UNIVERSIDAD DE CANTABRIA
Departamento de Física Moderna
y
C.S.I.C.-UNIVERSIDAD DE CANTABRIA
Instituto de Física de Cantabria

Análisis de las anisotropías de la Radiación Cósmica del Fondo de Microondas mediante el uso de Filtros Adaptados

Memoria presentada por el Licenciado
Diego Herranz Muñoz
para optar al título de Doctor en Ciencias Físicas
2001

José Luis Sanz Estévez, Doctor en Ciencias Físicas y
Catedrático de la Universidad de Cantabria

CERTIFICA

que la presente memoria
"*Análisis de las anisotropías de la Radiación Cósmica del Fondo de Microondas mediante el uso de Filtros Adaptados*",
ha sido realizada por
Diego Herranz Muñoz
bajo mi dirección. Considero que esta memoria contiene
aportaciones suficientes para constituir la Tesis Doctoral del
interesado.

Santander, 20 de Diciembre de 2001

José Luis Sanz Estévez

A mis padres

Agradecimientos

Se suele decir en mi tierra que *"es de bien nacidos el ser agradecidos"*. He aquí, en desordenada lista, aquellos que en estos momentos me hacen sentir la persona mejor nacida del mundo:

A Amancia y Valentín, mis padres. Todo se lo debo a ellos. Ellos me han enseñado mucho más de lo que se puede aprender durante una tesis. Gracias por todo el apoyo que me han dado, pese a que sé que en el fondo ellos hubiesen preferido que su hijo se dedicara a una profesión respetable, como por ejemplo actor de culebrones o sexador de pollos; espero que el ver este trabajo terminado les haga sentirse un poco mejor en ese sentido. También quisiera dar las gracias al resto de mi familia, que siempre está allí aunque a veces le haga menos caso del que merece. A la memoria de mis abuelos y mi primo Antonio.

A mi director de tesis, José Luis Sanz, por darme la oportunidad de dedicarme a lo que siempre quise hacer y por todo el esfuerzo (y la paciencia) que ha puesto en que llegue este momento. ¡Espero que se me hayan pegado de él otras cosas aparte de la letra! También quisiera dar las gracias a Enrique Martínez, que aun estando continuamente ocupado en mil cosas siempre encuentra tiempo para dar un consejo, hacer una crítica constructiva o aportar una idea. Le agradezco a Laura Cayón los numerosos consejos y, no menos importante, los ánimos que siempre me ha dado. A Txitxo y Matilde, por su hospitalidad durante mis estancias en Berkeley, y a Tom Broadhurst, que hizo en gran parte posible dichas estancias. Quisiera agradecer a Wayne Hu su amabilidad al facilitarme la figura 1.1.

A mis compañeros y amigos del IFCA, especialmente Patricio y Belén, con quienes da gusto compartir no sólo despacho, sino conversación y todo lo que lleva implícita la amistad. Sin olvidarnos de todos los demás: Chema, Ignacio, Sebastián, Silvia, Teo, Aurora, Julio, Javi, Jose, Rafa, Alejandro, Marco... y, en general, a todos aquellos que quedarán inmortalizados en las páginas de *'el caso del aceite de ricino'* [1]. Por algo será.

Por supuesto, no puedo pasar sin enviar un cariñoso recuerdo a mis amigos de Segovia: Juan, Edu, Víctor, Iñaky, Ángel, Pedro, Jorge, Mario, Cristina y todos los demás. Volver al terruño es doble placer gracias a ellos. Y espero que siga siendo así por muchos años.

A todos los buenos amigos que he ido haciendo en mi vida itinerante, especialmente a Juanjo, a Juan Benéitez y a Saulo, que aunque les vea una vez al año siguen siendo los mejores amigos que se pueda tener. A mis compañeros de Físicas de la Complutense, especialmente a los del grupo de Astrofísica: ¡lo logré, chicos! A los colegas del 'trol' y a otros elementos dispersos e inclasificables que rondaron por el 'Jaime' en esos cinco memorables y ridículos años.

Por último, quisiera desmentir la injusta fama de Santander como ciudad fría e ingrata con los forasteros. Y que conste que no lo digo para hacerle la pelota a Patri ni nada de eso. Es cierto: la ciudad está poblada de días grises de lluvia y de un ejército de gaviotas tristes; pero, por suerte, con el tiempo he ido encontrando aquí gente que me ha hecho sentir tan a gusto como en mi casa. Gracias a todos ellos, y sobre todo a Alberto, por todo lo bueno que han aportado a mi vida en esta ciudad.

[1] http://www.ifca.unican.es/∼herranz/ricino.html

"En verdad, es más fácil que un camello pase por el ojo de una aguja que el que un científico atraviese una puerta. Y aunque la puerta sea más grande que el portón de un granero o de una iglesia, lo más prudente que puede hacer es consentir en comportarse como un hombre ordinario, y entrar sin más, en vez de esperar a que todas las dificultades implicadas en una entrada realmente científica queden resueltas."

Sir Arthur Eddington

"El que sabe no habla,
el que habla no sabe."

Lao Zi

Analysis of the anisotropies in the Cosmic Microwave Background Radiation using Adaptive Filters

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics.

by

Diego Herranz Muñoz

Departamento de Física Moderna
Universidad de Cantabria

&

Instituto de Física de Cantabria
Consejo Superior de Investigaciones Científicas-Universidad de Cantabria

2001

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The discovery of the cosmological background radiation was one of the milestones in observational Cosmology during the XXth century. Together with the expansion of the Universe, it constitutes the foundation stone of modern cosmological models and the reference point of all theories that try to describe the origin and evolution of the Universe as a whole. The study of cosmological background radiation has profound implications in areas as diverse as Cosmology, particle physics and Galactic and extragalactic Astronomy. The detection of the cosmic microwave background (CMB) itself in 1964 and the discovery of its anisotropies (together with the confirmation of the black body nature of its power spectrum) by the COBE satellite in 1992 were the landmarks of Cosmology during the second half of last century. Therefore, it is no surprising the extensive interest this topic has generated during the last 40 years. Last decade, in particular, has seen an overwhelming growth in the literature devoted to the CMB. At this very moment, several experiments are been carried on in order to detect and analyze the CMB radiation with unpreceded sensitivity and resolution. Yet more experiments are in preparation for the near future, that will provide full-sky pictures in the mm and sub-mm regime, both in intensity and polarization, leading to huge amounts of data to be analyzed.

In this context, the development and testing of analysis techniques adapted to the kind of data generated by CMB experiments is a pressing need. Some of the problems to be resolved in any CMB experiments are: the choice of efficient map-making and data representation algorithms, destriping, power spectrum and

higher moments estimation, denoising and the identification and separation of the different physical components that constitute the CMB radiation (*c*omponent separation). Of all these problems, this work will focus in the last two points.

In this work different filtering techniques aiming at the detection of signals embedded in CMB data are explored. These signals, such as extragalactic point sources and galaxy clusters, can be considered as foregrounds that contaminate the underlying CMB signal and, therefore, are needed to be removed from the data. But, unlike instrumental noise, the two considered examples have scientific relevance on themselves; what is a leftover from the point of view of CMB physics can be extremely useful from the point of view of extragalactic Astronomy. Hence, it is relevant to perform the separation as accurately and efficiently as possible. A new technique, *s*cale-adaptive filters, is presented in this thesis. The philosophy of scale-adaptive filters is to maximize the probability of detection of a signal with a known profile while paying maximum attention to its characteristic scale, in order to avoid false detections. The performance of scale-adaptive filters will be tested with worked examples and with one-dimensional and two-dimensional realistic simulations of the future ESA's Planck mission.

The layout of this thesis is as follows: in this first introductory section the basic ideas on CMB theory and observations are briefly summarized. There also a general description of filtering techniques is included. Chapter 2 introduces the idea of scale-adaptive filter. A full derivation of the scale-adaptive filter formalism can be found in Appendix A. As a first test, in Chapter 2 the scale-adaptive filter is applied to simulated one-dimensional Gaussian sources embedded in artificial noise and compared with the well-known Gaussian and Mexican Hat filters. Chapter 3 is devoted to the generalization of the scale-adaptive filters to 2D image processing and more general spatial profiles. In particular, the chapter is focused on a filter adapted to the multiquadric family of functions (that is a good approximation of the real profile of galaxy clusters as they are seen at X-ray and microwave wavelengths). Again, some simple simulations are presented.

Chapters 4 and 5 apply scale-adaptive filters to more sophisticated, realistic simulations. The simulations accurately reproduce the conditions of the incoming data produced by the future ESA's Planck mission. Planck has the highest sensitivity, resolution and sky coverage of all the experiments scheduled for a near future

and therefore is the ideal testing battlefield for an analysis technique. In Chapter 4 the scale-adaptive filter is applied to unidimensional arrays of data corresponding to a 6-month scan of one of the 30 GHz Planck detectors.

Chapter 5 presents a further generalization of filtering techniques in the case where information from different frequency channels is available and the frequency response of the signal to be detected is known ($m$ultifiltering). An immediate application of multifiltering is the detection of galaxy clusters through the Sunyaev-Zel'dovich effect. Using multifiltering, the clusters are detected in a robust fashion, directly from the data, assuming only a previous knowledge about the shape of the clusters.

Finally, Chapter 6 shows a brief review of the conclusions of this thesis and the prospective work and projects. A summary of this thesis in Spanish is included in Appendix B.

1.1 The Cosmic Microwave Background Radiation

The cosmic microwave background (CMB) is a weak, uniform radiation that permeates the Universe and arrives to the Earth from all directions. The existence of a cosmic background radiation was predicted in 1948 by Gamow and his collaborators when studying the synthesis of light elements in the primordial Universe. In the framework of the Big Bang, these elements formed when the Universe was much denser and hotter than in the present. Gamow predicted the existence of a relic radiation of about $5K$ (Gamow 1948a,b, Alpher & Herman 1948). The Princeton group made a similar prediction (Dicke *et al.* 1965) and were preparing a radiometer to detect this residual of the 'primeval fireball' just before its actual discovery by Penzias & Wilson in 1964 (Penzias & Wilson 1965). Penzias and Wilson, while working for the Bell Company in the construction of a new detector, found with their horn antenna an excess of noise with a temperature of $\sim 3K$ coming from all directions and without daily or seasonal variation. After discarding all possible sources of artificial contamination as origin of this noise, they considered its astronomical nature. Soon it was interpreted as the radiation predicted by Gamow and his collaborators.

Two facts are the most conspicuous about this radiation: in first place, its electromagnetic spectrum corresponds exactly with a black body at $T_o = 2.73K$, and in second place, its radiation shows an extraordinary homogeneity and isotropy. As far as we know, only processes produced by a system in thermal equilibrium can produce radiation following the black body law. The isotropy of the CMB radiation tell us about a process that involves all the observed Universe as a whole. This suggests a cosmological origin of the CMB; it could be easily explained if the Universe went through a phase of thermal equilibrium in the past. This is precisely what Big Bang models had predicted, and this is why the discovery of CMB was so crucial to the acceptance of these models.

CMB isotropy is not perfect. Very small fluctuations ('anisotropies') are present at different angular scales. As we will see, the statistical properties of the anisotropies are determined by the cosmological model. Therefore, CMB anisotropies are very useful. This is so mainly because they are small. For a given cosmological model, this allows to calculate the anisotropies using only linear perturbation theory. This means that if we were able to measure CMB anisotropies with a precision of few percent, we would be able to determine cosmological parameters with a precision of the same order.

The homogeneity of the CMB contrasts with the clumpy Universe in which we live in the present, full of non-linear structures such as galaxies, clusters of galaxies and voids. How the Universe has evolved from an almost featureless state to its present form is one of the most challenging questions we can formulate. In § 1.1.3 we will see how phenomena at intermediate redshifts can leave an imprint on CMB anisotropies and how such imprint can help to improve the knowledge about the evolutionary Universe.

1.1.1 The origin of the CMB

After the Big Bang the Universe was extremely hot and dense. Under the conditions of pressure and temperature that reigned in that stage, the matter was totally ionised and the baryons and the photons were tightly coupled. Frequent collisions between them ensured complete thermodynamic equilibrium of matter and radiation. As the young Universe expanded, the pressure and temperature

decreased. After $\sim 3 \times 10^5$ years (redshift $z \sim 1000$) the temperature had dropped to $T \sim 3000K$. Below this temperature, electrons and protons combined to form neutral atoms and the photons were free to propagate. The Universe became transparent. This is why this epoch is called either *recombination* or *decoupling* time. After the decoupling, matter and radiation follow different paths. Most of the photons never interacted again with the matter, preserving the black body spectrum of the primordial thermal equilibrium. The temperature of the CMB has dropped since decoupling due to the expansion of the Universe according to $T(z) = (1 + z)T_o$, being T_o the actual temperature of $2.73K$. Since before the recombination the Universe was opaque, the CMB can be considered as the first and oldest picture of the Universe.

For any observer, the origin of the CMB photons arriving at each time can be located on a spherical surface centered in the observer and with a radius corresponding to the travel time of the light emitted at the recombination time. This (fictitious) surface is called *last scattering surface*. Unless the Universe has experimented a strong reionisation at some redshift $0 < z < 1000$, most of the CMB photons we receive come from the last scattering surface. CMB anisotropies can be studied as projections over this spherical surface.

1.1.2 Observables

The basic observable of the CMB is its intensity as a function of frequency and direction on the sky. Since the CMB spectrum is an extremely good black body, its usual to describe it in terms of the temperature fluctuation $\frac{\Delta T}{T}(\vec{n}) = \frac{T(\vec{n})-T_o}{T_o}$, where $\vec{n}$ is a unit vector on the sphere. It is useful to describe the random field $\Delta T/T$ in spherical harmonics:

$$\frac{\Delta T}{T}(\vec{n}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\vec{n}) \tag{1.1}$$

where the $a_{\ell m}$ coefficients are independent random variables with $< a_{\ell m} >= 0$.

$$a_{\ell m} = \int d\vec{n}\, Y^*_{\ell m}(\vec{n}) \frac{\Delta T}{T}(\vec{n}) \tag{1.2}$$

On small sections of the sky -where the curvature can be neglected- the spherical harmonic analysis becomes ordinary Fourier analysis in two dimensions and ℓ be-

comes the Fourier wavenumber. Small ℓ's correspond to large separations in the sky, whereas $\ell \sim 10^2$ represents degree scale separations.

If the temperature fluctuations are statistically isotropic, the variance of the $a_{\ell m}$ coefficients does not depend of m:

$$< a_{\ell m} a^*_{\ell' m'} >= C_\ell \delta_{\ell \ell'} \delta_{m m'} \,, \tag{1.3}$$

where the averages are to be taken over statistical ensembles. The quantities C_ℓ are the *angular power spectrum* of the temperature field.

