Chapter 2
Unravelling the iSW Effect Through the Matter Distribution

2.1 CMB Cross-Correlation with Tracers of Matter

The end of the last chapter painted the iSW effect of CMB photons as a clever probe that has the potential to independently prove the existence of Dark Energy in the Universe. However, the faintness of this signal makes it almost impossible to detect, at least using only the CMB as such. The story does not end here though, thanks to the resourcefulness of cosmologists at finding ways to detect and exploit the iSW effect more efficiently. Among the methods devised, one of them stands as the leading technique in the literature: the cross-correlation of the CMB with the distribution of matter in the Universe.

2.1.1 The Principle

Before going into the details of this approach, let us recall the main ideas it is based on. As discussed in Chap. 1, the iSW effect felt by the CMB photons in a ΛCDM universe is a result of the stretching of the large scale potentials in the Universe, caused by the acceleration of the expansion which is itself due to the presence of Dark Energy. Of course, one has to remember that these gravitational potentials originate from the presence of matter in the form of large Dark Matter halos in which sit clusters of galaxies. Following this, it is reasonable to think that there exists a certain degree of correlation between the distribution of matter and the resulting pattern of the iSW effect that it generates in the CMB across the sky. The link between these two elements is of course not trivial since, as its name emphasizes, the iSW effect is an integrated effect: CMB photons that come from a particular direction in sky have been affected by the matter distribution along the whole line of sight. This is then further complicated by two competing phenomena: on the one hand the iSW effect is expected in the ΛCDM paradigm to be redshift-dependent with the most recent structures in the Universe yielding a more pronounced effect as they are

locally more DE-dominated. However, these very structures are also closer and closer
to us so that their number becomes limited by the available volume around us. At
some point, the stronger iSW effect of the closest large structures becomes balanced
by their increasing scarcity and therefore gets progressively harder to detect. We
therefore intuit already that there will be an optimal redshift for the detection of the
iSW effect, at some point between $z = 0$ and the start of the era in which Dark
Energy became cosmologically important.

Now that we have a grasp of the motivations behind the use of cross-correlation
and its associated features, we will have in the next subsection a closer look at the
theoretical tools used to describe this correlation.

2.1.2 Theoretical Framework

As seen in Sects. 1.2 and 1.4, the anisotropies that are generated by the iSW effect
are directly correlated to the distribution of matter through the evolution of the
gravitational potential $\dot{\Phi}$. Exploiting this correlation and detecting the iSW effect can
be done via several approaches; I will focus here on the most widely used method
which is done in spherical harmonic space.

All cross-correlations methods suppose that we have first a map of the tempera-
ture of the CMB $T(\hat{\mathbf{n}})$ at hand, or rather, a map of the *relative* fluctuations of this
background:

$$\delta_T(\hat{\mathbf{n}}) = \frac{T(\hat{\mathbf{n}}) - \bar{T}}{\bar{T}} \tag{2.1}$$

with $\bar{T}$ the mean temperature of the CMB (with the latest studies indicating
$\bar{T} = 2.7260 \pm 0.0013$, cf. [15]). On the other hand, we need a survey that traces
the distribution of galaxies, from which we derive a map of the projected galaxy
overdensity field:

$$\delta_g(\hat{\mathbf{n}}) = \frac{N(\hat{\mathbf{n}}) - \bar{N}}{\bar{N}} \tag{2.2}$$

where $N(\hat{\mathbf{n}})$ is the number of galaxies in the pixel corresponding to the direction $\hat{\mathbf{n}}$ and
$\bar{N}$ is the mean number of galaxies per pixel. Both maps are therefore in dimensionless
units—which is always welcomed to simplify calculations. Then, for the approach
that I consider here, we use the fact that any field can be decomposed into a series
of functions which form an orthonormal set, as do the spherical harmonic functions
$Y_{\ell m}(\theta, \phi)$. It follows that both the CMB temperature map (δ_T) and the overdensity
map (δ_g) can be decomposed into

$$\delta_X(\theta, \phi) = \sum_{\ell,m} a_{\ell m}^X Y_{\ell m}(\theta, \phi), \tag{2.3}$$

where $a_{\ell m}^X$ $(X = g, T)$ are the spherical harmonic coefficients of the field resulting from its decomposition. The main quantity of interest for us is the 2-point galaxy-temperature cross-correlation as a function of the multipole ℓ, whose estimator is:

$$\hat{C}_{Tg}(\ell) = \frac{1}{(2\ell + 1)} \sum_m \mathcal{R}e \left[a_{\ell m}^g (a_{\ell m}^T)^* \right] = \frac{1}{(2\ell + 1)} \sum_m \mathcal{R}e \left[a_{\ell m}^T (a_{\ell m}^g)^* \right], \quad (2.4)$$

It traces the degree of correlation between the two maps in harmonic space. We can intuit that in the absence of Dark Energy and hence of the iSW effect, the correlation between these two will be reduced to only fortuitous coincidences. But now that we have defined its computation, let us take a more precise look at the expected cross-correlation signal $C_{Tg}(\ell)$.

Firstly, the fluctuations of the CMB temperature $\delta_T(\hat{\mathbf{n}})$ are known to be composed of several contributions, often categorised into primordial and secondary anisotropies. However, for the large scales that we consider the only ones that are correlated to the distribution of matter in the Universe are the secondary anisotropies generated through the late iSW effect. In the remainder of this section, I will associate the notation $\delta_T(\hat{\mathbf{n}})$ to these iSW fluctuations only. One way of expressing these temperature fluctuations on a particular line of sight is written as the following redshift-integral from the surface of last scattering (SLS) to us:

$$\delta_T(\hat{\mathbf{n}}) = \int_{z_{\mathrm{SLS}}}^{0} e^{-\tau(z)} (\dot{\Phi} - \dot{\Psi})[\hat{\mathbf{n}}, z] \, \mathrm{d}z. \quad (2.5)$$

where the dot denotes here differentiation with respect to z (details about the equation and its terms can be found in Sect. 1.4). Since the matter density is related to the gravitational potentials Φ and Ψ by the Poisson equation, these iSW temperature fluctuations will be related to the observed galaxy density contrast, given by:

$$\delta_g(\hat{\mathbf{n}}) = \int_{z_{\mathrm{SLS}}}^{0} b_g(z) \frac{\mathrm{d}N}{\mathrm{d}z}(z) \delta_m(\hat{\mathbf{n}}, z) \, \mathrm{d}z. \quad (2.6)$$

In this expression, $\mathrm{d}N/\mathrm{d}z$ is called the selection function of the survey (from which the density map is derived) and represents simply the redshift distribution of galaxies in the survey. More accurately, this function gives the number of galaxies contained in a shell of width $\mathrm{d}z$ at redshift z; it is often normalised and then describes the fraction of objects per redshift (an example of such selection function is shown in Fig. 2.1). The term δ_m corresponds to the matter density perturbations, which are related to the galaxy overdensities by a factor b_g: indeed, although we correlate in practice the CMB map with a galaxy map, we aim at probing in reality the correlation with the underlying distribution of Dark Matter. This "galaxy bias" can theoretically evolve in time and be a function of scale. However, it is generally assumed to be
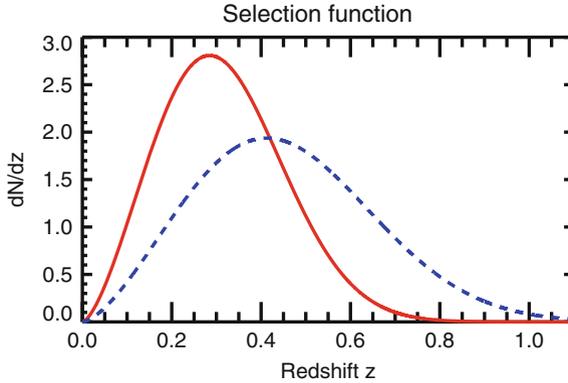
Selection function

Fig. 2.1 Redshift distributions of the two fictitious surveys considered (see text for details). These selection functions follow a widely-used analytical expression that describes well such quantities: $dN/dz(z) = Az^m \exp -(z/z_0)^\beta$, where A is chosen so that the integral of the function is normalized to unity. The parameters m and β control the slope of the rise and fall respectively, while z_0 fixes the median redshift of the distribution z_{med} ($= z_0\sqrt{2}$). Here I chose $z_0 = 0.48287$, $m = 1.51964$ and $\beta = 2.34207$ for the SDSS-like distribution (*red solid curve*), and increased z_0 to 0.7 for the second one (*blue dashed curve*)

time and scale independent for simplicity. For our purposes, a time dependent bias would be equivalent to changing the selection function of the survey. A possible scale dependence of the bias is more problematic, but on the very large scales (>10 Mpc) we are considering, the scale dependence is expected to be weak (see e.g. [8, 33]).

We are here interested in the correlation between δ_T and δ_g: their angular cross-correlation function in real space is defined as:

$$C^{Tg}(\vartheta) \equiv \langle \delta_T(\hat{\mathbf{n}}_1)\delta_g(\hat{\mathbf{n}}_2) \rangle \tag{2.7}$$

with the average carried over all the pairs at the same angular distance $\vartheta = |\hat{\mathbf{n}}_1 - \hat{\mathbf{n}}_2|$. As mentioned before, it is often preferred to work in the harmonic space and study the cross-correlation spectrum $C_{Tg}(\ell)$ instead of the function $C^{Tg}(\vartheta)$. Those two quantities are related through the Legendre polynomials P_ℓ:

$$C^{Tg}(\vartheta) = \sum_{l=2}^{\infty} \frac{2\ell + 1}{4\pi} C_\ell^{Tg} P_\ell[\cos(\vartheta)]. \tag{2.8}$$

After some calculations (for detailed steps, see e.g. [16]), it follows that the cross-correlation power spectrum is given by:

$$C_\ell^{Tg} = 4\pi \int \frac{dk}{k} \Delta^2(k) I_\ell^{iSW}(k) I_\ell^g(k), \tag{2.9}$$

where $\Delta(k)$ is the almost scale invariant initial matter power spectrum $\Delta^2(k) \equiv 4\pi k^3 P(k)/(2\pi)^3$ and the two integrands are respectively:

$$I_\ell^{\mathrm{iSW}}(k) = -2 \int e^{-\tau(z)} (\dot{\Phi}_k - \dot{\Psi}_k) j_\ell[k\chi(z)]dz \qquad (2.10)$$

$$I_\ell^g(k) = \int b_g(z) \frac{dN}{dz}(z) \delta_m(k,z) j_\ell[k\chi(z)]dz, \qquad (2.11)$$

where Φ_k, Ψ_k and $\delta_m(k,z)$ are the Fourier components of the gravitational potentials and matter perturbations for the wavenumber k, $j_\ell(x)$ are the spherical Bessel functions and χ is the comoving distance. The two integrands I_ℓ^{iSW} and I_ℓ^g can then be calculated for a given cosmological model using numerical codes that compute all the relevant quantities needed (the details of the corresponding equations can be found in [16]).

Of interesting note is that in the linear theory of perturbations (an approximation which should be more than valid here considering the large scales involved) the growth of the gravitational potentials is directly proportional to the ratio $D(a)/a$ (cf. Sect. 1.4). As a result, in a flat, matter-dominated universe the terms $\dot{\Phi}_k$ and $\dot{\Psi}_k$ in Eq. (2.10) would be equal to 0 and so would the integral—and the cross-correlation spectrum C_ℓ^{Tg} itself. Therefore, we confirm here that in the absence of Dark Energy (or more accurately, if the dominant component of the Universe were pressureless matter) there would be no iSW effect and no correlation between the CMB temperature and the distribution of matter.

2.1.3 Application to Detectability and Test Case

We reviewed in the previous section the analytic expression of the correlation between the CMB anisotropies and the distribution of matter through their cross-correlation power spectrum. The question now arises of the use of these theoretical predictions for the study of the Dark Energy.

The first application of these theoretical tools is to give the possibility to predict the detectability of the aforementioned correlation. Assuming a given cosmological model and that both the CMB temperature and the galaxy maps behave as Gaussian random fields, the covariance of the iSW cross-correlation signal (in the absence of noise) can be calculated by:

$$\sigma^2[C_\ell^{Tg}] = \frac{(C_\ell^{Tg})^2 + C_\ell^{gg} C_\ell^{TT}}{2\ell + 1}, \qquad (2.12)$$

where C_ℓ^{TT} is the full CMB temperature-temperature power spectrum (and not only the secondary anisotropies generated by the iSW effect) and C_ℓ^{gg} is the galaxy auto-correlation function that can be calculated theoretically:

$$C_\ell^{gg} = 4\pi \int \frac{dk}{k} \Delta^2(k) I_\ell^g(k) I_\ell^g(k) \tag{2.13}$$

(see Eq. (2.9) for a description of the terms). We can identify two sources of variance in Eq. (2.12): the cosmic variance of the correlation itself (the $(C_\ell^{Tg})^2$ term), and the fortuitous coincidences that arise between the CMB temperature and the galaxy distribution (the $C_\ell^{gg} C_\ell^{TT}$ term).

