

Universidade de São Paulo
Instituto de Física

Cosmologia no cone de luz: uma análise de múltiplos traçadores do espectro de potência angular

Ian Lucas Tashiro

Orientador: Prof. Dr. Luís Raul Weber Abramo

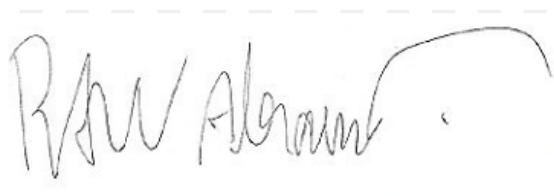
Dissertação de mestrado apresentada ao Instituto de Física da Universidade de São Paulo, como requisito parcial para a obtenção do título de Mestre(a) em Ciências.

Banca Examinadora:

Prof(a). Dr(a). Luís Raul Weber Abramo (Orientador - IFUSP)

Prof(a). Dr(a). Thiago dos Santos Pereira (UEL)

Prof(a). Dr(a). Valerio Marra (UFES)

A handwritten signature in blue ink, enclosed in a dashed rectangular box. The signature appears to read 'L. R. W. Abramo'.

São Paulo
2022

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Tashiro, Ian Lucas

Cosmologia no cone de luz: uma análise de múltiplos traçadores do espectro de potência angular. São Paulo, 2022.

Dissertação (Mestrado) – Universidade de São Paulo. Instituto de Física. Depto. de Física Matemática.

Orientador: Prof. Dr. Luís Raul Weber Abramo
Área de Concentração: Cosmologia

Unitermos: 1. Cosmologia; 2. Galáxias; 3. Estrutura do universo.

USP/IF/SBI-020/2022

University of São Paulo
Physics Institute

Cosmology in the lightcone: a multi-tracer analysis of the angular power spectrum

Ian Lucas Tashiro

Supervisor: Prof. Dr. Luís Raul Weber Abramo

Dissertation submitted to the Physics Institute of the University of São Paulo in partial fulfillment of the requirements for the degree of Master of Science.

Examining Committee:

Prof. Dr. Luís Raul Weber Abramo (Supervisor - IFUSP)

Prof. Dr. Thiago dos Santos Pereira (UEL)

Prof. Dr. Valerio Marra (UFES)

São Paulo
2022

Abstract

The upcoming next-generation of wide-area cosmological surveys will provide us with an unprecedented amount of high-quality data. This will allow us to constrain cosmological parameters with sufficient precision to distinguish between many competing theories that try to explain the different mechanisms through which the large scale structure of the universe forms, and maybe even point to new paths we haven't explored yet. However, these surveys will also bring a new set of challenges on how to properly extract the maximum amount of information possible from these huge data sets. Many commonly used approximations will no longer be valid, and in particular the standard Fourier power spectrum approach will not be able to capture the full information about the two-point statistics of tracers of the large scale structure in redshift space. An alternative is to work in harmonic space with the angular power spectrum $C_\ell(z, z')$, which naturally includes wide-angle effects and has the potential to provide a unified framework to combine these surveys in the future. In this work we develop an important tool towards that goal, by deriving an exact expression for the multi-tracer Fisher matrix of the angular power spectrum, both in real and redshift spaces.

Keywords: Large scale structure; Fisher matrix; Angular power spectrum; Multi-tracer; Redshift space distortions.

Resumo

A próxima geração de levantamentos cosmológicos em grandes áreas do céu vão nos proporcionar uma enorme quantidade de dados de alta qualidade. Isso nos permitirá restringir parâmetros cosmológicos com precisão suficiente para distinguir entre diversas teorias que tentam explicar os diferentes mecanismos que regem a formação da estrutura em larga escala do universo, e talvez até apontar para novos caminhos que ainda não foram explorados. Entretanto, esses levantamentos também trazem novos desafios sobre como extrair o máximo de informação possível desses enormes conjuntos de dados. Muitas aproximações comumente utilizadas deixarão de ser válidas e em particular a abordagem usual, que se utiliza do espectro de potência de Fourier, não é capaz de capturar toda a informação sobre a função de correlação de dois pontos dos traçadores da estrutura em larga escala incluindo distorções de redshift. Uma alternativa é trabalhar no espaço harmônico com o espectro de potência angular $C_\ell(z, z')$, que naturalmente inclui efeitos causados por largas separações angulares, e que tem o potencial de fornecer um formalismo unificado para combinar esses levantamentos no futuro. Nesta dissertação, uma ferramenta importante para alcançar esse objetivo é desenvolvida derivando uma expressão exata para a matriz de Fisher do espectro de potência angular para múltiplos traçadores, tanto no espaço real quanto no espaço de redshift.

Palavras-chaves: Estrutura em larga escala; Matriz de Fisher; Espectro de potência angular; Múltiplos traçadores; Distorções no espaço de redshift.

Agradecimentos

Agradeço primeiramente ao meu orientador, Raul, por toda a paciência e confiança em mim, não apenas no mestrado, mas também durante boa parte da minha graduação. Agradeço por todos os conselhos, tanto acadêmicos quanto pessoais em momentos em que eu estava completamente perdido, e por todo o apoio e disposição em me ajudar com tudo que precisei. Agradeço pela motivação e inspiração que me forneceu nesses últimos 4 anos e meio, espero um dia ter pelo menos 5% do seu conhecimento.

Agradeço ao meu colega, João, pela colaboração neste projeto, por todas as discussões sobre cosmologia, por toda a ajuda com conceitos que achei confusos. Aos meus colegas Lucas e Ana, por todas as discussões sobre as listas absurdas de quântica durante aquele curso online para o qual ninguém estava preparado.

Aos colegas da graduação, Nicolas, Leopoldo, Iuri, Alan, Giulia, Martim, Carol, Ulisses e muitos outros que tornaram minha experiência na USP mais agradável. É uma pena que praticamente todo o meu mestrado tenha sido realizado durante uma pandemia global que impediu a nossa convivência durante os últimos dois anos.

Aos meus amigos da república, Igor, Adriano, Fernanda, Magnim, Laura e Hadrien, pela companhia nesses anos de isolamento. Aos amigos do Paraná, Gabriel, Bruno, e Fernanda, por ainda manterem contato depois de todos esses anos, mesmo quando eu desapareço por meses.

Agradeço à toda a minha família, que sempre me apoiou e torceu por mim, meus avós, minhas tias, meu padrasto, meus irmãos, e especialmente minha mãe, e minha vó por ter me recebido em São Paulo de braços abertos. Sei que às vezes pode não parecer, mas sou muito grato a todos.

Às minhas amigas de São Paulo de fora da USP, Sara, Sayuri e Laís, agradeço pelas conversas aleatórias, pelas risadas, e principalmente por rirem das minhas piadas sem graça. Aos meus amigos de fora do Brasil, Christian, Klaarika, Emils, Oscar, Carolina, obrigado por me trazerem um pouco do resto do mundo.

Agradeço, em especial, à minha amiga Júlia, por ser essa pessoa absolutamente incrível, que me inspirou a ser um aluno melhor, uma pessoa melhor, e me ajudou a manter a sanidade em momentos particularmente difíceis. Mais do que ela sabe.

Por fim, agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo financiamento deste trabalho.

This page intentionally not left blank.

"Don't write in yellow."

-Tom Kibble

This page also intentionally not left blank.

Contents

1	Introduction	1
2	Current cosmological surveys	6
3	Galaxy Clustering	12
3.1	The Fourier power spectrum	13
3.2	The angular power spectrum	17
3.2.1	Angular spectra estimation	19
3.2.2	Comparing the harmonic and Fourier approaches	22
3.2.3	A Hybrid approach for the angular power spectrum	23
3.3	Sources of statistical errors	24
4	Redshift space distortions	28
4.1	Peculiar velocities and linear RSD	28
4.2	The angular power spectrum in redshift space	33
5	Fisher matrix in real space	35
5.1	Fisher matrix in the lightcone	37
5.2	The harmonic Fisher matrix in real space	39
5.3	Inverting the data covariance	43
5.4	The Fisher matrix in harmonic space	46
6	Fisher matrix in redshift space	48
6.1	Integrating products of spherical Bessel functions	49
6.2	Inverting the data covariance in redshift space	57
6.3	Fisher and covariance matrices in redshift space	65
7	Conclusions and future prospects	66
	Appendices	68
	Appendix A Derivation of Eq. (4.20)	68
	Appendix B FFTLog	71
	B.1 The algorithm	72

1 Introduction

Until a few decades ago, cosmology suffered from a significant deficiency of data and the tools used to analyze them were adequate to deal only with small areas and volumes. In fact, many different theories on structure formation were consistent with the scarce data available, but cosmological surveys in the past few decades quickly changed this picture. First, we have the cosmic microwave background (CMB), which was famously discovered by accident in 1964 by astronomers Penzias and Wilson, and the first accurate measurements of its temperature anisotropies was performed by COBE (Bennett et al., 1996), which showed that it had a near perfect black body spectrum. Then there is dark matter, which we have known about for almost a century now, ever since Zwicky described in 1934 the need of some sort of invisible *dark* matter to hold the Coma cluster together, but this anomaly was largely ignored for a long time. Only a few decades ago more compelling evidence for dark matter was gathered (see e.g. Bertone et al., 2005 for a review), and welcomed particle physicists into cosmology: currently the most promising candidates are very weakly interacting elementary particles that we haven't been able to detect so far. Finally, two independent teams in the late 90s (Riess et al., 1998; Perlmutter et al., 1999) found evidence of some dark energy that accelerates the expansion of the universe. The simplest explanation for it is a cosmological constant Λ added to Einstein's field equations, but an alternative to introducing a new energy component is modifying our theory of gravity itself.

These groundbreaking discoveries allowed us to discard many of our theories, and were revolutionary to such a degree that we now have a simple and quite elegant standard model of cosmology with only six parameters: Λ CDM. This model has been supported by an overwhelming amount of evidence and extended accordingly in the past couple of decades. It explains a wide variety of different phenomena and observations, from a fraction of a second after the Big Bang to today: the structure of the CMB, the distribution of galaxies on large scales, the accelerated expansion of the universe, the observed abundances of light elements.

Nevertheless, despite its incredible success, passing multiple tests over the past couple of decades, it is incomplete and does not offer a universally accepted solution to all of our problems: the true nature of arguably the two biggest mysteries in cosmology, dark matter and dark energy, still eludes us. On top of that, as the era of precision cosmology began with data from probes such as WMAP in the early 2000s, tensions between large and small scale data sets started to emerge, pointing to possible gaps in this model. The most well known of these tensions is probably the discrepancy between the present value of

Hubble constant H_0 inferred from Planck’s measurements of the CMB, and the value obtained from late-time supernovae observations by the SH0ES collaboration (Di Valentino et al., 2021). Another new generation of surveys such as the upcoming J-PAS (Benitez et al., 2014), DESI (Levi et al., 2019), Euclid (Amendola et al., 2018), PFS (Ellis et al., 2012), LSST (Ivezić et al., 2019), and the recently released DES-Y3 (Sevilla-Noarbe et al., 2021), will map even larger fractions of the observable universe and allow us to constrain cosmological parameters with unprecedented precision, and maybe even allow us to solve these tensions between different data sets and point to new physics beyond the current standard Λ CDM model.

However, nothing in life comes for free. With the much larger sky coverage and increased depth of these surveys, many of the approximations commonly used will no longer be valid (see Bellomo et al., 2020), and the much larger data sets themselves also present another challenge. Much effort has been dedicated to this problem in recent years, but some key aspects are still not very well understood. So it is now more important than ever to find optimal ways to take advantage of the upcoming wealth of data. A very powerful tool in this context is the Fisher matrix formalism, which provides us a way to predict how well we will be able to constrain cosmological parameters with future surveys. Hence, we can test and compare different data analysis pipelines and binning strategies even before any data is collected.

In the current standard cosmological model Λ CDM, we expect a sort of hierarchical formation of structure: smaller objects form first and they merge into larger and larger structures overtime. Gravity makes even the smallest of anisotropies in the density field grow, and today enough time has passed since the primordial universe that we can identify a clear pattern of cosmic filaments and voids that stretch across considerable fractions of the observable universe (see Figure 1). On large scales, however, due to the standard assumption of cosmology that the universe is homogeneous and isotropic, the position of an individual object has no relevance. But due to gravitational attraction, matter tends to agglomerate in regions of higher density, so the correlation between the positions of two galaxies carries physical information contained in the two-point function of the density contrast, and its Fourier transform, the power spectrum. Therefore, we can map out the 3D distribution of matter distribution in the universe, by combining measurements of redshift and angular position of astronomical objects, offering us a way to estimate cosmological parameters and better understand the mechanisms that rule structure formation in the universe.

In order to actually test cosmological models, we must first estimate some quantity from the data which can be predicted by our theories. Considering only one tracer, Feld-

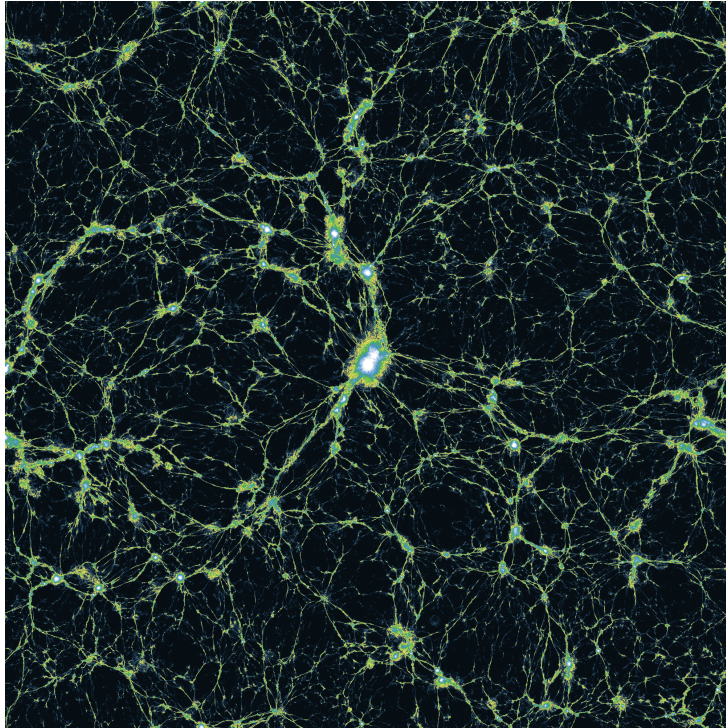


Figure 1: The large scale distribution of dark matter from the most ambitious hydrodynamical simulation to date by the Illustris TNG Collaboration.

man, Kaiser & Peacock (FKP) showed in Feldman et al. (1993) that there is a unique galaxy weighting scheme that leads to an ideal (minimum variance) estimator of the Fourier power spectrum. However, while galaxy positions have been our main source of information due to their sheer numbers, we can also study the large scale structure with a variety of different objects, such as quasars, voids and the Lyman- α forest. These highly nonlinear objects result from very complicated formation processes over long periods of time, which can cause their distribution to differ from that of dark matter. Therefore, these objects are biased *tracers* of the underlying matter density field, as first suggested by Kaiser (1984). This bias is a function of redshift and can depend on the properties of the objects such as luminosity, color and morphology, and can also be scale dependent, though this dependence is weak and tends to a constant on large scales (Mann et al., 1998). Since different tracers aren't independent, and actually map the same underlying distribution of dark matter (although in different ways, with different biases), one might hope to combine them in the analysis to extract as much information as possible from our data sets.

In fact, it has become evident that a multi tracer analysis offers many benefits. For instance, the two main sources of uncertainty in our measurements are shot noise, which arises from the fact that the distribution of galaxies is basically a discrete Poisson process,

and cosmic variance, a fundamental statistical limitation due to the fact that we cannot recreate the big bang in a laboratory with the same initial conditions as our own and wait billions of years to measure other universes. Nevertheless, we can overcome the limitations of cosmic variance on bias-sensitive parameters (Seljak, 2009; Seljak et al., 2009) by combining multiple tracers of the underlying density field.

Of course, designing a multi tracer survey isn't trivial, as different tracers may require different observational techniques for target selection, and different treatment of systematic effects. Fortunately, we can split a single tracer survey into multiples samples by color or luminosity (Blake et al., 2013; Ross et al., 2014), or combine observations made by different surveys (Gil-Marín et al., 2016), although this is somewhat limited by the fact that most surveys are designed to be complimentary to others in terms of sky coverage and redshift range, so the overlap might be too small for any practical purposes.

It has also been shown in Abramo et al. (2016) that the multi-tracer optimal estimator (MTOE) is a superior weighting scheme for the density fields of any number of tracers. While this is especially interesting for large scale effects, which suffer the most from cosmic variance, it's been shown that the MTOE has advantages in small scales as well (Montero-Dorta et al., 2020). This estimator allows one to simultaneously fit both the matter power spectrum as well as the biases of the tracers, as opposed to the PVP estimator (Percival et al., 2004) for which one must fix the biases. It was also shown to have a better performance than the FKP estimator, to which it reduces in the case of a single tracer.

So the problem has pretty much been now solved for the simplified scenario where we compute the Fourier spectrum of a volume at fixed redshift (a "snapshot" mapped through Cartesian coordinates). Since we can measure redshifts with outstanding precision, to which we can assign a distance, this standard approach of using the 3D Fourier matter power spectrum to constrain cosmological parameters with clustering data has served us well for many years. However, this approach presents many issues, which are becoming ever more apparent now. First, we can't directly measure distances to astrophysical objects, we can only infer them after choosing a specific cosmological model to translate from redshift space to real space, which can introduce a bias in the results, and not many studies have been done on the consequences of making a bad choice. More importantly, the Fourier approach is only valid in the flat-sky limit, which will not be the case for upcoming surveys, and it also requires the use of large redshift bins, effectively ignoring cosmic evolution and suppressing the effects of redshift space distortions caused by peculiar velocities (see Section 4).

One can circumvent the problems of the Fourier spectrum by decomposing the 3D lightcone into 2D redshift slices and work directly in harmonic space, with the formalism

of the angular power spectrum, which has already been widely developed and applied by the CMB community (e.g. Ade et al., 2014), and has also the advantage of having, in an idealized full sky survey, diagonal covariance matrices because the harmonic modes are independent on linear scales. The consistencies between galaxy clustering and cosmic shear analysis in harmonic and real space were verified in Andrade-Oliveira et al. (2021) and Doux et al. (2021), respectively.

Another advantage of working in harmonic space is that wide-angle and redshift space distortion (RSD) effects are naturally included. These are relevant effects that must be taken into account in future clustering analyses to avoid biasing our results, but RSDs are also especially important as a tool to help us understand dark energy, since they allow us to constrain the growth rate of large scale structure, which is different for each model.

However, it is still not clear how to weigh the different tracers in spherical shells (i.e., for different redshift slices on the past lightcone). In a series of papers that advocate for this transition to a tomographic pipeline, Tanidis and Camera (2019, 2021) discuss the importance of properly including redshift space distortions in angular power spectrum analyses, and also showcase the power of the multi tracer technique by constraining the sum of neutrino masses. The advantages of working in harmonic space were also showcased in Loureiro et al. (2019), where it was shown that this approach yields comparable, if not better constraints, even for spectroscopic data.

In this work, we further develop our collective knowledge on angular clustering statistics by deriving an exact expression for the multi tracer Fisher matrix of the angular power spectra, including RSD and without the commonly used flat-sky and Limber's approximations. This allows us to predict how well a survey can constrain cosmological parameters, and is also the basis necessary for a future work that will tell us the best way to combine multiple tracers of the large scale structure over the same volumes, in a way that maximizes the cosmological information we can extract from the data.

2 Current cosmological surveys

The two main probes that have been used to study the large scale structure of the universe are measurements of cosmic shear and galaxy clustering, and in this work we focus on the latter. Current galaxy surveys rely on basically two observational techniques: broad band imaging and spectroscopy. With the former, our instruments are designed to collect the light of as many objects as possible in multiple frequency bands, without any discrimination other than some lower flux limit to ensure a high quality of the data. These photometric imaging surveys give us relatively poor measurements on redshifts (with a precision of around $\sim 4\%$), in exchange for mapping a huge amount of objects, which allows the selection of specific galaxies or quasars to which we can point our spectroscopic telescopes and measure their spectrum much more precisely for a better estimation of their redshifts ($\sim 0.1\%$) in a follow up survey. This, of course, introduces some selection bias to our analysis and also comes at a cost of longer observational time, representing a trade-off between more accurate information about redshift and the quantity of objects we can map.

Since the speed of light is finite, measurements of distances also correspond to different times in the history of the universe, so we can study the evolution of large scale structures by observing their growth over our past lightcone. These observations contain valuable information about the nature of both dark matter and dark energy. The former acts in a way to pull matter together and accelerate the formation of large scale structures, while the latter has an opposite effect and effectively tries to push matter apart as the universe expands, slowing down this structure forming process, in a perpetual game of cosmic tug of war.

One of the arguably most important still ongoing surveys is the Sloan Digital Sky Survey (SDSS), which began in the year 2000 its mission of imaging the sky in five broad bands (u, g, r, i, z), and measuring the spectra of galaxies, quasars and stars. It was extended multiple times to also explore the structure of the Milky Way and observe supernova Ia events, and continued to map even larger fractions of the sky. The 7th Data Release (DR7) marked the completion of its initial goals, after cataloging over 18 million galaxies over an area of 8423 deg^2 ($1/5$ of the sky) in the northern hemisphere and measuring the spectra of over 930,000 galaxies and 120,000 quasars over 8000 deg^2 . The first angular power spectrum analysis for the full SDSS galaxy sample was then performed in Hayes et al. (2011), though the constraints achieved were weak (but consistent with others) due to computational limitations.

The survey was then extended for a third phase and SDSS-III's Baryon Oscillation Spectroscopic Survey (BOSS) was designed to measure the imprints left by the Baryon Acoustic Oscillations (BAO) in the late time galaxy distribution, which provides a standard ruler for the measurement of angular diameter distances as a function of redshift and also the Hubble constant, giving us another way to constrain the accelerated expansion of the universe completely independent from supernovae. It consisted of two simultaneous spectroscopic surveys: one measured the redshifts of 1.5 million red luminous galaxies (LRGs) extending to $z = 0.7$ over an area of 9329 deg^2 , and the other observed 150,000 quasars in the redshift range $2.15 \leq z \leq 3.5$ for a $Ly\alpha$ forest analysis, which allowed BOSS to successfully obtain the first reliable measurements of the BAO peaks. Its final DR12 (Alam et al., 2017) spectroscopic galaxy sample was used by Loureiro et al. (2019) to show that a harmonic approach provides competitive cosmological constraints for Λ CDM, w CDM and the sum of neutrino masses using the full shape of the angular power spectrum, whilst having numerous advantages over the Fourier method (see Section 3).

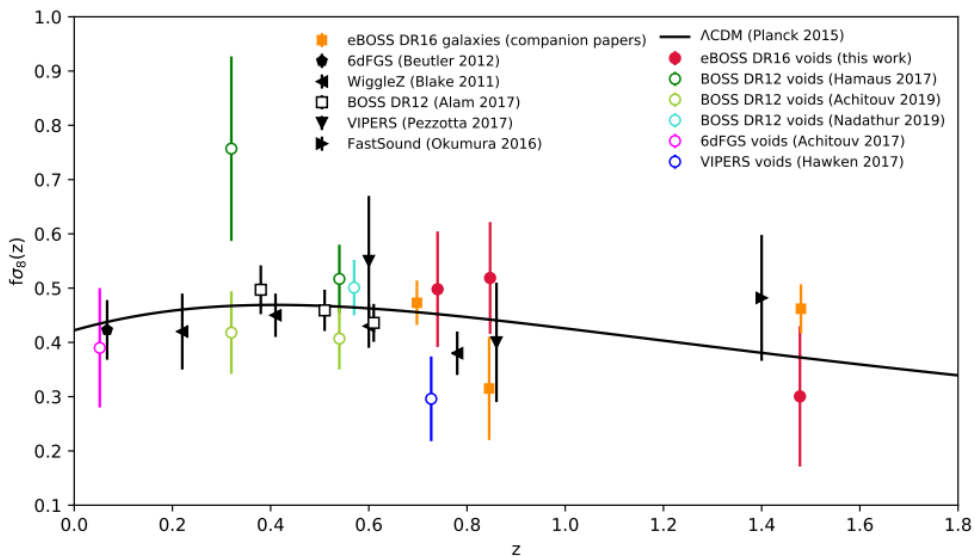


Figure 2: Constraints on $f\sigma_8$ obtained for different redshifts from multiple surveys. Extracted from Aubert et al. (2020)

BOSS was further extended in SDSS-IV with eBOSS, which complemented the previous measurements by mapping an even larger volume. This new survey targeted 300,000 LRGs in a different redshift range of $0.6 < z < 1.0$, over 6000 deg^2 , and was also capable of observing 175,000 emission line galaxies (ELGs) up to $z = 1.1$ over 1200 deg^2 . It also more than tripled the number of observed quasars, to a total of $\sim 500,000$, and greatly increased the redshift range of these observations to $0.8 < z < 3.5$. This was the first

survey to measure the redshifts of ELGs and it is currently the only multi tracer sample with a large overlapping volume from a single survey. A study on BAO and RSD was performed in Zhao et al. (2021) using data from the previous release (DR16) and combining these two tracers in Fourier space, which raised the significance level of a nonzero Ω_Λ to $\sim 11\sigma$. Combined with previous phases of SDSS and Planck’s CMB measurements, these observations allowed us to precisely measure 80% of the expansion history of the universe and constrain cosmological parameters such as Ω_Λ , H_0 and σ_8 with a precision better than 1%, and the growth rate to 6% (See Figure 2), and showed no need to extend the standard cosmological model beyond the flat Λ CDM. With its latest DR17 (Alam et al., 2021), SDSS provided the most detailed three-dimensional map of the universe to date.

Obtaining such precise spectra is, however, a very time consuming process, because we must target each object individually and observe it for a sufficient time to collect enough light and obtain its precise spectrum. Additionally, we have an increasing role being played by photometric surveys, that sacrifice some precision in their redshift measurements in order to map a much larger quantity of objects, with better signal-to-noise ratios. These surveys use a technique known as photo-z estimation, which relies on the observation of the sky in multiple frequency bands. For a given object, we can combine its intensity information from different broad-band filters to obtain an estimate for their redshifts. These are obviously far less precise than spectroscopic ones, but this method is significantly faster and increases the number of galaxies we can observe in a given amount of time by a factor of a few hundred.

The Dark Energy Survey is a fine example of a purely photometric survey, designed to help uncover the secrets of the accelerated expansion of the universe through type Ia supernovae, BAOs, galaxy clusters and weak gravitational lensing. It observed the galaxy distribution in southern sky in five photometric bands, with long integration times over 6 years (2013-2019) to detect fainter galaxies than previous surveys, as well as type Ia supernovae in a smaller patch of the sky. With its recently released three years of data (Y3) (Sevilla-Noarbe et al., 2021), DES mapped nearly 226 million galaxies fit for cosmological analysis, covering redshifts of approximately $z \sim 0.2-1.2$ over an area of $\sim 5000 \text{ deg}^2$, and provided the largest photometric data set we currently have.

