forecast [Tegmark, 1997], DALI forecast [Sellentin et al., 2014; Sellentin, 2015] or Markov Chain Monte Carlo (MCMC) analyses are advised: we leave these ambitious steps for future works. We underline that the covariance matrix computation is anyway a key step for all the techniques listed above. However, we would like to have an idea of the information content achievable by analysing the bispectrum on top of the power spectrum and have insights on the level of correlations between them. For a joint analysis we modify the signal-to-noise (S/N) ratio expression (3.1) to account for correlations up to a maximum binned measured angular multipole $\ell_{\rm max}$. The S/N as cumulative function of the maximum multipole included can be written as

$$\left(\frac{S}{N}\right)^2_D = \sum_{i,j}^{\ell(i),\ell(j)<\ell_{\rm max}} \mathrm{D}_i \left[C^D\right]^{-1}_{ij} \mathrm{D}_j, \tag{3.39}$$

where we need 1) the vector of observables (3.3) and 2) the associated covariance matrix. Schematically

$$\mathbf{D} = \{\mathbf{P},\, \mathbf{B}\}, \tag{3.40}$$

$$C^D = \begin{pmatrix} C^{PP} & C^{PB} \\ C^{PB} & C^{BB} \end{pmatrix}. \tag{3.41}$$

The matrices C^{PP}, C^{BB} and C^{PB}, are respectively the covariance of the power spectrum (2.99), bispectrum (2.100) and the cross-covariance between the two (2.101). When interested in accessing the cosmological information in the single probe, we will simply assume $\mathbf{D} = \mathbf{P}$ (resp. $\mathbf{B}$) and $C^D = C^{PP}$ (resp. C^{BB}). We will exploit the S/N 1) to understand how well the overall parameter space is constrained by our observations up to a given angular scale $\ell_{\rm max}$, 2) to test how much the uncertainties of our theoretical model can degrade the information content 3) to understand the impact of different approximations to the covariance matrices. On top of that we will use this quantifier 4) to verify the robustness of summary statistics when compared with the most complete analysis.

3.2.1 Power spectrum signal-to-noise ratio

In Fig. 3.2.1 we start our analysis from the simplest case: the power spectrum. We can see that, starting from a Gaussian covariance not contaminated by shot noise (dashed violet line), the main impact on the maximum information content is produced by accounting for the shot noise which degrades the S/N on all the scales. This is expected being a scale-independent contributions to the errors. We then compare the S/N as reconstructed from different approximations to the covariance, with and without the shot noise. We will follow the notation introduced in Eq. (2.99) labelling as G, NGins and NGssc respectively the Gaussian, the intra-survey and the super-survey component of the covariance. In both cases, and with respect to the Gaussian approximation, the biggest loss of information is induced by the super-sample covariance NGssc. Focusing on the analysis including

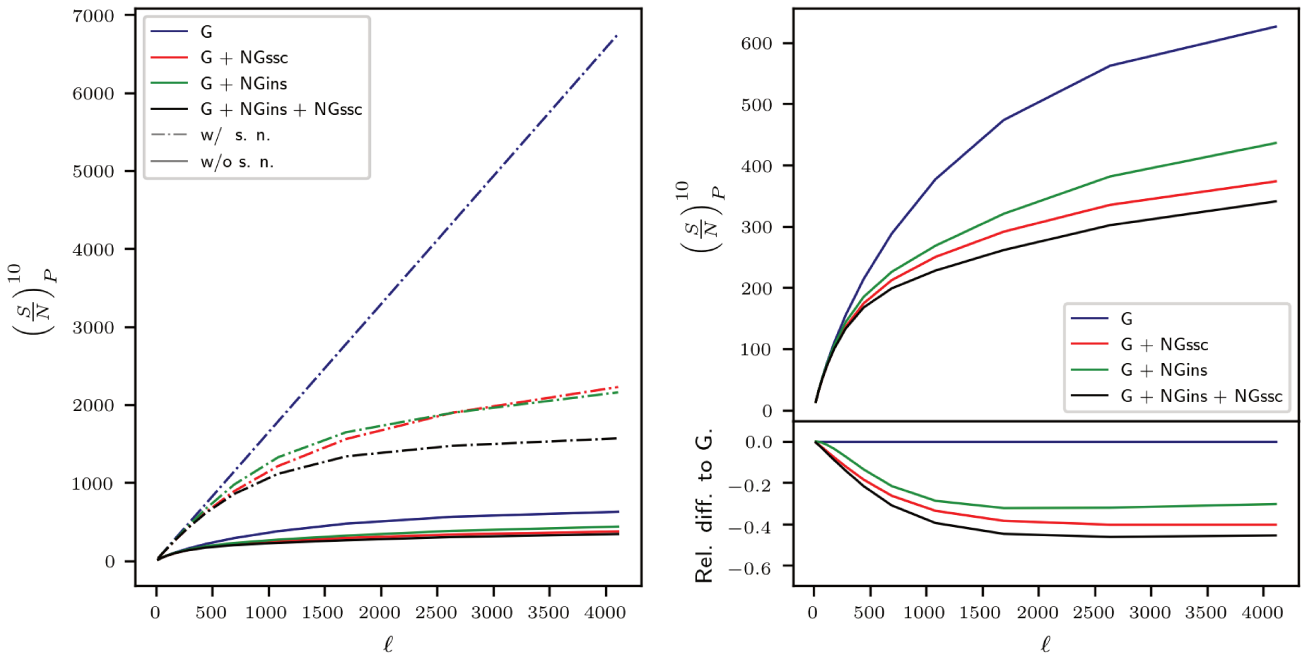

Figure 3.2.1: *Left:* S/N analysis for the power spectrum of the weak lensing convergence field when considering different contributions to the covariance matrix. To label them, we follow the notation given in Eq. (2.99). For every case, we plot 1) the signal assuming our observations not being contaminated by shot noise (dashed lines) and 2) the signal when the shot noise is considered (solid lines). *Top-right:* S/N analysis when the shot noise is considered, for different approximations to the covariance matrix. *Bottom-right:* fractional differences between the S/N (shot noise included) for a purely Gaussian covariance and the S/N as computed from different approximations to the covariance.

the shot noise (the more realistic case), the NGssc induces a 40% reduction on the maximum S/N. The in-survey component NGins (alone) leads to a degradation of about 30%. Finally, the combined effect of the two leads to a loss of about 45 %. In Fig. 3.2.2 we investigate at the level of power spectrum a first way to compress our data by looking at the behaviour of the maximum information content depending on the redshift binning of the sources. From the original 10 Euclid-like redshift bins[6], we test a first possible compression by binning the sources into respectively 5, 2 and 1 broader equipopulated intervals. In the case of the power spectrum analysis our data vector will have dimension 770, 210, 42 and 14 when considering 10, 5, 2 or 1 tomographic bins respectively. For all the possible approximations to the covariance, we see that the maximum information content decreases (as expected) while moving to smaller numbers of bins. Also, the way the cumulative S/N is affected by the binning is independent on which covariance approximation is considered. The maximum S/N in the 3 situations proposed in Fig. 3.2.1 (from left to the right), is reduced by roughly 1%, 5% and 20% when compared to the 10 bin case, whose maximum value is represented by red lines within the 3 panels on the right. Therefore, this analysis shows that a forecast based on a 5 bin tomography will not spoil our knowledge on the cosmological parameters for more than 1 % while simplifying a lot our computation. Indeed, the vector of power spectra in the 5 and 10 bins cases is

[6]In terms of the redshift extrema: {0.001, 0.418, 0.560, 0.678, 0.789, 0.900, 1.019, 1.155, 1.324, 1.576, 2.500}.

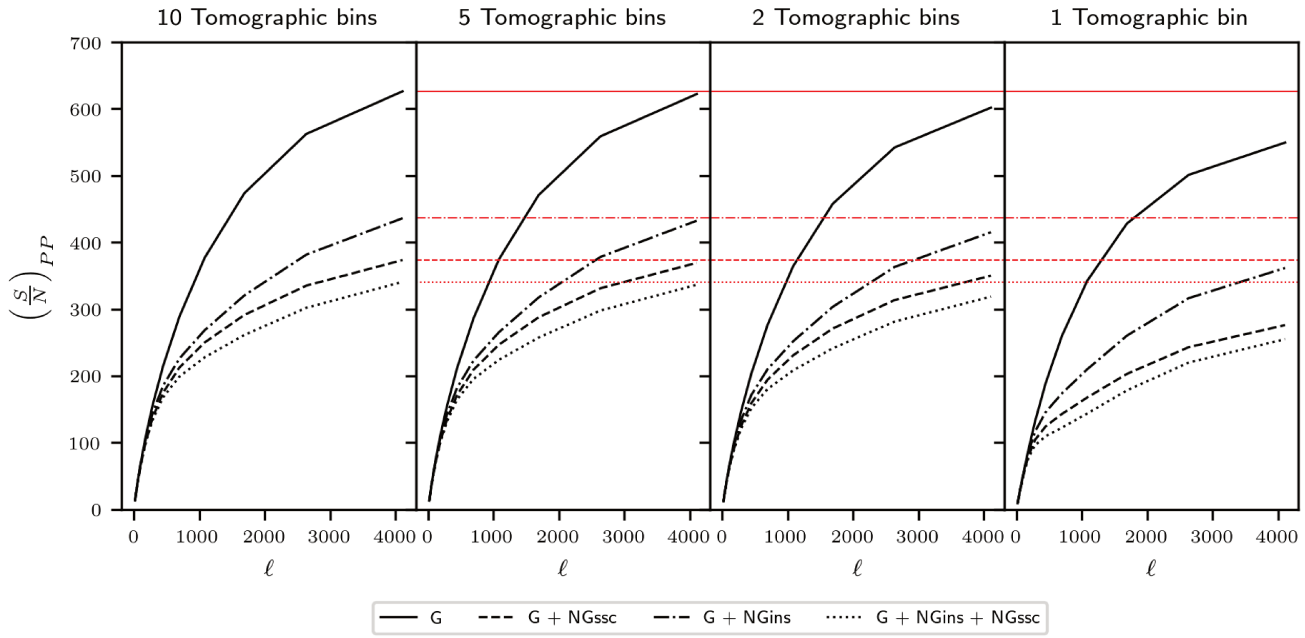

Figure 3.2.2: S/N analysis for the power spectrum of weak lensing convergence when considering different approximation to the covariance matrix. For them, we will follow the notation introduced in Eq. (2.99). Also, we allow for different binning in redshift of the expected sources. The number of bins indicated on the title to each panel are constructed as equipopulated, starting from the survey specificities listed in Sec. 3.1.2. As for the angular multipole, we keep our usual binning in $\log l$ (14 regularly spaced bins from $l = 10$ to $l = 5000$) and all the survey related parameters as chosen in Sec. 3.1.2. Specifically, going from the second panel on the left to the last panel on the right, the choice of tomographic bins is (in terms of their upper and lower limits) : $(0.001, 0.560, 0.789, 1.019, 1.324, 2.500)$ [5 bins case], $(0.001, 0.900, 2.500)$ [2 bins case] and $(0.001, 2.500)$ [1 bin case].

made of 210 and 770 elements respectively, so that the 5 bins case gives a $\sim$ 3.7 reduction of the data vector size and a $\sim$ 13.4 reduction of the covariance matrix size.

3.2.2 Bispectrum signal-to-noise ratio

In Fig. 3.2.3 we show for the bispectrum alone a similar analysis as the one proposed in Fig. 3.2.2 for the power spectrum. On top of that, we are also considering different approximations to the bisepctra in the data vector **B**, namely excluding different multi-halo contributions to its description. By comparing the solid lines (no approximations made on the data vector), we can inspect instead the impact of different approximations to the covariance matrix. Also, we investigate the possibility to compress the data vector via re-binning of the sources.

