



*Facultad
de
Ciencias*

**Theoretical derivation of the one- and
two-point PDFs of Line-Intensity Mapping
measurements**

(Derivación teórica de las PDFs a uno y dos puntos de
medidas de Line-Intensity Mapping)

Trabajo de Fin de Grado
para acceder al

GRADO EN FÍSICA

Autor: Alejandro González Ríos

Director: José Luis Bernal Mera

Septiembre-2025

Agradecimientos

Este trabajo supone la finalización de mi etapa como estudiante del Doble Grado de Matemáticas y Física. Han sido cinco años de mucho trabajo y esfuerzo, pero sin duda ha merecido totalmente la pena.

Me gustaría agradecer en primer lugar a mi tutor, Jose, toda su ayuda ofrecida (que no ha sido poca) para la realización de este trabajo y que pudiera llegar a completarse. No solo me ha explicado con claridad todos los conceptos nuevos necesarios, sino que también ha resuelto todas las dudas que me fueron surgiendo. Gracias a él, he podido profundizar en el campo de la cosmología y conocer cómo es trabajar en investigación científica.

Por otro lado, me siento un privilegiado por contar con el apoyo incondicional de mi familia. A mis padres, Marta y Alejandro, a mi hermana, María, y a mi novia, Carmen, quiero expresar mi más sincero agradecimiento por siempre creer en mí. Los cuatro sabéis que sin vosotros todo me hubiese resultado mucho más complicado.

Por último, no quiero olvidarme de ninguno de los amigos y amigas con los que he tenido el placer de compartir estos cinco años. Escribiría una lista de todos, tanto de quienes conocí en la Universidad como de quienes conozco desde antes, pero creo que entonces superaría la extensión máxima permitida en este documento. A todos, esta etapa no habría sido lo mismo sin vosotros, siempre recordaré los buenísimos momentos vividos juntos.

Abstract:

The aim of this work is to derive a theoretical formalism to compute the one- and two-point Probability Distribution Functions (PDFs) of the brightness temperature measured through Line-Intensity Mapping (LIM) experiments. We begin by introducing the conceptual framework, presenting the standard cosmological model together with the fundamentals of LIM and its significance in modern cosmology. We then describe in full detail the mathematical construction of the formalism. Focusing on the one-point PDF, we validate the formalism by comparison with the temperature distribution of matter realizations on some simple scenarios, finding very good agreement between them. This demonstrates the robustness of the theoretical formalism and sets the stage for extending the validation to the two-point PDF. The results we obtain show that this approach consistently probes the statistical properties of the LIM signal, opening the possibility of applying it to more complex and realistic scenarios.

Keywords: cosmology, dark-matter halos, power spectrum, Line-Intensity Mapping, galaxy clustering.

Resumen:

El objetivo de este trabajo es desarrollar un formalismo teórico para calcular la función de distribución de probabilidad (PDF) a uno y dos puntos de la temperatura de brillo medida mediante experimentos de Line-Intensity Mapping (LIM). Se comienza introduciendo el marco conceptual, presentando el modelo cosmológico estándar además de los fundamentos de LIM y su importancia en cosmología moderna. Posteriormente, se describe con todo detalle la construcción matemática del formalismo. Centrándose en la PDF a un punto, se valida el formalismo por comparación con la distribución de temperatura de realizaciones de materia en algunos casos simples, encontrando muy buena concordancia entre ellos. Esto demuestra la robustez del formalismo teórico y sienta las bases para extender la validación a la PDF a dos puntos. Los resultados obtenidos muestran que este enfoque explora de manera consistente las propiedades estadísticas de la señal de LIM, abriendo la posibilidad de aplicarlo a escenarios más complejos y realistas.

Palabras clave: cosmología, halos de materia oscura, espectro de potencias, Line-Intensity Mapping, agrupamiento de galaxias.

Contents

1	Introduction	1
1.1	Motivation	1
1.2	Standard Λ Cold Dark Matter cosmological model	2
1.3	Comoving coordinates and the FLRW metric	3
1.4	Distance measures	3
1.5	Large-scale structure	4
1.6	Line-Intensity Mapping	8
1.6.1	What is Line-Intensity Mapping? Its main target lines	8
1.6.2	LIM signal	10
1.6.3	Voxel Intensity Distribution	10
1.7	Ergodic hypothesis	11
2	Theoretical formalism	13
2.1	One-point probability distribution function	13
2.2	Two-point probability distribution function	18
2.2.1	One-halo term	19
2.2.2	Two-halo term	26
3	Computational implementation	31
3.1	Computational approach	31
3.2	VIDs from realizations	33
4	Validation of the theoretical formalism	39
4.1	Results for the VID	39
4.1.1	Poissonian VID validation	40
4.1.2	Smoothed Poissonian VID validation	41
4.1.3	Gaussian VID validation	43
4.1.4	Smoothed Gaussian VID validation	43
5	Conclusions and future work	47
	References	49

Chapter 1

Introduction

1.1 Motivation

The study of the Universe gives us the opportunity to test our understanding of fundamental physics, from which the union of theoretical and observational efforts has produced the standard cosmological model. However, much remains to be known, and the large-scale structure provides a very good testing ground for this. The best way to study the large-scale structure is by using various summary statistics, which are statistical descriptions of its properties. The most used summary statistics, because they are easy to calculate theoretically and measure robustly, are the two-point correlations such as the correlation function or the power spectrum, but they do not capture all the information. The Voxel Intensity Distribution (VID), which we use in this work, has much of this information. Here it appears Line-Intensity Mapping, which is emerging as a promising technique because it is not based on the resolution of individual galaxies but on the mapping of integrated spectral line emission. In this context, LIM provides access to epochs that are otherwise difficult to analyze with older methods.

The statistical description of intensity maps such as the ones LIM offers is essential for extracting cosmological information. Particularly, the one- and two-point Probability Distribution Functions of the LIM signal are useful to evidence the global properties of the temperature distribution. In this work, we develop a theoretical formalism to compute these functions. Furthermore, we validate this formalism for some specific and simplified cases, comparing the temperature distributions obtained from it with the ones obtained making Poissonian and Gaussian realizations on the sky.

This work is structured as follows. In Chapter 1 we introduce what is the Λ Cold Dark Matter cosmological model, the FLRW metric and how the Universe is structured at large scales. Moreover, we introduce what is LIM, which is the technique we based the work on, describing its main target lines and on what consists its measurable signal and defining what we understand as VID. Chapter 2 is the bulk of this work, where we describe with full detail the derivation of the one- and two-point PDF of the brightness temperature. In Chapter 3 we introduce some assumptions we made in order to simplify the computational implementation of the theoretical formalism and we explain how we made realizations of the sky so that we can validate the formalism for four different cases. In Chapter 4 we carry out this validation, making a comparison between the theoretical VID and the one obtained from realizations for the four cases analyzed. Finally, in Chapter 5 we make some conclusions of the work, introducing as well some future work to be done to continue verifying the validity of the formalism developed.

1.2 Standard Λ Cold Dark Matter cosmological model

The standard Λ Cold Dark Matter (Λ CDM) cosmological model is the theoretical model most favored by the data we have nowadays to describe the components of the Universe ([1], [2]), since it is the one which best reproduces the observations while being the simplest. It is defined by a set of assumptions, such as that general relativity is the correct theory that describes gravity on cosmological scales, and establishes that the Universe is dominated today by non-baryonic cold dark matter (CDM) and a cosmological constant, with initial perturbations generated by inflation in the very early Universe. In addition to CDM and the cosmological constant, the Universe also consists of radiation (photons, neutrinos at early times) and ordinary matter (baryons and leptons). This model of cosmology has become known as Λ CDM. First, the word “cold” in CDM comes from requiring the dark-matter particles to clump efficiently in the early Universe. Second, in relation to the cosmological constant, evidence from several sources suggests that there must be dark energy besides ordinary matter and radiation which, unlike dark matter, does not cluster strongly. The most believed option nowadays is that this form of energy remains constant with time, but there is still a lot of controversy over its value, because the value we can approximate making use of theory and vacuum fluctuations is much larger than the value required to explain cosmological observations. For this reason, the cosmological constant remains a puzzle for modern cosmology. Third, inflation is the mechanism that most likely generated the initial perturbations that grew into the structure observed today, and it consists of a brief epoch very early in the Universe during which it expanded very rapidly.

Besides, as explained in [3], the Cosmological Principle establishes that the Universe is isotropic and homogeneous. On the one hand, observations of several different cosmological probes show evidence that we observe the same properties looking at different directions on the sky. On the other hand, there is nothing to believe that we are located at a privileged point of the Universe. Therefore, the Universe must be isotropic around all observers in any point of the Universe. This means that, in addition to isotropic, the Universe must be homogeneous (invariant under translations). Regarding general relativity, time is also another coordinate. It is important to remark that the Cosmological Principle applies only for spatial coordinates. In fact, we have plenty of evidence that the Universe may evolve in time, but this is not in contradiction with the Cosmological Principle because the speed of light is finite. In this way, the information we obtain from distant objects corresponds to information in the past, since light has to travel until we observe it.

Isotropy and homogeneity are found by a set of observers at rest with respect to the cosmic fluid, which are referred to as fundamental observers. Cosmic fluid is a mathematical abstraction that represents all the energy and matter in the Universe. Note that, however, an observer with proper motion with respect to the cosmic fluid will find velocity-dependent observational effects, as for example a Doppler boost.

As we have explained, the Λ CDM cosmological model assumes general relativity as the theory that describes gravity and, particularly, spacetime. In this context, they are also assumed two postulates ([3]): the laws of physics take the same form in all inertial frames of reference, and the speed of light is finite and constant. As a consequence of the second postulate, if we consider an object with a relative velocity with respect to a given observer, the observed wavelength and the wavelength emitted by the object are related by a factor known as the redshift z , which is defined as

$$1 + z = \frac{\lambda_{\text{obs}}}{\lambda_{\text{emit}}} = \frac{\nu_{\text{emit}}}{\nu_{\text{obs}}}. \quad (1.1)$$

In this way, if the object is moving away from the observer, we have $\lambda_{\text{obs}} > \lambda_{\text{emit}}$ and the redshift is positive, but if the object is moving towards the observer, we have $\lambda_{\text{obs}} < \lambda_{\text{emit}}$ and the redshift is negative.

1.3 Comoving coordinates and the FLRW metric

Assuming a homogeneous and isotropic Universe, we can use the fundamental observers to define a cosmic time as a parameter with respect to which the spatial hypersurfaces that are homogeneous and isotropic for them evolve ([3]). In this way, we can use the position of the fundamental observers to define the spatial coordinate system and the direction of their four-velocities to define the temporal coordinate, so that this reference frame moves with the cosmic fluid because the fundamental observers are at rest with it. This set of coordinates are known as comoving coordinates.

The geometry of an isotropic and homogeneous, expanding spacetime is described by the Friedman-Lemaître-Robertson-Walker (FLRW) metric, which in spatial spherical coordinates is given by

$$ds^2 = c^2 dt^2 - a^2(t) \left[\frac{dr^2}{1 - \kappa r^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right], \quad (1.2)$$

where κ is the sign of the curvature of the Universe (spherical if $k = +1$, flat if $k = 0$, hyperbolic if $k = -1$) and $a(t)$ is the scale factor.

The scale factor relates the comoving distance given by the comoving coordinates with the actual physical distance between two objects in the expanding Universe. The scale factor is defined to be $a_0 = 1$ today, so that if the comoving distance today between two points is x_0 , the physical distance between them at some earlier time t was $a(t)x_0$ ([3]). Therefore, it can be used as an evolution coordinate because, for a monotonically expanding Universe, each value of the scale factor univocally corresponds to a moment in the evolution of the Universe. Likewise, the light that we would receive today from an emitter in our past lightcone is redshifted due to the expansion of the Universe. This is what is called the cosmological redshift.

It is easy to show ([4]) through the FLRW metric (1.2) that, since for radiation we have $ds^2 = 0$, the cosmological redshift can be expressed as

$$1 + z = \frac{a(t_0)}{a(t_e)} = \frac{1}{a(t_e)}, \quad (1.3)$$

where t_0 is the time for an observer at the present so that we have assumed $a(t_0) = 1$ and t_e is the time when a photon was emitted.

Note that in this work we assume flat Universe and the standard Λ CDM cosmology, with best fit parameter values from the full *Planck* analysis [5].

1.4 Distance measures

There are different cosmological distance measures we can deal with and give formulae for them, which are all defined in [6]. In last section, we introduced the comoving coordinates due to the fact that, since the Universe is in constant expansion, the distance between two objects is constantly changing. Although there are several distance measures that are function of time (or equivalently, redshift) and are crucial to derive the brightness temperature expression which we use as LIM signal as explained in section 1.6, in

this work we only focus on one, the comoving line-of-sight (or radial) distance, which is denoted by χ .

If we consider two nearby objects in the Universe, the small comoving line-of-sight distance $\delta\chi$ between them is the distance between them which remains constant with time if the objects are in rest with the cosmic fluid. In other words, since the cosmological redshift is defined by equation (1.3), it is the proper distance between the two objects multiplied by the factor $(1+z)$. Therefore, the total comoving line-of-sight distance between us and a distant object is computed by integrating the infinitesimal contributions between nearby objects along the radial ray from $z=0$ (us, the observer) to the object. It is explained in [6] that the comoving line-of-sight distance can be obtained by the following expression

$$\chi(z) = c \int_0^z \frac{dz'}{H(z')}, \quad (1.4)$$

where c is the speed of light and $H(z)$ is the Hubble parameter, defined by $H(z) = \dot{a}/a$ where $\dot{a}$ denotes the derivative of the scale factor with respect to cosmic time. It is common to write it as function of time instead of as function of redshift, but this is equivalent as demonstrated by equation (1.3).

1.5 Large-scale structure

In this section we make a brief description about how we can understand the Universe to be structured at large scales, as well as an introduction of what dark-matter halos are and the most relevant aspects for the statistical studies of clustering.

As discussed in [7], nowadays it has become clear that much of the mass in the Universe is dark, and that this mass was initially rather smoothly distributed. For this reason, we can think of the luminous galaxies we see today as biased tracers of the dark-matter distribution, which means that the relation between the number of galaxies in a randomly placed cell and the amount of dark matter the same cell contains may be complicated. Linear and higher order perturbation theory descriptions of gravitational clustering from Gaussian initial fluctuations have been developed, because there is evidence that the initial fluctuation field was very close to a Gaussian random field. These describe the evolution and slightly non-linear clustering of the dark matter but they do not work when the clustering is highly non-linear, and they do not provide rigorous framework for describing how the clustering of galaxies differs from that of the dark matter, either.

A way to study the non-linear evolution of the dark-matter distribution is using numerical simulations of the large-scale structure clustering process, which has been done extensively. As shown in Figure 1.1, these simulations demonstrate that an initially smooth matter distribution evolves into a complex network of lines and knots. The dense knots are often called dark-matter halos, which correspond to situations in which dark-matter overdensities collapse in a gravitationally bound structure. If distinct halos can be identified, then it is likely that they are small compared to the typical distances between them. Simulations such as the cited show that the halo abundance, spatial distribution, and internal density profiles are closely related to the properties of the initial fluctuation field, so that these provide a way to analyze the spatial statistics of the dark-matter density field from the linear to highly non-linear regimes.

The idea that galaxies form within these dark-matter halos has gained increasing credence, so that a halo based description of the dark-matter distribution of large-scale structure is very useful. The physical properties of galaxies are determined by the halos in which

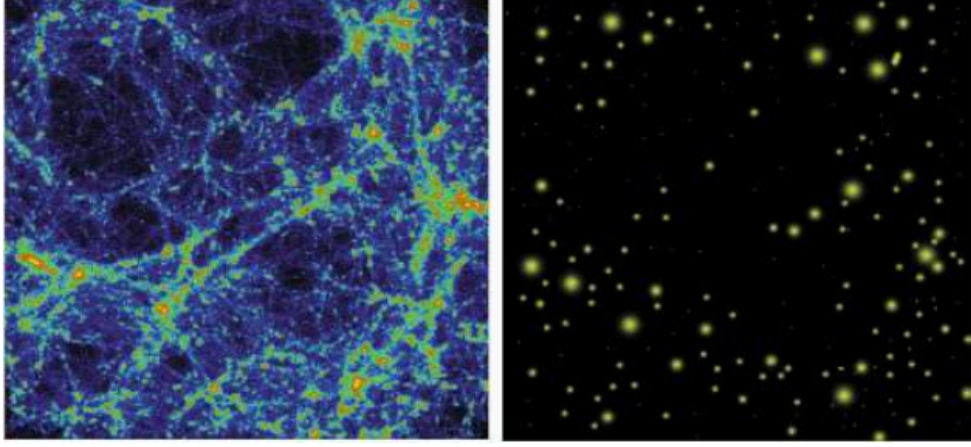


Figure 1.1: Left: complex distribution of dark matter found in numerical simulations. Right: distribution of dark-matter halos, where all the information is contained. Obtained from [7].

they form, hence the statistical properties of a given galaxy population are determined by the properties of the respective halo population. As explained in [3], although the properties and dynamics of individual halos can be very complicated, their average properties are simple and can be characterized by only their mass at a given time. For this reason, it is very useful to predict the abundance of dark-matter halos for each mass. This is known as the halo mass function, which will be denoted by dn/dM in the following, and represents the differential number density of halos of a given infinitesimal interval of halo mass. This quantity, since halos formed from regions in the initial density field which were sufficiently dense that they later collapsed, can be estimated from the number density of regions in the initial fluctuation field which were dense enough to collapse ([3], [7]). The halo mass function will be very important in next chapter to derive the probability of having a halo with some specific mass affecting to a specific point in a region of the Universe.

Another important quantity that we will use throughout this work is the halo bias. As we have seen, dark-matter halos are collapsed objects that form in the highest peaks of the density field. Therefore, although their clustering is expected to be related with the clustering of matter, in practice it cannot be the same. For this reason, we consider a relation between the matter perturbations and the halo number density perturbations, which we call bias ([3]):

$$\delta_h = \mathcal{F}[b]\delta_m. \quad (1.5)$$

In this equation, $\mathcal{F}[b]$ is a general functional of the bias. Although non-linear terms should be considered in order to faithfully reproduce the halo distribution, in this work we only consider a linear bias in order to simplify calculations, making sure that a linear bias approximation is accurate enough for our objective.

Now, we are going to introduce two important quantities that are used in the statistical description of any density distribution at large scales. To deal with these quantities, we will sometimes need to work in Fourier space, so we first introduce the convention followed throughout. For a d -dimensional Fourier space variable $\mathbf{u}$ being the conjugate of a configuration space variable $\mathbf{v}$, the direct and inverse Fourier transforms of a function f and its Fourier counterpart $\tilde{f}$ will be given by

$$\tilde{f}(\mathbf{u}) = \int d^d \mathbf{v} f(\mathbf{v}) e^{-i\mathbf{v} \cdot \mathbf{u}} \quad (1.6)$$

$$f(\mathbf{v}) = \frac{1}{(2\pi)^d} \int d^d \mathbf{u} \tilde{f}(\mathbf{u}) e^{i\mathbf{v} \cdot \mathbf{u}}, \quad (1.7)$$

where the tilde denotes Fourier space functions.

The quantities we are introducing are two-point quantities, this means they are defined between any two points in space, and their values depend on the position of these two points. However, these quantities can also be generalized to be defined between any n points in space, where n is any natural number. The first quantity is the so-called two-point correlation function ([7]), and is defined by

$$\langle \delta(\mathbf{x}) \delta(\mathbf{x}') \rangle = \xi(\mathbf{x} - \mathbf{x}'), \quad (1.8)$$

where δ represents the number density perturbations (or overdensity field) which can be related to the average number density as explained in next chapter, and the brackets denote the average over realizations of δ . The two-point correlation function represents the probability of finding two contributions to the density separated by a distance vector $\mathbf{x} - \mathbf{x}'$ with respect to a uniform distribution, and can be related to different types of physical content of the Universe depending on what density we consider, such as halo density or any kind of matter density. Regarding the number density perturbations $\delta(\mathbf{x})$ in configuration space, they can also be represented as a sum over Fourier modes of its Fourier transform $\tilde{\delta}(\mathbf{k})$.