A combination of the DES-Y3 3x2pt analysis with CMB, BAO, RSD, and SNe Ia data allowed us to obtain the most precise constraints on the Λ CDM and w CDM models to date (see Figure 3). The full observations over its 6 years of operations will deliver data with roughly double the exposure time and continue to support the development of new data processing techniques for even more ambitious surveys in the future.

One of these next generation surveys is the Dark Energy Spectroscopic Instrument, which began operating in 2019 and is considered the successor to BOSS. It will target over 30 million objects and measure redshifts across a wide area of $14,000 \text{ deg}^2$, including LRGs up to $z = 1$, quasars in the range $2.1 < z < 3.5$, as well as ELGs up to $z = 1.7$ which will represent its largest sample. It is expected to significantly improve the precision on BAO and the sum of neutrino masses, and place strong constraints on theories of modified gravity and inflation by measuring the spectral index n_s .

The LSST (Ivezic et al., 2019), on the other hand, is the most ambitious wide area photometric survey ever proposed, set to begin collecting data in 2022 in six band filters (u, g, r, i, z, y). The main program will take 90% of its observing time and will be dedicated to mapping a $18,000 \text{ deg}^2$ region over 800 times during its 10 years of operations, which will allow us to observe much fainter galaxies than existing surveys and probe much deeper redshifts than ever before. This will give us databases with several billion galaxies up to $z \sim 3$, as well as millions SNe Ia up to $z \sim 1$. The remaining 10% will be dedicated to

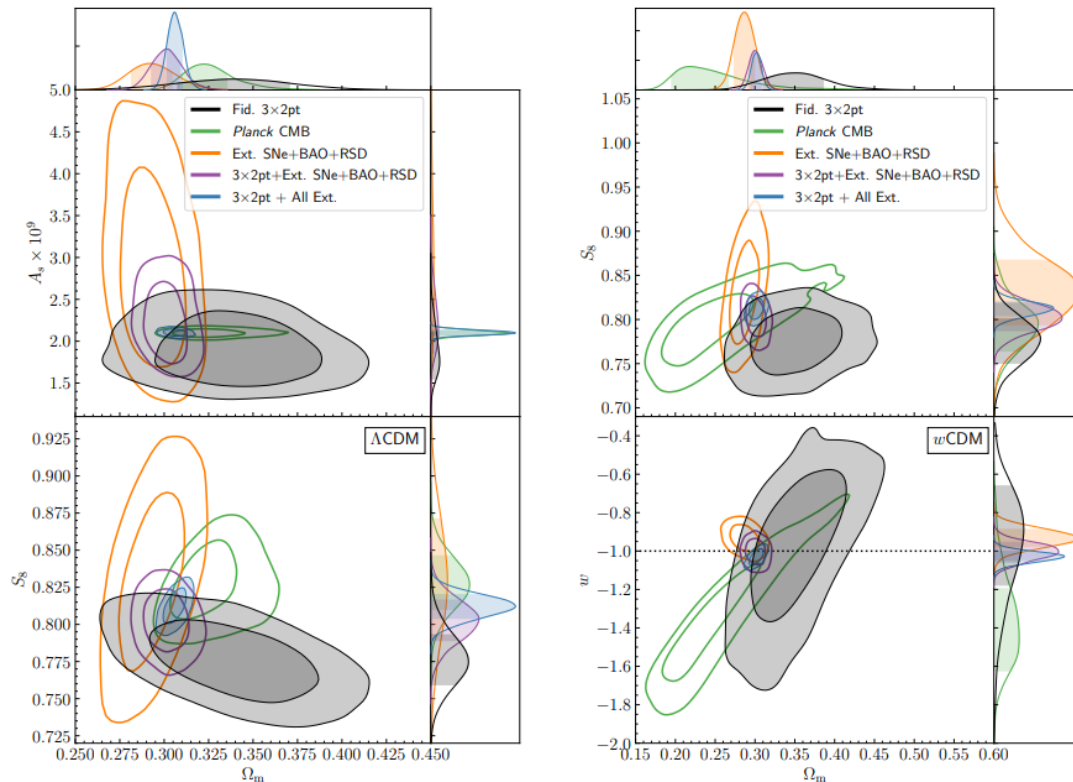


Figure 3: Marginalized constraints on parameters in the Λ CDM (left) and w CDM (right) models, from the DES 3x2pt, *Planck* CMB and external BAO+ RSD+SNe Ia data. The combinations of these data sets are shown in blue and represent the best constraints on these cosmological parameters to date. Extracted from Abbott et al. (2022)

other programs to improve the sky coverage for specific regions, such as the ecliptic and galactic planes, which will further increase the total coverage to over half of the whole sky.

Another upcoming mission is the Euclid space telescope (Moneti et al., 2022), whose launch is scheduled for 2023. It will be complemented by ground-based telescopes in both hemispheres to obtain photometric redshifts in 7 different broad optical filters (g,r,i,z from the ground, and Y,J,H from Euclid) in an area of $15,000 \text{ deg}^2$. It is expected to measure over 30 million spectroscopic redshifts to be used for galaxy clustering and BAO in the range $0.7 < z < 1.8$, and observe over 1.5 billion sources suitable for weak lensing analysis up to $z = 2$, with a precision 50 times better than current ground-based surveys. Its main objective is to help us understand the physical origins of the accelerated expansion of the universe by, e.g., measuring the growth index γ with a 1σ precision, enough to distinguish between a large variety of modified gravity theories and general relativity.

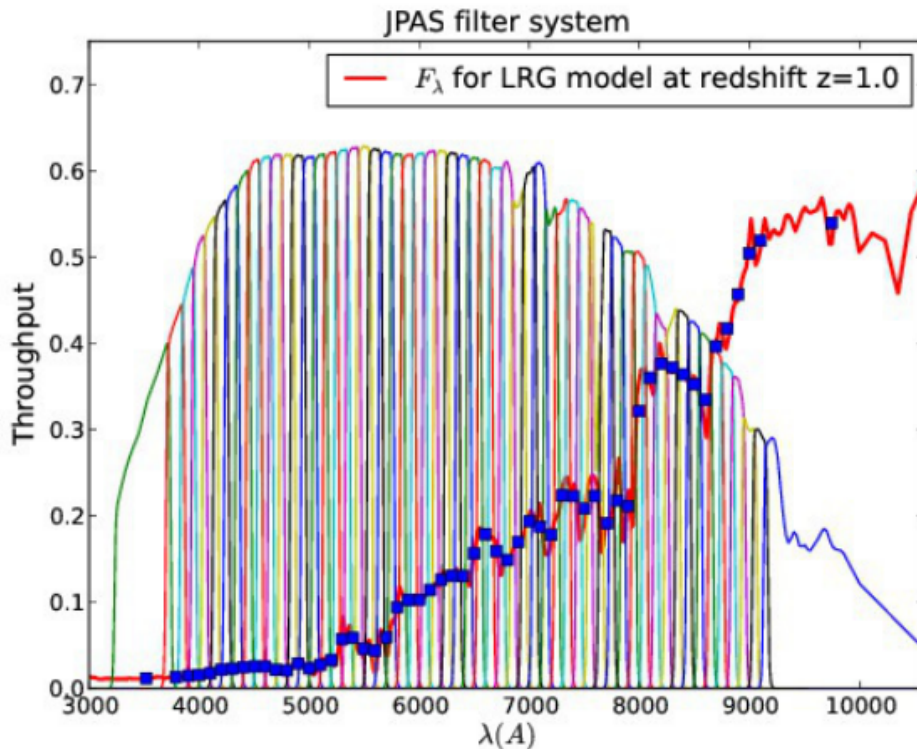


Figure 4: The J-PAS filter system, extracted from Benitez et al. (2014). The narrow filters have a FWHM of 145\AA and are spaced by about 100\AA , meaning they have a significant overlap, which allows for a better quality of the estimated photo- z .

While all of these surveys are based on either broad band imaging or moderate resolution spectroscopy, each having their advantages and shortcomings, the recently started J-PAS (Benitez et al., 2014) proposes to unite the best of both worlds using 56 narrow

band (see Figure 4) filters instead. It will be able to detect several tracers such as millions of LRGs (up to $z \sim 1$), ELGs (up to $z \sim 1.4$) and quasars (up to $z \sim 5$), and by mapping a large area of 8500 deg^2 ($\sim 1/5$ of the sky) while still maintaining near-spectroscopic accuracies on redshifts ($\sim 0.3\%$). This will allow for a competitive measurement of both the angular and radial components BAOs, which will provide an independent determination of the equation of state parameter w with precisions of order $< 4\%$. Recently, the J-PAS pathfinder survey miniJPAS Bonoli et al. (2021), confirmed the quality of that data set in terms of depth, photometric redshift accuracy and the ability to distinguish between the different types of objects using the pseudo-spectra provided by the narrow-band filters.

This high quality data will, however, come at a cost: the LSST alone is expected to bring in 20 TB of data per night of observation. So we must develop an efficient data analysis pipeline to extract as much information as possible from the upcoming wealth of data. Some form of data compression has typically been used to make our analysis computationally realistic (Tegmark et al., 1998; Alonso, 2018), however, care must be taken to avoid losing precision or even suppressing signs of new physics (Heavens et al., 2020). The near full-sky coverage also means that many of the commonly used approximations used for past surveys will no longer be valid. I will discuss these limitations in the following section.

3 Galaxy Clustering

One of the main ways we have been trying to study the distribution of matter in the universe is through the clustering of galaxies, because they are the easiest to detect and most abundant objects we can observe. For that reason, we build very large telescopes to measure the positions of galaxies and construct a three dimensional map of their positions in the sky. On large scales, galaxies trace the dark matter distribution up to a factor of proportionality known as bias, $b(z)$, which is in principle a function of scale, but is often assumed to be constant to simplify calculations. This assumption is somewhat justified by works such as Verde et al. (2002) who found that, on large scales, galaxies have a linear bias close to unity, but we might have to measure this quantity more precisely for the full picture.

It is, however, hopeless to try and create a model for the exact positions of galaxies in our universe and describe the distribution of matter in the universe in a deterministic way. Instead, we try to describe the statistical properties of the matter density field. Let $\rho_g(\vec{r})$ be the galaxy density at a position $\vec{r}(z)$ in real space¹, and $\bar{\rho}_g$ the mean density over all space. The galaxy density contrast is defined by

$$\delta_g(\vec{r}, z) = \frac{\rho_g(\vec{r}, z) - \bar{\rho}_g}{\bar{\rho}_g} \quad (3.1)$$

These density fluctuations grow over time because galaxies interact gravitationally with each other, so regions of high density tend to become denser with time, while underdense regions will become even less dense. Two-point functions of these overdensities have proven to be very useful summary statistics because of the near Gaussian nature of the matter density field on large scales, and there are a few different ways we can go about it.

In this Section I will briefly review the approaches typically used to analyze clustering data: the Fourier power spectrum $P(k)$ used mostly with spectroscopic data, and the angular spectrum C_ℓ , which has been in widespread use by the CMB community but has also been used with photometric data, which sacrifice some precision in the radial information in order to map the angular position of a much larger number of objects. Here I focus on the latter because that is the formalism we further developed with this work, for reasons that will become clear with this section.

¹This definition is easier if we think of cells of volume ΔV , in that case $\rho_g(\vec{x}_i) = N_g(\vec{x}_i)/\Delta V$, where \vec{x}_i is the position of any given cell.

3.1 The Fourier power spectrum

Primordial perturbations in the early universe are considered the seeds of all structure we observe today. They are a consequence of quantum fluctuations during inflation, therefore we can think of the matter distribution in the large scale structure of the universe as the result of a stochastic process, which is described by the theory of random fields (see e.g. *Primordial Cosmology* by Peter and Uzan). A fundamental assumption in cosmology is that this field is statistically homogeneous and isotropic, and hence that it is completely determined by its irreducible moments, or correlation functions, with the first two being the mean and the covariance, also called the two-point correlation function in the context of large scale structure.

Our observations further justify working with the simple case of a linear Gaussian field, whose Fourier modes evolve independently, These fields are completely described by the mean and the two-point correlation function, defined as the expectation value:

$$\xi_g(|\vec{r} - \vec{r}'|) = \langle \delta_g(r) \delta_g^*(r') \rangle \quad (3.2)$$

Physically, this expresses the excess probability of finding any two galaxies separated by a distance $\Delta\vec{r} = \vec{r} - \vec{r}'$. Hence, it is a measure of the distribution of matter. Such perturbative quantities are usually assumed to be statistically homogeneous and isotropic, that is, invariant under translation and rotations. The former implies that the correlation is a function of only the distance between the galaxies, and only of the absolute value of that distance, as a consequence of the latter.

In Fourier space, many of our equations have a simpler form, so it is useful to take the Fourier transform of the galaxy density contrasts. Here we have to invoke the cosmological principle, and assume that the density $\rho(\vec{r})$ represents the real cosmic density field reasonably well such that it can be taken to be periodic inside the survey volume. In that case, the Fourier transform of the density fluctuations is:

$$\hat{\delta}_g(\vec{k}) = \int d^3r e^{i\vec{k}\cdot\vec{r}} \delta_g(\vec{r}); \quad \delta_g(\vec{r}) = \int \frac{d^3k}{(2\pi)^3} e^{-i\vec{k}\cdot\vec{r}} \hat{\delta}_g(\vec{k}) \quad (3.3)$$

which can also be written in terms of the growth function $D(z)$ and the scale-independent bias $b(z)$ to make the dependence on redshift more explicit:

$$\delta_g(\vec{k}, z) = D(z) \delta_g(\vec{k}, 0) = b(z) D(z) \delta(\vec{k}, 0) \quad (3.4)$$

From this definition, we write a related quantity, the galaxy power spectrum, $P_g(k)$, which is often more convenient to work with because its covariance matrix is more diagonal than that of the correlation function, and it is the quantity actually predicted by our theories of structure formation. It is also more intuitive physically because it cleanly separates well understood large scale processes from non-linear small scale ones. It is defined by the covariance of the Fourier modes:

$$\begin{aligned} \langle \delta_g(\vec{k}, z) \delta_g^*(\vec{k}', z') \rangle &= \int d^3r d^3r' \exp(i\vec{k}\vec{r} - i\vec{k}'\vec{r}') \xi_g(r, r') \\ &= (2\pi)^3 \delta(\vec{k} - \vec{k}') P_g(k, z) \end{aligned} \quad (3.5)$$

with

$$P_g(k, z) \equiv \int d^3r e^{i\vec{k}\cdot\vec{r}} \xi_g(r), \quad \xi_g(r) = \int \frac{d^3k}{(2\pi)^3} e^{-i\vec{k}\cdot\vec{r}} P_g(k, z) \quad (3.6)$$

The Dirac delta function in Eq. (3.5) expresses the assumed statistical homogeneity of galaxy clustering, while the fact that $P_g(k)$ depends only on the absolute value of the wavevector \vec{k} expresses isotropy. This quantity maps the galaxy distribution in the universe, and can be related to the power spectrum of the underlying matter density field through the galaxy bias:

$$P_g(k, z) = b(z)^2 P(k) \quad (3.7)$$

If the density field is Gaussian, then the matter power spectrum $P(k)$ provides a complete statistical description of the matter distribution. It should be noted, however, that this approximation eventually collapses once the fluctuations grow large enough, for they cannot remain Gaussian because the density must always be positive. Hence, for small, non-linear scales, we must include higher order terms which are then non-vanishing. But in this work we focus on the linear regime, where inhomogeneities are still small enough to be treated in linear order. Also note that this formalism isn't exclusive to galaxies and can be straightforwardly applied to any other tracers of matter.

Let's now consider the total power spectrum of the actual matter density field, $P(k)$, and think about their connection to the primordial fluctuations in the early universe. They are typically described by the primordial power spectrum, $P_0(k)$, which can be predicted by our theory. These perturbations grow as a consequence of the dynamical interactions

of matter over time, and their linear evolution can be encoded in a transfer function, $T(k)$, which also describes the physics of dark matter and electromagnetism interactions. The linear power spectrum in the present day can then be written as:

$$P_{lin}(k) = T(k)^2 P_0(k) \quad (3.8)$$

The primordial spectrum is usually assumed to follow a power law, $P_0(k) \propto k^n$, while the computation of $T(k)$ is very complicated, but can be performed with numerical codes such as CMBFAST or CAMB. So by measuring the linear power spectrum today, we can straightforwardly test our theoretical predictions and fit cosmological parameters. Though in this work we only consider the linear regime, it is worth noting that the non-linear power spectrum has no analytical solutions. They can be modeled with N-body simulations (see e.g. Peacock and Dodds, 1996). However, this requires a huge amount of computing time, so what we do instead is use the *Halo model* formalism (see Seljak, 2000) to generate analytic fits (*Halofit* in CAMB) for the non-linear spectra of many N-body simulations.

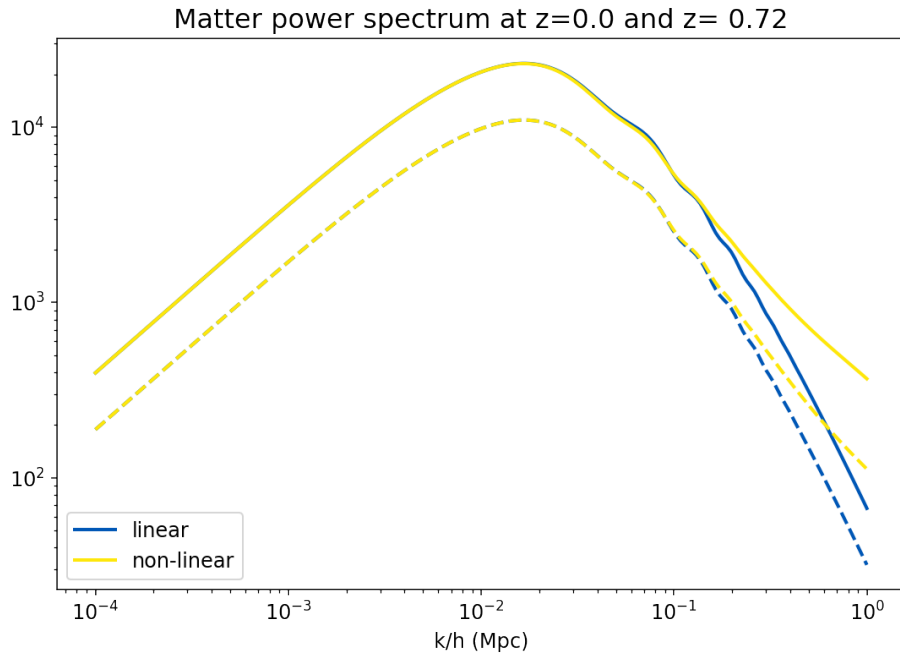


Figure 5: The linear (blue) and non-linear (yellow) matter power spectrum for CAMB’s fiducial cosmology at redshifts $z = 0$ (solid) and $z = 0.72$ (dashed).

One can see in Figure 5 that the linear and non-linear spectra are essentially identical until we get to small scales ($k \gtrsim 0.1 h^{-1} \text{ Mpc}$). It is also possible to see the imprints of the BAOs as wiggles in the matter power spectrum. These oscillations provide strong con-

straints on both the angular diameter distance $d_A(z)$ and the Hubble parameter $H(z)$ with completely linear physics, unlike SNIa which involve highly non-linear stellar explosions.

Not only it facilitates the comparison between theory and data, the power spectrum is also useful as a form of data compression, since it reduces the data set from millions of positions of the tracers down to just a few hundred band powers. But how do we actually go from a given data set consisting of hundreds of millions of positions of galaxies to an actual power spectrum which we can compare with our theory?

In order to extract information from our observations, we use something called an estimator, which is a function of the data and is also a random variable itself. We want our estimator to be both precise and accurate: the first requirement means that it should give us a distribution around the ensemble average as narrow as possible, that is, with minimum variance. The second one means that this average of our measurements must be as close as possible to the true value of the quantity we are trying to estimate. Ideally, we want this average to be exactly equal to the true value, and if that is the case, the estimator is said to be unbiased.

The most widely known estimators for the Fourier power spectrum are perhaps the FKP (Feldman et al., 1993) and PVP (Percival et al., 2004) estimators which are, in fact, both unbiased and have minimal variance. The first one, however, assumes that all types of galaxies trace the underlying density field in the same way as dark matter does, and we have known for a long time that this is not the case. While the second one takes into account the fact that different tracers will map the matter distribution in a distinct manner, one must first fix the biases and RSD parameters. A generalization of these estimators was obtained in Abramo et al. (2016), using the full Fisher matrix for the Fourier power spectrum for surveys of multiple tracers, which was previously derived from counts of galaxies in Abramo (2012). This optimal estimator can also simultaneously measure the scale dependent bias, RSD and non-gaussianities.

Since redshifts can be measured with exquisite precision by galaxy surveys, the standard approach has been to use the Fourier power spectrum. Despite certain advantages related to the radial information, $P(k)$ is not directly observable, so it is necessary to make a number of assumptions, such as choosing a cosmological model to translate from redshift to real space. This is usually done recursively: one first chooses a set of fiducial parameters, computes the power spectrum assuming they correctly describe the background cosmology, then estimates a new set of cosmological parameters to be subsequently used as the new fiducial cosmology, and the process is repeated until convergence is reached. This can introduce a bias into our results, and estimating the errors is not straightforward because not only $P(k)$, but also its argument k depend on cosmological parameters.

Another problem is that the Fourier power spectrum approach doesn't include lensing, and it also assumes that the sky is flat, which isn't realistic for the next generation of cosmological surveys because they will probe deeper redshifts and have a much larger sky coverage. This assumption also implies that the survey volume is a large 3D Cartesian box at a large distance from the observer, but since the speed of light is finite, we don't observe the Universe as a snapshot. Instead, what we see are successive 2D surfaces of simultaneity over our past lightcone. Thus, we don't have access to a 3D map of the universe in a given instant, and we can't observe the time evolution over a cosmologically relevant time (> 100 Myr) of a specific region in the sky either. Therefore, physical observables are more naturally expressed in spherical coordinates $\{z, \theta, \varphi\}$, in the lightcone, as opposed to our theory, which is best expressed in Cartesian coordinates $\{t, x, y, z\}$ due to our basic hypothesis of large-scale homogeneity and isotropy.

Works such as Yamamoto et al. (2006) and Scoccimarro (2015) have tried to take into account the curvature of the sky and they are able to deal with this problem up to a point, but these approximations start to fail at intermediate scales. Therefore, this method is quickly becoming unreliable to study the large scale structure through the clustering of galaxies, and we need an alternative.

3.2 The angular power spectrum

Our observations of the universe consist basically of spherical shells of constant redshift, hence we can use a decomposition in spherical harmonics $Y_{\ell m}$, since they constitute a natural orthonormal basis on the sphere. Let's consider again the density contrast of galaxies, which can be written as:

$$\delta_g(\theta, \varphi) = \sum_{l,m} a_{\ell m} Y_{\ell m}(\theta, \varphi) \quad (3.9)$$

with the harmonic coefficients given by:

$$a_{\ell m} = \int d\Omega Y_{\ell m}^*(\theta, \varphi) \delta_g(\theta, \varphi). \quad (3.10)$$

Alternatively, $\delta_g(\theta, \varphi)$ can be seen as the galaxy density $\delta(\vec{r})$, Eq. (3.3), integrated

along the line of sight, such that:

$$\begin{aligned}
a_{\ell m} &= \int d\Omega Y_{\ell m}(\theta, \varphi) \int dz \delta_g(\vec{r}(z)) n(z) \\
&= \int d\Omega Y_{\ell m}(\theta, \varphi) \int dz n(z) \int \frac{d^3 k}{(2\pi)^3} b(z) \hat{\delta}_g(\vec{k}) 4\pi \sum_{\ell' m'} i^{\ell'} j_{\ell'}(kr) Y_{\ell' m'}(\hat{k}) Y_{\ell' m'}^*(\hat{r}) \\
&= \frac{4\pi}{(2\pi)^3} \int dz b(z) n(z) \int d^3 k D(z) \hat{\delta}(\vec{k}, 0) \sum_{\ell' m'} i^{\ell'} j_{\ell'}(kr) Y_{\ell' m'}(\hat{k}) \delta_{\ell m} \delta_{\ell' m'} \\
&= \frac{4\pi}{(2\pi)^3} \int dz b(z) n(z) D(z) \int d^3 k \delta(\vec{k}, 0) i^{\ell} j_{\ell}(kr) Y_{\ell m}(\hat{k})
\end{aligned} \tag{3.11}$$

where $n(z)$ is the normalized selection function, and j_{ℓ} is the ℓ -th order spherical Bessel function. In the second line, the plane wave expansion and the spherical harmonic addition theorem were used to rewrite the exponential:

$$e^{i\vec{k}\vec{r}} = 4\pi \sum_{\ell m} i^{\ell} j_{\ell}(kr) Y_{\ell m}(\hat{k}) Y_{\ell m}^*(\hat{r}) \tag{3.12}$$

and the orthonormality of the spherical harmonics, $\int d\Omega_s Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}) = \delta_{\ell\ell'} \delta_{mm'}$

The angular power spectrum is then defined as the variance of these harmonic coefficients

$$\langle a_{\ell m}(r) a_{\ell' m'}^*(r') \rangle = \delta_{\ell\ell'} \delta_{mm'} C_{\ell}(r, r') \tag{3.13}$$

Being a two-point function, the angular spectrum represents the lowest order deviation from isotropy. This is an actual observable, as opposed to the Fourier power spectrum, and has the advantage of being fully model independent.

The dependencies on z of Eq. (3.10) can also be collected in a window function,

$$W(kr) = \int dz b(z) n(z) D(z) j_{\ell}(kr(z)) \tag{3.14}$$

such that a simple expression for the angular power spectrum can be found:

$$C_\ell^{ij}(r, r') = \frac{2}{\pi} \int_0^\infty dk k^2 W^i(kr) W^j(kr') P(k, 0) \quad (3.15)$$

where the superscripts i, j were introduced to denote different redshifts bins because the C_ℓ 's at different z are not independent, and their correlations actually contain very important information about the Baryon Acoustic Oscillations (BAOs). Therefore, the equation above defines both auto ($i = j$) and cross-correlations ($i \neq j$). This allows us to recover three dimensional information as well as compare data with theory. Note that this expression only includes the main contribution of the density perturbations, but unlike in the case of the Fourier spectrum, effects such as lensing can be naturally included in the window function as an additional term:

$$W(kr) = \int dz b(z) n(z) D(z) j_\ell(kr(z)) + 2[\mathcal{Q}(z) - 1] \ell(\ell + 1) \kappa(z) \quad (3.16)$$

where $\kappa(z)$ is the weak lensing convergence.

3.2.1 Angular spectra estimation

Here I will briefly discuss some basic measurement techniques for the angular power spectrum, and the issues that come with it. First, note that our theories describe most quantities of interest as continuous functions of position, but in practice, our observations are intrinsically discrete and we can only assign some value (e.g. galaxy number counts) at a finite number of points on the sphere. Hence, a key intermediate step in the data analysis process is to convert the raw data obtained from our detectors to a pixelized sky map. It is important to use a good pixelization scheme that is as smooth and regular as possible, and which also exploits symmetries of the spherical harmonics to simplify the subsequent analysis.