Starting from the possibility of data compression, we can see that a forecast based on 5 equipopulated redshift bins still allows to recover the full information content for a Euclid-like survey. In the case of the bispectrum, the data vector **B** has dimension 72280, 9790, 776 and 130 when considering 10, 5, 2 or 1 tomographic bins respectively. The 5 bin tomography then allows for a reduction of the size of the data vector and of the covariance matrix

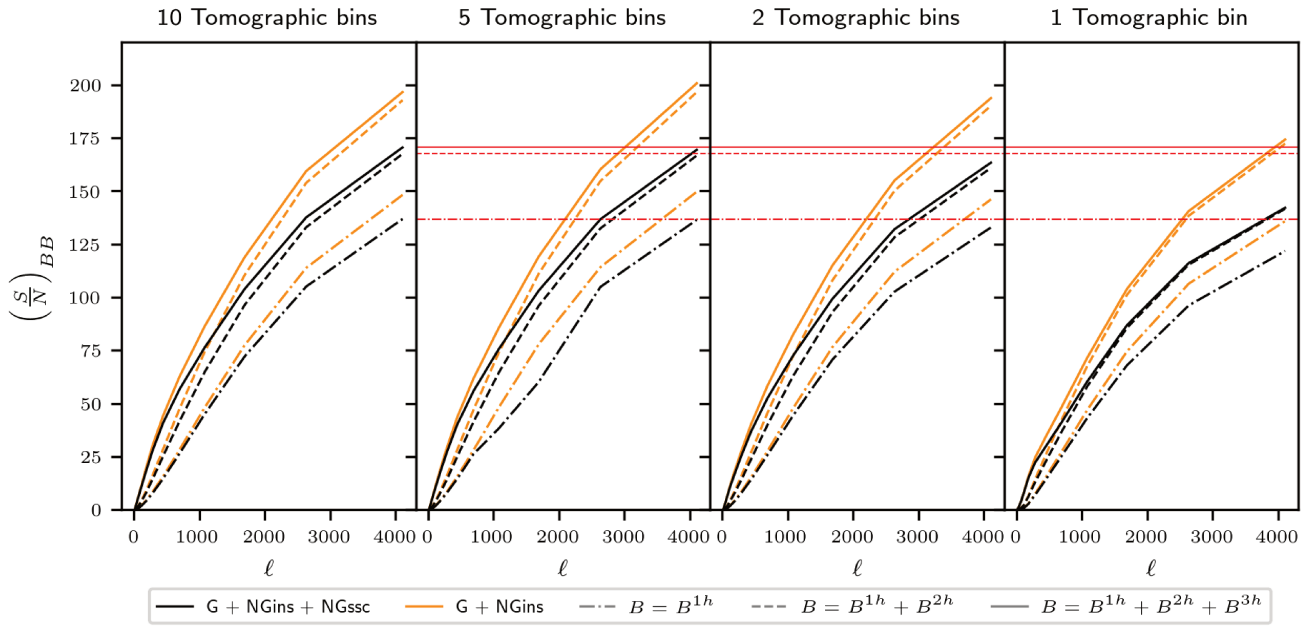

Figure 3.2.3: S/N analysis for the bispectrum alone of weak lensing convergence when considering different approximations to the covariance matrix. We are here inspecting different effects. As in Fig. 3.2.2, we are looking 1) at the impact of different approximations to the covariance matrix and 2) at the behaviour of the information content as function of the binning of the sources in redshift. For the first point, we label the different components of the covariance matrix according to the notation introduced in Eq. (2.100). On top of that, we study the impact of approximating the data vector $\mathbf{B}$ by neglecting different multi-halo configurations. In this analysis, we employ the same binning in ℓ and in redshift as for the respective panels in Fig. 3.2.2. The shot noise is included in all the forecasts.

by a factor ~ 7.4 and ~ 54.8 respectively.

Secondly, we analyse the impact on the maximum S/N of different approximations both at the level of covariance (while not having any approximations on the data vector $\mathbf{B}$ - solid lines) and at the level of data vector $\mathbf{B}$ (while keeping the whole covariance - lines of the same colour). Independently of the tomographic analysis, neglecting the 3-halo term leads to a very small effect on the information content at all scales ($\sim 1\%$). The analysis is instead more sensitive to the 2-halo term. The super-sample covariance clearly leads to a major degradation of the information content in the bispectrum. By comparing the analyses when all the multi-halo configurations are included in the bispectra (solid lines), we can see that the super-sample covariance reduces the information content by $\sim 20\%$ for the case of 1 bin tomography, $\sim 15\%$ for the 2 and 5 bins cases and of $\sim 13\%$ for the 10 bins tomography.

3.2.3 Joint signal-to-noise ratio

In Fig. 3.2.4, we finally move to the joint study of the probes. A joint analysis allows to improve the forecast by accessing the cosmological information that was lost in mode couplings. Motivated by our previous findings, we

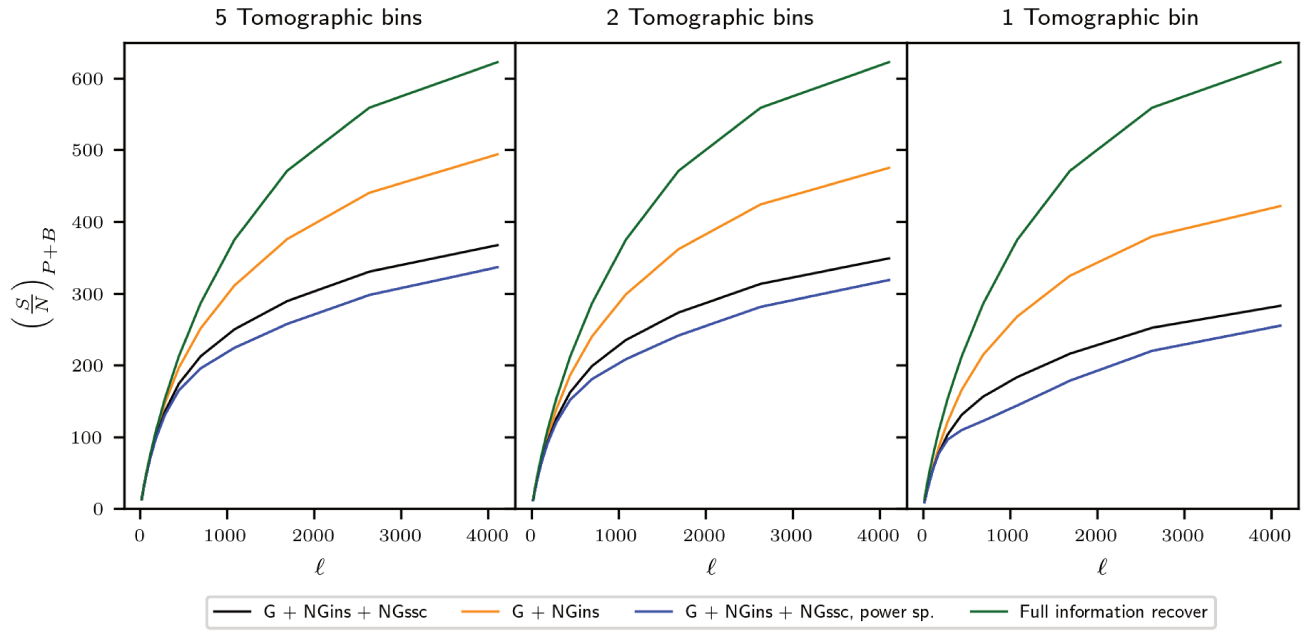

Figure 3.2.4: S/N analysis for the weak lensing convergence power spectrum-bispectrum joint probe. We compare the analyses when including (black solid line) and excluding (yellow solid line) the NGssc component. The G and NGins components are kept in both cases. We are including all the required multi-halo configurations in the bispectrum data vector $\mathbf{B}$. As in Fig. 3.2.2, we are also focusing on the behaviour of the information content as function of the binning of the sources in redshift. We employ the same bins in ℓ and the same survey properties as in the three respective panels in Fig. 3.2.2 and in Fig. 3.2.3. In all the three panels we also depict the cumulative S/N for the power spectrum probe, when including all the components to the associated covariance. The green solid line represent the hypothetical case of a pure Gaussian weak lensing convergence field contaminated with shot noise.

restrict the joint analysis to 5 tomographic bins. In the case of the joint power spectrum-bispectrum S/N our data vector has dimension 10000, 818 and 144 when considering 5, 2 or 1 tomographic bins respectively. Comparing the joint S/N with the power spectrum alone case (blue solid line), we find that the maximum information content increases by $\sim$10% for all the considered tomographic analysis. This confirms the need for the inclusion of the bispectrum analysis for future weak lensing analyses. Similarly to the previous paragraph, we address the effect of the super-sample covariance: when included, the maximum achievable information content is reduced by about 30% in the 1 bin tomographic case and by about 25% in the 2 and 5 bin analyses. As a final remark on this joint analysis, we compare the information content in all the 3 tomographic cases with the hypothetical case of a Gaussian field contaminated with shot noise (dark green solid lines in Fig. 3.2.4). If the convergence field were Gaussian, this line would represent a perfect reconstruction of the cosmological information in the field. It can be observed that, in the most informative case with 5 tomographic bins, we recover about 60% of this ideal cosmological information.

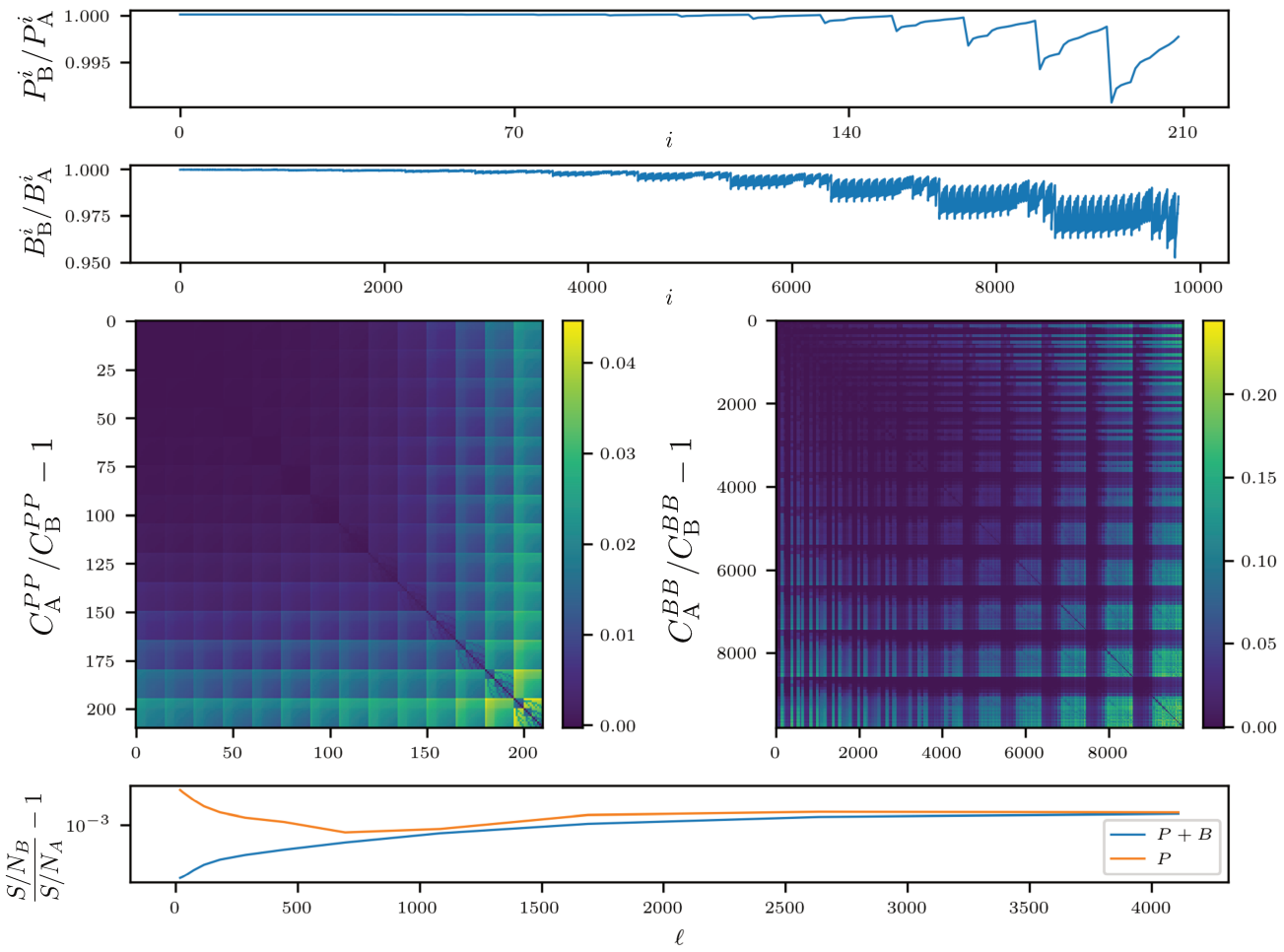

Figure 3.2.5: Fractional differences between observables (data vector and covariance) as derived from the model A and B (defined in the main text). From the top to the bottom: 1) vector of power spectra $\mathbf{P}$, 2) vector of bispectra $\mathbf{B}$, 3 (left)) power spectrum covariance matrix , 3 (right)) bispectrum covariance matrix, 4) power spectrum and joint power spectrum-bispectrum cumulative S/N. For this analysis we keep our usual binning in $\log \ell$ (14 regularly spaced bins from $\ell = 10$ to $l\ell = 5000$) and all the survey related parameter as chosen in Sec. 3.1.2. Specifically, we use a 5 bin tomography for the sources.