Angular power spectrum of the CMB

If the fluctuations are Gaussian, then the angular power spectrum C_ℓ fully characterizes the temperature field. The power spectrum is usually displayed as the power per logarithmic interval in wavenumber:

$$\Delta^2_T = \frac{\ell(\ell+1)}{2\pi} C_\ell T^2 \,. \tag{1.4}$$

The C_ℓ's can be accurately calculated for the inflationary models as a function of the cosmological parameters (e.g. Seljak & Zaldarriaga 1996, Hu *et al.* 1998). If the fluctuations are non-Gaussian, higher order moments must be studied. However, the angular power spectrum is still a fundamental test of the viability of those theories.

The behaviour of the angular power spectrum of the CMB as a function of ℓ can be splitted into three major regions:

- The large angular scales (low ℓ) correspond to physical sizes on the last scattering surface larger than the horizon at the epoch of decoupling. Since no physical process could affect these regions as a whole due to causality, no particular structure is expected in this region of the power spectrum. This is reflected in the angular power spectrum as a *plateau* at low ℓ.

- Intermediate regions ($\ell \simeq 30$ to $\ell \simeq 1000$) correspond to regions with physical size less than the horizon. Matter and radiation inside these regions could be affected by causal processes. Before the decoupling time, acoustic oscillations occurred in the baryon-photon fluid, giving rise to *acoustic (Doppler) peaks*

in the case of inflationary-line Universe models where the oscillations are in phase. The position of these peaks is mainly determined by the geometry of the Universe (Kamionkowski *et al.* 1994), the first one located around $\ell = 200$ (corresponding to the size of the horizon at the epoch of decoupling). In the case of some non inflationary models (such as topological defects), the oscillations are not in phase and a large bump is expected, but no multiple peaks.

- At small scales (arcminute scales and below), temperature fluctuations are damped due to the fact that decoupling is not instantaneous, that is, the last scattering surface has a finite thickness $\Delta z \sim 100$ (Jones & Wyse 1985). Thus, fluctuations with smaller scales than the thickness of the last scattering surface will be reduced by averaging over photons coming from different parts of the surface. Other mechanism responsible for the fall-off of the fluctuations is the so-called 'Silk damping' (Silk 1968): photons diffuse out from overdense regions, dragging the baryons with them and consequently smoothing the density fluctuations of both photons and baryons.

A real observer is limited to one Universe and one sky with its one set of $a_{\ell m}$'s, that is, $2\ell+1$ numbers for each ℓ. For Gaussian fluctuations each C_ℓ is drawn from a χ^2 distribution with $2\ell+1$ degrees of freedom. This leads to an inevitable variance of $\frac{2}{2\ell+1}C_\ell^2$ that affects mainly at low ℓ's. This fundamental limitation is called *cosmic variance*. Another effect that reduces our ability to accurately measure the angular power spectrum is the *sample variance*, which is due to partial coverage of the sky. If the fraction of sky covered is f_{sky}, then the errors increase by a factor of $f_{sky}^{-1/2}$. Even full-sky coverage experiments can be affected by sample variance, since highly contaminated parts of the data (i.e. the Galactic plane) may be needed to discard. Cosmic and sample variance are present independently of the resolution and sensitivity of the experiment.

The shape of the angular power spectrum for the working cosmological model with the present experimental results are shown in Figure 1.1. The plateau at low ℓ's and the acoustic peaks can be clearly seen. The low ℓ region was determined by the early measurements of COBE (Smoot *et al.* 1992). Only very recently the region of the acoustic peaks has become accessible with experiments such as the

balloon-borne BOOMERanG (Mauskopf *et al.* 2000, Netterfiled *et al.* 2001) and MAXIMA (Lee *et al.* 2001) and the ground based interferometer DASI (Halverson *et al.* 2001). The results of those experiments clearly show the multiple acoustic peak structure and strongly disfavour topological defects as seeds for the structure formation in the Universe. The favoured model is dominated by dark energy (around 70%) and a 30% of matter containing mainly non-baryonic dark matter and an amount of baryons in agreement with Big-Bang nucleosynthesis predictions and light element abundances measurements. The results agree very well with another estimations of the cosmological parameters obtained with completely different methods such as type Ia supernovae, light element abundances, etc. A complete reference list with all the CMB experiments shown in Figure 1.1 can be found in the recent review by Hu & Dodelson (2001).

1.1.3 Temperature anisotropies

The angular size of the anisotropies depend on their size in the last scattering surface and on the geometry of the Universe along the path of the photons to the observer. Hence, mapping the CMB anisotropies is a powerful cosmological test. CMB anisotropies can be classified as *primary*, if they are origined until the decoupling time, or as *secondary*, if they are due to processes in the way of the photons from the last scattering surface to us. Besides, emission from several kind of sources (*foregrounds*) is superimposed to cosmic microwave radiation.

Only a very brief summary on the principal facts about CMB anisotropies will be included in the following paragraphs. For a more extensive description, see the recent reviews of Durrer (2001) and Hu & Dodelson (2001).

Primary anisotropies

The ultimate origin of the CMB anisotropies can be found in the initial density fluctuations of matter and radiation fluid at the decoupling time. It is easy to understand that the fluctuations in the photon density at the decoupling time $\delta_{\gamma d}$ must be strongly correlated with with the fluctuation in the intensity of the CMB. Those fluctuations serve also as seeds for the structure formation in the Universe. Two different cases of initial density fluctuations are distinguished: *adiabatic* and

isocurvature. Adiabatic fluctuations are characterized by a null fluctuation of specific entropy associated to each component at each point:

$$\delta(\frac{n_b}{n_\gamma}) = \delta(\frac{n_X}{n_\gamma}) = 0\,, \tag{1.5}$$

where the subscripts γ, b and X denote photons, baryons and non-baryonic dark matter, respectively. Equation (1.5) implies the following relation at the initial time:

$$\delta_\gamma = \frac{4}{3}\delta_b = \frac{4}{3}\delta_X \tag{1.6}$$

Isocurvature fluctuations are characterized by a null fluctuation of total energy at each point, i.e., $\delta(\rho_\gamma + \rho_b + \rho_X) = 0$, what keeps constant the space curvature. In addition it is usually assumed that the entropy per baryon remains constant $\delta(n_b/n_\gamma) = 0$, when there exist a non-baryonic dark matter component. This leads to the following relation at the initial time:

$$\delta_\gamma = -\frac{4\rho_X}{3\rho_b + 4\rho_\gamma}\delta_X \tag{1.7}$$

There are two main scenarios that try to explain the origin of these initial seeds: inflation and topological defects. Inflationary models (Guth 1982, Linde 1982, 1983, Albrecht & Steindhart 1982) consider that the fluctuations were originated from quantum fluctuations that were boosted is scale during the exponential expansion that characterizes the inflationary epoch (Hawking 1982, Guth & Pi 1982, Starobinskii 1982, Bardeen *et al.* 1983). Inflationary models favour adiabatic fluctuations (e.g. Kolb & Turner 1990) drawn from a realization of a homogeneous and isotropic Gaussian random field. If the fluctuations are Gaussian, they are fully characterized by their power spectrum $P(k)$. Inflation favour power spectra of the type

$$P(k) = Ak^n\,, \tag{1.8}$$

where A is the power spectrum normalization and n is the spectral index. Standard inflation predicts $n = 1$ (Harrison-Zel'dovich spectrum, Harrison 1970, Zel'dovich 1972). In non-standard inflationary models isocurvature fluctuations (Efstathiou & Bond 1986, Peebles 1999a,b) and deviations from Gaussianity (Salopek 1992, Peebles 1999a,b) are also possible. Topological defects scenarios consider cosmic

strings, textures and monopoles that might be formed during symmetry breaking phase transitions in the early Universe as seeds to initial perturbations (for a review see Vilenkin & Shellard 1994). CMB fluctuations induced by topological defects are non-Gaussian. Hybrid scenarios that combine inflation and topological defects have also been proposed (Jeannerot 1996, Linde *et al.* 1997, Avelino *et al.* 1998).

Photon density fluctuations are not the only source of primary anisotropies. Other contributions arise from the gravitational redshift of photons climbing out of potential wells in their way from the last scattering surface to us (Sachs-Wolfe effect, Sachs & Wolfe 1967) and the Doppler effect due to the peculiar velocities of the last scatters of the photons. Then, if recombination occurs in the matter-dominated era and $\Omega \gtrsim 0.1$,

$$\frac{\Delta T}{T}(\vec{n}) \approx \frac{1}{4}\delta_{\gamma d} + \frac{1}{3}\phi_d - \vec{n}\vec{v}_d \tag{1.9}$$

where $\vec{n}$ is the direction given by the line of sight, the subscript d indicates quantities at the decoupling time and the units have been chosen so that $c = 8\pi G \equiv 1$. The first term in eq. (1.9) corresponds to density fluctuations. The second term corresponds to the Sachs-Wolfe effect and dominates at scales larger than the horizon size at decoupling, where the CMB anisotropies are directly related to the initial power spectrum of matter density fluctuations. The third term represents the Doppler effect. The three terms together cause the behaviour of the angular power spectrum described in § 1.1.2.

Secondary anisotropies

The secondary anisotropies provide information about the evolution of the Universe after decoupling. Two kind of phenomena are mainly responsible for these anisotropies: gravitational and scattering effects.

Gravity affects photons in two ways:

- *Gravitational shifts.* If the photons cross a region with a potential well that is evolving they experiment a variation in their energy due to the variations in gravitational redshift/blueshift along the path of the photon (Martínez-González *et al.* 1990). The integrated effect is

$$\frac{\Delta T}{T}(\vec{n}) = \int \frac{\partial \phi}{\partial t} dt\,. \tag{1.10}$$

In the special case where the potential well is due to linear density perturbations in a Universe different from the Einstein-de Sitter one, this effect is known as the Integrated Sachs-Wolfe effect (ISW). In an Einstein-de Sitter Universe the ISW is zero. Four different regimes can be distinguished for this mechanism: *early ISW effect* when the potential decays between the decoupling time and the full matter domination, *late ISW effect* when the potential decay due to the natural Universe expansion once matter no longer dominates the expansion (Λ or open models), *Rees-Sciama effect* which is due to non-linear evolution of collapsing structures and finally all the other gravitational redshift effects, including gravitational waves and topological defects.

- *Gravitational Lensing.* Gravitational shifts described above change the energy of the photons but not their trajectories. Gravitational lensing is the deviation of the light due to a gravitational well without changing the energy of the photons. Matter density fluctuations along the path of the photons smear the angular power spectrum on scales typically small (Martínez-González *et al.* 1997). Although this effect is expected to be small, it could be detectable for future CMB experiments and could be used to trace the large scale structure of matter.

Scattering. If the Universe suffer a reionisation after decoupling (due to quasars, massive formation of first generation stars, etc.), then the CMB anisotropies will be dramatically influenced by scattering with electrons. Scattering can change the direction of the photons (Thomson scattering) and their energy (Compton scattering). In the first case, a smearing of the light will occur affecting scales smaller than the horizon size at the redshift of the scattering. If the Universe went through a global reionisation at a given redshift z_r, then a dumping in the fluctuations would happen due to photon path mixing. Local ionised regions can also affect the CMB. For example, inverse Compton scattering of microwave photons by hot electrons in the intracluster gas of a cluster of galaxies produces spectral distortions of the black body spectrum of the CMB, known as the *thermal Sunyaev-Zel'dovich effect*. The Sunyaev-Zel'dovich (SZ) effect will be further discussed in § 1.1.4.

1.1.4 Foregrounds

Besides the cosmological radiation described in the previous sections, other objects and physical processes produce emission in the microwave region of the electromagnetic spectrum. The set of astronomical microwave sources other than the CMB are known as *foregrounds*. In any direction of the sky, a detector will measure the photons coming from the last scattering surface mixed with photons coming from foregrounds situated along the line of sight. A good understanding of the different foregrounds is necessary for a suitable analysis of the CMB. Moreover, the knowledge of the foregrounds is a major goal on itself because they contain themselves very valuable information on astronomical phenomena. There are two main contributions to the foregrounds: Galactic and extragalactic. Galactic foregrounds are typically extense and grow stronger towards the Galactic Plane; they can dominate at low ℓ's and be a painful problem to estimate the first C_ℓ's. Extragalactic foregrounds are local contaminants, meaning that they are centered in the positions of galaxies and clusters of galaxies surrounding our own and, due to the relatively small angular scale of such objects as seen from Earth, the emission from these sources is located in small areas of the sky. This work will focus mainly on the detection of extragalactic foregrounds.

The main Galactic foregrounds are:

Synchrotron

Synchrotron emission is produced by relativistic electrons that are accelerated in magnetic fields (for a recent review see Smoot 1999). Hence, it depends on the energy spectrum of the electrons and the strengths of the magnetic field. Supernova remnants are strong synchrotron sources. There also exists a diffuse Galactic synchrotron emission associated to the Galactic magnetic field and electrons freed by older supernovae. Synchrotron emission is highly polarized and dominates in the low frequency range (< 30 GHz). There is a complete sky survey at 408 MHz (Haslam *et al.* 1982) and another half sky survey at 1024 MHz (Reich & Reich 1986) which have been used to extrapolate synchrotron emission to higher frequencies relevant to CMB observations.

Dust emission

Dust grains in our Galaxy are heated by the interstellar radiation field, absorbing UV photons and re-emitting the energy in the far IR. In general the equilibrium temperature between absorption and emission is $T \approx 20K$ in the limit of large dust grains. The spectrum of this light can be well approximated by a modified black body spectrum $I_\nu \propto B_\nu(T_\alpha)\nu^\alpha$ where $\alpha \sim 2$ (Draine & Lee 1984). This emission peaks at ~ 140 μm but extends out to microwave wavelengths. Dust dominates over the other components at high frequencies ($\nu > 300$ GHz). Wright (1991) and Reach *et al.* (1995) found that two components were necessary to fit the Galactic dust emission, including a cold component with $T \sim 7K$. If this cold component exists, it could dominate the dust emission at low frequencies. The Berkeley-Durham dust map (Schlegel *et al.* 1998) is based on infrared radiation from dust, observed by IRAS at 100 μm and DIRBE at 100 μm and 240 μm. This map can be used to clean up future CMB data. There could be another component of dust emission due to spinning dust grains that would have a peak in the low frequency channels of Planck. In Oliveira-Costa *et al.* (1999) they have shown evidence for this hypothesis. They have cross-correlated Tenerife data with Galactic templates and they found a turnover in the spectrum at 10 and 15 GHz which supports the spinning dust hypothesis. However, its existence is still unclear.