Many different schemes were proposed over the past decades, but today the widely used methodology for discretization and fast numerical analysis on a sphere is HEALPix, which was originally developed for processing CMB data by Gorski et al. (1999), but can generally be applied for any function on the sphere.

I consider here only the auto-correlations $r = r'$ for simplicity, and denote $C_\ell(r, r) \equiv C_\ell$. Noting that we can only measure $(2\ell + 1)$ m -modes for a given multipole ℓ , the

simplest estimator for the angular power spectrum is

$$\hat{C}_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 \quad (3.17)$$

For an idealized full sky coverage, it is unbiased, that is:

$$\langle \hat{C}_\ell \rangle = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} \langle a_{\ell m} a_{\ell m}^* \rangle \rightarrow C_\ell \quad (3.18)$$

and in this case the harmonic modes are independent, so the covariance matrices are diagonal on linear scales, unlike in real space, where the angular correlation function on different scales are correlated. However, this is never the case since we will always have at the very least a region of the sky masked by the Milky Way, and for a realistic, partial sky coverage the spherical harmonics are no longer orthogonal, so different ℓ modes are coupled and contribute to the true power spectrum. Nevertheless, we can deal with this by approximating the integral over the sky by a discrete sum over the pixels in the survey region, such that the harmonics coefficients of Eq. (3.10) are now given by

$$\begin{aligned} \tilde{a}_{\ell m} &= \int d\Omega Y_{\ell m}^*(\theta, \varphi) \delta_g(\theta, \varphi) \\ &\approx \sum_p \Omega_p Y_{\ell m}^*(\theta_p, \varphi_p) \delta_g(\theta_p, \varphi_p) \end{aligned} \quad (3.19)$$

where (θ_p, φ_p) are the coordinates of the pixel p and Ω_p is the surface area of that pixel. A better estimator for the angular spectra is then the *pseudo- C_ℓ* , or simply PC_ℓ , defined as (see Hivon et al., 2001):

$$\tilde{C}_\ell = \frac{1}{1 + 2\ell} \sum_{m=-\ell}^{\ell} |\tilde{a}_{\ell m}|^2 \quad (3.20)$$

This is a nearly optimal estimator and very fast to compute from a pixelated map. The original algorithm has been extended and implemented in public software packages, such as NAMASTER (Alonso et al., 2019), which provides a framework capable of computing the cross-spectra for galaxy clustering, lensing or shear, and to deal with any number of linear systematics. Note that the pseudo- C_ℓ gives us a biased estimate of the angular spectra, but the full-sky C_ℓ can then be recovered with a coupling matrix M resulting

from the cut sky:

$$\langle \tilde{C}_\ell \rangle = \sum_{\ell'} M_{\ell\ell'} \langle C_{\ell'} \rangle \quad (3.21)$$

This coupling matrix M depends only on the geometry of the survey mask, and is computed by comparing the spectrum of a lognormal mock simulation of the full sky to the spectrum of the survey mask. This can easily be done numerically in terms of Wigner 3j symbols with NAMASTER.

One then needs to invert the coupling matrix to recover the true C_ℓ , however this can't generally be done directly due to the intrinsic loss of information when masking the sky. A possible solution is to bin the coupled PC_ℓ into bandpowers q , defined as a set of N_q multipoles $\ell^q \equiv (\ell_q^1, \dots, \ell_q^{N_q})$ and a set of weights for these multipoles normalized such that $\sum_\ell w_q^\ell = 1$. The binned angular power spectrum is then given by:

$$\tilde{C}_q = \sum_\ell w_q^\ell \tilde{C}_\ell = \sum_{q'} \mathcal{M}_{qq'} C_{q'} \quad (3.22)$$

where it should be understood that the sum is implicitly performed over the multipoles in the bins $\ell \in q$, and the binned coupling matrix is

$$\mathcal{M}_{qq'} = \sum_\ell \sum_{\ell'} w_q^\ell M_{\ell\ell'} \quad (3.23)$$

which is much more stable and easier to invert than the original unbinned matrix. In order to finally compare the estimated angular power spectrum with our predictions, the theoretical spectrum must also be binned, following the same steps outlined for the estimator above:

- i) couple the multipoles;
- ii) bin the spectra into bandpowers;
- iii) decouple the bandpowers.

These steps are encoded into a filter matrix:

$$\mathcal{F}_{q\ell} = \sum_{q'} \mathcal{M}_{qq'}^{-1} \sum_{\ell'} w_{q'}^\ell M_{\ell'\ell} \quad (3.24)$$

such that the binned theoretical angular power spectrum is given by:

$$C_q^{th} = \sum_{\ell} \mathcal{F}_{q\ell} C_{\ell}^{th} \quad (3.25)$$

Note that the PC_{ℓ} does not assume Gaussianity, so it gives us unbiased estimates of the angular power spectrum for realistic, non-Gaussian fields as well.

3.2.2 Comparing the harmonic and Fourier approaches

There are some difficulties in applying angular power spectrum estimators to spectroscopic data. First, even with a fine binning strategy, it is not straightforward to guarantee that all of the radial information is contained in the projected redshift bins. And more importantly, these surveys require targeting of specific galaxies with fiber spectrographs and longer integration times are needed, resulting in a much lower galaxy density and, therefore, a lower signal to noise ratio (SNR). Nevertheless, it has been successfully done in Loureiro et al. (2019) for the BOSS DR12 catalog, where the authors obtained cosmological constraints comparable to the results obtained by the BOSS collaboration with a standard analysis, while not having to assume a fiducial cosmological model.

For photometric surveys, on the other hand, a tomographic analysis is naturally performed since they lack the precision on the radial information, though they are still competitive due to having a much higher number of objects. Besides, it is possible to recover the 3D clustering information from the 2D tomography using angular auto and cross correlations between different redshift bins (e.g. Asorey et al., 2012), and both approaches were shown to be equivalent. The harmonic approach also does not suffer from the flat-sky limitations and can naturally include wide-angle effects, which will be relevant for the upcoming surveys.

Of course, the harmonic approach has its own drawbacks. In principle, to compare with our theoretical predictions, we must include all cross-correlation bins. But since the spherical Bessel integrals are slowly converging due to their highly oscillatory nature, even with some form of data compression, such as the Karhunen-Loeve method described in Tegmark et al. (1998), this is computationally impractical for our future data sets.

It has been common practice to use the Limber approximation, which makes those integrals much easier to compute by replacing the spherical Bessel functions with a Dirac

delta:

$$j_\ell(kx) \rightarrow \sqrt{\frac{\pi}{2\ell+1}} \delta_D(\ell + 1/2 - kx) \quad (3.26)$$

This approximation is good if the rest of the integrand varies slowly compared to the Bessel function, and this is typically true for large ℓ (for an in depth discussion, see LoVerde and Afshordi, 2008). Although the effects on parameter estimation are negligible with current data, Fang et al. (2020b) argue that for future data sets like the upcoming LSST, the Limber approximation can bias our results significantly. To deal with this problem, they propose an extension to the FFTLog method to efficiently compute the angular power spectra that is also able to deal with integrals involving derivatives of the spherical Bessel function, which show up in RSD and other high-order corrections. The authors also point out that one might get away with using the Limber approximation for galaxy-galaxy lensing, but not for galaxy clustering.

Other numerical methods for the harmonic approach such as ANGPOW (Campagne et al., 2017), 2-FAST (Gebhardt and Jeong, 2017), SuperFaB (Gebhardt and Doré, 2021) have been proposed to speed up those integrals, but here we review a simple pipeline that helps alleviate this issue without sacrificing precision.

3.2.3 A Hybrid approach for the angular power spectrum

An important step in the process of extracting information from our surveys is the proper binning of the data. One can in principle even take varying bin sizes to better sample scales with a higher signal-to-noise ratio. In the Fourier power spectrum approach, we must use redshift bins thick enough (e.g. $\Delta z \approx 0.1$) to contain all the k modes we are interested in. All quantities are assumed to be constant in that bin, and the correlation between different bins are ignored, effectively giving us diagonal covariance matrices by construction. But for the angular power spectrum C_ℓ , we lose significant information by taking a redshift slice as large as Δz , which can suppress the signal along the line of sight, so we must use much smaller bin sizes. However, they must also be larger than the error in the redshift measurements, and in this case we must include the cross-correlations between the different bins, and this can quickly become computationally unfeasible for large data sets.

A hybrid method proposed by Camera et al. (2018) consists in combining the two methods: first, the data is divided in thick $\Delta z = 0.1$ bins, as in the Fourier approach. Then, they are subdivided into 10 thin top-hat bins of $\delta z = 0.01$, which are convolved

with a Gaussian of width $\sigma_z = 0.001$, making the bins overlap slightly to account for non-negligible errors in redshift estimation. Each thick bin is considered a separate survey to which we apply the tomographic approach, considering all the cross correlations between the thin bins. Then, all the Fisher matrices obtained from each thick bin are summed.

This hybrid approach is far less computationally expensive than the others: for $0.6 \leq z \leq 2.0$, it requires the computation of only 770 spectra, as opposed to the 9870 needed with the regular tomographic approach. This method works because most of the information comes from correlations at small separations, since the correlation coefficient between different bins quickly falls off as a function of Δz . Thus, treating every thick bin as a separate survey and neglecting the cross correlations between distant bins is a reasonable approximation. Although it neglects lensing, it was also shown that this binning strategy has a constraining power comparable to the standard $P(k)$ analysis, and presents a significant information gain over the tomographic approach. The hybrid method, therefore, optimizes the angular power spectrum computations for spectroscopic surveys.

3.3 Sources of statistical errors

In this subsection, I discuss the two main statistical issues we encounter when trying to measure the power spectrum or the correlation function of a given galaxy sample: shot noise and cosmic variance.

Cosmological surveys can only probe a fraction of the density field of our universe, and they do so imperfectly. Although it would be ideal, our telescopes can't detect all the objects in a given region, only those brighter than some lower intensity threshold. Thus, a survey is characterized by its selection function, $\bar{n}(\vec{r})$, which is the expected mean number of galaxies at position \vec{r} given the selection criteria of the survey, such as the flux limit. Now remember that the Fourier transform of the density contrast, Eq. (3.1), can be written as:

$$\delta(\vec{k}) = \frac{1}{V} \int d^3r \delta(\vec{r}) e^{i\vec{k}\cdot\vec{r}} \quad (3.27)$$

where V is a volume sufficiently large to represent the real underlying density field. By definition, the power spectrum is simply

$$P(k) \equiv \langle |\delta(\vec{k})|^2 \rangle \quad (3.28)$$

Since we don't observe a smooth field, the finite number of observed galaxies are assumed to be drawn from a Poisson sampling of the true density field, then the selection function can be interpreted as the probability of including a galaxy at position \vec{r} in the survey, in units of number of galaxies per unit volume. In this Poisson process, the number density distribution of objects is given by:

$$n(\vec{r}) = \sum_i \delta_D(\vec{r} - \vec{r}_i) \quad (3.29)$$

where \vec{r}_i is the position a galaxy i . The continuous densities ρ are then replaced by this discrete number density and the contrast becomes:

$$\delta_d(\vec{r}) = \frac{n(\vec{r}) - \bar{n}(\vec{r})}{\bar{n}(\vec{r})} \quad (3.30)$$

Now we split the volume V into infinitesimal volume elements dV_i where there are either 0 or 1 galaxies ($n_i = 0, 1$), such that $\langle n_i \rangle = \langle n_i^2 \rangle$, and we take the Fourier transform of the density contrast with this discrete distribution:

$$\delta_d(\vec{k}) = \frac{1}{\bar{n}} \sum_i n_i e^{i\vec{k} \cdot \vec{r}_i} - \delta_{\vec{k},0}^K \quad (3.31)$$

where δ^K is the Kronecker delta and the subscript d represents the discretized contrast. Now we can compute the ensemble average

$$\begin{aligned} \langle \delta_d(\vec{k}) \delta_d^*(\vec{k}') \rangle = & \left\langle \frac{1}{\bar{n}^2} \sum_{ij} n_i n_j e^{i(\vec{k} \cdot \vec{r}_i - \vec{k}' \cdot \vec{r}_j)} - \frac{1}{\bar{n}} \sum_i n_i e^{i\vec{k} \cdot \vec{r}_i} \delta_{\vec{k}',0}^K \right. \\ & \left. - \frac{1}{\bar{n}} \sum_j n_j e^{i\vec{k}' \cdot \vec{r}_j} \delta_{\vec{k},0}^K + \delta_{\vec{k},0}^K \delta_{\vec{k}',0}^K \right\rangle \end{aligned} \quad (3.32)$$

Remember that for a Poisson distribution:

$$\langle n_i \rangle = \bar{n} dV_i \quad (3.33)$$

$$\langle n_i n_j \rangle_{i \neq j} = \bar{n}^2 dV_i dV_j [1 + \delta(\vec{r}_i), \delta(\vec{r}_j)], \quad (3.34)$$

then after a bit of algebra, we find that the power spectrum we measure for a given dis-

tribution of galaxies will have a noise contribution, given by the inverse of their number density:

$$\langle \delta_d(\vec{k}) \delta_d^*(\vec{k}') \rangle = P(k) + \frac{\delta_{\vec{k}\vec{k}'}}{\bar{n}}. \quad (3.35)$$

Intuitively, this makes sense because we should be able to extract more accurate information if we can observe a higher density of galaxies: in the limit $\bar{n} \rightarrow \infty$, the shot noise is null and we perfectly recover the correlation function. In principle, instrumental noise should also be added here, but here we assume it is much smaller than the shot noise so that it can be neglected.

This term also appears for the angular power spectrum and, in general the expectation value of the observed overdensities is a sum of the true correlation function and this diagonal *shot noise* term:

$$\langle \delta(r) \delta^*(r') \rangle = C(r, r') = \xi(r, r') + \frac{\delta_D(r - r')}{\bar{n}(r)} \quad (3.36)$$

Now, remember that the scales $k \sim 1/\lambda$ we can study with a given survey depend on its volume. Therefore, some modes are beyond reach because we can only map a finite volume, and the largest modes we can probe are very poorly determined because there is only one realization of the matter density field available, that is, we only have one universe to observe. This limitation is known as *cosmic variance* and is the other main source of statistical uncertainty we have in power spectrum estimation, being more relevant at larger scales.

We can estimate how much this affects our error bars by considering the ideal scenario of a full sky survey, in which case Eq. (3.17) is unbiased and optimal. If we further assume that the harmonic density contrast $a_{\ell m}$ on the sphere are Gaussian variables, we can use Wick's theorem to compute the estimator's variance:

$$\begin{aligned} Var(\hat{C}_\ell, \hat{C}_{\ell'}) &= \langle \hat{C}_\ell \hat{C}_\ell \rangle - \langle \hat{C}_\ell \rangle \langle \hat{C}_\ell \rangle = \frac{1}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} \langle a_{\ell m} a_{\ell m}^* a_{\ell' m'} a_{\ell' m'}^* \rangle - C_\ell C_{\ell'} \\ &= \frac{1}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} \left(C_\ell C_{\ell'} + (-1)^{-(m+m')} \delta_{\ell\ell'} \delta_{mm'} C_\ell^2 + \delta_{\ell\ell'} \delta_{mm'} C_{\ell'}^2 \right) \\ &\quad - C_\ell C_{\ell'} \end{aligned} \quad (3.37)$$

$$= \frac{1}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} 2C_\ell^2 \delta_{\ell\ell'} \delta_{mm'} = \frac{2C_\ell^2}{2\ell + 1} \delta_{\ell\ell'} = \sigma_c^2$$

While shot noise can, in principle, be reduced by increasing the number density of the observed tracers, cosmic variance is limited by the volume of the survey, and that's a problem especially on very large scales since we only have one universe to observe at any particular time. This in principle represents a fundamental statistical limit for our measurements about the large scale structure of the universe.

However, with a multi tracer approach, it is possible to overcome this limitation (see e.g. Seljak, 2009; Abramo and Leonard, 2013) and measure some bias sensitive parameters with a precision unconstrained by the cosmic variance σ_c , as long as the shot noise of all concerning tracers is sufficiently small. This is done by combining two or more tracers of the same underlying density field with different biases and comparing their relative clustering. This increase in constraining power is a direct consequence of the multi tracer Fisher matrix, whose eigenvalues are the relative clustering amplitudes between these tracers (see Abramo and Leonard, 2013).

4 Redshift space distortions

In this section I briefly review an important effect that affects our measurements: linear redshift space distortions (RSD), first derived in Kaiser (1987). This is an important effect caused by peculiar velocities of galaxies and results in the discrepancy between their positions in real and redshift space, which complicates the interpretation of galaxy clustering and must be taken into account. On the other hand, it also offers a new independent way of measuring the growth rate of structure formation, which can be used for constraining the density parameters of the universe, and also as a test of general relativity on large scales. A more complete review on the effects of linear RSD can be found in Hamilton (1998), or Percival et al. (2011) for a more recent but less thorough one. A description of non-linear RSD effects can be found in e.g. Taruya et al. (2010).

4.1 Peculiar velocities and linear RSD

Although this can change in the following years, or decades, with the refinement of our gravitational waves detection network, or even some other innovative method, currently the best way we have to estimate cosmological distances is the recessional velocity of galaxies, to which we can attribute a distance through the Hubble–Lemaître law:

$$v(z) = H(z)r(z) \quad (4.1)$$

For relatively close galaxies, the Hubble parameter H can be taken as approximately constant such that the expression above can be written as:

$$v \approx H_0 r \quad (4.2)$$

where H_0 is the Hubble constant today. So up to a few hundred megaparsecs from us, the recessional velocity v of a galaxy is linearly proportional to its real distance r .

In practice, our measurements are done by identifying emission or absorption lines of particular elements in the spectra of the observed objects, and comparing to the known wavelengths of the same lines for a source at rest. Then the redshift can be found through

the relation:

$$z = \frac{\lambda_o - \lambda_e}{\lambda_e} = \frac{\Delta\lambda}{\lambda_e} = \sqrt{\frac{1 + v/c}{1 - v/c}} - 1 \approx \frac{v}{c} \quad (4.3)$$

where λ_o and λ_e are the wavelengths of the observed and emitted light, respectively, and the approximation is valid for redshifts $z \ll 1$. Therefore, we can use Eq. (4.2) to infer distances on scales for which the traditional distance ladder doesn't work. Note that this relies on the assumption that atomic physics does not change over cosmological time scales, such that these spectra were actually the same billions of years ago.

Since it is possible to measure redshifts with an incredibly good accuracy, unlike distances, large cosmological surveys were devised to map the position of galaxies in redshift space. However, our measurements are susceptible to deviations from the Hubble flux, caused by inhomogeneities in the gravitational field around the galaxy, which can give rise to the so called peculiar velocities. In other words, if the universe was not expanding, galaxies would still have some velocity, generally in the direction of a region of high density around it, which would shift their measured spectra. The observed velocity of a given galaxy is then given by a combination of Hubble flow (which is proportional to their physical distance) and their peculiar velocity, $\vec{v}_p = \vec{v} \cdot \hat{r}$:

$$\vec{v}_o = H(z)\vec{r} + \vec{v}_p \quad (4.4)$$

Therefore, the distances \vec{s} we actually observe in a redshift map are distorted by a factor proportional to their peculiar velocity, and to good approximation only in the radial direction. In comoving coordinates, we have:

$$\vec{s} = \vec{r} + \frac{v_p(\vec{r})}{a(z)H(z)}\hat{r} \quad (4.5)$$

where $a(z) = (1 + z)^{-1}$ is the scale factor and $H(z)$ is the Hubble parameter. Note that both the position in real space, $\vec{r} = \vec{r}(z)$, and in redshift space, $\vec{s} = \vec{s}(z)$, are functions of the redshift z , but henceforth we shall generally suppress this dependence from our notation for simplicity. Other factors, such as weak lensing and the integrated Sachs-Wolfe effect, also affect the observed redshift of galaxies, but here we consider only distortions caused by peculiar velocities.

This effect is quite intuitive: the expansion of the universe makes galaxies move away from us at a rate proportional to their distance. Hence, if their peculiar motion makes

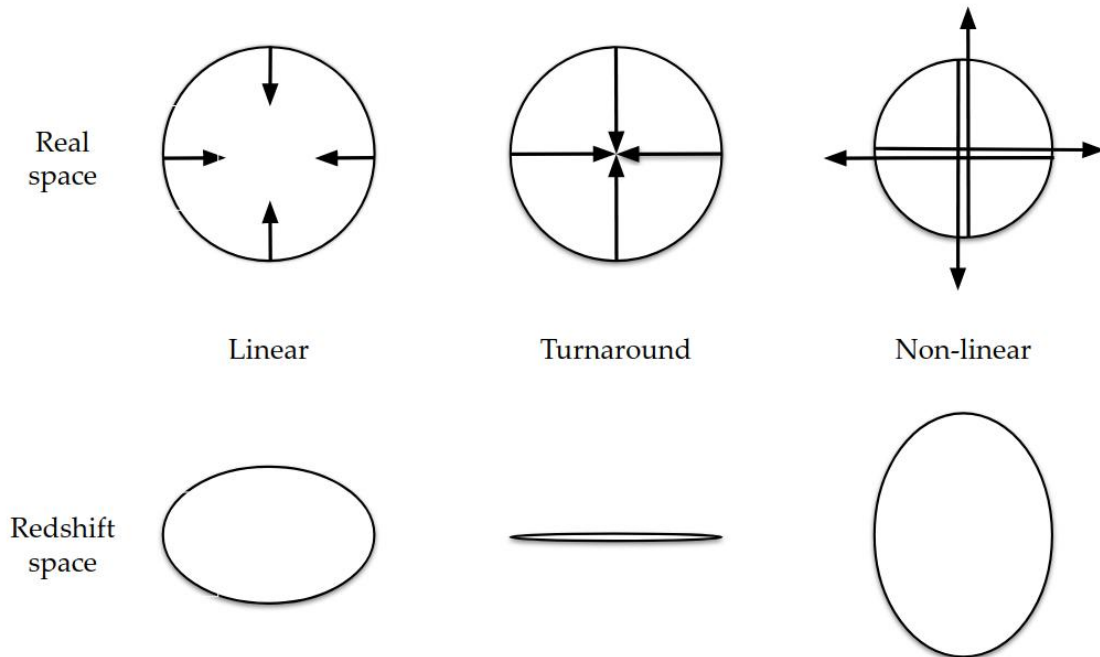


Figure 6: A simple schematic showing how peculiar velocities (represented by the arrows) affect our observations of a spherical distribution of matter.

them recede at a even faster rate, they will look farther away than they actually are. By contrast, if their peculiar velocity is anti-parallel to the line of sight, they will seem closer to us. To see how this affects a group or cluster of galaxies, let's consider a spherical shell of galaxies falling towards its center.

A schematic is presented in Figure 6 to illustrate how peculiar velocities distort overdensities along the line of sight. At large scales, the peculiar velocity of galaxies in a spherical shell falling towards a central region of overdensity is small compared to its radius, so what we observe is a squashed shell. Hence, the clustering is artificially amplified along the line of sight and this is known as the Kaiser effect. On the other hand, when we go to smaller scales, the infall peculiar velocity tends to get relatively larger. At the turnaround point, when the peculiar velocity just cancels out the general Hubble expansion, we would observe a shell collapsed to a single velocity in redshift space. And at even smaller, non-linear scales, the relatively large velocities make the galaxy distribution appear elongated along the line of sight, giving rise to the so called fingers-of-god.

This is a significant effect that must be taken into account for any observable that involves radial distance, such as galaxy number counts. For small areas in the sky, the Fourier transform of the density contrast in redshift space is related to the real-space

contrast by

$$\delta^s(\vec{k}) = (1 + \beta\mu^2)\delta(\vec{k}) \quad (4.6)$$

where the superscript s denotes a quantity in redshift space, $\mu = \hat{k} \cdot \hat{z}$ is the cosine of the angle between \vec{k} and the line of sight \hat{z} , and β is the linear redshift distortion parameter, which, in a standard pressureless FLRW cosmology with bias b , can be related to the present day cosmological density Ω_m of the universe by

$$\beta = \frac{f(\Omega_m)}{b} \approx \frac{\Omega_m^{0.54}}{b} \quad (4.7)$$

where $f = d \ln \delta / d \ln a$ is the dimensionless linear growth rate. Note that β is not really a true cosmological parameter such as, say, Ω_0 , because it only exists within the context of linear biasing, which is most likely wrong and just a good approximation. Nevertheless, this distortion parameter can be measured through the amplitude of distortions on large scales, and was originally proposed to provide an independent way to determine the density parameter Ω_m itself, though the particularly simple approximation in (4.7) requires one to assume that Λ CDM, and thus general relativity, are correct. As a matter of fact, it is possible to measure this growth index γ in $f(\Omega_m) = \Omega_m^\gamma$, and nowadays the distortion parameter is also a powerful tool to test models of modified gravity (see e.g. Hernández-Aguayo et al., 2019, and references therein).

Equation (4.6) was first derived in Kaiser (1987) and is known as the Kaiser formula, but this approximation is not suitable for the next generation of surveys and we can do better. We start from the continuity equation, which can be used to relate the density and velocity fields:

$$\dot{\rho} = -\vec{\nabla} \cdot (\rho \vec{v}) \quad (4.8)$$

where the dot denotes a time derivative and $\vec{\nabla} \cdot$ is the divergence. In linear perturbation theory, we can write the density as $\rho = \rho_0[1 + \delta(\vec{r})]$, with ρ_0 constant. Performing a change of variables $d/dt = aH(d/da)$, we have:

$$-aH \frac{d\delta(\vec{r})}{da} = \vec{\nabla} \cdot \vec{v} = -H\delta(\vec{r}) \frac{d \ln \delta(\vec{r})}{d \ln a} = -H(z)f(z)\delta(\vec{r}) \quad (4.9)$$

In Fourier space, notorious for simplifying vectorial computations, the divergence of the velocity field takes a much simpler form, so we take a quick detour and Fourier transform

the equation above to write the peculiar velocity as:

$$v_p(\vec{k}) = \vec{v}(\vec{k}) \cdot \hat{r} = if(z)H(z)\delta(\vec{k}, z)\frac{(\vec{k} \cdot \hat{r})}{k^2} \quad (4.10)$$

Note that $i(\vec{k} \cdot \hat{r})/k^2 \equiv -\frac{d}{dr}\vec{\nabla}^{-2}$, so we can go back to real space, where the peculiar velocity is now given by (see e.g. Hamilton and Culhane, 1995):

$$v_p(\vec{r}) = -H(z)f(z)\frac{d}{dr}\vec{\nabla}^{-2}\delta(\vec{r}) \quad (4.11)$$

where $\vec{\nabla}^{-2}$ is the inverse Laplacian. This comes from the fact that, in linear perturbation theory, an overdensity creates a peculiar velocity field (see e.g. Peebles, 2020), which can be derived from the a potential function whose Laplacian is the overdensity itself.