3.2.4 Uncertainty of the theoretical model: concentration parameter as a random variable

We investigate the robustness of our S/N-based forecasts against the uncertainties of the halo profile properties. In particular, we want to test the impact of the convolution of the halos with the probability density function $p\left(c, m, z\right)$ for the concentration parameter c introduced in Eq. (1.163). In this section we will make the redshift dependence explicit again. We will compare the observables as built from two models. In the first model (model A), we will evaluate I^{β}_{μ} (3.15) from the full integration over the range $(c_{\rm v}^{\rm Min}, c_{\rm v}^{\rm Max})$. In the second model (model B) we will simply assume $p\left(c_{\rm v}, m, z\right) = \delta_{\rm D}\left(c_{\rm v} - \bar{c}_{\rm v}\left(m, z\right)\right)$. This last approximation is the one mostly used in the

literature. The mean value $\bar{c}_v\,(m,z)$ is given by the following fitting formula [Bullock et al., 2001]

$$\bar{c}_v\,(m,z) = K\,\frac{a(z)}{a_c}, \tag{3.42}$$

$$m_\star\,(a_c) = F\,m, \quad \nu\,(m_\star, a_c) = 1 \tag{3.43}$$

where a_c is the epoch at which the typical collapsing mass $m_\star$ (defined by $\nu = 1$) equals a fixed fraction of the halo mass at the same epoch. The best-fit values for the free parameters $F,\ K$ are respectively 0.01 and 4.0 for a ΛCDM Cosmology. In Fig. 3.2.5 (starting from the top), we can see that the vector of power spectra **P** differs at maximum by 1% between the two models (first panel). The data vector **B** is more affected with a peak at 5% (second panel) for the smallest scale. We recall from Sec. 3.1.1 that the observables in our data vector are ordered for increasing value of the multipoles used within the configurations. The jagged profile instead corresponds to different tomographic dispositions of the sources per fixed spatial configuration. In the third row we compare the power spectrum (left panel) and the bispectrum (right panel) covariance matrix as derived from the two models. They differ at maximum by 4% and 25%, respectively. However, there is no impact at all at the level of reconstructed S/N: we can see in the bottom panel in Fig. 3.2.5 that the fractional differences between the two models is well within the numerical precision of our pipeline ($\sim$1%).

For completeness, we refer to Cooray & Hu [2001] for a similar analysis. In this work, the impact of marginalising over the concentration parameter was performed at the level of three-dimensional, 1-halo power spectrum and trispectrum. Specifically they found an impact of $\sim$ 5% and $\sim$ 20% respectively for a distribution $p\,(c_\nu, m, z)$ of variance $\sigma_{\ln c} = 0.2$.

3.2.5 Principal component analysis

The bispectrum only brings a relatively small improvement of information, mainly due to the effect of the super-sample covariance. Thus, it seems desirable to simplify our problem and restrict our data vector to the most informative modes. The principal component analysis (PCA) is a simple way to exhibit which observables, or linear combination of them, are the most informative. It can be achieved via an eigenvalue decomposition of the $n \times n$ covariance C onto an orthonormal basis of vectors $\mathbf{v}_{i=1,\dots,n}$

$$C_{ij} = \sum_a S_{ai}\,S_{aj}\,\lambda_a. \tag{3.44}$$

In the above equation, $S_{aj} \equiv v_{a,j}$ and λ_a is the a^{th} eigenvalue associated to the mode $\mathbf{v}_a$. Also, we are omitting the supescrit D: from now on we will apply this method only to the joint covariance implying $C = C^{\text{D}}$. The decomposition (3.44) allows to find linear combinations of the original observables in the data vector $\{D_i\}_{i=1,\dots,n}$

which defines a new set of uncorrelated observables $\{\hat{D}_i\}_{i=1,\dots,n}$

$$\hat{D}_i = \sum_j S_{ij} D_j. \tag{3.45}$$

At this point, the signal-to-noise ratio for each of them will simply be the ratio of the signal and its own error

$$\left(\frac{S}{N}\right)^2_a = \frac{\hat{D}^2_a}{\lambda_a}. \tag{3.46}$$

In the PCA formalism, the eigenvalue λ_a is indeed the variance associated to $\hat{D}_a$. From a much more interesting perspective, we can analyse the information content in the different modes $\mathbf{v}_a$ of the covariance as cumulative function of the new observables

$$\left(\frac{S}{N}\right)^2_{\mathrm{c},a_{\max}} \equiv \sum_{a=1}^{a_{\max}} \left(\frac{S}{N}\right)^2_a, \tag{3.47}$$

up to a mode $\mathbf{v}_{a_{\max}}$. In the following, we will analyse the cosmological information both as quantified from a single mode (Eq. (3.46)) and as cumulative function of a set of eigenmodes (Eq. (3.47)). In this kind of analyses, the smallest eigenvalues (associated to the most important eigenmodes in terms of information reconstruction) are potentially affected by numerical errors. Due to the wide dynamics of the observables used for this work, the covariance matrices have large condition numbers: the eigenvalues span over a range of about 30 orders of magnitude and the smallest ones can be affected by numerical errors if a too naive eigenmode decomposition algorithm is used. For this reason we relied on a specific high performance routine able to search for the eigenvalues (and associated eigenmodes) in a large dynamical range as described in Sec. 3.1.4.

We present the main results of the PCA analyses we did in Fig. 3.2.6 and in Fig. 3.2.7 where the eigenmodes $\mathbf{v}_a$ are ordered respectively by increasing value of their variance λ_a and decreasing information content S/N_a. The analyses have been applied on the full joint covariance, including super-sample contributions and all the multi-halo configurations for the bispectrum vector. In the first row of both figures, we display the reconstruction of the signal-to-noise ratio as cumulative function of the eigenmodes included in the covariance. In the second row of both figures instead, we show the information content per eigenmode S/N_a. We refer to the corresponding captions for more details. From a parallel study of Fig. 3.2.6 and Fig. 3.2.7 we can investigate the possibility of further reducing the dimensionality of our analysis while preserving the maximum information content. These figures ought to be analysed in the light of Fig.3.2.8 where we show the ratio between the (original) observables and their standard deviations (for a 2 bin tomography, for simplicity). In particular, Fig.3.2.8 would represent the information content of our observables if they were not correlated. Even though it does not provide a reliable insight on the strength of our measurements in the regime we are exploring, it can still help in understanding what are the configurations expected to be more (less) informative. Also, it helps in giving a physical understanding to the results of the PCA analyses.

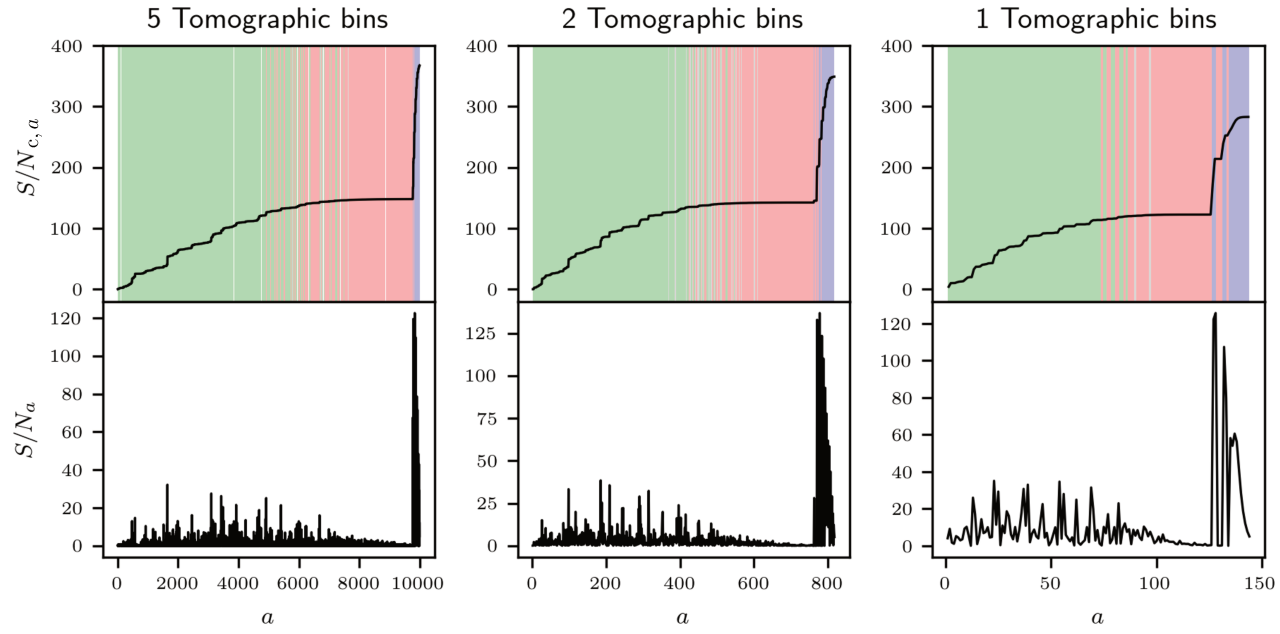

Figure 3.2.6: PCA analysis for the joint 5 tomographic bin covariance matrix with no approximations on the covariance nor on the bispectrum data vector. *Top:* $S/N_{c,a}$ (3.47) as cumulative function of the eigenmodes $\mathbf{v}_a$. The eigenmodes $\mathbf{v}_a$ are ordered for increasing value of the associated variance λ_a. The different colours on the background label modes which corresponds to observables $\hat{D}_a$ which are mostly combinations of power spectra (*blue*), bispectra in the linear/mildly non-linear regime $[\ell < 400]$ (*red*) and bispectra in the non-linear regime $[\ell > 400]$ (*green*). We perform this classification for a given observable $\hat{D}_a$ by looking at which are the most representative observables D_i contributing via the matrix S_{ai}. Considering only observables D_i for which the corresponding matrix element $|S_{ai}| \geqslant 0.2$, the mode $\mathbf{a}$ is then classified according to which of the above classes represents more than the 90% of them. If there is not a specific preference, we assign a *grey* colour. *Bottom:* S/N ratio per eigenmode. The eigenmodes $\mathbf{v}_a$ are ordered for increasing value of the associated variance λ_a.

First of all, by looking at the bottom row in Fig. 3.2.6, we can identify a fraction of the eigenmodes which have a very poor information content S/N_a. On the top row, they correspond to a red plateau, i.e. eigenmodes mainly associated to bispectra in the linear/mildly non-linear regime via the linear combination (3.45) (see caption for more details). This feature is present for all the tomographic redistributions of the sources. This result can be easily explained by looking at Fig.3.2.8. As a matter of fact, the bispectra sourcing these modes have a very a low signal-to-noise ratio since at these scales the matter field has a very weak deviation from the Gaussian statistics.

Going back to the top-row panels of Fig. 3.2.6, the eigenmodes located in blue regions are characterised by a large variance, being on the right end of all the panels. However, they are also the most important in terms of recovering the total information content, carrying the largest signal-to-noise ratio per mode (bottom row). In particular, they are at the left end of the panels in Fig. 3.2.7. Following the procedure described in detail within the caption of Fig. 3.2.6, we find that these modes are mainly linear combinations of power spectra. Then, we can easily understand the location of these modes within both Fig. 3.2.6 and Fig. 3.2.7. Power spectra have an absolute standard deviation much larger than the bispectra ($\sim 10^{-9} - 10^{-13}$ and $\sim 10^{-12} - 10^{-20}$ respectively), which motivates their location on the right end of Fig. 3.2.6 (we recall that, approximately, $C^{PP} \sim P^2$). On the other

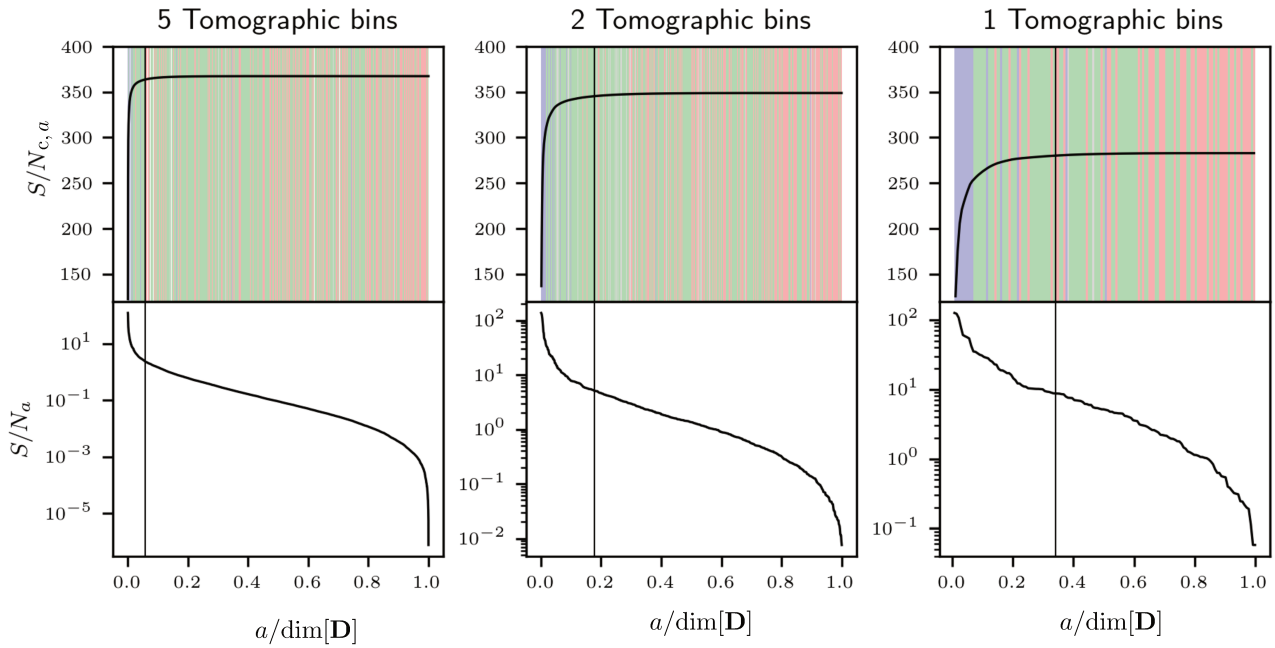

Figure 3.2.7: The analysis here proposed is the same as in Fig. 3.2.6. However, the eigenmodes are now ordered for decreasing value of the associated S/N_a and and we are reporting our results in terms of the fraction of the total number of eigenmodes used. This is just motivated by graphical reasons. *Top:* S/N as cumulative function of the eigenmodes (3.47). The colour code is the same as in Fig. 3.2.6. *Bottom:* S/N_a per eigenmode (3.45). The vertical lines in the different panels indicate the fraction of eigenmodes required for recovering the 99% of the maximum achievable S/N.

hand, power spectra have a much larger signal-to-noise ratio when compared to all the other observables in the vector **D**. By comparing the first and the second row in Fig.3.2.8 we clearly see that power spectra can be measured with a much smaller statistical uncertainty when compared to the bispectra.