Free-free emission

Free electrons, when accelerated by ions in the interstellar gas, radiate due to thermal bremsstrahlung. This emission is located in hot regions ($T \gtrsim 10^4K$) where the hydrogen is ionised. Therefore, H_α regions are thought to be good tracers of free-free emission. Free-free is the less well known Galactic foreground due to the fact that it only dominates over a small range of frequencies ($25 - 75$ GHz), where the total Galactic emission is minimal. However, the correlation of free-free emission with H_α can be used to create free-free template maps. The Wisconsin H_α Mapper (WHAM) survey (Reynolds *et al.* 1998) will produce a map in H_α of the northern sky with 1° resolution. This survey can be combined with the southern celestial hemisphere H_α survey (McCullogh *et al.* 1999). Free-free emission is partially correlated with dust as the H_α emission has long been

known to be. A possible source of confusion exists between dust-correlated free-free emission and spinning dust (see above). Future CMB experiments should be able to solve this uncertainty.

Figure 1.2 shows the frequency behaviour of the strength of the different Galactic foregrounds. Dust is the dominant contribution at high frequencies whereas synchrotron dominates the low frequency regime. The CMB and its dipole and quadrupole are also shown.

Extragalactic foregrounds are of very diverse nature. Almost all of them are unresolved at the resolutions of present and near future CMB experiments. All these foregrounds receive the generic name of *extragalactic point sources*. Resolved extragalactic sources include nearby galaxies (such as the Magellanic Clouds) and some galaxy clusters. Even though the Sunyaev-Zel'dovich imprint in the CMB should be considered as a secondary anisotropy (see § 1.1.3), it will be discussed in this section due to the similarity between galaxy clusters and point sources regarding their spatial profile.

Extragalactic point sources

There are different source populations that can produce emission in the microwave domain. The 'zoo' of extragalactic objects has many species; hence, the systematic study of extragalactic point sources has many uncertainties. Toffolatti *et al.* (1998) and Guiderdoni (1999) have performed a detailed study of these extragalactic point sources at the frequencies of the Planck mission. From these works, fluctuations from point sources are well below the expected amplitude of CMB fluctuations in the frequency range $50 - 200$ GHz on all angular scales covered by the Planck mission. The number counts of extragalactic point sources are dominated at high frequencies by IR selected sources (such as starburst and late type galaxies at intermediate to low redshift and high redshift ellipticals). On the other hand, radio selected sources (flat-spectrum radiogalaxies, radio-loud quasars, BL Lacs, etc., mostly at substantial redshift) are expected to dominate the counts in the Planck low frequency channels. At intermediate frequencies (353, 217 and 143 GHz) the channels are almost free of point source counts (being the infrared population the dominating among these few counts). Since the sources are, as a first approach, randomly distributed in the sky, their angular power spectrum can be approximated by white noise, i.e., C_ℓ = constant for all scales. Thus, confusion from point sources mainly will affect small angular scales. The contributions to fluctuations owing to clustering of both radio and far-IR sources is found to be generally small in comparison with the Poisson term; however, the relative importance of the clustering contribution increases and may eventually become dominant if sources are identified and substracted down to low flux limits (Toffolatti *et al.* 1998).

Detection and substraction of bright point sources is essential before any other

analysis of the CMB is performed. Vielva *et al.* (2001a) have shown that Planck will detect thousands of these point sources. Due to the uncertainties in the knowledge about the different source populations as well as their variability, it is extremely difficult to perform detection techniques based in the frequency dependence of the angular power spectrum of the sources. Fortunately, all the point sources have a characteristic in common: their size. As we will see through this thesis, filter and wavelet techniques are specially well suited to detect point sources (Cayón *et al.* 2000, Vielva *et al.* 2001a, 2001b, Sanz *et al.* 2001, Herranz *et al.* 2001a, 2001b).

The Sunyaev-Zel'dovich effect

As mentioned above, the Sunyaev-Zel'dovich effect (SZE) is a secondary anisotropy source and therefore should be treated in § 1.1.3. In spite of this fact, it is included here for two reasons: first, because its origin can be related to actual astronomical objects, as is the case of extragalactic point sources, and second because galaxy clusters, except for the nearest ones, share with point sources their small angular sizes, close or clearly below the resolution of CMB experiments. Hence, the same detection techniques that apply to point sources can be used to detect the SZE of galaxy clusters, albeit in this last case the frequency dependence can be used to improve the results.

The physics of the SZE is simple (for a recent review see Birkinshaw 1999). Gas in hydrostatic equilibrium within a cluster's gravitational potential well corresponding to typical cluster masses must have an electron temperature T_e of several keV. At this temperature, the gas is ionised an thermal emission from the gas appears in the X-ray part of the spectrum. Electrons in the intracluster gas can scatter photons of the CMB. If the electrons have enough energy they can even transfer some energy to the photons. This is known as *inverse Compton scattering*. When a photon is scattered by an electron, its energy and momentum is altered. The change in energy and direction of the photon is given by the Compton formula:

$$\epsilon' = \frac{\epsilon}{1 + (1 - \cos\phi)\frac{\epsilon}{m_e c^2}} \tag{1.11}$$

in the rest frame of the electron before the interaction, and where ϵ and ϵ' are the photon energies before and after the scattering and ϕ is the deflection angle of the

photon. For low energy photons and mildly relativistic or non-relativistic electrons, $\epsilon \ll m_e c^2$ and the scattering is almost elastic ($\epsilon \simeq \epsilon'$). This limit is appropiate for the scatterings in cluster of galaxies that produce the SZE and causes a considerable simplification in the physics. When an ensemble of electrons produce scattering the final effect must be calculated by averaging over the electron distribution. An approximation usually done is to assume that the electron velocities follow a Maxwellian distribution.

When describing the SZE is useful to define the *Compton parameter*:

$$y_c = \frac{k_B T_e}{m_e c^2} \tau \tag{1.12}$$

where $\tau = \sigma_T \int n_e dl$ is the *optical depth* (the optical depth is the line integral along the line of sight of the electron density times the Thomson scattering cross-section σ_T). It can be found that the induced change in the CMB temperature is expressed as:

$$\frac{\Delta T(x)}{T_0} = g(x) y_c \tag{1.13}$$

where $x = h\nu / k_B T_{CMB}$ and $g(x) = (x \coth(x/2) - 4)$ is the *spectral shape factor*. Although the most common way to express the SZE is the corresponding intensity change which is given by:

$$\Delta I(x) = I_0 f(x) y_c \,, \; f(x) = \frac{x^4 e^x}{(e^x - 1)^2} g(x) \tag{1.14}$$

Note that eqs. (1.13) and (1.14) are independent of redshift, meaning that the temperature distortion produced by a galaxy cluster does not depend on the distance to the clusters.

The characteristic shape of the SZE frequency dependence is shown in Figure 1.3. At $x = 3.81$ ($\nu = 217$ GHz) the intensity change is zero. Below this frequency clusters will produce a decrement on the CMB intensity whereas above it clusters will enhance the CMB intensity. Typical temperature distortions due to the SZE are $\delta T/T \sim 10^{-4}$, one order bigger than the intrinsic CMB anisotropies. For very hot clusters ($T > 10$ keV) relativistic corrections should be taken into account.

The SZE discussed above is called *thermal SZE* due to the fact that the velocities of the electrons producing the scatterings are drawn from thermal agitation.

But if the bulk of the cluster gas is moving relative to the the Hubble flow with a peculiar velocity v_r along the line of sight, then a *kinematic* contribution to the SZE appears. This contribution is given by:

$$\left(\frac{\Delta T}{T}\right)_{kin} = -\frac{v_r}{c}\tau \tag{1.15}$$

with τ the optical depth. This contribution is independent both on the redshift of the cluster and the frequency; this fact is interesting because at $\nu = 217$ GHz the only contribution to the SZE comes from the kinematic SZE, providing this a method to measure the radial component of the peculiar velocities of the clusters. However, the detection of the kinematic SZE is a difficult task since typically it is about two orders of magnitude below the thermal SZE.

For a recent review on the current status of SZE measurements, see Rephaeli (2001). The SZE can be used as a powerful probe in Cosmology due to its lack of dependence of the redshift of the clusters. In that sense, SZE surveys can reach even the most distant clusters, allowing to detect the first clusters formed in the Universe. The knowledge of the number counts of clusters as a function of redshift can be used as a test to cosmological and structure formation models. Moreover, SZE data can be combined with X-ray data to determine the distance to clusters. Finally, kinematic SZE can be used to estimate the peculiar velocities of galaxy clusters. For all these reasons and for the sake of cleaning up CMB maps, the detection of SZE is one of the most promising fields in Astronomy.

1.1.5 The component separation problem

In the previous sections the main facts about the CMB and the microwave foregrounds have been quickly reviewed. From this discussion it is clear the big relevance of studying the microwave sky. Data collection, analysis and posterior interpretation are difficult tasks, especially when we consider the huge amounts of data that are necessary to store and handle in typical CMB experiments. The statistical pipeline which confronts the data with the theory include mapping, filtering, comparing, cleaning, compressing, forecasting, estimating and many other tasks, most of them out of the scope of this thesis. But before (or while) performing any analysis is necessary to know what is CMB and what is not; in other

words, we must separate the different components that contribute to the data. In that sense, noise can be considered as a independent component that we need to remove (*denoising*) while the other components (i.e., the CMB and the different foregrounds) must be separated and kept apart for their posterior analysis (*component separation*).

Component separation of CMB maps has been widely discussed over the last years. The different methods proposed include Wiener filtering (WF, Tegmark & Efstathiou 1996, Bouchet et al. 1997), maximum entropy method (MEM, Hobson et al. 1998, 1999), fast independent component analysis (FastICA, Maino et al. 2001), Mexican Hat Wavelet analysis (MHW, Cayón et al. 2000, Vielva el al. 2001a) and matched filter analysis (MF, Tegmark & Oliveira-Costa 1998). For a comparison of different methods applied to CMB, see Tegmark (1997) and Jones (1998). WF, MF and MHW techniques will be further discussed along this work.

Among the numerous data analysis techniques that have been applied to the study of CMB and in particular to the component separation problem, *filtering* is specially well-suited to deal with the detection and extraction of signals embedded in a noisy background. WF and MF are two different filtering techniques. The Mexican Hat Wavelet can be used as a wavelet (e.g. in multiresolution analysis) or as a linear filter; wavelets and filters are deeply related between them. The following section is devoted to introduce the concept of 'filter' and the basic formalism that will be used through this thesis.

1.2 Filters

1.2.1 What is a filter?

A classical, maybe somewhat engineer-oriented definition of filter is: "a filter is a device in the form of a piece of physical hardware or computer software that is applied to a set of noisy data in order extract information about a prescribed quantity of interest" (Haykin 1996). This reflects the intuitive idea of a filter: a tool that, provided a certain data input, gives an output that have some desirable properties. From the signal processing point of view, a filter is a kind of

system, that is, a process that results in the transformation of a signal. From the mathematical point of view, a filter is an *operator*:

$$L : f(t) \longrightarrow g(t) = Lf(t) \tag{1.16}$$

where f is the input, g is the output and t is the independent variable (for the sake of simplicity we will discuss only the one-dimensional case, though generalization to the n-dimensional case is straightforward), which let us call 'time' (it could also be space, temperature or any other quantity). The filter is *linear* if the filtered quantity g is a linear functional of the inputs. The filter is *time-invariant* when if the input is delayed by τ the output is also delayed by τ, that is, $g(t-\tau) = Lf(t-\tau)$. Most of the filters used in a wide range of disciplines are linear and time-invariant.

Time-invariance is a very desirable property of a filter. Let δ be the Dirac distribution. Then, if f is continuous, its value at t is obtained by the integral:

$$f(t) = \int_{-\infty}^{\infty} f(u)\delta(t-u)du \tag{1.17}$$

The continuity and linearity of L imply that

$$Lf(t) = \int_{-\infty}^{\infty} f(u)L\delta(t-u)du \tag{1.18}$$

Let h be the *impulse response* of L:

$$h(t) = L\delta(t) \tag{1.19}$$

Due to the time-invariance of the operator, $L\delta(t-u) = h(t-u)$ and hence

$$Lf(t) = \int_{-\infty}^{\infty} f(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)f(t-u)du = h \otimes f(t) \tag{1.20}$$

where the symbol $\otimes$ denotes convolution. *A time-invariant linear filter is thus equivalent to a convolution with the impulse response h.* The convolution property is quite useful when going to Fourier space. Let us choose the following convention for the Fourier transform:

$$\hat{f}(q) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{iqt}dt \tag{1.21}$$

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(q)e^{-iqt}dq \tag{1.22}$$

Making use of the convolution theorem (see Appendix A):

$$Lf(t) = g(t) = h \otimes f(t) = \int_{-\infty}^{\infty} \hat{h}(q)\hat{f}(q)e^{-iqt}dq \tag{1.23}$$

The Fourier transform $\hat{h}$ of the impulse response h is known as the *transfer function* of the filter.

This discussion is valid for continuous functions (signals). For discrete data it is easy to generalize the definitions above, obtaining equivalent results in terms of sums instead of integrals.

1.2.2 Why filtering?

In the last section we have seen that filtering with a linear time-invariant filter is equivalent to multiply the Fourier transform of the data with a transfer function that, at the end, modulates the spectral properties of the input. In other words, a linear time-invariant filter is a frequency-selective device. This suggests some immediate applications. The first of them is *denoising*: as the noise usually manifests in Fourier space in the high frequency domain, a filter that takes values close to zero in that region and close to unity in the low frequency domain (low-pass filter) will remove the noise without affecting the large scale features of the signal. A similar application of low-pass filters is the smoothing of data. On the contrary, a high-pass filter (one that is zero near the origin and close to unity for high frequencies) will erase all large scale features while preserving the rapidly variable features. Other possibility is to select a signal which is known to be in a certain frequency domain (band-pass filter) or to do exactly the opposite: to remove an interference located at a certain frequency ('notch' filter). Additionally, filters can be used for prediction, that is, to estimate the value of an underlying signal at a time (or position) where we have no direct measurement.

The filters described above allow to easily manipulate the spectral behaviour of the data, selecting some frequency domains and removing others. If a signal is know to lie in between certain frequency limits, it can be enhanced with respect to the other components of the data. Moreover, these filters are linear, which is often a desirable characteristic from the point of view of mathematical analysis. Finally, the convolution property provides a easy recipe to quickly filter large data

sets. For all these reasons, filters have become a powerful tool in signal processing and data analysis.