Making use of the number density perturbations in Fourier space, $\tilde{\delta}(\mathbf{k})$, we can introduce the second quantity that helps in the statistical description of any density distribution, which is known as the power spectrum ([7]) and is defined by

$$\langle \tilde{\delta}(\mathbf{k}) \tilde{\delta}(\mathbf{k}') \rangle = (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}') P(\mathbf{k}), \quad (1.9)$$

where δ_D represents the three-dimensional Dirac delta. The power spectrum is a way of measuring the magnitude (or variance) of the density distribution fluctuations as a function of Fourier modes, and as stated in [1] if there are lots of very under- and overdense regions it will be large, whereas it is small if the density distribution is smooth.

There exists a relation between the two-point correlation function and the power spectrum, which is that they are Fourier transform pairs. Therefore, considering $\langle \delta(\mathbf{x}) \delta(\mathbf{x} + \mathbf{r}) \rangle = \xi(\mathbf{r})$ we can write

$$\xi(\mathbf{r}) = \frac{1}{(2\pi)^3} \int d^3 \mathbf{k} e^{i\mathbf{k} \cdot \mathbf{r}} P(\mathbf{k}). \quad (1.10)$$

Since the power spectrum $P(\mathbf{k})$ has units of k^{-3} , assuming that it is isotropic (i.e. $P(\mathbf{k}) = P(k)$ and it only depends on the magnitude of the modes), to make representations of $P(k)$ as function of the modes k we often use the dimensionless quantity

$$\Delta(k) = \frac{k^3 P(k)}{2\pi^2},$$

which is a good indicator of the clumpiness on scale k ([1]). Figure 1.2 shows $\Delta(k)$ as function of k for a galaxy distribution, where it can be compared the theoretical prediction with the measured data. It can be seen that on large scales (small k) the variance is smaller than unity, whereas in small scales (larger k) the variance is close to unity.

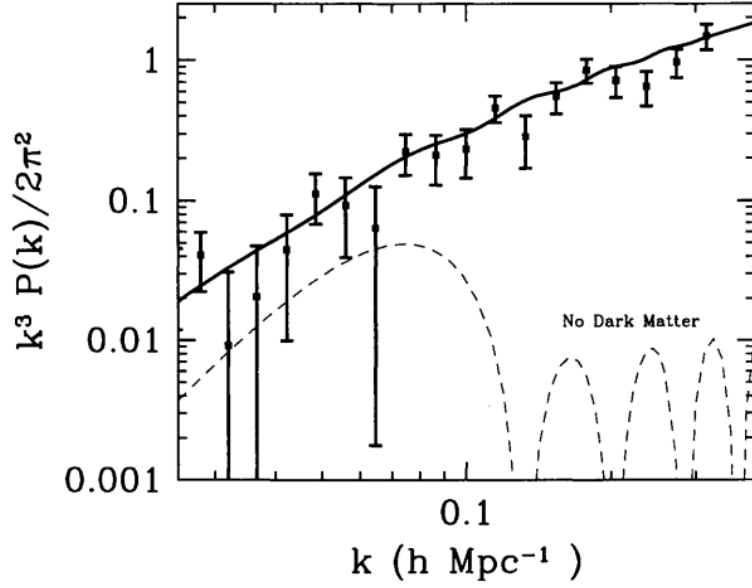


Figure 1.2: The variance $\Delta(k)$ of the Fourier transform of the galaxy distribution as a function of scale. The solid line is the theoretical prediction from a model in which the Universe contains dark matter, a cosmological constant, with perturbations generated by inflation. The dashed line is a theory with only baryons and no dark matter. Obtained from [1].

Finally, as explained in [7], under the approach that the mass in the Universe is partitioned up into halos and that the halos are small compared to the usual distance there exist between them, the distribution of the mass can be studied in two steps. On the one hand, the statistics of the mass density field on small scales are determined by the spatial distribution within halos, in a way that the precise manner in which the halos themselves are organized into large-scale structures is not important. On the other hand, on large scales what it is important is the spatial distribution of the halos, in a way that the details of the internal structure of the halos are not significant if the scales are larger than a typical halo. The halo model takes into account this distinction, and its formalism separate terms between those which describe the distribution of mass within each halo, and those which describe the spatial distribution of the halos themselves. In this way, regarding the two-point correlation function and the power spectrum, they are separated into two terms, which are called the one- and two-halo term, and account for the case in which the two contributions to the density are from the same halo and for the case in which the two contributions are from different halos, respectively. This means the two-point correlation function and the power spectrum can be expressed as

$$\xi(\mathbf{x} - \mathbf{x}') = \xi^{1h}(\mathbf{x} - \mathbf{x}') + \xi^{2h}(\mathbf{x} - \mathbf{x}') \quad \text{and} \quad P(\mathbf{k}) = P^{1h}(\mathbf{k}) + P^{2h}(\mathbf{k}),$$

where the superscript $1h$ means *one – halo* and the superscript $2h$ means *two – halo*. We will follow a similar formalism in next chapter when we derive the two-point PDF of the brightness temperature measured by LIM experiments.

As we have discussed, matter distribution in the Universe is determined by the Universe's evolution and the underlying physics that drive all interactions. However, we can not measure directly the mass distribution because it is mostly dark matter, hence we

need tracers. One of them is galaxy clustering and Line-Intensity Mapping is another, as we describe in next section.

1.6 Line-Intensity Mapping

In this section we are going to introduce what is Line-Intensity Mapping, why it is useful to extract information from the observable Universe and what it is based on. The references consulted here are [8], [9], [10], [11] and [12], where it is possible to find further and deeper information than the presented in this work.

1.6.1 What is Line-Intensity Mapping? Its main target lines

Line-intensity mapping is an exciting and emerging technique to survey the Universe. Unlike galaxy surveys, which determine the large scale distribution of mass by locating huge numbers of galaxies, intensity mapping measures the integrated emission of several spectral lines originating from individually unresolved galaxies and the intergalactic medium (IGM) with relatively low-aperture instruments. Mapping the intensity fluctuations of an array of lines offers a unique opportunity to probe redshifts well beyond the reach of other cosmological observations (such as galaxy surveys) over potentially huge three-dimensional volumes of the Universe, access regimes that cannot be explored otherwise and exploit the potential of cross-correlations with other measurements. The reason why LIM can use smaller aperture instruments compared to the ones used by galaxy surveys is because it does not require very high-resolution detections since it uses all incoming photons from any source within the field of view, obtaining tomographic line-of-sight information from targeting a known spectral line at different frequencies. Furthermore, this allows LIM experiments to be carried out with modest experimental budgets.

To help uncover the potential of LIM, we introduce here an example of its capability compared to that of the galaxy surveys. Figure 1.3 shows a comparison between the galaxies detected in a small sky patch by the Very Large Array (VLA), an advanced radio telescope observatory, and the intensity map detected by the carbon-monoxide (CO) intensity mapping instrument (COMAP). While the VLA detects only a small amount of the total number of CO-emitting galaxies, COMAP produces a map of the intensity fluctuations, which is sensitive to emission throughout the field.

As it has been shown in Figure 1.3, LIM holds promise to become a key tool in advancing our understanding of cosmology and astrophysics, as the Cosmic Microwave Background (CMB) and galaxy surveys have been until now. Moreover, in terms of access, LIM is great positioned to probe crucial epochs in the history of the Universe. For example, it is able to directly probe the epoch of reionization (EoR, the transition from a Universe filled with neutral hydrogen to one that is mostly ionized as it is today), the IGM, the interstellar medium (ISM) and the formation of stars. Figure 1.4 shows the epochs of the Universe that LIM is able to probe. Furthermore, extending the reach of a survey to higher redshifts increases the volume observed significantly. This enables the exploration of larger scales, potentially reaching scales of the order of the horizon, where imprints of inflation might be detectable.

Regarding the spectral emission lines that can be targets of LIM experiments, there exists a variety of them ranging from the microwave to the ultraviolet (UV) bands, which can be used to probe different phases of the IGM and ISM. We now briefly describe some of them. First, we consider CO emission lines. CO is the most common molecule in Universe besides diatomic molecular hydrogen (H_2) and is the most used tracer of molecular gas.

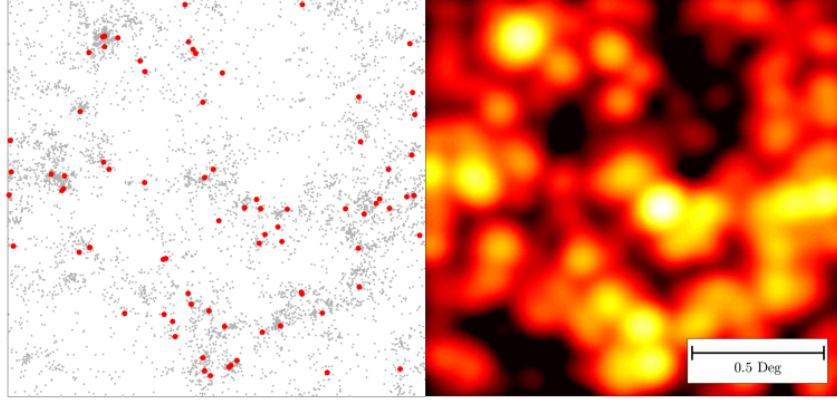


Figure 1.3: Left: a simulated 2.5 deg^2 field with galaxy positions, where the most bright galaxies detected by the VLA are marked in red. Right: CO intensity map detected by COMAP. Obtained from [10].

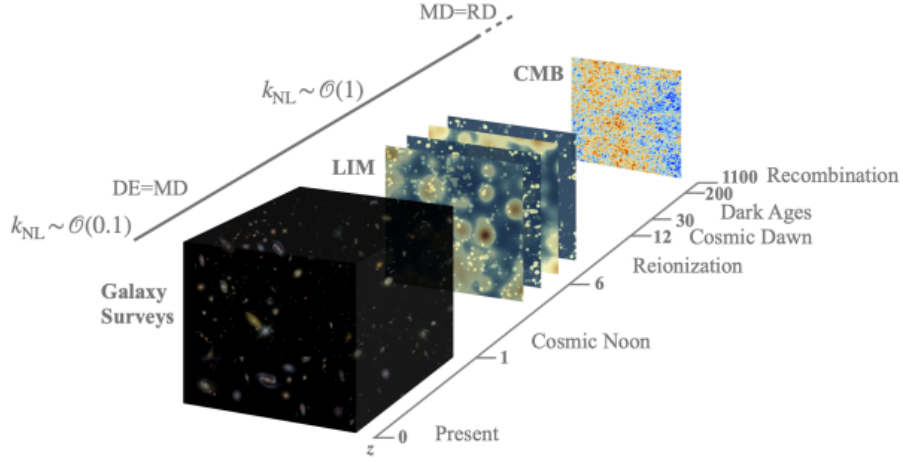


Figure 1.4: Different redshift volumes within reach of galaxy surveys, LIM and the CMB. Intensity mapping of multiple line emissions provides access to volumes otherwise inaccessible. MD=RD and DE=MD denote the redshifts of matter-radiation and dark energy-matter equality, respectively, and k_{NL} is an estimation of the scale at which matter clustering becomes non linear, specified in units of Mpc^{-1} . Obtained from [9].

Some of the brightest rotational line emissions in galactic spectra are those of CO and can be observed by terrestrial telescopes targeting the sub-mm wavelength range. Although one of the drawbacks CO emission has is that it is very sensitive to the environment, it can be useful to estimate the amount of stellar mass and the star-formation rate in galaxies. Second, we consider ionized carbon (CII) lines. Atomic and ionic fine-structure lines in the infrared are important drivers of the cooling process of interstellar gas, among which the [CII] $158 \mu\text{m}$ fine-structure line is the brightest. It provides a target for LIM experiments to trace the star-formation history, specially at high redshifts due to its brightness. This line suffers from CO lines contamination as foreground line-interlopers because its frequency lies just above the ladder of CO lines. Third, other atomic fine-structure lines provide

different targets for LIM experiments, such as silicon SIII lines, oxygen OI and OII lines, and nitrogen NII lines, which can be used to probe ISM physics. These lines and the ratios between them and other lines are useful to measure different parameters to characterize galaxies. Fourth and finally, several optical and ultraviolet hydrogen lines provide targets for LIM experiments, too. Among them, we highlight the Ly α , H α and H β lines, emitted in the ultraviolet (UV) and optical and redshifted to wavelengths down to the infrared. These lines can be used to extract information about star formation within galaxies, among other uses.

1.6.2 LIM signal

LIM experiments, as explained in [8] and [9], measure the specific intensity I_ν per unit of observed frequency ν_{obs} , which can be derived from the line-luminosity density ρ_L per comoving volume given the line luminosity L . This way, I_ν can be calculated from

$$I_\nu(\mathbf{x}) = \frac{c}{4\pi\nu H(z)} \rho_L(\mathbf{x}), \quad (1.11)$$

where $\nu = \nu_{\text{emit}}$. This equation allows to obtain the specific intensity as function of the line luminosity density ρ_L . However, experiments covering small frequencies (below some tens of GHz) usually employ the brightness temperature using the Rayleigh-Jeans relation

$$T(\mathbf{x}) = \frac{c^2 I_\nu(\mathbf{x})}{2k_B \nu_{\text{obs}}^2} = \frac{c^2 (1+z)^2}{2k_B \nu^2} I_\nu(\mathbf{x}), \quad (1.12)$$

where k_B is the Boltzmann constant. Through equation (1.11), we can express the brightness temperature as function of the line luminosity density as

$$T(\mathbf{x}) = \frac{c^3 (1+z)^2}{8\pi k_B \nu^3 H(z)} \rho_L(\mathbf{x}) = X_{\text{LT}} \rho_L(\mathbf{x}), \quad (1.13)$$

where we have defined X_{LT} as a redshift-dependent multiplicative factor to simplify the expressions. During this work we will use the brightness temperature as variable to describe line-intensity maps, but our approach is equally applicable to specific intensities.

1.6.3 Voxel Intensity Distribution

As we have explained in last section, the correlation function and the power spectrum are two useful statistical quantities that help analyze density distributions at large scales. However, since they are used to compute the variance of the distributions, they can only describe the Gaussian part of these distributions. This poses a problem when working with line-intensity maps, because as described in [8], since line-intensity fluctuations trace the non-linear, non-Gaussian large-scale structure, they are very non-Gaussian. For this reason, a significant part of the information contained in line-intensity maps is not captured by power spectrum measurements, hence it is not useful anymore. Accessing non-Gaussian information motivate the development of alternative summary statistics. In this context, one-point statistics are one of the best solutions to this problem, because they depend directly on the LIM PDF, hence on the full distribution of non-Gaussian intensity fluctuations and the whole line-luminosity function.

The one-point statistics that we are going to use is the actual estimator of the PDF, known as Voxel Intensity Distribution in the context of LIM. Imagine we have a region of

the Universe partitioned into smaller regions, known as voxels, as we will explain in next chapter when we introduce the practical considerations. In this situation, the VID is the histogram of the measured brightness temperature in each voxel normalized by the total number of voxels, and its relation with the observed PDF is

$$\mathcal{B}(\Delta T_i) = \int_{\Delta T_i} d\Delta T \mathcal{P}_{\text{obs}}(\Delta T), \quad (1.14)$$

where the integral is limited to the temperature interval centered on ΔT_i . Here, we normalize by the total number of voxels in order to deal with an intensive quantity, the value of which does not depend on the size of the survey. Furthermore, the observed PDF ($\mathcal{P}_{\text{obs}}$ in equation (1.14)) is a convolution between the astrophysical PDF ($\mathcal{P}_{\text{astro}}$), which represents the PDF of the line-luminosity signal, and the instrumental noise PDF ($\mathcal{P}_{\text{noise}}$), since there is no perfect experiment without noise, so that

$$\mathcal{P}_{\text{obs}}(\Delta T) = (\mathcal{P}_{\text{astro}} * \mathcal{P}_{\text{noise}})(\Delta T).$$

Hence, $\mathcal{B}$ in equation (1.14) represents the total VID measured in the experiment.

The VID, as described in [9], is complementary to the power spectrum beyond its access to the non-Gaussian information in the map. Combining the power spectrum and the VID can help break the degeneracy between cosmology and astrophysics and, as cited in [8], several studies have highlighted the sensitivity of the VID not only to astrophysical parameters but also to beyond- Λ CDM cosmologies and physics beyond the standard model.

1.7 Ergodic hypothesis

An important hypothesis we are going to assume to be true throughout this work in the halo/galaxy structure we are going to analyze is the ergodic hypothesis. This hypothesis, as described in [13], basically establishes that, given a dynamical system, it is ergodic if and only if the fraction of time the system spends in a certain region in phase space is equal to the phase space average of that region. In other words, the ergodic hypothesis states that, given an observable of the system, averaging it over time is equivalent to averaging over many copies of the system in different microstates.

In cosmology, it can be understood as that the average over potential realizations of the Universe for the state in a point is equivalent to the average over the states in all points. Particularly, in this work we are going to apply the ergodic hypothesis when we deal with the Fourier transform of the PDF of the brightness temperature for a single emitter, such as for example in equation (2.1). We are going to use it to write the average over realizations of the brightness temperature $T(\boldsymbol{x})$ as an average of the brightness temperature over the specific observed volume V_{obs} of the Universe.

Chapter 2

Theoretical formalism

This chapter presents the bulk of the work done for this project. Here we derive the formalism for the prediction of the one- and two-point PDFs of the brightness temperature in LIM experiments. For the latter, we separate it in two terms: the one-halo term and the two-halo term, depending on whether the temperatures measured correspond both to positions within the same halo or within two different halos.

In this work, given an emission line whose brightness temperature desires to be studied, we will neglect its broadening because we assume the spectral resolution is lower than the width of the line, hence we approximate the latter as a Dirac delta. However, we will consider the effect of angular and spectral resolutions. The first one refers to the capability of the temperature measuring instrument to distinguish between two objects separated by a small angle in the sky (therefore it is transverse to the line of sight), and the latter refers to the ability of the instrument to distinguish between two closely spaced spectral lines in wavelength or frequency (therefore it goes along the line of sight, with varying redshift). We consider spherical coordinates (r, ϕ, θ) in configuration space with origin at each galaxy, where $\phi \in [0, 2\pi)$ and $\theta \in [0, \pi]$, and we define $\mu := \hat{x} \cdot \hat{z}$ as the cosine of the angle between $\mathbf{x}$ and the z -axis. The z -axis is chosen to be the line of sight direction. The observed map is the LIM signal (brightness temperature) convoluted with the resolution and some experimental noise. Therefore, the measurable signal is not a three-dimensional Dirac delta, but it will have the actual spatial distribution of the resolution kernel. In this context, we assume that the angular resolution is the same in every polar direction, this means the brightness temperature $T(\mathbf{x}) = T(r, \phi, \theta)$ is independent of the angle ϕ (it has azimuthal symmetry), and that the spectral resolution is symmetric with respect to frequency (or redshift), therefore the brightness temperature has even parity with respect to the variable μ . Hence, the brightness temperature can be simply expressed as $T(\mathbf{x}) = T(r, \mu)$. These symmetries are going to be significant in the derivation of the one-halo term of the two-point PDF. Analogously, in Fourier space we can express $\tilde{T}(\mathbf{k}) = \tilde{T}(k, \mu_k)$, where we define $\mu_k := \hat{k} \cdot \hat{z}$ as the cosine of the angle between $\mathbf{k}$ and the z -axis.

2.1 One-point probability distribution function

The aim of this first section of the chapter is to model the LIM PDF of measuring a certain brightness temperature in an observed volume of the Universe. In order to do so, the two main articles we based our formalism on are [8] and [14], but going beyond them in some aspects.

Since brightness temperature is an additive quantity, the PDF of the aggregate emission is the convolution of the PDF of each emitter that exists in the observed volume, hence we need to take into account clustering to measure the PDF of many emitters. This calculation is much more tractable in Fourier space taking into account the convolution theorem. This theorem states that the Fourier transform of the convolution of two or more functions is equal to the product of the Fourier transforms of each function separately, and can be easily proved as done in [15]. For this reason, we will always try to work in Fourier space, so that convolutions become multiplications. In this context, τ is defined as the Fourier conjugate of the brightness temperature T . Hence, the Fourier transform of a given PDF $\mathcal{P}$ of the brightness temperature is

$$\tilde{\mathcal{P}}(\tau) = \int dT \mathcal{P}(T) e^{-iT\tau} = \langle e^{-iT\tau} \rangle = \frac{1}{V_{\text{obs}}} \int_{V_{\text{obs}}} d^3\mathbf{x} e^{-iT(\mathbf{x})\tau}, \quad (2.1)$$

where in the second equality angle brackets denote average over realizations. In the last equality, the ergodic hypothesis can be invoked, taking the average over the observed volume V_{obs} .