It follows then that the theoretical signal-to-noise ratio (S/N) of this iSW detection for a given multipole range $[\ell_{\min}, \ell_{\max}]$ is:

$$\left(\frac{S}{N}\right)^2 = \sum_{\ell=\ell_{\min}}^{\ell_{\max}} \frac{(C_\ell^{Tg})^2}{\sigma^2[C_\ell^{Tg}]} = \sum_{\ell=\ell_{\min}}^{\ell_{\max}} (2\ell+1) \frac{(C_\ell^{Tg})^2}{(C_\ell^{Tg})^2 + C_\ell^{gg} C_\ell^{TT}}. \tag{2.14}$$

The cumulative character of the S/N is due to the fact that, in the ideal case we consider here (Gaussian fields, full-sky maps), the power spectrum estimates at different multipoles are independent from one another (no off-diagonal terms in the covariance matrix).[1] For illustration purposes, I will consider two fictitious full-sky surveys, one with a selection function (see Eq. 2.6) similar to the eighth data release (DR8) of the Sloan Digital Sky Survey (SDSS, [4]) and one with a higher median redshift (both are shown in Fig. 2.1). Assuming the best-fit ΛCDM cosmology from *Planck* [38], I calculate the theoretical auto-correlation (C_ℓ^{gg}) and cross-correlation (C_ℓ^{Tg}) spectra for such surveys. To do so, I use a modified version of the cosmological code CMBFAST [42] named CROSS_CMBFAST [12]. This code computes the iSW-correlation power spectrum and the 2-point angular iSW-correlation function for a given galaxy window function, and is limited to flat geometry. It includes dark energy models specified by a constant equation of state or a linear parametrisation in the scale factor and a constant sound speed. In addition to the auto- and cross-correlation spectra, the computation also includes the usual CMBFAST outputs, i.e. the CMB temperature and polarisation spectra.

I illustrate in Fig. 2.2 the resulting cross-correlation spectra and the ideal S/N estimation defined in Eq. (2.14): as we can see, the signal peaks in the $\ell = 10$–20 multipole range and quickly falls off at smaller scales. Similarly, the largest contributions to the signal-to-noise ratio come mainly from the lowest multipoles, with the total S/N quickly reaching a plateau: this shows that most of the significant signal is below $\ell = 100$, at that is it pointless to consider higher multipoles in cross-correlation studies. It is important to notice here the crucial influence of the redshift range (through the selection function) covered by the surveys considered, as already intuited at the end of Sect. 2.1.1. Here, we can witness that the S/N is substantially higher for the survey with the higher median redshift compared to the SDSS-like one (4.5 vs. 3.3 for the cumulative S/N).

---

[1] This is often not exactly true when working with real datasets (due to partial sky coverage, non-Gaussian contamination, etc.) and has to be accounted for properly, e.g. using Monte-Carlo simulations (see later in this chapter).
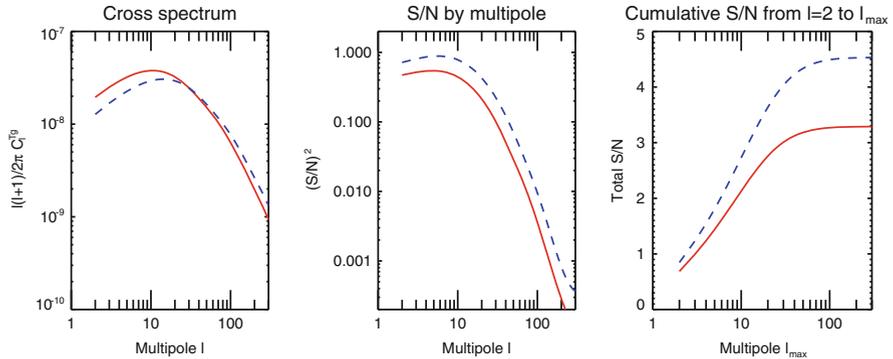
Fig. 2.2 Theoretical cross-correlation results for the SDSS-like survey (*red solid curves*) and a second one with higher median redshift (*blue dashed curves*). *Left panel* theoretical angular cross power spectra of the CMB-galaxy correlation. *Middle panel* contribution to the total squared S/N of the signal as a function of the multipole ℓ. *Right panel* total cumulative S/N of the cross-correlation signal for the multipole range $\ell = [2, \ell_{max}]$ as a function of ℓ_{max}

The above test cases give us an idea about the expected characteristics of the cross-correlation in terms of the signal itself and its detectability; however, the predictive tools used here can be extended to any given survey (through the selection function) and to a large variety of cosmologies (included in the CROSS_CMBFAST code). We can now wonder: how can we use these tools to constrain cosmological models and learn more about the Dark Energy? There are several tests that we can think of and that are used in the literature; I will briefly present the most widely used without diving too deep into calculations (a review of these methods can be found in [14]). The first one is actually independent of the cosmology and does not assume any kind of DE model (besides its existence itself): it consists in measuring the cross-correlation on the data and checking how much it departs from a scenario with no correlations at all, i.e. without Dark Energy. This is the so-called "null hypothesis", and basically consists in performing a χ^2 test on the measured correlation with respect to a "null model", i.e. a model with zero correlation. If the test shows a significant deviation from the null model, then it constitutes a proof of the existence of the Dark Energy (assuming a flat Universe) but does not give any additional information beyond that. In a second time, we can go further and try to check how the data fares against the prediction from a given cosmological model: using a similar χ^2 test, we can compare the measured correlation with the theoretical (non-zero) one predicted beforehand using the framework that I described above. Although this test cannot give a definitive answer on whether a given model is "the right one", inversely it can invalidate the model if the computed χ^2 is too high. A third and last approach, called "amplitude fitting" (or "template matching"), combines both previous ideas: considering a cosmological model (often the fiducial one that we try to (in)validate), we construct a "template" by multiplying its associated theoretical correlation by an "amplitude factor". The method consists then in computing the amplitude

(and associated error) that fits the best the measured correlation: a value close to 1 then indicates an agreement with the underlying model. If this amplitude simultaneously shows a sufficiently large S/N (i.e. a small associated error), it intrinsically disproves the null hypothesis and strongly hints at the presence of Dark Energy.

We can already intuit here that the presence of sources of noise and partial sky coverage will complicate the extraction of the iSW signal. The next paragraph will be an overview of the current state of the detection of the iSW effect through cross-correlation techniques.

2.1.4 State-of-the-Art of the iSW Detection

In the literature, many attempts have been made to detect an iSW signal through the cross-correlation of the CMB with galaxy surveys, with varying degrees of success. I do not intend to summarize here all the results so far, since such a type of compilation has already been made in the past: for reference, in Table 1 of [14] the authors present a "meta-analysis" of iSW detections (up to the publication of their work) and their reported statistical significance. Over the last decade a large variety of surveys has been explored, exploiting the whole spectrum of light: X-ray [9, XRB survey]; optical [1, 3, SDSS], near infrared [23, 2MASS]; radio [10, NVSS]. The cross-correlation studies also followed the evolution of CMB observations, from the first detected anisotropies of the Cosmic Microwave Background Explorer [5, COBE] to the succession of releases by the Wilkinson Microwave Anisotropy Probe [43, WMAP] and the latest publication of the *Planck* satellite [37].

What strikes the most when reviewing the current results of the literature are their wide diversity, as the reported significances range from negligible [40] to $4.4\,\sigma$ [18]. Although these differences can be partly attributed to the specific features[2] of each survey (and CMB maps) and to the methods[3] used by the authors of each work, some puzzling discrepancies are still present. Indeed, a couple of analyses based of very similar (if not identical) datasets have yielded contradictory conclusions on the level of detection of an iSW signal (see [14], again for reference), while other works reported a signal at odds with ΛCDM expectations ($1\,\sigma$ excess in [18], $2\,\sigma$ in [21]).

This intriguing situation may find its source in the characteristics of the current (and past) generation of surveys, or rather, their shortcomings. Indeed, for each of them, a trade-off has often been made between the deepness of the survey and its coverage, resulting in the end in the observation of a small total volume of the Universe. Unfortunately, this is a crucial point for the iSW studies and has a dramatic impact on the potential S/N of the detection: considering a smaller volume results

---

[2] In terms of the redshift range of the surveyed objects and the fraction of the sky covered by the survey.

[3] Not all of the cross-correlation works base their analysis in harmonic space, as some prefer working in real space or using wavelets; although the tools and estimators are different, the different tests (χ^2, null hypothesis,…) remain however the same and similar to those described in Sect. 2.1.3.

simultaneously in a decrease of the theoretical signal and an increase of its variance. Coupled to this is the intrinsic weakness of the signal, and the presence of galactic foregrounds in both the CMB and the LSS maps that also mask crucial large scale data and can introduce spurious correlations. Therefore, any method claiming to detect the iSW effect may not be as thorough as required in accounting for missing data. Ideally, any reported detection level should be independent of any assumption or particular cosmology. Very recent works aimed at a better control of systematics (see e.g. [20], for a thorough review of potential contaminants in SDSS), while others made use of the latest releases of both CMB and LSS data [36]. In both cases, the correlations were found to have much less discrepancies with ΛCDM and a good agreement with their expected amplitude (cf. "amplitude fitting" tests in Sect. 2.1.3), but with relatively low significances ($1.5 - 3\,\sigma$ range).

It is quite established that current surveys are far from ideal: to further the point, [2] showed that obtaining a near optimal iSW detection (at a $\sim 5\,\sigma$ level) would require an all-sky survey with about 10 million galaxies almost uniformly distributed within $0 < z < 1$, with systematics below $\sim$0.1 % and systematic errors in redshift estimates $<$0.05—again, far from the characteristics of currently available datasets. However, the next generation of probes will address many of the current shortcomings: surveys such as Euclid, Pan-STARRS or LSST will get much closer to the ideal requirements for the detection of the iSW effect (see [13, 14] for forecasts) and will also—needless to say—improve our knowledge of a vast range of cosmological topics. For the time being, considering the landscape depicted above, there are a few possible orientations that can be taken regarding iSW studies: the first is to make the most of the current available data and has already been explored quite extensively, notably through the combination of surveys in the context of LSS/CMB cross-correlation (see e.g. [18, 21]). However, some innovative ways of using the LSS and CMB are also explored, such as the study of the impact of individual structures in the CMB that I will develop in Chap. 3. A second possible line of research would be to anticipate the future release of data from the next generation of surveys and develop optimised tools in advance, in order to make the most of the improved accuracy and avoid potential biases and caveats: this will be the topic of the next section of this chapter (Sect. 2.2). Finally, considering the great time scales involved in the schedule of the aforementioned surveys, an interesting approach would also be to look for alternative, promising, and already available tracers of matter (and gravitational potentials) to cross-correlate to the CMB: this will be the central point of the Sect. 2.3, through the use of the Cosmic Infrared Background.

2.2 Optimising the Cross-Correlation for iSW Detection

The work that I will present in this section originates from a pre-thesis internship: it aims at exploring, optimising and proof-testing a protocol that I devised for the exploitation of the iSW effect, through CMB/LSS cross-correlation and in the context of next-generation surveys. With this prospective work I have no pretension of giving

here a complete and exhaustive method for the analysis of these datasets, but I will focus on a few crucial points that allowed me to point out some interesting features and potential problems of this kind of analysis.

2.2.1 Context and Tools

Being entirely theoretical in nature, this study required me to simulate realistic datasets and analyse them thoroughly—a task that required the use of accurate numerical code and tools. The first requirement was to be able to predict, for a given "next-gen-like" survey and a cosmology, the theoretical cross-correlation between the CMB and these LSS, as well as their auto-correlation. If we assume that these two fields are Gaussian in nature (which I will do), the only quantities that we need to compute are their angular cross- and (respective) auto-correlation power spectra in harmonic space. To do so I used the numerical code CROSS_CMBFAST that I already mentioned in Sect. 2.1.3.

To develop a little further on this particular code, its specificities, and qualities, it has the particular feature of adopting a top-down approach, starting from the primordial spectrum of fluctuations. This differs from other similar studies where the starting point is the present matter power spectrum (which is evolved backward in order to find its correlation with the CMB). The approach of CROSS_CMBFAST unifies the treatment of CMB and matter power spectra, and is more convenient for taking full account of possible fluctuations in the Dark Energy. Moreover, since all perturbations are evolved numerically with the CMBFAST code, it does not resort to the frequently used approximate analytical expressions for the growth function (mentioned at the end of Sect. 2.1.2), nor to the also commonly used "Limber approximation" for small angles [28].