1.2.3 Filter design

Ideal low-pass, high-pass, band-pass and notch transfer functions can be easily constructed in Fourier space simply by setting the value of the transfer function equal to zero outside the desired frequencies. For example, an ideal low-pass filter has a transfer function:

$$\hat{h}_{lp}(q) = \begin{cases} 1 & \text{if } \mid q \mid < q_c\,, \\ 0 & \text{if } \mid q \mid \geq q_c\,, \end{cases} \tag{1.24}$$

where q_c is a cutoff frequency. The region $\mid q \mid < q_c$ is the *passband* of the filter, whereas the region $\mid q \mid \geq q_c$ is called *stopband*. The impulse response of this filter is

$$h_{lp}(t) = 2q_c \frac{\sin(q_c t)}{t} \tag{1.25}$$

This function has a central lobe surrounded by decreasing ripples. This kind of pattern is usually produced by strong discontinuities in the transfer function. That means that after filtering with the filter given by eq. (1.24) the data will show undesirable ring-like features. Besides, in many filtering contexts the signals to be separated do not lie in totally disjoint frequency bands. For example, the spectrum of point sources in CMB maps extends to frequencies dominated by white noise and other backgrounds. In order to avoid ringing effects and to improve the filtering of signals with overlapping spectra, a gradual transition from stopband to passband is generally preferable. Such filters are called *non-ideal filters*. In filtering theory, non-ideal is generally better than ideal!

Having assured the goodness of the filters with continuous transfer functions, the question is: 'what filter is the best fitted for a given problem?' Filter design is in many senses an art. For a given data set, there are as many filters as possible questions about the data we could formulate. There would be filters suitable for denoising, filters for selecting some signals according to certain signal properties, filters for smoothing, filters for predicting the future behaviour of the data, etc. Regarding component separation, that is, signal recovering, the general problem can be summarized in the following diagram:

$$\text{SIGNAL + NOISE} \longrightarrow \text{FILTER} \longrightarrow \text{SIGNAL + reduced noise}$$

In the optimistic case, the 'reduced noise' should tend to zero. *Optimal filtering* is the art of reducing the noise in the output as much as possible. All of this is clear from the qualitative point of view. Going from the qualitative to the quantitative requires to make use of statistics; in other words, before trying to design an optimal filter is necessary to define what is meant by 'optimal' using statistical terms. Different statistical criteria will lead to different optimal filters. In the next section we will discuss two different filters, Wiener filter and the 'matched' filter, that are optimal in different senses. In Chapter 2 a new criterion for optimal will be introduced that will lead to a new family of filters, called *scale-adaptive filters*.

1.2.4 Some examples of filters

To illustrate how the choice of a criterion or another can lead to very different filters that are optimal in a certain way let us discuss two very popular examples of filters: the Wiener filter and the matched filter.

Wiener filter

Let us consider the following situation: there is some underlying, uncorrupted signal $u(t)$ that we want to measure. The measurement process is imperfect, however, and what comes out is a corrupted signal $c(t)$. This corruption may have arisen from two different causes: first, the apparatus may have not a perfect Dirac-δ response, so that the signal is smeared out by some known response function $r(t)$ to give a smeared signal $s(t)$,

$$s(t) = \int_{-\infty}^{\infty} r(\tau)u(t-\tau)d\tau\,, \tag{1.26}$$

and second, the measured signal $c(t)$ may contain an additional component of noise $n(t)$, so that

$$c(t) = s(t) + n(t) = r \otimes u(t) + n(t). \tag{1.27}$$

Let $U(q)$, $R(q)$, $C(q)$ and $N(q)$ be the Fourier transforms of $u(t)$, $r(t)$, $c(t)$ and $n(t)$, respectively. To obtain an estimator $\tilde{u}(t)$ as close as possible to $u(t)$ (or, equivalently, a $\tilde{U}(q)$ as close as possible to $U(q)$) we want to find a filter $\phi(t)$ (or

$\Phi(q)$) which, when applied to $c(t)$ (or $C(q)$) and then deconvolved by $r(t)$ or $R(q)$, produces $\tilde{u}(t)$.

$$\tilde{U}(q) = \frac{C(q)\Phi(q)}{R(q)} \tag{1.28}$$

In what sense is $\tilde{U}$ to be close to U? We ask that they be close *in the least-square sense*, i.e.,

$$\int_{-\infty}^{\infty} |\tilde{u}(t) - u(t)|^2 \, dt = \int_{-\infty}^{\infty} \left|\tilde{U}(q) - U(q)\right|^2 dq \tag{1.29}$$

is minimimum. The equality in eq. (1.29) comes from Plancherel's formula. Substituting equations (1.27) and (1.28) into the right-hand side of eq. (1.29) we have

$$\int_{-\infty}^{\infty} \left| \frac{[S(q) + N(q)]\Phi(q)}{R(q)} - \frac{S(q)}{R(q)} \right|^2 dq \quad = \\ \int_{-\infty}^{\infty} |R(q)|^{-2} \left\{ |S(q)|^2 \left[1 - \Phi(q)\right]^2 + |N(q)|^2 |\Phi(q)|^2 \right\} dq \tag{1.30}$$

For the right hand of the equality it has been used the fact that the signal S and the noise N are uncorrelated and therefore their cross product, when integrated over q, gives zero. Minimizing the lower integral of eq. (1.30) with respect to Φ gives

$$\Phi(q) = \frac{|S(q)|^2}{|S(q)|^2 + |N(q)|^2} \tag{1.31}$$

This is the formula for the *Wiener filter* (WF) $\Phi(q)$. Equation (1.31) does not contain the true signal U. This makes for an important simplification: the optimal filter can be determined independently of the deconvolution function that relates S and U. The other filters we are going to consider in this work can be constructed so that this is also true for them. Therefore, from now on the smearing term will be put sideways in the discussion.

The filter in eq. (1.31) need previous knowledge on $|S|^2$ and $|N|^2$ (or, at least, on $|S|^2$, since $|C|^2$=$|S|^2$+$|N|^2$). There is no way to do this from the measured signal C alone without some other information, or some assumption or guess. Luckily, the extra information is easy to obtain in most cases. Even a rough guess of $|S|^2$ can give excellent results when it is applied to data. Unfortunately, this can not be used as a basis for an iterative method: successive iterations of Wiener filter converge to a signal of $S(q) = 0$.

Wiener filter has several drawbacks. Firstly, it is easy to see that as an estimator of the signal S it is biased, both in real and in Fourier space; in particular, the estimation of the power spectrum of the signal is increasingly biased for high q's. It is possible to construct unbiased estimators based on WF. For a review on this topic, see Rybicki & Press (1992). The second limitation is that WF is based on a minimal variance approach; therefore, it should provide a powerful reconstruction technique for random systems whose statistical behaviour does not depend on moments higher than the second, namely the variance. If the signal is strongly non-Gaussian, for example, the performance of WF is expected to be poor. Even in the framework of Gaussian random fields, it depends on an adequate modelling of the data and an assumed knowledge of the correlation function (or, equivalently, the power spectrum) of the underlying field and the nature of the statistical uncertainties. Finally, the WF always yields a conservative estimate of the underlying field, replacing noise by the zero field in the absence of good data.

Despite its limitations, Wiener filter has been extensively used in signal processing and image reconstruction. Currently, WF is usually overlooked in favor of other methods, such as Maximum Entropy, particularly in the field of image processing. However, in certain fields of Astronomy where the underlying signal is close to be a Gaussian random field it proves to be optimal or near optimal. It can be proven that, when the field to be reconstructed is a Gaussian random field, the Wiener estimator coincides with the Bayesian estimator designed to maximize the *a posteriori* probability of the field (Zaroubi *et al.* 1995). WF has been successfully applied to the study of large-scale structure (e.g. Zaroubi *et al.* 1995) and to component separation in CMB maps (Tegmark & Efstathiou 1996, Bouchet *et al.* 1995, 1999). The technique proposed by Tegmark & Efstathiou (1996) and in parallel by Bouchet *et al.* (1995) is a generalization of WF to multiple frequency channels and multiple components:

Let us consider a set of data measured at a certain frequency ν. Each measurement is a combination of m physical components $\mathbf{s} = (s_1(\nu), s_2(\nu), \cdots, s_m(\nu))$. Let now consider we have n frequency channels. In a general case, after the measurement the output data vector $\mathbf{d} = (d_1, d_2, \cdots, d_n)$ can be written as a linear convolution of the signal:

$$\mathbf{d} = \mathbf{R}\mathbf{s} + \mathbf{n} \tag{1.32}$$

where $\mathbf{R}$ is some known $n \times m$ *response* matrix and $\mathbf{n} = (n_1(\nu), n_2(\nu), \cdots, n_n(\nu))$ is a random noise vector. The noise and the signal are taken to have zero mean and we assume knowledge of the covariance matrices:

$$\begin{aligned} \mathbf{S} &= < \mathbf{s}\mathbf{s}^t > , \\ \mathbf{N} &= < \mathbf{n}\mathbf{n}^t > \ . \end{aligned} \tag{1.33}$$

In addition, signal and noise are taken to be uncorrelated:

$$< \mathbf{n}\mathbf{s}^t > = 0 \tag{1.34}$$

Following the same ideas of equations (1.29) and (1.30) we find the WF form for the multifrequency case:

$$\mathbf{W} = \mathbf{S}\mathbf{R}^t(\mathbf{R}\mathbf{S}\mathbf{R}^t + \mathbf{N})^{-1} \tag{1.35}$$

The filter in eq. (1.35) has the same drawbacks that filter (1.31). In particular, it is biased as an estimator of the power spectrum of the underlying data. Tegmark & Efstathiou (1996) proved that the modified filter

$$\mathbf{W} = \mathbf{\Lambda}\mathbf{R}^t(\mathbf{R}\mathbf{S}\mathbf{R}^t + \mathbf{N})^{-1} , \tag{1.36}$$

where the matrix $\mathbf{\Lambda} \equiv \mathrm{diag}\{\lambda_1, \lambda_2, \cdots, \lambda_m\}$ has elements $\lambda_i = 1/(\mathbf{R}^t[\mathbf{R}\mathbf{S}\mathbf{R}^t + \mathbf{N}]^{-1}\mathbf{R})_{ii}$, can be used as an unbiased estimator of the power spectrum of the underlying signal. This modified WF requires only assumptions about the power spectra of the foregrounds and the CMB. The availability of multiple frequency channels is essential: when the number of channels is reduced to one, the method degenerates not to standard WF, but to no substraction at all.

The matched filter

The Wiener filter derived in last section was obtained as an optimal filter in the least-squares sense, but there are other possible criteria to consider a filter as optimal. Let us consider a single signal $s(t)$ with a maximum at a certain t_0 (for simplicity let $t_0 = 0$) embedded in a noisy background $n(t)$. We can define the *signal to noise* ratio of the signal as the ratio between the maximum value of the signal and the variance of the field:

$$s/n = \frac{s(0)}{\sigma} \tag{1.37}$$

The variance σ includes the contributions from the noise and the signal. Our ability to detect the signal will be directly proportional to its s/n ratio. In order to maximize the chance of detecting the signal, we are interested in make the s/n ratio as large as possible. By introducing a linear filter with impulse $\phi(t)$ we modify the data in a certain way so that the new signal to noise ratio is $(s/n)_\phi = \frac{\phi \otimes s(0)}{\sigma_\phi}$, where now σ_ϕ is the variance of the filtered field. To simplify the notation, we will call $s_\phi = \phi \otimes s$. Let us define the *gain* of the filter as the ratio:

$$g = \frac{\left(\frac{s_\phi(0)}{\sigma_\phi}\right)}{\left(\frac{s(0)}{\sigma}\right)} \tag{1.38}$$

One way to maximize the gain of the filter is to make $s_\phi(0) = s(0)$ while minimizing σ_ϕ. It is useful to perform this minimization in Fourier space, where it can be proven that the filter that satisfies this minimization is

$$\Phi(q) \propto \frac{s^*(q)}{P(q)} \tag{1.39}$$

where $s(q)$ is the Fourier transform of the signal profile $s(t)$ and $P(q)$ is the power spectrum of the data. A complete proof of eq. (1.39) with the correct proportionality factors is given in Appendix A as a simplification of the derivation of scale-adaptive filters (to be introduced in Chapter 2). For the moment the value of the proportionality constant is not important for this discussion. This filter is called the *matched filter* for the signal s.

To illustrate how the matched filter works let us consider an example. Imagine a Dirac-δ signal located at $t = 0$ and then consider that the signal is measured with a detector whose instrumental point spread function (psf) is a Gaussian $G_\sigma(t)$, where σ is the width of the Gaussian. The smeared signal is then $s(t) = G_\sigma(t)$. This could be the case of a star seen through a telescope with a Gaussian psf or an extragalactic point source detected with an antenna with Gaussian beam in CMB maps. Due to inherent imperfections of the measure, the detector's output includes a certain amount of white instrumental noise that corrupts the signal. Imagine that the noise dominates so that the power spectrum of the data is well approximated by the power spectrum of white noise, that is, $P(q) \simeq$ constant. Then, the matched filter given by eq. (1.39) is

$$\Phi(q) \propto G_\sigma(q)\,. \tag{1.40}$$

This is not a surprise: it is an old and well-known result that a Gaussian filter is optimal for denoising signals with a Gaussian profile embedded in white noise. The Gaussian filter whips out the small-scale noise while preserving the features with scale σ or larger.

The previous example is useful because it illustrates another interpretation of the filtering process: to pass a filter on a data set can be considered as to calculate the correlation of the data with the impulse response of the filter. The matched filter of the example works well because its obvious correlation with the shape of the signal: this correlation is maximum where the data resemble a Gaussian and close to zero elsewhere.

Note that to produce a matched filter it is not necessary to know the signal power spectrum. The only prerequisite is the knowledge of the profile of the signal. This makes the matched filter a specially useful tool when the data contain sets of identical signals which need to be detected, specially if the signals are located in relatively small regions of the data. Maximizing the gain as defined in eq. (1.38) is equivalent to maximizing the probability of detection of the sources. For that reason, matched filter is optimal in the detection sense.

Point sources in CMB maps are ideal signals for a matched filter. Knowing the shape of the beam, it is straightforward to build a matched filter directly from the data without making any other assumption (Tegmark & Oliveira-Costa 1998, Herranz *et al.* 2001a). Though the matched filter may be less powerful than the Wiener filter in image reconstruction (and certainly less useful since it only recovers a single component while WF can be designed to separate all components simultaneously), it is more robust because makes less assumptions about the data. Besides, the matched filter can be normalised to be an unbiased estimator of the amplitude of the signal.

1.2.5 Filters and wavelets

In the last ten years the world of signal processing has suffered an authentic revolution due to the fast development of wavelet techniques. Wavelets are known to be very efficient dealing with problems of data compression, denoising and pattern recognition. The property that makes wavelets so interesting is that they keep

a good space-frequency localization. Unlike the Fourier transform, the wavelet transform allows to have information about the importance of different scales at each position.

The discussion of wavelet theory is out of the scope of this thesis. However, some basic ideas on wavelets are useful to better understand the concept of *scale* that will repeatedly appear in this work. We will find that the scale-adaptive filters that are introduced in Chapter 2 resemble wavelets. In certain particular cases, they *are* wavelets. This fact is not casual: we will use filters that select more or less narrow frequency bands in order to detect compact signals (sources) that are located in very precise positions in the real space. Therefore, these filters should inherit some of the space-frequency localization properties of wavelets.