Finally, the modes located in the green bands are mainly sourced by bispectra evaluated on modes deep into the non-linear regime (see caption to Fig. 3.2.6). Compared to the modes associated to the bispectra in the linear/mildly non-linear regime (red), they have a lower variance, as they are located at the left end of the plots in Fig. 3.2.6 while carrying more information: they represents a transition between the blue and the red modes in top panels in Fig. 3.2.7 and they are crucial for improving the information carried by the first modes associated to the power spectra (blue). Once again, we can understand these dynamics with the help of Fig.3.2.8, where the bispectrum at non-linear scales has a higher signal-to-noise ratio compared to the one at larger scales. On the other hand, the smaller variance of these configurations is due to the fact that the bispectrum signal is much weaker the more we measure it in the non-linear regime, as we can see in Fig. 3.1.2 (we recall that, approximately, $C^{BB} \sim P^3 + B^2$).

The vertical lines in Fig. 3.2.7 indicate the fraction of eigenmodes required to recover 99% of the full information content, once the PCA-modes have been reordered by decreasing value of the associated signal-to-noise ratio. The result is remarkable: the higher is the number of tomographic bins, the higher is the compression efficiency.

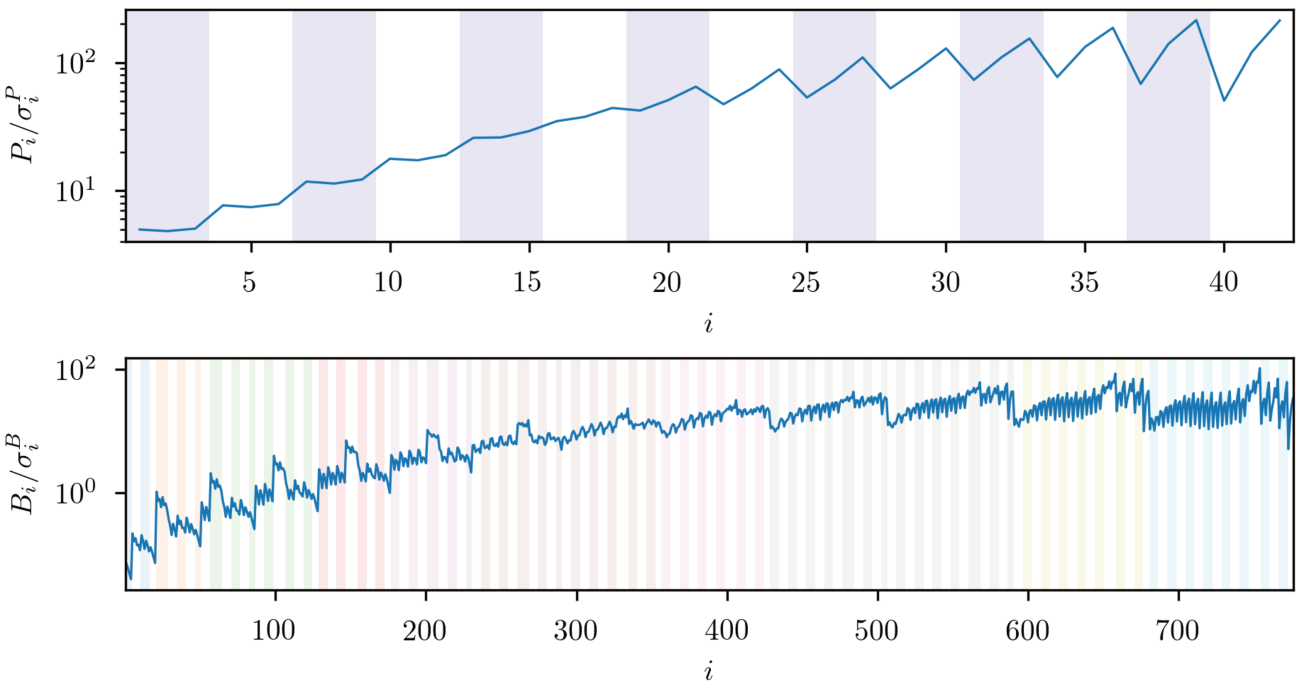

Figure 3.2.8: Classical signal-to-noise ratio analysis for the original data vector $\mathbf{D}$. We diplay the ratio between the value of an observable, distinguishing between power spectra and bispectra, and the associated standard deviation $\sigma_i^{P(B)} \equiv \sqrt{C_{ii}^{PP(BB)}}$. We assumed a 2 bin tomography. For every configuration in Fourier space, the different tomographic distributions of the sources are indicated by the same background colour. *Top*: binned tomographic power spectrum. *Bottom*: binned tomographic bispectrum. In the latter case, different colours on the background correspond to triangular configurations sharing the largest edge. Therefore, according to the ordering of the data vector defined in Sec. 3.1.1, within each of these bends we move from squeezed to equilateral configurations.

Specifically, just $\lesssim 40\%$, $\lesssim 20\%$ and $\lesssim 10\%$ of the modes are required respectively for the 1, 2 and 5 bin analyses.

In Fig. 3.2.9 and in Fig. 3.2.12 we visualise the absolute value of the elements within the projection matrix S_{ai} (3.44) for different choices of a (left and central panel) and the total matrix itself (right panel). The index a indicate a specfic mode $\mathbf{v}_a$ of the original covariance matrix. The modes are ordered for decreasing value of the information content S/N_a. The index i runs instead over the different elements of the data vector $\mathbf{D}$. Therefore, we want here to visualise how much each of the original observables contributes to the newly defined ones $\hat{\mathbf{D}}$. The colours on the background of the left and central panels have to be read as follows. Different tiny bends refer to the same configurations in Fourier space, then spanning over the different tomographic contributions to it. We recall that in our data vector $\mathbf{D}$ these contributions are contiguous. In the region of the x-axis corresponding to the sub-vector $\mathbf{B}$ (on the right of the vertical violet line), different colours of the bends indicate Fourier configurations sharing the same larger multipole. In particular, while moving from left to the right within each of these macro-bends, we are actually spanning bispectrum configurations from squeezed triangles to equilateral ones.

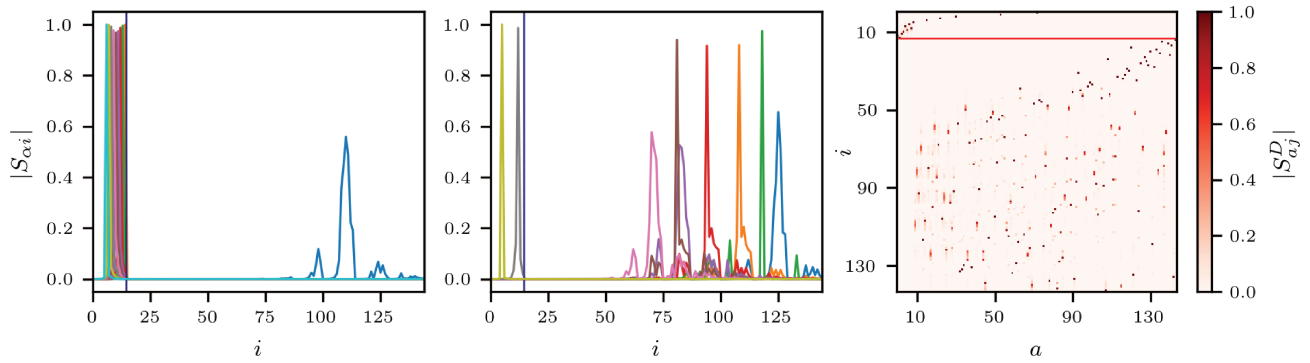

Figure 3.2.9: *Left*: values of the projection matrix $|S_{ai}|$ for different choices of a. For this plot, we picked the most informative modes $\mathbf{v}_a$ which allow to recover the 90% of the total information content. We depict our results for the 1 tomographic bin joint covariance matrix. The vertical blue lines separates the vectors $\mathbf{P}$ (on the left) and $\mathbf{B}$ (on the right) within the vector of observables, labelled by i. *Center*: same analysis as in the left panel. Here we picked the most informative modes $\mathbf{v}_a$ that allows us to recover an extra 5% of the total information on top of the modes displayed in the left panel. *Right*: representation of the full projection matrix $|S_{ai}|$, a being ordered for decreasing value of the information content per mode S/N_a. The red horizontal line separate power spectra and bispectra in the data vector.

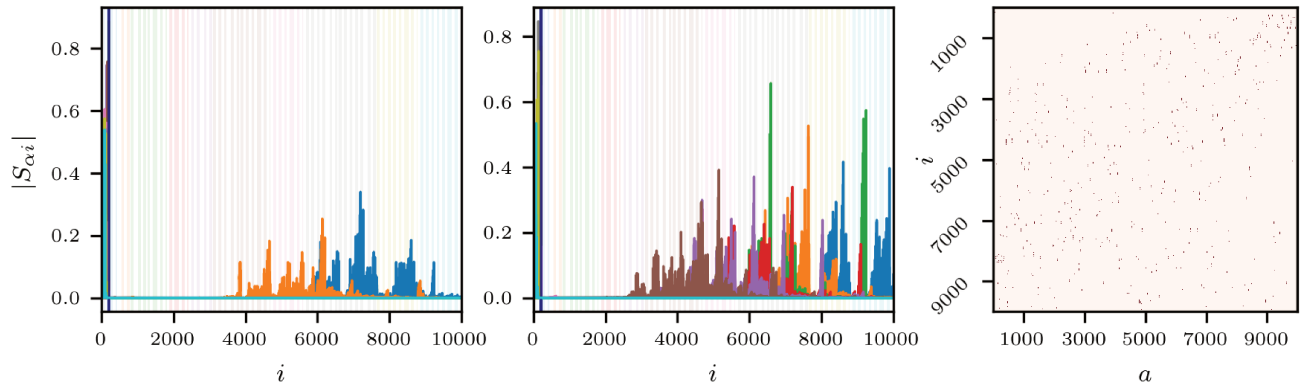

Figure 3.2.10: *Left-Center*: same analysis as in the left-central panels in Fig. 3.2.9 but for the 5 tomographic bin covariance matrix. The tiny different bends on the background include different tomographic contributions for the same Fourier configuration. Different colours on the background correspond to triangular configuration sharing the largest multipole. Therefore, according to the ordering of the data vector $\mathbf{D}$ 3.1.1, within each of these bends we move from squeezed to equilateral configurations (more detailes within the main text). *Right*: sparsity pattern for the full projection matrix $|S_{ai}|$ with a being ordered for decreasing value of the information content per mode S/N_a. The points have been selected whenever $|S_{ai}| \geqslant 0.05$.

In Fig.3.2.9, we first focus on the 1 tomographic bin joint covariance matrix. In the left panel we depict the values $|S_{ai}|$ as function of the index i for the most informative modes $\mathbf{v}_a$ recovering the 90% of the total information content. We can see that they are mainly sourced by the power spectrum in the non-linear regime. In the central panel we show the values $|S_{ai}|$ as function of the index i for the modes $\mathbf{v}_a$ required to add an extra 5% on top of the previous ones. They are mainly linear combinations of bispectra in the non-linear regime. Furthermore, the projection matrix shown in the rightmost panel is extremely sparse, peaking on just few configurations: overall,

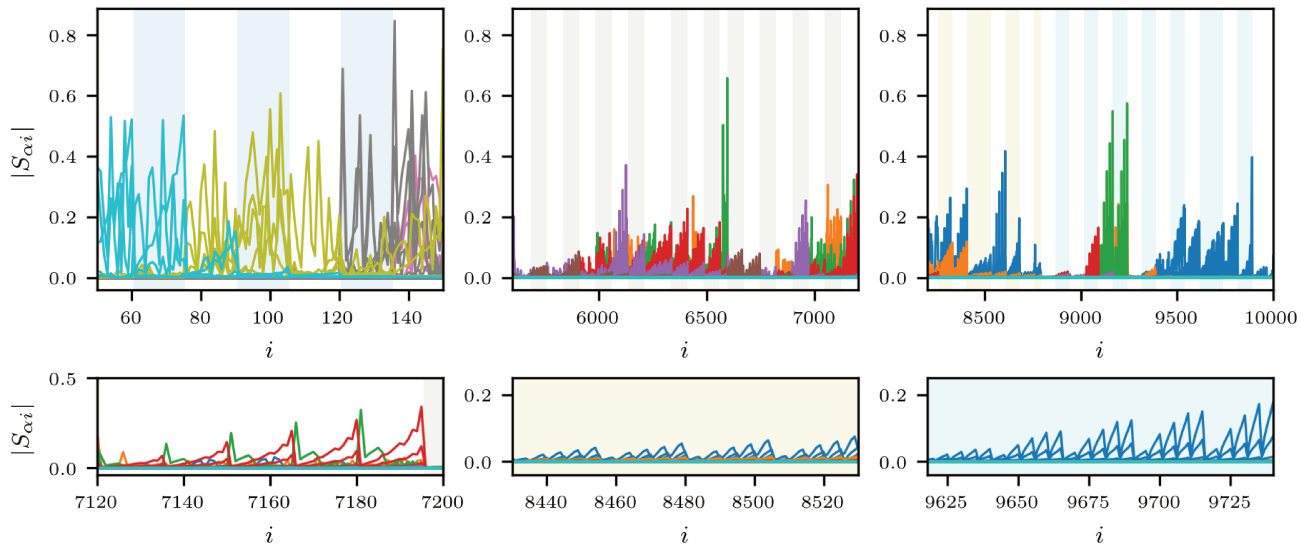

Figure 3.2.11: Same analysis as in the left and central panel in Fig. 3.2.10. In particular we are considering a 5 tomographic bin joint covariance matrix. We are now zooming on just one triangular configuration per panel. The colour map used in these plots is the same as in the left and central panels in Fig. 3.2.10, to facilitate the comparison.

just 53 observables (36%) are involved in recovering 95% of the information content[7]. This open a possibility for a second level of data compression beyond the one identified at the previous step at the level of modes.