In order to thoroughly test my protocol (described in the next section) against "reallife" situations, I also needed to work with realistic simulations of CMB and LSS maps. For the creation and efficient exploitation of such maps I used the HEALPix (Hierarchical Equal Area isoLatitude Pixelization) package [19]: this suite of codes is based on a specific pixelization scheme of a spherical surface that is particularly suited for fast and accurate statistical and astrophysical analysis of massive full-sky data sets. It contains all the required tools to produce maps based on Gaussian realisations of a given power spectrum, as well as perform a decomposition of a map into its spherical harmonic coefficients (and consequently its angular power spectrum).

2.2.2 Simulations, Covariance Estimation and Parameter Recovery

The main idea that motivated my work presented here was to investigate a way to recover as much information as possible on the Dark Energy from a single survey
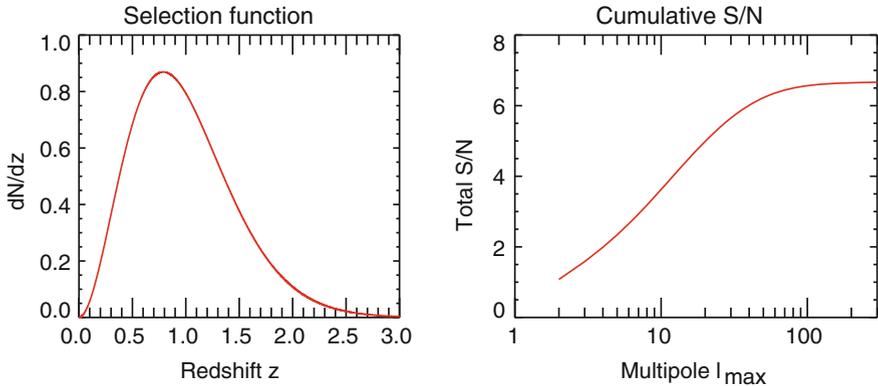
Fig. 2.3 *Left panel* normalised redshift distribution of my fiducial survey, following a similar analytical form as the one mentioned in Fig. 2.1. The parameters chosen here are $z_0 = 0.95$, $m = 1.9$ and $\beta = 1.5$. *Right panel* total cumulative S/N of the theoretical cross-correlation signal for this survey as a function of ℓ_{max}

with ideal characteristics (full-sky coverage, large redshift range, noiseless) and its cross-correlation with the CMB. I started by fixing a reference cosmological model, that I chose to be at that time the best-fit ΛCDM model derived from the WMAP 7-year data [27]. In the meantime, I also fixed the characteristics of the mock survey that I considered, through the choice of its selection function—the only quantity needed to model a survey (in the linear regime) beside the cosmology; I illustrate this survey in Fig. 2.3.

In this particular context, the information that we are trying to obtain from CMB/LSS correlation becomes simply the density parameter Ω_Λ of the Dark Energy. From there, the protocol that I devised for the Ω_Λ reconstruction is the following:

- I consider a CMB map and a galaxy overdensity map;
- I assume that the selection function of the galaxy survey has already been estimated beforehand by other means;
- I use the tools of the HEALPix package to extract the cross-correlation power spectrum C_ℓ^{Tg}(data) from the two "data" maps;
- I use the CROSS_CMBFAST code to compute the expected theoretical cross power spectra $C_\ell^{Tg}(\Omega_\Lambda)$ for a large range of Ω_Λ values (between 0.3 and 0.9);
- Using a χ^2 test, I look for the model that fits best the measured spectrum and keep the corresponding Ω_Λ as the reconstructed DE density parameter.

Although this protocol is fairly straightforward, there is a specific point that requires a particular attention: the computation of the χ^2, or more specifically the computation of the covariance matrix involved in the calculation. Indeed, for a given measured spectrum C_ℓ^{data}, the χ^2 of a given model C_ℓ^{theo} is given by:

$$\chi^2 = {}^{\mathrm{T}}\left(C_\ell^{\text{data}} - C_\ell^{\text{theo}}\right) \mathcal{M}^{-1} \left(C_\ell^{\text{data}} - C_\ell^{\text{theo}}\right), \tag{2.15}$$

$\mathcal{M}$ being the covariance matrix, each of its elements containing the covariance of the cross-correlation power spectrum C_ℓ^{data} between two multipoles, i.e.:

$$\mathcal{M}_{ij} = \text{Covar}\left(C_{\ell_i}^{\text{data}}, C_{\ell_j}^{\text{data}}\right). \qquad (2.16)$$

The tricky part here is that there are several methods for computing this matrix, that differ quite substantially from each other. For the simplest approach, we suppose that the survey considered is quite ideal and that there is no correlation between multipoles, which is generally true only for full-sky surveys. It follows that the covariance matrix is diagonal and that an analytical approach is sufficient to compute its elements using the theoretical covariance already shown in Eq. (2.12) of Sect. 2.1.3. However, this approach is inadequate for more realistic scenarii, since galaxy surveys are often plagued with partial sky coverage, and contaminants that may further complicate the analysis. Although I do consider an ideal survey here, it is still necessary to go beyond this simple case, especially from the perspective of the potential application of the protocol to real datasets.

For these reasons, two alternative methods have been devised, often called "MC1" and "MC2" in the literature (see e.g. [17] for an application of these approaches). They are both based on Monte-Carlo techniques (hence the "MC") and consist in generating a very large number of as-accurate-as-possible maps (CMB and/or LSS) of which we derive the covariance matrix of the cross-correlation. To be more precise, let us start with the MC2 method: assuming a fiducial model, the principle of the method is to generate N pairs[4] of galaxy and CMB maps with the same characteristics as the original ones and their expected cross-correlation. In practice, these maps are obtained by generating Gaussian realisations of the theoretical power spectra derived from the fiducial cosmology and the knowledge of the considered survey (for the details of the calculations, see e.g. [17]). We then reproduce in these maps the known features of the original maps, whether it is a partial sky coverage (with a mask) or a known contaminant. From there, we compute the cross power spectrum for each of these N pairs of maps, and then derive each element of the covariance matrix according to:

$$\mathcal{M}_{ij} = \text{Covar}\left(C_{\ell_i}, C_{\ell_j}\right) = \frac{1}{N}\sum_{k=1}^{N}[C_{\ell_i}^k - \overline{C}_{\ell_i}][C_{\ell_j}^k - \overline{C}_{\ell_j}] \qquad (2.17)$$

where the C_ℓ^k are the cross power spectra of each simulated pair, and $\overline{C}_\ell$ is the average of these power spectra over the N pairs. Now the only difference between the two Monte-Carlo approaches is that in the MC1 method, only maps of the CMB are generated. They are then correlated to the true galaxy map, and from these correlations are computed the elements of the covariance matrix.

[4] With N typically of the order of ten thousands.

Fig. 2.4 Normalised histograms of the reconstructed values of Ω_Λ by the protocol described in the text, using either the MC1 (*blue*) or the MC2 (*green*) method for the computation of the covariance matrix involved in the χ^2 test. The true input value of Ω_Λ is indicated by the *red vertical line* and is equal to the best-fit value of the parameter from WMAP7, $\Omega_\Lambda = 0.734$
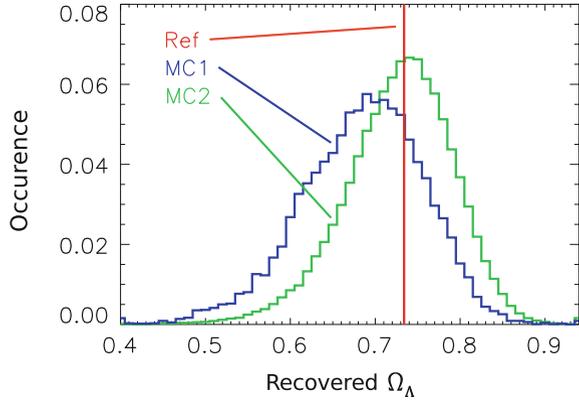
The MC1 is the most widely used estimator in the literature: reasonably fast to implement, it accounts for the cosmic variance and the accidental CMB/LSS correlations, supposedly the primary source of error. However it does not account for the variance in the density maps since only CMB maps are randomly generated. The MC2 method alleviates these problems by also generating random density maps (based on a fiducial cosmology and the selection function). However, this method is more time demanding (more maps to generate) and requires a robust knowledge of the features of the considered survey (including its selection function and systematics) to accurately simulate realisations of it. Finally, a shortcoming common to both approaches is that they are model dependent and could fail if the data model is poorly understood (e.g. non-Gaussianity of the maps).

In the remainder of this section, I will focus on how these two MC methods fare in the context of the objective of my protocol, i.e. the reconstruction of the Ω_Λ parameter. After fixing the input cosmology, I thoroughly tested each method by repeating the following steps thousands of times:

- Simulation of pair of CMB/density maps with the correlation expected from the chosen cosmology;
- Application of my protocol described earlier to search for the best-fitting Ω_Λ;
- Storage of the reconstructed Ω_Λ value in a histogram.

A comparison of the recovered Ω_Λ values for the two MC methods, along with the input value (here $\Omega_\Lambda^{\text{input}} = 0.734$, WMAP7 best-fit), is shown in Fig. 2.4. The most striking feature here is the bias of the MC1 method towards smaller Ω_Λ values: although the width of both distributions is quite large (with a noticeably larger trend for MC1), the peak for the MC1 method is still clearly shifted (to $\Omega_\Lambda \sim 0.69$), while the MC2 method shows a much better agreement (peak at $\Omega_\Lambda \sim 0.74$) with the input value.

We can intuit an explanation for the observed difference between the two methods, which revolves around the asymmetrical nature of the MC1 approach. In the ideal case considered here (no noise, full-sky maps), there are but very little correlations

between multipoles: the covariance matrix is nearly diagonal, and the values of its elements tend to the analytical expressions given by Eq. (2.12). Due to the intrinsic weakness of the iSW signal (i.e. C_ℓ^{Tg} is relatively small), we can simplify this equation even further and write:

$$\sigma^2[C_\ell^{Tg}] \sim \frac{C_\ell^{gg} C_\ell^{TT}}{2\ell + 1}. \tag{2.18}$$

In the MC2 method, the matrix is computed once and for all using thousands of simulated maps, and is completely independent of the data considered. However, in the MC1 approach, only one density map (the data) is used in the computation of the covariance matrix. As a consequence, the resulting variance will be directly proportional to the spectrum of this particular map, according to Eq. (2.18). Over the thousands of repeated tests that we performed, the Gaussian realisations of this spectrum will be sometimes lower, sometimes higher than the fiducial model, in equal proportions. Then, as said above the "low" realisations will give a smaller covariance matrix (i.e. smaller errors bars), pulling down the Ω_Λ value preferred by the χ^2 test. On the other hand, the "high" realisations will have larger error bars, and therefore will not have the same "pull" towards high Ω_Λ values. This asymmetry in the MC1 method is the most likely explanation for the over-abundance of smaller recovered Ω_Λ values when we repeat the test many times.

It stems from the previous analysis that the MC2 method should be the preferable choice in order to avoid a possible bias on the covariance matrix. However the MC1 approach remains more time-efficient and does not require a deep knowledge of the considered survey and its systematics, a point which is often not properly controlled. In any case, whether we want to generate CMB or density maps, some form of model dependence is always present and has to be accounted for in the interpretation of any result. In the rest of Sect. 2.2, I will use the MC2 method for the covariance matrix, bearing in mind its advantages (no bias in the recovered parameters) and limitations (additional computation time and good knowledge of the considered survey required).

2.2.3 Tomography for iSW Studies

By design, the classic approach to the cross-correlation intrinsically misses part of the information about the iSW effect itself, as it looks for an integrated signal along the whole line of sight with no regard for its redshift dependence. Such loss of information could prove critical, especially in the context of time-varying models of Dark Energy (e.g. quintessence). But even for a cosmological constant, the iSW signal recovered by CMB/LSS cross-correlation will have a redshift dependence due to sheer volume effects (already discussed at the end of Sect. 2.1.1). In the case of the mock survey that I used in the previous section, the large redshift coverage and potential iSW S/N (illustrated in Fig. 2.3) might also be squandered by the classic approach described previously. Now the question that I will develop in the present section is the following: Could it be possible to modify my protocol of

Fig. 2.5 Illustration of the redshift slicing of the ideal survey (*blue solid line*) considered in Fig. 2.3. The selection function is divided using "top-hat" functions, into five redshift intervals (*red dashed lines*): $z \in [0, 0.5], [0.5, 1], [1, 1.5], [1.5, 2]$ and $[2, \infty]$
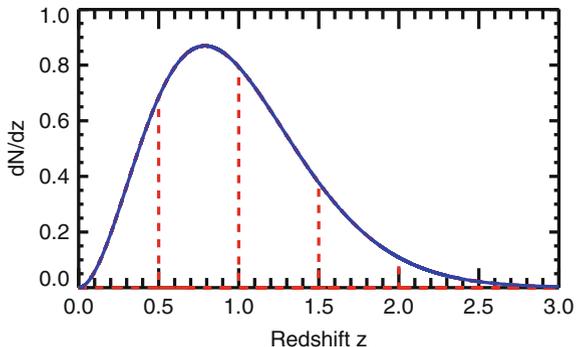
cross-correlation analysis in order to exploit as much as possible the full potential of such surveys?