Now, to find the relationship between the real and redshift space overdensities, remember that the mass must be conserved, such that:

$$\rho_s d^3 s = \rho_r d^3 r \Rightarrow 1 + \delta^s(\vec{s}) = \frac{d^3 r}{d^3 s} [1 + \delta(\vec{r})] \quad (4.12)$$

where the Jacobian can be computed from Eq. (4.5):

$$\frac{d^3 r}{d^3 s} \equiv J = \left[\left(1 + \frac{1}{H(z)} \frac{d}{dr} v_p(\vec{r}) \right) \left(1 + \frac{v_p(\vec{r})}{H(z)r} \right)^2 \right]^{-1} \approx \left(1 + \frac{1}{H(z)} \frac{d}{dr} v_p(\vec{r}) \right)^{-1} \quad (4.13)$$

where the approximation follows from the fact that the correction from the derivative of the peculiar velocity, is larger than the other correction by a factor of kr : for a plane wave perturbation, it is of order kv/H_0 , while the second one is of order $v/H_0 r$. Kaiser argued that r is of the order of the size of the survey, while k is of the order of the Fourier modes we try to measure. But there are very few modes with wavelength that large within the survey volume and they are very poorly determined, hence we are only generally interested in modes with $kr \gg 1$. Therefore, in linear approximation, we have:

$$1 + \delta^s(\vec{s}) = \left(1 - \frac{1}{H(z)} \frac{d}{dr} v_p(\vec{r}) \right) [1 + \delta(\vec{r})] \quad (4.14)$$

Now remember the expression we found for the peculiar velocity, Eq. (4.11), then the relation between the density contrast in real and redshift space, without the small angle

approximation, becomes

$$\delta^s(\vec{s}) = \left(1 + f(z) \frac{d^2}{dr^2} \vec{\nabla}^{-2} \right) \delta(\vec{r}) \quad (4.15)$$

4.2 The angular power spectrum in redshift space

We can now find an exact expression for the angular power spectrum in redshift space. From now on, I will drop the superscript in the notation of the overdensity, $\delta^s(\vec{s}) \rightarrow \delta(\vec{s})$, because it should be clear from their argument whether they refer to real or redshift space. First, we decompose the overdensities in spherical harmonics:

$$\delta(\vec{s}) = \sum_{l,m} a_{\ell m}(s) Y_{\ell m}(\hat{s}); \quad a_{\ell m}(s) = \int d\Omega_s Y_{\ell m}^*(\hat{s}) \delta(\vec{s}) \quad (4.16)$$

where we add a subscript on $d\Omega_s$ to make it clear that the angular integral is performed over the unprimed coordinates. The harmonic correlation in redshift space is then given by:

$$\begin{aligned} \langle a_{\ell m}(s) a_{\ell' m'}^*(s') \rangle &= \int d\Omega_s d\Omega_{s'} Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}') \langle \delta(\vec{s}) \delta^*(\vec{s}') \rangle \\ &= \int d\Omega_s d\Omega_{s'} Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}') \\ &\quad \times \left\langle \left[1 + f(z) \frac{d^2}{dr^2} \vec{\nabla}^{-2} \right] \delta(\vec{r}) \left[1 + f(z') \frac{d^2}{dr'^2} \vec{\nabla}^{-2} \right] \delta^*(\vec{r}') \right\rangle \end{aligned} \quad (4.17)$$

Now, remember that the real space overdensity is:

$$\delta(\vec{r}) = \int \frac{d^3 k}{(2\pi)^3} e^{-i\vec{k}\vec{r}} \hat{\delta}(\vec{k}) \quad (4.18)$$

so we can use the plane wave expansion and after some lengthy but straightforward alge-

bra (see Appendix A for an explicit derivation), we find that

$$\begin{aligned}
\langle \delta_{\ell m}(s) \delta_{\ell' m'}^*(s') \rangle &= \frac{2}{\pi} \int dk k^2 [j_\ell(kr) - f(z)j_\ell''(kr)] \\
&\quad \times [j_{\ell'}(kr') - f(z')j_{\ell'}''(kr')] P(k) \delta_{\ell\ell'} \delta_{mm'} \quad (4.19) \\
&\equiv C_\ell^s(r, r') \delta_{\ell\ell'} \delta_{mm'}
\end{aligned}$$

where the angular power spectrum in redshift space is

$$C_\ell^s(r, r') \equiv \frac{2}{\pi} \int dk k^2 [j_\ell(kr) - f(z)j_\ell''(kr)] [j_{\ell'}(kr') - f(z')j_{\ell'}''(kr')] P(k) \quad (4.20)$$

Note that this expression does not take lensing into account, though it can be naturally included in a window function. In this work, however, it is a complication we will not deal with.

5 Fisher matrix in real space

When estimating how accurately we can measure cosmological parameters and constrain our models, it is useful to use the Fisher information matrix, first proposed nearly 80 years ago by Fisher (1935) and introduced to the cosmology community by Tegmark et al. (1997).

Suppose we want to constrain a set of parameters θ , under the assumption of a certain model M given some data set \mathbf{x} . Before the measurements, we can think of \mathbf{x} as a random variable with some probability distribution $p(\mathbf{x}|\theta)$ that is a function of the parameters. The Fisher matrix offers us a simple way of measuring how much information the observable \mathbf{x} carries about a parameter θ . In the frequentist view, it is defined as the expectation value of the Hessian of log likelihood $\mathcal{L} = -\ln L$:

$$F_{\alpha\beta} = \left\langle \frac{\partial^2 \ln \mathcal{L}}{\partial \theta_\alpha \partial \theta_\beta} \right\rangle \quad (5.1)$$

while in the Bayesian approach the data is not seen as random variables, so the Fisher matrix is then evaluated at the maximum likelihood:

$$F_{\mu\nu} = \left. \frac{\partial^2 \mathcal{L}}{\partial \theta_\alpha \partial \theta_\beta} \right|_{\theta=\bar{\theta}} \quad (5.2)$$

These two definitions coincide for Gaussian data if the mean and variance depend on the parameters in a linear way. The Fisher matrix is the inverse of the covariance matrix, and indeed, the Cramér-Rao inequality tells us that it provides a lower bound on the error of a parameter θ_μ ,

$$\sigma(\theta_\alpha) \leq \sqrt{(F^{-1})_{\alpha\alpha}} \quad (5.3)$$

which means that this formalism allows us to find out the absolute minimum error bars we could possibly obtain for a parameter θ_α with a given unbiased estimator, regardless of the method used to analyze the data. This is useful when planning future experiments, since no data or even simulations are necessary. We can forecast how well multiple experiments will do, knowing only our model and measurement uncertainties, and compare their precision versus the cost to make a better informed decision.

In the case of a Gaussian distribution with mean $\boldsymbol{\mu}$, and covariance:

$$C = \langle (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T \rangle, \quad (5.4)$$

and defining the data matrix:

$$D = (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T, \quad (5.5)$$

we have:

$$\begin{aligned} 2\mathcal{L} &= \ln \det C + (\mathbf{x} - \boldsymbol{\mu})C^{-1}(\mathbf{x} - \boldsymbol{\mu})^T \\ 2\mathcal{L} &= \text{Tr}[\ln C + C^{-1}D], \end{aligned} \quad (5.6)$$

where we used the identity $\ln \det C = \text{Tr} \ln C$. Taking the partial derivative of the equation above gives:

$$2 \frac{\partial \mathcal{L}}{\partial \theta_\alpha} = \text{Tr} \left[C^{-1} \frac{\partial C}{\partial \theta_\alpha} - C^{-1} \frac{\partial C}{\partial \theta_\alpha} C^{-1} D + C^{-1} \frac{\partial D}{\partial \theta_\alpha} \right] \quad (5.7)$$

Now remember that $\langle \mathbf{x} \rangle = \boldsymbol{\mu}$ and $\langle \mathbf{x}\mathbf{x}^T \rangle = C + \boldsymbol{\mu}\boldsymbol{\mu}^T$, so

$$\begin{aligned} \langle D \rangle &= C \\ \left\langle \frac{\partial D}{\partial \theta_\alpha} \right\rangle &= 0 \\ \left\langle \frac{\partial D^2}{\partial \theta_\alpha \partial \theta_\beta} \right\rangle &= \frac{\partial \boldsymbol{\mu}}{\partial \theta_\alpha} \frac{\partial \boldsymbol{\mu}^T}{\partial \theta_\beta} + \frac{\partial \boldsymbol{\mu}}{\partial \theta_\alpha} \frac{\partial \boldsymbol{\mu}^T}{\partial \theta_\beta} \end{aligned} \quad (5.8)$$

With this we see that $\langle \partial \mathcal{L} / \partial \theta_\alpha \rangle = 0$. Using the chain rule on (5.7), we get:

$$\begin{aligned} 2 \frac{\partial \mathcal{L}^2}{\partial \theta_\alpha \partial \theta_\beta} &= \text{Tr} \left[C^{-1} \frac{\partial C^2}{\partial \theta_\alpha \partial \theta_\beta} - C^{-1} \frac{\partial C}{\partial \theta_\alpha} C^{-1} \frac{\partial C}{\partial \theta_\beta} \right. \\ &\quad \left. + C^{-1} \left(\frac{\partial C}{\partial \theta_\alpha} C^{-1} \frac{\partial C}{\partial \theta_\beta} + \frac{\partial C}{\partial \theta_\beta} C^{-1} \frac{\partial C}{\partial \theta_\alpha} \right) C^{-1} D \right] \end{aligned}$$

$$\begin{aligned}
& - C^{-1} \left(\frac{\partial C}{\partial \theta_\alpha} C^{-1} \frac{\partial D}{\partial \theta_\beta} + \frac{\partial C}{\partial \theta_\beta} C^{-1} \frac{\partial D}{\partial \theta_\alpha} \right) \\
& - C^{-1} \frac{\partial C^2}{\partial \theta_\alpha \partial \theta_\beta} C^{-1} D + C^{-1} \frac{\partial D^2}{\partial \theta_\alpha \partial \theta_\beta} \Big] \tag{5.9}
\end{aligned}$$

Therefore, for a Gaussian distribution with $\boldsymbol{\mu} = 0$, the Fisher matrix (5.1) is simply:

$$F_{\alpha\beta} = \frac{1}{2} \text{Tr} \left[C^{-1} \frac{\partial C}{\partial \theta_\alpha} C^{-1} \frac{\partial C}{\partial \theta_\beta} \right] \tag{5.10}$$

The beauty of this formalism is that we can combine multiple independent experiments, as well as any prior knowledge we have about the parameters, by adding their Fisher matrices. Then we simply invert the summed matrix to get the uncertainties of the joint analysis. The Fisher matrix can also in principle be computed analytically, making it a much faster alternative for error estimation than Monte Carlo simulations, which must explore the whole multidimensional parameter space.

In addition to being a very powerful prediction tool before any data is even gathered, the Fisher matrix can also be used to maximize the information we can extract from the actual data, by paving the path to obtaining optimized analytical maximum likelihood (ML) estimators.

5.1 Fisher matrix in the lightcone

We wish to compute the Fisher matrix for the key physical observable that we have on redshift slices: the angular power spectrum, $C_\ell(z, z')$. This corresponds to angular averaging two slices, one at redshift z and the other at redshift z' . Due to the azimuthal symmetry, the dependence on m is cancelled and only the total angular momentum number ℓ survives the angular integration. This is in contrast with the 3D picture in Cartesian coordinates, where we start with the density contrast $\delta^i(\vec{x})$ for a tracer μ , and after Fourier transforming and averaging over the modes, we end up with the power spectrum $P_i(\vec{k})$, or more generically, with spectra for any tracers i and j , $P^{ij}(\vec{k})$.

In Abramo et al. (2016) it was shown how to optimally combine different tracers over the same volume, in order to maximize the signal/noise of the observables $P^i(\vec{k})$. Since the optimal weights are simply a realization of the idea of maximizing signal (or minimizing noise), similarly to what is done with Wiener filters, the key intermediate step

towards obtaining those weights is an analytical computation of the Fisher matrix. This was done in, e.g., Abramo (2012); Abramo and Leonard (2013), where a simple analytical expression for the Fisher matrix for the observables $P^i(\vec{k})$ was obtained:

$$F^{ij} = \frac{\delta^{ij}}{P^i P^j} \int_{v_i} \frac{d^3x d^3k}{(2\pi)^3} \mathcal{F}^{ij}, \quad (5.11)$$

where the spatial integration is performed over the volume where the tracers have been mapped, and which are used for the computation of the power spectra. Inside the integral we have the Fisher information matrix per unit of phase space volume (Abramo and Leonard, 2013):

$$\mathcal{F}^{ij} = \frac{1}{4} \frac{\delta^{ij} \mathcal{P}^i \mathcal{P}^j (1 + \mathcal{P}) + \mathcal{P}^i \mathcal{P}^j (1 - \mathcal{P})}{(1 + \mathcal{P})^2} \quad (5.12)$$

where we defined the effective (adimensional) clustering strength in units of shot noise, $\mathcal{P}^i = \bar{n}^i(\vec{x}) P^i(\vec{k})$, and the total effective clustering strength of all the N_t tracers, $\mathcal{P} = \sum_{i=1}^{N_t} \mathcal{P}^i$. Shot noise is assumed to be Poissonian, that is $P_{shot}^i = 1/\bar{n}^i$, where $\bar{n}^i(\vec{x})$ is the number density of the tracer α inside the survey volume, In other words, $\mathcal{P}^i = P^i/P_{shot}^i$.

This means that equation (5.12) is in fact the Fisher information density per unit of phase space volume for $\log \mathcal{P}^i$ (Abramo, 2012):

$$F[\log \mathcal{P}^i(\vec{x}, \vec{k}), \log \mathcal{P}^j(\vec{x}', \vec{k}')] = (2\pi)^3 \delta_D(\vec{x} - \vec{x}') \delta_D(\vec{k} - \vec{k}') \mathcal{F}^{ij}(\vec{x}, \vec{k}) \quad (5.13)$$

In bins of finite volume we have:

$$F[\log \mathcal{P}^i, \log \mathcal{P}^j] = \delta^{ij} \mathcal{F}^{ij} \quad (5.14)$$

which means that separate volume carry effectively independent information. One can easily check that the multi-tracer Fisher matrix of equation (5.11) reduces to the FKP Fisher matrix when there is only one type of tracer.

Several approximations were used in order to arrive at these expressions. Amongst them, the assumption that the volumes were large and separate, which is what allow us to assume the statistical independence between them. In practice, one basically must map a volume v_i that is large is sufficiently spatially separated from another volume v_j , such that the measurements of the power spectra in each volume sample statistically independent realizations of the modes $\tilde{\delta}_i(\vec{k})$.

The main contribution of this work is the generalization of the approach used in Abramo (2012); Abramo and Leonard (2013); Abramo et al. (2016) in order to derive the Fisher matrix for the angular power spectra $C_\ell(z, z')$:

$$F_{\ell_1, \ell_2}^{ij}(z_1, z'_1; z_2, z'_2) = F[C_{\ell_1}^i(z_1, z'_1), C_{\ell_2}^j(z_2, z'_2)]. \quad (5.15)$$

We are particularly interested in the limit where the redshift slices are very thin and can be even adjacent, e.g.: $z'_i = z_i + \Delta z_i$, with $\Delta z_i \ll 1$, and $z_2 = z_1 + \Delta z_{12}$, with Δz_{12} also small, in which case it will not be necessarily true that $C_{\ell_1}^{i1}(z_1, z'_1)$ is still statistically independent of $C_{\ell_2}^{i2}(z_2, z'_2)$.

5.2 The harmonic Fisher matrix in real space

The remainder of this dissertation will focus on presenting the main results of this work (Abramo et al., 2022): we derive exact expressions for the covariance and Fisher matrices of the angular power spectrum for two redshift slices with radii \bar{x} and \bar{y} , $C_\ell^{ij}(\bar{x}, \bar{y})$, working under the assumption of linear biasing for the tracers.

First, let's derive the Fisher matrix from first principles in real space, that is, excluding redshift space distortions. The degrees of freedom available in a survey are the positions of the galaxies, or other tracers of the underlying matter distribution. The measurement of the number densities $n^i(\vec{x})$ of tracer species i , over some volume around the position \vec{x} , reflects some mean density of those tracers as well as a fluctuation $\delta n^i = n^i - \bar{n}^i$ around this mean. From these observables we can compute the main object that carries information about cosmology, the data covariance, defined as:

$$\Sigma^{ij}(\vec{x}, \vec{y}) = \langle \delta n^i(\vec{x}) \delta n^j(\vec{y}) \rangle = \bar{n}^i(\vec{x}) \bar{n}^j(\vec{y}) \xi^{ij}(\vec{x}, \vec{y}) + \bar{n}^i(\vec{x}) \delta^{ij} \delta_D(\vec{x} - \vec{y}) \quad (5.16)$$

where $\xi^{ij}(\vec{x}, \vec{y})$ is the 2-point correlation function and the last term is the contribution from shot noise, which we assume can be modeled by a Poisson process. The multi-tracer 2-point correlation function is generally assumed to be related to the matter correlation function, $\xi^{(m)}(\vec{x}, \vec{y})$, through some knowable relations such as tracer bias, redshift-space distortions, etc. In real space, the matter two-point correlation function can be written in terms of the matter power spectrum $P_m(k)$ as:

$$\xi^{(m)}(\vec{x}, \vec{y}) = \xi^{(m)}(|\vec{x} - \vec{y}|) = \int \frac{d^3k}{(2\pi)^3} e^{-i\vec{k} \cdot (\vec{x} - \vec{y})} P_m(k) \quad (5.17)$$

If we want to estimate the constraints on some set of parameters θ^μ , which can be derived from the two-point correlation functions, we start by constructing the Fisher information matrix for that set of parameters:

$$\begin{aligned}
F[\theta^\mu, \theta^\nu] &= F^{\mu\nu} = \frac{1}{2} \text{Tr} \left[\frac{\partial \Sigma}{\partial \theta^\mu} \Sigma^{-1} \frac{\partial \Sigma}{\partial \theta^\nu} \Sigma^{-1} \right] \\
&= \frac{1}{2} \sum_{ij i' j'} \int d^3 x \int d^3 y \int d^3 x' \int d^3 y' \\
&\quad \times \left\{ \frac{\partial \Sigma^{ij}(\vec{x}, \vec{y})}{\partial \theta^\mu} [\Sigma^{j i'}(\vec{y}, \vec{x}')]^{-1} \frac{\partial \Sigma^{i' j'}(\vec{x}', \vec{y}')}{\partial \theta^\nu} [\Sigma^{j' i}(\vec{y}', \vec{x})]^{-1} \right\}
\end{aligned} \tag{5.18}$$

where the compact notation for the trace in the definition of the first line denotes both of a sum over all tracers, as well as spatial integrals over all positions. As hinted by this expression, the tracer indices are linked to the spatial positions where we measure those same tracers: $i \leftrightarrow \vec{x}$, $j \leftrightarrow \vec{y}$, $i' \leftrightarrow \vec{x}'$, $j' \leftrightarrow \vec{y}'$. Therefore, we can use a compact notation where the sum over a tracer index also denotes the integral over its corresponding positions in space.

After computing the Fisher matrix we can invert it to find the covariance matrix for the parameters θ^μ :

$$\text{Cov}[\theta^\mu, \theta^\nu] = \text{Cov}_{\mu\nu} = [F_{\mu\nu}]^{-1} \tag{5.19}$$

Now, remember that we do not observe the Universe in snapshots, but over the past light-cone, and the surfaces of constant time are not flat, but 2D spherical shells. Therefore, to include all the physics that depends on evolution or redshift-space effects, it is necessary to describe them using spherical coordinates.

So we would like to express the number of galaxies that occupy some spherical shell of width Δx and radius $x \in [\bar{x} - \Delta x/2, \bar{x} + \Delta x/2]$ in terms of an expansion over spherical harmonics,

$$N_{\ell m}^i(\bar{x}) = \int_{\bar{x}} d^3 x Y_{\ell m}^*(\hat{x}) n^i(\vec{x}) = \int_{\bar{x} - \Delta x/2}^{\bar{x} + \Delta x/2} dx x^2 \int d^2 \hat{x} Y_{\ell m}^*(\hat{x}) n^i(\vec{x}) \tag{5.20}$$

From this, we define the harmonic space 2-point correlation function (i.e. the angular power spectrum) in terms of the expectation value of the counts inside the spherical shells

of radii \bar{x} and \bar{y} :

$$\begin{aligned} \langle \delta N_{\ell m}^i(\bar{x}) \delta N_{\ell' m'}^{j*}(\bar{y}) \rangle &= \Sigma_{\ell}^{ij}(\bar{x}, \bar{y}) \delta_{\ell\ell'} \delta_{mm'} \\ &= \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \int d^2 \hat{x} \int d^2 \hat{y} Y_{\ell m}^*(\hat{x}) Y_{\ell' m'}(\hat{y}) \Sigma_{\ell}^{ij}(\bar{x}, \bar{y}) \end{aligned} \quad (5.21)$$

where symmetry under rotations leads to the diagonal structure of the correlation function in harmonic space. Notice that the harmonic data covariance above reduces to the angular power spectrum only after subtracting shot noise, $\Sigma_{\ell}^{ij}(\bar{x}, \bar{y}) \rightarrow C_{\ell}^{ij}(\bar{x}, \bar{y})$.

Let's take a step back and look at the covariance matrix in configuration space. Note that it can be written in terms of the harmonic space covariance by summing over the spherical harmonics:

$$\begin{aligned} \Sigma^{ij}(\bar{x}, \bar{y}) &= \frac{4\pi}{\Delta V_{\bar{x}}} \frac{4\pi}{\Delta V_{\bar{y}}} \sum_{\ell m} Y_{\ell m}(\hat{x}) Y_{\ell m}^*(\hat{y}) \Sigma_{\ell}^{ij}(\bar{x}, \bar{y}) \\ &= \frac{4\pi}{\Delta V_{\bar{x}}} \frac{4\pi}{\Delta V_{\bar{y}}} \sum_{\ell} \frac{2\ell + 1}{4\pi} \mathcal{L}_{\ell}(\hat{x} \cdot \hat{y}) \Sigma_{\ell}^{ij}(\bar{x}, \bar{y}) \end{aligned} \quad (5.22)$$

where the volume of the spherical shell around \bar{x} is given by $\Delta V_{\bar{x}} = \int_{\bar{x}} d^3x = 4\pi \bar{x}^2 \Delta x$, and the last line makes it clear that the correlation function depends on the angle only through the combination $\mu = \hat{x} \cdot \hat{y}$ that appears in the argument of the Legendre polynomial \mathcal{L}_{ℓ} . Now, if we are able to find an expression for the inverse of the data covariance in harmonic space, such that:

$$\sum_j \sum_{\bar{y}} [\Sigma_{\ell}^{ij}(\bar{x}, \bar{y})]^{-1} \Sigma_{\ell}^{j' i'}(\bar{y}, \bar{x}') = \delta_{ii'} \delta_{\bar{x}\bar{x}'} \quad (5.23)$$

then we could define the inverse of the data covariance in real space in analogy to Eq. (5.22):

$$[\Sigma^{ij}(\bar{x}, \bar{y})]^{-1} = \frac{\Delta V_{\bar{x}}}{4\pi} \frac{\Delta V_{\bar{y}}}{4\pi} \sum_{\ell m} Y_{\ell m}(\hat{x}) Y_{\ell m}^*(\hat{y}) [\Sigma_{\ell}^{ij}(\bar{x}, \bar{y})]^{-1} \quad (5.24)$$

Taking the limit back to the continuum by making the volume elements infinitesimally small, it is then straightforward to show, with the help of Eq. (5.23) and the orthogonality

of the spherical harmonics,

$$\sum_{\ell m} Y_{\ell m}(\hat{x}) Y_{\ell m}^*(\hat{y}) = \delta_D(\hat{x} - \hat{y}),$$

that the expression above, Eq. (5.24), is indeed the inverse of the covariance:

$$\sum_j \int d^3 y [\Sigma^{ij}(\vec{x}, \vec{y})]^{-1} \Sigma^{j' i'}(\vec{y}, \vec{x}') = \delta_{ii'} \delta_D(\vec{x} - \vec{x}') \quad (5.25)$$

We can now go back to the expression for the Fisher matrix, Eq. (5.18), and plug in the the covariance and its inverse, Eqs. (5.22) and (5.24), to obtain, after some trivial algebra:

$$\begin{aligned} F^{\mu\nu} = & \frac{1}{2} \sum_{\ell} (2\ell + 1) \sum_{i, \bar{x}} \sum_{j, \bar{y}} \sum_{i', \bar{x}'} \sum_{j', \bar{y}'} \frac{\Delta V_{\bar{x}}}{4\pi} \frac{\Delta V_{\bar{y}}}{4\pi} \frac{\Delta V_{\bar{x}'}}{4\pi} \frac{\Delta V_{\bar{y}'}}{4\pi} \\ & \times \frac{\partial \Sigma_{\ell}^{ij}(\bar{x}, \bar{y})}{\partial \theta^{\mu}} \left[\Sigma_{\ell}^{j' i'}(\bar{y}, \bar{x}') \right]^{-1} \frac{\partial \Sigma_{\ell}^{i' j'}(\bar{x}', \bar{y}')}{\partial \theta^{\nu}} \left[\Sigma_{\ell}^{j' i}(\bar{y}', \bar{x}) \right]^{-1} \end{aligned} \quad (5.26)$$

Notice that the angular integrals cancel all the dependence on the spherical harmonics, such that only one set of (ℓ, m) survives, resulting in the factor $\sum_{m=-\ell}^{\ell} = 2\ell + 1$ since no term depends on m . This shows that the Fisher matrix can be decomposed into linearly independent spherical harmonic components:

$$F^{\mu\nu} = \sum_{\ell} F_{\ell}^{\mu\nu} \quad (5.27)$$

where we define:

$$F_{\ell}^{\mu\nu} = \frac{2\ell + 1}{2} \text{Tr} \left\{ \frac{\partial \Sigma_{\ell}^{ij}}{\partial \theta^{\mu}} \left[\Sigma_{\ell}^{j' i'} \right]^{-1} \frac{\partial \Sigma_{\ell}^{i' j'}}{\partial \theta^{\nu}} \left[\Sigma_{\ell}^{j' i} \right]^{-1} \right\} \quad (5.28)$$

with the trace denoting, in short-hand notation, both sums over tracer indices as well as integrals over the corresponding radii, $\sum_{\bar{x}} \Delta V_{\bar{x}}/4\pi(\dots) \rightarrow \int d\bar{x} \bar{x}^2(\dots)$. Since we will choose our parameters $\theta^{\mu} \rightarrow C_{\ell}^{ij}$, and $\partial C_{\ell}^{ij}/\partial C_{\ell'}^{i' j'} \sim \delta_{\ell\ell'}$ the full Fisher matrix for the correlation functions in harmonic space is given by Eq. (5.28). Note that in this derivation we have assumed that the whole sky was available, but as discussed in Section 2, this is never the case and with partial sky coverage $f_{sky} = \Delta\Omega/4\pi$, where $\Delta\Omega$ is the survey

angular area, and collecting different ℓ 's inside a bin $\bar{\ell}$, the expression becomes (Hu and Jain, 2004):

$$F_{\bar{\ell}}^{\mu\nu} = \frac{f_{sky}}{2} \sum_{\ell \in \bar{\ell}} (2\ell + 1) \text{Tr} \left\{ \frac{\partial \Sigma_{\ell}^{ij}}{\partial \theta^{\mu}} \left[\Sigma_{\ell}^{jj'} \right]^{-1} \frac{\partial \Sigma_{\ell}^{i'j'}}{\partial \theta^{\nu}} \left[\Sigma_{\ell}^{j'i} \right]^{-1} \right\} \quad (5.29)$$