To illustrate the concept of scale, let us consider the *discrete wavelet transform* (DWT). The wavelet basis is constructed from dilations and translations of the *mother* (or *analyzing*) wavelet function ψ and a second related function $\phi(t)$ called the *father* (or *scaling*) function:

$$\begin{aligned} \psi_{j,l} &= 2^{j/2}\psi\left(2^j t - l\right) , \\ \phi_{j,l} &= 2^{j/2}\phi\left(2^j t - l\right) , \end{aligned} \tag{1.41}$$

where j and l are integer numbers denoting the dilation and translation indices, respectively. The functions ϕ and ψ have generally compact support and they usually are chosen to be orthogonal. They must together satisfy some mathematical relations, as first shown by Daubechies (1988). In particular, two straightforward requirements are:

$$\begin{aligned} \int \psi(t)dt &= 0 , \\ \int \phi(t)dt &= 1 . \end{aligned} \tag{1.42}$$

The reconstruction of the signal $f(t)$ using the wavelet basis is given by

$$f(t) = a_{0,0}\phi_{0,0}(t) + \sum_j \sum_l w_{j,l}\psi_{j,l}(t) \tag{1.43}$$

being a, w the wavelet coefficient defined as:

$$\begin{aligned} a_{0,0} &= \int f(t)\phi_{0,0}(t)dt , \\ w_{j,l} &= \int f(t)\psi_{j,l}(t)dt . \end{aligned} \tag{1.44}$$

Equation (1.43) can be interpreted as the sum of a low resolution, smoothed function plus a series of consecutive refinements that carry information about the details of the function $f(t)$. The difference between the refinement level j and the next tell us about the structure of f at the scale j.

So the scaling functions carry information about structures of a certain scale inside a compact region. This is precisely what makes them so useful from the point of view of source detection. It suggests the idea of using wavelets as filters in order to select structures which have the same characteristic scale than the wavelet.

For the purposes of this work we are more interested in the *continuous wavelet transform* (CWT). Rather than restricting translation and dilation by the set of integers, we can allow them to vary continuously. For $R > 0$, $b \in I\!R$, define

$$\psi_{(R,b)}(t) = R^{-1/2}\psi\left(\frac{t-b}{R}\right) . \tag{1.45}$$

The continuous wavelet transform is thus defined to be:

$$Wf(R,b) = \int_{\infty}^{\infty} f(t)\psi_{(R,b)}(t)\,dt = f \otimes \bar{\psi}_R(b)\,, \tag{1.46}$$

where

$$\bar{\psi}_R(t) = R^{-1/2}\psi\left(\frac{-t}{R}\right) . \tag{1.47}$$

This transform has the form of a convolution with a scaled function. As we will see, the scale-adaptive filters to be defined in Chapter 2 have the same structure than eq. (1.45), except for the normalisation factor. As in the DWT case, continuous wavelets are localized in both real and Fourier spaces and can be used to decompose a signal into elements with different resolution, although the set of such elements is now continuous. In this case, the generalisation of eq. (1.43) is a double integral on b and R instead of a sum over indices j, l.

An example of wavelet generally used in the CWT is the *Mexican Hat Wavelet* (MHW):

$$\psi_{MHW}(t) = \frac{2}{\pi^{1/4}\sqrt{3\sigma}}\left(\frac{t^2}{\sigma^2} - 1\right)e^{-t^2/2\sigma^2} . \tag{1.48}$$

This wavelet is obtained from the second derivative of a Gaussian of width σ. Because its relation with the Gaussian (the exponential is an eigenvector of the derivation operator), the MHW gives a high correlation when it is used to filter

a Gaussian signal. This is why it has been used to detect Gaussian signals in many fields of Astronomy. MHW have been successfully used to detect/extract point sources from CMB data alone (Cayón *et al.* 2000, Vielva *et al.* 2001a) or in combination with other methods such as Maximum Entropy (Vielva *et al.* 2001b). Recently, a generalization of the Mexican Hat Wavelet to the sphere has been presented (Cayón *et al.* 2001).

Besides the detection of localized sources, wavelets have been used on CMB data to perform denoising (Sanz *et al.* 1999a,b, Tenorio *et al.* 1999) and to detect non-Gaussianity in CMB maps (Barreiro & Hobson 2001, Cayón *et al.* 2001, Martínez-González *et al.* 2001).

There are many books and reviews devoted to wavelets; for example, see Mallat (1998), Odgen (1997) or Burrus *et al.* (1998).

1.2.6 Filtering of discrete data

Before concluding this introduction, a word has to be said about the filtering of discrete data. Along all section § 1.2 it has been assumed that both the data and the filter are continuous functions. But real data are often discrete, as is the case of pixelised CMB maps or discrete time ordered data. In that case, one can choose two different ways to operate.

- One can directly choose the framework of discrete signal processing and consider the discretised form of the filters or wavelets needed. All results given in the last sections can be easily generalized to the discrete case, basically by substituting integrals by sums over the number of data points.

- Or one can consider that the discrete data are finely sampled enough and proceed with the integrals.

At the end, no matter what choice have we done, the discreteness of the data is unavoidable: in practise we will perform operations with pixels, (discrete) Fast Fourier Transforms, etc. However, the continuous notation is generally more simple and allows to reach analytical results more easily. Through the following chapters filters are always treated as continuous functions. How can harmonize this with the discreteness of the data we are going to filter?

The answer is that in most cases this is not very important. When the conditions of the sampling theorem are met, the overall system for filtering a discrete data set with a discretised version of a continuous filter is equivalent to filter the data with a 'good' discrete filter. That means that, if the mentioned conditions apply, we can take a continuous filter, evaluate it in the same grid where the discrete data are defined, and proceed. And the conditions of the sampling theorem are:

- Both the signal S and the filter Φ must be band-limited in Fourier space with $S(q) = \Phi(q) = 0$ for $|q| > q_M$.
- The data and the filter must be sampled with a sampling frequency $\omega_s > 2q_M$, where $\omega_s = 2\pi/T$ and T is the sampling interval.

Conversely, similar conditions must be met when changing the places of the Fourier space and the real space, in order to allow deconvolutions. So the data and the filter must be both of finite support (or periodic) and band-limited. The price for not satisfying these conditions is *aliasing*: a contamination of high frequencies invading the region of low frequencies.

Let us assume that the data satisfy the sampling theorem conditions. This is always true from the practical point of view when we are working with CMB maps: we have no way to perceive the data outside the measured area and with greater resolution than the pixel, so there is no other alternative than convince ourselves that the quantity we have measured is zero (or a mirror reflection) outside the measured area and that the underlying function is some interpolation of the sampled points. This automatically make the sampling theorem conditions 'true'. Then, it is only necessary to make the filter met the conditions. For this is enough that the filter is zero outside the region covered by the data in both real and Fourier space. In fact, it is sufficient if the filter drops rapidly near the limits of the allowed area.

In practise, a continuous filter will work well when applied to discrete data if:

- It decreases quickly in real space.
- It decreases quickly in Fourier space.

- Its typical scale of variation is of the order of the pixel or more.

The two first conditions remind us of wavelets. The third condition is to assure that the filter is not going to loose its shape when it is discretised. For example, it is a bad idea to filter with a Gaussian with a width much smaller than the pixel because it becomes a Heaviside step with the size of the pixel.

All the filters described in the following chapters satisfy the first two conditions mentioned above. And the third condition is an exercise of common sense for the filter user.

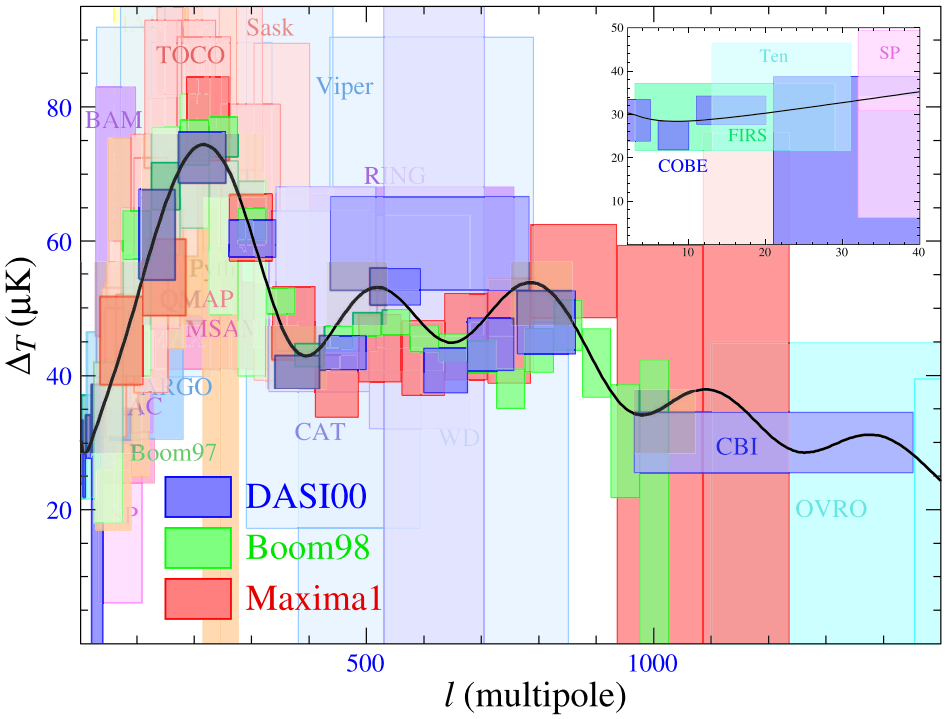

Figure 1.1: Anisotropy detections in the CMB. The boxes represent the 1-σ errors and the approximate ℓ-bandwidth of the different experiments. The solid line show the prediction of the working cosmological model ($\Omega_{tot} = 1$, $\Omega_\Lambda = 2/3$, $\Omega_b h^2 = 0.02$, $\Omega_m h^2 = 0.16$, $n = 1$, $E_i = 2.2 \times 10^{16}$ GeV and $z_{ri} = 7$), complete with the acoustic peaks. Figure reproduced with permission from Hu & Dodelson (2001).

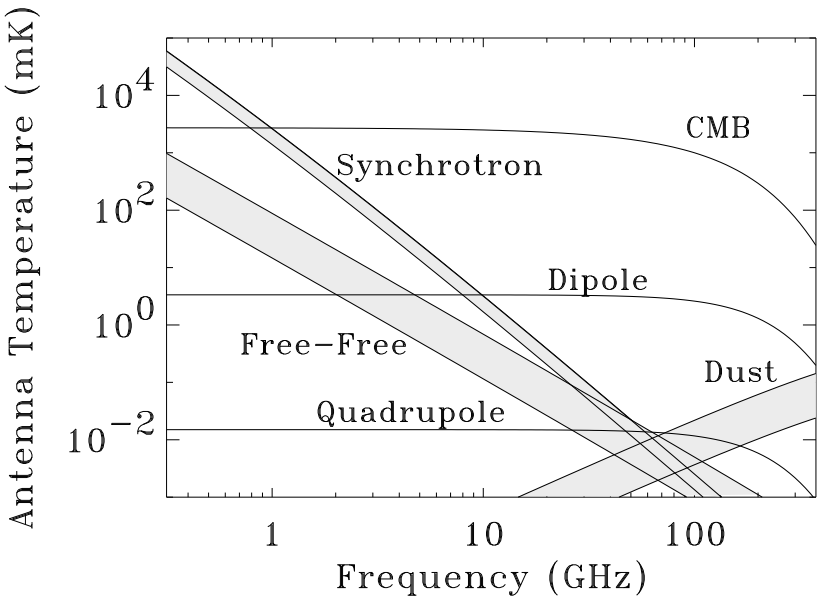

Figure 1.2: The frequency dependence and relative strength of Galactic synchrotron, free-free and dust as well as of the CMB, the dipole and the quadrupole are shown. Taken from Smoot (1999).

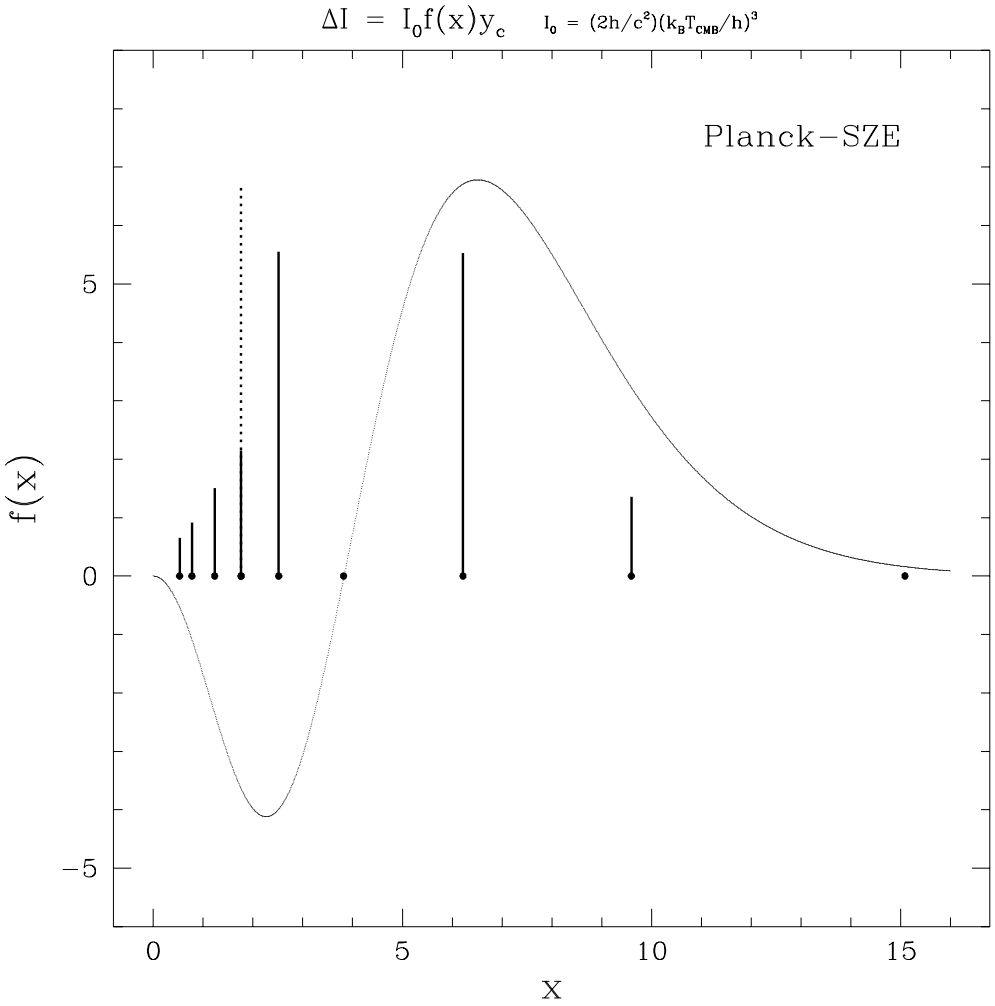

Figure 1.3: Characteristic frequency dependence of the thermal Sunyaev-Zel'dovich. At $x = 3.81$ (217 GHz) the thermal SZE is zero. Vertical lines show the relative Planck's sensitivity to detect the SZE at its different frequency channels. Figure taken with permission from Diego *et al.* (2001b).