In Fig.3.2.10 we perform the same analysis for the 5 bin tomography, reaching similar conclusions. In particular the tomography does not affect the possibility of dimensional reduction: we can discard 2093 configurations (i.e. 80% of the total) and still recover 95% of the information[7]. In Fig. 3.2.11, we show the profiles $|S_{ai}|$ zooming onto smaller fraction of the data vector $\mathbf{D}$ (top row) and onto single triangular configurations (bottom row). The first panel on the left in the top row focus on the region of x-axis associated to the vector $\mathbf{O}$. These panels are useful in order to understand the distribution of the cosmological information among different tomographic bins, per fixed spatial configuration. While it is not possible to identify dominant tomographic configurations at the level of the power spectra (top-leftmost panel), at the bispectrum level the tomographic configurations that are most informative are, in general, those involving high redshift sources. We can clearly see this feature in the bottom row of Fig. 3.2.11 where we focus on single spatial configurations. The jagged profile mirrors the ordering of the tomographic dsitributions of the sources within the data vector. We recall that for the bispectrum, the observables associated to the tomographic bins (i, j, k) are ordered such that k is the fastest index and i the slowest index. Then we can see a peak in the profile whenever, for a fixed value of $k > i, j$, i and j move closer to k. The maximum information for a given spatial configuration happens to be when all the tomographic indices have reached the maximum value corresponding to the furthest sources.

Finally, in Fig. 3.2.12, we show the profile of the projection matrix $|S_{ai}|$ for a single mode a, picked among the

[7]We consider configurations D_i with an associated weight $|S_{ai}| < 0.05$ as negligible. The same threshold has been applied to reproduce the sparsity pattern for the 5 tomographic bin covariance matrix in Fig. 3.2.10.

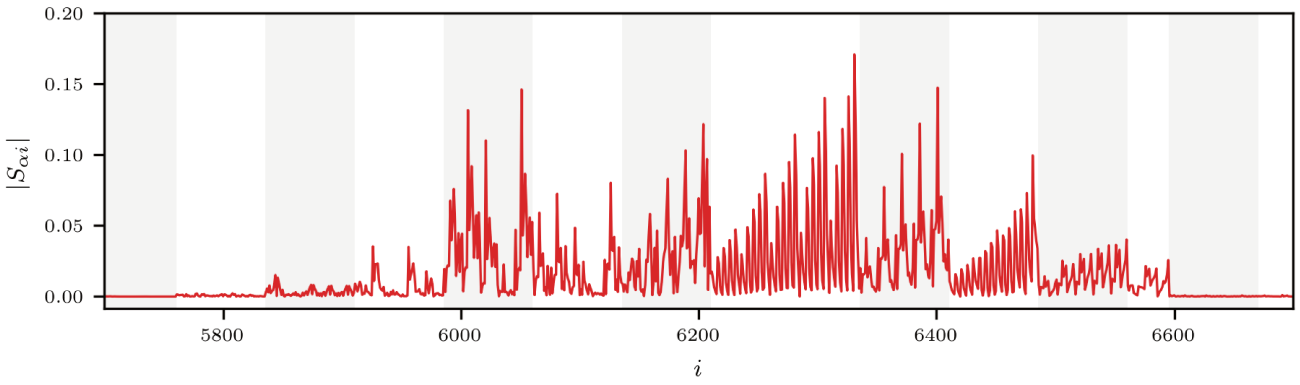

Figure 3.2.12: Profile of the projection matrix $|S_{ai}|$ for a specific choice of mode $\mathbf{v}_a$ picked from the most informative ones accounting for the 95% of the total information content. The colour code is the same as in Fig. 3.2.10 in order to facilitate the comparison.

most informative ones recovering the 95% of the total information content in a 5 bin tomography. In particular we picked a mode mainly dominated by bispectra to inspect the possibility of reducing the computational burden for this part of the data vector. Unfortunately, we can see that there are no dominant configurations and bispectra evaluated on several different spatial configurations are required to access the information content of the mode analysed. Also, for a specific triangular configuration (single background bend), it is not possible to identify a dominant contribution from a specific distribution of the sources.

3.3 Discussions and conclusions

3.3.1 Summary of the approximations

As part of our conclusions, we summarise here the approximations adopted throughout the project along with the level of urgency for their validation. As we will explain, the precision of the halo model in describing the matter clustering will pretty often represent a bottleneck for the accuracy of the whole forecast. Therefore, while lacking a better semi-analytical model for the gravitational collapse in the non-linear regime, it is important not to forget that the halo model itself is an approximation. To start with, in deriving the statistical properties of the projected fields we assumed the Limber and the flat-sky approximations. Kilbinger et al. [2017] thoroughly explored these approximations at the level of 2-point statistics and compared their predictions against a full-sky approach for the cosmic shear field with a CFHTLenS-like galaxy distribution [Kilbinger et al., 2013]. For the joint approximation under exam, they found it to be accurate to better than 10% for $\ell > 3$, converging slowly to the true projection with percent level precision at $\ell > 100$. Given that the accuracy of the halo model predicted polyspectra is in general much lower, we consider the above performance satisfactory for our work. Secondly, while computing the covariance matrix for the power spectrum and bispectrum we assumed that the trispectra, 5- and 6-order polyspectra in Eqs. (2.122), (2.147) and (2.149) (resp.) are slowly varying within the considered ℓ-bins. To our knowledge, no

tests in the literature were performed on this matter for the observables and the scales of interests. However, given the dynamics of the spectra in Fig. 3.1.2 (left panel) we advise further tests of this assumption, especially for high ℓ. On the other hand, the high number of configurations required for our analyses makes the full integrations over the bins width unfeasible given the computational resources at our disposal. Still at the level of covariance matrix we approximated all the polyspectra with their 1-halo component (starting from the 3-point one). Once again, we are not aware of any studies on the impact of neglecting these higher halo terms. Given our discussion at the beginning of Sec. 3.1.3, we do not expect them to be important for cosmological analyses even though they may have a role on regularising the covariance matrix reducing errors at the inversion. Finally, in the computation of the super-sample covariance (2.228), we are again assuming the joint flat-sky Limber approximation for both the intra- and super-survey modes. At the power spectrum level, Barreira et al. [2018] analysed the impact of these approximations against a full spherical analysis of the super-survey modes. For Euclid- and LSST-like tomographies and at the power spectrum covariance level, they report that for surveys covering less than $\sim 5\%$ of the sky the two results agree to better than 1%. However, when moving to the expected realistic coverage of $f_{\rm sky} \approx 0.3$–0.4 the use of the flat-sky expressions results in an underestimation of the SSC contribution of about 10%. Again, this performance has to be considered satisfactory for the present work given that the precision bottleneck is mainly given by the accuracy of the matter clustering model.

3.3.2 Comparison with similar works in the literature

We compare the analysis performed in this chapter with similar works in the literature. The impact of the correlations between observed and super-sample modes has already been addressed in the literature. Specifically, Kayo & Takada [2013] performed a preliminary study on the information content of the weak lensing bispectrum, which was further developed in Kayo et al. [2013] where 3 tomographic bins where considered for cosmological parameter forecast. However, they included the super-sample correlations at very small scales. These terms are known in the literature as halo sample variance and they contribute to the correlations when all the points in a given configuration are inside the same halo. Formally, the halo sample variance is just a part of the components we obtained via the response approach described in Sec. 2.3.2. In particular they corresponds to the last terms in Eq. (3.28) and Eq. (3.29) which are the responses of the 1-halo term to the long mode $\delta_{\rm b}$. In a different study, Barreira [2019] accounts for the super-sample covariance in the PT response formalism by making different approximations. In particular, he considers only correlations between bispectra on squeezed configurations. This approximation results in a negligible impact on the signal-to-noise ratio when adding the super-sample covariance. Our study shows that this is not the case when all the triangular configurations are included in the analysis.

3.3.3 Not only the bispectrum

Besides the bispectrum, different techniques have been advocated in order to recover the cosmological information not accessible via the 2-point statistics of the weak lensing field. In the following paragraph, we would like to

propose a summary of the most current alternatives approaches in the literature.

Minkowski functionals This approach proposes to study the morphological properties of a thresholded n-dimensional scalar field (e.g. for the convergence field $n = 2$). The Minkowski functionals are $n + 1$ functions of the threshold ν which provide full description of the morphological properties of the field [Mecke et al., 1994]. For the weak lensing convergence field, the full characterisation is then given by 3 Minkowski functionals. They are respectively related to 1) the total area of the domain supporting the field above the threshold, to 2) the total circumference of these regions and to 3) the Euler characteristic of the field. Minkowski functionals have been shown to be sensitive to the underlying Cosmology and can be exploited to constrain cosmological parameters (Matsubara [2000]; Matsubara & Jain [2001]; Sato et al. [2001]; Guimarães [2002]; Kratochvil et al. [2012]; Munshi et al. [2012]; Petri et al. [2015]; Liu & Madhavacheril [2019]). For the weak lensing convergence, they have been study first by Sato et al. [2001] using the signal as measured from ray-tracing simulations in the weakly non-linear regime. Sato et al. [2001] showed that the accuracy of the Ω_m inference can be improved by $\sim$20% when using the Minkowski functionals with respect to other conventional methods, like direct measurements of the convergence skewness. Kratochvil et al. [2012] analysed the information content of the 3 Minkowski functionals (both jointly and separatly) in comparison with the power spectrum. Tomographic binning of the sources and different smoothing scales were considered. Interestingly, Minkowsky functionals already recover all of the information carried by the power spectrum which does not improve the constraints any further when added on top of the formers. In particular the results were reported in terms of the 1σ contours of the marginal posteriors for the parameters (Ω_m, w, σ_8). We finally point the reader to the recent paper from Petri et al. [2015] where MCMC analyses were performed on the publicly available data from the 154 deg^2 CFHTLenS survey [Van Waerbeke et al., 2013]. The data vector considered in this analysis were formed by moments of the convergence field (up to fourth order: two quadratic, three cubic and four quartic), Minkowski functionals and power spectrum. As results from the cited work, the Minkowski functionals' constraints on the doublet (Ω_m, σ_8) are incompatible with cosmic microwave background anisotropies measurements [Hinshaw et al., 2013; Planck Collaboration, 2014] while the moments' constraints do not show such tension. We may therefore think that systematics are sill present in the modelling of Minkowski functional descriptors. As far as the dark energy equation of state is concerned, the considered data set is insufficient to constrain w to a reasonable precision with any of the probes.