The main idea here is to take advantage of the deepness in redshift of the survey: one of the ways that we can think of in order to do so is to "slice" the considered survey. More precisely, instead of integrating the whole distribution of matter contained within it to get one density map, I divided the survey into redshift bins and constructed a density map for each. We intuit here already that some care has to be taken when choosing the shape and the number of bins. The simplest way to slice a survey would be to simply divide its selection function into redshift intervals ("top-hat" cuts), as shown in Fig. 2.5. However, one has to remember that the redshift of survey objects often have errors in their determination. Hu and Scranton [22] supposed that photometric redshift estimates are Gaussian distributed with an RMS fluctuation $\sigma(z)$ that increases with redshift as $\sigma(z) = \sigma_{max}(1 + z)/(1 + z_{max})$, i.e. the bin sizes are chosen to increase proportionally to the error. The two free parameters z_{max} and σ_{max} roughly correspond to the end of the redshift range of the survey and the redshift estimate errors around z_{max}, respectively. Then, a top-hat cut in the $[z_1, z_2]$ interval in estimated redshift becomes a smooth overlapping distribution in actual redshift:

$$n_{z_1 \to z_2}(z) = \frac{1}{2} \frac{dN}{dz}(z) \left[\text{erfc} \left(\frac{z_1 - z}{\sigma(z)\sqrt{2}} \right) - \text{erfc} \left(\frac{z_2 - z}{\sigma(z)\sqrt{2}} \right) \right], \qquad (2.19)$$

where dN/dz is the original, full selection function of the survey. An illustration of this redshift slicing is shown in Fig. 2.6, which used the same redshift intervals as the slicing shown in Fig. 2.5. Given this formalism, one could a priori divide a given survey in as many slices as wanted.

Let us now refocus on our objective: the exploitation of a survey for CMB/LSS correlations. We have now a different starting point than in the previous section, as we have now as many density maps as there are redshift slices. From there, the general progression of the protocol remains the same: if we suppose that we divided the considered survey in N redshift bins, we first extract the N angular cross-correlation power spectra of each density map with the CMB. In parallel to this, we compute

Fig. 2.6 Redshift slicing of
the same survey of Fig. 2.3
(*blue solid line*), using the
expression in Eq. (2.19)
for the redshift bins. The
five divisions (*red dashed
lines*) have the same bounds
(the z_1 and z_2 in Eq. (2.19)
expression) as the slices in
Fig. 2.5
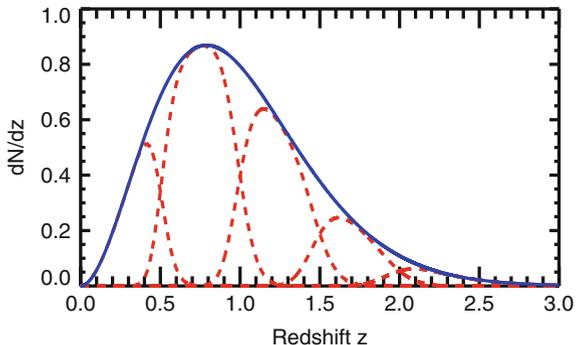
the expected theoretical cross-correlation of each slice using the knowledge on the
survey selection function coupled to the expression in Eq. (2.19). And finally, we
perform a χ^2 test on all the spectra simultaneously to determine the Ω_Λ that best
fits all the data. However, one has to be careful once again in the calculation of the
covariance matrix: indeed, as it can be seen clearly in Fig. 2.6, there exists a non-
negligible overlap between the different redshift slices. This overlap is practically
unavoidable as it originates from the (always present) errors in redshift estimations.
It creates therefore correlations between the resulting slices of density maps, which
in turns induces a form of redundancy when cross-correlating all these partial maps
with the CMB. One can then end up with unrealistic results with underestimated
errors bars on the recovered parameters. To account for this "leakage" of the redshift
bins into one another, I proceeded with some additional steps in the computation
of the covariance matrix: when applying the MC2 method, after generating the N
partial density maps (and the CMB map), not only did I correlate each map with
the CMB, but I also correlated each possible pair of density maps. The end result
of this is a non-trivial covariance matrix, that is no longer diagonal even in the case
of an ideal survey, as the non-diagonal terms contain the correlations between the
overlapping redshift slices of the matter distribution.

Now in order to assess the impact of the redshift slicing on the efficiency of
the Ω_Λ reconstruction, I subjected my revised protocol to the same repeated test
that I performed in the previous section. A particular point that I did not mention
in the previous paragraph is that the simulation of accurate data maps for this test
also has to nclude all the correlations between redshift slices: I therefore had to
extensively modify the CROSS_CMBFAST code so that it would compute not only
the theoretical cross-correlation between the CMB and density slices, but also the
one between each pair of slices. After that was taken care of, I tested my protocol and
explored several slicing strategies, focusing especially on the choice of the number
of slices: I compared five numbers of slices ranging from 0 to 10, and present the
results of the Ω_Λ reconstruction in Fig. 2.7.

As we can see there, even a simple division of the survey into two redshift slices
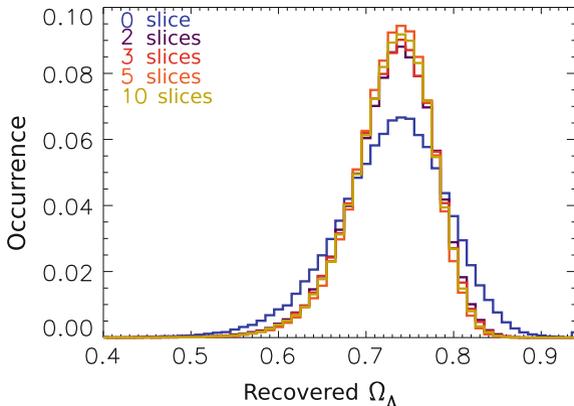yields an appreciable improvement of the Ω_Λ reconstruction, with a significant

Fig. 2.7 Normalised histograms of the reconstructed values of Ω_Λ by the protocol described in the text, in the context of several slicing choices of the original survey. The number of slices are 0 (meaning the whole redshift range, *blue* histogram), 2 (*purple*), 3 (*red*), 5 (*orange*) and 10 (*yellow*), with the histograms becoming thinner as the number of slices increases. The true input value of Ω_Λ is again the WMAP7 best-fit value $\Omega_\Lambda = 0.734$

tightening of the distribution: the standard deviation decreases from $\Delta\Omega_\Lambda \sim 0.120$ to $\Delta\Omega_\Lambda \sim 0.093$ (drop of $\sim 25\%$). The width of the distribution gets even smaller when I further increase the number of slices, although the improvement becomes less and less noticeable, reaching a plateau at some point (particularly visible also in Fig. 2.8). This shows that there is no need to divide the original survey into too many redshift bins, as the information that we can recover on the Ω_Λ parameter seems to reach a limit.

One could possibly ignore this fact (as it has no harmful effect on the reconstruction) and still carelessly consider as many slices as possible, but another important parameter has to be taken into account. As the number of redshift bins increases, so does the number of "density" maps that we have to consider, especially when computing the covariance matrix. In more practical terms, the more maps we have, the more auto- and cross-correlations we have to perform: one CMB maps and N redshift bins yield N (auto) and $N(N+1)/2$ (cross) spectra. This has two different effects, one of them being the increase in computation time for the matrix, roughly proportional to N^2 (times the number of simulations). The less obvious effect is that each underlying auto- and cross-spectrum gets weaker as we further divide the survey: this has the impeding consequence of requiring a higher number of simulated maps in order for the computed covariance matrix to converge. This particular effect is illustrated in Fig. 2.8, where I show how the width of the distribution of reconstructed Ω_Λ evolves with the number of simulated maps used for the covariance matrix computation. We witness there that the number of simulations required for reaching a plateau does indeed increase for larger numbers of slices, from only 2,000 for the original survey to 2–3 times as much for five redshift slices. Combined to the observed convergence of the recovered Ω_Λ distribution, we understand that some form of balance has to be

Fig. 2.8 Standard deviation
of the distribution of recon-
structed Ω_Λ as a function
of the number of sets of
simulated maps used in the
computation of the covari-
ance matrix. The same slicing
choices as in Fig. 2.7 (with
their corresponding colours)
are explored. The *curves* get
lower as the number of slices
increases
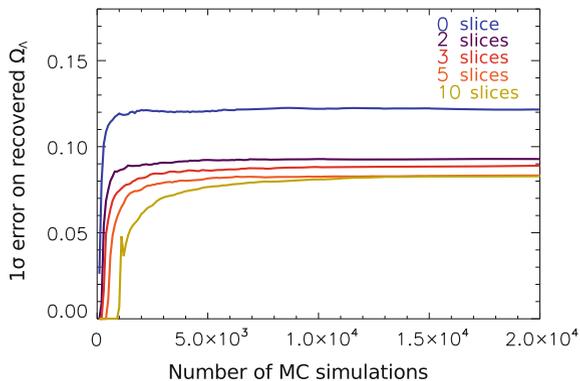
found between the choice of the number of slices (i.e. the precision of the constraints
on Ω_Λ) and the number of required simulations (i.e. the computation time needed).

In the end, in this section I only scratched the surface of the topic, as there are
many improvements to my protocol that I could consider for a future work. First, in
the context of real maps and real surveys, a lot of importance has to be given to the
often partial coverage of the sky that has some drastic and intricate effects that affect
the cross-correlation, as well as the presence of foregrounds and contaminants that
further complicates the interpretation of the signal. On the other hand one could be
interested in assessing the full constraining power of the iSW effect on cosmology,
without using any underlying model, using the potential of the future generation of
surveys. Also in the context of an optimal exploitation of future datasets, the exact
choice of redshift slices (shapes) would need a particular tailoring to be as adapted as
possible to each surveys (this kind of work has been explored recently by [24]). As
a final note, although single surveys are not appropriate for it as of now, the slicing
study that I performed may already be applicable to current data in the context of
the combination of surveys for iSW detection: indeed, considering several surveys
together with their own selection functions is quite similar to considering several
slices in redshift (sometimes overlapping) of the same underlying distribution of
matter, although some additional care is required to account for the calibration and
systematics of each survey.

2.3 The Cosmic Infrared Background and the iSW Effect

In Sect. 2.2, I briefly explored some methods of exploiting the potential of the next
generation of surveys, optimising the amount of information extracted while taking
care of the possible complications that may appear along the way.

However, no matter how promising these future missions may look, the time scales
involved in their schedules does not allow much work other than in the realm of

simulations and predictions. As an example, one of the most exciting future satellite, *Euclid*, has its most optimistic launch date scheduled for 2020! The ground-based LSST also has a similar schedule. Other experiments may have a closer "due date" (e.g. in the course of 2015 for DES) but still measured in terms of years. On the other end of the spectrum, recent works have been aimed at exploiting as much as possible the current generation of surveys, including attempts at combining several datasets and revising their work at each new data release. In this context, I chose to sit in the middle ground: I searched for an alternative tracer of matter that could be correlated to the CMB and that would be both already available in some form, but also with a better potential that the currently used datasets. Such alternatives have already been explored in the literature, for example through the correlation between the iSW effect and the thermal Sunyaev-Zel'dovich effect [44], the latter being the result of a boost received by CMB photons from high energy cluster electrons by inverse Compton scattering: a non-zero correlation is indeed expected between the two effects as these electrons sit in the gravitational potentials that produce the iSW effect. An interesting aspect of this approach is that it constitutes a "CMB-only" detection of the iSW effect (although with some underlying assumptions). In a similar fashion, the [36] explored for the first time the correlation of the CMB with the reconstructed gravitational lensing potential extracted from the CMB data itself.

As for myself, using some of the assets and expertise available at the IAS, I set my eyes on an original tracer of matter, which is one of the other few backgrounds of the Universe: the Cosmic Infrared Background, whose characteristics and exploitation for iSW studies I will describe in the next sections.