5.3 Inverting the data covariance

From Eq. (5.28), it is clear that in order to compute the Fisher matrix, we must first invert the data covariance, and this turns out to be the hardest part. For any realistic scenario, it is hopeless to invert such a matrix numerically due to its massive size. So we must find an analytical expression for this inverse, defined such that:

$$\sum_j \sum_{\bar{y}} \Sigma_{\ell}^{ij}(\bar{x}, \bar{y}) \left[\Sigma_{\ell}^{j'i'}(\bar{y}, \bar{x}') \right]^{-1} = \delta_{ii'} \delta_{\bar{x}\bar{x}'} \quad (5.30)$$

Let's start by going back to the covariance in harmonic space, Eq. (5.21), and writing it in the form:

$$\begin{aligned} \Sigma_{\ell}^{ij} &= \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \int d^2 \hat{x} \int d^2 \hat{y} Y_{\ell m}(\hat{x}) Y_{\ell m}^*(\hat{y}) \langle \delta n^i(\vec{x}) \delta n^j(\vec{y}) \rangle \\ &= \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \int d^2 \hat{x} \int d^2 \hat{y} Y_{\ell m}(\hat{x}) Y_{\ell m}^*(\hat{y}) \left[\bar{n}^i(\vec{x}) \bar{n}^j(\vec{y}) \xi^{ij}(\vec{x}, \vec{y}) \right. \\ &\quad \left. + \delta_{ij} \bar{n}^i(\vec{x}) \delta_D(\vec{x} - \vec{y}) \right] \end{aligned} \quad (5.31)$$

We can identify the first term with the correlation function of the counts in shells in harmonic space (i.e. the angular power spectrum), and the second term is the shot noise. For simplicity, here we will assume that the mean numbers of tracers can be pulled outside the integral, though they can vary as a function of the angular coordinates due to masks or selection functions. As discussed before, we can also simplify the notation, $\bar{n}^i = \bar{n}^i(\bar{x})$,

$\bar{n}^j = \bar{n}^j(\bar{y})$. The angular power spectrum is then:

$$\begin{aligned} C_\ell^{ij} &= \bar{n}^i \bar{n}^j \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \int d^2 \hat{x} Y_{\ell m}(\hat{x}) \int d^2 \hat{y} Y_{\ell m}^*(\hat{y}) \xi^{ij}(\vec{x}, \vec{y}) \\ &= \bar{n}^i \bar{n}^j \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \int d^2 \hat{x} Y_{\ell m}(\hat{x}) \int d^2 \hat{y} Y_{\ell' m'}^*(\hat{y}) \int \frac{d^3 k}{(2\pi)^3} e^{i\vec{k}(\vec{x}-\vec{y})} P^{ij}(\vec{k}) \end{aligned} \quad (5.32)$$

where we assumed homogeneity, and that the correlation function is computed at a hypersurface of constant time. We now use the Rayleigh expansion of a plane wave in spherical harmonics:

$$e^{i\vec{k}\cdot\vec{x}} = 4\pi \sum_{\ell m} i^\ell j_\ell(kx) Y_{\ell m}(\hat{k}) Y_{\ell m}^*(\hat{x}) \quad (5.33)$$

and after some trivial algebra (see the real space term in Appendix A) we get:

$$C_\ell^{ij} = \bar{n}^i \bar{n}^j \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \times \frac{2}{\pi} \int_0^\infty dk k^2 j_\ell(kx) j_\ell(ky) P^{ij}(k) \quad (5.34)$$

In the linear regime we have $P^{ij}(k) = b^i b^j D^i D^j P_m(k)$, and if the redshift slices are thin enough, we can take the growth functions and biases of the tracers outside the volume integral and simplify the expression above:

$$\begin{aligned} C_\ell^{ij}(\bar{x}, \bar{y}) &= \bar{n}^i \bar{n}^j b^i b^j D^i D^j \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \times \frac{2}{\pi} \int_0^\infty dk k^2 j_\ell(kx) j_\ell(ky) P_m(k) \\ &= \bar{n}^i \bar{n}^j b^i b^j D^i D^j C_\ell^{(m)}(\bar{x}, \bar{y}) \end{aligned} \quad (5.35)$$

where $C_\ell^{(m)}(\bar{x}, \bar{y})$ is the matter angular power spectrum. This expression can also be extended to take into account non-linearities (see, e.g., Grasshorn Gebhardt and Jeong, 2020).

If the spherical shells are sufficiently thin, such that the argument of the spherical Bessel functions can be approximated by the mean radius, $x \rightarrow \bar{x}$, then the angular power

spectrum further simplifies to:

$$C_\ell^{ij}(\bar{x}, \bar{y}) \approx \bar{N}^i \bar{N}^j \times \frac{2}{\pi} \int_0^\infty dk k^2 j_\ell(k\bar{x}) j_\ell(k\bar{y}) P^{ij}(k) \quad (5.36)$$

where in the second line we defined the mean number of tracers *per unit solid angle*, $\bar{N}_x^i = \bar{n}^i(\bar{x}) \Delta V_x / 4\pi$. Though we will not rely on this approximation, it will be useful later to simplify the notation.

The result presented in Eq. (5.34), together with the closure relation for spherical Bessel functions,

$$\int_0^\infty dt t^2 j_\ell(at) j_\ell(bt) = \frac{\pi}{2} \frac{\delta(a-b)}{a^2} \quad (5.37)$$

allows us to rewrite the harmonic data covariance, Eq. (5.31), as:

$$\Sigma_\ell^{ij} = \bar{n}^i \bar{n}^j \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \times \frac{2}{\pi} \int_0^\infty dk k^2 j_\ell(kx) j_\ell(ky) M^{ij}(k) \quad (5.38)$$

where we defined the matrix:

$$M^{ij}(k) = P^{ij}(k) + \frac{\delta_{ij}}{\sqrt{\bar{n}^i \bar{n}^j}} \quad (5.39)$$

It is easy to show that the inverse data covariance is given by:

$$\left[\Sigma_\ell^{j'j}(y, x') \right]^{-1} = \frac{1}{\bar{n}^j} \frac{1}{\bar{n}^{j'}} \times \frac{2}{\pi} \int_0^\infty dk' k'^2 j_\ell(k'y) j_\ell(k'x') \left[M^{j'j'}(k') \right]^{-1} \quad (5.40)$$

which, together with the expression for the covariance Eq. (5.31), satisfies the definition of the inverse, Eq. (5.23). So now we need to find an expression for the inverse of the matrix M^{ij} . First, we rewrite it as:

$$M^{ij}(k) = \frac{1}{\sqrt{\bar{n}^i}} \left[\delta^{ij} + \mathcal{P}^{ij} \right] \frac{1}{\sqrt{\bar{n}^j}} \quad (5.41)$$

where $\mathcal{P}^{ij} = \sqrt{\bar{n}^i \bar{n}^j} P^{ij}$. Notice also that this adimensional quantity is separable, in the sense that $\mathcal{P}^{ij} \mathcal{P}^{i'j'} = \mathcal{P}^{ii'} \mathcal{P}^{jj'}$, and $\sum_l \mathcal{P}^{il} \mathcal{P}^{jl} = \mathcal{P}^{ij} \mathcal{P}$, where we defined the trace $\mathcal{P} = \sum_i \mathcal{P}^{ii}$.

We will assume here that the biases, growth rate and number densities of the tracers vary slowly with the radii, compared to the typical scales where we measure clustering. Under these approximations, one can easily check that the inverse of Eq. (5.41) is given by:

$$[M^{ij}]^{-1} = \sqrt{\bar{n}^i} \left[\delta_{ij} - \frac{\mathcal{P}^{ij}}{1 + \mathcal{P}} \right] \sqrt{\bar{n}^j} \quad (5.42)$$

The fundamental reason we were able to invert this expression is that the data covariance in real space can be written as a function of its Fourier conjugate:

$$\Sigma_\ell^{ij}(\bar{x}, \bar{y}) = N_{\bar{x}}^i N_{\bar{y}}^j \frac{2}{\pi} \int_0^\infty dk k^2 \frac{2}{\pi} \int_0^\infty dk' k'^2 j_\ell(k\bar{x}) \Sigma_\ell^{ij}(k, k') j_\ell(k'\bar{y}) \quad (5.43)$$

where we defined

$$\Sigma_\ell^{ij}(k, k') = \frac{\pi}{2} \frac{\delta(k - k')}{k^2} \left[\frac{\delta^{ij}}{\bar{n}^i} + P^{ij}(k) \right] = \frac{\pi}{2} \frac{\delta(k - k')}{k^2} M^{ij}(k) \quad (5.44)$$

which is perfectly diagonal in the Fourier modes. So the problem was reduced to finding the trivial inverse of the matrix M^{ij} . A similar approach will be used in Section 6 when we generalize the results of this section to include redshift space distortions.

5.4 The Fisher matrix in harmonic space

Now that we have expressions for the inverse of the harmonic data covariance, we can return to Eq. (5.28) and write the Fisher matrix for the observables of interest, the angular power spectra, $\theta^\mu \rightarrow C_\ell^{ij}(\bar{x}, \bar{y})$.

But first, it is worth noting that these angular spectra are symmetric, $C_\ell^{ij} = C_\ell^{ji}$, and so are the corresponding data covariances. Therefore, if we simply identify the parameters of the Fisher matrix like that, we would be counting the cross-correlations twice. We handle this double-counting by defining non-degenerate spectra as:

$$C_\ell^{[ij]} = \begin{cases} C_\ell^{ij} & i \leq j \\ 0 & i > j \end{cases} \quad (5.45)$$

Note that we can write the angular power spectrum in terms of their non-degenerate

counterpart:

$$C_\ell^{ij}(\bar{x}, \bar{y}) = C_\ell^{[ij]}(\bar{x}, \bar{y}) + C_\ell^{[ji]}(\bar{y}, \bar{x}) - \delta_{ij} C_\ell^{ij}(\bar{x}, \bar{y}) \quad (5.46)$$

where the last term of this expression is such that $C_\ell^{ii}(\bar{x}, \bar{y}) = C_\ell^{ii}(\bar{y}, \bar{x})$, by definition.

We can now evaluate the partial derivatives, remembering that a given tracer index always refers to the same radial bin, so we can use the short-hand notation $\bar{\delta}_{i' i'} = \delta_{i' i'} \delta_{\bar{x} \bar{x}'}$, leading to:

$$\frac{\delta C_\ell^{ij}}{\delta C_\ell^{[i' j']}} = \bar{\delta}_{i' i'} \bar{\delta}_{j j'} + \bar{\delta}_{i j'} \bar{\delta}_{j' i'} - \bar{\delta}_{i j} \bar{\delta}_{j' i'} \bar{\delta}_{i' j'} \bar{\delta}_{j' i} \quad (5.47)$$

Substituting this into Eq. (5.28) finally gives us the Fisher matrix for the angular power spectrum in real space:

$$F[C_\ell^{[ij]}, C_\ell^{[i' j']}] = \frac{2\ell + 1}{4} (2 - \bar{\delta}_{ij}) (2 - \bar{\delta}_{i' j'}) \left\{ [\Sigma_\ell^{ii'}]^{-1} [\Sigma_\ell^{jj'}]^{-1} + [\Sigma_\ell^{ij'}]^{-1} [\Sigma_\ell^{i' j}]^{-1} \right\} \quad (5.48)$$

And the covariance matrix in harmonic space follows directly from it:

$$\text{Cov}[C_\ell^{[ij]}, C_\ell^{[i' j']}] = \frac{1}{2\ell + 1} \left[\Sigma_\ell^{i' i'} \Sigma_\ell^{j j'} + \Sigma_\ell^{i j'} \Sigma_\ell^{i' j} \right] \quad (5.49)$$

6 Fisher matrix in redshift space

In this section we will take into account the linear redshift space distortions described in Section 4 to generalize the procedure of Section 5 and compute the Fisher matrix directly in redshift space. Analogously to the real space case, the two-point correlation function is defined here as the expectation value of the counts inside the spherical shells of radii \bar{x} and \bar{y} , but now including the effects of peculiar velocities which adds some extra terms, as shown in the derivation leading to Eq. (4.20):

$$C_{\ell,s}^{ij}(\bar{x}, \bar{y}) = \bar{n}^i(\bar{x})\bar{n}^j(\bar{y}) \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \frac{2}{\pi} \int_0^\infty dk k^2 \times [j_\ell(kx) - \beta^i(z)j_\ell''(kx)] [j_\ell(ky) - \beta^j(z')j_\ell''(ky)] P^{ij}(k) \quad (6.1)$$

with $\beta^i = f(z)/b^i$, where $f(z)$ is the growth rate. Notice that making $\beta^i \rightarrow 0$ is equivalent to cancelling the effects of the peculiar velocities, and this expression indeed reduces to the real space correlation function given by Eq. (5.34). Therefore, in redshift space the harmonic data covariance is simply a generalization of Eq. (5.38):

$$\Sigma_{\ell,s}^{ij}(\bar{x}, \bar{y}) = \bar{n}^i\bar{n}^j \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \frac{2}{\pi} \int_0^\infty dk k^2 \times \left\{ [j_\ell(kx) - \beta^i j_\ell''(kx)] [j_\ell(ky) - \beta^j j_\ell''(ky)] P^{ij}(k) + j_\ell(kx)j_\ell(ky) \frac{\delta_{ij}}{\sqrt{\bar{n}^i\bar{n}^j}} \right\} \quad (6.2)$$

Just like in real space, the most difficult step in the computation of the Fisher matrix is finding the inverse of the data covariance in the same sense of Eq. (5.23), that is:

$$\sum_j \sum_{\bar{y}} [\Sigma_{\ell,s}^{ij}(\bar{x}, \bar{y})]^{-1} \Sigma_{\ell,s}^{j' i'}(\bar{y}, \bar{x}') = \delta_{i' i} \delta_{\bar{x}' \bar{x}} \quad (6.3)$$

In order to accomplish this, in analogy to what was done in real space, we will make use of analytical solutions to the integrals of products of spherical Bessel functions that appear in Eq. (6.2). These solutions, however, were nowhere to be found in the literature,

so we have to derive them.

6.1 Integrating products of spherical Bessel functions

First, note that the term without any derivatives is simply the closure relation, Eq. (5.37), which was used in real space. The difficulty arises with the terms involving the second derivatives, for which we recall that the spherical Bessel functions satisfy two important recursion relations:

$$j_\ell(z) = \frac{\ell - 1}{z} j_{\ell-1}(z) - \partial_z j_{\ell-1}(z) \quad (6.4)$$

$$j_\ell(z) = \frac{2\ell - 1}{z} j_{\ell-1}(z) - j_{\ell-2}(z) \quad (6.5)$$

These relations can be used to find a differential equation for the spherical Bessel functions:

$$z^2 j_\ell''(z) + 2z j_\ell'(z) + [z^2 - \ell(\ell + 1)] j_\ell(z) = 0 \quad (6.6)$$

which can be combined again with the first recursion relation above to write an expression for the second derivative as a function of spherical Bessel functions:

$$j_\ell''(z) = \frac{1}{z^2} \{ [\ell^2 - \ell - z^2] j_\ell(z) + 2z j_{\ell+1}(z) \} \quad (6.7)$$

Before continuing, let's quickly draw a distinction between our path and that chosen since Fisher, Scharf and Lahav (1994) first wrote an expression for the angular power spectrum in redshift space. After that seminal paper, many others (e.g. Padmanabhan et al., 2007) expressed the second derivative of the spherical Bessel function in terms of Bessel functions of different orders:

$$j_\ell''(z) = \frac{\ell(\ell - 1)}{(2\ell + 1)(2\ell - 1)} j_{\ell-2}(z) + \frac{(\ell + 1)(\ell + 2)}{(2\ell + 1)(2\ell + 3)} j_{\ell+2}(z) - \frac{(2\ell^2 + 2\ell - 1)}{(2\ell + 3)(2\ell - 1)} j_\ell(z) \quad (6.8)$$

In this work, on the other hand, we make use of the alternative expression given by Eq. (6.7), which, as we will see, leads to analytical integrals.

Let's first solve the term with only one derivative of the spherical Bessel function:

$$A = \int dk k^2 j_\ell(kx) j_\ell''(ky) \quad (6.9)$$

by using the expression above, Eq (6.7), to split it into three integrals, with no more derivatives involved, which should simplify our problem:

$$\begin{aligned} A &= \int \frac{dk k^2 j_\ell(kx)}{(ky)^2} \left\{ [\ell^2 - \ell - k^2 y^2] j_\ell(ky) + 2ky j_{\ell+1}(ky) \right\} \\ &= \frac{(\ell^2 - \ell)}{y^2} \underbrace{\int dk j_\ell(kx) j_\ell(ky)}_{A_1} - \underbrace{\int dk k^2 j_\ell(kx) j_\ell(ky)}_{A_2} + \frac{2}{y} \underbrace{\int dk k j_\ell(kx) j_{\ell+1}(ky)}_{A_3} \end{aligned} \quad (6.10)$$

For the first term, we use the following well known result (see e.g. Bloomfield et al. 2017):

$$A_1 = \int dk j_\ell(kx) j_\ell(ky) = \frac{\pi}{2} \frac{1}{(2\ell + 1)r_>} \left(\frac{r_<}{r_>} \right)^\ell \quad (6.11)$$

where $r_>$ and $r_<$ are, respectively, the larger and smaller value between x and y . The second term is, once again, simply the closure relation:

$$A_2 = \int_0^\infty dk k^2 j_\ell(kx) j_\ell(ky) = \frac{\pi}{2} \frac{\delta_D(x - y)}{x^2} \quad (6.12)$$

And for the last one, we remember that the spherical Bessel function is defined in terms of the usual Bessel function of the first kind:

$$j_\ell(z) = \sqrt{\frac{\pi}{2z}} J_{\ell+1/2}(z) \quad (6.13)$$

so we can rewrite that term as:

$$A_3 = \int dk k j_\ell(kx) j_{\ell+1}(ky) = \frac{\pi}{2} \frac{1}{\sqrt{xy}} \int dk J_{\ell+1/2}(kx) J_{\ell+3/2}(ky) \quad (6.14)$$

This integral is also well known and can be found in e.g. the classic book *A Treatise on the Theory of Bessel Functions* by Watson (1944), and gives us:

$$A_3 = \int dk k j_\ell(kx) j_{\ell+1}(ky) = \frac{\pi}{2} \times \begin{cases} \frac{1}{2x^2} & \text{if } x = y \\ 0 & \text{if } x > y \\ \frac{x^\ell}{y^{\ell+2}} & \text{if } x < y \end{cases} \quad (6.15)$$

Plugging these results back into (6.10), we obtain the first analytical integral we will need in order to invert the data covariance:

$$\int dk k^2 j_\ell(kx) j_\ell''(ky) = -\frac{\pi}{2x^2} \delta_D(x - y) + H_\ell(x, y) \quad (6.16)$$

where:

$$H_\ell(x, y) = \frac{\pi}{2y^2} \begin{cases} \frac{1}{x} \left[\frac{2(2\ell + 1) + \ell(\ell - 1)}{2(2\ell + 1)} \right] & \text{if } x = y \\ \frac{1}{x} \left(\frac{y}{x} \right)^\ell \left[\frac{2(2\ell + 1) + \ell(\ell - 1)}{(2\ell + 1)} \right] & \text{if } x > y \\ \frac{1}{y} \frac{\ell(\ell - 1)}{(2\ell + 1)} \left(\frac{x}{y} \right)^\ell & \text{if } x < y \end{cases} \quad (6.17)$$

Notice that, curiously, this result is discontinuous, but that does not turn out to be a problem for our purposes. As we will see later on, this integral appears twice, with interchanged arguments in the Bessel functions $\bar{x} \leftrightarrow \bar{y}$. In particular, when we consider the auto-correlations of the same tracer, their symmetric combination is a continuous expression (see Figure 9):

$$\tilde{H}_\ell = \frac{1}{2} [H_\ell(x, y) + H_\ell(y, x)] = \frac{\pi}{2r_>} \left(\frac{r_<}{r_>} \right)^\ell \left[\frac{2(2\ell + 1)r_<^2 + \ell(\ell - 1)(x^2 + y^2)}{2(2\ell + 1)x^2y^2} \right] \quad (6.18)$$

Now, we still need to find an analytical expression for the term in Eq (6.2) which contains two derivatives of the spherical Bessel function:

$$B = \int dk k^2 j_\ell''(kx) j_\ell''(ky) \quad (6.19)$$

In similar fashion to what we did for Eq (6.9), we rewrite the derivatives and split this integral into nine parts:

$$\begin{aligned}
B &= \int \frac{dk k^2}{(kx)^2 (ky)^2} \{ [\ell^2 - \ell - k^2 y^2] j_\ell(ky) + 2ky j_{\ell+1}(ky) \} \\
&\quad \times \{ [\ell^2 - \ell - k^2 x^2] j_\ell(kx) + 2kx j_{\ell+1}(kx) \} \\
&= \frac{(\ell^2 - \ell)^2}{(xy)^2} \underbrace{\int \frac{dk}{k^2} j_\ell(kx) j_\ell(ky)}_{B1} - \frac{(\ell^2 - \ell)}{y^2} \underbrace{\int dk j_\ell(kx) j_\ell(ky)}_{B2} \\
&\quad + 2 \frac{(\ell^2 - \ell)}{xy^2} \underbrace{\int \frac{dk}{k} j_{\ell+1}(kx) j_\ell(ky)}_{B3} - \frac{(\ell^2 - \ell)}{x^2} \underbrace{\int dk j_\ell(kx) j_\ell(ky)}_{B4} \quad (6.20) \\
&\quad + \underbrace{\int dk k^2 j_\ell(kx) j_\ell(ky)}_{B5} - \frac{2}{x} \underbrace{\int dk k j_{\ell+1}(kx) j_\ell(ky)}_{B6} \\
&\quad + 2 \frac{(\ell^2 - \ell)}{yx^2} \underbrace{\int \frac{dk}{k} j_{\ell+1}(ky) j_\ell(kx)}_{B7} - \frac{2}{y} \underbrace{\int dk k j_{\ell+1}(ky) j_\ell(kx)}_{B8} \\
&\quad + \frac{4}{xy} \underbrace{\int dk j_{\ell+1}(kx) j_{\ell+1}(ky)}_{B9}
\end{aligned}$$

We already know how to compute most of these terms, except for $B1$ and $B3$ (which is identical to $B7$). Although they have an analytical solution in terms of hypergeometric functions (see Watson (1944)), more suitable expressions can be found for integer values of ℓ . We start by rewriting the integral in $B1$ as:

$$B1 = r_{>} \int_0^\infty \frac{dq}{q^2} j_\ell(q) j_\ell \left(q \frac{r_{<}}{r_{>}} \right) \quad (6.21)$$

with $q = kr_{>}$, and by solving this integral numerically for some multipoles, we can find the pattern:

$$B1 = \int_0^\infty \frac{dk}{k^2} j_\ell(kx) j_\ell(ky)$$

$$= \frac{\pi 2^\ell (\ell + 1)! (2\ell - 3)!!}{2 (2\ell + 3)!} r_{>} \left(\frac{r_{<}}{r_{>}} \right)^\ell \left[2\ell + 3 - (2\ell - 1) \left(\frac{r_{<}}{r_{>}} \right)^2 \right] \quad (6.22)$$

Similarly, for the integral in B_3 , we obtain:

$$B_3 = \int_0^\infty \frac{dk}{k} j_{\ell+1}(kx) j_\ell(ky) = \frac{\pi}{2} \begin{cases} \frac{1}{(2\ell + 1)(2\ell + 3)} & \text{if } x = y \\ \frac{1}{(2\ell + 1)(2\ell + 3)} \left(\frac{x}{y} \right)^{\ell+1} & \text{if } x < y \\ \frac{1}{2} \left(\frac{y}{x} \right)^\ell \left[\frac{1}{2\ell + 1} - \frac{1}{2\ell + 3} \left(\frac{y}{x} \right)^2 \right] & \text{if } x > y \end{cases} \quad (6.23)$$

These closed forms of the integrals in B_1 and B_3 are new results, as far as we are aware. We checked the validity of these expressions by numerically performing those integrals with the help of the FFTLog algorithm (see Appendix B), which basically computes a Hankel transform – Eq B.1. In Figures 7 and 8, I plot a comparison between the analytical expressions, Eqs. (6.22) and (6.23), and the numerical solution to those integrals, from $\ell = 2$ up to $\ell = 50$. Though extensions to the original FFTLog have been developed to include integrals involving derivatives of Bessel functions Fang et al. (2020b) and to deal with non-Gaussian covariances Fang et al. (2020a), here I use a simple implementation of the algorithm to perform these integrals.

We can clearly see that our analytical results clearly match the numerical Hankel transform to exquisite accuracy. It is important to note that one might run into numerical issues with our analytical expression for the first integral, Eq. (6.22), at high multipoles because then the argument of the factorials become very large. *Python*, for instance, does not naturally support arbitrarily large float, but in principle, this problem can be solved by using e.g. the *decimal* module.

On the other hand, when $x \neq y$, the FFTLog result also becomes very unstable after a certain value of ℓ . Increasing the number of points for the integration interval does help alleviate this issue, but this defeats the whole point, since to get a stable result up to $\ell = 60$, the FFTLog algorithm for these particular integrals takes as long as a simple quadrature routine, while for comparison the analytical result was nearly 2000 times faster without losing precision. Some extension to the code might be required to properly perform those integrals for higher multipoles, but here I will not dwell on that since at this point it is already clear that the analytical and numerical results agree with each other.