From now on, we will denote $\boldsymbol{\vartheta}$ as the set of astrophysical properties given a halo mass, as discussed in [16]. The halo mass M is explicitly separated from $\boldsymbol{\vartheta}$, and we consider an infinitesimal mass bin centered at M . Under these assumptions, the brightness temperature PDF for the bin is

$$\mathcal{P}^{(M)}(T) = \int d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \left[\mathcal{P}_{N=0}^{(M,\boldsymbol{\vartheta})} \delta_D(T) + \mathcal{P}_{N=1}^{(M,\boldsymbol{\vartheta})} \mathcal{P}_1^{(M,\boldsymbol{\vartheta})}(T) \right]. \quad (2.2)$$

In equation (2.2), we marginalize over the conditional multidimensional distribution $\mathcal{P}(\boldsymbol{\vartheta}|M)$. Moreover, $\mathcal{P}_{N=x}^{(M,\boldsymbol{\vartheta})}$ is the PDF of having x emitters (halos) of mass M and set of properties $\boldsymbol{\vartheta}$ contributing to a specific point, and $\mathcal{P}_x^{(M,\boldsymbol{\vartheta})}(T)$ is the PDF of finding a temperature T in a point in the space receiving contributions from x emitters with such properties. If there is no emitter, this is $N = 0$, then there is no signal and $\mathcal{P}_0^{(M,\boldsymbol{\vartheta})}(T) = \delta_D(T)$ is the Dirac delta centered at zero. For an infinitesimal mass bin we can consider $\mathcal{P}_{N>1}^{(M,\boldsymbol{\vartheta})} \equiv 0$, hence it becomes $\mathcal{P}_{N=0}^{(M,\boldsymbol{\vartheta})} = 1 - \mathcal{P}_{N=1}^{(M,\boldsymbol{\vartheta})}$.

Therefore, we can express $\mathcal{P}^{(M)}(T)$ in equation (2.2) as

$$\begin{aligned} \mathcal{P}^{(M)}(T) &= \int d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \left[(1 - \mathcal{P}_{N=1}^{(M,\boldsymbol{\vartheta})}) \delta_D(T) + \mathcal{P}_{N=1}^{(M,\boldsymbol{\vartheta})} \mathcal{P}_1^{(M,\boldsymbol{\vartheta})}(T) \right] = \\ &= \int d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \delta_D(T) + \int d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}_{N=1}^{(M,\boldsymbol{\vartheta})} \left(\mathcal{P}_1^{(M,\boldsymbol{\vartheta})}(T) - \delta_D(T) \right). \end{aligned} \quad (2.3)$$

With this, taking into account that $\int d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) = 1$ and calculating the Fourier transform of $\mathcal{P}^{(M)}(T)$, we obtain

$$\tilde{\mathcal{P}}^{(M)}(\tau) = 1 + \int d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}_{N=1}^{(M,\boldsymbol{\vartheta})} \left(\tilde{\mathcal{P}}_1^{(M,\boldsymbol{\vartheta})}(\tau) - 1 \right). \quad (2.4)$$

In relation to $\tilde{\mathcal{P}}_1^{(M,\boldsymbol{\vartheta})}(\tau)$, it is the Fourier transform of $\mathcal{P}_1^{(M,\boldsymbol{\vartheta})}(T)$. Although the profile of the emission extends arbitrarily in space, in practice we can truncate it at some distance large enough that there is no sizable signal loss. Under this assumption, the signal profile only covers a finite volume V_{prof} which depends on M and $\boldsymbol{\vartheta}$, so that we can express

$$\begin{aligned} \tilde{\mathcal{P}}_1^{(M,\boldsymbol{\vartheta})}(\tau) &= \int d^3\mathbf{x} dL \mathcal{P}(\mathbf{x}|M, \boldsymbol{\vartheta}) \mathcal{P}(L|M, \boldsymbol{\vartheta}) e^{-iT(\mathbf{x})\tau} = \\ &= \frac{1}{V_{\text{prof}}} \int_{V_{\text{prof}}} d^3\mathbf{x} dL \mathcal{P}(L|M, \boldsymbol{\vartheta}) e^{-iT(\mathbf{x})\tau}. \end{aligned} \quad (2.5)$$

Here, we marginalize over the position (for which $\mathcal{P}(\mathbf{x}|M, \boldsymbol{\vartheta})$ is uniform over V_{prof} and zero otherwise as in the second equality) and $\mathcal{P}(L|M, \boldsymbol{\vartheta})$ accounts for any distribution of the line luminosity given the halo mass and the astrophysical properties.

Besides that, assuming Poisson statistics and neglecting clustering from now, $\mathcal{P}_{N=1}^{(M, \boldsymbol{\vartheta})}$ can be expressed as

$$\mathcal{P}_{N=1}^{(M, \boldsymbol{\vartheta})} = dM \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}), \quad (2.6)$$

where dn/dM is the halo mass function. This is, the number density of halos in the bin times the volume occupied, which depends on M and $\boldsymbol{\vartheta}$. For infinitesimal bins we can assume $\mathcal{P}_{N=1}^{(M, \boldsymbol{\vartheta})} \ll 1$, so it is possible to interpret equation (2.4) as the linear expansion of the exponential and write

$$\tilde{\mathcal{P}}^{(M)}(\tau) = \exp \left\{ \int d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}_{N=1}^{(M, \boldsymbol{\vartheta})} \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau) - 1 \right) \right\}. \quad (2.7)$$

However, this is the Fourier transform of the PDF for a single halo mass M . We need to extend this to all halos (i.e., to all possible masses). Considering the convolution theorem, since the PDF for the whole population of halos is the convolution of the PDFs for each halo, the Fourier transform of the PDF for the whole population is the product of the individual Fourier transforms. Hence, the Fourier transform of the PDF for the whole population without accounting for clustering is

$$\begin{aligned} \tilde{\mathcal{P}}^{(u)}(\tau) &= \prod \tilde{\mathcal{P}}^{(M)}(\tau) = \\ &= \exp \left\{ \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau) - 1 \right) \right\}, \end{aligned} \quad (2.8)$$

where we have substituted the sum of exponents by the integral limit in the second equality and the superscript u means unclustered, and $\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau)$ is given by equation (2.5).

Now, we are including the effect of clustering. For this, first we assume that the astrophysical properties are uncorrelated with clustering. Clustering varies in scales much larger than the observed intensity profiles of the sources. Hence, for a specific realization or position, we can include its effect by adding the halo overdensity field

$$\delta_h = \frac{n_h - \langle n_h \rangle}{\langle n_h \rangle} \quad (2.9)$$

to the halo mass function, where n_h is the local halo number density. Accordingly, we promote

$$\frac{dn}{dM} \rightarrow \frac{dn}{dM} (1 + \delta_h(\mathbf{x}, M)), \quad (2.10)$$

where $\mathbf{x}$ is a specific position of the observed volume, and then the PDF of having one emitter of mass M and set of properties $\boldsymbol{\vartheta}$ contributing to a specific point, this is $\mathcal{P}_{N=1}^{(M, \boldsymbol{\vartheta})}$, is

$$\mathcal{P}_{N=1}^{(M, \boldsymbol{\vartheta})} = dM \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) (1 + \delta_h(\mathbf{x}, M)). \quad (2.11)$$

Therefore, we can rewrite equation (2.8) taking into account equation (2.11) to obtain the

Fourier transform of the PDF considering the effect of clustering as

$$\begin{aligned}
\tilde{\mathcal{P}}^{(\delta)}(\tau) &= \exp \left\{ \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} (1 + \delta_h(\mathbf{x}, M)) V_{\text{prof}}(M, \boldsymbol{\vartheta}) \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau) - 1 \right) \right\} = \\
&= \exp \left\{ \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau) - 1 \right) \right\} \times \\
&\times \exp \left\{ \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau) - 1 \right) \delta_h(\mathbf{x}, M) \right\} = \\
&= \tilde{\mathcal{P}}^{(u)}(\tau) \exp \left\{ \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau) - 1 \right) \delta_h(\mathbf{x}, M) \right\}. \quad (2.12)
\end{aligned}$$

It is important to remark that $\tilde{\mathcal{P}}^{(\delta)}(\tau)$ only accounts for clustering for a single realization.

Now, we need to calculate the Fourier transform of the global PDF, which is obtained performing the average over the realizations of $\tilde{\mathcal{P}}^{(\delta)}$. Since $\tilde{\mathcal{P}}^{(u)}$ does not depend on the overdensities, we can take it out of the average and we are left with the exponential of the term including δ_h in equation (2.12) as

$$\begin{aligned}
\tilde{\mathcal{P}}(\tau) &= \langle \tilde{\mathcal{P}}^{(\delta)}(\tau) \rangle = \\
&= \tilde{\mathcal{P}}^{(u)} \langle \exp \left\{ \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau) - 1 \right) \delta_h(\mathbf{x}, M) \right\} \rangle. \quad (2.13)
\end{aligned}$$

At this point, we invoke the moment-generating function (see [17]), which states that for a random variable X ,

$$\langle e^X \rangle = \exp \left\{ \sum_{p=1}^{\infty} \langle X^p \rangle / p! \right\}. \quad (2.14)$$

By definition in equation (2.9), $\langle \delta_h \rangle = 0$. Although gravitational collapse induces non Gaussianities and higher-order terms should be included, we take a first approximation and truncate the moment-generating function at $p = 2$. Hence, equation (2.13) can be written as

$$\begin{aligned}
\tilde{\mathcal{P}}(\tau) &= \tilde{\mathcal{P}}^{(u)} \exp \left\{ \left[\int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau) - 1 \right) \right]^2 \times \right. \\
&\quad \left. \times \frac{\langle \delta_h(\mathbf{x}, M) \delta_h(\mathbf{x}, M) \rangle}{2} \right\}, \quad (2.15)
\end{aligned}$$

where we have applied the considerations described. Furthermore, we relate δ_h to the underlying matter density field δ_m with a linear, mass-dependent halo bias b_h so that $\delta_h(\mathbf{x}, M) = b_h(M) \delta_m(\mathbf{x})$. This linear bias formula, as developed in [7], is only accurate on large scales as the considered in this work. Thus, we can calculate the second cumulant

of the halo distribution as

$$\begin{aligned}
\langle \delta_h(\mathbf{x}, M) \delta_h(\mathbf{x}, M') \rangle &= \langle b_h(M) \delta_m(\mathbf{x}, M) b_h(M') \delta_m(\mathbf{x}, M') \rangle = \\
&= b_h(M) b_h(M') \langle \frac{1}{(2\pi)^3} \int d^3\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{x}} \tilde{\delta}_m(\mathbf{k}, M) \times \\
&\times \frac{1}{(2\pi)^3} \int d^3\mathbf{k}' e^{i\mathbf{k}'\cdot\mathbf{x}} \tilde{\delta}_m(\mathbf{k}', M') \rangle = \\
&= b_h(M) b_h(M') \frac{1}{(2\pi)^6} \int d^3\mathbf{k} d^3\mathbf{k}' e^{i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{k}'\cdot\mathbf{x}} \langle \tilde{\delta}_m(\mathbf{k}, M) \tilde{\delta}_m(\mathbf{k}', M') \rangle = \\
&= b_h(M) b_h(M') \frac{1}{(2\pi)^6} \int d^3\mathbf{k} d^3\mathbf{k}' e^{i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{k}'\cdot\mathbf{x}} (2\pi)^3 P(\mathbf{k}) \delta_D(\mathbf{k} + \mathbf{k}') = \\
&= b_h(M) b_h(M') \frac{1}{(2\pi)^3} \int d^3\mathbf{k} P(\mathbf{k}) = b_h(M) b_h(M') \sigma^2, \tag{2.16}
\end{aligned}$$

where we have written $\delta_m(\mathbf{x})$ for each mass as the inverse Fourier transform of $\tilde{\delta}_m(\mathbf{k})$ in each case. Moreover, in the fourth equality we have made use of the identity (2.3) in [18], which allows to relate the second cumulant of the matter distribution in Fourier space with the power spectrum. In the last equality we have also considered σ^2 , which is the zero-lag variance of the matter distribution, as the two-point correlation function evaluated at $\mathbf{r} = \mathbf{0}$, this is $\xi(\mathbf{0}) = \sigma^2$. With equation (2.16), we can finally obtain the Fourier transform of the global PDF from equation (2.15), which is given by

$$\tilde{\mathcal{P}}(\tau) = \tilde{\mathcal{P}}^{(u)} \exp \left\{ \left[\int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau) - 1 \right) b_h(M) \right]^2 \frac{\sigma^2}{2} \right\}. \tag{2.17}$$

In this equation, we assume $M = M'$ without loss of generality, hence we have $b_h(M) b_h(M') = (b_h(M))^2$.

It is important to highlight that this allows to calculate the Fourier transform of the PDF of the brightness temperature in a specific point of the Universe. However, as we will see in next chapter, in practice we deal with a discretized map, a voxelized volume in which voxels are the basic unit of information we have access to. For this reason, halo overdensities are smoothed over scales of the size of the voxel. Similar to what is done in [8], we can model this convolving the overdensity field with a window function W_{vox} of the voxel, which is a normalized ($\int d^3\mathbf{x} W_{\text{vox}}(\mathbf{x}) = 1$) top-hat function with the extent of the voxel. Furthermore, in practice the signal we can measure in LIM experiments is not a Dirac delta, it is a Gaussian beam characterized by the angular resolution. This also induces some smoothness on the overdensity field, which can be modeled also convolving the overdensity field with another window function W_{smooth} . This function, as we have cited, is a normalized Gaussian centered at zero and with standard deviation given by the angular resolution. Therefore, in practice we can calculate the overdensity field $\delta_h^v(\mathbf{x})$ taking into account these window functions as

$$\delta_h^v(\mathbf{x}) = \delta_h(\mathbf{x}) * W_{\text{vox}}(\mathbf{x}) * W_{\text{smooth}}(\mathbf{x}).$$

This also has implications when calculating the zero-lag variance of the matter distribution. It is easy to show, making use of the convolution theorem and the definition of the power spectrum given by equation (1.9), that the zero-lag variance if we consider the described window functions becomes to be

$$\sigma_v^2 = \frac{1}{(2\pi)^3} \int d^3\mathbf{k} \tilde{W}_{\text{vox}}^2(\mathbf{k}) \tilde{W}_{\text{smooth}}^2(\mathbf{k}) P(\mathbf{k}), \tag{2.18}$$

where $\tilde{W}_{\text{vox}}(\mathbf{k})$ is the Fourier transform of W_{vox} , which is a sinc function, and $\tilde{W}_{\text{smooth}}(\mathbf{k})$ is the Fourier transform of W_{smooth} , which is another Gaussian function. In addition to this, taking into account the window functions also modifies the power spectrum. In a similar way to the zero-lag variance, the power spectrum can be expressed as

$$P_v(\mathbf{k}) = \tilde{W}_{\text{vox}}^2(\mathbf{k}) \tilde{W}_{\text{smooth}}^2(\mathbf{k}) P(\mathbf{k}). \quad (2.19)$$

We will see an example of how considering the window functions changes the shape of the power spectrum but applied to the angular power spectrum in next chapter.

With all this, we can obtain the PDF of the brightness temperature in a point by computing the inverse Fourier transform of equation (2.17) above,

$$\mathcal{P}(T) = \frac{1}{2\pi} \int d\tau \tilde{\mathcal{P}}(\tau) e^{iT\tau}. \quad (2.20)$$

Finally note, as we explain in detail in next chapter, that in practice we compute the VID, not directly the PDF, and it is important to discuss here about its covariance. It is usually assumed ([8]) that the covariance of the VID is that of a multinomial distribution, this is

$$C_{ab}^{\text{multinomial}} = \begin{cases} \frac{\mathcal{B}_a}{N_{\text{vox}}} (1 - \mathcal{B}_a) & \text{if } a = b \\ -\frac{\mathcal{B}_a \mathcal{B}_b}{N_{\text{vox}}} & \text{if } a \neq b, \end{cases} \quad (2.21)$$

where N_{vox} is the number of voxels in our discretized map. However, as shown in [8], the numerical covariance of the VID only agrees well with the multinomial variance in the diagonal terms. The off-diagonal do not agree well because there exist physical correlations. In this context, the two-point PDF offers an alternative to compute the theoretical analytic covariance of the VID, as shown in [14]. We now start to derive the theoretical formalism of the two-point PDF.

2.2 Two-point probability distribution function

The aim of this second section is to model the LIM PDF of measuring two different temperatures in an observed volume corresponding to two different positions separated by a certain distance vector. This distance vector is denoted as $\mathbf{r}$ in the following. The main reference used in this section is [14], but most of the work is new compared to this article. As in the previous section, in order to calculate convolutions in an easier way we will always work in Fourier space so that they become multiplications. In this context, τ_a is defined as the Fourier conjugate of a brightness temperature T , and τ_b is defined as the Fourier conjugate of a brightness temperature T' . Therefore, the Fourier transform of a given two-point PDF of the brightness temperature $\mathcal{P}(T, T'; \mathbf{r})$ is

$$\begin{aligned} \tilde{\mathcal{P}}(\tau_a, \tau_b; \mathbf{r}) &= \int dT dT' \mathcal{P}(T, T'; \mathbf{r}) e^{-iT\tau_a} e^{-iT'\tau_b} = \langle e^{-iT\tau_a} e^{-iT'\tau_b} \rangle = \\ &= \frac{1}{V_{\text{obs}}} \int_{V_{\text{obs}}} d^3\mathbf{x} e^{-iT(\mathbf{x})\tau_a} e^{-iT'(\mathbf{x}+\mathbf{r})\tau_b}, \end{aligned} \quad (2.22)$$

where we have invoked the ergodic hypothesis again as in equation (2.1) and we have related T and T' through the equality $T'(\mathbf{x}) = T(\mathbf{x} + \mathbf{r})$ due to the fact that they are separated by the distance vector $\mathbf{r}$.

We assume each temperature contribution comes from one and only one halo, and we will separate the two-point PDF into a one- and a two-halo term so that the two-point PDF in configuration space becomes

$$\mathcal{P}(T, T'; \mathbf{r}) = \mathcal{P}_{1h}(T, T'; \mathbf{r}) * \mathcal{P}_{2h}(T, T'; \mathbf{r}) \quad (2.23)$$

and the two-point PDF in Fourier space becomes

$$\tilde{\mathcal{P}}(\tau_a, \tau_b; \mathbf{r}) = \tilde{\mathcal{P}}_{1h}(\tau_a, \tau_b; \mathbf{r}) \tilde{\mathcal{P}}_{2h}(\tau_a, \tau_b; \mathbf{r}) \quad (2.24)$$

due to the convolution theorem. Firstly, the one-halo term (denoted by $\mathcal{P}_{1h}$) will account for the probability of having two different positions with two different temperatures T and T' , respectively, separated by $\mathbf{r}$ within the same halo of mass M . Secondly, the two-halo term (denoted by $\mathcal{P}_{2h}$) will account for the probability of having two different positions with two different temperatures T and T' , respectively, separated by $\mathbf{r}$ within two different halos of masses M and M' . A sketch of the situation that represents each of the terms is provided in Figure 2.1.

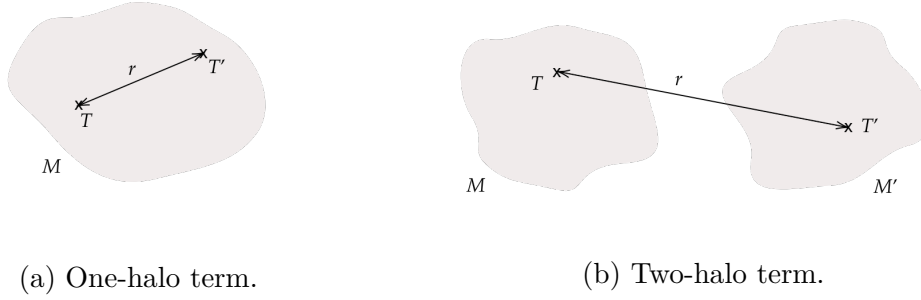


Figure 2.1: Two-point PDF terms.

In this way, we first compute the one-halo term of the two-point PDF in section 2.2.1 and then the two-halo term of the two-point PDF in section 2.2.2.

2.2.1 One-halo term

As described before, the one-halo term accounts for the possibilities for which the two different positions of temperatures T and T' belong to the same halo of mass M .