2.3.1 The Blurry Red Light

First discovered by [39], the Cosmic Infrared Background (CIB hereafter) is visible roughly from 10 to 1,000 μm in wavelength and is one of the backgrounds present in the Universe (see Fig. 2.9). This particular background, present in every survey that covers infrared wavelengths, arises from the accumulated electromagnetic emissions from star-forming galaxies distributed across a large redshift range. It finds its origin in the smallest and farthest galaxies (which cannot be resolved by telescopes and their finite resolution) and in the densest population of galaxies (whose observation is confusion-limited), which thus appear as a blurry background. The earliest epoch for the production of the CIB is thought to be when star formation first began at the end of the Dark Ages (and the onset of the reionisation epoch); contributions continued through the present epoch, including our current Dark Energy dominated era. Similarly to the CMB, the CIB also features anisotropies (first detected and discussed in [26, 29]) that are shaped by the distribution and clustering of these infrared galaxies in the Universe.

Ever since its discovery, many efforts have been deployed to detect the CIB and its anisotropies with increasing precision, as they contain a lot of information about the star and galaxy formation histories, including their clustering processes.
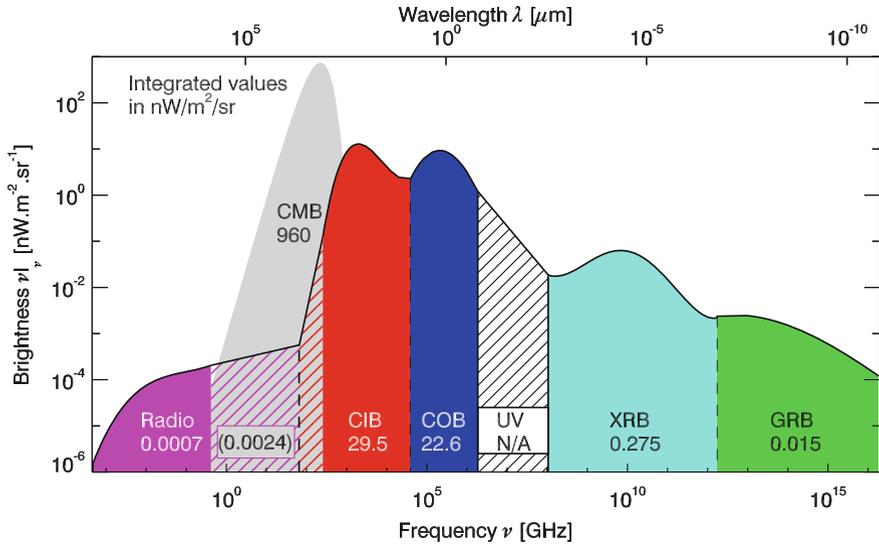
Fig. 2.9 Schematic spectral energy distributions (SED) of the most important (by intensity) backgrounds in the Universe, and their approximate *brightness* in nW· m^{-2}· sr^{-1} written in their respective zones. From *right* to *left* the radio background, the CMB, the CIB, the cosmic optical background (COB), the UV background, the X-ray background (XRB) and the gamma-ray background (GRB). Image courtesy of H. Dole

The most recent papers on the CIB anisotropies use sophisticated models which compute the halo occupation distribution (HOD, see e.g. [11, 31]) and the Dark Matter halos properties, in order to predict the power spectrum of these anisotropies (see e.g. [32]).

2.3.2 The CIB as an Alternative Tracer of Matter

We can easily understand that the anisotropies of the CIB are underlined by the galaxy density field and thus the matter density fluctuations: it is therefore reasonable to expect that the CIB has a positive correlation with the CMB through the iSW effect. We should also bear in mind that it contains contributions from galaxies over a very large range of redshift (up to $z \sim 7$, with a peak of emission around $z \sim 2$), and that it is contained in many past and current surveys (such as IRAS or the newly released *Planck*, both covering IR frequencies) that have often a large coverage of the sky. These two points are the reasons that motivated my choice of the CIB as a tracer of matter, as they make the CIB a particularly suitable and promising candidate for the iSW detection by cross-correlation with the CMB. However, the extraction of the CIB from a survey remains a challenging and delicate task that is still ongoing to this

day: this led me to focus first on a theoretical study of the CIB-CMB cross-correlation through the iSW effect, and an assessment of its detectability.

As mentioned before, the most recent works focused on the studies of CIB anisotropies used sophisticated HOD to predict and describe them. Such models are particularly useful when describing the small, non-linear scales of the CIB. In the context of the iSW effect, we focus on much larger scales which is why I used a simpler model for the CIB, similar to the description made by [25]. The general definition of the CIB anisotropies at a given frequency ν and in a given direction $\hat{\mathbf{n}}$ can be then written as the following line-of-sight integral:

$$\delta T_{\mathrm{CIB}}(\hat{\mathbf{n}}, \nu) = \int_{z_{\mathrm{far}}}^{0} dz \, a(z) \, \delta j(\hat{\mathbf{n}}, \nu, z) \tag{2.20}$$

with δj being the emissivity anisotropies of the CIB. The integration is made from some initial time z_{far} before star formation began to our location at $z = 0$. In their work, Knox et al. hypothesized that the CIB anisotropies are direct tracers of the matter density fluctuations $\delta = \delta \rho_m / \bar{\rho}_m$, up to a bias factor. Therefore, the previous expression becomes an integral of the product between a mean far infrared (FIR) emissivity and the matter density fluctuation field:

$$\delta T_{\mathrm{CIB}}(\hat{\mathbf{n}}, \nu) = \int_{z_{\mathrm{far}}}^{0} dz \, a(z) \, b_j(\nu, z) \, \bar{j}(\nu, z) \, \delta(\hat{\mathbf{n}}, z). \tag{2.21}$$

Here the quantity $b_j(\nu, z)$ is some form of bias that links the matter distribution and the emissivity. It is frequency- and redshift-dependent and is here defined by:

$$\frac{\delta j(\hat{\mathbf{n}}, \nu, z)}{\bar{j}(\nu, z)} = b_j(\nu, z) \, \delta(\hat{\mathbf{n}}, z) \tag{2.22}$$

with $\bar{j}(\nu, z)$ being the mean emissivity per comoving unit volume at frequency ν as a function of redshift z. We can observe here some similarities between Eq. (2.21) and the expression in Eq. (2.6) of the galaxy density contrast for the usual surveys. However the selection function has been replaced by the emissivity function, with an additional scale factor as we no longer consider individual objects that we count on the sky, but rather the light that the objects emit (which of course gets redshifted in the time it takes to reach us). There is also here an added dependence on the frequency through the emissivity: although overlapping, it is different populations of objects that we observe when surveying different wavelengths. This emissivity function is derived using an empirical, parametric model based on number counts of galaxies: at the time of this study I used the work of [6], although a recent update has been published since (see [7]).
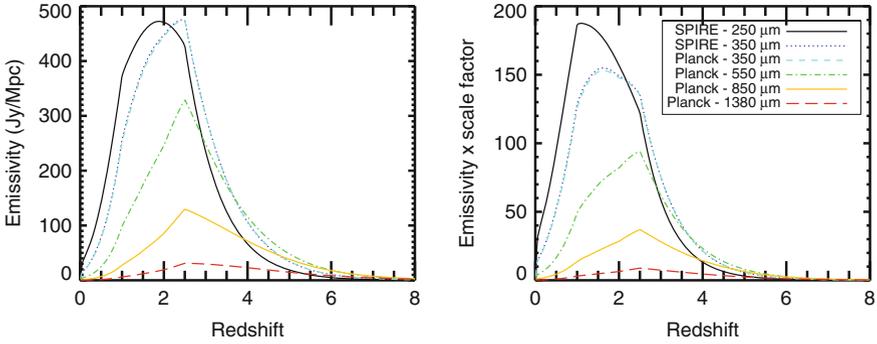
Fig. 2.10 *Left panel* emissivities as functions of redshift for six experiment/frequency pairs, computed using the model of [7]. Notice the very slight difference between instruments at 350 μm due to a difference in bandpass. *Right panel* same functions multiplied by the scale factor $a(z)$, or equivalently $1/(1 + z)$, for comparison purposes with selection functions of galaxy surveys

Examples of such emissivities are shown in Fig. 2.10 (left panel): it should be noted that they not only depend on the frequency but also on the experiment considered, because each instrument possesses a unique bandpass for each frequency and has its own resolution. This results in the observation of a "unique" CIB map by every pair of frequency and instrument, corresponding to a population of objects just as unique. The depth in redshift of the CIB is also illustrated in this figure, with emissivity functions that theoretically reach much higher redshifts that any current (and even some future) surveys. We should however keep in mind that, according to Eq. (2.21), these functions have to be multiplied by the scale factor (right panel of Fig. 2.10) in order to be compared to the selection function of classic galaxy surveys. This causes a decrease at high z but even then, they still cover a larger range of redshift than typical galaxy surveys (see e.g. Figs. 2.1 and 2.3 for comparison). We also note that Fig. 2.10 shows significant overlaps in redshift between the emissivity functions at different frequencies which will consequently induce correlations between CIB observations, similarly to those between redshift slices of a same survey (discussed in Sect. 2.2.3). At this point, this overlap already indicates that a combined use of CIB observations at several frequencies may not yield improvements in the detection of the iSW effect.

A last element that needs to be pointed out is the previously mentioned linear bias[5] $b_j(\nu, z)$ present in Eq. (2.21) that I chose constant in redshift here: $b_j(\nu, z) = b_{\text{lin}}(\nu)$. To obtain it for each frequency that I considered, I computed the value of b_{lin} that gives the best agreement between my calculation of the linear CIB power spectrum and those obtained from the *Planck* data [34]. This point will be discussed in more detail in the next subsection.

---

[5] This bias here represents our matter-emissivity bias in Eqs. (2.21) and (2.22) and should not be confused with the widely used galaxy-Dark Matter bias, though ours *does* contain information about how the emitting objects populate Dark Matter halos.

Now, starting from Eq. (2.21) we can use the same formalism as for the CMB-galaxy correlation in order to derive the theoretical CMB-CIB cross-correlation at any given frequency. Following the same steps from Eqs. (2.7) to (2.11), we end up with similar expressions for both the cross- and auto-correlation spectrum of the CIB, except that the galaxy bias is replaced by the emissivity one, and the mean emissivity (times the scale factor) plays the rôle of the selection function. The next section will be dedicated to the computation of the expected value of these modified equations in several contexts.

2.3.3 Predictions on the CIB-CMB Cross-Correlation

2.3.3.1 Computing the Expected Correlations

To compute the expressions mentioned at the end of the previous section, I once again adapted the CROSS_CMBFAST code to use it for this CIB study, performing a number of modifications to suit this new context. In my customised version, for a given cosmology and emissivity function $\bar{j}(\nu, z)$ my code calculates the CIB-CMB angular cross-correlation power spectrum, as well as the predicted auto-correlation power spectrum of the CIB fluctuations.

In Fig. 2.11, I present predictions for the CIB-CMB cross-correlation, at several FIR wavelengths and for different instruments, namely: IRAS at 100 µm, *Herschel* SPIRE at 250, 350 and 500 µm and *Planck* HFI at 350, 550, 850, 1380, and 2097 µm. We note that at 350 µm the SPIRE- and *Planck*-predicted spectra differ slightly from each other, once again due to slight differences in wavelength bandwidth of the two instruments (hence a difference in the emissivities used, see Fig. 2.10).

In a fashion similar to previous galaxy-iSW cross-correlations (see e.g. Fig. 2.2), we note that the cross-correlation peaks around $\ell \simeq 10$–30, and quickly vanishes at higher multipoles. Comparing the signal at the different wavelengths shows that the amplitude of the cross-correlation signal is maximum at a wavelength $\simeq 250$ µm. This is not entirely surprising, since this wavelength roughly corresponds to the maximum of the observed CIB spectral energy distribution (cf. Fig. 2.9).