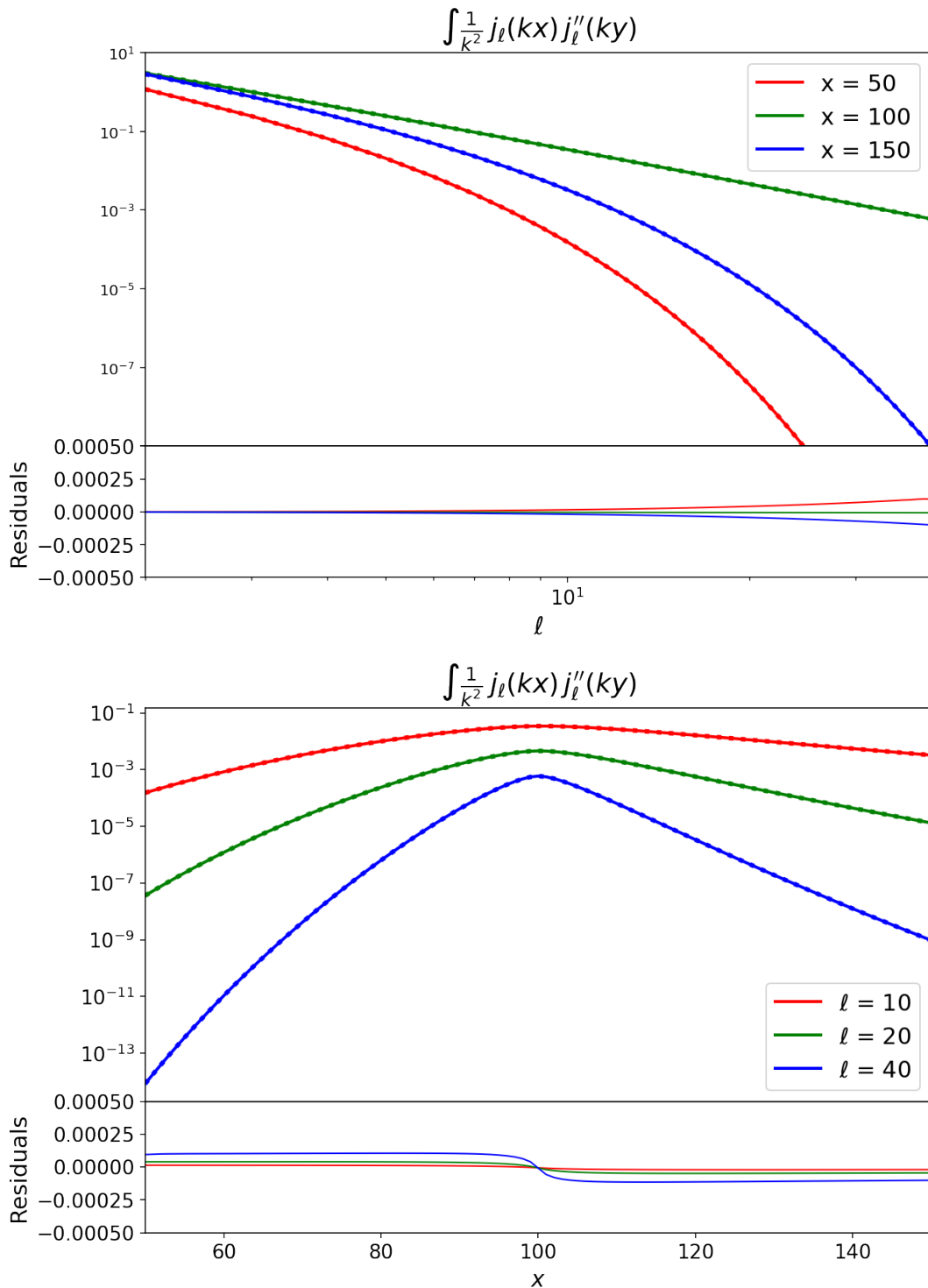


Figure 7: Integrals of the product of spherical Bessel functions with the k^{-2} term, using the FFTlog algorithm (dotted lines) and the analytical (solid lines) expression, Eq. (6.22), with relative residuals. We fix $y = 100$, for some $x = 50, 100, 200$ and vary ℓ in the top plot, while in the bottom plots we vary x for some fixed $\ell = 10, 20, 40$.

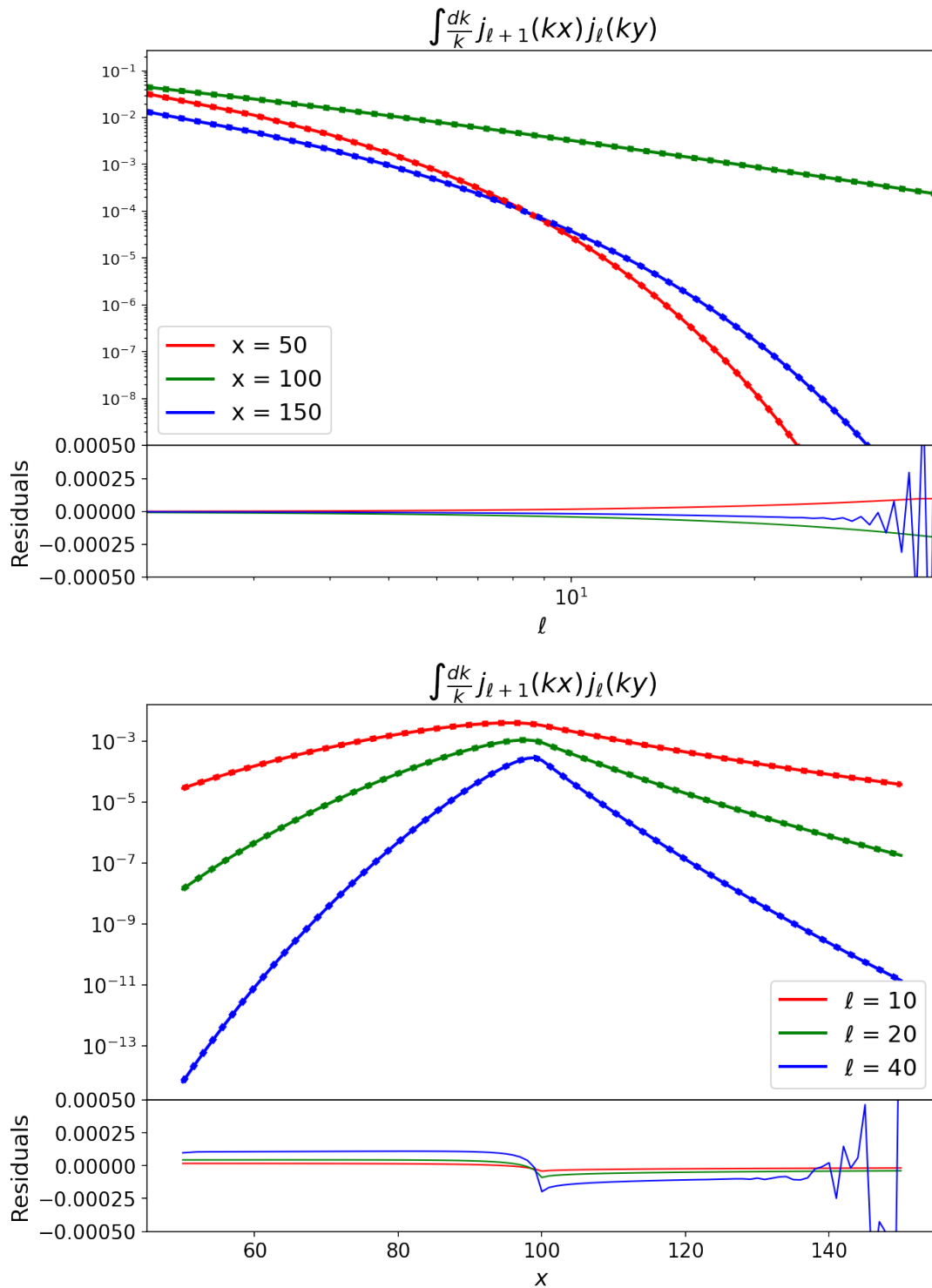


Figure 8: Integrals of the product of spherical Bessel functions with the k^{-1} term, using the FFTlog algorithm (dotted lines) and the analytical (solid lines) expression, Eq. (6.23), with relative residuals. We fix $y = 100$, for some $x = 50, 100, 200$ and vary ℓ in the top plot, while in the bottom plots we vary x for some fixed $\ell = 10, 20, 40$.

Note that these plots also show that those integrals quickly become negligible when $x \neq y$. This is a consequence of the fact that the spherical Bessel functions oscillate out of phase when they have different arguments, especially for higher multipoles. This hints at the possibility of not having to include very high multipoles and cross-correlations of widely separated bins, though how far is far enough will depend on the application and precision desired.

Using these new results, Eq. (6.22) and (6.23), along with equations (6.11), (6.12) and (6.15), we can finally write an analytical expression for the last integral we need — in addition to Eq. (6.16) — in order to invert the redshift space data covariance and ultimately compute the Fisher matrix:

$$\int_0^\infty dk k^2 j_\ell''(kx)j_\ell''(ky) = \frac{\pi}{2} \frac{\delta(x-y)}{x^2} + G_\ell(x, y) \quad (6.24)$$

where we defined:

$$\begin{aligned} G_\ell(x, y) = & \frac{\pi}{2} \left\{ \frac{4}{r_>^3} \frac{1}{(2\ell+3)} \left(\frac{r_<}{r_>}\right)^\ell - 2 \frac{r_<^\ell}{r_>^{\ell+3}} - \left(\frac{1}{r_>^2} + \frac{1}{r_<^2}\right) \frac{1}{r_>} \left(\frac{r_<}{r_>}\right)^\ell \frac{(\ell^2 - \ell)}{(2\ell+1)} \right. \\ & + \frac{2(\ell^2 - \ell)}{r_>^2 r_<} \left(\frac{r_<}{r_>}\right)^{\ell+1} \frac{1}{(2\ell+1)(2\ell+3)} \\ & + \frac{2(\ell^2 - \ell)}{r_<^2 r_>} \left(\frac{r_<}{r_>}\right)^\ell \left[\frac{1}{(2\ell+1)} - \frac{1}{(2\ell+3)} \left(\frac{r_<}{r_>}\right)^2 \right] \\ & \left. + \frac{(\ell^2 - \ell)^2}{(r_< r_>)^2} \frac{2^\ell (\ell+1)! (2\ell-3)!!}{(2\ell+3)!} r_> \left(\frac{r_<}{r_>}\right)^\ell \left[2\ell+3 - (2\ell-1) \left(\frac{r_<}{r_>}\right)^2 \right] \right\} \end{aligned} \quad (6.25)$$

Though this expression in Eq. (6.25) might look daunting, it is very easy and fast to compute numerically. In Figure 9, I plot this term, as well as the continuous version of $H_\ell(x, y)$ given in Eq. (6.18), with a fixed value of $y = 100$ and varying x for some multipoles $\ell = 10, 20, 40$. We can see that there is a clear peak at $x = y$, which becomes more and more pronounced for higher multipoles and seems to tend to a Dirac delta in the limit $\ell \rightarrow \infty$. This will lead to a semi diagonal structure for the redshift space data covariance and is what will allow us to find its inverse, to which we turn now.

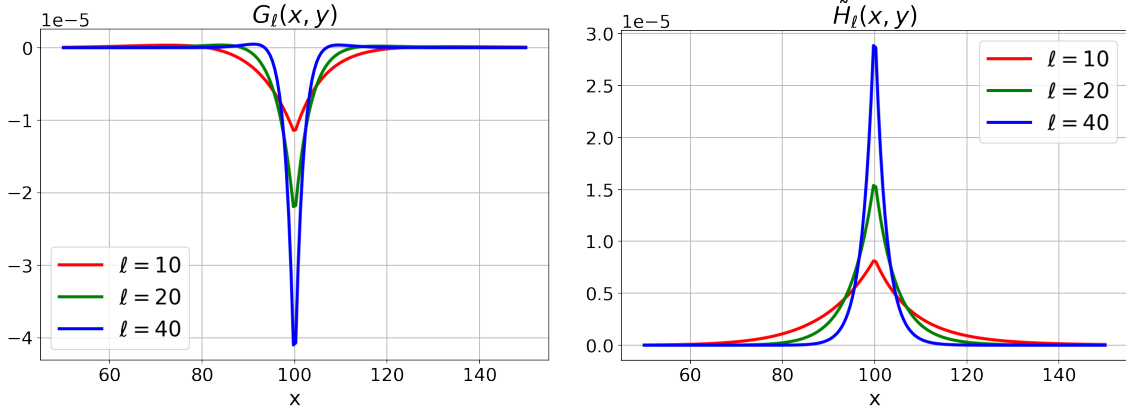


Figure 9: Plots of Eqs. (6.18) and (6.25) as a function of x with a fixed $y = 100$, for multipoles $\ell = 10, 20, 40$.

6.2 Inverting the data covariance in redshift space

In this section, I will show how to invert the expression for the data covariance. Like in real space, this turn out to be the hardest step of computing the Fisher matrix, and now it is even more complicated because there are three extra terms in the covariance, which we rewrite here as:

$$\begin{aligned}
 \Sigma_{\ell,s}^{ij}(\bar{x}, \bar{y}) &= \bar{n}^i(\bar{x}) \bar{n}^j(\bar{y}) \int_{\bar{x}} dx x^2 \int_{\bar{y}} dy y^2 \frac{2}{\pi} \int_0^\infty dk k^2 \left\{ [j_\ell(kx) - \beta^i j_\ell''(kx)] \right. \\
 &\quad \times [j_\ell(ky) - \beta^j j_\ell''(ky)] P^{ij}(k) + j_\ell(kx) j_\ell(ky) \frac{\delta^{ij}}{\sqrt{\bar{n}^i(x) \bar{n}^j(y)}} \left. \right\} \\
 &\approx \bar{N}^i \delta_{ij} \delta_{\bar{x}\bar{y}} \\
 &\quad + \bar{N}^i \bar{N}^j \frac{2}{\pi} \int_0^\infty dk k^2 [j_\ell(\bar{x}) - \beta^i j_\ell''(k\bar{x})] [j_\ell(k\bar{y}) - \beta^j j_\ell''(k\bar{y})] P^{ij}(k)
 \end{aligned} \tag{6.26}$$

where remember that the tracer and bin indices are linked and $\bar{N}^i = \bar{N}_{\bar{x}}^i = \bar{n}^i(\bar{x}) \Delta V_{\bar{x}} / 4\pi$ is the mean angular number density of tracers in the radial bin \bar{x} . In the second line, we simply used the approximation of Eq. (5.36) to shorten the notation, but our results do not depend on it.

Note that we don't have analytical expressions for these integrals, since the $P^{ij}(k)$ could, in principle, be any complicated function of k . But all our hard work was not in

vain, because we can rewrite our problem by using a condensed notation:

$$\frac{\pi}{2} \frac{\delta_D(k - k')}{k^2} \rightarrow \delta^{kk'} \quad , \quad \frac{2}{\pi} \int dk k^2 \rightarrow \sum_{[k]} \quad (6.27)$$

and defining a normalized data covariance:

$$\Sigma_{\ell,s}^{ij} \rightarrow \tilde{\Sigma}_{\ell,s}^{ij} \equiv \frac{1}{\sqrt{N^i N^j}} \Sigma_{\ell,s}^{ij} \quad (6.28)$$

Now let's also define the auxiliary ‘‘dual vectors’’:

$$V_\ell^i(k, \bar{x}) \equiv \sqrt{\frac{\Delta V_{\bar{x}}}{4\pi}} \mathcal{P}^{ii}(k) [j_\ell(k\bar{x}) - \beta^i j_\ell''(k\bar{x})] \quad (6.29)$$

where we used the definition of the clustering strength introduced in the previous section, Eq. (5.41), $\mathcal{P}^{ij}(k) = \sqrt{\bar{n}^i \bar{n}^j} P^{ij}(k)$. Using the associative property of the spectra $P^{ii}(k)P^{jj}(k') = P^{ij}(k)P^{ij}(k')$ (valid at least to first approximation, in the linear regime), we can write the data covariance as:

$$\begin{aligned} \tilde{\Sigma}_{\ell,s}^{ij}(\bar{x}, \bar{y}) &= \delta_{ij} \delta_{\bar{x}\bar{y}} + \sum_{[k],[k']} V_\ell^i(k, \bar{x}) \delta^{kk'} V_\ell^j(k', \bar{y}) \\ &= \delta_{ij} \delta_{\bar{x}\bar{y}} + \sum_{[k]} V_\ell^i(k, \bar{x}) V_\ell^j(k, \bar{y}) \end{aligned} \quad (6.30)$$

Now, we use the Woodbury matrix identity,

$$\begin{aligned} M &= A + XBY^{tr} \\ \Rightarrow M^{-1} &= A^{-1} + A^{-1}X(B^{-1} + Y^{tr}A^{-1}X)^{-1}Y^{tr}A^{-1} \end{aligned} \quad (6.31)$$

with $A \rightarrow \delta_{ij} \delta_{\bar{x}\bar{y}}$, $B \rightarrow \delta_{ij} \delta^{kk'}$ and $X, Y \rightarrow V_\ell^i(k, \bar{x}), V_\ell^j(k', \bar{y})$, to obtain:

$$[\tilde{\Sigma}_{\ell,s}^{ij}(\bar{x}, \bar{y})]^{-1} = \delta_{ij} \delta_{\bar{x}\bar{y}} \quad (6.32)$$

$$- \sum_{[k]} \sum_{[k']} V_\ell^i(k, \bar{x}) \left[\delta_{ij} \delta^{kk'} + \sum_{\bar{x}} V_\ell^i(k, \bar{x}) V_\ell^j(k', \bar{x}) \right]^{-1} V_\ell^j(k', \bar{y})$$

The matrix inside the square brackets can be thought of as the Fourier space analogous of the data covariance:

$$\tilde{\Sigma}_{\ell,s}^{ij}(k, k') = \delta_{ij} \delta^{kk'} + \sum_{\bar{x}} V_{\ell}^i(k, \bar{x}) V_{\ell}^i(k', \bar{x}) \quad (6.33)$$

Notice that the redshift space distortions introduce some off-diagonal structure in the Fourier data covariance, as opposed to the real space case where it was perfectly diagonal — see Eq. (5.44). Nevertheless, it is still dominated by the diagonal terms, so it is feasible to invert this object and then transform it back to the real space representation and obtain the inverse covariance in terms of the spherical shells.

Hence, the problem of inverting the multi tracer data covariance in redshift space is reduced to inverting Eq. (6.33). We use the fact that the correlation function is computed at a hypersurface of constant time, and the sum over the radii \bar{x} above corresponds to an integration over the radial coordinate in that hypersurface. Therefore, the quantities \bar{n}^i , b^i and $\beta^i = b^i/f$ can be considered constants in this expression, and we remember that $\sum_{\bar{x}} \int_{\bar{x}} dx = \int_0^{\infty} dx$, so that we can write the Fourier covariance in terms of the integrals we computed in the previous subsection, Eqs. (6.25) and (6.17):

$$\begin{aligned} \tilde{\Sigma}_{\ell}^{ij}(k, k') &= \delta_{ij} \delta^{kk'} + \sqrt{\mathcal{P}^{ii}(k) \mathcal{P}^{jj}(k')} \\ &\times \sum_{\bar{x}} \int_{\bar{x}} dx x^2 [j_{\ell}(kx) - \beta^i j_{\ell}''(kx)] [j_{\ell}(k'x) - \beta^j j_{\ell}''(k'x)] \\ &= \delta^{kk'} \left[\delta_{ij} + (1 + \beta^i)(1 + \beta^j) \sqrt{\mathcal{P}^{ii}(k) \mathcal{P}^{jj}(k')} \right] \\ &+ \sqrt{\mathcal{P}^{ii}(k) \mathcal{P}^{jj}(k')} [\beta^i \beta^j G_{\ell}(k, k') - \beta^i H_{\ell}(k, k') - \beta^j H_{\ell}(k', k)] \end{aligned}$$

Now it is convenient to gather the terms in the last line of Eq. (6.34) into the mixing matrix:

$$T_{\ell}^{ij}(k, k') \equiv \sqrt{\mathcal{P}^{ij}(k) \mathcal{P}^{ij}(k')} [\beta^i \beta^j G_{\ell}(k, k') - \beta^i H_{\ell}(k, k') - \beta^j H_{\ell}(k', k)] \quad (6.34)$$

where we used the definition of clustering strength \mathcal{P} introduced in the previous section,

$$\mathcal{P}^{ij}(k) = b^i b^j \sqrt{\bar{n}^i \bar{n}^j} P(k) = \sqrt{\bar{n}^i \bar{n}^j} P^{ij}(k) \quad (6.35)$$

We also define an analogous redshift-space clustering strength:

$$\mathcal{P}_{(s)}^{ij} \equiv (1 + \beta^i)(1 + \beta^j)\mathcal{P}^{ij} \quad (6.36)$$

and note that in the linear regime, the clustering strength (as well as $\mathcal{P}_{(s)}^{ij}$) can be factored, $\mathcal{P}^{ii}(k)\mathcal{P}^{jj}(k') = \mathcal{P}^{ij}(k)\mathcal{P}^{ij}(k')$, so we can rewrite the Fourier covariance as:

$$\begin{aligned} \tilde{\Sigma}_{\ell}^{ij}(k, k') &= \delta^{kk'} \left[\delta^{ij} + \mathcal{P}_{(s)}^{ij} \right] + T_{\ell}^{ij}(k, k') \\ &= \delta^{kk'} \left[\delta^{ij} + \mathcal{P}_{(s)}^{ij} + \lambda_{\ell}^{ij}(k) \right] + \tilde{T}_{\ell}^{ij}(k, k') \\ &= \delta^{kk'} \mathcal{S}_{\ell}^{ij}(k) + \tilde{T}_{\ell}^{ij}(k, k') \end{aligned} \quad (6.37)$$

where $\lambda_{\ell}^{ij}(k) \sim T_{\ell}^{ij}(k, k)$, and $\tilde{T}_{\ell}^{ij}(k, k')$ collects all the off-diagonal terms, and we defined

$$\mathcal{S}_{\ell}^{ij}(k) = \delta^{ij} + \lambda_{\ell}^{ij} + \mathcal{P}_{(s)}^{ij} \quad (6.38)$$

As it turns out, the off-diagonal terms of the matrix $T_{\ell}^{ij}(k, k')$ are indeed negligible in relation to the diagonal ones. In the top row of Fig. 10 we plot this normalized matrix $T_{\ell}^{norm}(k, k') = T_{\ell}^{ij}(k, k') / \sqrt{T_{\ell}^{ii}(k, k)T_{\ell}^{jj}(k', k')}$ in the case of a single tracer, for the multipoles $\ell = 10$ and $\ell = 40$. In the lower panel, we show the lines of the (normalized) mixing matrices with $\ell = 10, 20$, and 40 , for fixed values $k' = 0.03, 0.05, 0.07$ and $0.09 h \text{ Mpc}^{-1}$.

These two plots clearly show that the matrix elements fall off as we move away from the diagonal, and at a faster rate for higher multipoles ℓ . Note that the normalized mixing matrix is determined entirely by the functions H_{ℓ} and G_{ℓ} , and has no dependence at all on the shape of the power spectrum or on the number of tracers. Therefore, it should be a very good approximation to discard all non-diagonal terms of the matrix $T_{\ell}^{ij}(k, k')$ altogether, except for the lowest multipoles. Then we can find an expression for the inverse of the covariance in terms of a series expansion around $\tilde{T}_{\ell}^{ij}(k, k')$.

We can now proceed to computing the inverse of this data covariance in harmonic Fourier space. First, notice that:

$$(1 + \mathcal{P} + \lambda)^{-1} = (1 + \mathcal{P})^{-1} \left[1 + (1 + \mathcal{P})^{-1}\lambda \right]^{-1}$$

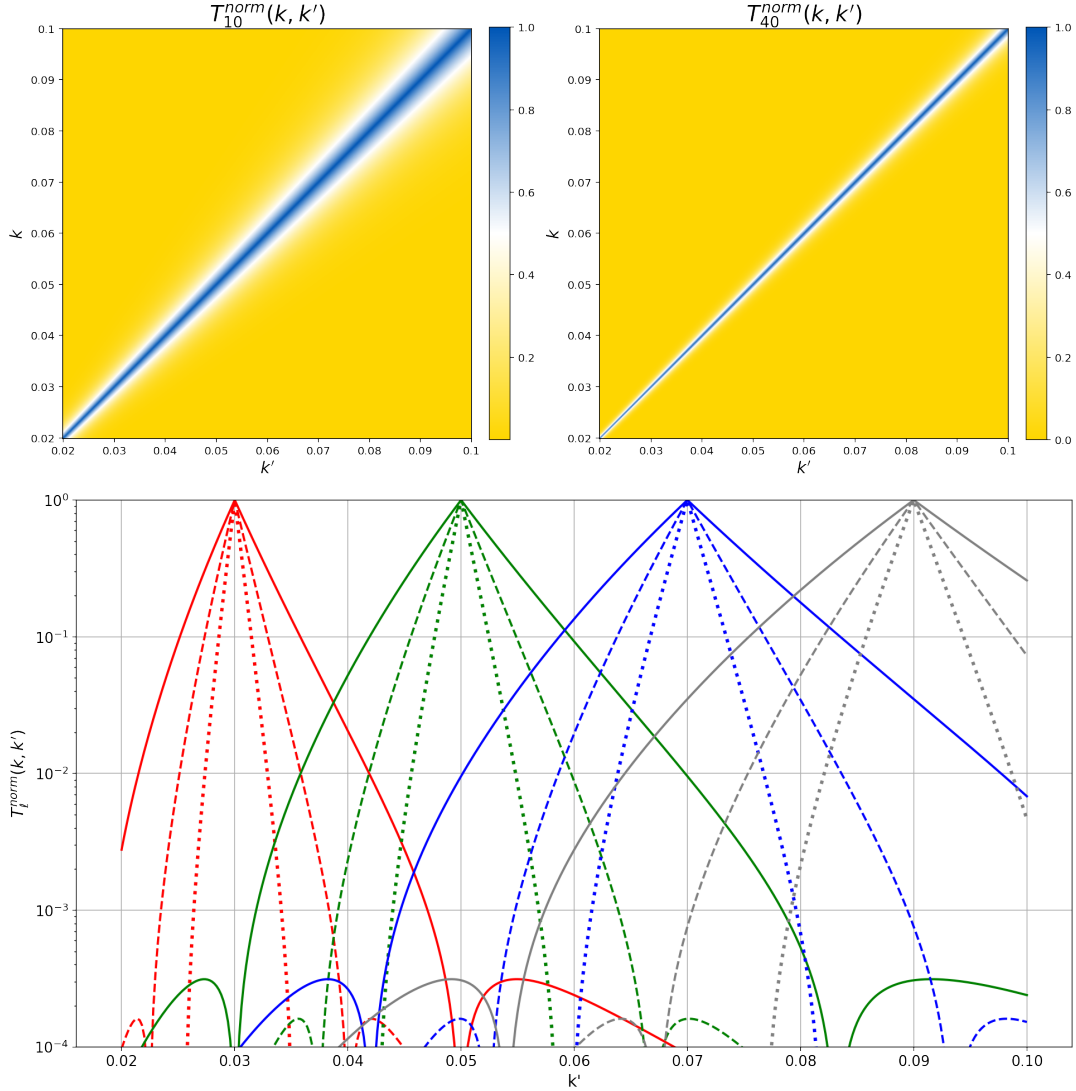


Figure 10: The radial (Fourier) mixing matrix $T_\ell^{ij}(k, k')$, normalized by its diagonal, $T_\ell^{norm}(k, k') = T_\ell^{ij}(k, k') / \sqrt{T_\ell^{ii}(k, k) T_\ell^{jj}(k', k')}$, for visualization purposes. Top: mixing matrix for $\ell = 10$ (left) and $\ell = 40$ (right). Bottom: rows of the normalized mixing matrices, with fixed $k' = 0.03, 0.05, 0.07,$ and 0.09 Mpc^{-1} (from left to right, respectively). Notice that for $k = k'$ the normalized mixing matrix is always equal to 1. The different lines corresponds to the multipoles $\ell = 10$ (solid), 20 (dashed) and 40 (dotted).