Following the same procedure as in modelling the one-point PDF, we denote $\boldsymbol{\vartheta}$ as the set of astrophysical properties given a halo mass and the halo mass M is explicitly separated from them. We consider an infinitesimal mass bin centered at M , so that the one-halo term of the two-point PDF of the brightness temperature for the bin is

$$\mathcal{P}_{1h}^{(M)}(T, T'; \mathbf{r}) = \int d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \left[\mathcal{P}_{N=0}^{(M, \boldsymbol{\vartheta})} \delta_D(T) \delta_D(T') + \mathcal{P}_{N=1}^{(M, \boldsymbol{\vartheta})} \mathcal{P}_1^{(M, \boldsymbol{\vartheta})}(T, T'; \mathbf{r}) \right], \quad (2.25)$$

where we again marginalize over the conditional multidimensional distribution $\mathcal{P}(\boldsymbol{\vartheta}|M)$ and $\mathcal{P}_{N=x}^{(M, \boldsymbol{\vartheta})}$ has the same meaning as in equation (2.2). However, now $\mathcal{P}_x^{(M, \boldsymbol{\vartheta})}(T, T'; \mathbf{r})$ is the two-point PDF of finding temperatures T and T' separated by a distance vector $\mathbf{r}$ in the space receiving contributions from x emitters with set of properties $\boldsymbol{\vartheta}$. If there is no emitter, this is $N = 0$, then there is no signal and $\mathcal{P}_0^{(M, \boldsymbol{\vartheta})}(T, T'; \mathbf{r}) = \delta_D(T) \delta_D(T')$ is the

product of two Dirac deltas both centered at zero. In the same way as in the one-point PDF, since we are assuming in the one-halo term that the two temperatures T and T' are both observed within the same halo, for an infinitesimal mass bin we can consider $\mathcal{P}_{N>1}^{(M,\vartheta)} \equiv 0$ so that $\mathcal{P}_{N=0}^{(M,\vartheta)} = 1 - \mathcal{P}_{N=1}^{(M,\vartheta)}$. Therefore, we can express $\mathcal{P}_{1h}^{(M)}(T, T'; \mathbf{r})$ in equation (2.25) as

$$\begin{aligned} \mathcal{P}_{1h}^{(M)}(T, T'; \mathbf{r}) &= \int d\vartheta \mathcal{P}(\vartheta|M) \left[(1 - \mathcal{P}_{N=1}^{(M,\vartheta)}) \delta_D(T) \delta_D(T') + \mathcal{P}_{N=1}^{(M,\vartheta)} \mathcal{P}_1^{(M,\vartheta)}(T, T'; \mathbf{r}) \right] = \\ &= \int d\vartheta \mathcal{P}(\vartheta|M) \delta_D(T) \delta_D(T') + \\ &+ \int d\vartheta \mathcal{P}(\vartheta|M) \mathcal{P}_{N=1}^{(M,\vartheta)} \left(\mathcal{P}_1^{(M,\vartheta)}(T, T'; \mathbf{r}) - \delta_D(T) \delta_D(T') \right). \end{aligned} \quad (2.26)$$

Taking into account that $\int d\vartheta \mathcal{P}(\vartheta|M) = 1$ and calculating the Fourier transform of $\mathcal{P}_{1h}^{(M)}(T, T'; \mathbf{r})$, we obtain

$$\tilde{\mathcal{P}}_{1h}^{(M)}(\tau_a, \tau_b; \mathbf{r}) = 1 + \int d\vartheta \mathcal{P}(\vartheta|M) \mathcal{P}_{N=1}^{(M,\vartheta)} \left(\tilde{\mathcal{P}}_1^{(M,\vartheta)}(\tau_a, \tau_b; \mathbf{r}) - 1 \right). \quad (2.27)$$

Regarding $\tilde{\mathcal{P}}_{1h}^{(M)}(\tau_a, \tau_b; \mathbf{r})$, it is the Fourier transform of $\mathcal{P}_{1h}^{(M)}(T, T'; \mathbf{r})$. Following the assumption made before when modelling the one-point PDF, we can truncate the profile of the emission at some distance large enough that there is no sizable signal loss. Hence, the signal profile only covers a finite volume V_{prof} , so that we can express

$$\begin{aligned} \tilde{\mathcal{P}}_1^{(M,\vartheta)}(\tau_a, \tau_b; \mathbf{r}) &= \int d^3\mathbf{x} dL \mathcal{P}(\mathbf{x}|M, \vartheta) \mathcal{P}(L|M, \vartheta) e^{-iT(\mathbf{x})\tau_a} e^{-iT(\mathbf{x}+\mathbf{r})\tau_b} = \\ &= \frac{1}{V_{\text{prof}}} \int_{V_{\text{prof}}} d^3\mathbf{x} dL \mathcal{P}(L|M, \vartheta) e^{-iT(\mathbf{x})\tau_a} e^{-iT(\mathbf{x}+\mathbf{r})\tau_b}. \end{aligned} \quad (2.28)$$

In this equation, we have again marginalized over the position, for which $\mathcal{P}(\mathbf{x}|M, \vartheta)$ is uniform over V_{prof} and zero otherwise, and $\mathcal{P}(L|M, \vartheta)$ accounts for any distribution of the line luminosity given the halo mass and the astrophysical properties. Now, we define

$$K_a(\mathbf{x}) := e^{-iT(\mathbf{x})\tau_a} - 1 \longrightarrow e^{-iT(\mathbf{x})\tau_a} = 1 + K_a(\mathbf{x})$$

and analogously we define $K_b(\mathbf{x} + \mathbf{r})$, so that we can write equation (2.28) as

$$\begin{aligned} \tilde{\mathcal{P}}_1^{(M,\vartheta)}(\tau_a, \tau_b; \mathbf{r}) &= \frac{1}{V_{\text{prof}}} \int_{V_{\text{prof}}} d^3\mathbf{x} dL \mathcal{P}(L|M, \vartheta) (1 + K_a(\mathbf{x})) (1 + K_b(\mathbf{x} + \mathbf{r})) = \\ &= \frac{1}{V_{\text{prof}}} \int_{V_{\text{prof}}} d^3\mathbf{x} dL \mathcal{P}(L|M, \vartheta) \times \\ &\times (1 + K_a(\mathbf{x}) + K_b(\mathbf{x} + \mathbf{r}) + K_a(\mathbf{x})K_b(\mathbf{x} + \mathbf{r})). \end{aligned} \quad (2.29)$$

With this, following the same steps as in the one-point PDF, $\mathcal{P}_{N=1}^{(M,\vartheta)}$ is given by equation (2.6) and for infinitesimal bins we can assume $\mathcal{P}_{N=1}^{(M,\vartheta)} \ll 1$, so it is possible to interpret equation (2.27) as the linear expansion of the exponential and write

$$\tilde{\mathcal{P}}_{1h}^{(M)}(\tau_a, \tau_b; \mathbf{r}) = \exp \left\{ \int d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}_{N=1}^{(M, \boldsymbol{\vartheta})} \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau_a, \tau_b; \mathbf{r}) - 1 \right) \right\}. \quad (2.30)$$

However, again this is the Fourier transform of the PDF for a single halo mass M . To extend this to the whole population of halos, we need to make the convolution of the PDFs for each halo, which in Fourier space becomes a product. Therefore, we obtain

$$\begin{aligned} \tilde{\mathcal{P}}_{1h}(\tau_a, \tau_b; \mathbf{r}) &= \prod \tilde{\mathcal{P}}_{1h}^{(M)}(\tau_a, \tau_b; \mathbf{r}) = \\ &= \exp \left\{ \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \left(\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau_a, \tau_b; \mathbf{r}) - 1 \right) \right\}, \end{aligned} \quad (2.31)$$

where we have substituted the sum of exponents by the integral limit in the second equality. It is important to remark that this equation represents the one-halo term of the two-point brightness temperature PDF. In this term we do not include the effect of clustering as we did in the one-point PDF. This is because this effect, which originates from the fact that two or more halos may be close in space, is already taken into account in the two-halo term of the two-point brightness temperature PDF, as we will discuss in the next section.

Now, taking into account equation (2.29), which gives the expression of $\tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau_a, \tau_b; \mathbf{r})$, we observe that the terms of the integral in the last equality corresponding to $K_a(\mathbf{x})$ and $K_b(\mathbf{x} + \mathbf{r})$ are easy to compute, but the term involving the product $K_a(\mathbf{x})K_b(\mathbf{x} + \mathbf{r})$ is more difficult to calculate computationally because it is a convolution in three dimensions. In order to make it easier to compute we develop it now alone using Fourier transforms, properties of the Dirac delta and developments in Legendre polynomials.

It is important to remark that the spatial integral in equation (2.29) is done only over V_{prof} as a consequence of the application of the ergodic hypothesis, but that we are going to consider from now on the integral over the entire tridimensional space. This does not lead to any problem, because we can always multiply the integrand by a window function which is zero out of V_{prof} in order to transform the integral over V_{prof} to a integral over the entire space.

We can regard, firstly, the functions $K_a(\mathbf{x})$ and $K_b(\mathbf{x} + \mathbf{r})$ as the inverse Fourier transforms of their counterparts in Fourier space $\tilde{K}_a(\mathbf{k})$ and $\tilde{K}_b(\mathbf{k}')$, respectively. Moreover, we can use the definition of the three-dimensional Dirac delta, δ_D , as stated in [18], which gives

$$\int d^3\mathbf{x} e^{i\mathbf{x} \cdot (\mathbf{k} + \mathbf{k}')} = (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}').$$

These expressions can be substituted into the convolution term of the integral in equation (2.29) to obtain

$$\begin{aligned} \int d^3\mathbf{x} K_a(\mathbf{x}) K_b(\mathbf{x} + \mathbf{r}) &= \int d^3\mathbf{x} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \int \frac{d^3\mathbf{k}'}{(2\pi)^3} \tilde{K}_a(\mathbf{k}) \tilde{K}_b(\mathbf{k}') e^{i\mathbf{k} \cdot \mathbf{x}} e^{i\mathbf{k}' \cdot (\mathbf{x} + \mathbf{r})} = \\ &= \frac{1}{(2\pi)^6} \int d^3\mathbf{x} \int d^3\mathbf{k} d^3\mathbf{k}' \tilde{K}_a(\mathbf{k}) \tilde{K}_b(\mathbf{k}') e^{i\mathbf{x} \cdot (\mathbf{k} + \mathbf{k}')} e^{i\mathbf{k}' \cdot \mathbf{r}} = \\ &= \frac{1}{(2\pi)^6} \int d^3\mathbf{k} d^3\mathbf{k}' \tilde{K}_a(\mathbf{k}) \tilde{K}_b(\mathbf{k}') e^{i\mathbf{k}' \cdot \mathbf{r}} \int d^3\mathbf{x} e^{i\mathbf{x} \cdot (\mathbf{k} + \mathbf{k}')} = \\ &= \frac{1}{(2\pi)^6} \int d^3\mathbf{k} d^3\mathbf{k}' \tilde{K}_a(\mathbf{k}) \tilde{K}_b(\mathbf{k}') e^{i\mathbf{k}' \cdot \mathbf{r}} (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}') = \\ &= \frac{1}{(2\pi)^3} \int d^3\mathbf{k} \tilde{K}_a(\mathbf{k}) \tilde{K}_b(-\mathbf{k}) e^{-i\mathbf{k} \cdot \mathbf{r}}. \end{aligned} \quad (2.32)$$

Now, since the functions K_a and K_b have azimuthal symmetry because they are functions of $T(\mathbf{x})$, which have azimuthal symmetry, they do not depend explicitly on the azimuthal angle, so we can express $\tilde{K}_a(\mathbf{k})$ simply as $\tilde{K}_a(k, \mu_k)$ and $\tilde{K}_b(\mathbf{k})$ as $\tilde{K}_b(k, \mu_k)$, where k is the module of the vector $\mathbf{k}$. Therefore, the product $\tilde{K}_a(\mathbf{k})\tilde{K}_b(-\mathbf{k})$ can be redefined as a function $\Gamma(\mathbf{k}) = \Gamma(k, \mu_k) := \tilde{K}_a(\mathbf{k})\tilde{K}_b(-\mathbf{k})$.

Developing in Legendre polynomials we can now express the function $\Gamma(\mathbf{k})$ as

$$\Gamma(\mathbf{k}) = \Gamma(k, \mu_k) = \sum_{l=0}^{\infty} \Gamma_l(k) \mathcal{L}_l(\mu_k), \quad (2.33)$$

where $\mathcal{L}_l(\mu_k)$ is the Legendre polynomial of degree l depending on the cosine of the angle between $\mathbf{k}$ and the z -axis, μ_k , and $\Gamma_l(k)$ is given by

$$\Gamma_l(k) = \frac{2l+1}{2} \int_{-1}^1 d\mu_k \tilde{K}_a(\mathbf{k}) \tilde{K}_b(-\mathbf{k}) \mathcal{L}_l(\mu_k). \quad (2.34)$$

This is done in order to eliminate the direct dependence of the function Γ on the variable μ_k . Therefore, continuing from equation (2.32) we can write

$$\begin{aligned} \int d^3\mathbf{x} K_a(\mathbf{x}) K_b(\mathbf{x} + \mathbf{r}) &= \frac{1}{(2\pi)^3} \int d^3\mathbf{k} \Gamma(k, \mu_k) e^{-i\mathbf{k} \cdot \mathbf{r}} = \\ &= \frac{1}{(2\pi)^3} \int d^3\mathbf{k} \sum_{l=0}^{\infty} \Gamma_l(k) \mathcal{L}_l(\mu_k) e^{-i\mathbf{k} \cdot \mathbf{r}} = \\ &= \frac{1}{(2\pi)^3} \int dk k^2 \sum_{l=0}^{\infty} \Gamma_l(k) \int d\Omega_k \mathcal{L}_l(\mu_k) e^{-i\mathbf{k} \cdot \mathbf{r}}, \end{aligned} \quad (2.35)$$

where $d\Omega_k = \sin \theta_k d\theta_k d\phi_k$. Here θ_k and ϕ_k are the polar and azimuthal angles of $\mathbf{k}$, respectively. This is, $\mu_k = \cos(\theta_k)$.

At this point, we make use of the identity (A12) in [18], which allows to write the integral over the solid angle Ω_k as a product of a spherical Bessel function and a Legendre polynomial. This identity is

$$\int d\Omega_x e^{i\mathbf{x} \cdot \mathbf{y}} \mathcal{L}_l(\mathbf{x} \cdot \mathbf{z}) = 4\pi i^l j_l(xy) \mathcal{L}_l(\hat{\mathbf{y}} \cdot \hat{\mathbf{z}}), \quad (2.36)$$

where x is the module of the vector $\mathbf{x}$, y is the module of the vector $\mathbf{y}$ and $j_l(xy)$ denotes the spherical Bessel function of order l in the variable xy . In addition to this, note that

$$\mathcal{L}_l(-x) = (-1)^l \mathcal{L}_l(x), \quad (2.37)$$

which is going to be useful in the next steps of the derivation. This property means that the Legendre polynomial of degree n is an even function if n is an even number, and it is an odd function if n is an odd number. Now, continuing from equation (2.35) and taking

advantage of the identity from equation (2.36) we can obtain

$$\begin{aligned}
& \frac{1}{(2\pi)^3} \int dk k^2 \sum_{l=0}^{\infty} \Gamma_l(k) \int d\Omega_k \mathcal{L}_l(\mu_k) e^{-i\mathbf{k} \cdot \mathbf{r}} = \\
& = \frac{1}{(2\pi)^3} \int dk k^2 \sum_{l=0}^{\infty} \Gamma_l(k) 4\pi i^l j_l(kr) \mathcal{L}_l(-\hat{\mathbf{r}} \cdot \hat{\mathbf{z}}) = \\
& = \frac{1}{(2\pi)^3} \int dk k^2 \sum_{l=0}^{\infty} \Gamma_l(k) 4\pi (-i)^l j_l(kr) \mathcal{L}_l(\hat{\mathbf{r}} \cdot \hat{\mathbf{z}}) = \\
& = \frac{1}{(2\pi)^3} \int dk k^2 \sum_{l=0}^{\infty} \frac{2l+1}{2} \int d\mu_k \tilde{K}_a(\mathbf{k}) \tilde{K}_b(-\mathbf{k}) \mathcal{L}_l(\mu_k) 4\pi (-i)^l j_l(kr) \mathcal{L}_l(\hat{\mathbf{r}} \cdot \hat{\mathbf{z}}), \quad (2.38)
\end{aligned}$$

where we have also used the property of the Legendre polynomials given by equation (2.37) and the expression of the coefficients $\Gamma_l(k)$ given by equation (2.34). It is possible to develop both expressions $\tilde{K}_a(\mathbf{k})$ and $\tilde{K}_b(-\mathbf{k})$ in equation (2.38) as a sum of Legendre polynomials as

$$\tilde{K}_a(\mathbf{k}) = \sum_{l'=0}^{\infty} \tilde{K}_a^{(l')}(k) \mathcal{L}_{l'}(\mu_k) \quad (2.39)$$

$$\tilde{K}_b(-\mathbf{k}) = \sum_{l''=0}^{\infty} \tilde{K}_b^{(l'')}(k) \mathcal{L}_{l''}(-\mu_k), \quad (2.40)$$

where the coefficients $\tilde{K}_a^{(l')}(k)$ and $\tilde{K}_b^{(l'')}(k)$ are given by

$$\tilde{K}_a^{(l')}(k) = \frac{2l'+1}{2} \int d\mu_k \tilde{K}_a(\mathbf{k}) \mathcal{L}_{l'}(\mu_k) \quad (2.41)$$

$$\tilde{K}_b^{(l'')}(k) = \frac{2l''+1}{2} \int d\mu_k \tilde{K}_b(-\mathbf{k}) \mathcal{L}_{l''}(-\mu_k). \quad (2.42)$$

We can now substitute these expansions in Legendre polynomials in equation (2.38) and, reordering the sums, we obtain

$$\begin{aligned}
& \frac{4\pi}{(2\pi)^3} \int dk k^2 \sum_{l=0}^{\infty} \sum_{l'=0}^{\infty} \sum_{l''=0}^{\infty} \frac{2l+1}{2} \times \\
& \times \int d\mu_k \tilde{K}_a^{(l')}(k) \mathcal{L}_{l'}(\mu_k) \tilde{K}_b^{(l'')}(k) (-1)^{l''} \mathcal{L}_{l''}(\mu_k) \mathcal{L}_l(\mu_k) (-i)^l j_l(kr) \mathcal{L}_l(\mu) = \\
& = \frac{1}{2\pi^2} \int dk k^2 \sum_{l=0}^{\infty} \sum_{l'=0}^{\infty} \sum_{l''=0}^{\infty} \frac{2l+1}{2} \times \\
& \times \tilde{K}_a^{(l')}(k) \tilde{K}_b^{(l'')}(k) (-1)^{l''} (-i)^l j_l(kr) \mathcal{L}_l(\mu) \int d\mu_k \mathcal{L}_{l'}(\mu_k) \mathcal{L}_{l''}(\mu_k) \mathcal{L}_l(\mu_k). \quad (2.43)
\end{aligned}$$

At this point, we can make use of the property proved in [19] which allows to evaluate the definite integral of the product of three Legendre polynomials, giving the result

$$\int d\mu_k \mathcal{L}_{l'}(\mu_k) \mathcal{L}_{l''}(\mu_k) \mathcal{L}_l(\mu_k) = 2 \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix}^2, \quad (2.44)$$

where $\begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix}$ is a Wigner's 3-j symbol, whose general expression is also shown in [19]. There, the integral is done in the variable θ_k , but it is analogous in the variable μ_k because

$d\mu_k = -\sin\theta d\theta$ and there is another minus one coming from the change of variable in the limits of integration. It is important to highlight that, since $m_1 = m_2 = m_3 = 0$ in the expression of the general Wigner's 3-j symbol, in order the Wigner's 3-j symbol in equation (2.44) to be different from zero, it is needed the sum $l + l' + l''$ to be an even integer. Hence, this Wigner's 3-j symbol is invariant under any permutation of its indexes, as we should expect since it is obtained from a definite integral of a product of three polynomials, whose result does not change if the order of the polynomials is changed.