To get some form of validation for my computed spectra, I compared my predictions for CIB autocorrelation to the measurements of [34], taking the opportunity to address a concern mentioned in the previous section—namely, the determination of the linear emissivity bias. To obtain it at each frequency, we compute the value of this bias that gives the best agreement between my linear CIB power spectrum and the corresponding one obtained from the *Planck* data. I chose to fit the two spectra in the range of multipoles $\ell \in [10, 50]$, where most of the iSW signal resides. This is illustrated in Fig. 2.12 where I plotted the biased and non-biased CIB linear spectra and compared them to the ones from [34] at their four frequencies. Overall, the two sets of spectra show good agreement over the multipoles of interest; the spectra deviate at higher ℓs (starting from $\simeq 100$) due to the rise of non-linearities that I did not account for in my linear model—namely the small-scale correlations between
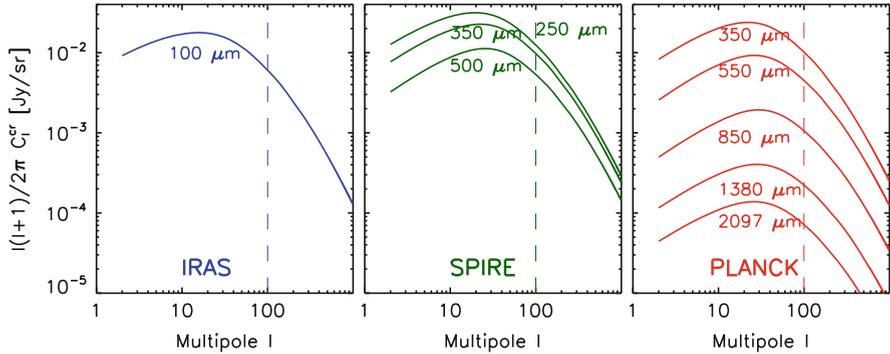
Fig. 2.11 Theoretical angular cross power spectrum of the CIB-CMB correlation calculated for IRAS at 100 μm (*left-hand panel*), for *Herschel* SPIRE between 250 and 500 μm (*central panel*) and *Planck* HFI between 350 and 2,097 μm (*right panel*). The linear bias, b_{lin}, is fixed here to 1 at all frequencies in order to compare the non-biased CIB power spectra. The *vertical dashed line* on each panel marks the upper limit of the multipoles used in our analysis: this choice comes from the vanishing of the iSW signal (see Fig. 2.13) and the rise of non-linearities at higher ℓ

galaxies inside the same halos. The linear bias we obtain this way increases with the wavelength: this is coherent with the fact that as we go deeper into the infrared, the galaxies probed are more luminous at higher z. They reside in more massive and rarer halos, and are therefore more biased.

It should be also noted that the results on the cross-correlation are also not exact at the highest ℓs, as hinted at by the deviation observed in the auto-correlation spectrum. Indeed, the non-linear counterpart to the iSW effect, called the Rees-Sciama effect, contributes at those scales (see [41] for a discussion). However, in our case the linear part of the iSW largely dominates at the peak observed in Fig. 2.11.

2.3.3.2 Detectability Assessment

Following the same progression as for the CMB-galaxy case, I then investigated the detection level of the iSW effect using CMB-CIB cross-correlation by performing a signal-to-noise ratio analysis. Using the power spectra that I computed in the previous section, I can express for each given frequency ν the total signal-to-noise ratio of the iSW detection as follows:

$$\left[\frac{S}{N}\right]^2(\nu) = \sum_{\ell=2}^{\ell_{\text{max}}} (2\ell + 1) \frac{[C_\ell^{\text{cr}}(\nu)]^2}{[C_\ell^{\text{cr}}(\nu)]^2 + C_\ell^{\text{CIB}}(\nu) \times C_\ell^{\text{CMB}}} \qquad (2.23)$$

where $C_\ell^{\text{cr}}(\nu)$ is the CMB-CIB cross-correlation spectrum, and $C_\ell^{\text{CIB}}(\nu)$ the CIB auto-correlation spectra, both at a given frequency ν, while C_ℓ^{CMB} is the CMB auto-correlation spectrum. The total (cumulative) signal-to-noise is summed over
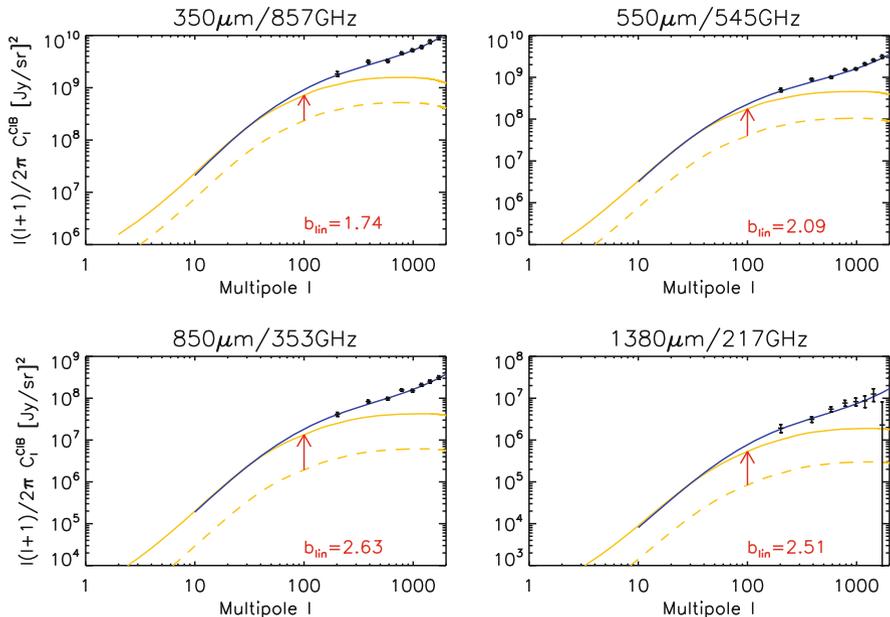
Fig. 2.12 Angular power spectra of the CIB fluctuations at four frequencies of the *Planck*-HFI instrument, as fitted by the *Planck* team (*blue uppermost continuous line*) and by our non-biased models (*dashed yellow line*). For each frequency, we provide in *red* the linear bias which gives the best agreement between the two models, and plot our models taking into account this bias (*solid yellow line*). The data points correspond to measurements obtained by the *Planck* team [34]

multipoles between $\ell = 2$ and $\ell_{max} \leqslant 100$ where the signal has its major contribution (see Fig. 2.11).

In my analysis I considered first the ideal situation where the CIB and CMB maps used for cross-correlation are noiseless and cover the whole sky. With these assumptions we obtain the highest possible signal-to-noise ratio, the only limitation being the cosmic variance. In Fig. 2.13 I present the predictions for the CIB-CMB cross-correlation in the case of a full-sky CIB map, provided by the previously mentioned instruments and frequencies.

With these optimistic assumptions, I obtained high levels of detection for the CIB-CMB correlation which reach $\simeq 7\sigma$, on par with (if not better than) the most promising surveys of the generation to come (for detailed S/N results, see Table 2.1 in Sect. 2.3.4). Interestingly, it should be mentioned that these results in the ideal case are independent of the previously discussed linear bias, even if it boosts the correlation signal. This can be understood from Eq. (2.23) where the linear bias can be factorized from each term (one for C_ℓ^{cr} and a squared one for C_ℓ^{CIB}) and therefore cancels out.

As evoked before, we see that the largest contribution to the S/N comes from multipoles smaller than $\simeq 50$. On the other hand, the most interesting feature of these
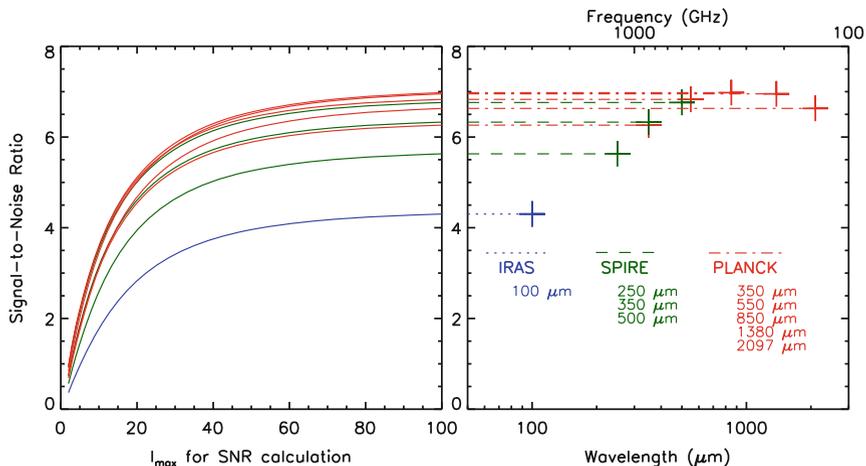
Fig. 2.13 *Left-hand panel* cumulated S/N as a function of ℓ_{max} (defined in Eq. 2.23) for the CMB-CIB cross-correlation, at our chosen frequencies and instruments. *Right-hand panel* total S/N with $\ell_{max} = 100$ as a function of frequency/wavelength. *Dotted line* is for IRAS, the *dashed* ones are for SPIRE, and the *dot-dashed* ones for *Planck*

Table 2.1 Total signal-to-noise ratio of the CIB-CMB cross-correlation for four of the CIB frequencies of *Planck*-HFI

CIB frequency (GHz)	857	545	353	217
CIB wavelength (μm)	350	550	850	1380
Ideal case, single correlation S/N	6.26	6.83	6.98	6.95
Joint S/N			7.12	
Realistic case n°1, single correlation S/N				
($f_{sky} = 0.75$, $f_{CMB} = 0.01$,	5.36	5.73	5.39	3.56
$\mathcal{A}_{fore.} = 0.01$)				
Joint S/N			5.88	
Realistic case n°2, single correlation S/N				
($f_{sky} = 0.15$, $f_{CMB} = 0.01$,	2.40	2.56	2.41	1.59
$\mathcal{A}_{fore.} = 0.01$)				
Joint S/N			2.63	

The results are given for each frequency and for the joint cross-correlation, first for the ideal case discussed in Sect. 2.3.3.2 and then for two more realistic cases

results is that contrary to what could be intuited from Fig. 2.11, the total S/N peaks around 850 μm instead of 250 μm for the cross-correlation signal itself. The reason for this is actually quite subtle: it comes from the shape of the "noise" term in the S/N expression in Eq. (2.23), as a function of ℓ, namely:

$$[N_\ell]^2(\nu) \equiv ([C_\ell^{cr}(\nu)]^2 + C_\ell^{CIB}(\nu)C_\ell^{CMB})/(2\ell + 1). \tag{2.24}$$
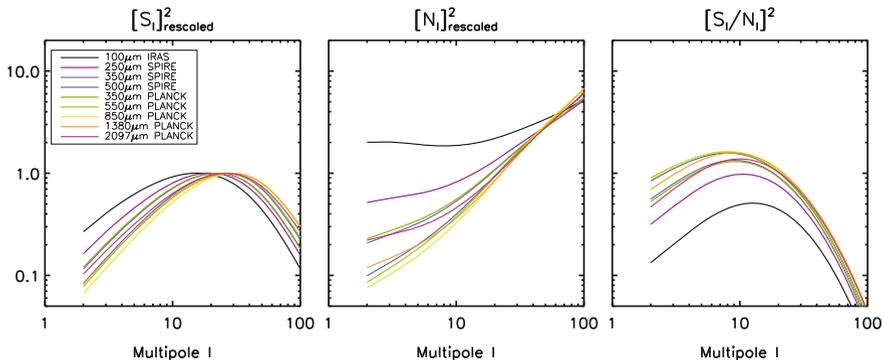
Fig. 2.14 "Signal" terms (*left-hand panel*, rescaled to unity) and "noise" terms (*middle panel*, same rescaling factor as the "signal") of the S/N as functions of ℓ (see text for details) for our chosen frequencies and instruments. The quotient of the two terms, used in the calculation of the S/N itself, is shown in the *right-hand panel* the main difference throughout the frequencies comes from the shape of the "noise" term

For all the frequencies studied here, this "noise" has roughly the same amplitude *relatively* to its corresponding "signal":

$$[S_\ell]^2(\nu) \equiv [C_\ell^{\mathrm{cr}}(\nu)]^2. \qquad (2.25)$$

This is illustrated in Fig. 2.14, where I plotted in the left panel all the $[S_\ell(\nu)]^2$ terms with their respective maximum rescaled to unity. In the middle panel, I applied the same rescaling factor of each $[S_\ell(\nu)]^2$ term to the corresponding $[N_\ell(\nu)]^2$ term. By doing this, it is possible to compare the results from all frequencies without changing their associated signal-to-noise ratios. On the resulting graph, we see that at $\ell = 100$ the rescaled noise amplitude is roughly the same, while the signal has the same shape at all frequencies, except for a small shift in ℓ. However there is a major difference in the shape of the noise power spectrum from one frequency to another: its slope changes depending on the frequency, with the steepest one for *Planck* 850 μm. Therefore its amplitude goes down more quickly than the others as ℓ approaches zero where coincidently the signal is strong, which then boosts the S/N at the low multipoles, and the total S/N.

In light of these results, the optimal frequency for iSW detection appears to be around 353 GHz/850 μm with a maximum S/N reaching 7σ. However in practice, the CIB extraction at this frequency might prove challenging since the CMB becomes dominant here, and increasingly so as we go down in frequency. Therefore the possible residuals in our extracted CIB map have to be accounted for, and other sources of noise as well, which is the purpose of the next subsection.