Chapter 2

Scale-Adaptive Filters

In the first chapter of this work it was introduced the importance of component separation in CMB studies. Filtering techniques were suggested as a possible way to overcome this problem. But the usefulness of filtering is not limited to the CMB field of work. Many other disciplines in Astronomy, Physics in general, engineering, image processing and all sciences involving data analysis use filters. This chapter is devoted to introduce a new kind of filter, the *scale-adaptive filter*, in a general way. Examples will be taken from Astronomy problems when necessary.

2.1 The role of filtering in Astronomy

One of the challenges in the analysis of 1D spectra, 2D images or 3D volumes in Astrophysics and Cosmology is to overcome the problem of separating a localized signal from a background. In particular, we are interested in localized signals (sources, from now on) with central symmetry and a background that we shall assume with the properties of homogeneity and isotropy and it will be characterized by a power spectrum. Typical cases in the 1D case include: a) the spectra of QSOs, how to detect/extract absorption lines from a noisy spectrum once a profile is assumed, b) time series analysis where a localized emission is superposed on a background, c) point and extended sources to be detected in time-ordered data where the scanning strategy (for satellites like Planck) is affected by the well-known $1/f$ noise. In the 2D case, we mention as typical cases: a) cleaning of images to detect typical astrophysical sources on a white noise background, b) the

detection/extraction of point sources and extended sources (clusters of galaxies) on microwave/IR maps where the background is dominated by white noise or intrinsic CMB signal or galactic emission. In the 3D case, we remark as an example: a) the detection of clusters and large-scale structure in 3D catalogs.

The classical treatment to remove the background has been *filtering*. Low and high-pass filters reduced the contribution from the high and low frequencies present in the spectrum or image. In general, this process is extremely inefficient in dealing with very localized sources. The reason for that is that a very localized source can be decomposed in Fourier space but many waves are needed (infinite for a delta distribution!), so if low/high-pass filters are applied then at the end many artifacts (rings in 2D images) usually appear surrounding the sources.

A very important application of these principles is the detection of sources in two-dimensional astronomical images. Several packages are commonly used for this task, such as DAOfind[1], used to find stellar objects in astronomical images, and SExtractor (Bertin & Arnouts 1996). When trying to detect sources, two different problems have to be solved: first, it is necessary to take account for the background variations across the image. In addition, if instrumental noise (i. e. white noise) appears, it should be removed as far as possible in order to increase the signal to noise ratio (SNR). SExtractor estimates the local background on a grid of size set by the user and then approximates it by a low-order polynomial. Once such a background is defined the detection of sources at a certain level can be established through sets of connected pixels above certain threshold. DAOfind implicitly assumes that the background is smooth and its characteristic scale is very much larger than the scale of the stars. But in the case where the characteristic scale of variation of the background is approximately the scale of the structures the previous schemes obviously fail. An incorrect estimation of the background leads to a biased estimation of the amplitude of the sources. An example is a typical image of the cosmic microwave background radiation (CMB) at a resolution of several arcmin. If the intrinsic signal is due to a cold dark matter model (i. e. the characteristic scale of the background is of the order of ~ 10 arcmin), then the separation of point sources with the same characteristic scale becomes a very

[1]P.B. Stetson, User's Manual for Daophot II, is available at http://www.ast.cam.ac.uk/ mbt/daophot.

difficult task.

To deal with the instrumental noise the traditional procedure is filtering (e. g. Gaussian window). DAOfind filters with an elliptical Gaussian that mimics the stellar psf and then it performs the detection looking for peaks above certain threshold. SExtractor includes the possibility of filtering with several kind of filters (Top Hat, Gaussian, Mexican Hat and even self-made filters for every particular situation). An obvious advantage of this procedures (background estimation plus filtering) is that no a priori information on the structures is needed. A serious drawback is that the choice of the filter will have a great influence on the final result. The choice of filter depends on many factors, including in most cases personal preferences. In this context, it is necessary to find a systematic way to determine the optimal filter for every case.

Other methods have been used to separate different components given several microwave maps: Wiener filtering (WF) and maximum entropy methods (MEM). Regarding point sources, WF assumes a mean spectral flux dependence together with other assumptions for the other components (Tegmark & Efstathiou, 1996; Bouchet & Gispert 1999) whereas MEM assumes that the point sources are distributed like noise (Hobson et al. 1999). These methods are powerful regarding extended sources like clusters of galaxies because they use the concrete spectral dependence for the Sunyaev-Zeldovich effect. It is clear that the unknown spectral dependence for point sources remark the inefficiency of the previous methods.

A possible solution to overcome this problem came with the usage of *wavelets*. These are localized bases that allow a representation of a local object due, in general, to their basic properties: spatial and frequency localization (as opposed to the trigonometric functions appearing in Fourier decomposition). We remark at this point the success of such a technique dealing with simulated microwave maps: the "Mexican Hat" wavelet can be used in a nice way (no extra assumptions about the background) to detect/extract point sources and create a simulated catalog (Cayón et al. 2000, Vielva et al. 2001a,2001b). Two advantages emerge: on the one hand, one localizes the structures associated to the maxima of the wavelet coefficients and, what is more remarkable, we gain in the detection (as compared to real space) due to the amplification effect because at the scale of the source the background is not contributing to the dispersion. One relevant question concerns

the possibility to find *optimal* filters. Tegmark & Oliveira-Costa (1998) introduced a matched filter in order to minimize the variance in the map. With this method one can identify a big number of point sources in CMB maps. However, they failed to introduce the appropriate constraints in the minimization problem, i. e. the fact that we have a maximum at the source position at the scale defined by the source in order not to have spurious identifications.

Thus, this type of analysis lead us to the following questions: is there an optimal filter that, given the source profile and the power spectrum of the background, is the best adapted to the profile and characteristic scale of the sources?, Is the "Mexican Hat" wavelet the optimal filter (in the previous sense) dealing with point sources? In order to answer these questions, we will assume that the sources can be approximated by structures with central symmetry given by a profile $\tau(x)$, $x \equiv |\vec{x}|$ with a characteristic scale (e.g. a single maximum at its center and rapid decay at large distances). If the nD image contains different types of sources, a number of filters adapted to each profile must be used to detect them. The background will be modelled by a homogeneous and isotropic random field given in terms of the power spectrum $P(q)$, $q \equiv |\vec{q}|$. In particular, we shall explore a scale-free spectrum $P(q) \propto q^{-\gamma}$ that includes the cases of white noise ($\gamma = 0$), $1/f$ noise ($\gamma = 1$), etc. Moreover, any spectrum of physical interest often can be locally approximated by a power-law. If the characteristics of the noise are not known a priori it can be always estimated directly from the nD image. We consider the n-dimensional case and make special emphasis on the analysis of spectra ($n = 1$), 2D images ($n = 2$) and 3D volumes ($n = 3$). In all the calculations we assume that the overlapping by nearby sources is negligible and also that their contribution to the total power spectrum is also negligible. All of this is a very good approximation at least above a certain flux level.

2.2 The Scale-Adaptive Filter

Let us consider an n-dimensional ($n = 1, 2, 3$) image with data values defined by

$$y(\vec{x}) = s(x) + n(\vec{x}), \tag{2.1}$$

where x is the spatial coordinate (in the 1D case x can be also time, when we are dealing with time-ordered data sets) and $s(x), x \equiv |\vec{x}|$, represents a source with central symmetry placed at the origin with a characteristic scale (e. g. a single maximum at its center and rapid decay at large distances) and $n(\vec{x})$ a homogeneous & isotropic background (random field) with mean value $\langle n(\vec{x})\rangle = 0$ and characterized by the power spectrum $P(q), q \equiv |\vec{q}|$ (this can represent instrumental noise and/or a real background),

$$\langle n(\vec{q}) n^*(\vec{q}\,')\rangle = P(q)\delta^n(\vec{q} - \vec{q}\,'), \tag{2.2}$$

where $n(\vec{q})$ is the nD Fourier transform ($n(\vec{q}) = (2\pi)^{-n/2} \int d\vec{x}\, e^{i\vec{q}\vec{x}} n(\vec{x})$), symbol n^* represents the complex conjugate of n, $\vec{q}$ is the wave vector and δ^n is the nD Dirac distribution.

Let us introduce a spherical (centrally symmetric) filter, $\Psi(\vec{x}; \vec{b}, R)$, dependent on $n+1$ parameters:

$$\Psi(\vec{x}; R, \vec{b}) = \frac{1}{R^n}\psi\left(\frac{|\vec{x} - \vec{b}|}{R}\right), \tag{2.3}$$

where $\vec{b}$ defines a translation whereas R defines a scaling. Note the similarity of this expression with the continuous wavelet transform given by eq. (1.45). Then, we define the filtered field $w(R, \vec{b})$:

$$w(R, \vec{b}) = \int d\vec{x}\, y(\vec{x}) \Psi(\vec{x}; \vec{b}, R). \tag{2.4}$$

We do not assume a priori the positiveness of Ψ. The previous convolution can be written as a product in Fourier space, in the form

$$w(R, \vec{b}) = \int d\vec{q}\, e^{-i\vec{q}\vec{b}} y(\vec{q}) \psi(Rq). \tag{2.5}$$

where $y(\vec{q})$ and $\psi(q)$ are the Fourier transforms of $y(\vec{x})$ and $\psi(\vec{x})$, respectively. Because of the central symmetry assumed for the filter, $\psi(q)$ depends only on the modulus of $\vec{q}$. A simple calculation -taking into account eqs. (2.1) and (2.2)- gives the average at the origin $\vec{b} = 0$, $\langle w(R, \vec{0})\rangle$, and the variance, $\sigma_w^2(R) = \langle w^2(R, \vec{b})\rangle - \langle w(R, \vec{b})\rangle^2$, of the filtered field:

$$\begin{aligned} \langle w(R, \vec{0})\rangle &= \alpha \int dq\, q^{n-1}\, s(q)\psi(Rq), \\ \sigma_w^2(R) &= \alpha \int dq\, q^{n-1}\, P(q)\psi^2(Rq). \end{aligned} \tag{2.6}$$

where $q = |\vec{q}|$, $\alpha = 2, 2\pi, 4\pi$ for $n = 1, 2, 3$, respectively, (for n-dimensions $\alpha = 2\pi^{n/2}\Gamma^{-1}(n/2)$) and the limits in the integrals go from 0 to ∞.

Now, we are going to express the conditions in order to obtain an optimal filter for the detection of the source $s(x)$ at the origin. One basic idea is to find a filter such that when the original image is filtered with a scale $R_o \approx R_s$ -being R_s the characteristic scale of the source- one obtains the maximum *detection level* $\mathcal{D}_w$

$$\mathcal{D}_w \equiv \frac{\langle w(R, \vec{0})\rangle}{\sigma_w(R)}. \tag{2.7}$$

Taking into account the fact that the source is characterized by a single scale R_s, other basic idea is to generate a filter giving the maximum contribution at the center of the source at a filtering scale $R_o \approx R_s$. Finally, we would like to estimate directly the amplitude of the source by the previous number. Therefore, taking into account these basic ideas we will introduce from the mathematical point of view the optimal filters we are looking for.

By definition a filter will be called *scale-adaptive filter* (SAF) if the following conditions are satisfied:

1. There exists a scale R_o such that $\langle w(R, \vec{0})\rangle$ has a maximum at that scale.
2. $\langle w(R_o, \vec{0})\rangle = s(0) \equiv A$, i. e. $w(R, \vec{0})$ is an unbiased estimator of the amplitude of the source.
3. The variance of $w(R, \vec{b})$ has a minimum at the scale R_o, i. e. we have an efficient estimator.

As a by-product, the ratio given by eq. (2.7) will be maximum. We remark that no other information about the source profile is assumed, so "optimal" must be understood in the previous sense.

By introducing the profile $\tau(x)$ of the source, $s(x) = A\tau(x)$, the condition A.10 and the equation (2.6) give the constraint:

$$\int dq\, q^{n-1}\, \tau(q)\psi(R_o q) = \frac{1}{\alpha}, \tag{2.8}$$

whereas the condition A.9 gives the constraint:

$$\int dq\, q^{n-1}\, \tau(q)\psi(R_o q) \left(n + \frac{d\ln\tau}{d\ln q}\right) = 0. \tag{2.9}$$

So, the problem is reduced to the functional minimization (with respect to ψ) of $\sigma_w^2(R)$ given by equation (2.6) with the constrains given by equations (2.8) and (2.9). This minimization incorporate these constraints through a couple of Lagrangian multipliers. The full derivation of the solution is described in Appendix A. The solution (*scale-adaptive filter*) is found to be:

$$\tilde{\psi}(q) \equiv \psi(R_o q) = \frac{1}{\alpha}\frac{\tau(q)}{P(q)}\frac{1}{\Delta}\left[nb + c - (na+b)\frac{d\,ln\tau}{d\,lnq}\right], \tag{2.10}$$

$$\begin{aligned} \Delta &= ac - b^2, \\ a &\equiv \int dq\, q^{n-1}\, \tfrac{\tau^2}{P}, \\ b &\equiv \int dq\, q^{n-1}\, \tfrac{\tau}{P}\tfrac{d\,\tau}{d\,lnq}, \\ c &\equiv \int dq\, q^{n-1}\, \tfrac{1}{P}\left[\tfrac{d\,\tau}{d\,lnq}\right]^2. \end{aligned} \tag{2.11}$$

Therefore, we have obtained analytically the functional form of the scale-adaptive filter (its shape and characteristic scale are associated to the source profile and power spectrum). It is clear that assuming an adimensional dependence $\tau(x/R_s)$, where R_s is the characteristic scale of the source, then such scale will appear explicitly in the form $\tilde{\psi}(qR_s)$. Obviously, we assume all the differentiable and regularity conditions at $q \to 0, \infty$ for τ and $P(q)$ in order to have finite expressions for a, b, c. Let us remark that if we assume the behavior $\frac{\tau(q)}{P(q)} \to 0$, $\frac{1}{P}\frac{d\,\tau}{d\,lnq} \to 0$ for $q \to 0$ then $\psi(q) \to 0$ and Ψ is a "compensated" filter, i. e. $\int d\vec{x}\,\Psi = 0$. Strictly speaking there is another condition to be satisfied to get the reconstruction of the image and thus to have a wavelet: $\int dq q^{-1}\psi^2(q) < \infty$ (the admissibility condition).