Therefore, substituting in equation (2.43) the result of the integral from equation (2.44), we can write

$$\begin{aligned} \int d^3\mathbf{x} K_a(\mathbf{x})K_b(\mathbf{x} + \mathbf{r}) &= \\ &= \frac{1}{2\pi^2} \int dk k^2 \sum_{l=0}^{\infty} \sum_{l'=0}^{\infty} \sum_{l''=0}^{\infty} (2l+1) \tilde{K}_a^{(l')}(k) \tilde{K}_b^{(l'')}(k) (-1)^{l''} (-i)^l j_l(kr) \mathcal{L}_l(\mu) \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix}^2. \end{aligned} \quad (2.45)$$

Regarding the functions $\tilde{K}_a^{(l')}(k)$ and $\tilde{K}_b^{(l'')}(k)$, they are the coefficients of the decomposition in Legendre polynomials of the functions $\tilde{K}_a(\mathbf{k})$ and $\tilde{K}_b(-\mathbf{k})$, respectively, as shown in equation (2.39) and equation (2.40). $K_a(\mathbf{x})$ and $K_b(\mathbf{x})$ are both even functions with respect to the variable μ since they are functions of $T(\mathbf{x})$. Hence, $\tilde{K}_a(\mathbf{k})$ and $\tilde{K}_b(-\mathbf{k})$, which are their direct Fourier transforms, are both even functions, too, because the Fourier transform of an even function is an even function. Therefore, taking into account the expressions of the coefficients $\tilde{K}_a^{(l')}(k)$ and $\tilde{K}_b^{(l'')}(k)$ of equation (2.41) and equation (2.42), these are going to be equal to zero if l' and l'' are both odd integers since then the Legendre polynomials $\mathcal{L}_{l'}(\mu_k)$ and $\mathcal{L}_{l''}(-\mu_k)$ would be both odd functions, so that the integrals in those equations would be zero, considering that the variable μ_k goes from minus one to one. In conclusion, in order to be the expression in equation (2.45) different from zero, we need l' and l'' to be both even integers.

From now on, the indexes l' and l'' will be considered to be both even integers, and since $l + l' + l''$ must be an even integer so that the Wigner's 3-j symbol is different from zero, the index l will be even, too. Hence, the expression in equation (2.45) can be written as

$$\begin{aligned} \int d^3\mathbf{x} K_a(\mathbf{x})K_b(\mathbf{x} + \mathbf{r}) &= \\ &= \frac{1}{2\pi^2} \int dk k^2 \sum_{l=0}^{\infty} \sum_{l'=0}^{\infty} \sum_{l''=0}^{\infty} (2l+1) \tilde{K}_a^{(l')}(k) \tilde{K}_b^{(l'')}(k) i^l j_l(kr) \mathcal{L}_l(\mu) \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix}^2. \end{aligned} \quad (2.46)$$

This expression is a function of the position $\mathbf{r}$, so that we can denote it as $B(\mathbf{r})$ and we can decompose it as a sum of Legendre polynomials, leading to

$$\begin{aligned} \int d^3\mathbf{x} K_a(\mathbf{x})K_b(\mathbf{x} + \mathbf{r}) &= \\ &= \frac{1}{2\pi^2} \int dk k^2 \sum_{l=0}^{\infty} \sum_{l'=0}^{\infty} \sum_{l''=0}^{\infty} (2l+1) \tilde{K}_a^{(l')}(k) \tilde{K}_b^{(l'')}(k) i^l j_l(kr) \mathcal{L}_l(\mu) \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix}^2 = \\ &= B(\mathbf{r}) = \sum_{l=0}^{\infty} B^{(l)}(r) \mathcal{L}_l(\mu), \end{aligned} \quad (2.47)$$

where r denotes the module of the vector $\mathbf{r}$. Moreover, since the Legendre polynomials form a basis of the functions defined in the interval $[-1, 1]$, it is possible to obtain the coefficients $B^{(l)}(r)$ of the decomposition of $B(\mathbf{r})$ directly as

$$B^{(l)}(r) = \frac{2l+1}{2\pi^2} i^l \sum_{l'=0}^{\infty} \sum_{l''=0}^{\infty} \int dk k^2 \tilde{K}_a^{(l')}(k) \tilde{K}_b^{(l'')}(k) j_l(kr) \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix}^2 \quad (2.48)$$

From here, in order to compute equation (2.48) in an easier way, we can obtain relations between the coefficients $\tilde{K}_a^{(l')}(k)$ and $\tilde{K}_b^{(l'')}(k)$ and their counterparts in configuration space, $K_a^{(l')}(x)$ and $K_b^{(l'')}(x)$, respectively. To do this, consider a general function depending on the vector $\mathbf{x}$, $K(\mathbf{x})$, and take its direct Fourier transform, given by

$$\begin{aligned} \tilde{K}(\mathbf{k}) &= \tilde{K}(k, \mu_k) = \int d^3\mathbf{x} K(x, \mu) e^{-i\mathbf{k}\cdot\mathbf{x}} = \int dx x^2 \sum_{l=0}^{\infty} K^{(l)}(x) \int d\Omega_x e^{-i\mathbf{k}\cdot\mathbf{x}} \mathcal{L}_l(\mu) = \\ &= \int dx x^2 \sum_{l=0}^{\infty} K^{(l)}(x) 4\pi(-i)^l j_l(kx) \mathcal{L}_l(\mu_k), \end{aligned} \quad (2.49)$$

where the function $K(\mathbf{x})$ has been decomposed in Legendre polynomials as $K(\mathbf{x}) = \sum_{l=0}^{\infty} K^{(l)}(x) \mathcal{L}_l(\mu)$ and we have also used the property in equation (2.36). Developing also $\tilde{K}(\mathbf{k}) = \sum_{l=0}^{\infty} \tilde{K}^{(l)}(k) \mathcal{L}_l(\mu_k)$ as a sum of Legendre polynomials and comparing with equation (2.49), we get

$$\tilde{K}^{(l)}(k) = 4\pi(-i)^l \int dx x^2 K^{(l)}(x) j_l(kx), \quad (2.50)$$

which gives the relation desired between the coefficients in Fourier space and the coefficients in configuration space. Therefore, we obtain

$$\tilde{K}_a^{(l')}(k) = 4\pi(-i)^{l'} \int dx x^2 K_a^{(l')}(x) j_{l'}(kx) \quad (2.51)$$

and analogously we obtain $\tilde{K}_b^{(l'')}(k)$. Considering that l' and l'' are both even integers and substituting the expressions of $\tilde{K}_a^{(l')}(k)$ and $\tilde{K}_b^{(l'')}(k)$ in equation (2.48), we can express $B^{(l)}(r)$ as

$$\begin{aligned} B^{(l)}(r) &= \frac{2l+1}{2\pi^2} i^l \sum_{l'=0}^{\infty} \sum_{l''=0}^{\infty} \int dk k^2 \int dx x^2 K_a^{(l')}(x) j_{l'}(kx) 4\pi(-i)^{l'} \times \\ &\times \int dx' x'^2 K_b^{(l'')}(x') j_{l''}(kx') 4\pi(-i)^{l''} j_l(kr) \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix}^2 = \\ &= 8(2l+1) i^l \sum_{l'=0}^{\infty} \sum_{l''=0}^{\infty} i^{l'+l''} \times \\ &\times \int dx dx' x^2 x'^2 K_a^{(l')}(x) K_b^{(l'')}(x') \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix}^2 \int dk k^2 j_l(kr) j_{l'}(kx) j_{l''}(kx'). \end{aligned} \quad (2.52)$$

As we can observe, an integral of a product of three Bessel functions appears. This integral can be computed using the theory developed in [20], where it is shown that

it can be expressed in terms of a sum of associated Legendre polynomials. This way, the convolution term in equation (2.29) can be computed calculating the coefficients in equation (2.52) and making the sum multiplying by the Legendre polynomials in equation (2.47).

2.2.2 Two-halo term

Now, we are going to find an expression for the two-halo term of the LIM two-point PDF of the brightness temperature. As described before, the two-halo term accounts for the possibilities for which the two different positions of temperatures T and T' belong to two different halos of masses M and M' , respectively. This situation can be observed in Figure 2.1.

Following the same procedure as in modelling the one-point PDF and the one-halo term of the two-point PDF and since we are now dealing with two different halos, we denote by $\boldsymbol{\vartheta}$ the set of astrophysical properties of the halo of mass M and by $\boldsymbol{\vartheta}'$ the set of astrophysical properties of the halo of mass M' . We consider two infinitesimal mass bins, one centered at M and another centered at M' , so that the two-halo term of the two-point PDF of the brightness temperature for the bins is

$$\begin{aligned} \mathcal{P}_{2h}^{(M,M')}(T, T'; \mathbf{r}) = & \int d\boldsymbol{\vartheta} d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}(\boldsymbol{\vartheta}'|M') \times \\ & \times \left[\mathcal{P}_{N=0}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')} \delta_D(T) \delta_D(T') + \mathcal{P}_{N=2}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')} \mathcal{P}_1^{(M,\boldsymbol{\vartheta})}(T) \mathcal{P}_1^{(M',\boldsymbol{\vartheta}')} (T') \right]. \end{aligned} \quad (2.53)$$

In this equation, we have made some assumptions. Firstly, we marginalize over the two conditional multidimensional distributions $\mathcal{P}(\boldsymbol{\vartheta}|M)$ and $\mathcal{P}(\boldsymbol{\vartheta}'|M')$ to account for the astrophysical properties of both halos. Secondly, we assume that $\mathcal{P}_{N=1}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')}$ is equal to zero, because as defined before this is the PDF of having one emitter contributing to a specific point, but in this term we consider two emitters (halos) contributing. That is the reason why we now have $\mathcal{P}_{N=2}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')}$, which is the PDF of measuring, given a temperature T in a specific position within a halo of mass M and set of properties $\boldsymbol{\vartheta}$, another temperature T' separated by a distance vector $\mathbf{r}$ from T within a different halo of mass M' and set of properties $\boldsymbol{\vartheta}'$. Thirdly, we assume that the probability of observing a temperature within one halo is independent of the probability of observing another temperature within another halo, hence $\mathcal{P}_2^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')} (T, T'; \mathbf{r}) = \mathcal{P}_1^{(M,\boldsymbol{\vartheta})}(T) \mathcal{P}_1^{(M',\boldsymbol{\vartheta}')} (T')$. Besides these assumptions, if there is no emitter, this is $N = 0$, then there is no signal and $\mathcal{P}_0^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')} = \delta_D(T) \delta_D(T')$ is the product of two Dirac deltas both centered at zero.

Furthermore, since we are assuming in this term that the two temperatures T and T' are observed within two different halos, for two infinitesimal mass bins we can consider $\mathcal{P}_{N>2}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')} \equiv 0$ so that, since we also have that $\mathcal{P}_{N=1}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')}$ is equal to zero, $\mathcal{P}_{N=0}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')} = 1 - \mathcal{P}_{N=2}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')}(\mathbf{r})$. Therefore, we can write $\mathcal{P}_{2h}^{(M,M')}(T, T'; \mathbf{r})$ as

$$\begin{aligned}
\mathcal{P}_{2h}^{(M,M')}(T, T'; \mathbf{r}) &= \int d\boldsymbol{\vartheta} d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}(\boldsymbol{\vartheta}'|M') \times \\
&\times \left[\left(1 - \mathcal{P}_{N=2}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')}\right) \delta_D(T) \delta_D(T') + \mathcal{P}_{N=2}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')}\mathcal{P}_1^{(M,\boldsymbol{\vartheta})}(T) \mathcal{P}_1^{(M',\boldsymbol{\vartheta}')} (T') \right] = \\
&= \int d\boldsymbol{\vartheta} d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}(\boldsymbol{\vartheta}'|M') \delta_D(T) \delta_D(T') + \int d\boldsymbol{\vartheta} d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}(\boldsymbol{\vartheta}'|M') \mathcal{P}_{N=2}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')}\times \\
&\times \left(\mathcal{P}_1^{(M,\boldsymbol{\vartheta})}(T) \mathcal{P}_1^{(M',\boldsymbol{\vartheta}')} (T') - \delta_D(T) \delta_D(T') \right). \tag{2.54}
\end{aligned}$$

With this, taking into account that $\int d\boldsymbol{\vartheta} d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}(\boldsymbol{\vartheta}'|M') = 1$ and calculating the Fourier transform of $\mathcal{P}_{2h}^{(M,M')}(T, T'; \mathbf{r})$, we obtain

$$\begin{aligned}
\tilde{\mathcal{P}}_{2h}^{(M,M')}(\tau_a, \tau_b; \mathbf{r}) &= 1 + \int d\boldsymbol{\vartheta} d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}(\boldsymbol{\vartheta}'|M') \mathcal{P}_{N=2}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta}')}\times \\
&\times \left(\tilde{\mathcal{P}}_1^{(M,\boldsymbol{\vartheta})}(\tau_a) \tilde{\mathcal{P}}_1^{(M',\boldsymbol{\vartheta}')}(\tau_b) - 1 \right). \tag{2.55}
\end{aligned}$$

In this equation, as in the one-halo term τ_a is the Fourier conjugate of T and τ_b is the Fourier conjugate of T' .

In relation to $\tilde{\mathcal{P}}_1^{(M,\boldsymbol{\vartheta})}(\tau_a)$ and $\tilde{\mathcal{P}}_1^{(M',\boldsymbol{\vartheta}')}(\tau_b)$, these are the Fourier transforms of the one-point PDFs of T and T' within the halo of mass M and within the halo of mass M' , respectively. Consequently, we can calculate them as we did when we modeled the one-point PDF, this is

$$\begin{aligned}
\tilde{\mathcal{P}}_1^{(M,\boldsymbol{\vartheta})}(\tau_a) &= \int d^3\mathbf{x} dL \mathcal{P}(\mathbf{x}|M, \boldsymbol{\vartheta}) \mathcal{P}(L|M, \boldsymbol{\vartheta}) e^{-iT(\mathbf{x}-\mathbf{x}_1)\tau_a} = \\
&= \frac{1}{V_{\text{prof}}} \int_{V_{\text{prof}}} d^3\mathbf{x} dL \mathcal{P}(L|M, \boldsymbol{\vartheta}) e^{-iT(\mathbf{x}-\mathbf{x}_1)\tau_a} \tag{2.56}
\end{aligned}$$

and analogously for $\tilde{\mathcal{P}}_1^{(M',\boldsymbol{\vartheta}')}(\tau_b)$. This way, V_{prof} is the volume which the signal profile corresponding to T covers, V'_{prof} is the volume which the signal profile corresponding to T' covers, $\mathbf{x}_1$ is the center position of the halo of mass M and $\mathbf{x}_2$ is the center position of the halo of mass M' . There exists a difference with respect to the analogous definition in the one-point PDF, which is that in each case we center the temperature profile in the position where it is the center of the halo. Besides that, we marginalize over the position, for which $\mathcal{P}(\mathbf{x}|M, \boldsymbol{\vartheta})$ is uniform over V_{prof} and zero otherwise and $\mathcal{P}(\mathbf{x}'|M', \boldsymbol{\vartheta}')$ is uniform over V'_{prof} and zero otherwise. We also marginalize over the line luminosity, for which $\mathcal{P}(L|M, \boldsymbol{\vartheta})$ and $\mathcal{P}(L'|M', \boldsymbol{\vartheta}')$ account for any distribution of the line luminosity given the halo mass and the astrophysical properties, in each case.

Now, assuming Poisson statistics, we can define $\mathcal{P}_{N=2}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta'})}$ in equation (2.55) as

$$\begin{aligned}
\mathcal{P}_{N=2}^{(M,M',\boldsymbol{\vartheta},\boldsymbol{\vartheta'})} &= dM \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) dM' \frac{dn}{dM'} V'_{\text{prof}}(M', \boldsymbol{\vartheta}') \times \\
&\times \int d^3\mathbf{x}_1 \int d^3\mathbf{x}_2 \mathcal{P}(\mathbf{x} - \mathbf{x}_1, M, \boldsymbol{\vartheta}) \mathcal{P}(\mathbf{x}' - \mathbf{x}_2, M', \boldsymbol{\vartheta}') (1 + \xi(\mathbf{x}_1 - \mathbf{x}_2)). \tag{2.57}
\end{aligned}$$

This is, as stated before, the probability of measuring a temperature T' within a halo of mass M' given a temperature T in a specific position within a halo of mass M , with T and

T' separated by a distance vector $\mathbf{r} = \mathbf{x} - \mathbf{x}'$. In equation (2.57), the term $1 + \xi(\mathbf{x}_1 - \mathbf{x}_2)$ accounts for the possibilities of finding two halos separated by a distance vector $\mathbf{x}_1 - \mathbf{x}_2$, so that it is equal to one if we do not consider clustering and the distribution of the halos is therefore uniform, since $\xi(\mathbf{x}_1 - \mathbf{x}_2)$ is the halo two-point correlation function. Note that if we consider the window functions W_{vox} and W_{smooth} as we explained in the one-point PDF section, since the power spectrum is modified due to the effect of these functions and becomes to be $P_v(\mathbf{k})$, taking into account equation (1.10) the two-point correlation function would be affected by these window functions, too. Moreover, in equation (2.57) we integrate over $\mathbf{x}_1$ and $\mathbf{x}_2$ to account for any position of the halos of masses M and M' , and we multiply by $\mathcal{P}(\mathbf{x} - \mathbf{x}_1, M, \boldsymbol{\vartheta})$ and $\mathcal{P}(\mathbf{x}' - \mathbf{x}_2, M', \boldsymbol{\vartheta}')$, which are given by

$$\mathcal{P}(\mathbf{x} - \mathbf{x}_1, M, \boldsymbol{\vartheta}) = \begin{cases} 1/V_{\text{prof}} & \text{if } \|\mathbf{x} - \mathbf{x}_1\| \leq V_{\text{prof}} \\ 0 & \text{otherwise} \end{cases} \quad (2.58)$$

$$\mathcal{P}(\mathbf{x}' - \mathbf{x}_2, M', \boldsymbol{\vartheta}') = \begin{cases} 1/V'_{\text{prof}} & \text{if } \|\mathbf{x}' - \mathbf{x}_2\| \leq V'_{\text{prof}}, \\ 0 & \text{otherwise,} \end{cases} \quad (2.59)$$

so that they act as window functions over V_{prof} and V'_{prof} , respectively. Furthermore, we can assume for infinitesimal bins that $\mathcal{P}_{N=2}^{(M, M', \boldsymbol{\vartheta}, \boldsymbol{\vartheta}')} \ll 1$, so it is possible to interpret equation (2.55) as the linear expansion of the exponential and write

$$\begin{aligned} \tilde{\mathcal{P}}_{2h}^{(M, M')}(\tau_a, \tau_b; \mathbf{r}) &= \\ &= \exp \left\{ \int d\boldsymbol{\vartheta} d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}(\boldsymbol{\vartheta}'|M') \mathcal{P}_{N=2}^{(M, M', \boldsymbol{\vartheta}, \boldsymbol{\vartheta}')} \left(\mathcal{P}_1^{(M, \boldsymbol{\vartheta})}(T) \mathcal{P}_1^{(M', \boldsymbol{\vartheta}')} (T') - 1 \right) \right\}. \end{aligned} \quad (2.60)$$