2.3.4 Accounting for Realistic Conditions

The contaminants of the CIB and obstacles to its extraction are many: first the signal is completely dominated on a large part of the sky by emissions from our own galaxy. The contamination from this foreground in the galactic plane is several orders of magnitude above the CIB level and prevents us from extracting the CIB, therefore reducing the "usable" fraction of the sky by at least ∼25 %. Furthermore, the rest of the sky is also quite polluted—from a CIB point-of-view—by these foregrounds full of galactic dust. These will have to be removed from our maps although some residuals might remain in the final CIB map used for the cross-correlation. There may even be a significant CMB residual (especially at *Planck* frequencies) in this map due to an imperfect separation of components, which could have a dramatic impact on the cross-correlation and easily induce false detections.

In the light of these elements, it appeared clearly to me that I needed to carry a more realistic study by including these possible sources of contamination and assess their impact on the iSW detection. To account for these effects on the detectability of the CIB-CMB cross-correlation, I used a more complete formulation of the signal-to-noise ratio, by adding new elements to the noise term. The expression of the S/N therefore becomes at a given frequency ν:

$$\left[\frac{S}{N}\right]^2(\nu) = f_{\mathrm{sky}} \sum_{\ell=2}^{\ell_{\max}} (2\ell + 1)$$

$$\times \frac{[C_\ell^{\mathrm{cr}}(\nu)]^2}{[C_\ell^{\mathrm{cr}}(\nu) + N_\ell^{\mathrm{cr}}(\nu)]^2 + [C_\ell^{\mathrm{CIB}}(\nu) + N_\ell^{\mathrm{CIB}}(\nu)][C_\ell^{\mathrm{CMB}} + N_\ell^{\mathrm{CMB}}]} \quad (2.26)$$

where f_{sky} is the fraction of the sky common to the CMB and the CIB maps, and N_ℓ^{cr}, N_ℓ^{CIB} and N_ℓ^{CMB} are the noise contributions respectively in the cross, CIB and CMB signal. Since the CMB is expected to be only variance-limited at the multipoles of interest, I fixed here $N_\ell^{\mathrm{CMB}} = 0$. However we still have to take into account the CIB contamination.

To do so, I first break the CIB noise power spectrum into several independent parts:

$$N_\ell^{\mathrm{CIB}}(\nu) = R_\ell^{\mathrm{CMB}}(\nu) + R_\ell^{\mathrm{fore.}}(\nu) + N_\ell^{\mathrm{instr.}}(\nu) + N_\ell^{\mathrm{correl.}}(\nu) \quad (2.27)$$

where these four different terms represent, from left to right, the power spectra of the CMB residual, the galactic foreground residuals, the instrumental noise and finally the noise due to a correlation between residuals and the CIB (which appears when autocorrelating the final CIB map).

I quantify the CMB residual in the CIB map as a fraction f_{CMB} of the total CMB map, which affects both the cross-correlation and CIB noise; this approach assumes a "global" CMB contamination (on the whole sky) without any spatial dependence. This consequently defines the noise in the cross signal:

$$N_\ell^{\mathrm{cr}}(\nu) = f_{\mathrm{CMB}}(\nu) \times C_\ell^{\mathrm{CMB}} \tag{2.28}$$

and the following two contributions:

$$R_\ell^{\mathrm{CMB}}(\nu) = f_{\mathrm{CMB}}^2(\nu) \times C_\ell^{\mathrm{CMB}} \tag{2.29}$$

$$N_\ell^{\mathrm{correl.}}(\nu) = 2 f_{\mathrm{CMB}}(\nu) \times C_\ell^{\mathrm{cr}}(\nu) \tag{2.30}$$

In the literature, the power spectrum of foregrounds such as dust emissions are often modelled as (and often found to be close to) power laws: here in my analysis I define the spectrum of the foreground residuals with the following expression:

$$R_\ell^{\mathrm{fore.}}(\nu) = \mathcal{A}_{\mathrm{fore.}}(\nu) \times C_{\ell=10}^{\mathrm{CIB}}(\nu) \left(\frac{\ell}{10}\right)^\alpha , \tag{2.31}$$

This power law has an amplitude defined relatively to the real CIB signal through a chosen constant $\mathcal{A}_{\mathrm{fore.}}$, which defines the quantity:

$$\mathcal{A}_{\mathrm{fore.}}(\nu) = R_{\ell=10}^{\mathrm{fore.}}(\nu)/C_{\ell=10}^{\mathrm{CIB}}(\nu) , \tag{2.32}$$

i.e. the ratio between the foreground residuals and the CIB spectrum at the multipole $\ell = 10$, approximatively where the cross-signal is at its maximum. The slope of the spectrum α is fixed here for all frequencies; previous analysis of infrared maps [30, 45] found it to be $\simeq -3$ for foregrounds at high galactic latitudes. Finally, the instrumental noise power spectra $N_\ell^{\mathrm{instr.}}$ at each frequency are taken from the first ten months of *Planck* data in [35], and extrapolated to the thirty months, i.e. the end of the fourth *Planck* full-sky survey.

In this section I focused on four of the five previously described *Planck* HFI frequencies, from 217 to 857 GHz. I chose to discard the fifth 143 GHz as the CMB completely dominates the CIB signal there. I also put aside the IRAS frequency here because of its weaker significance (even in the ideal case), and the SPIRE frequencies since the instrument was not scheduled to ever cover very large regions of the sky (i.e. $f_{\mathrm{sky}} \ll 1$), dramatically decreasing its associated S/N—as it is proportional to the square root of f_{sky} in Eq. (2.26).

At this point, the framework that I devised for the S/N calculation has three input parameters at each of the four frequencies: f_{sky}, f_{CMB} and $\mathcal{A}_{\mathrm{fore.}}$, whose values can be chosen freely. The next step would have been to explore this 3D parameter space at each frequency and compute the corresponding S/N at each point. Considering the very large number of possible combinations of parameters, it would not have been practical to perform and display the complete results of such exploration. Therefore I made the decision of fixing f_{sky} to two values of interest:

- $f_{\mathrm{sky}} = 0.75$, which corresponds to an optimistic case where the only part of the sky discarded is the galactic plane; this is an optimistic scenario in the sense that there are other highly contaminated regions where from the component separation techniques might not be able to extract the CIB.
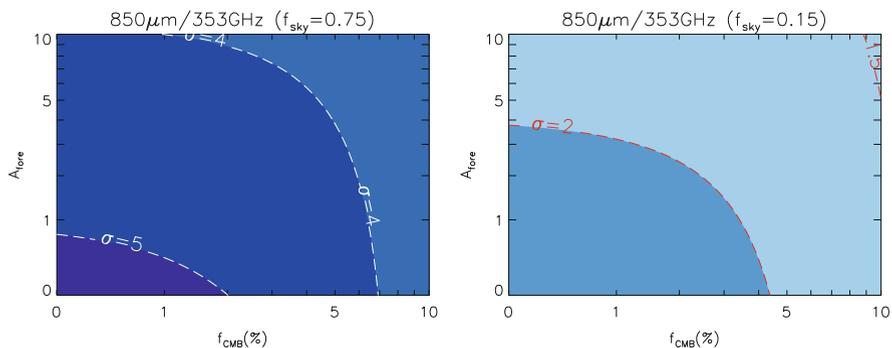
850μm/353GHz (f_sky=0.75)   850μm/353GHz (f_sky=0.15)

Fig. 2.15 Total signal-to-noise ratio of the CIB-CMB cross-correlation at 353 GHz, as a function of the CMB residuals (in percentage of the total CMB signal) and the foregrounds residuals (through the parameter $\mathcal{A}_{fore.}$). *Left panel* $f_{sky} = 0.75$, the results go from less than 4 to more than 5, from the *brightest colored area* to the *darkest. Right panel* $f_{sky} = 0.15$, the S/N goes from slightly less than 1.5 to more than 2, again from the *brightest* to the *darkest area*

- $f_{sky} = 0.15$, which is a low estimate of the area of the sky where the current data allow for an efficient CIB extraction. The methods currently employed are based on the use of HI maps as a tracer of the galactic dust, though it only remains valid for an HI column density lower than a specific threshold (see [34] for details on these methods).

As far as the two other input parameters are concerned, I limited their ranges to reasonable values, with $f_{CMB} \in [0, 0.1]$ and $\mathcal{A}_{fore.} \in [0, 10]$.

Among the four *Planck* frequencies that I kept here, I started by studying the effects of the noise for the frequency that gave the best S/N results in the ideal case (850 μm/353 GHz)—having in mind that a significant drop in S/N at this wavelength would not bode well for the rest of them. The corresponding results are presented in Fig. 2.15 which shows the contour levels of the S/N in the (f_{CMB}, $\mathcal{A}_{fore.}$) parameter space. The influence of the CMB is clearly visible at this frequency, quickly reducing the S/N as its residual level increases. This effect is even more pronounced at 1,380 μm/217 GHz, where the S/N is typically twice as low as in the ideal case (see Table 2.1), due to the fact that we get closer to the maximum of the SED of the CMB. It makes this frequency far less significant for the iSW detection than in the ideal case. The presence of instrumental noise—whose effect cannot be appreciated with Fig. 2.15 alone—becomes significant at the two lowest frequencies (217 and 353 GHz), again reducing their value in the cross-correlation. As expected the galactic foreground residuals also decrease the S/N, though their influence is roughly the same at all frequencies as they are defined relatively to the CIB spectrum in this analysis. Lastly, the biggest influence comes from the fraction of the sky through the f_{sky} parameter, as the total S/N scales as $\sqrt{f_{sky}}$. This makes it a crucial requirement for future applications to have the largest possible coverage to minimize this effect—very similarly to the requirement of galaxy surveys.
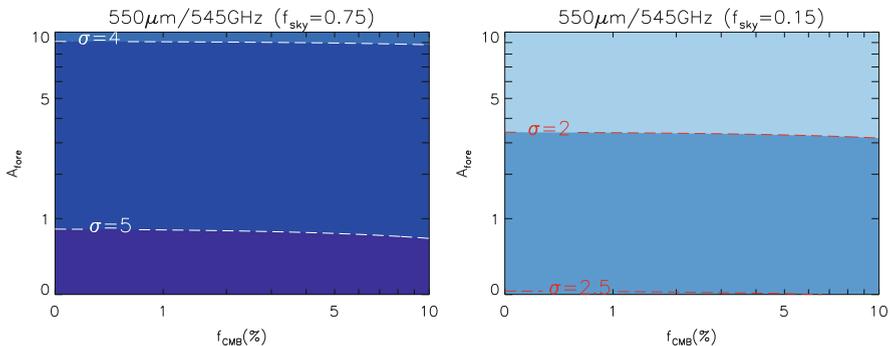
Fig. 2.16 Total signal-to-noise ratio of the CIB-CMB cross-correlation at 545 GHz as a function of the CMB residuals and the foregrounds residuals. *Left panel* $f_{sky} = 0.75$, the results go from slightly less than 4 to more than 5, from the *brightest* to the *darkest area*. *Right panel* $f_{sky} = 0.15$, the S/N goes from less than 2 to slightly more than 2.5

Taking all these remarks into account and after some exploration of the parameter space, the optimal frequency that stands out in these more realistic scenarii is 545 GHz/550 μm. Indeed, it is weakly influenced by instrumental noise and CMB residuals and also has a higher "original" S/N (in the ideal case) than the other remaining frequency 857 GHz/350 μm. The S/N analysis at 545 GHz is presented in Fig. 2.16: for the case of a large but realistic coverage, the S/N still reaches high and promising values around $4.5\,\sigma$. Even in a more pessimistic scenario, the significance of the detection stays around a $2.5\,\sigma$ level, comparable with the constraints from current galaxy surveys.

However, one may wonder here if it is really necessary to argue and determine which is the most suited frequency for giving the best results for the CMB-CIB cross-correlation. Indeed, in the same manner as I combined the constraints from redshift slices in Sect. 2.2.3, it is reasonable to consider the possibility of combining the cross-correlation signals from the CIB at several frequencies to improve the S/N—a approach that I will explore in the last section of this chapter.

2.3.5 Multi-frequency Joint Analysis

Until now I have only considered a detection at a single CIB frequency and its associated significance. In practice, when handling real data, we will most likely have several maps of the CIB at different frequencies, hence as many cross spectra. For example, in the case of *Planck* we should eventually be able to extract the CIB at four different frequencies on a large fraction of the sky. This could potentially allow to increase the total signal-to-noise ratio of the iSW detection by combining the constraints from all available frequencies. However, the improvement brought by this approach will be limited by the possible intrinsic correlations (and redundant

information) between the CIB maps at different frequencies—once again in the same manner as the correlations between redshift slices of a galaxy survey limited the amount of information available.