Taking into account eq. (2.5) the amplitude will be estimated as:

$$A = w(R_o, \vec{0}) = \int d\vec{q}\, y(\vec{q})\, \tilde{\psi}(q), \tag{2.12}$$

where $\tilde{\psi}$ is given by eq. (2.10).

2.2.1 Scale-Adaptive filters on real space

The equation (2.10) can be written on real space as follows:

$$\Psi(\vec{x}; R_o, \vec{0}) = \frac{1}{R_o^n}\psi\left(\frac{|\vec{x}|}{R_o}\right) = \frac{1}{\alpha\Delta}\left[(nb+c)F(x) - (na+b)G(x)\right], \tag{2.13}$$

where F and G are the inverse Fourier transform of $\frac{\tau}{P}$ and $\frac{1}{P}\frac{d\tau}{dlnq}$, respectively.

For a flat background, i.e. $P =$ constant, and assuming the behaviour $q^{\frac{1}{2}(n+1)}\tau \rightarrow 0$ when $q \rightarrow \infty$, one obtains

$$b = -\frac{n}{2}a, \quad F = \frac{\tau(x)}{P}, \quad G(x) = -\frac{1}{P}[n\tau + \frac{d\tau}{dlnx}]. \tag{2.14}$$

If we also assume a Gaussian profile, i.e. $\tau = e^{-\frac{x^2}{2\theta^2}}$, one finds the scale-adaptive filter:

$$\Psi(\vec{x};\theta,\vec{0}) = \frac{1}{\pi^{n/2}\theta^n} e^{-\frac{x^2}{2\theta^2}} \left[1 + \frac{n}{2} - \left(\frac{x}{\theta}\right)^2\right], \tag{2.15}$$

that is a useful formula to be used for the detection of nD-Gaussian structures on nD-images (e.g. spectra, 2D images or 3D volumes) on a flat background.

On the other hand, if one assumes a Gaussian profile but a non-flat spectrum one can easily find

$$\Psi(\vec{x};\theta,\vec{0}) = \frac{1}{\alpha\Delta}[(nb+c)F(x) - (na+b)\theta^2\nabla^2 F(x)], \tag{2.16}$$

being $F(x)$ the Fourier transform of $\frac{\tau}{P}$.

2.3 Detection level, gain and reliability

Taking into account the previous expression (2.10), one can calculate the detection level (see equation (2.7))

$$\mathcal{D}_w = A\left(\frac{\alpha\Delta}{n^2a + 2nb + c}\right)^{1/2}. \tag{2.17}$$

On the other hand, we can calculate the dispersion of the field $n(\vec{x})$

$$\sigma_b^2 = \alpha \int dq\, q^{n-1} P(q), \tag{2.18}$$

that allows to define a detection level on real space as

$$\mathcal{D} = \frac{A}{\sigma_b}. \tag{2.19}$$

The *gain* going from real space to filter space is defined by

$$g \equiv \frac{\mathcal{D}_w}{\mathcal{D}} = \frac{\sigma_b}{\sigma_w(R_o)}. \tag{2.20}$$

If the background has a characteristic scale different from the scale of the structures (sources) to be detected, it is obvious that $\sigma_w < \sigma_b$, so that we have a real gain going from real space to filter space.

The identification of sources as peaks above a high threshold (e. g. $3\sigma_w$) in filter space gives a low probability of false detections (reliability) because if the background has a characteristic scale different from the sources then everything detected with our method is real, but if both scales are comparable one can give an estimate based on the fluctuations of the background. For instance, in the case of a Gaussian background, false detections above $3\sigma_w$, due to the Gaussian background, appear with a probability $\simeq 1.5 \times 10^{-3}$ based on the formula

$$Pr(w > A) = \frac{1}{2}\,\mathrm{erfc}\left(\frac{A}{2^{1/2}\sigma_w}\right). \tag{2.21}$$

To select the false detections from the real ones one can study the scale-adaptive filter profile nearby any real source (see the last paragraph of § 2.7).

Regarding the completeness (i. e. how many real sources we miss with our method), this is a complicated topic because the background can slightly modify the location of the peaks and their amplitude. We will address this problem via numerical simulations (see § 2.6).

2.4 Gaussian source on a background

In many physical applications the standard response of the instruments can be approximated by a Gaussian function. In particular, the Point Spread Function (PSF) for many telescopes is of Gaussian type. Other more specific applications are related to the Cosmic Microwave Background (CMB), where the antennas used are well approximated by a Gaussian. Dealing with absorption systems associated to QSOs, if the absorption line is not saturated and is dominated by thermal motions, then the line is usually approximated by an inverted Gaussian inserted in a continuum plus noise.

Let us assume that the source and the background can be represented by

$$s(x) = A\,e^{-\frac{x^2}{2\theta^2}}, \quad P = Dq^{-\gamma}. \tag{2.22}$$

The structure to be detected could have an intrinsic Gaussian profile or it could be a point source in n-dimensions observed with an instrument that can be

modelled through a Gaussian pattern with a beam size θ. The background can be described by a scale-free spectrum. In this case: $\tau(q) = \theta^n e^{-\frac{1}{2}(q\theta)^2}$, and equations (2.11) give

$$a = \frac{\theta^{n-\gamma}}{2D}\Gamma(m), \quad b = -ma, \quad c = m(1+m)a, \quad m \equiv \frac{n+\gamma}{2} \tag{2.23}$$

and the scale-adaptive filter is

$$\tilde{\psi}(q) = \frac{1}{\alpha\Gamma(m)}(q\theta)^\gamma e^{-\frac{1}{2}(q\theta)^2}\left[2+\gamma-n+\frac{n-\gamma}{m}(q\theta)^2\right]. \tag{2.24}$$

Taking into account the q behaviour in this formula, one obtains a compensated filter (i. e. $\psi(q=0)=0$) if $\gamma > 0$ or $\gamma = n-2$. In figure 2.1 appear the optimal scale-adaptive filters for the 1D, 2D and 3D cases, respectively, and scale-free power spectrum with indexes $\gamma = 0, 1, 2, 3$. There is a degeneration in the case $n = 3$, where the $\gamma = 1$ line overlaps with the $\gamma = 3$ line, and in the case $n = 2$, where the $\gamma = 0$ line overlaps with the $\gamma = 2$ one. This degeneration can be deduced directly from equation (2.23).

On the other hand, the detection level in filter space is given by equation (2.18)

$$\mathcal{D}_w = A\left[\frac{2\alpha}{D}\frac{\Gamma(1+m)}{4m+(n-\gamma)^2}\right]^{1/2}\theta^{\frac{n-\gamma}{2}}. \tag{2.25}$$

Finally, it is interesting to remark that the cases $\gamma = n$ and $\gamma = n-2$ give the same scale-adaptive filter

$$\tilde{\psi}(q) = \frac{2}{\alpha\Gamma(n)}(q\theta)^n e^{-\frac{x^2}{2\theta^2}}. \tag{2.26}$$

Therefore, in the cases $n = 2, \gamma = 0, 2$ the Mexican Hat is found to be the scale-adaptive filter. This justify the use of this wavelet to detect point sources in CMB maps (Cayón 2000, Vielva 2001a, 2001b).

2.4.1 Gaussian source and white noise

This sub-case corresponds to $P =$ constant or $\gamma = 0$, and the scale-adaptive filter is

$$\tilde{\psi}(q) = \frac{2}{\alpha\Gamma(\frac{n}{2})}e^{-\frac{1}{2}(q\theta)^2}\left[1-\frac{n}{2}+(q\theta)^2\right], \tag{2.27}$$

that gives a scale-adaptive filter for the analysis in the different dimensions except for $n = 2$ that gives the Mexican Hat wavelet

$$\psi = \frac{1}{\pi} e^{-\frac{1}{2}(q\theta)^2} (q\theta)^2. \tag{2.28}$$

The detection level is given by equation (2.25):

$$\mathcal{D}_w = A \left[\frac{\alpha \Gamma(\frac{n}{2})}{2D(1+n/2)} \right]^{1/2} \theta^{n/2}. \tag{2.29}$$

We have calculated the contribution of sources to the power spectrum in order to estimate their influence in calculating the scale-adaptive filter. We arrive to the conclusion that if the signal/noise ratio (i. e. dispersion associated to the sources over dispersion associated to the background) is $(\sigma_s/\sigma_b) < 0.6\,(l_p/\theta)$, being l_p the pixel scale and θ the width of the source, then the extra contribution to the scale-adaptive filter coefficients a, b, c is less than a 10%.

2.4.2 Gaussian source and $1/f$ noise

Let us assume a source with a Gaussian profile and a background represented by $1/f$ noise, i.e. $P = Dq^{-1}$ or $\gamma = 1$. In this case $\tau(q) = \theta^n e^{-\frac{1}{2}(q\theta)^2}$, and equation (2.24) gives the scale-adaptive filter

$$\tilde{\psi}(q) = \frac{1}{\alpha \Gamma(\frac{n+1}{2})} e^{-\frac{1}{2}(q\theta)^2} (q\theta) \left[3 - n + 2\frac{n-1}{n+1} (q\theta)^2 \right], \tag{2.30}$$

For instance, in the case $n = 1$ one has the wavelet $\psi = (q\theta) e^{-\frac{1}{2}(q\theta)^2}$, that is the scale-adaptive filter to be used to detect a Gaussian signal on 1D spectra. In this case the detection level is (see equation(2.25)) $\mathcal{D}_w = AD^{-1/2}$.

2.5 Exponential source on a background

Typical example in Astrophysics is the exponential disk associated to spiral galaxies. Another interesting application could be in some areas of physics where the profile expected for the signal associated to the detection of some particles could be of exponential type.

Let us assume that the source and background can be represented by

$$s(x) = e^{-\frac{x}{\lambda}}, \quad P = Dq^{-\gamma}. \tag{2.31}$$

In this case:

$$\tau(q) = \beta\lambda^n \left[1 + (q\lambda)^2\right]^{-\frac{n+1}{2}}, \tag{2.32}$$

where $\beta = (\frac{2}{\pi})^{1/2}, 1, 2(\frac{2}{\pi})^{1/2}$ for $n = 1, 2, 3$, respectively, and equations (2.11) give

$$a = \frac{\beta^2\lambda^{n-\gamma}}{2D} \frac{\Gamma(m)\Gamma\left(1 + \frac{n-\gamma}{2}\right)}{\Gamma(n+1)}, \tag{2.33}$$

$$b = -ma, \quad c = \frac{n+1}{n+2} m(1+m)a, \quad m \equiv \frac{n+\gamma}{2} \tag{2.34}$$

and the scale-adaptive filter is

$$\tilde{\psi}(q) = \frac{2L}{\alpha\beta}(q\lambda)^\gamma \left[1 + (q\lambda)^2\right]^{-\frac{n+1}{2}} \left[1 + \frac{\gamma - n}{2}(n+1) + M\frac{(q\lambda)^2}{1 + (q\lambda)^2}\right], \tag{2.35}$$

$$\begin{aligned} L &\equiv \frac{\Gamma(n+1)}{\Gamma\left(\frac{n+\gamma}{2}\right)\Gamma\left(2 + \frac{n-\gamma}{2}\right)}, \\ M &\equiv \frac{n-\gamma}{n+\gamma}(n+1)(n+2). \end{aligned} \tag{2.36}$$

In figure 2.2 appear the scale-adaptive filters for the 1D, 2D and 3D cases, respectively, and power spectrum with indexes $\gamma = 0, 1, 2, 3$. The filter profiles are more extended than in the Gaussian source, as one can expect from the more gentle fall of the exponential source.

An interesting case is $\gamma = n$, then the scale-adaptive filter is

$$\tilde{\psi}(q) = \frac{2n}{\alpha\beta}(q\lambda)^n \left[1 + (q\lambda)^2\right]^{-\frac{n+1}{2}}. \tag{2.37}$$

2.5.1 Exponential source and white noise

For this sub-case $\gamma = 0$, then equation (2.35) leads to the scale-adaptive filter

$$\tilde{\psi}(q) = \frac{2}{\alpha\beta} \frac{\Gamma(n+1)}{\Gamma(\frac{n}{2})\Gamma(2 + \frac{n}{2})} \left[1 + (q\lambda)^2\right]^{-\frac{n+1}{2}} \left[1 + \frac{n}{2}(n+1) + \frac{(n+1)(n+2)(q\lambda)^2}{1 + (q\lambda)^2}\right]. \tag{2.38}$$

For instance, for an exponential structure to be optimally detected in a 1D spectrum, we must use

$$\tilde{\psi}(q) = \frac{8}{(2\pi)^{1/2}} \frac{(q\lambda)^2}{\left[1 + (q\lambda)^2\right]^2}. \tag{2.39}$$

2.5.2 Exponential source and $1/f$ noise

An interesting case is $n = \gamma = 1$, equation (2.35) gives

$$\tilde{\psi}(q) = \left(\frac{\pi}{2}\right)^{1/2} \frac{(q\lambda)}{1 + (q\lambda)^2}. \tag{2.40}$$

2.6 Simulations of one-dimensional Gaussian sources and $1/f$ noise

In order to test some of the ideas proposed in previous sections, we simulated the case of one-dimensional Gaussian sources on a background. The kind of background simulated is the well-known $1/f$ noise. This kind of noise appears very often in many devices in experimental physics. Further simulations with 2-dimensional data and realistic realizations of noise will be carried on in further chapters of this work.

For the sake of simplicity, all the simulated sources have the same amplitude and that is set to be 1 (in arbitrary units). 100 of these sources were deployed over a 32768 pixel 'field'. The number of sources and the size of the field were selected in order to have enough sources for statistical studies, to avoid (as far as possible) the overlapping of the sources and to minimize the contribution of the sources to the total dispersion of the simulations. The width of the Gaussian profiles was chosen to be $\theta \simeq 3\theta_p$: this is the case for a pixel of $1.5'$ and a Gaussian with a FWHM of $5'$. Noise was added so that the SN ratio of the sources, defined as the ratio A/σ_b (where A is the amplitude of the source and σ_b is the standard deviation of the noise), assumes values of $2, 3, 4$ and 5. Finally, the optimal filter, given by eq. (2.30) with $n = 1$, was applied to the image.

To compare with a more traditional filtering scheme, we filtered the images also with a Gaussian of width equal to θ and a Mexican Hat wavelet of width equal to θ.