At this point, we need to make an assumption in order to obtain a computable expression for $\tilde{\mathcal{P}}_{2h}^{(M, M')}(\tau_a, \tau_b; \mathbf{r})$. Following the argument described in [7], if $\xi(\mathbf{r})$ varies slowly on scales which are larger than the typical extent of a halo, as we assume in this work, $\mathcal{P}(\mathbf{x} - \mathbf{x}_1, M, \boldsymbol{\vartheta})$ and $\mathcal{P}(\mathbf{x}' - \mathbf{x}_2, M', \boldsymbol{\vartheta}')$ can be considered to be Dirac delta functions centered each one at the respective center position of the halo as they become to be extremely sharpened compared to the scales in which $\xi(\mathbf{r})$ varies. Hence, we can assume $\mathcal{P}(\mathbf{x} - \mathbf{x}_1, M, \boldsymbol{\vartheta}) \approx \delta_D(\mathbf{x} - \mathbf{x}_1)$ and $\mathcal{P}(\mathbf{x}' - \mathbf{x}_2) \approx \delta_D(\mathbf{x}' - \mathbf{x}_2)$. Moreover, like in the one-point PDF and the one-halo term of the two-point PDF, $\tilde{\mathcal{P}}_{2h}^{(M, M')}(\tau_a, \tau_b; \mathbf{r})$ is the two-halo term of the two-point PDF of only two halos of masses M and M' . In order to obtain the two-halo term of the two-point PDF of the whole population of halos, we need to make a convolution with each of the masses M and M' we can have. Therefore, as we are working in Fourier space, we need to multiply by the PDF of each of the masses we can have. This means we obtain

$$\begin{aligned}
\tilde{\mathcal{P}}_{2h}(\tau_a, \tau_b; \mathbf{r}) &= \prod \tilde{\mathcal{P}}_{2h}^{(M, M')}(\tau_a, \tau_b; \mathbf{r}) = \\
&= \exp \left\{ \int dM dM' d\boldsymbol{\vartheta} d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}(\boldsymbol{\vartheta}'|M') \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \frac{dn}{dM'} V'_{\text{prof}}(M', \boldsymbol{\vartheta}') \times \right. \\
&\times \int d^3\mathbf{x}_1 d^3\mathbf{x}_2 \left[\left(\int_{V_{\text{prof}}, V'_{\text{prof}}} d^3\mathbf{x} d^3\mathbf{x}' \mathcal{P}(\mathbf{x} - \mathbf{x}_1, M, \boldsymbol{\vartheta}) \mathcal{P}(\mathbf{x}' - \mathbf{x}_2, M', \boldsymbol{\vartheta}') \times \right. \right. \\
&\times (1 + \xi(\mathbf{x}_1 - \mathbf{x}_2)) \frac{e^{-iT(\mathbf{x} - \mathbf{x}_1)\tau_a} e^{-iT(\mathbf{x}' - \mathbf{x}_2)\tau_b}}{V_{\text{prof}} V'_{\text{prof}}} \Big) - \\
&\left. \left. - \mathcal{P}(\mathbf{x} - \mathbf{x}_1, M, \boldsymbol{\vartheta}) \mathcal{P}(\mathbf{x}' - \mathbf{x}_2, M', \boldsymbol{\vartheta}') (1 + \xi(\mathbf{x}_1 - \mathbf{x}_2)) \right] \right\} = \\
&= \exp \left\{ \int dM dM' d\boldsymbol{\vartheta} d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}(\boldsymbol{\vartheta}'|M') \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \frac{dn}{dM'} V'_{\text{prof}}(M', \boldsymbol{\vartheta}') \times \right. \\
&\times \left[\int_{V_{\text{prof}}, V'_{\text{prof}}} d^3\mathbf{x} d^3\mathbf{x}' (1 + \xi(\mathbf{r})) \frac{e^{-iT_{\mathbf{x}=\mathbf{x}_1}(M)\tau_a} e^{-iT_{\mathbf{x}'=\mathbf{x}_2}(M')\tau_b}}{V_{\text{prof}} V'_{\text{prof}}} - (1 + \xi(\mathbf{r})) \right] \Big\} = \\
&= \exp \left\{ \int dM dM' d\boldsymbol{\vartheta} d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}|M) \mathcal{P}(\boldsymbol{\vartheta}'|M') \frac{dn}{dM} V_{\text{prof}}(M, \boldsymbol{\vartheta}) \frac{dn}{dM'} V'_{\text{prof}}(M', \boldsymbol{\vartheta}') \times \right. \\
&\times (e^{-iT_{\mathbf{x}=\mathbf{x}_1}(M)\tau_a} e^{-iT_{\mathbf{x}'=\mathbf{x}_2}(M')\tau_b} - 1) (1 + b_h(M) b_h(M') \xi_m(\mathbf{r})) \Big\}. \tag{2.61}
\end{aligned}$$

In the second equality we have substituted the expressions for $\mathcal{P}_{N=2}^{(M, M', \boldsymbol{\vartheta}, \boldsymbol{\vartheta}')} , \tilde{\mathcal{P}}_1^{(M, \boldsymbol{\vartheta})}(\tau_a)$ and $\tilde{\mathcal{P}}_1^{(M', \boldsymbol{\vartheta}')}(\tau_b)$, and in the third equality we have first integrated over $\mathbf{x}_1$ and $\mathbf{x}_2$ considering that $\mathcal{P}(\mathbf{x} - \mathbf{x}_1, M, \boldsymbol{\vartheta})$ and $\mathcal{P}(\mathbf{x}' - \mathbf{x}_2, M', \boldsymbol{\vartheta}')$ are assumed to be Dirac delta functions. Therefore, $\xi(\mathbf{x}_1 - \mathbf{x}_2)$ becomes $\xi(\mathbf{r}) = \xi(\mathbf{x} - \mathbf{x}')$, $T(\mathbf{x} - \mathbf{x}_1)$ becomes $T_{\mathbf{x}=\mathbf{x}_1}(M)$, which is the temperature measured in the center of the halo of mass M , and $T(\mathbf{x}' - \mathbf{x}_2)$ becomes $T_{\mathbf{x}'=\mathbf{x}_2}(M')$, which is the temperature measured in the center of the halo of mass M' . Finally, in the last equality we have integrated over $\mathbf{x}$ and $\mathbf{x}'$, which allows to simplify V_{prof} and V'_{prof} since the integrand does not depend on either $\mathbf{x}$ or $\mathbf{x}'$. Besides that, we have related the two-point correlation function $\xi(\mathbf{r})$ to the underlying matter correlation function $\xi_m(\mathbf{r})$ with a linear, mass-dependent halo bias as we did in the one-point PDF. This means we have considered

$$\xi(\mathbf{r}) = b_h(M) b_h(M') \xi_m(\mathbf{r}).$$

Finally, regarding the expression obtained in the last equality of equation (2.61), we can define some quantities to make it more compact. Firstly, we define four quantities, which have the subscript a if they depend on the variable τ_a and have the subscript b if they depend on the variable τ_b . We define

$$\alpha_a := \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}} b_h(M) e^{-iT_{\mathbf{x}=\mathbf{x}_1}(M)\tau_a} \tag{2.62}$$

and analogously for α_b , and we define

$$\beta_a := \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}} e^{-iT_{\mathbf{x}=\mathbf{x}_1}(M)\tau_a} \tag{2.63}$$

and analogously for β_b . Secondly, we define an integrated bias over the halo mass M in V_{prof} as

$$b_{\text{tot}} := \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}} b_h(M) \quad (2.64)$$

and analogously we define an integrated bias b'_{tot} over the halo mass M' in V'_{prof} . Therefore, we can write equation (2.61) as

$$\begin{aligned} \tilde{\mathcal{P}}_{2h}(\tau_a, \tau_b; \mathbf{r}) &= \exp \left\{ \beta_a \beta_b - \int dM d\boldsymbol{\vartheta} \mathcal{P}(\boldsymbol{\vartheta}|M) \frac{dn}{dM} V_{\text{prof}} \int dM' d\boldsymbol{\vartheta}' \mathcal{P}(\boldsymbol{\vartheta}'|M') \frac{dn'}{dM'} V'_{\text{prof}} + \right. \\ &\quad \left. + \alpha_a \alpha_b \xi_m(\mathbf{r}) - b_{\text{tot}} b'_{\text{tot}} \xi_m(\mathbf{r}) \right\} = \\ &= \exp \left\{ \beta_a \beta_b - \langle n V_{\text{prof}} \rangle \langle n' V'_{\text{prof}} \rangle + \alpha_a \alpha_b \xi_m(\mathbf{r}) - b_{\text{tot}} b'_{\text{tot}} \xi_m(\mathbf{r}) \right\}, \end{aligned}$$

where n and n' are the number of galaxies per unit of volume in V_{prof} and V'_{prof} , respectively.

Once we have calculated both the one-halo term and the two-halo term of the two-point PDF in Fourier space, we can simply multiply them to obtain the global two-point brightness temperature PDF and make the inverse Fourier transform in both variables τ_a and τ_b to obtain the two-point PDF in configuration space, this is

$$P(T, T'; \mathbf{r}) = \frac{1}{(2\pi)^2} \int d\tau_a d\tau_b \tilde{\mathcal{P}}(\tau_a, \tau_b; \mathbf{r}) e^{iT\tau_a} e^{iT'\tau_b}. \quad (2.65)$$

Chapter 3

Computational implementation

In this chapter we are going to explain first how we computationally implement the theoretical model developed (in previous chapter) in order to calculate the PDF of the brightness temperature measured in LIM experiments. Second, we describe how we make realizations of the population of galaxies for a projected field on the sky, so that we can later obtain the PDF from them and compare the realizations with the predictions made by our theoretical model, in order to validate whether it is correct or not, which is our final goal.

Note that, although the realizations made can be used to validate both the theoretical model for the one-point PDF and the theoretical model for the two-point PDF, in this work we restrict ourselves to analyze only the one-point PDF of the brightness temperature, due to the scope of this work and the time available. The two-point PDF will remain simply as a theoretical model developed in previous chapter, and its validation through realizations will be left for future work.

3.1 Computational approach

In this section, we explain how we manage to computationally approach our work, describing the assumptions made and the cosmological parameters considered.

First highlight that, although the theoretical model developed in previous chapter is valid for computing the PDF of the brightness temperature in a three-dimensional region of the Universe, in order to facilitate its computational implementation and validation with realizations we will consider a projection in two dimensions of a three-dimensional slice of the Universe, eliminating the dimension along the line of sight. In order to do this, we will need to adapt the derivation in previous chapter to the two-dimensional case. Furthermore, we will assume several properties (simplifications) that make the implementation simpler and faster. These assumptions are that all galaxies in the sky have the same luminosity, the same mass, the same bias relating galaxy density and matter density and the same redshift (they are observed from the same distance).

Similar to what is done in [8], the derivation in previous chapter allows computing the PDF of the brightness temperature in a specific point of the Universe (or of the sky, as it is done in this work). In practice, however, we work with a discretized map (it is actually the only thing we can measure) in which each cell or voxel (two- or three-dimensional pixels, depending on if we work with a region of the Universe or with the sky) corresponds to a comoving volume, and where we measure the brightness temperature from observations. As we will see, discretizing the map will be relevant in several stages of the computational

implementation of the model, as it has been shown in previous chapter when we considered the pixel window function affecting the computation of the zero-lag variance and the power spectrum, for example.

The VID, as explained before ([8]), depends not only on the intrinsic signal but also on the experimental setup of the LIM experiment, through the noise, the resolution and the voxel volume, for example. For this reason, it is crucial to take into account the experimental limitations the LIM survey has, giving values to its defining parameters. Particularly, in this work we consider a spectral line emitted at $\nu_{\text{emit}} = 115.271$ GHz (it would correspond to the CO(1-0) line) and observed at $\nu_{\text{obs}} = 76.847$ GHz ($z = 0.5$) with $\theta_{\text{FWHM}} = 1$ deg, which corresponds to the angular resolution. Since we are considering a slice of the Universe projected on the sky, eliminating the information along the line of sight, we do not have any spectral resolution. The redshift width of the slice of the Universe considered is $DZ = 0.01$, which will be useful to calculate comoving volumes. Moreover, we consider a noise-per-voxel standard deviation of $\sigma_N = 0.2 \mu\text{K}$. Regarding the line we focus on, we assume we can distinguish it from any other, so that we do not have any line interloper nor correlations with other emission lines (no contaminants other than noise).

At this point, note that in order to make the realizations we need to obtain the distribution of the brightness temperature in the sky, we will use Hierarchical Equal Area isoLatitude Pixelization (HEALPIX) of the sphere. In this pixelization, the sphere is tessellated into curvilinear quadrilaterals in a way that all pixels at a given resolution have the same area and allowing to make faster computations than other pixelizations. The lowest resolution partition is comprised of twelve base pixels, and the resolution increases dividing each pixel into more new ones. At each resolution increment, each pixel is divided into a grid of $N_{\text{side}} \times N_{\text{side}}$ new pixels. Therefore, N_{side} is a number that characterizes the resolution of each pixelization, and is always a power of two. In this way, the total number of pixels in which the sphere is divided at each resolution is $N_{\text{pix}} = 12N_{\text{side}}^2$. Particularly, in this work we will consider $N_{\text{side}} = 256$, so that $N_{\text{pix}} = 786432$. With this pixelization, the side length of each pixel is $\text{pix_side} = 13.751$ arcmin.

Through all these parameters, we obtain the following characterization of the distribution of galaxies in the sky. First, the mean number density of galaxies is $n = 0.125$ galaxies/Mpc³, and the linear bias relating galaxy density with matter density is $b_h = 1.362$. Second, the mean temperature of the sky is $T_{\text{mean}} = 1.863 \mu\text{K}$. Taking into account the redshift and the redshift width of the slice of the Universe considered, we get a comoving volume of the full sky of $V = 4\pi\chi^2(z)[\chi(z + DZ/2) - \chi(z - DZ/2)] = 1.612 \cdot 10^9$ Mpc³, hence a total number of galaxies equal to $N_{\text{gal}} = 2.015 \cdot 10^8$ galaxies. In addition to this, regarding the pixelization chosen, we obtain a mean number density of galaxies per pixel of $\bar{n} = 255.668$ galaxies/pixel, hence each galaxy has a luminosity of $L_{\text{gal}} = 29006.41 L_{\odot}$, where $L_{\odot}$ refers to the luminosity of the Sun. Finally, each pixel has a volume of $V_{\text{pix}} = 2049.821$ Mpc³, which allows to calculate the total contribution of each galaxy to the temperature measured in a pixel through equation (1.13), obtaining $T_{\text{tot}} = 0.007 \mu\text{K}$.

Lastly, since we are working with the two-dimensional sky instead of with a three-dimensional volume, the power spectrum explained in Chapter 1 and Chapter 2 becomes to be slightly different because now it measures the variance of a density distribution in two dimensions instead of three. This new power spectrum is known as angular power spectrum. As explained in [3], if we consider matter perturbations as function of position $\hat{n}$ on the sky (similar to what we did in three dimensions with $\delta(\mathbf{x})$), we can decompose

this perturbation distribution $\delta(\hat{\mathbf{n}})$ in spherical harmonics so that

$$\delta(\hat{\mathbf{n}}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}). \quad (3.1)$$

This way, the $a_{\ell m}$ are the coefficients of the decomposition of the matter perturbation distribution, about which we can only predict their ensemble average, not the value of a specific coefficient. The angular power spectrum, in this case, is given by the covariance of the coefficients $a_{\ell m}$:

$$\langle a_{\ell m} a_{\ell' m'}^* \rangle = C_{\ell} \delta_{\ell \ell'} \delta_{m m'}. \quad (3.2)$$

In practice, it is usually represented the quantity $\ell(\ell+1)C_{\ell}/2\pi$ as function of ℓ in order to predict the variance of the perturbation distribution on different scales (large scales are equivalent to low ℓ and small scales are equivalent to high ℓ). Furthermore, as we described in Chapter 2 for the power spectrum, we also need to take into account the analogous in two dimensions of W_{vox} and W_{smooth} , the voxel and smoothing window functions. In two dimensions, they depend on ℓ and will be written as W_{ℓ}^{pix} and W_{ℓ}^{smooth} , respectively. It is explained in [21] that they can be calculated by

$$W_{\ell}^{\text{pix}} = \text{sinc}\left(\frac{\ell \cdot \text{pix_side}}{2\pi}\right) \quad (3.3)$$

$$W_{\ell}^{\text{smooth}} = e^{-\ell(\ell+1)\sigma_{\text{FWHM}}^2/2}, \quad (3.4)$$

where σ_{FWHM} is the standard deviation of the Gaussian beam profile. As an example of how taking into account the window functions changes the shape of the angular power spectrum, we consider the value of *pix_side* given and $\sigma_{\text{FWHM}} = 0.007$ rad given by θ_{FWHM} . We first calculate the angular power spectrum through Code for Anisotropies in the Microwave Background (CAMB) given the cosmological parameters assumed in this work. This angular power spectrum is what we call $C_{\ell,1}$. After that, we calculate the angular power spectrum $C_{\ell,2} = (W_{\ell}^{\text{pix}})^2 C_{\ell,1}$ smoothed with W_{ℓ}^{pix} , and the angular power spectrum $C_{\ell,3} = (W_{\ell}^{\text{pix}} W_{\ell}^{\text{smooth}})^2 C_{\ell,1}$ smoothed with both W_{ℓ}^{pix} and W_{ℓ}^{smooth} . These three different angular power spectra are shown in Figure 3.1, where we represent the adimensional quantities $\ell(\ell+1)C_{\ell,i}/2\pi$ with $i \in \{1, 2, 3\}$ as function of ℓ in order to compare their shapes. As we can observe, as we add window functions to the angular power spectrum it decreases faster with increasing ℓ . This is due to the fact that if we add window functions which smooth the overdensity field, there are less fluctuations hence the variance decreases, specially in small scales (large ℓ).

3.2 VIDs from realizations

In this section we explain how we make realizations populating the sky with galaxies with the parameters described before and how we obtain the brightness temperature distribution from them. We will distinguish four different cases: Poissonian realization, smoothed Poissonian realization, Gaussian realization and smoothed Gaussian realization. We make Poissonian and Gaussian realizations in order to analyze how the clustering affects the temperature distribution. On the one hand, Poissonian realizations do not include the effect of clustering, assuming that there are no spatial correlations. On the other hand, the Gaussian realizations do include the effect of clustering because they are made from

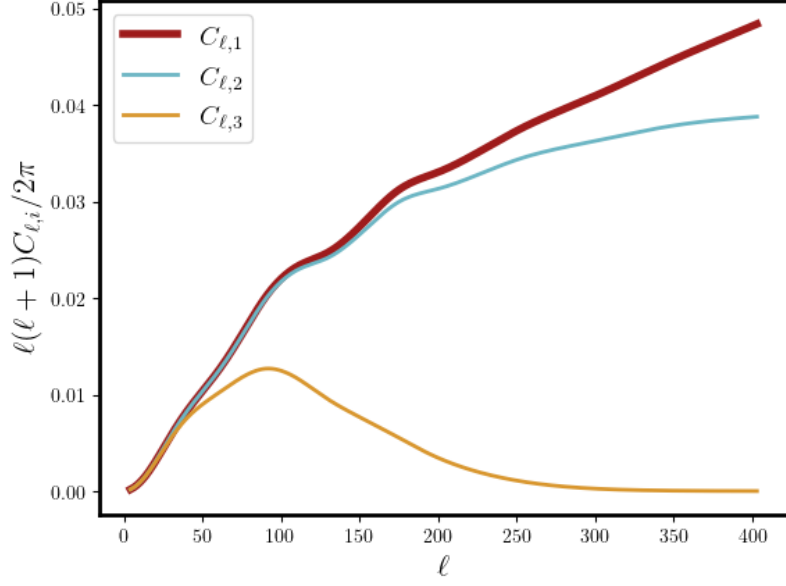


Figure 3.1: Comparison between $l(l+1)C_{l,i}/2\pi$ with $i \in \{1, 2, 3\}$ as function of ℓ . $C_{\ell,1}$ is the angular power spectrum with no window functions, $C_{\ell,2}$ is the angular power spectrum considering W_{ℓ}^{pix} and $C_{\ell,3}$ is the angular power spectrum considering both W_{ℓ}^{pix} and W_{ℓ}^{smooth} .

the angular matter power spectrum, which accounts for the spatial correlations, hence the clustering, too.

First, note that we assume there exists some noise coming from the LIM experiment, and we add it to all realizations we make. In this case, we compute it by taking random samples from a normal distribution with zero average and a standard deviation given by σ_N and adding them to the temperature map we have in each case, so that we add one value accounting for noise to each pixel.

For the first two maps, which correspond to the Poissonian realization and the smoothed Poissonian realization, we start from a null temperature map and what we do is populate it randomly, in a manner that we add the N_{gal} galaxies to random pixels in the map. Since $N_{\text{gal}} > N_{\text{pix}}$, we will have pixels with more than one galaxy, but we could have pixels with no galaxies at all, too. Furthermore, for the smoothed Poissonian realization, we smooth the map considering that the signal profile is a Gaussian beam with standard deviation given by θ_{FWHM} , this is $\sigma_{\text{FWHM}} = 0.007$ rad. This is different from the Poissonian realization, in which we assume the signal profile is a Dirac delta centered at each galaxy. In reality, the non-smoothed case is not possible to have because the LIM experiment will always have some finite angular resolution. The full-sky temperature maps obtained through HEALPix for the Poissonian realization and the smoothed Poissonian realization are shown in Figure 3.2 and Figure 3.3, respectively.