Peak statistics This probe aims at constraining cosmological parameters via the analysis of the statistical distribution of peaks in the two-dimensional convergence field. Potentially, such peaks are motivated by the projection of single, discrete, massive collapsed object in the foreground and altogether these objects are tracers of the underlying matter field. In turn, their statistical distribution is determined by the corresponding mass function which can be either calibrated in simulations or derived from first principles. In both cases, the mass function has a strong dependence on the underlying Cosmology which can then be probed via the detection of the aforementioned peaks. Clearly, the possibility of one-to-one relation between peaks and underlying objects is idealistic: high peaks in the convergence can be induced by aligned structures while low peaks maybe due to the lensing from

a given object being compensated by an under-dense region in the foreground. The possibility to probe cosmology at non-linear scales via peak counts is therefore a complex subject and its potential as complementary probe to the power spectrum led to several studies over the past years (e.g. Marian et al. [2009]; Maturi et al. [2010]; Dietrich & Hartlap [2010]; Kratochvil et al. [2010]; Yang et al. [2011]; Marian et al. [2013]; Lin & Kilbinger [2015]; Liu et al. [2015]; Lin & Kilbinger [2018]). We only discuss in the following a few interesting works. Dietrich & Hartlap [2010] inspected the constraining power of peak statistics by looking at the S/N of the tomographic aperture mass field [Schneider, 1996] as measured from ray-tracing simulations. Interestingly, this novel probe is shown to provide tighter constraints on the doublet (Ω_m, σ_8) than standard tomographic shear measurements (2-point correlation function). Respectively the two approaches lead to FoM for (Ω_m, σ_8) of 123 and 71. The joint analysis allows to increase the FoM by $\sim$40% when compared to the cosmic-shear tomography alone. Subsequently, Yang et al. [2011] used *N*-body simulations to also analyse the relation between the height of the peaks in the convergence field and the number of background halos. High peaks, with amplitude $\gtrsim 3.5\,\kappa/\sigma_\kappa$ (σ_κ being the r.m.s. of the convergence field κ), turn out to be dominated by a single massive halo while medium-height peaks ($\sim 0.5 - 1.5\,\kappa/\sigma_\kappa$) cannot be attributed to a single collapsed one. Cosmological analyses are mainly driven by the statistical distribution of the formers. In terms of forecast of cosmological parameters considered (Ω_m, σ_8, w), the joint analysis peak statistics-power spectrum (where 2 possible redshift are considered for the sources, i.e. $z_s = 1$ and $z_s = 2$) allows to reduce the 1σ constraints by, overall, a factor of $\sim$3 compared to the power spectrum alone (we are in particular comparing here the 4th and the 7th rows - from the top - in Table XIV form the original paper [Yang et al., 2011]). Finally, Liu et al. [2015] explored the constraining power of peak statistics on the publicly available data from the 154 deg^2 CFHTLenS survey [Van Waerbeke et al., 2013] performing MCMC analyses on the usual three-dimensional parameters space (Ω_m, σ_8, w) and using as data vector the histogram of the peaks as function of the height of the convergence field and the convergence power spectrum (jointly and separately). Liu et al. [2015] found that constraints from peak counts are comparable to those from the power spectrum and somewhat tighter when different smoothing scales are combined. When the power spectrum and peak counts are combined, the parameter constraints are improved by an overall factor of $\sim$2 when compared to the power spectrum alone.

Machine Learning techniques Machine Learning saw a dramatic growth in terms of exploitation within cosmological and astronomical applications over the last years, especially under the forms of its specific declination into Neural Networks (NNs) (e.g. Lochner et al. [2016]; Möller et al. [2016]; Charnock & Moss [2017]; Ravanbakhsh et al. [2017]; Fluri et al. [2018]; Lanusse et al. [2018]; Charnock et al. [2018]; Gupta et al. [2018]; Perraudin et al. [2019]; Mustafa et al. [2019]; Ribli et al. [2019]; Kodi Ramanah et al. [2019]; He et al. [2019]). This was mainly motivated by the perfect match between supply and demand. From one side, we are entering the big data era for Cosmology with the up-coming launches of modern galaxy surveys such as DES [Flaugher, 2005], the Hyper SuprimeCam Survey (HSC, Aihara et al. [2018]), the Large Synoptic Survey Telescope (LSST, Ivezić et al. [2019]) and Euclid [Laureijs et al., 2011; Amendola et al., 2013]. On the other hand, the recent technological developments have been making GPUs affordable for the scientific community and NNs can provide short-cuts

to the complex non-linear physics involved in LSS formation. We list here few recent works concerning the application of these novel techniques to weak lensing data sets. Gupta et al. [2018] inspected the performance of Convolutional Neural Networks (CNN) (the employed architecture is described in Sec. 2 of the original paper) in constraining the cosmological parameters (Ω_m, σ_8) from noiseless ray-tracing simulations. Their final claim is that CNNs outperform the power spectrum by a factor of ~ 5 and peak counts by a factor of ~ 4 in terms of FoM for the 2 parameters above. The (Gaussian) posterior analyses from the CNN were performed using as observables within the data vector the cosmological parameters $(\Omega_m[\bar{\Omega}_m], \sigma_8[\bar{\sigma}_8])$ as estimated from the CNN from maps sharing the same underlying Cosmology $(\bar{\Omega}_m, \bar{\sigma}_8)$. Ribli et al. [2019], similarly to Fluri et al. [2018] however without the Limber approximation, applied CNNs on noisy simulated maps. An error analysis similar to the one described above (but with a varying covariance in the MCMC sampler) allowed them to show that the CNN they employed (claimed to be more accurate than the one in Gupta et al. [2018]) produces constraints one the cosmological parameters (Ω_m, σ_8) which are ~ 8 and ~ 13 times smaller than those obtained from peak counts and power spectrum analyses, in terms of 1σ area. When the shape noise is included, the improvements is respectively of a factor of $\sim 1.05 - 1.42$ and $\sim 2.4 - 2.8$ for a Euclid-like survey ($n_{\rm tot} \approx 30$ gal arcmin^{-2}). Interestingly, in Appendix C of Fluri et al. [2018] the authors tested the architecture on a model excluded from the training set resulting in a negligible impact at the level of confidence regions. Finally, and for completeness, we would like to cite the work of Mustafa et al. [2019] where, while not focusing on cosmological parameter estimation, an application of Generative Adversarial Networks is proposed for a fast production of weak lensing convergence maps which result being characterised by the same summary statistics as the fully simulated maps. This is of major interest given our discussion in Sec. 2.2.1 where we underlined the computational challenge required by standard N-body simulations whenever used for the computation of highly dimensional covariance matrices.

Given the high performance of the novel probes described in the previous paragraphs, we wonder if the bispectrum can still represent a worthy tool for cosmological analyses. We remind that one of our most important results is that the bispectrum allows us to enhance the maximum achievable S/N by 10% when compared to the power spectrum alone. We can clearly see that, in terms of cosmological parameter forecasts, especially peak counts and machine learning-related techniques have the potential to outperform the bispectrum. However we are strongly convinced that the present status of the art does not rule out the bispectrum as meaningful cosmological probe. First of all, the analyses performed in this thesis are substantially related to the S/N. It is not possible to predict yet how different cosmological parameters will be constrained once a full MCMC will be run. The work from Petri et al. [2015] also showed the potential presence of systematics in the modelling of the Minkowski functional descriptors. Secondly, thanks to the tomographic analysis we are capable to access a time-dependent information which is paramount in order to constraint dark energy equation of state parametrizations. With this regards, analyses performed on the publicly available data from the 154 deg^2 CFHTLenS survey led to poor constraints on w both when exploiting peak counts and Minkowski functionals. Also, we are not aware of any published work tackling the forecast of time-dependent dark energy equation of state with NNs. Finally, all the probes listed in this last section are strongly dependent on cosmological simulations being our theoretical understanding of them still

not sufficient for the required forecast precision. Therefore, all the issues highlighted in Sec. 2.2.1 contaminate the present applicability of these techniques for unbiased cosmological parameter forecasts. Moreover, peak counts strongly depend on pixel-size measurements making this probe extremely dependent on the specific characteristics of a given simulation or on the detection properties of a realistic survey.

The bispectrum, in particular if estimated via analytical forward modelling, is still a competitive probe. As a matter of fact it allows us to model and have a full physical understanding of its covariance and signal without relying on simulations. In particular, they eventually miss important super-sample correlations and can be extremely costly whenever we want to meet the precision requirements from the upcoming generation of galaxy surveys.

3.3.4 Conclusions

In this chapter, we have presented the first full joint analysis of the information content for the convergence weak lensing power spectrum and bispectrum for a Euclid-like survey. We modelled the observables with the halo model and developed a high performance code capable to output fast and precise covariance matrices for the binned tomographic spectra. The covariance matrices calculated for the present work have been used for the signal-to-noise ratio analysis, but can benefit any forecast based on a Gaussian likelihood. We remind that the content of this chapter has been previously presented in Rizzato et al. [2018] (submitted to the *Monthly Notices of the Royal Astronomical Society*).

At the power spectrum level and including shot noise, we were capable to show that the super-sample covariance is the main source of error, leading to a reduction of 40% of the maximum achievable signal-to-noise ratio compared to the Gaussian case. The non-Gaussian cross-correlations between in-survey modes account for a loss of information of about 30% instead when compared to the Gaussian case. The combined effect of these two sources of error leads to a loss of about 45% on the signal-to-noise ratio. An important result of our analysis is the possibility to recover the cosmological information content of a Euclid-like survey by using 5 equipopulated tomographic redshift bins, instead of 10. This results is in particular insensitive to the angular scale and to the components included in the error budget of the observables and does not account for systematics like PSF, photo-z, blending or intrinsic alignments. The same phenomenology was found at the bispectrum level. On the bispectrum signal-to-noise ratio the super-sample covariance has an impact of about $\sim$13% (10 bin tomography) when the observables have been estimated via all the multi-halo configurations. We tested the impact of these configurations in the modelling of the vector of bispectra and we found that neglecting the 3-halo terms is a good approximation for our analysis, both when including or excluding the super-sample covariance.

Motivated by our previous findings, we performed a joint power spectrum-bispectrum analysis on 5 equipopulated tomographic redshift bins. We proved that this combined approach can improve the information content by $\sim$10% with respect to the power spectrum alone. The super-sample covariance of the bispectrum cannot be ignored and reduces the maximum information achievable of about $\sim$25%. In these analysis the 2- and 3-halo terms have been taken into account for the bispectra in the vector, but we restricted the computation of the co-

variance to the 1-halo terms which dominates the covariance for the most relevant configurations. These are the most important results of this work.

We found that the halo modelling uncertainty due to the scatter of the concentration parameter in the simulations does not affect the signal-to-noise ratio of the power spectrum by more than 1%. This makes our future forecasts robust to errors on this parameter.

In this work, a preliminary study on the possibility of further compressing our covariance matrices while preserving the cosmological information is considered. In particular, by performing a principal component analysis on our covariance matrices, we found that a very small fraction of the eigenmodes ($\lesssim$ 10% for 5 bin tomography) carries most of the information, and that not all the configurations equally contribute to the full information content. Indeed, only 20% of the data vector is included in the linear combinations that forms this 10% of eigenmodes, and thus a large fraction of the vector of the observables ($\sim$ 80% for 5 bin tomography) is not significant to reconstruct the signal-to-noise ratio at the different scales considered in this work.

Conclusions and perspectives

The main subject of this thesis was the development of precise and fast numerical tools ready to be employed in the analysis of key aspects involved in cosmological parameter forecasts from future weak lensing surveys. The computation of remarkably big covariance matrices is a key step in all Gaussian-based forecast techniques. This is in particular true both when performing analytical approximations (Fisher approximation) and when sampling the parameter space (Monte Carlo Markov Chain methods). In this thesis and in the associated paper [Rizzato et al., 2018], we tackled this challenging numerical task in the case of the weak lensing convergence field whose formalism has been described in Sec. 2.1. In Sec. 2.2.1, we argued that data-driven approximations to the covariances may not be suitable for future surveys being unable to describe the super-survey correlations. Therefore, we opted for the forward modelling of them. While providing an analytical approximation to the covariance, this approach can be extremely costly, especially when combining 2- and 3-point statistics. In weak lensing studies, like the one here proposed, this complexity is exacerbated when accounting for the tomographic distribution of the sources. We developed a high performance code that allows a highly parallelised computation of the covariance matrices for the weak lensing convergence power spectrum-bispectrum joint probe, assuming a Euclid-like photometry. While we deferred a precise parameter forecast for up-coming works, still we used the covariance matrices produced with our algorithm to have insights one the level of correlations within our data vector. We employed and studied the signal-to-noise ratio as definition of the information content in the probes and we made use of the principal component analysis to explore possibilities for data compression. The results of our work have been summarised in Sec. 3.3.

Starting form the results presented in this manuscript, several possible research paths are open. We collect them in four possible groups whose overall focus is the refinement and advanced cosmological applications of different key aspects of the tools here developed. Altogether, they have to be considered while getting ready for the development of likelihood codes for modern galaxy surveys. From a more general perspective, this last achievement has to be considered as the long term goal at which this work is aiming at.

Implementation of galaxy clustering. So far, the main object of our study has been the cosmological weak lensing signal. However, parameter forecasts can improve a lot if weak lensing is properly combined with galaxy clustering, as it is actually the scientific program of missions like Euclid and WFIRST. While the for-

malism required is quite different, still the cosmological exploitation of the galaxy clustering probe (up to the 3-point statistics) implies numerical challenges very similar to those tackled within this manuscript. The statistical isotropy being broken in redshift space, galaxy clustering bispectra will not depend solely on the length of the edges for a desired triangular configuration, but they will also depend on its orientation with respect to a chosen line-of-sight. The main consequence is an easily diverging complexity of the data vector and of the related correlations [Yankelevich & Porciani, 2019]. The core part of the code developed being a fast and precise evaluation of remarkably big covariance matrices, we will be in the position to implement galaxy clustering observables and related errors without increasing the complexity of the computation. Therefore, we consider the analysis of the joint weak lensing-galaxy clustering probe as a milestone for the scientific program started with the present work.