In a similar way to what was done in real space, we can use the Woodbury identity, which leads to:

$$\left[\delta^{ij} + \mathcal{P}_{(s)}^{ij} \right]^{-1} = \delta^{ij} - \frac{\mathcal{P}_{(s)}^{ij}}{1 + \mathcal{P}_{(s)}} \quad (6.39)$$

where $\mathcal{P}_{(s)} = \sum_i \mathcal{P}_{(s)}^{ii}$. The inverse of \mathcal{S}_ℓ^{ij} is therefore given by:

$$[\mathcal{S}_\ell^{ij}]^{-1} = \sum_l \left[\delta^{il} - \frac{\mathcal{P}_{(s)}^{il}}{1 + \mathcal{P}_{(s)}} \right] \left[\delta^{lj} + \sum_n \left(\delta^{ln} - \frac{\mathcal{P}_{(s)}^{ln}}{1 + \mathcal{P}_{(s)}} \right) \lambda_\ell^{nj} \right]^{-1} \quad (6.40)$$

Let's now go back to Eq. (6.37) to check that the redshift-space corrections λ_ℓ^{ij} are quite small. In practical applications, we use discrete Fourier bins of width Δk , so in this case we have $\delta(k - k') \rightarrow \delta_{kk'}^K / \Delta k$, where $\delta_{kk'}^K$ is the Kronecker delta for the bins. The covariance in Fourier space is then given by:

$$\tilde{\Sigma}_\ell^{ij}(k, k') \rightarrow \frac{\pi}{2} \frac{\delta_{kk'}^K}{k^2 \Delta k} \mathcal{S}_\ell^{ij}(k) + \tilde{T}_\ell^{ij}(k, k'), \quad (6.41)$$

where now we identify:

$$\lambda_\ell^{ij}(k) \rightarrow \frac{2}{\pi} k^2 \Delta k T_\ell^{ij}(k, k). \quad (6.42)$$

These terms are therefore proportional to the Fourier bin width, while from Eq. (6.34) and the definitions of H_ℓ and G_ℓ , Eqs. (6.17) and (6.25), it is straightforward to see that, up to factors of the redshift-space parameter β , we have $T_\ell^{ij}(k, k) \sim \alpha_\ell \mathcal{P}^{ij}(k) / k^3$, where α_ℓ is a coefficient that could be very large for larger values of ℓ . This means that the term $\lambda_\ell^{ij} \sim \alpha_\ell \mathcal{P}^{ij}(k) \Delta k / k$ is not negligible in general. If in a very particular case λ_ℓ^{ij} happens to be small compared with δ^{ij} and \mathcal{P}_s^{ij} [see Eq. (6.38)], then the inverse of the matrix \mathcal{S}_ℓ^{ij} , Eq. (6.40), could be written in terms of a series expansion:

$$[\mathcal{S}_\ell^{ij}]^{-1} \approx \delta^{ij} - \frac{\mathcal{P}_{(s)}^{ij}}{1 + \mathcal{P}_{(s)}} - \sum_{l,n} \left(\delta^{il} - \frac{\mathcal{P}_{(s)}^{il}}{1 + \mathcal{P}_{(s)}} \right) \lambda_\ell^{ln} \left(\delta^{nj} - \frac{\mathcal{P}_{(s)}^{nj}}{1 + \mathcal{P}_{(s)}} \right) + \mathcal{O}(\lambda^2) \quad (6.43)$$

In any case, $\mathcal{S}_\ell^{ij}(k)$ is a simple $N_t \times N_t$ object (where N_t is the number of tracers), which can easily be inverted numerically for each value of k and multipole ℓ . Once $[\mathcal{S}_\ell^{ij}(k)]^{-1}$ has been determined, we can finally obtain the inverse covariance. Since the off-diagonal matrix $\tilde{T}^{ij}(k, k')$ is small compared with the diagonal part of the data covariance of Eq. (6.37), we can invert that expression in terms of a series expansion around

$\tilde{T}^{ij}(k, k')$:

$$[\tilde{\Sigma}_\ell^{ij}(k, k')]^{-1} \approx \delta^{kk'} [\mathcal{S}_\ell^{ij}(k)]^{-1} - \sum_{n, n'} [\mathcal{S}_\ell^{in}(k)]^{-1} \tilde{T}_\ell^{nm'}(k, k') [\mathcal{S}_\ell^{n'j}(k')]^{-1} + \mathcal{O}(\tilde{T}^2) \quad (6.44)$$

As discussed above, it is often sufficient to consider only the first term in the expression above, neglecting the off-diagonal matrix \tilde{T}_ℓ^{ij} , which leads to corrections of order $\sim 10^{-2} - 10^{-3}$, depending on the multipole. This is a manifestation of the fact that the correlation function is much more diagonal in Fourier space than it is in real space, and is what ultimately allows us to invert the harmonic data covariance.

Remember that the inverse must be such that:

$$\begin{aligned} \sum_j \sum_{[k'']} [\tilde{\Sigma}_\ell^{ij}(k, k'')]^{-1} \tilde{\Sigma}_\ell^{ji'}(k'', k') &= \sum_j \frac{2}{\pi} \int_0^\infty dk'' k''^2 [\tilde{\Sigma}_\ell^{ij}(k, k'')]^{-1} \tilde{\Sigma}_\ell^{ji'}(k'', k') \\ &= \delta^{kk'} \delta_{ii'} = \frac{\pi}{2} \frac{\delta(k - k')}{kk'} \delta_{ii'} \end{aligned} \quad (6.45)$$

We check the accuracy of the inversion of the data covariance by considering the case of a single tracer, such that we have:

$$\mathcal{S}_\ell = 1 + \lambda_\ell + \mathcal{P}_{(s)} \quad , \quad \mathcal{S}_\ell^{-1} = \frac{1}{1 + \lambda_\ell + \mathcal{P}_{(s)}}$$

Note that, after discretizing the Dirac delta as a Kronecker delta, $\delta(k - k') \rightarrow \delta_{kk'}^K / \Delta k$, we can rewrite Eq. (6.45) as:

$$M_\ell(k, k') = \frac{2}{\pi} k k' \Delta k \sum_{[k'']} [\tilde{\Sigma}_\ell(k, k'')]^{-1} \tilde{\Sigma}_\ell(k'', k') = \delta_{kk'}^K \quad (6.46)$$

For simplicity, I now take constant $\beta = 1$, $\bar{n} = 10^{-3}$. Using a simple power spectrum from CAMB's fiducial cosmology, I plot in Figure 11 some rows of $M_\ell(k, k')$ for fixed $k = 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09$, with $k \in [0.02, 0.10]$ and a bin size of $\Delta k = 0.0002$. We can clearly see that the approximation works really well, especially for higher multipoles. Even better precision can be achieved by taking smaller bins, at the cost of a longer computing time, but only up to a certain point where the $\mathcal{O}^2(\tilde{T})$ correction in Eq. (6.44) becomes relevant. We found, however, that these corrections are completely

negligible, with the largest elements being at most of order $\sim 10^{-5}$ for $\ell \lesssim 100$ and scales $k \lesssim 1 h \text{ Mpc}^{-1}$.

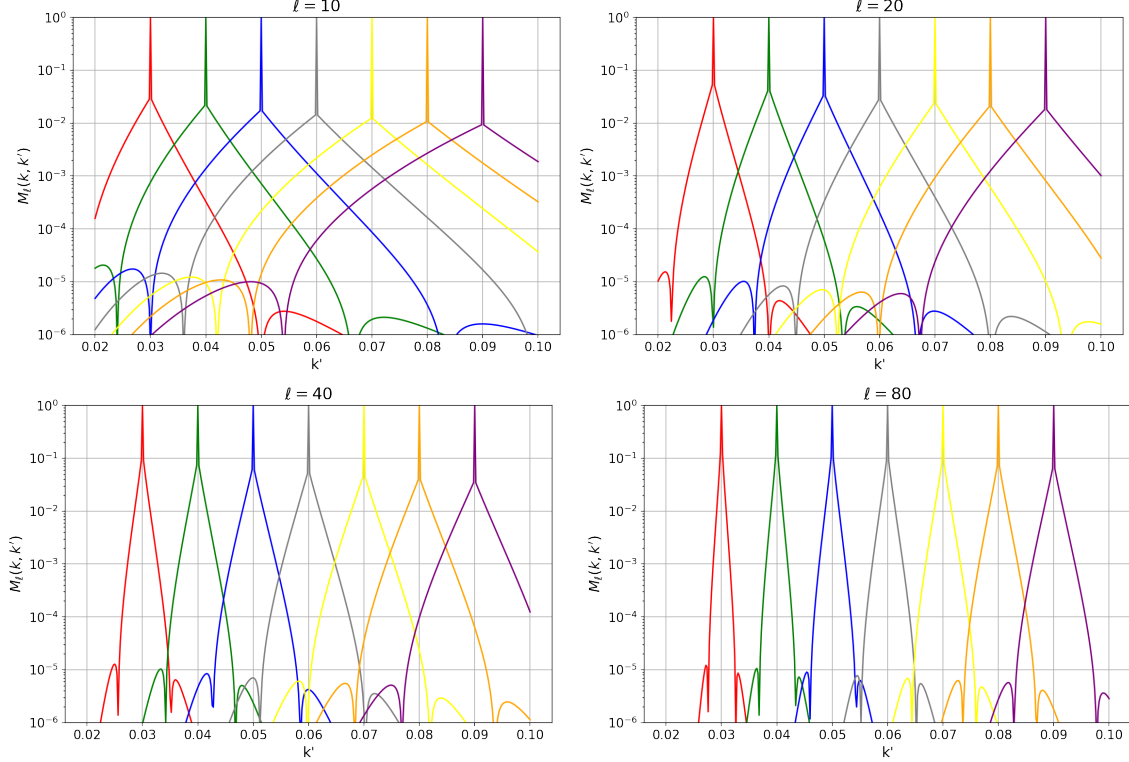


Figure 11: Rows of $M(k, k')$ for different ℓ with a bin size of $\Delta k = 0.0002$.

Hence, after reverting back to our original notations so that we can compare with the configuration space result, we finally find that the inverse of the harmonic-space data covariance is:

$$\begin{aligned}
 [\Sigma_{\ell, s}^{ij}(\bar{x}, \bar{y})]^{-1} &\approx \frac{1}{N^i(\bar{x})} \delta_{ij} \delta_{\bar{x}\bar{y}} & (6.47) \\
 &- \frac{2}{\pi} \int dk k^2 [j_\ell(k\bar{x}) - \beta^i j_\ell''(k\bar{x})] [j_\ell(k\bar{y}) - \beta^j j_\ell''(k\bar{y})] P^{ij}(k) [\mathcal{S}^{ij}(k)]^{-1} \\
 &+ \left(\frac{2}{\pi}\right)^2 \int dk k^2 \int dk' k'^2 [j_\ell(k\bar{x}) - \beta^i j_\ell''(k\bar{x})] [j_\ell(k'\bar{y}) - \beta^j j_\ell''(k'\bar{y})] \\
 &\times \sum_{i', j'} \sqrt{P^{ii}(k)} [\mathcal{S}^{ii}(k)]^{-1} \tilde{T}_\ell^{i'j'}(k, k') [\mathcal{S}^{j'j}(k')]^{-1} \sqrt{P^{jj}(k')} + \dots
 \end{aligned}$$

As a sanity check, notice that for $\beta^i \rightarrow 0$ we also get $T_\ell^{kk'} \rightarrow 0$, and we recover the real-space inverse, Eq. (5.40). Remember that T_ℓ^{ij} and S_ℓ^{ij} are given in terms of analytical

expressions — see Eqs. (6.34) and (6.38) —, so the inverse covariance above can be computed in terms of $3 \times N_t(N_t + 1)/2$ Fourier integrals involving Bessel functions, which can be handled by, e.g., FFTLog (see Appendix B for a basic description of the algorithm, and references of more efficient extensions). If a more accurate calculation is needed, we could drop the approximation used in Eq. (6.26) and then we would also compute another $3 \times N_t(N_t + 1)/2$ double integrals.

As discussed above, it is often a good approximation to neglect the matrix \tilde{T}_ℓ^{ij} , in which case the expressions above would simplify further, to:

$$[\tilde{\Sigma}_\ell^{ij}(k, k')]^{-1} \rightarrow \delta^{kk'} [\mathcal{S}_\ell^{ij}(k)]^{-1}, \quad (6.48)$$

and hence we can keep only the single Fourier integral in Eq. (6.47). Although this approximation is excellent for high multipoles ($\ell \gtrsim 10$), one may need to perform the double integral in Eq. (6.47) for very low multipoles.

6.3 Fisher and covariance matrices in redshift space

The inverse of the data covariance was the final missing piece of the puzzle, and we are finally in a position to return to the Fisher matrix. It is straightforward to see that all the expressions derived in Section 5 in real space still hold, and can be trivially generalized to redshift space. The Fisher matrix for the angular power spectra in redshift space is then given by an expression completely analogous to Eq. (5.48):

$$F[C_{\ell,s}^{[ij]}, C_{\ell,s}^{[i'j']}] = \frac{2\ell + 1}{4} (2 - \bar{\delta}_{ij})(2 - \bar{\delta}_{i'j'}) \left\{ [\Sigma_{\ell,s}^{ii'}]^{-1} [\Sigma_{\ell,s}^{jj'}]^{-1} + [\Sigma_{\ell,s}^{ij'}]^{-1} [\Sigma_{\ell,s}^{i'j}]^{-1} \right\} \quad (6.49)$$

where we used the notation introduced in Section 4 that connects the tracer index and the radial bin index implicitly, i.e., $\bar{\delta}_{ij} = \delta_{ij} \delta_{\bar{x}\bar{y}}$. The inversion of this expression is also identical to the case in real space, which leads to the harmonic covariance matrix in redshift space that generalizes Eq. (5.49):

$$\text{Cov}[C_{\ell,s}^{[ij]}, C_{\ell,s}^{[i'j']}] = \frac{1}{2\ell + 1} \left[\Sigma_{\ell,s}^{ii'} \Sigma_{\ell,s}^{jj'} + \Sigma_{\ell,s}^{ij'} \Sigma_{\ell,s}^{i'j} \right] \quad (6.50)$$

7 Conclusions and future prospects

An outstanding amount of data is set to be delivered by state of the art cosmological surveys in the upcoming years. But the way we have usually handled clustering data in the past with the Fourier power spectrum is not gonna cut it for these new generation surveys, because some crucial approximations for that method will not be realistic anymore. Hence, a more robust analysis that can properly take into account the curvature of the sky for large angular scales is necessary. This naturally leads us to a description of the matter distribution in terms of spherical harmonics, with the natural observable that contains all the information in the past lightcone being the angular power spectrum.

Motivated by the need to understand how much information we will actually be able to extract from these nearly full-sky surveys, we have derived an expression for the multi tracer Fisher matrix of the angular power spectrum, including cross-correlations, both in real and redshift spaces. In order to accomplish this, we found new analytical expressions for integrals involving products of spherical Bessel functions in terms of polynomials, which allowed us to find an analytical expression for the inverse of the harmonic data covariance, including redshift space distortions. This was the key step that in turn gave us semi-analytical, easy-to-compute, expressions for the Fisher and covariance matrices for the angular power spectrum, which are essentially exact, since the commonly used Limber and flat-sky approximations were not employed in this derivation. In fact, the only approximation we resorted to was a first order expansion for the inverse of the data covariance, and we showed that this approximation works incredibly well.

We now also have the foundation necessary to compute the optimal weights for the quadratic estimators that combine different tracers for a better estimate of the angular power spectrum. The procedure can be found in detail in Tegmark et al. (1998) or Abramo et al. (2016), and should be a straightforward calculation once the Fisher matrix is known. For this next step, it is important to note that in this derivation, we have assumed a full-sky coverage, which is not realistic because at least a portion of the sky will always be masked by our own galaxy. Therefore, before we can apply these estimators to real data sets and compare with others, we will have to take into account the partial sky coverage to generalize our results in a similar way to what is done for the pseudo- C_ℓ estimator described in Section 3.

Additionally, here we limited our calculations to the linear regime of redshift space distortions, as described in Section 4. The angular power spectrum, however, requires the use of narrow redshift bins to fully encapsulate wide angle effects and redshift evolution,

and not wash out the BAO and RSD information. The increased precision on photo- z measurements and higher tracer densities from spectroscopic surveys will allow us to use smaller bins and increase the RSD signal, but when they are comparable to deviations caused by peculiar velocities (typically of order $750 \text{ km s}^{-1} \rightarrow \Delta z \approx 0.005$) the non-linear RSD effects become important and have to be taken into account.

Nevertheless, this work represents an important step towards a unified formalism for studying the large scale structure with a spherical harmonic description, suitable for the overwhelming amount of data coming our way this decade. Much work has yet to be done to fully take advantage of the incredible collaborative efforts in mapping the universe with such precision. In our never ending quest to understand the cosmos, only one thing is certain: with more knowledge comes more questions.

Appendices

A Derivation of Eq. (4.20)

For simplicity, we split the harmonic correlation, Eq. (4.17), into four terms and compute each one separately:

$$\begin{aligned}
\langle a_{\ell m}(s) a_{\ell' m'}^*(s') \rangle &= \int d\Omega_s d\Omega_{s'} Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}') \\
&\times \left\langle \left[1 + f(z) \frac{d^2}{dr^2} \vec{\nabla}^{-2} \right] \delta(\vec{r}) \left[1 + f(z') \frac{d^2}{dr'^2} \vec{\nabla}^{-2} \right] \delta^*(\vec{r}') \right\rangle \quad (\text{A.1}) \\
&= f_1(s, s') + f_2(s, s') + f_3(s, s') + f_4(s, s')
\end{aligned}$$

The first term is just the correlation in real space:

$$\begin{aligned}
f_1(s, s') &= \langle a_{\ell m}^*(s) a_{\ell' m'}(s') \rangle \\
&= \int d\Omega_s d\Omega_{s'} Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}') \langle \delta(\vec{r}) \delta^*(\vec{r}') \rangle \\
&= \int d\Omega_s d\Omega_{s'} Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}') \left\langle \int \frac{d^3 k}{(2\pi)^3} e^{-i\vec{k}\vec{r}} \tilde{\delta}(\vec{k}) \int \frac{d^3 k'}{(2\pi)^3} e^{i\vec{k}'\vec{r}'} \tilde{\delta}^*(\vec{k}') \right\rangle \\
&= \int \frac{d^3 k}{(2\pi)^3} P(\vec{k}) \int d\Omega_s d\Omega_{s'} Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}') e^{-i\vec{k}(\vec{r}-\vec{r}')} \quad (\text{A.2})
\end{aligned}$$

where in the last line we used the definition of the Fourier power spectrum, Eq. (3.5). To move forward, we use the expansion in plane waves to rewrite the exponentials:

$$e^{i\vec{k}\vec{r}} = \sum_{\ell} (2\ell + 1) i^{\ell} j_{\ell}(kr) P_{\ell}(\hat{k} \cdot \hat{r}) \quad (\text{A.3})$$

where the P_{ℓ} are the Legendre polynomials. With the addition theorem for the spherical

harmonics, we can write the equation above as:

$$e^{i\vec{k}\vec{r}} = 4\pi \sum_{\ell m} i^\ell j_\ell(kr) Y_{\ell m}(\hat{k}) Y_{\ell m}^*(\hat{r}) \quad (\text{A.4})$$

This is known as the Rayleigh formula, and we can use this result together with the orthonormality of the spherical harmonics, $\int d\Omega_s Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}) = \delta_{\ell\ell'} \delta_{mm'}$, to write f_1 as:

$$\begin{aligned} f_1(s, s') &= \int \frac{d^3k}{(2\pi)^3} P(\vec{k}) \int d\Omega_s d\Omega_{s'} Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}') \\ &\quad \times 4\pi \sum_{l_1 m_1} (-i)^{l_1} j_{l_1}(kr) Y_{l_1 m_1}^*(\hat{k}) Y_{l_1 m_1}(\hat{s}) \\ &\quad \times 4\pi \sum_{l_2 m_2} i^{l_2} j_{l_2}(kr') Y_{l_2 m_2}(\hat{k}) Y_{l_2 m_2}^*(\hat{s}') \quad (\text{A.5}) \\ &= \frac{2}{\pi} \int dk k^2 \int d\Omega_k P(k) i^{\ell-\ell'} j_\ell(kr) j_{\ell'}(kr') Y_{\ell m}(\hat{k}) Y_{\ell' m'}^*(\hat{k}) \\ &= \frac{2}{\pi} \int dk k^2 j_\ell(kr) j_{\ell'}(kr') P(k) \delta_{\ell\ell'} \delta_{mm'} \end{aligned}$$

Hence, the real space part of the harmonic correlation is:

$$f_1(s, s') = C_\ell^r(r, r') \delta_{\ell\ell'} \delta_{mm'} \quad (\text{A.6})$$

where the real space angular power is given in terms of the Fourier spectrum through the following relation:

$$C_\ell^r(r, r') \equiv \frac{2}{\pi} \int_0^\infty dk k^2 j_\ell(kr) j_\ell(kr') P(k) \quad (\text{A.7})$$

This expression is found all over the literature and is enough for the flat sky limit, since all the other extra RSD terms in this approximation are just multiplicative constants. But for a complete description of RSDs on the sphere, we must compute the other terms in Eq. (A.1). Let's now see how to deal with the last one, then the remaining two will be

trivial:

$$f_4(s, s') = \int d\Omega_s d\Omega_{s'} Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}') \left\langle \frac{d^2}{dr^2} \vec{\nabla}^{-2} \delta(\vec{r}) \frac{d^2}{dr'^2} \vec{\nabla}'^{-2} \delta^*(\vec{r}') \right\rangle f(z) f(z') \quad (\text{A.8})$$

Here it is once again convenient to work in Fourier space, where $e^{i(\vec{k}\vec{r}-\vec{k}'\vec{r}')}$ we have :

$$\begin{aligned} f_4(s, s') &= f(z) f(z') \int d\Omega_s d\Omega_{s'} Y_{\ell m}^*(\hat{s}) Y_{\ell' m'}(\hat{s}') \\ &\times \int \frac{d^3 k d^3 k'}{(2\pi)^6} \frac{d^2}{dr^2} \frac{d^2}{dr'^2} \vec{\nabla}^{-2} \vec{\nabla}'^{-2} e^{i(\vec{k}\vec{r}-\vec{k}'\vec{r}')} \langle \delta(\vec{k}) \delta^*(\vec{k}') \rangle \end{aligned} \quad (\text{A.9})$$

Using the definition of the power spectrum and the Rayleigh formula once again, Eq. (A.4), and making the derivative now be taken with respect to the argument, $j_\ell''(kr) = \frac{d^2 j_\ell(kr)}{d(kr)^2}$, we get:

$$\begin{aligned} f_4(s, s') &= \sum_{\ell_1 m_1} \sum_{\ell_2 m_2} i^{\ell_1 - \ell_2} \int d\Omega_s Y_{\ell m}^*(\hat{s}) Y_{\ell_1 m_1}(\hat{s}) \int d\Omega_{s'} Y_{\ell' m'}(\hat{s}') Y_{\ell_2 m_2}^*(\hat{s}') \\ &\times \frac{2}{\pi} f(z) f(z') \int \Omega_k Y_{\ell_1 m_1}^*(\hat{k}) Y_{\ell_2 m_2}(\hat{k}) \int dk k^2 P(k, z) j_{\ell_1}''(kr) j_{\ell_2}''(kr') \\ &= \frac{2}{\pi} f(z) f(z') \sum_{\ell_1 m_1} \sum_{\ell_2 m_2} i^{\ell_1 - \ell_2} \delta_{\ell \ell_1} \delta_{m m_1} \delta_{\ell' \ell_2} \delta_{m' m_2} \delta_{\ell_1 \ell_2} \delta_{m_1 m_2} \\ &\times \int dk k^2 P(k, z) j_{\ell_1}''(kr) j_{\ell_2}''(kr') \end{aligned} \quad (\text{A.10})$$

After summing over all the indices, we finally get:

$$f_4(s, s') = \frac{2}{\pi} f(z) f(z') \int dk k^2 j_\ell''(kr) j_{\ell'}''(kr') P(k) \delta_{\ell \ell'} \delta_{m m'} \quad (\text{A.11})$$

The derivation of the second and third terms is exactly the same, except with only one derivative, and they are related to each other by simply interchanging their arguments. Hence, in full analogy to the derivation of f_4 , it is easy to show that:

$$f_2(s, s') = -\frac{2f(z)}{\pi} \int dk k^2 j_\ell''(kr) j_\ell(kr') P(k) \delta_{\ell\ell'} \delta_{mm'} \quad (\text{A.12})$$

$$f_3(s, s') = -\frac{2f(z')}{\pi} \int dk k^2 j_\ell(kr) j_\ell''(kr') P(k) \delta_{\ell\ell'} \delta_{mm'} \quad (\text{A.13})$$

Putting Eqs. (A.6) and (A.11-A.13) together, we finally get

$$\begin{aligned} \langle a_{\ell m}(s) a_{\ell' m'}^*(s') \rangle &= \frac{2}{\pi} \int dk k^2 [j_\ell(kr) - f(z) j_\ell''(kr)] \\ &\quad \times [j_\ell(kr') - f(z') j_\ell''(kr')] P(k) \delta_{\ell\ell'} \delta_{mm'} \end{aligned} \quad (\text{A.14})$$

as quoted in the main text. Then Eq. (4.20) follows trivially from the definition of the angular power spectrum as the 2-point correlation function of the harmonic coefficients.

B FFTLog

Here, we briefly describe the FFTLog algorithm, widely used to quickly compute integrals involving Bessel functions. We start by discussing the context in which this is a useful tool. First, remember that the Hankel transform is an integral transformation whose kernel is a Bessel function:

$$g(r) = \int_0^\infty dk f(k) (kr)^q J_\mu(kr) r \quad (\text{B.1})$$

It often shows up problems with cylindrical or spherical symmetries, when we write the Fourier transform in spherical coordinates, and in cosmology we often go back and forth between real and Fourier space. This transform most commonly appears in the form of Eq. (B.1). We usually describe large-scale clustering by the configuration-space 2-point correlation function, and its Fourier transform, the power spectrum, and those two quantities are related by a Hankel transform.

Therefore, we need to be able to perform these transformations quickly and with good accuracy. But doing so is not so trivial, first because the Bessel function is highly oscillatory, which means that integrating with any of the most common quadrature methods can give us very inaccurate results. Another problem is that we usually want to cover several orders of magnitude, so if we used something like the FFT algorithm, which requires linearly spaced points, we would have to deal with unnecessarily large matrices and just

waste resources.

These problems may be solved by using an FFT algorithm originally proposed by Talman (1978), for applications in atomic physics, and was implemented in cosmology by Hamilton (2000), who called it the FFTLog algorithm, since it is analogous to the regular FFT, but for sequences uniformly spaced and periodic in logarithmic space.

B.1 The algorithm

Consider a function $f(k)$ periodic in logarithmic space $\ln k$ with period L ,

$$f(ke^L) = f(k) \quad (\text{B.2})$$

This periodicity implies that the Fourier transform of $f(k)$ contains only discrete Fourier modes, $\exp(2\pi im \ln(k/k_0)/L)$, with integer m . Take the fundamental interval to be $[\ln k_0 - L/2, \ln k_0 + L/2]$ and suppose that $f(k)$ is smooth in the sense that it is some linear combination of the N lowest frequency modes,

$$f(k) = \sum_m c_m e^{2\pi im \ln(k/k_0)/L} \quad (\text{B.3})$$

The sampling theorem tells us the coefficients c_m are

$$c_m = \frac{1}{N} \sum_n f_n e^{-2\pi imn/N} \quad (\text{B.4})$$

where

$$f_n = f(k_n) = \sum_m c_m e^{2\pi imn/N} \quad (\text{B.5})$$

with $k_n = k_0 e^{nL/N}$.