Regarding the maps in Figure 3.2 and Figure 3.3, we would expect the one corresponding to the smoothed case to have less contrast in colors than the non-smoothed one, because the smoothness makes the highest temperatures to drop and the lowest temperatures to rise. As we can observe in those figures, this is visible especially in the legend showing the range of temperatures present, because the range in the smoothed case is shorter than in the non-smoothed one. In addition to this, regarding the VIDs that we

will obtain from these maps, the VID from the smoothed one is expected to be narrower and higher than the VID from the non-smoothed one. This is because as we have explained the highest and lowest temperatures disappear, and the temperatures close to the mean temperature become more numerous.

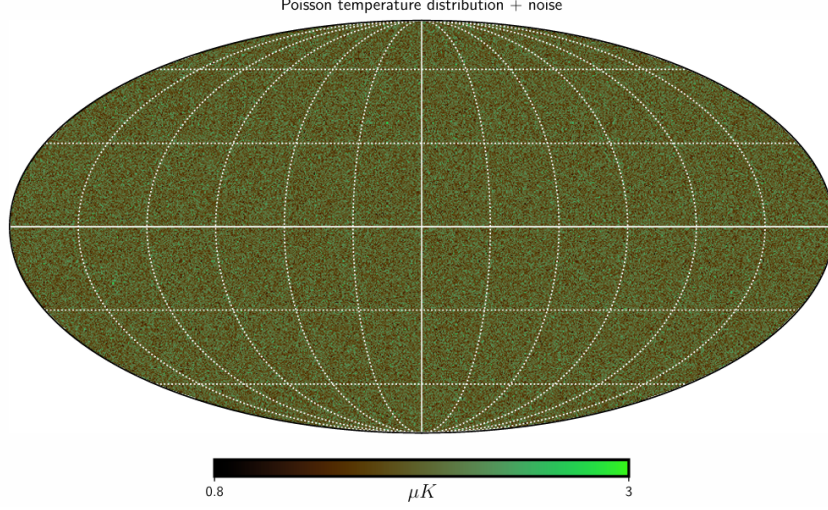


Figure 3.2: Full-sky temperature map for the Poissonian realization, including the experimental noise given by σ_N .

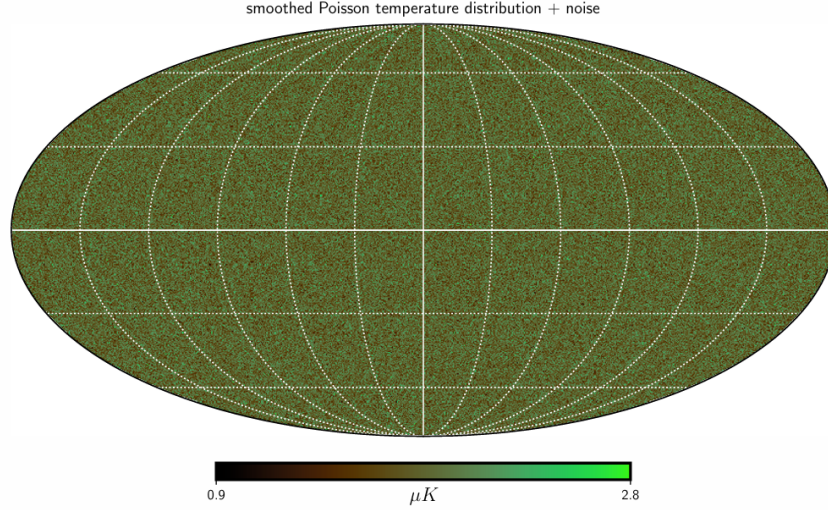


Figure 3.3: Full-sky temperature map for the smoothed Poissonian realization, including the experimental noise given by σ_N .

For the last two maps, which correspond to the Gaussian realization and the smoothed Gaussian realization, we start from the angular power spectrum $C_{\ell,1}$ shown in Figure 3.1 and we create a full-sky perturbation map making a Gaussian realization with HEALPix with angular correlations given by $C_{\ell,1}$. After that, we populate the pixels of this map with galaxies with luminosity equal to L_{gal} . In order to do that, similar to what is done in [22], we calculate the expected number density of galaxies per pixel as $\bar{n}(1 + b_h \delta(\hat{\mathbf{n}}))$,

where $\delta(\hat{n})$ is the matter perturbation map obtained from the angular power spectrum and b_h is the galaxy bias relating galaxies with matter distribution. With this, we make a Poissonian realization with parameter given by this expected number density of galaxies per pixel, so that each pixel is associated with a number of galaxies with temperature T_{tot} . In addition to this, just like we did in the smoothed Poissonian realization, for the smoothed Gaussian realization we smooth the map considering that the signal profile is a Gaussian beam with standard deviation equal to σ_{FWHM} . These two full-sky temperature maps are shown in Figure 3.4 (Gaussian realization) and Figure 3.5 (smoothed Gaussian realization).

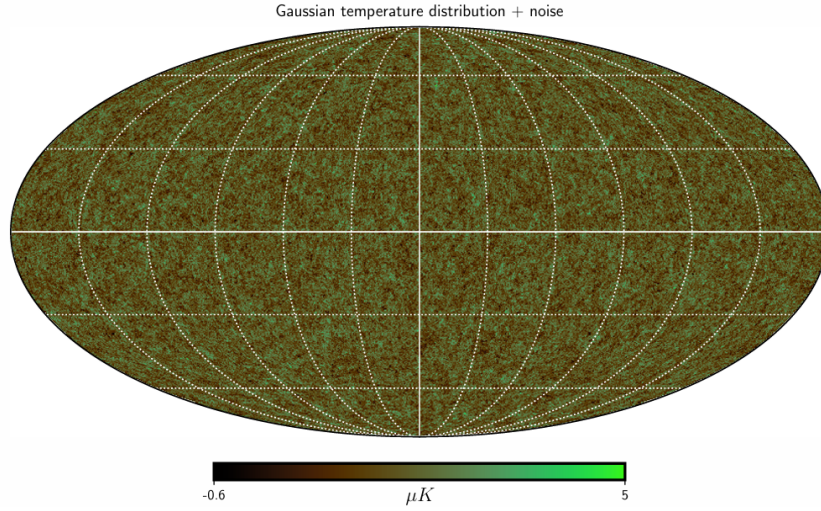


Figure 3.4: Full-sky temperature map for the Gaussian realization, accounting for clustering and including the experimental noise given by σ_N .

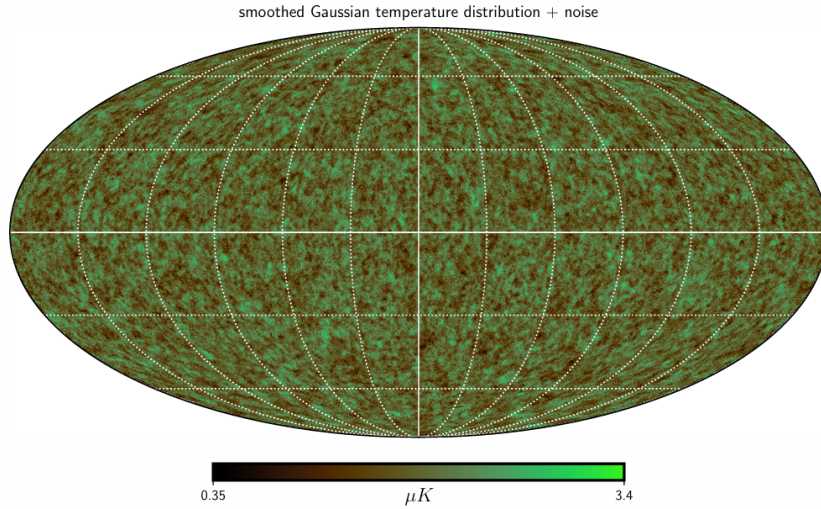


Figure 3.5: Full-sky temperature map for the smoothed Gaussian realization, accounting for clustering and including the experimental noise given by σ_N .

Similar to what happened in the Poissonian realizations, the map which represents the

Gaussian realization has more contrast in colors than the smoothed one. It is even clearer than in the Poissonian realizations that the map which represents the smoothed Gaussian realization appears to be blurred compared to the non-smoothed one. Moreover, if we observe the legend showing the range of temperatures present, it is visible that the range in the smoothed case is way shorter than the range in the non-smoothed case. Furthermore, if we compare the maps representing the Gaussian realizations with the maps which represent the Poissonian realizations, we observe that the ranges of temperatures in the cases of Gaussian realizations are longer than in the case of Poissonian realizations, something that is also directly visible in the maps, since the maps for the Gaussian realizations have more contrast in colors than the maps for the Poissonian realizations. These differences between the Poissonian and Gaussian realizations will be also observable in the VIDs, in a way that the VIDs of the Gaussian realizations are expected to be lower and wider than the VIDs of the Poissonian realizations. This is because since the Gaussian realizations consider clustering of matter, there are areas with less matter, hence less galaxies and lower temperatures (and conversely, too), while the Poissonian realizations has all pixels with the same expected number of galaxies.

Once we have the four different maps, we can calculate the VIDs simply computing the histogram of each map. In this way, we obtain the VID for each of the maps of the Poissonian and Gaussian realizations. These VIDs, normalized by the number of pixels, are shown in Figure 3.6 so that we can compare them and analyze their differences.

As we expected, if we observe Figure 3.6 in both Poissonian and Gaussian realizations the VID for the smoothed case is higher and narrower than the VID for the non-smoothed case. Furthermore, comparing the Poissonian VIDs with the Gaussian ones, we observe as we explained that the Gaussian VIDs are wider and lower due to the fact that the ranges of temperatures in these cases are longer than in the cases of the Poissonian realizations. We will return to these VIDs in next chapter, when we compare them with the ones obtained through the theoretical formalism so that we can validate it .

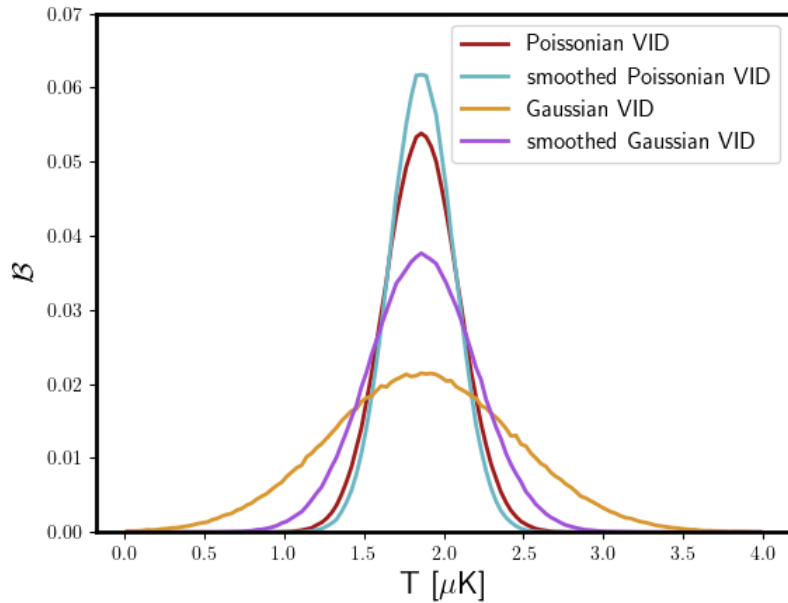


Figure 3.6: Voxel intensity distributions obtained from the four maps representing the Poissonian and Gaussian realizations made.

Chapter 4

Validation of the theoretical formalism

We have obtained in last chapter the VIDs calculated from four different realizations of a full-sky temperature map. The ultimate goal, which is discussed in this chapter, is to validate the VIDs we can compute from the theoretical formalism explained in Chapter 2 with these ones. For this, we first analyze the assumptions considered and the realizations carried out in Chapter 3 in order to obtain explicit theoretical equations which allow to compute the VIDs from this formalism. Later, we compare these VIDs with the ones obtained from realizations, validating in this way whether the theoretical formalism is correct or not. To do that, we divide the discussion in four sections, one for each of the realizations made, in increasing complexity.

4.1 Results for the VID

We start from equations (2.5) and (2.17), which allow to compute the Fourier transform of the one-point PDF of the brightness temperature. First, as we have assumed to be working with the sky instead of with a three-dimensional volume, we need to replace the three-dimensional vector $\mathbf{x}$ with the two-dimensional angle on the sky $\hat{\mathbf{n}}$. Now, we analyze the assumptions made over the galaxies considered in order to simplify the cited equations. On the one hand, since all galaxies have the same luminosity, equation (2.5) can be rewritten as

$$\tilde{\mathcal{P}}_1(\tau) = \int d^2\hat{\mathbf{n}} \mathcal{P}(\hat{\mathbf{n}}|M, \boldsymbol{\vartheta}) e^{-iT(\hat{\mathbf{n}})\tau}, \quad (4.1)$$

where we have taken the conditional probability distribution $\mathcal{P}(L|M, \boldsymbol{\vartheta})$ to be a Dirac delta centered on L_{gal} . On the other hand, since all galaxies have the same mass, the integral over the masses of galaxies multiplied by the halo mass function in the expression of the Fourier transform of the unclustered one-point PDF of equation (2.8) can be replaced directly by n , the mean number density of galaxies. Moreover, $b_h(M)$ can be replaced directly by b_h because we only consider one possible galaxy mass. Therefore, under these assumptions we can write

$$\tilde{\mathcal{P}}^{(u)}(\tau) = \exp \left\{ nV_{\text{prof}}(\tilde{\mathcal{P}}_1(\tau) - 1) \right\}. \quad (4.2)$$

Analogously, the Fourier transform of the one-point PDF including the effect of clustering given by equation (2.17) can be rewritten as

$$\tilde{\mathcal{P}}(\tau) = \tilde{\mathcal{P}}^{(u)} \exp \left\{ \left[nV_{\text{prof}}(\tilde{\mathcal{P}}_1(\tau) - 1)b_h \right]^2 \frac{\sigma^2}{2} \right\}. \quad (4.3)$$

Furthermore, we have added to all realizations made in Chapter 3 some noise coming from the LIM experiment, which we have modeled to be Gaussian. To take into account this noise when computing the theoretical temperature distribution, we convolute the PDF of the brightness temperature in each case with a normalized Gaussian function in one dimension with null average and standard deviation given by σ_N , which becomes a multiplication in Fourier space. In this way we have

$$\tilde{\mathcal{P}}_{\text{noise}}(\tau) = e^{-\tau^2 \sigma_N^2},$$

which is the Fourier transform of the Gaussian function characterizing the noise, and we multiply it by $\tilde{\mathcal{P}}(\tau)$ to obtain the total one-point PDF. We do this in each of the four cases analyzed.

Note that once we have the Fourier transforms of the four one-point PDFs taking also into account the Gaussian noise, we need to make their inverse Fourier transform and obtain the VIDs from them so that we can compare with the ones obtained from realizations. In order to do so we use the library FINUFFT, which computes efficiently Nonuniform Fast Fourier Transforms (NUFFT) in up to three dimensions. Particularly, we use a code which allows computing type 1 NUFFT, which starts from nonuniform (arbitrary) values of the Fourier conjugate τ and obtains the inverse Fourier transform of a given function evaluated in uniform (equally spaced) values of the temperature T . We use NUFFT instead of simply Fast Fourier Transforms (FFT) because we usually need to cover large variable ranges, hence the regular FFT would be computationally too expensive. This way, we can in the end obtain the desired PDFs in configuration space of the variable T , from which we compute the VIDs easily implementing equation (1.14). We now proceed to validate the four cases analyzed.

4.1.1 Poissonian VID validation

In this first case, as we explained, clustering is not considered, this means $\sigma^2 = 0$ and we assume the signal profile is a Dirac delta centered at each galaxy. Therefore, at the center of each galaxy there is a temperature equal to T_{tot} and we obtain

$$\tilde{\mathcal{P}}_1(\tau) = \int d^2 \hat{\mathbf{n}} \delta_D(\hat{\mathbf{n}} - \hat{\mathbf{n}}_i) e^{-T(\hat{\mathbf{n}})\tau} = e^{-T(\hat{\mathbf{n}}_i)\tau} = e^{-T_{\text{tot}}\tau}, \quad (4.4)$$

where $\hat{\mathbf{n}}_i$ refers to the position of any galaxy. Furthermore, since the smallest volume we can resolve is V_{pix} and the signal profile is a Dirac delta, we get in this case $V_{\text{prof}} \approx V_{\text{pix}}$ so that

$$\tilde{\mathcal{P}}(\tau) = \exp \left\{ nV_{\text{pix}}(\tilde{\mathcal{P}}_1(\tau) - 1) \right\} = \exp \left\{ \bar{n}(\tilde{\mathcal{P}}_1(\tau) - 1) \right\}. \quad (4.5)$$

This way we obtain the Fourier transform of the theoretical one-point PDF for the Poissonian case. As explained, all that remains to be done is to add the Gaussian noise, to do the inverse Fourier transform and to compute the VID from it. Figure 4.1 shows this theoretical VID obtained as well as the VID obtained from the Poissonian realization so that we can compare them.

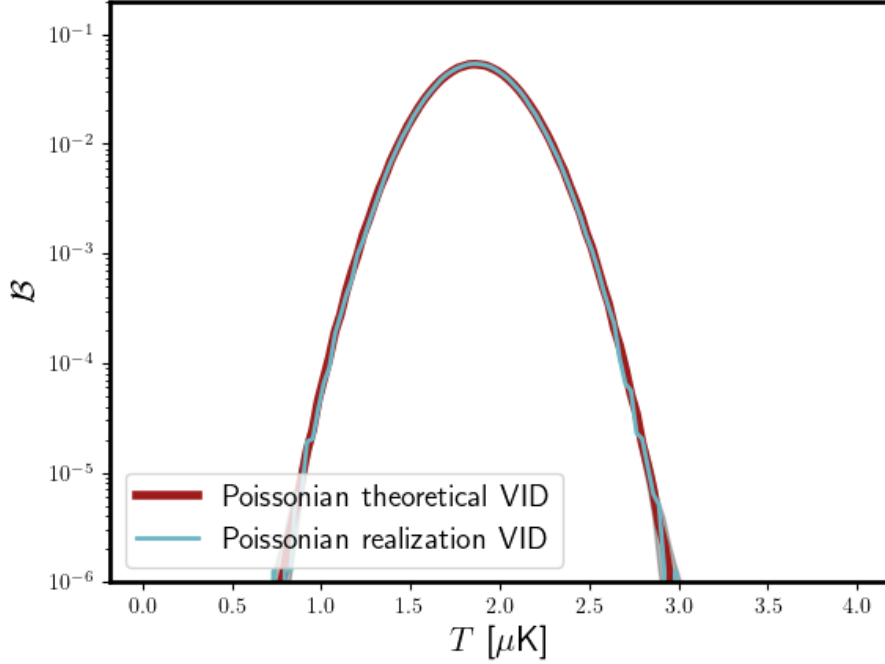


Figure 4.1: Comparison between the theoretical VID obtained for the Poissonian case and the VID obtained from the Poissonian realization. The y-axis is on logarithmic scale and we plot the error band on the same axis.

As we can observe, both VIDs match almost perfectly, validating for this particular case the theoretical formalism followed. The error band, which we obtain filling the region between the theoretical VID minus the theoretical error and the theoretical VID plus the theoretical error, is only slightly observable on the lowest and highest temperatures. The theoretical error, particularly, is obtained calculating the square root of the diagonal terms of the covariance of the VID, given by equation (2.21).

4.1.2 Smoothed Poissonian VID validation

In this second case we assume $\sigma^2 = 0$, too, but now we assume the signal profile is T_{tot} times a normalized Gaussian beam with null mean and standard deviation given by σ_{FWHM} . Here, we compute equation (4.1) in a different way than the case before. We consider the Gaussian profile to be covering a specific number of pixels *counts* (we assume a radius equal to five times the standard deviation), and we do the integral in the equation cited in a discrete way. We first calculate the temperature in each pixel of the profile by

$$T_{\text{pix}}^{(j)} = \int_{A_{\text{pix}}^{(j)}} d^2\hat{\mathbf{n}} T(\hat{\mathbf{n}}), \quad (4.6)$$

where $T(\hat{\mathbf{n}})$ is the Gaussian beam multiplied by T_{tot} , because the temperature integrated over the profile needs to be equal to T_{tot} . In this equation, $A_{\text{pix}}^{(j)}$ denotes the area of each pixel, i.e. the region we integrate over. This integral is that of a two-dimensional Gaussian, which can be separated as a multiplication of two integrals in one dimension. Since the integral of a Gaussian in one dimension is analytical, we can easily compute in this way

the temperature for each of the pixels in the profile. Note that we center the pixels in the origin, where the galaxy is in each case. Then, we calculate $\tilde{\mathcal{P}}_1$ as a normalized sum by

$$\tilde{\mathcal{P}}_1(\tau) = \sum_j \frac{counts^{(j)}}{counts} e^{-iT_{\text{pix}}^{(j)}\tau}, \quad (4.7)$$

where $counts^{(j)}$ accounts for the number of times a pixel has temperature $T_{\text{pix}}^{(j)}$, so that we only sum over the different temperatures of the pixels. We need to compute $\tilde{\mathcal{P}}_1(\tau)$ like this because T_{tot} depends on the pixelization, so we can not do the integral of equation (4.1) the same way we did in the Poissonian case.