I expanded my previous S/N formalism to express the theoretical joint significance of a set of n cross-correlations (i.e. CIB at n frequencies, each correlated to the same CMB):

$$\left(\frac{S}{N}\right)^2_{\text{Total}} = X^T \mathcal{M}^{-1} X \tag{2.33}$$

with X (X^T) being the column (row) vector of all the cross-correlations:

$$X^T = \left(X^T(v_1) \ \ldots \ X^T(v_n)\right)$$

where $X^T(v_i)$ contains the cross-spectrum at the frequency v_i, from $\ell = 2$ to 100:

$$X^T(v_i) = \left(C^{\text{cr}}_{\ell=2}(v_i) \ \ldots \ C^{\text{cr}}_{\ell=100}(v_i)\right)$$

The block matrix $\mathcal{M}$ is the covariance matrix, containing $n \times n$ blocks. Each one of them represents the covariance of two cross-spectra at different CIB frequencies, depending on the position of the block. At the ith line and jth column, the block $\mathcal{M}^{ij}$ is written as:

$$\mathcal{M}^{ij} = \begin{pmatrix} \mathcal{M}^{ij}_{\ell=2} & & 0 \\ & \ddots & \\ 0 & & \mathcal{M}^{ij}_{\ell=100} \end{pmatrix}$$

The diagonality of $\mathcal{M}^{ij}$ comes from the assumption that the different multipoles are uncorrelated. This could prove no longer true for small fractions of the sky but gives an upper bound on the S/N. In the noiseless case discussed in Sect. 2.3.3, the elements of each block can be expressed as follows:

$$\mathcal{M}^{ij}_{\ell} = \text{Covar}(C^{\text{cr}}_{\ell}(v_i), C^{\text{cr}}_{\ell}(v_j))$$
$$= \frac{C^{\text{cr}}_{\ell}(v_i)C^{\text{cr}}_{\ell}(v_j) + C^{\text{CMB}}_{\ell}C^{\text{crCIB}}_{\ell}(v_i, v_j)}{2\ell + 1}$$

We can see here the dependence on the aforementioned possible correlation between the CIB at frequency v_i and the CIB at frequency v_j, through the cross-spectrum $C^{\text{crCIB}}_{\ell}(v_i, v_j)$. To perform a more advanced analysis, it is easy to modify this expression to account for the possible sources of noise discussed in the previous section.

Once again, the large number of possible combinations of noise parameters makes it unpractical to present a complete study of the joint correlation. Instead I focused on a few particular cases, motivated by my previous findings. A summary of my results on single and joint correlations is presented in Table 2.1. Going back first to the ideal case, I quantified the impact of the joint detection. I found a relatively small gain, as

it increases the total S/N by a mere $\simeq 0.15$ compared to the maximum significance of a single detection. This can be attributed to the high correlations between the CIB at its different observed frequencies, which limits the usefulness of the joint cross-correlation.

Considering now more realistic situations, with the presence of instrumental noise, I choose to fix some of the free parameters, with $f_{sky} = 0.75$ and $f_{sky} = 0.15$. A reasonable confidence in component separation techniques allows us to hope for small enough residuals, so that we choose $f_{CMB} = 0.01$ and $\mathcal{A}_{fore.} = 0.01$. In these cases, the joint correlation has a limited interest (respectively a $\simeq 0.15$ and $\simeq 0.07$ gain for $f_{sky} = 0.75$ and 0.15) due to the correlations in both the CIB signals but also in the astrophysical noise contributions—CMB and dust—between frequencies.

In the end, in the context of my model, the joint analysis of several CIB frequencies does not yield a significant improvement over the use of the best single frequency. This answers our interrogation from the end of the previous section: looking for the optimal frequency for the iSW detection is enough and justified here in order to focus the effort in the right direction (especially for the tricky part of the process: the CIB extraction). Nonetheless, the use of additional frequencies—if available—could have a non-negligible impact in the presence of a source of uncorrelated noise (between frequencies), as it would allow to get rid of most of it.

As a conclusion of this first ever investigation of the CIB-CMB cross-correlation, I found very promising results on the iSW effect and its detectability under various observational situations. Expected realistic significances range from ∼2.5 to 5.5 depending on the frequency, the levels of noise and the fraction of the sky available for analysis: these show a great potential compared to even the most promising galaxy surveys (cf. [2, 13]). The results of this work will be valuable in the forthcoming years of analysis and exploitation of the *Planck* data. The formalism I developed provides an accurate and flexible forecast of the expected results of the CIB-CMB cross-correlation and allows to constrain the requirements for a significant iSW detection.

References

1. J.K. Adelman-McCarthy et al., The sixth data release of the sloan digital sky survey. Astrophys. J. Suppl. Ser. **175**, 297–313 (2008)
2. N. Afshordi, Integrated Sachs-Wolfe effect in cross-correlation: the observer's manual. Phys. Rev. D **70**(8), 083536 (2004)
3. M.A. Agüeros et al., Candidate isolated neutron stars and other optically blank X-ray fields identified from the ROSAT all-sky and sloan digital sky surveys. Astron. J. **131**, 1740–1749 (2006)
4. H. Aihara et al., The eighth data release of the sloan digital sky survey: first data from SDSS-III. Astrophys. J. Suppl. Ser. **193**, 29 (2011)
5. C.L. Bennett, G.F. Smoot, A. Kogut, COBE DMR maps of the microwave sky. In Bull. Am. Astron. Soc. **22**, 336 (1990)
6. M. Béthermin, H. Dole, G. Lagache, D. Le Borgne, A. Penin, Modeling the evolution of infrared galaxies: a parametric backward evolution model. Astron. Astrophys. **529**, A4 (2011)

7. M. Béthermin et al., A unified empirical model for infrared galaxy counts based on the observed physical evolution of distant galaxies. Astrophys. J. **757**, L23 (2012)
8. M. Blanton, R. Cen, J.P. Ostriker, M.A. Strauss, The physical origin of scale-dependent bias in cosmological simulations. Astrophys. J. **522**, 590–603 (1999)
9. E. Boldt, The cosmic X-ray background. Phys. Rep. **146**, 215–257 (1987)
10. J.J. Condon et al., The NRAO VLA sky survey. Astron. J. **115**, 1693–1716 (1998)
11. A. Cooray, R. Sheth, Halo models of large scale structure. Phys. Rep. **372**, 1–129 (2002)
12. P.-S. Corasaniti, T. Giannantonio, A. Melchiorri, Constraining dark energy with cross-correlated CMB and large scale structure data. Phys. Rev. D **71**(12), 123521 (2005)
13. M. Douspis, P.G. Castro, C. Caprini, N. Aghanim, Optimising large galaxy surveys for ISW detection. Astron. Astrophys. **485**, 395–401 (2008)
14. F.-X. Dupé, A. Rassat, J.-L. Starck, M.J. Fadili, Measuring the integrated Sachs-Wolfe effect. Astron. Astrophys. **534**, A51 (2011)
15. D.J. Fixsen, The temperature of the cosmic microwave background. Astrophys. J. **707**, 916–920 (2009)
16. J. Garriga, L. Pogosian, T. Vachaspati, Forecasting cosmic doomsday from CMB-LSS cross-correlations. Phys. Rev. D **69**(6), 063511 (2004)
17. T. Giannantonio et al., Combined analysis of the integrated Sachs-Wolfe effect and cosmological implications. Phys. Rev. D **77**(12), 123520 (2008)
18. T. Giannantonio, R. Crittenden, R. Nichol, A.J. Ross, The significance of the integrated Sachs-Wolfe effect revisited. Mon. Not. R. Astron. Soc. **426**, 2581–2599 (2012)
19. K.M. Górski et al., HEALPix: a framework for high-resolution discretization and fast analysis of data distributed on the sphere. Astrophys. J. **622**, 759–771 (2005)
20. C. Hernández-Monteagudo et al., The SDSS-III baryonic oscillation spectroscopic survey: constraints on the integrated Sachs-Wolfe effect. Mon. Not. R. Astron. Soc. **438**, 1724–1740 (2014)
21. S. Ho, C. Hirata, N. Padmanabhan, U. Seljak, N. Bahcall, Correlation of CMB with large-scale structure. I. Integrated Sachs-Wolfe tomography and cosmological implications. Phys. Rev. D **78**(4), 043519 (2008)
22. W. Hu, R. Scranton, Measuring dark energy clustering with CMB-galaxy correlations. Phys. Rev. D **70**(12), 123002 (2004)
23. T.H. Jarrett et al., 2MASS extended source catalog: overview and algorithms. Astron. J. **119**, 2498–2531 (2000)
24. G. Jürgens, B.M. Schäfer, Integrated Sachs-Wolfe tomography with orthogonal polynomials. Mon. Not. R. Astron. Soc. **425**, 2589–2598 (2012)
25. L. Knox, A. Cooray, D. Eisenstein, Z. Haiman, Probing early structure formation with far-infrared background correlations. Astrophys. J. **550**, 7–20 (2001)
26. L. Lagache, J.L. Puget, Detection of the extra-galactic background fluctuations at 170 mu m. Aston. Astrophys. **355**, 17–22 (2000)
27. D. Larson et al., Seven-year Wilkinson microwave anisotropy probe (WMAP) observations: power spectra and WMAP-derived parameters. Astrophys. J. Suppl. Ser. **192**, 16 (2011)
28. D.N. Limber, The analysis of counts of the extragalactic nebulae in terms of a fluctuating density field II. Astrophys. J. **119**, 655 (1954)
29. H. Matsuhara et al., ISO deep far-infrared survey in the "Lockman Hole". II. Power spectrum analysis: evidence of a strong evolution in number counts. Astron. Astrophys. **361**, 407–414 (2000)
30. M.-A. Miville-Deschênes, G. Lagache, J.-L. Puget, Power spectrum of the cosmic infrared background at 60 and 100 μm with IRAS. Astron. Astrophys. **393**, 749–756 (2002)
31. J.A. Peacock, R.E. Smith, Halo occupation numbers and galaxy bias. Mon. Not. R. Astron. Soc. **318**, 1144–1156 (2000)
32. A. Pénin, O. Doré, G. Lagache, M. Béthermin, Modeling the evolution of infrared galaxies: clustering of galaxies in the cosmic infrared background. Astron. Astrophys. **537**, A137 (2012)
33. W.J. Percival et al., The shape of the sloan digital sky survey data release 5 galaxy power spectrum. Astrophys. J. **657**, 645–663 (2007)

34. Planck Collaboration, Planck early results. XVIII. The power spectrum of cosmic infrared background anisotropies. Astron. Astrophys. **536**, A18 (2011a)
35. Planck Collaboration, Planck early results. VI. The high frequency instrument data processing. Astron. Astrophys. **536**, A6 (2011b)
36. Planck Collaboration, Planck 2013 results. XIX. The integrated Sachs-Wolfe effect (2013a). ArXiv:1303.5079
37. Planck Collaboration, Planck 2013 results. I. Overview of products and scientific results (2013b). ArXiv:1303.5062
38. Planck Collaboration, Planck 2013 results. XVI. Cosmological, parameters (2013c). ArXiv:1303.5076
39. J.-L. Puget et al., Tentative detection of a cosmic far-infrared background with COBE. Astron. Astrophys. **308**, L5 (1996)
40. U. Sawangwit et al., Cross-correlating WMAP5 with 1.5 million LRGs: a new test for the ISW effect. Mon. Not. R. Astron. Soc. **402**, 2228–2244 (2010)
41. B.M. Schäfer, A.F. Kalovidouris, L. Heisenberg, Parameter estimation biases due to contributions from the Rees-Sciama effect to the integrated Sachs-Wolfe spectrum. Mon. Not. R. Astron. Soc. **416**, 1302–1310 (2011)
42. U. Seljak, M. Zaldarriaga, A line-of-sight integration approach to cosmic microwave background anisotropies. Astrophys. J. **469**, 437 (1996)
43. D.N. Spergel et al., First-year Wilkinson microwave anisotropy probe (WMAP) observations: determination of cosmological parameters. Astrophys. J. Suppl. Ser. **148**, 175–194 (2003)
44. N. Taburet, C. Hernández-Monteagudo, N. Aghanim, M. Douspis, R.A. Sunyaev, The ISW-tSZ cross-correlation: integrated Sachs-Wolfe extraction out of pure cosmic microwave background data. Mon. Not. R. Astron. Soc. **418**, 2207–2218 (2011)
45. E.L. Wright, Angular power spectra of the COBE DIRBE maps. Astrophys. J. **496**, 1 (1998)

The Large Scale Structures
A Window on the Dark Components of the Universe
Ilić, S.
2014, XXI, 140 p. 60 illus., 24 illus. in color., Hardcover