Covariance matrices compression The matter non-linear clustering induces correlations between different cosmological scales and probes. We used the principal component analysis (PCA) to determine which subsets of the observables, or linear combinations of them, are the most informative. While an overwhelming amount of data will be produced over the next decade, PCA allowed us to prove that we can reduce the computational burden involving the weak lensing bispectrum via the compression of the covariance matrix. In this thesis we have qualitatively identified three possible levels of dimensionality reduction: re-binning of the sources, PCA of the whole covariance matrix, exclusion of the least informative observables. Starting from here, we advise to exactly test their performance by looking at the actual amount of recovered information from the progressively reduced data vector. Beside the possibilities here explored, it will be extremely interesting to push the study of summary statistics for future galaxy surveys even further via the implementation of some existing approaches which so far have not been applied in this specific context (cosmic shear full nulling [Bernardeau et al., 2014], Komatsu-Spergel-Wandelt (KSW) estimator [Komatsu et al., 2005], skew-C_l approach [Munshi & Heavens, 2010], modal bispectrum [Fergusson et al., 2012], binned bispectrum [Bucher et al., 2010]) and the design of new ones.

Test of the approximations In this exploratory work we considered different approximations and they are summarised in the conclusions to Chapter 3. Not expecting major impacts on the final signal-to-noise ratio analyses, we employed them to reduce the computational burden associated to the forward modelling of covariance matrices, extremely challenging task by itself. However, bearing in mind the more valuable task of realistic likelihood implementations, the errors on the observables have to be refined exploring the validity range of the results we obtained. At the level of systematics, we expect photometric errors to enhance the cross-correlations between different tomographic redshift bins, potentially reducing the information content. Secondly, we will also need to account for intrinsic alignments, sources of correlations between galaxy shapes which are not caused by lensing from matter large-scale structures. One has to model all these effect not to bias the cosmological analyses for modern galaxy surveys. Also, they can potentially alter the efficiency of the compression techniques listed in the previous paragraph. At the level of covariance matrix, while the Limber and the flat-sky approximations are well motivated at the scales of interest, the assumption of slowly varying polyspectra within the multipole bins appears to be quite conservative. On the other hand, the level of complexity entailed by eventually relaxing this

assumption would be unsuitable for forecast analyses like those we presented. In this sense, the data compression techniques mentioned above will play a key role in reducing as much as possible the complexity of the data vector allowing a much easier test of these assumptions. Finally, we advise to study and account for the effect of sources of secondary non-Gaussianities which are not related to the matter clustering itself, namely the Born approximation, neglecting lens-lens coupling or using the reduced shear instead of the shear itself can lead to secondary contributions to higher order correlations.

Cosmological parameter forecast Summary statistics and tests of approximations are important steps towards the major goal of a precise cosmological parameter forecast. The signal-to-noise ratio analysis just provides an overall idea of the strength of our signal compared to the level of correlations. Further, to exactly understand the improvement in terms of error bars for the single parameters, a full posterior exploration is required. Fisher forecast and Markov Chain Monte Carlo (MCMC) methods are the most popular methods employed by the community. The former is computationally less expensive compared to the latter and it can eventually be used to get a fast feedback on the performance of the different actions we can take on the data vector and on the covariance (relaxation of approximations and data compression). However it relies on the assumption of a multivariate Gaussian posterior for the cosmological parameters and on the possibility to precisely compute the numerical derivatives of the observables with respect to the cosmological parameters themselves. Both these assumptions are rarely verified in real world applications. For example, the first is poorly realised whenever we want to measure parameters for the first time and the data are not expected to be extremely constraining. In this case posteriors will not be sharply peaked around the best fit value, around which a Guassian approximation could eventually work. The validity of the second approximation is hindered by our ignorance on the exact physical description of the matter non-linear clustering: polyspectra are well fitted by models available on the market but their derivatives are poorly reproduced [Reimberg et al., 2018]. MCMC methods instead allow deviations from this ideal situation but with the price of generating the vector of observables (and eventually the covariance matrix) at every point in parameter space. In this sense, we consider dimensionality reduction crucial for a precise forecast within up-coming galaxy surveys communities. While we will employ Fisher based analyses to access the performance of approximations and dimensionality reduction, a full posterior exploration is required to obtain quantitative results on the constraining power of our probes.

A possible alternative to be considered in parallel to MCMC analyses is represented by the recently introduced forecasting technique *Derivative Approximation for LIkelihoods* (DALI) [Sellentin et al., 2014]. It allows to approximate non-Gaussian posteriors with a positive definite and normalizable analytical expression. More specifically, the complex non-Gaussian features of the posterior are captured by a derivative expansion around its peak, summarising this information via tensors of given symmetries. Being an analytical approach, it inherits all the valuable aspects of a Fisher-like analysis. The formalism has been proved extremely flexible in reproducing strong correlations in parameter space. We are presently testing its actual capacity at approximating realistic posteriors. In case of great performance, we will capable to extend analytical studies like the one described in this manuscript to the tensors required by the DALI expansion. Indeed, the have to be thought as a new kind of summary statistics

beyond the simple covariance matrix or the theory vector (expectation value of the observables). At a more mature level of development for the DALI formalism, we will compare its forecasts with those coming from an actual MCMC for the joint power spectrum-bispectrum probe. We will apply this study to both weak lensing and galaxy clustering. In case of successful test, we may think to include the analytically-derived DALI tensors mentioned above as part of future likelihood codes for up-coming galaxy surveys, marking a turning point for cosmological data pipelines.

References

Aguena M., Lima M., 2018, Phys. Rev. D, 98, 123529

Aihara H., et al., 2018, PASJ, 70, S8

Amanullah R., et al., 2010, ApJ, 716, 712

Amendola L., et al., 2013, Living Reviews in Relativity, 16, 6

Bacon D. J., Refregier A. R., Ellis R. S., 2000, MNRAS, 318, 625

Baldauf T., Seljak U., Senatore L., Zaldarriaga M., 2011, J. Cosmology Astropart. Phys., 2011, 031

Baldauf T., Seljak U., Desjacques V., McDonald P., 2012, Phys. Rev. D, 86, 083540

Baldauf T., Seljak U., Senatore L., Zaldarriaga M., 2016, J. Cosmology Astropart. Phys., 2016, 007

Bardeen J. M., 1980, Phys. Rev. D, 22, 1882

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Barreira A., 2019, J. Cosmology Astropart. Phys., 2019, 008

Barreira A., Krause E., Schmidt F., 2018, J. Cosmology Astropart. Phys., 6, 015

Barrow J. D., Bhavsar S. P., Sonoda D. H., 1984, MNRAS, 210, 19P

Bartelmann M., Maturi M., 2017, Scholarpedia, 12, 32440

Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291

Bennett C. L., et al., 2013, ApJS, 208, 20

Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1

Bernardeau F., Crocce M., Scoccimarro R., 2008, Phys. Rev. D, 78, 103521

Bernardeau F., Bonvin C., Vernizzi F., 2010, Phys. Rev. D, 81, 083002

Bernardeau F., Pitrou C., Uzan J.-P., 2011, J. Cosmology Astropart. Phys., 2011, 015

Bernardeau F., Crocce M., Scoccimarro R., 2012, Phys. Rev. D, 85, 123519

Bernardeau F., Nishimichi T., Taruya A., 2014, MNRAS, 445, 1526

Bertacca D., Bartolo N., Bruni M., Koyama K., Maartens R., Matarrese S., Sasaki M., Wands D., 2015, Classical and Quantum Gravity, 32, 175019

Bertschinger E., 1992, in Martínez V. J., Portilla M., Sáez D., eds, New Insights into the Universe. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 65–126

Bertschinger E., 2000, arXiv e-prints, pp astro–ph/0101009

Blandford R. D., Narayan R., 1992, ARA&A, 30, 311

Blazek J., MacCrann N., Troxel M. A., Fang X., 2017, arXiv e-prints, p. arXiv:1708.09247

Bond J. R., Efstathiou G., 1984, ApJ, 285, L45

Bond J. R., Myers S. T., 1996, ApJS, 103, 1

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440

Bothun G. D., Geller M. J., Beers T. C., Huchra J. P., 1983, ApJ, 268, 47

Bucher M., van Tent B., Carvalho C. S., 2010, MNRAS, 407, 2193

Bull P., et al., 2016, Physics of the Dark Universe, 12, 56

Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559

Carroll S. M., 2004, Spacetime and geometry. An introduction to general relativity

Chan K. C., Scoccimarro R., Sheth R. K., 2012, Phys. Rev. D, 85, 083509

Chan K. C., Moradinezhad Dizgah A., Noreña J., 2018, Phys. Rev. D, 97, 043532

Chang C., et al., 2013, MNRAS, 428, 2695

Charnock T., Moss A., 2017, ApJ, 837, L28

Charnock T., Lavaux G., Wandelt B. D., 2018, Phys. Rev. D, 97, 083004

Chiang C.-T., Wagner C., Schmidt F., Komatsu E., 2014, J. Cosmology Astropart. Phys., 5, 048

Clifton T., Ferreira P. G., Padilla A., Skordis C., 2012, Phys. Rep., 513, 1

Cole S., Kaiser N., 1989, MNRAS, 237, 1127

Collaboration T. D. E. S., 2005, preprint, (arXiv:0510346)

Cooray A., Hu W., 2001, ApJ, 554, 56

Cooray A., Sheth R., 2002, Phys. Rep., 372, 1

Crocce M., Scoccimarro R., 2006a, Phys. Rev. D, 73, 063519

Crocce M., Scoccimarro R., 2006b, Phys. Rev. D, 73, 063520

Crocce M., Scoccimarro R., 2006c, Phys. Rev. D, 73, 063520

Crocce M., Scoccimarro R., 2008, Phys. Rev. D, 77, 023533

Crocce M., Scoccimarro R., Bernardeau F., 2012, MNRAS, 427, 2537

DESI Collaboration 2016, arXiv e-prints, p. arXiv:1611.00036

Dai L., Pajer E., Schmidt F., 2015, J. Cosmology Astropart. Phys., 2015, 059

Dark Energy Survey Collaboration et al., 2016, MNRAS, 460, 1270

Desjacques V., Jeong D., Schmidt F., 2018, Phys. Rep., 733, 1

Dietrich J. P., Hartlap J., 2010, MNRAS, 402, 1049

Dodelson S., 2003, Modern Cosmology. Academic Press, Elsevier Science

Doré O., Lu T., Pen U.-L., 2009, arXiv e-prints, p. arXiv:0905.0501

Efron B., 1982, The Jackknife, the Bootstrap and other resampling plans

Eifler T., 2011, MNRAS, 418, 536

Eisenstein D. J., Hu W., 1998, ApJ, 496, 605

Fergusson J. R., Liguori M., Shellard E. P. S., 2012, J. Cosmology Astropart. Phys., 2012, 032

Figueroa D. G., Sefusatti E., Riotto A., Vernizzi F., 2012, J. Cosmology Astropart. Phys., 2012, 036

Flaugher B., 2005, International Journal of Modern Physics A, 20, 3121

Fluri J., Kacprzak T., Refregier A., Amara A., Lucchi A., Hofmann T., 2018, Phys. Rev. D, 98, 123518

Fosalba P., Pan J., Szapudi I., 2005, ApJ, 632, 29

Friedrich O., Seitz S., Eifler T. F., Gruen D., 2016, MNRAS, 456, 2662

Fry J. N., Gaztanaga E., 1993, ApJ, 413, 447

Genel S., et al., 2014, MNRAS, 445, 175

Goroff M. H., Grinstein a., Rey S. J., Wise M. B., 1986, ApJ, 311, 6

Green J., et al., 2011, arXiv e-prints, p. arXiv:1108.1374

Guimarães A. C. C., 2002, MNRAS, 337, 631

Gupta A., Matilla J. M. Z., Hsu D., Haiman Z., 2018, Phys. Rev. D, 97, 103515

Hamana T., Miyazaki S., Okura Y., Okamura T., Futamase T., 2013, PASJ, 65, 104

Hamilton A. J. S., Kumar P., Lu E., Matthews A., 1991, ApJ, 374, L1

Hamilton A. J. S., Rimes C. D., Scoccimarro R., 2006, MNRAS, 371, 1188

Hartlap J., Hilbert S., Schneider P., Hildebrand t H., 2011, A&A, 528, A51

He S., Li Y., Feng Y., Ho S., Ravanbakhsh S., Chen W., Póczos B., 2019, Proceedings of the National Academy of Science, 116, 13825

Hearin A. P., Zentner A. R., Ma Z., 2012, J. Cosmology Astropart. Phys., 4, 034

Henry J. P., 2000, ApJ, 534, 565

Hikage C., Takada M., Hamana T., Spergel D., 2011, MNRAS, 412, 65

Hildebrandt H., et al., 2017, MNRAS, 465, 1454

Hinshaw G., et al., 2013, ApJS, 208, 19

Hoffmann K., Bel J., Gaztañaga E., 2015, MNRAS, 450, 1674

Hogg D. W., Eisenstein D. J., Blanton M. R., Bahcall N. A., Brinkmann J., Gunn J. E., Schneider D. P., 2005, ApJ, 624, 54

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581

Hu W., 2000, Phys. Rev. D, 62, 043007

Hu W., Dodelson S., 2002, ARA&A, 40, 171

Hu W., Kravtsov A. V., 2003, ApJ, 584, 702

Huterer D., Takada M., 2005, Astroparticle Physics, 23, 369

Ishak M., 2019, Living Reviews in Relativity, 22, 1

Ivezić Ž., et al., 2019, ApJ, 873, 111

Jain B., Bertschinger E., 1996, ApJ, 456, 43

Jarvis M., Jain B., 2008, J. Cosmology Astropart. Phys., 2008, 003

Jarvis M., et al., 2016, MNRAS, 460, 2245

Jasche J., Kitaura F. S., Ensslin T. A., 2009, arXiv e-prints, p. arXiv:0901.3043