Once we have $\tilde{\mathcal{P}}_1(\tau)$, we calculate $\tilde{\mathcal{P}}(\tau)$ from equation (4.2). To do so, we consider that nV_{prof} is the number of galaxies covered by the signal profile, so in this case we can calculate it multiplying $nV/4\pi$ (the number of galaxies per steradian) by the area of the pixels covered by the profile, which is obtained doing $pix_side^2 \cdot counts$ (pix_side^2 is the area of one pixel). The final expression for $\tilde{\mathcal{P}}(\tau)$ is

$$\tilde{\mathcal{P}}(\tau) = \exp \left\{ \frac{nV}{4\pi} pix_side^2 counts (\tilde{\mathcal{P}}_1(\tau) - 1) \right\}. \quad (4.8)$$

All that remains to be done is to add the Gaussian noise, to do the inverse Fourier transform and to compute the VID from it. Figure 4.2 shows this theoretical VID obtained as well as the VID obtained from the smoothed Poissonian realization so that we can compare them.

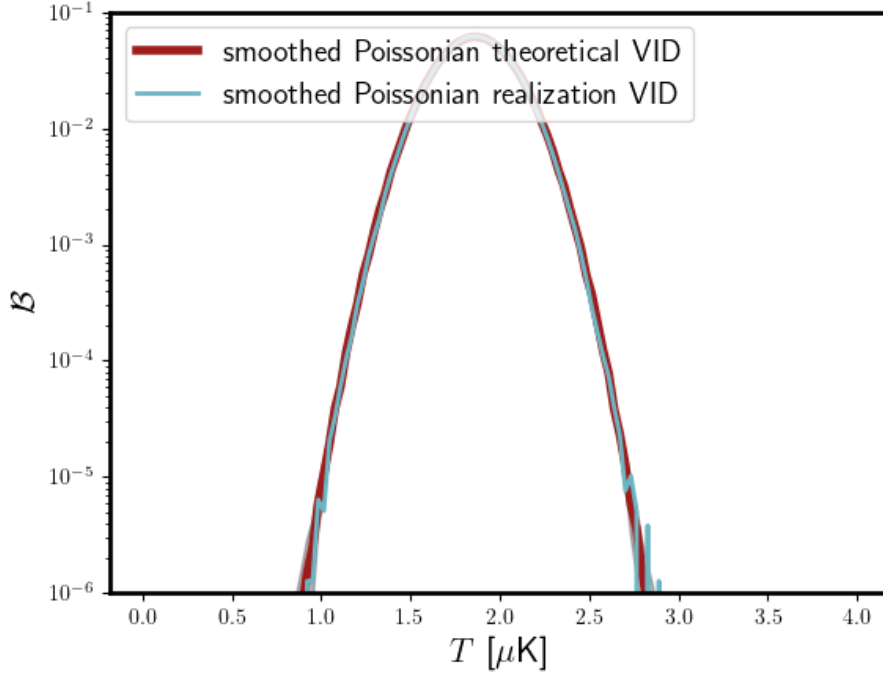


Figure 4.2: Comparison between the theoretical VID obtained for the smoothed Poissonian case and the VID obtained from the smoothed Poissonian realization. The y-axis is on logarithmic scale and we plot the error band on the same axis.

As we can observe, both VIDs match almost perfectly. Although there are some discrepancies in the highest temperatures (due to noise in the realization), this comparison

allows us to validate for the smoothed Poissonian case the theoretical formalism developed in Chapter 2. The error band is obtained with the same procedure as the previous case and it is only slightly observable on the lowest and highest temperatures.

4.1.3 Gaussian VID validation

In this third case, we again consider the profile to be a Dirac delta centered at each galaxy, so that we can obtain $\tilde{\mathcal{P}}_1(\tau)$ through equation (4.4) as we did in the Poissonian case. The difference is that now we need to consider clustering, so we have to calculate σ^2 and make use of equation (4.3), where $\tilde{\mathcal{P}}^{(u)}(\tau)$ can be calculated through equation (4.5).

In order to calculate σ^2 , we note that under the assumptions made it is the zero-lag variance in two dimensions instead of three, because the information along the line of sight is projected into the angular power spectrum. It is calculated from the so-called angular two-point correlation function $w(\theta)$, which is the inverse harmonic spherical transform of the angular power spectrum, simply evaluating $w(0)$. As explained in [1] and assuming isotropy and homogeneity, the angular two-point correlation function can be expressed as

$$w(\theta) = \frac{1}{2\pi} \int_0^\infty d\ell \ell C_\ell J_0(\ell\theta), \quad (4.9)$$

where J_0 is the ordinary Bessel function of order zero. However, as we work with a pixelated map of the sky, in practice we deal with the window function W_ℓ^{pix} given by equation (3.3), which smooths the angular power spectrum. This function can also be directly computed with HEALPY and it depends on ℓ due to the decomposition in spherical harmonics. Therefore, the zero-lag variance for this particular case is calculated by

$$\sigma^2 = w(0) = \frac{1}{2\pi} \int_0^\infty d\ell \ell C_\ell (W_\ell^{\text{pix}})^2 J_0(0), \quad (4.10)$$

and we obtain $\sigma^2 = 0.045$. With this value, we calculate the inverse Fourier transform of the one-point PDF by

$$\tilde{\mathcal{P}}(\tau) = \tilde{\mathcal{P}}^{(u)} \exp \left\{ \left[\bar{n}(\tilde{\mathcal{P}}_1(\tau) - 1) b_h \right]^2 \frac{\sigma^2}{2} \right\}, \quad (4.11)$$

where $\tilde{\mathcal{P}}^{(u)}(\tau)$ is calculated through equation (4.5). Then, we add the Gaussian noise, we do the inverse Fourier transform and we compute the VID from it. Figure 4.3 shows this theoretical VID obtained as well as the VID obtained from the Gaussian realization so that we can compare them.

As we can observe, both VIDs match perfectly, validating the theoretical formalism followed for this particular case. The error band, which is imperceptible, is obtained with the same procedure as before.

4.1.4 Smoothed Gaussian VID validation

In this last case, we again need to take into account clustering, but the difference with the previous case is that now we consider the signal profile to be T_{tot} times a normalized Gaussian beam as we did in the smoothed Poissonian case. Therefore, starting from equation (4.3), in this case $\tilde{\mathcal{P}}^{(u)}(\tau)$ is equal to the Fourier transform of the PDF given by equation (4.8), which corresponds to a uniform distribution without clustering. Then, the

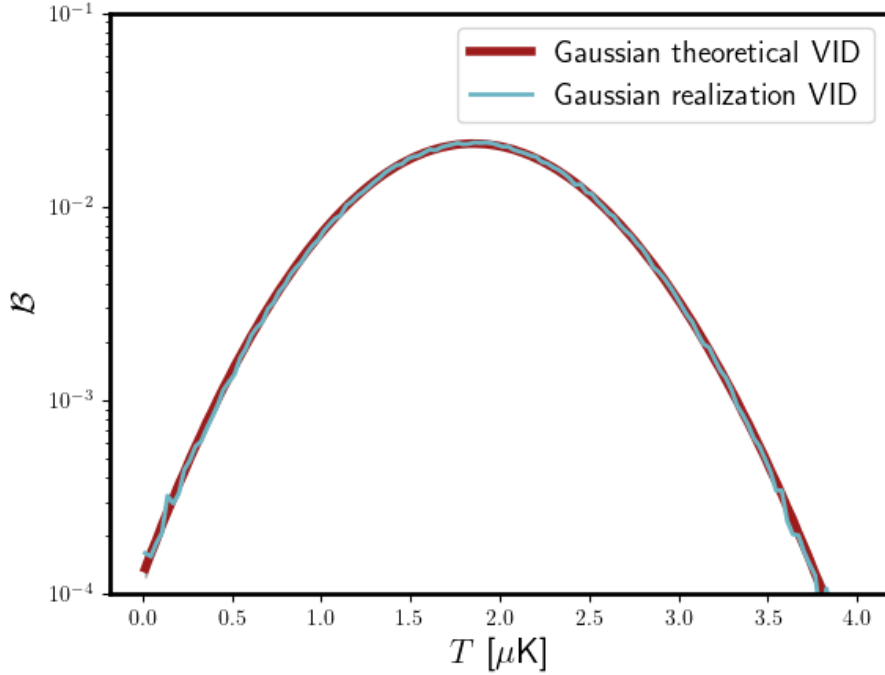


Figure 4.3: Comparison between the theoretical VID obtained for the Gaussian case and the VID obtained from the Gaussian realization. The y-axis is on logarithmic scale and we plot the error band on the same axis.

Fourier transform of the one-point PDF for the smoothed Gaussian case can be obtained by

$$\tilde{\mathcal{P}}(\tau) = \tilde{\mathcal{P}}^{(u)} \exp \left\{ \left[\frac{nV}{4\pi} \text{pix_side}^2 \text{counts} (\tilde{\mathcal{P}}_1(\tau) - 1) b_h \right]^2 \frac{\sigma^2}{2} \right\}, \quad (4.12)$$

where $\tilde{\mathcal{P}}_1(\tau)$ is obtained from equation (4.7).

However, σ^2 is different from the one computed in the Gaussian case, because now we need to take into account the window function originated from the smoothing, W_ℓ^{smooth} , in addition to W_ℓ^{pix} . Hence, in this case we calculate σ^2 from

$$\sigma^2 = w(0) = \frac{1}{2\pi} \int_0^\infty d\ell \ell C_\ell (W_\ell^{\text{pix}} W_\ell^{\text{smooth}})^2 J_0(0), \quad (4.13)$$

obtaining $\sigma^2 = 0.011$. It makes sense that the zero-lag variance in this case is smaller than in the Gaussian case, because adding a smoothing window function implies that structure on scales smaller than the beam size is undetectable as cited in [1], hence it reduces the variance. In a clumpy Universe, the smaller the scales we resolve the largest the variance, but as we smooth the matter distribution, the perturbations become smaller and so their variance.

Once we compute the Fourier transform of the PDF through equation (4.12), it only remains to add the Gaussian noise, to do the inverse Fourier transform and to compute the VID from it. Figure 4.4 shows this theoretical VID obtained as well as the VID obtained from the smoothed Gaussian realization so that we can compare them.

It is clear from the figure that both VIDS match perfectly, allowing us to validate also for this fourth case the theoretical formalism followed. The error band is again imperceptible and it is computed as before.

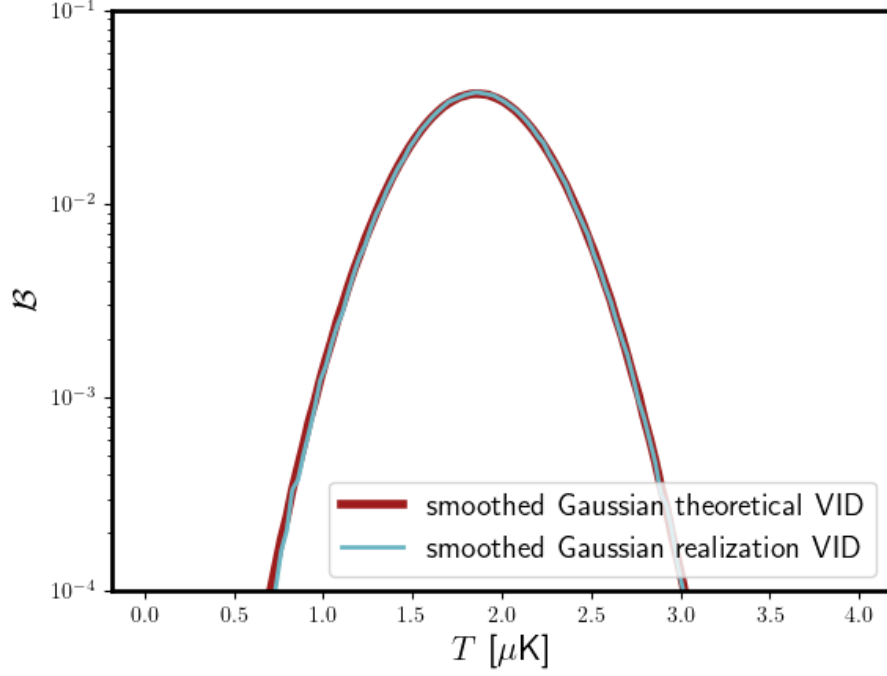


Figure 4.4: Comparison between the theoretical VID obtained for the smoothed Gaussian case and the VID obtained from the smoothed Gaussian realization. The y-axis is on logarithmic scale and we plot the error band on the same axis.

Chapter 5

Conclusions and future work

The main aim of this work has been to develop a theoretical formalism which enables to compute the one- and two-point PDFs of the brightness temperature measured in Line-Intensity Mapping experiments, and check its validity in cases that are simple to implement computationally.

After introducing all the necessary contents to understand the object of study, such as the principles of the standard Λ Cold Dark Matter cosmological model and how the LIM technique works, we have focused on describing the theoretical formalism which allows to obtain the desired temperature distributions. We have developed in two separated parts the one- and two-point PDFs, differentiating as well in the case of the latter between the one-halo term and the two-halo term. This work is the first step towards a robust modeling of the two-point PDF, for which there is no analytic modeling so far. As a first step, we have focused on the validation of the theoretical one-point PDF under some simplifications. These simplifications have been, among others, that we have assumed to be working with the projected full-sky instead of with a three-dimensional volume. We have computed Poissonian and Gaussian realizations of the distribution of galaxies on the sky so that we managed to obtain the temperature distributions from them to compare with the ones computed with the theoretical formalism, taking into account some Gaussian noise due to the experimental limitation of measuring instrumentation. In all four cases analyzed we have succeed on validating our theoretical formalism, clarifying that the temperature distributions obtained from it match almost perfectly with the ones we have computed from realizations. Our simplifications regarding the distribution of masses and luminosities were aimed for reducing the computing time and accelerate the validation, but there is nothing to indicate that accounting for such distributions (always that the distributions are precise) would make our modeling break. Highlight that, since we obtain an actual estimator of the PDF, we can extract non-Gaussian information from the temperature distribution unlike if for example we used the power spectrum, which only makes Gaussian descriptions of the fields.

This work shows one of the areas of application of LIM techniques, although there are many others as we cited in Chapter 1. Regarding the possible future work to be done, the first extension of this work would be to validate also the theoretical formalism for the two-point PDF of the brightness temperature under the same assumptions followed when validating the one-point PDF. Note, for example, that this means dealing with temperature distributions in two-dimensions, what we can understand as heat maps. Following this, it would be also interesting to analyze the case in which we still work with the sky instead of a volume of the Universe but eliminating some of the assumptions made re-

garding the properties of the galaxies. For example, considering different galaxy masses and luminosities and accounting for more complicated and realistic clustering and bias distributions.

The most difficult case to implement computationally would be considering a finite volume of the Universe, hence an extra dimension along the line of sight, which would mean working with three dimensional vectors and integrals as developed in the theoretical formalism of Chapter 2. In this case it would be interesting to reduce the computational cost of operations as much as possible, and that is why we do when deriving the one-halo term of the two-point PDF in Chapter 2 all the expansions in Legendre polynomials of the convolution term. It would be crucial, thus, to find more ways of easing the computational implementation and costs.

In conclusion, we have fulfilled the purpose of this work, developing a theoretical formalism for the one- and two-point PDFs of the brightness temperature and validating it for some simple cases. This paves the way to validating more complex cases, closer to what actually happens in the Universe, so that the significant role of LIM in this context becomes clearly proved.

References

- [1] S. Dodelson and F. Schmidt, “Modern Cosmology”, Academic Press, 2020.
- [2] L. Perivolaropoulos and F. Skara, “Challenges for Λ CDM: An update”, *New Astron. Rev.* **95**, (2022), [arXiv:2105.05208 \[astro-ph.CO\]](#).
- [3] J. L. Bernal, Lecture notes: “Cosmology”, Instituto de Física de Cantabria (CSIC-UC), 2022.
- [4] S. Weinberg, “Cosmology”, Oxford University Press, 2008.
- [5] **Planck** Collaboration, N. Aghanim *et al.*, “Planck 2018 results. VI. Cosmological parameters”, *Astron. Astrophys.* **641** (2020) A6, [arXiv:1807.06209 \[astro-ph.CO\]](#).
- [6] D. W. Hogg, “Distance measures in cosmology”, (2000), [arXiv:astro-ph/9905116](#).
- [7] A. Cooray and R. Sheth, “Halo models of large scale structure”, *Phys. Rept.* **372** 1-129, (2002), [arXiv:astro-ph/0206508](#).
- [8] J. L. Bernal, “Towards accurate modeling of line-intensity mapping one-point statistics: Including extended profiles”, *Phys. Rev. D* **109** no. 4, (2024) 043517, [arXiv:2309.06481 \[astro-ph.CO\]](#).
- [9] J. L. Bernal and E. D. Kovetz, “Line-intensity mapping: theory review with a focus on star-formation lines”, *Astron. Astrophys. Rev.* **30** no. 1, (2022) 5, [arXiv:2206.15377 \[astro-ph.CO\]](#).
- [10] E. D. Kovetz *et al.*, “Line-Intensity Mapping: 2017 Status Report”, (2017), [arXiv:1709.09066 \[astro-ph.CO\]](#).
- [11] J. L. Bernal, P. C. Breysse, H. Gil-Marín, E. D. Kovetz, “A User’s Guide to Extracting Cosmological Information from Line-Intensity Maps”, *Phys. Rev. D* **100** no. 12, (2019) 123522, [arXiv:1907.10067 \[astro-ph.CO\]](#).
- [12] T. M. Oxholm, “A Beginner’s Guide to Line Intensity Mapping Power Spectra”, *Proceedings of the Wisconsin Space Conference* **1** no. 1, (2021), [arXiv:2204.00685 \[astro-ph.CO\]](#).
- [13] P. Reichert, “The ergodic hypothesis: a typicality statement”, (2023), [arXiv:2307.05166 \[cond-mat.stat-mech\]](#).
- [14] L. Thiele, J. C. Hill, and K. M. Smith, “Accurate analytic model for the weak lensing convergence one point probability distribution function and its autocovariance”, *Phys. Rev. D* **102** no. 12, (2020) 123545, [arXiv:2009.06547 \[astro-ph.CO\]](#).

- [15] R. Bracewell, “The Fourier transform and its applications”, McGraw-Hill, 2000.
- [16] Y. Zhang, A. R. Pullen, R. S. Somerville, P. C. Breysse, J. C. Forbes, S. Yang, et al., “Characterizing the Conditional Galaxy Property Distribution using Gaussian mixture models”, *Astrophys. J.*, **950** no. 2, (2023) 159, [arXiv:2302.11166 \[astro-ph.GA\]](#).
- [17] G. Casella and R. L. Berger, “Statistical Inference”, Duxbury Thomson Learning, 2008.
- [18] O. H. E. Philcox and Z. Slepian, “Beyond the Yamamoto approximation: anisotropic power spectra and correlation functions with pairwise lines of sight”, *Phys. Rev. D* **103**, (2021) 123509, [arXiv:2102.08384 \[astro-ph.CO\]](#).
- [19] A. Edmonds, “Angular momentum in quantum mechanics”, Princeton University Press, 1957.
- [20] A. D. Jackson and L. C. Maximon, “Integrals of Products of Bessel Functions”, *SIAM J. Math. Anal.*, **3** no. 3, (1972), <https://doi.org/10.1137/0503043>.
- [21] R. M. Sullivan, L. T. Hergt and D. Scott, “Methods for CMB map analysis”, (2024), [arXiv:2410.12951 \[astro-ph.CO\]](#).
- [22] A. Agrawal, R. Makiya, C.-T. Chiang, D. Jeong, S. Saito and E. Komatsu, “Generating Log-normal Mock Catalog of Galaxies in Redshift Space”, *JCAP* **10** (2017) 003, [arXiv:1706.09195 \[astro-ph.CO\]](#).