The discrete Hankel transform of f_n is

$$\tilde{f}_n = \sum_m c_m u_m e^{-2\pi imn/N} \quad (\text{B.6})$$

with the coefficients

$$u_m(\mu, q) = (k_0 r_0)^{-2\pi i m / L} U_\mu \left(q + \frac{2\pi i m}{L} \right) \quad (\text{B.7})$$

and the complex valued function

$$U_\mu(x) = \int_0^\infty t^x J_\mu(t) dt = 2^x \frac{\Gamma[(\mu + 1 + x)/2]}{\Gamma[(\mu + 1 - x)/2]} \quad (\text{B.8})$$

The fundamental aspect of the algorithm lies in this function, U_μ , which is the Mellin transform of the Bessel function of first kind, so it can be computed analytically in terms of Gamma functions.

What the FFTLog algorithm actually computes is the discrete Hankel transform of a sequence of logarithmically spaced points, and consists of basically three steps:

1. Fourier transform the f_n to obtain the coefficients c_m (B.4).
2. Multiply it by u_m (B.7).
3. Fourier transform $u_m c_m$ back to get the Hankel transform \tilde{f}_n .

This algorithm suffers from some of the same problems of the regular FFT. Usually we're interested in the discrete Hankel transform as an approximation to the continuous one, so we truncate the function to a finite logarithmic interval, then we replicate the truncated function in logarithmic space, and these steps cause ringing and aliasing, respectively. Therefore, appropriate precautions must be taken to reduce these unwanted effects.

References

TMC Abbott, M Agüena, A Alarcon, S Allam, O Alves, A Amon, F Andrade-Oliveira, J Annis, S Avila, D Bacon, et al. Dark energy survey year 3 results: cosmological constraints from galaxy clustering and weak lensing. *Physical Review D*, 105(2):023520, 2022.

L Raul Abramo. The full Fisher matrix for galaxy surveys. *Monthly Notices of the Royal Astronomical Society*, 420(3):2042–2057, 2012.

L Raul Abramo and Katie E Leonard. Why multitracer surveys beat cosmic variance. *Monthly Notices of the Royal Astronomical Society*, 432(1):318–326, 2013.

L Raul Abramo, Lucas F Secco, and Arthur Loureiro. Fourier analysis of multitracer cosmological surveys. *Monthly Notices of the Royal Astronomical Society*, 455(4):3871–3889, 2016.

L. Raul Abramo, Ian Tashiro, João Ferri, and Arthur Loureiro. Fisher matrix for the angular power spectrum of multi-tracer galaxy surveys, 2022. URL <https://arxiv.org/abs/2204.05057>.

Katherine Accetta, Conny Aerts, Victor Silva Aguirre, Romina Ahumada, Nikhil Ajgaonkar, N Filiz Ak, Shadab Alam, Carlos Allende Prieto, Andres Almeida, Friedrich Anders, et al. The Seventeenth Data Release of the Sloan Digital Sky Surveys: Complete Release of MaNGA, MaStar and APOGEE-2 Data. *arXiv preprint arXiv:2112.02026*, 2021.

Peter AR Ade, N Aghanim, MIR Alves, C Armitage-Caplan, M Arnaud, M Ashdown, F Atrio-Barandela, J Aumont, H Aussel, C Baccigalupi, et al. Planck 2013 results. I. Overview of products and scientific results. *Astronomy & Astrophysics*, 571:A1, 2014.

Shadab Alam, Metin Ata, Stephen Bailey, Florian Beutler, Dmitry Bizyaev, Jonathan A Blazek, Adam S Bolton, Joel R Brownstein, Angela Burden, Chia-Hsun Chuang, et al. The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: cosmological analysis of the dr12 galaxy sample. *Monthly Notices of the Royal Astronomical Society*, 470(3):2617–2652, 2017.

Shadab Alam, Marie Aubert, Santiago Avila, Christophe Balland, Julian E Bautista, Matthew A Bershad, Dmitry Bizyaev, Michael R Blanton, Adam S Bolton, Jo Bovy, et al. Completed sdss-iv extended baryon oscillation spectroscopic survey: Cosmological implications from two decades of spectroscopic surveys at the apache point observatory. *Physical Review D*, 103(8):083533, 2021.

Charles Alcock and Bohdan Paczyński. An evolution free test for non-zero cosmological constant. *Nature*, 281(5730):358–359, 1979.

Daniel Aloni, Asher Berlin, Melissa Joseph, Martin Schmaltz, and Neal Weiner. A step in understanding the hubble tension. *arXiv preprint arXiv:2111.00014*, 2021.

David Alonso. Science-driven 3D data compression. *Monthly Notices of the Royal Astronomical Society*, 473(4):4306–4317, 2018.

David Alonso, Javier Sanchez, and Anze Slosar. A unified pseudo- \mathcal{Cl} framework. *Monthly Notices of the Royal Astronomical Society*, 484(3):4127–4151, 2019.

Luca Amendola, Stephen Appleby, Anastasios Avgoustidis, David Bacon, Tessa Baker, Marco Baldi, Nicola Bartolo, Alain Blanchard, Camille Bonvin, Stefano Borgani, et al. Cosmology and fundamental physics with the Euclid satellite. *Living reviews in relativity*, 21(1):2, 2018.

F Andrade-Oliveira, H Camacho, L Faga, R Gomes, R Rosenfeld, A Troja, O Alves, C Doux, J Elvin-Poole, X Fang, et al. Galaxy Clustering in Harmonic Space from the Dark Energy Survey Year 1 Data: Compatibility with Real-Space Results. *Monthly Notices of the Royal Astronomical Society*, 2021.

A Arbey and F Mahmoudi. Dark matter and the early Universe: a review. *Progress in Particle and Nuclear Physics*, 119:103865, 2021.

Jacobo Asorey, Martin Crocce, Enrique Gaztanaga, and Antony Lewis. Recovering 3D clustering information with angular correlations. *Monthly Notices of the Royal Astronomical Society*, 427(3):1891–1902, 2012.

Marie Aubert, Marie-Claude Cousinou, Stéphanie Escoffier, Adam J Hawken, Seshadri Nadathur, Shadab Alam, Julian Bautista, Etienne Burtin, Arnaud de Mattia, Héctor Gil-Marín, et al. The completed sdss-iv extended baryon oscillation spectroscopic survey: Growth rate of structure measurement from cosmic voids. *arXiv preprint arXiv:2007.09013*, 2020.

Nicola Bellomo, José Luis Bernal, Giulio Scelfo, Alvisè Raccanelli, and Licia Verde. Beware of commonly used approximations I: errors in forecasts. *arXiv preprint arXiv:2005.10384*, 2020.

N Benitez, R Dupke, M Moles, L Sodre, J Cenarro, A Marin-Franch, K Taylor, D Cristobal, A Fernandez-Soto, C Mendes de Oliveira, et al. J-PAS: the Javalambre-physics of the accelerated universe astrophysical survey. *arXiv preprint arXiv:1403.5237*, 2014.

Charles L Bennett, Anthony J Banday, Krzysztof M Górski, G Hinshaw, P Jackson, P Keegstra, A Kogut, George F Smoot, David T Wilkinson, and Edward L Wright. Four-year coBE* dmr cosmic microwave background observations: maps and basic results. *The Astrophysical Journal*, 464(1):L1, 1996.

Gianfranco Bertone, Dan Hooper, and Joseph Silk. Particle dark matter: evidence, candidates and constraints. *Physics Reports*, 405(5–6):279–390, Jan 2005. ISSN 0370-1573. doi: 10.1016/j.physrep.2004.08.031. URL <http://dx.doi.org/10.1016/j.physrep.2004.08.031>.

Chris Blake, Ivan K Baldry, Joss Bland-Hawthorn, Leonidas Christodoulou, Matthew Colless, Christopher Conselice, Simon P Driver, Andrew M Hopkins, Jochen Liske, Jon

Loveday, et al. Galaxy And Mass Assembly (GAMA): improved cosmic growth measurements using multiple tracers of large-scale structure. *Monthly Notices of the Royal Astronomical Society*, 436(4):3089–3105, 2013.

Jolyon K Bloomfield, Stephen HP Face, and Zander Moss. Indefinite integrals of spherical bessel functions. *arXiv preprint arXiv:1703.06428*, 2017.

S. Bonoli et al. The miniJPAS survey: A preview of the Universe in 56 colors. *Astron. Astrophys.*, 653:A31, 2021. doi: 10.1051/0004-6361/202038841.

Stefano Camera, José Fonseca, Roy Maartens, and Mário G Santos. Optimized angular power spectra for spectroscopic galaxy surveys. *Monthly Notices of the Royal Astronomical Society*, 481(1):1251–1261, 2018.

J-E Campagne, J Neveu, and S Plaszczynski. Angpow: a software for the fast computation of accurate tomographic power spectra. *Astronomy & Astrophysics*, 602:A72, 2017.

Julien Carron. On the assumption of Gaussianity for cosmological two-point statistics and parameter dependent covariance matrices. *Astronomy & Astrophysics*, 551:A88, 2013.

Dark Energy Spectroscopic Instrument Collaboration, Amir Aghamousa, Jessica Aguilar, Steve Ahlen, Shadab Alam, Lori E Allen, Carlos Allende Prieto, James Annis, Stephen Bailey, Christophe Balland, et al. The DESI Experiment part I: Science, targeting, and survey design. 2018.

University of Chicago Lawrence Berkeley National Laboratory Cerro-Tololo Inter-American Observatory Dark Energy Survey Collaboration: Fermilab, University of Illinois at Urbana-Champaign and Brenna Flaugher. The dark energy survey. *International Journal of Modern Physics A*, 20(14):3121–3123, 2005.

Enea Di Dio, Francesco Montanari, Ruth Durrer, and Julien Lesgourgues. Cosmological parameter estimation with large scale structure observations. *Journal of Cosmology and Astroparticle Physics*, 2014(01):042, 2014.

Eleonora Di Valentino, Olga Mena, Supriya Pan, Luca Visinelli, Weiqiang Yang, Alessandro Melchiorri, David F Mota, Adam G Riess, and Joseph Silk. In the realm of the hubble tension – a review of solutions. *arXiv preprint arXiv:2103.01183*, 2021.

Cyrille Doux, C Chang, B Jain, J Blazek, H Camacho, X Fang, M Gatti, E Krause, N MacCrann, S Samuroff, et al. Consistency of cosmic shear analyses in harmonic and real space. *Monthly Notices of the Royal Astronomical Society*, 503(3):3796–3817, 2021.

Richard Ellis, Olivier Le Fevre, Masashi Chiba, Masahiro Takada, Hiroaki Aihara, James Gunn, Timothy Heckman, Hideki Takami, Michael Seiffert, Hitoshi Murayama, et al. Extragalactic science and cosmology with the SUBARU Prime Focus Spectrograph (PFS). Technical report, 2012.

Xiao Fang, Tim Eifler, and Elisabeth Krause. 2D-FFTLog: efficient computation of real-space covariance matrices for galaxy clustering and weak lensing. *Monthly Notices of the Royal Astronomical Society*, 497(3):2699–2714, 2020a.

-
- Xiao Fang, Elisabeth Krause, Tim Eifler, and Niall MacCrann. Beyond Limber: Efficient computation of angular power spectra for galaxy clustering and weak lensing. *Journal of Cosmology and Astroparticle Physics*, 2020(05):010, 2020b.
- Hume A Feldman, Nick Kaiser, and John A Peacock. Power spectrum analysis of three-dimensional redshift surveys. *arXiv preprint astro-ph/9304022*, 1993.
- Ronald A Fisher. The logic of inductive inference. *Journal of the royal statistical society*, 98(1):39–82, 1935.
- Fisher, Scharf and Lahav. A spherical harmonic approach to redshift distortion and a measurement of $\Omega(0)$ from the 1.2-Jy IRAS Redshift Survey. *Monthly Notices of the Royal Astronomical Society*, 266:219, January 1994. doi: 10.1093/mnras/266.1.219.
- Henry S Grasshorn Gebhardt and Olivier Doré. Superfab: a fabulous code for spherical fourier-bessel decomposition. *arXiv preprint arXiv:2102.10079*, 2021.
- Henry S Grasshorn Gebhardt and Donghui Jeong. 2-FAST: Fast and accurate computation of projected two-point functions. *arXiv preprint arXiv:1709.02401*, 2017.
- Héctor Gil-Marín, Will J Percival, Licia Verde, Joel R Brownstein, Chia-Hsun Chuang, Francisco-Shu Kitaura, Sergio A Rodríguez-Torres, and Matthew D Olmstead. The clustering of galaxies in the SDSS-III Baryon Oscillation Spectroscopic Survey: RSD measurement from the power spectrum and bispectrum of the DR12 BOSS galaxies. *Monthly Notices of the Royal Astronomical Society*, page stw2679, 2016.
- Krzysztof M Gorski, Benjamin D Wandelt, Frode K Hansen, Eric Hivon, and Anthony J Banday. The healpix primer. *arXiv preprint astro-ph/9905275*, 1999.
- Krzysztof M Gorski, Eric Hivon, Anthony J Banday, Benjamin D Wandelt, Frode K Hansen, Mstvos Reinecke, and Matthia Bartelmann. Healpix: A framework for high-resolution discretization and fast analysis of data distributed on the sphere. *The Astrophysical Journal*, 622(2):759, 2005.
- Henry S. Grasshorn Gebhardt and Donghui Jeong. Nonlinear redshift-space distortions in the harmonic-space galaxy power spectrum. *Physical Review D*, 102(8):083521, October 2020. doi: 10.1103/PhysRevD.102.083521.
- AJS Hamilton. Linear redshift distortions: A Review. *The evolving universe*, pages 185–275, 1998.
- AJS Hamilton. Uncorrelated modes of the non-linear power spectrum. *Monthly Notices of the Royal Astronomical Society*, 312(2):257–284, 2000.
- AJS Hamilton and M Culhane. Spherical redshift distortions. *arXiv preprint astro-ph/9507021*, 1995.
- Samira Hamimeche and Antony Lewis. Likelihood analysis of CMB temperature and polarization power spectra. *Physical Review D*, 77(10):103013, 2008.

Brett Hayes and Robert Brunner. The SDSS DR7 galaxy angular power spectrum: volume limits and galaxy morphology. *Monthly Notices of the Royal Astronomical Society*, 428(4):3487–3496, 2013.

Brett Hayes, Robert Brunner, and Ashley Ross. The sdss dr7 galaxy angular power spectrum. *arXiv preprint arXiv:1112.5723*, 2011.

Alan Heavens. Statistical techniques in cosmology. *arXiv preprint arXiv:0906.0664*, 2009.

Alan F Heavens, Elena Sellentin, and Andrew H Jaffe. Extreme data compression while searching for new physics. *Monthly Notices of the Royal Astronomical Society*, 498(3):3440–3451, 2020.

César Hernández-Aguayo, Jiamin Hou, Baojiu Li, Carlton M Baugh, and Ariel G Sánchez. Large-scale redshift space distortions in modified gravity theories. *Monthly Notices of the Royal Astronomical Society*, 485(2):2194–2213, 2019.

Eric Hivon, Krzysztof M Górski, C Barth Netterfield, Brendan P Crill, Simon Prunet, and Frode Hansen. Master of the cmb anisotropy power spectrum: a fast method for statistical analysis of large and complex cmb data sets. *arXiv preprint astro-ph/0105302*, 2001.

Wayne Hu and Bhuvnesh Jain. Joint galaxy-lensing observables and the dark energy. *Physical Review D*, 70(4):043009, August 2004. doi: 10.1103/PhysRevD.70.043009.

Željko Ivezić, Steven M Kahn, J Anthony Tyson, Bob Abel, Emily Acosta, Robyn Allsman, David Alonso, Yusra AlSayyad, Scott F Anderson, John Andrew, et al. LSST: from science drivers to reference design and anticipated data products. *The Astrophysical Journal*, 873(2):111, 2019.

Nick Kaiser. On the spatial correlations of abell clusters. 1984.

Nick Kaiser. Clustering in real space and in redshift space. *Monthly Notices of the Royal Astronomical Society*, 227(1):1–21, 1987.

B Leistedt, A Rassat, Alexandre Refregier, and J-L Starck. 3dex: a code for fast spherical fourier-bessel decomposition of 3d surveys. *Astronomy & Astrophysics*, 540:A60, 2012.

Boris Leistedt, Hiranya V Peiris, Daniel J Mortlock, Aurélien Benoit-Lévy, and Andrew Pontzen. Estimating the large-scale angular power spectrum in the presence of systematics: a case study of Sloan Digital Sky Survey quasars. *Monthly Notices of the Royal Astronomical Society*, 435(3):1857–1873, 2013.

Michael E Levi, Lori E Allen, Anand Raichoor, Charles Baltay, Segev BenZvi, Florian Beutler, Adam Bolton, Francisco J Castander, Chia-Hsun Chuang, Andrew Cooper, et al. The dark energy spectroscopic instrument (DESI). *arXiv preprint arXiv:1907.10688*, 2019.

Arthur Loureiro, Bruno Moraes, Filipe B Abdalla, Andrei Cuceu, Michael McLeod, Lorne Whiteway, Sreekumar T Balan, Aurélien Benoit-Lévy, Ofer Lahav, Marc Manera, et al. Cosmological measurements from angular power spectra analysis of BOSS DR12 tomography. *Monthly Notices of the Royal Astronomical Society*, 485(1):326–355, 2019.

Marilena LoVerde and Niayesh Afshordi. Extended limber approximation. *Physical Review D*, 78(12):123506, 2008.

RG Mann, JA Peacock, and AF Heavens. Eulerian bias and the galaxy density field. *Monthly Notices of the Royal Astronomical Society*, 293(3):209–221, 1998.

Takahiko Matsubara. The correlation function in redshift space: General formula with wide-angle effects and cosmological distortions. *The Astrophysical Journal*, 535(1):1, 2000.

A Moneti, HJ Mccracken, M Shuntov, OB Kauffmann, P Capak, I Davidzon, O Ilbert, C Scarlata, S Toft, J Weaver, et al. Euclid preparation: Xvii. cosmic dawn survey: Spitzer space telescope observations of the euclid deep fields and calibration fields. *Astronomy & Astrophysics*, 658, 2022.

Antonio D Montero-Dorta, L Raul Abramo, Benjamin R Granett, Sylvain de la Torre, and Luigi Guzzo. The multi-tracer optimal estimator applied to VIPERS. *Monthly Notices of the Royal Astronomical Society*, 493(4):5257–5272, 2020.

Seshadri Nadathur, Paul M Carter, Will J Percival, Hans A Winther, and Julian E Bautista. Beyond bao: Improving cosmological constraints from boss data with measurement of the void-galaxy cross-correlation. *Physical Review D*, 100(2):023504, 2019.

Nikhil Padmanabhan, David J. Schlegel, Uroš Seljak, Alexey Makarov, Neta A. Bahcall, Michael R. Blanton, Jonathan Brinkmann, Daniel J. Eisenstein, Douglas P. Finkbeiner, James E. Gunn, David W. Hogg, Željko Ivezić, Gillian R. Knapp, Jon Loveday, Robert H. Lupton, Robert C. Nichol, Donald P. Schneider, Michael A. Strauss, Max Tegmark, and Donald G. York. The clustering of luminous red galaxies in the Sloan Digital Sky Survey imaging data. *Monthly Notices of the Royal Astronomical Society*, 378(3):852–872, July 2007. doi: 10.1111/j.1365-2966.2007.11593.x.

J. A. Peacock and S. J. Dodds. Non-linear evolution of cosmological power spectra. *Monthly Notices of the Royal Astronomical Society*, 280(3):L19–L26, Jun 1996. ISSN 1365-2966. doi: 10.1093/mnras/280.3.L19. URL <http://dx.doi.org/10.1093/mnras/280.3.L19>.

Phillip James Edwin Peebles. *The large-scale structure of the universe*, volume 98. Princeton university press, 2020.

Will J Percival, Licia Verde, and John A Peacock. Fourier analysis of luminosity-dependent galaxy clustering. *Monthly Notices of the Royal Astronomical Society*, 347(2):645–653, 2004.

Will J Percival, Lado Samushia, Ashley J Ross, Charles Shapiro, and Alvis Raccanelli. Redshift-space distortions. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 369(1957):5058–5067, 2011.

Saul Perlmutter, Goldhaber Aldering, Gerson Goldhaber, RA Knop, Peter Nugent, Patricia G Castro, Susana Deustua, Sebastien Fabbro, Ariel Goobar, Donald E Groom, et al. Measurements of ω and λ from 42 high-redshift supernovae. *The Astrophysical Journal*, 517(2):565, 1999.

Patrick Peter and Jean-Philippe Uzan. *Primordial cosmology*. Oxford University Press, 2009.

Marco Raveri. Is there concordance within the concordance Λ CDM model? *arXiv preprint arXiv:1510.00688*, 2015.

Martin Reinecke and Dag Sverre Seljebotn. Libsharp—spherical harmonic transforms revisited. *Astronomy & Astrophysics*, 554:A112, 2013.

Adam G Riess, Alexei V Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M Garnavich, Ron L Gilliland, Craig J Hogan, Saurabh Jha, Robert P Kirshner, et al. Observational evidence from supernovae for an accelerating universe and a cosmological constant. *The Astronomical Journal*, 116(3):1009, 1998.

Ashley J Ross, Lado Samushia, Angela Burden, Will J Percival, Rita Tojeiro, Marc Manera, Florian Beutler, J Brinkmann, Joel R Brownstein, Aurelio Carnero, et al. The clustering of galaxies in the SDSS-III DR10 Baryon Oscillation Spectroscopic Survey: no detectable colour dependence of distance scale or growth rate measurements. *Monthly Notices of the Royal Astronomical Society*, 437(2):1109–1126, 2014.

Salvador Salazar-Albornoz, Ariel G Sánchez, Jan Niklas Grieb, Martin Crocce, Roman Scoccimarro, Shadab Alam, Florian Beutler, Joel R Brownstein, Chia-Hsun Chuang, Francisco-Shu Kitaura, et al. The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: angular clustering tomography and its cosmological implications. *Monthly Notices of the Royal Astronomical Society*, 468(3):2938–2956, 2017.

Siddharth Satpathy, Shadab Alam, Shirley Ho, Martin White, Neta A Bahcall, Florian Beutler, Joel R Brownstein, Chia-Hsun Chuang, Daniel J Eisenstein, Jan Niklas Grieb, et al. The clustering of galaxies in the completed sdss-iii baryon oscillation spectroscopic survey: On the measurement of growth rate using galaxy correlation functions. *Monthly Notices of the Royal Astronomical Society*, 469(2):1369–1382, 2017.

Roman Scoccimarro. Fast estimators for redshift-space clustering. *Physical Review D*, 92(8):083532, 2015.

Uroš Seljak. Analytic model for galaxy and dark matter clustering. *Monthly Notices of the Royal Astronomical Society*, 318(1):203–213, 2000.

Uroš Seljak. Extracting primordial non-gaussianity without cosmic variance. *Physical Review Letters*, 102(2):021302, 2009.

Uros Seljak and Matias Zaldarriaga. A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies. *The Astrophysical Journal*, 469:437, Oct 1996. ISSN 1538-4357. doi: 10.1086/177793. URL <http://dx.doi.org/10.1086/177793>.

Uros Seljak, Nico Hamaus, and Vincent Desjacques. How to suppress the shot noise in galaxy surveys. *Physical Review Letters*, 103(9):091303, 2009.

I Sevilla-Noarbe, K Bechtol, M Carrasco Kind, A Carnero Rosell, MR Becker, A Drlica-Wagner, RA Gruendl, ES Rykoff, E Sheldon, B Yanny, et al. Dark Energy Survey Year 3 Results: Photometric Data Set for Cosmology. *The Astrophysical Journal Supplement Series*, 254(2):24, 2021.

James D Talman. Numerical Fourier and Bessel transforms in logarithmic variables. *Journal of computational physics*, 29(1):35–48, 1978.

Konstantinos Tanidis and Stefano Camera. Developing a unified pipeline for large-scale structure data analysis with angular power spectra–I. The importance of redshift-space distortions for galaxy number counts. *Monthly Notices of the Royal Astronomical Society*, 489(3):3385–3402, 2019.

Konstantinos Tanidis and Stefano Camera. Developing a unified pipeline for large-scale structure data analysis with angular power spectra–III. Implementing the multitracers technique to constrain neutrino masses. *Monthly Notices of the Royal Astronomical Society*, 502(2):2952–2960, 2021.

Atsushi Taruya, Takahiro Nishimichi, and Shun Saito. Baryon acoustic oscillations in 2D: Modeling redshift-space power spectrum from perturbation theory. *Physical Review D*, 82(6):063522, 2010.

Max Tegmark, Andy N Taylor, and Alan F Heavens. Karhunen-Loeve eigenvalue problems in cosmology: How should we tackle large data sets? *The Astrophysical Journal*, 480(1):22, 1997.

Max Tegmark, Andrew JS Hamilton, Michael A Strauss, Michael S Vogeley, and Alexander S Szalay. Measuring the galaxy power spectrum with future redshift surveys. *The Astrophysical Journal*, 499(2):555, 1998.

Max Tegmark, Daniel J Eisenstein, Michael A Strauss, David H Weinberg, Michael R Blanton, Joshua A Frieman, Masataka Fukugita, James E Gunn, Andrew JS Hamilton, Gillian R Knapp, et al. Cosmological constraints from the sdss luminous red galaxies. *Physical Review D*, 74(12):123507, 2006.

Licia Verde, Alan F Heavens, Will J Percival, Sabino Matarrese, Carlton M Baugh, Joss Bland-Hawthorn, Terry Bridges, Russell Cannon, Shaun Cole, Matthew Colless, et al. The 2dF Galaxy Redshift Survey: the bias of galaxies and the density of the Universe. *Monthly Notices of the Royal Astronomical Society*, 335(2):432–440, 2002.

Michael S Vogeley and Alexander S Szalay. Eigenmode analysis of galaxy redshift surveys I. theory and methods. *arXiv preprint astro-ph/9601185*, 1996.

G. N. Watson. *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, Cambridge, England, 1944.

Glen Wright, FX Coudert, Martin Bentley, Sylvain Deville, and Graham Steel. This Study is Intentionally Left Blank-A systematic literature review of blank pages in academic publishing. 2014.

Henrique S Xavier, Filipe B Abdalla, and Benjamin Joachimi. Improving lognormal models for cosmological fields. *Monthly Notices of the Royal Astronomical Society*, 459(4):3693–3710, 2016.

Kazuhiro Yamamoto, Masashi Nakamichi, Akinari Kamino, Bruce A Bassett, and Hiroaki Nishioka. A measurement of the quadrupole power spectrum in the clustering of the 2df qso survey. *Publications of the Astronomical Society of Japan*, 58(1):93–102, 2006.

Gong-Bo Zhao, Yuting Wang, Atsushi Taruya, Weibing Zhang, Héctor Gil-Marín, Arnaud de Mattia, Ashley J Ross, Anand Raichoor, Cheng Zhao, Will J Percival, et al. The completed SDSS-IV extended Baryon Oscillation Spectroscopic Survey: a multitracer analysis in Fourier space for measuring the cosmic structure growth and expansion rate. *Monthly Notices of the Royal Astronomical Society*, 504(1):33–52, 2021.