Jee M. J., Tyson J. A., Schneider M. D., Wittman D., Schmidt S., Hilbert S., 2013, ApJ, 765, 74

Jee M. J., Tyson J. A., Hilbert S., Schneider M. D., Schmidt S., Wittman D., 2016, ApJ, 824, 77

Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372

Joachimi B., Shi X., Schneider P., 2009, A&A, 508, 1193

Joachimi B., Semboloni E., Hilbert S., Bett P. E., Hartlap J., Hoekstra H., Schneider P., 2013, MNRAS, 436, 819

Joudaki S., et al., 2017, MNRAS, 465, 2033

Kaiser N., 1984, ApJ, 284, L9

Kaiser N., Wilson G., Luppino G. A., 2000, arXiv e-prints, pp astro–ph/0003338

Kamionkowski M., Kosowsky A., 1999, A, 49, 77

Kayo I., Takada M., 2013, preprint, (arXiv:1306.4684)

Kayo I., Takada M., Jain B., 2013, MNRAS, 429, 344

Khandai N., Di Matteo T., Croft R., Wilkins S., Feng Y., Tucker E., DeGraf C., Liu M.-S., 2015, MNRAS, 450, 1349

Kiessling A., et al., 2015, Space Sci. Rev., 193, 67

Kilbinger M., 2015, Rep. Prog. Phys., 78, 086901

Kilbinger M., et al., 2013, MNRAS, 430, 2200

Kilbinger M., et al., 2017, MNRAS, 472, 2126

Kitching T. D., Taylor A. N., 2011, MNRAS, 416, 1717

Klypin A. A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102

Kodi Ramanah D., Charnock T., Lavaux G., 2019, Phys. Rev. D, 100, 043515

Komatsu E., Spergel D. N., Wandelt B. D., 2005, ApJ, 634, 14

Kowalski M., et al., 2008, ApJ, 686, 749

Kratochvil J. M., Haiman Z., May M., 2010, Phys. Rev. D, 81, 043519

Kratochvil J. M., Lim E. A., Wang S., Haiman Z., May M., Huffenberger K., 2012, Phys. Rev. D, 85, 103513

Krause E., Eifler T., 2017, MNRAS, 470, 2100

Krause E., et al., 2017, arXiv e-prints, p. arXiv:1706.09359

LSST Science Collaboration et al., 2009, preprint, (arXiv:0912.0201)

Lacasa F., Kunz M., 2017, A&A, 604, A104

Lacasa F., Rosenfeld R., 2016, J. Cosmology Astropart. Phys., 8, 005

Lacasa F., Lima M., Aguena M., 2016, preprint, (arXiv:1612.05958)

Lanusse F., Ma Q., Li N., Collett T. E., Li C.-L., Ravanbakhsh S., Mandelbaum R., Póczos B., 2018, MNRAS, 473, 3895

Laureijs R., et al., 2011, preprint (arXiv:1110.3193)

Lazanu A., Liguori M., 2018, J. Cosmology Astropart. Phys., 2018, 055

Lazanu A., Giannantonio T., Schmittfull M., Shellard E. P. S., 2016, Phys. Rev. D, 93, 083517

Lazeyras T., Wagner C., Baldauf T., Schmidt F., 2016, J. Cosmology Astropart. Phys., 2016, 018

Lemaître G., 1931, MNRAS, 91, 483

Lemaître G., 1933, Annales de la Société Scientifique de Bruxelles, 53

Li Y., Hu W., Takada M., 2014, Phys. Rev. D, 89, 083519

Lifshitz E., 1946, J. Phys.(USSR), 10, 116

Lima M., Hu W., 2007, Phys. Rev. D, 76, 123013

Lin C.-A., Kilbinger M., 2015, A&A, 576, A24

Lin C.-A., Kilbinger M., 2018, A&A, 614, A36

Liu J., Madhavacheril M. S., 2019, Phys. Rev. D, 99, 083508

Liu J., Petri A., Haiman Z., Hui L., Kratochvil J. M., May M., 2015, Phys. Rev. D, 91, 063507

Lochner M., McEwen J. D., Peiris H. V., Lahav O., Winter M. K., 2016, ApJS, 225, 31

Loh J. M., 2008, ApJ, 681, 726

Loverde M., Afshordi N., 2008, Phys. Rev. D, 78, 123506

M IAglom A., A Silverman R., 1964, SERBIULA (sistema Librum 2.0), 71

Ma C.-P., Fry J. N., 2000, ApJ, 543, 503

Ma Z., Hu W., Huterer D., 2006, ApJ, 636, 21

Makino N., Sasaki M., Suto Y., 1992, Phys. Rev. D, 46, 585

Maldacena J., 2003, J. High Energy Physics, 2003, 013

Marcori O. H., Pereira T. S., 2017, J. Cosmology Astropart. Phys., 2017, 032

Marian L., Smith R. E., Bernstein G. M., 2009, ApJ, 698, L33

Marian L., Smith R. E., Hilbert S., Schneider P., 2013, MNRAS, 432, 1338

Matsubara T., 1995, ApJS, 101, 1

Matsubara T., 2000, arXiv Astrophysics e-prints,

Matsubara T., Jain B., 2001, ApJ, 552, L89

Maturi M., Angrick C., Pace F., Bartelmann M., 2010, A&A, 519, A23

Mcclelland J., Silk J., 1978, ApJS, 36

Mead A. J., Peacock J. A., Heymans C., Joudaki S., Heavens A. F., 2015, MNRAS, 454, 1958

Mecke K. R., Buchert T., Wagner H., 1994, A&A, 288, 697

Mo H. J., White S. D. M., 1996, MNRAS, 282, 347

Mo H. J., Jing Y. P., White S. D. M., 1997, MNRAS, 284, 189

Mohammed I., Seljak U., 2014, MNRAS, 445, 3382

Möller A., et al., 2016, J. Cosmology Astropart. Phys., 2016, 008

Monaco P., 1995, ApJ, 447, 23

Munshi D., Heavens A., 2010, MNRAS, 401, 2406

Munshi D., van Waerbeke L., Smidt J., Coles P., 2012, MNRAS, 419, 536

Mustafa M., Bard D., Bhimji W., Lukić Z., Al-Rfou R., Kratochvil J. M., 2019, Computational Astrophysics and Cosmology, 6, 1

Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563

Ntelis P., et al., 2017, J. Cosmology Astropart. Phys., 2017, 019

Pajer E., Schmidt F., Zaldarriaga M., 2013, Phys. Rev. D, 88, 083502

Peacock J. A., Dodds S. J., 1996, MNRAS, 280, L19

Peacock J. A., Heavens A. F., 1990, MNRAS, 243, 133

Peebles P. J. E., 1980, The large-scale structure of the universe

Perraudin N., Defferrard M., Kacprzak T., Sgier R., 2019, Astronomy and Computing, 27, 130

Petri A., Liu J., Haiman Z., May M., Hui L., Kratochvil J. M., 2015, Phys. Rev. D, 91, 103511

Planck Collaboration 2014, A&A, 571, A16

Planck Collaboration 2016, A&A, 594, A13

Planck Collaboration 2018, arXiv e-prints, p. arXiv:1807.06209

Planck Collaboration et al., 2016, A&A, 594, A18

Press W. H., Schechter P., 1974, ApJ, 187, 425

Pyne T., Birkinshaw M., 1996, ApJ, 458, 46

Ravanbakhsh S., Oliva J., Fromenteau S., Price L. C., Ho S., Schneider J., Poczos B., 2017, arXiv e-prints, p. arXiv:1711.02033

Reimberg P., Bernardeau F., Nishimichi T., Rizzato M., 2018, arXiv e-prints, p. arXiv:1811.02976

Ribli D., Ármin Pataki B., Zorrilla Matilla J. M., Hsu D., Haiman Z., Csabai I., 2019, arXiv e-prints, p. arXiv:1902.03663

Rimes C. D., Hamilton A. J. S., 2005, MNRAS, 360, L82

Rimes C. D., Hamilton A. J. S., 2006, MNRAS, 371, 1205

Rizzato M., Benabed K., Bernardeau F., Lacasa F., 2018, arXiv e-prints, p. arXiv:1812.07437

Robertson H. P., 1935, ApJ, 82, 284

Robertson H. P., 1936a, ApJ, 83, 187

Robertson H. P., 1936b, ApJ, 83, 257

Samtleben D., Staggs S., Winstein B., 2007, Annu. Rev. Nucl. Part. S, 57, 245

Sato J., Takada M., Jing Y. P., Futamase T., 2001, ApJ, 551, L5

Sato M., Hamana T., Takahashi R., Takada M., Yoshida N., Matsubara T., Sugiyama N., 2009, ApJ, 701, 945

Schmidt F., 2016, Phys. Rev. D, 93, 063512

Schmitz D. M., Hirata C. M., Blazek J., Krause E., 2018, J. Cosmology Astropart. Phys., 2018, 030

Schneider P., 1985, A&A, 143, 413

Schneider P., 1996, MNRAS, 283, 837

Schneider M. D., Bridle S., 2010, MNRAS, 402, 2127

Schneider P., van Waerbeke L., Kilbinger M., Mellier Y., 2002, A&A, 396, 1

Schneider A., et al., 2016, J. Cosmology Astropart. Phys., 2016, 047

Scoccimarro R., 2001, in Fry J. N., Buchler J. R., Kandrup H., eds, Vol. 927, The Onset of Nonlinearity in Cosmology. pp 13–23 (arXiv:astro-ph/0008277), doi:10.1111/j.1749-6632.2001.tb05618.x

Scoccimarro R., Colombi S., Fry J. N., Frieman J. A., Hivon E., Melott A., 1998, ApJ, 496, 586

Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, ApJ, 546, 20

Seljak U., 2000, MNRAS, 318, 203

Seljak U., Vlah Z., 2015, Phys. Rev. D, 91, 123516

Sellentin E., 2015, MNRAS, 453, 893

Sellentin E., Quartin M., Amendola L., 2014, MNRAS, 441, 1831

Semboloni E., Hoekstra H., Schaye J., van Daalen M. P., McCarthy I. G., 2011, MNRAS, 417, 2020

Sherwin B. D., Zaldarriaga M., 2012, Phys. Rev. D, 85, 103523

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Shirasaki M., Takada M., Miyatake H., Takahashi R., Hamana T., Nishimichi T., Murata R., 2017, MNRAS, 470, 3476

Sirko E., 2005, ApJ, 634, 728

Smith R. E., et al., 2003, MNRAS, 341, 1311

Smith R. E., Sheth R. K., Scoccimarro R., 2008, Phys. Rev. D, 78, 023523

Takada M., Bridle S., 2007, New J. Phys., 9, 446

Takada M., Hu W., 2013, Phys. Rev. D, 87, 123504

Takada M., Jain B., 2003, MNRAS, 340, 580

Takada M., Jain B., 2009, MNRAS, 395, 2065

Takahashi R., Sato M., Nishimichi T., Taruya A., Oguri M., 2012, ApJ, 761, 152

Taruya A., Bernardeau F., Nishimichi T., Codis S., 2012, Phys. Rev. D, 86, 103528

Taylor A., Joachimi B., Kitching T., 2013, MNRAS, 432, 1928

Tegmark M., 1997, Phys. Rev. Lett., 79, 3806

Tegmark M. e. a., 2004, ApJ, 606, 702

Tegmark M., Taylor A. N., Heavens A. F., 1997, ApJ, 480, 22

Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878

Troxel M. A., Ishak M., 2015, Phys. Rep., 558, 1

Valageas P., Nishimichi T., 2011a, A&A, 527, A87

Valageas P., Nishimichi T., 2011b, A&A, 532, A4

Valageas P., Nishimichi T., Taruya A., 2013, Phys. Rev. D, 87, 083522

Van Waerbeke L., et al., 2013, MNRAS, 433, 3373

Wagner C., Schmidt F., Chiang C.-T., Komatsu E., 2015a, MNRAS, 448, L11

Wagner C., Schmidt F., Chiang C.-T., Komatsu E., 2015b, J. Cosmology Astropart. Phys., 2015, 042

Weedman D. W., 1986, Quasar astronomy

Weinberg S., 1972, Gravitation and Cosmology. John Wiley and Sons, New York, http://www-spires.fnal.gov/spires/find/books/www?cl=QC6.W431

White M., Carlstrom J. E., Dragovan M., Holzapfel W. L., 1999, ApJ, 514, 12

Wittman D. M., Tyson J. A., Kirkman D., Dell'Antonio I., Bernstein G., 2000, Nature, 405, 143

Yang X., Kratochvil J. M., Wang S., Lim E. A., Haiman Z., May M., 2011, Phys. Rev. D, 84, 043529

Yankelevich V., Porciani C., 2019, MNRAS, 483, 2078

Zel'dovich Y. B., 1970, A&A, 5, 84

de Putter R., Wagner C., Mena O., Verde L., Percival W. J., 2012, J. Cosmology Astropart. Phys., 2012, 019