N	s/n	σ	$\bar{A}$	s/n_m	$\sigma_{A_{op}}$	$d_{3\sigma}$	$e_{3\sigma}$	$d_{5\sigma}$	$e_{5\sigma}$
1	2	0.5085	1.0670	2.0983	0.5075	19	12	0	0
2	3	0.3454	1.0195	2.9514	0.3432	47	9	1	0
3	4	0.2656	0.9994	3.7629	0.2582	72	6	10	0
4	5	0.2189	0.9890	4.5174	0.2070	90	9	36	1

Table 2.1: Simulations. The first column shows the number of the simulation. The second column indicates the original SN ratio, as explained in the text. Column 3 indicates the dispersion of the map. Column 4 shows the mean amplitude of the sources as measured from the map and column 6 indicates the variance of the source amplitudes. This quantity is not equal to the original SN ratio (column 2) because the sources contribute to the final dispersion of the image, thus lowering the final, measured SN ratio. Column 5 indicates the mean SN ratio of the sources (that is, the ratio between the quantities in columns 4 and 3). Columns 7 and 8 indicate the number of sources detected at the 3σ level and the number of 3σ detections that do not correspond to real sources, respectively. Finally, columns 9 and 10 indicate the number of sources detected at the 5σ level and the number of 5σ detections that do not correspond to real sources, respectively.

This is a rather naive usage of the Mexican Hat wavelet and the Gaussian source because the optimal width for these filters in the general case is not the source scale (Cayón 2000, Vielva 2001a), but it serves us well because what we intend is to compare how do filters work when we have no further information about the data (i.e., the optimal scale, which is different for each background). The result of these simulations is shown in tables 2.1 and 2.2. Table 2.1 refers to the original simulations. It shows the original SN ratio of each simulation as well as statistical quantities of interest such as the dispersion of the map and the mean amplitude of the sources in it. Finally, it shows the number of sources directly detected from the simulations above 3σ and 5σ thresholds and the number of spurious detections above these thresholds. As expected, only a few sources are detected, except for the most favorable cases (high original SN ratio and low detection threshold). The small bias in the mean measured amplitude is due to pixelization effects. Table 2.2 refers to the simulations in table 2.1 after filtering with a Gaussian of width θ, a Mexican Hat wavelet of width θ and the scale-adaptive filter. Each row in table 2.2 corresponds to the same row in table 2.1.

A Gaussian filter smoothes the image, removing small-scale noise. It also

N	σ_f	$\bar{A}_f$	s/n_f	σ_{A_f}	g	$d_{3\sigma}$	$e_{3\sigma}$	$d_{5\sigma}$	$e_{5\sigma}$
				Gaussian filter					
1	0.4816	0.7466	1.5504	0.4822	0.7389	8	3	0	0
2	0.3260	0.7237	2.2204	0.3216	0.7523	27	4	0	0
3	0.2494	0.7134	2.8602	0.2411	0.7601	46	3	0	0
4	0.2045	0.7076	3.4604	0.1929	0.7660	67	2	5	0
				Mexican Hat wavelet					
1	0.2507	1.0070	4.0169	0.2475	1.9144	77	31	17	2
2	0.1798	0.9762	5.4298	0.1665	1.8397	93	16	64	4
3	0.1471	0.9642	6.5548	0.1250	1.7419	96	5	94	4
4	0.1292	0.9587	7.4214	0.1008	1.6429	98	2	98	2
				Scale-Adaptive filter					
1	0.2288	1.0159	4.4408	0.2217	2.1164	87	13	26	2
2	0.1673	0.9919	5.9305	0.1496	2.0094	94	8	79	5
3	0.1394	0.9825	7.0474	0.1131	1.8728	97	3	97	3
4	0.1244	0.9778	7.8592	0.0918	1.7398	99	1	99	1

Table 2.2: Gaussian, Mexican Hat and Scale-Adaptive filter results. The first column shows the number of the simulation, the second shows the dispersion of the filtered map, the third column indicates the mean estimated amplitude of the sources, the fourth column shows the mean SN ratio of the sources, column 5 indicates the dispersion of the measured amplitudes of the sources, column 6 indicates the gain of the filter (calculated through $(s/n)_{filter}/(s/n)_{original}$) and columns 7 and 8 indicate the number of sources detected at a 3σ threshold and the number of 3σ detections that do not correspond to real sources, respectively. Finally, columns 9 and 10 indicate the number of sources detected at the 5σ level and the number of 5σ detections that do not correspond to real sources, respectively.

smoothes the source peaks, thus lowering the amplitude of detected sources. For the case of $1/f$ noise the dominant fluctuations appear at large scales and are not affected by the Gaussian filter. The large-scale features may contribute to contamination in two different ways: they can conceal sources in large 'valleys' and

can produce spurious peaks. None of these effects can be avoided with a Gaussian filter. On the other hand, the smoothing effect of the Gaussian filter takes place normally and lowers the amplitude of the sources. Therefore, the number of true detections is smaller than in the non-filtered image, and the spurious detections are not removed even in the highest s/n case. The gains, indicated in column 6, clearly reflects this situation ($g < 1$).

The Mexican Hat wavelet has a better performance under $1/f$ conditions. The Mexican Hat removes large-scale fluctuations, allowing the 'hidden' sources to arise above the detection threshold. For example, in the case of original $s/n = 2.95$ there were 47 sources above the 3σ level and only 1 above the 5σ level. After filtering with the Mexican Hat, there are 93 detections above 3σ level and 64 above the 5σ level, a significant improvement. The number of spurious sources remains almost the same.

The scale-adaptive filter also deals with the large-scale structure. It is constructed to enhance all fluctuations in the source scale, while removing fluctuations that arise in other scales. In addition, it is required to be unbiased with respect to the amplitude. In practice, the amplitude is slightly underestimated due to the propagation of pixelization effects. This small bias is lower than a 10% and can be calibrated in any case. In the $1/f$ case the number of true detections is higher than in the Mexican Hat case and the number of spurious sources is comparable or slightly reduced. Only in the case of low initial SN ratio the number of spurious detections is greater. This is due to the fact that this scale-adaptive filter enhances all fluctuations in the source scale. For the case of initial SN ratio of 2.95 we find 94 sources (of 100) and 8 spurious detections (a reliability close to 10%) above the 3σ level, a result very similar to the obtained with the Mexican Hat. Above the 5σ detection level the scale-adaptive filter finds 79 sources where the Mexican Hat found only 64. The number of spurious sources have not increased significantly (from 4 to 5). For higher initial SN ratios the completeness and reliability quickly improve.

The gain obtained with the scale-adaptive filter is greater than the one obtained with the Mexican Hat. It can be analytically calculated, using eqs. (2.6) and (2.17)

for the scale-adaptive filter and its equivalents for the Mexican Hat, leading to:

$$\frac{g_{op}}{g_{mh}} = \left[\frac{4}{\pi} \frac{\Gamma(\frac{1+\gamma}{2})\Gamma(\frac{5-\gamma}{2})}{1+\frac{(1-\gamma)^2}{2(1+\gamma)}} \right]^{1/2} \tag{2.41}$$

for the one-dimensional case. This formula holds while $\gamma \leq 1$. According to equation (2.41), the ratio g_{op}/g_{mh} is 1.41 in the case $\gamma = 0$ and 1.13 in the case $\gamma = 1$. The mean observed ratio in the simulations is 1.31 and 1.08 respectively, and fits well with our expectations. As a conclusion we have that scale-adaptive filter gives higher gains than the classical Mexican Hat filter.

In figure 2.3 an example of the simulations is shown. In the top panel there is a 500 pixel wide subsection of the $s/n = 3$, $\gamma = 1$ simulation. This subsection corresponds to a region in which the large-scale noise has a positive value. Four sources are present in this area, all of them arising above the 3σ level (indicated with a dotted line). The position of the sources are marked with an asterisk in the lower panel. Additionally, there are three peaks, corresponding to background fluctuations, that arise above the 3σ level. The second panel from the top shows the image after filtering with the (optimal) scale-adaptive filter. The large-scale features have been removed and also the small-scale noise is reduced. The sources have been amplified with respect to the original map and now all of them reach the 3σ level but the spurious peaks have been removed. The amplitudes of the sources remain unbiased and close to the true value of 1. In the third panel from the top there is the image after filtering with a Gaussian. The whole image has been smoothed and now one of the sources barely reaches the 3σ level. The large-scale fluctuations remain untouched and all the spurious peaks remain in the filtered image. In the bottom panel we see the image after filtering with the Mexican Hat. The large-scale fluctuations are also removed as well as the small-scale noise, as in the case of the scale-adaptive filter. Nonetheless, the gain is lower and only three of the sources reach the 3σ level. Additionally, it is found that the small-scale noise removal is less efficient in the case of the Mexican Hat.

2.7 Extraction of sources

The scale-adaptive filter gives the position and an unbiased estimator of the amplitude of the source. We propose to make the extraction of the source on real space, i. e. one subtracts the function $A\tau(|\vec{x}-\vec{x_o}|)$, being τ the given profile and A the estimated amplitude, at the position of the source $\vec{x_o}$.

From the practical point of view, in order to select the appropriate sources (with a given scale and avoiding to select spurious detections if the background and/or noise are manifest at scales comparable to the sources) we can operate with the scale-adaptive filter at other different scales R as given by equation (2.10) but with $\tilde{\psi}(qx), x \equiv R/R_o$. If the scale that gives the maximum do not correspond to the scale we are looking for then this is a spurious source (or another type of source with a different scale). As an additional check, we can calculate the source profile in the filter space nearby any real source, e. g. for a Gaussian profile the behaviour around the maximum $R_o \approx R_s$ must be

$$\langle w(R)\rangle = Ax^{\gamma}\left(\frac{2}{1+x^2}\right)^m\left[1+\frac{n-\gamma}{2}\frac{1-x^2}{1+x^2}\right], \quad x \equiv \frac{R}{R_o}, \quad m \equiv \frac{n+\gamma}{2}, \quad (2.42)$$

and an analogous behaviour can be found for the exponential profile. If a detected source do not follows such a behaviour then it would be consider as a false detection and must be deleted from the initial catalog.

2.8 Conclusions

In this chapter it has been introduced for the first time the concept of *scale-adaptive filter* to detect/extract spherical sources on a background modelled by a (homogeneous & isotropic) random field characterized by its power spectrum. It has been obtained a generic analytical formula that allows to calculate such a scale-adaptive filter either in Fourier or real space as a function of the source profile and power spectrum of the background. The scale-adaptive filter is an unbiased an efficient estimator of the amplitude of the sources.

The previous formula has been applied to the cases of a Gaussian and an exponential profile embedded in a background with scale-free spectra. In particular, the interesting cases of white noise and $1/f$ noise have received a careful treatment.

The detection level has been calculated for the physically interesting cases of spectra, images and volumes. For some particular cases, the scale-adaptive filters are wavelets (e. g. a Gaussian source embedded in white noise in the 2D case requires a scale-adaptive filter that is a Mexican Hat wavelet).

To test these ideas, simulations of Gaussian sources embedded in a $1/f$ noise have been performed. The scale-adaptive filters have been compared using these simulations with Gaussian filters and Mexican Hat wavelets. In the last two cases the gain is lower, the noise removal is less efficient and the number of real detections is smaller.

The extraction of the sources identified at a certain scale is proposed to be done directly on real space. At the location of the source $\vec{x_o}$ one subtracts the function $A\tau(|\vec{x} - \vec{x_o}|)$, being τ the given profile.

All the calculations assume that the overlapping of nearby sources is negligible and the contribution of the sources to the background is also negligible. This is a very good approximation in many cases of interest, at least above a certain flux level.

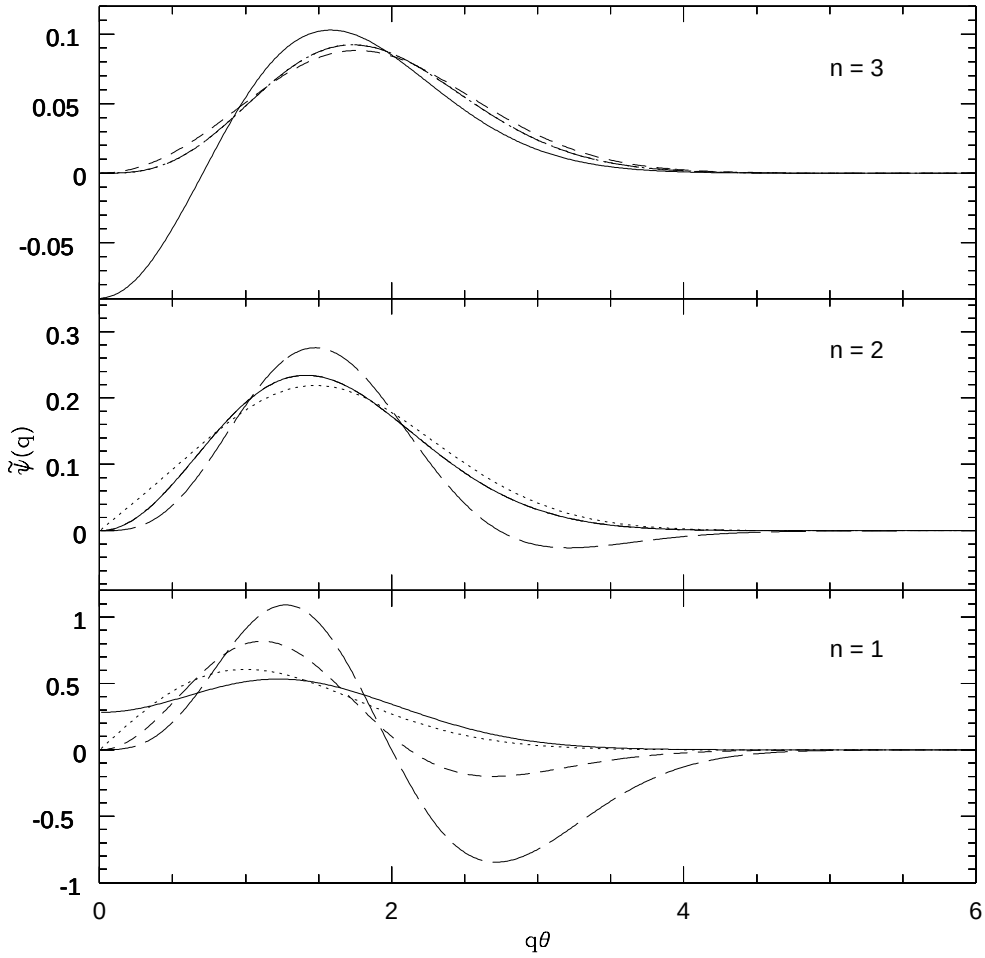

Figure 2.1: Scale-adaptive filters for a Gaussian source in a background $P(q) = Dq^{-\gamma}$ for $\gamma = 0$ (white noise, solid line), $\gamma = 1$ ($1/f$ noise, dotted line), $\gamma = 2$ (short-dashed line) and $\gamma = 3$ (large-dashed line). One, two and three-dimensional cases are represented. There is a degeneration in the case $n = 3$, where the $\gamma = 1$ line overlaps with the $\gamma = 3$ line, and in the case $n = 2$, where the $\gamma = 0$ line overlaps with the $\gamma = 2$ one.

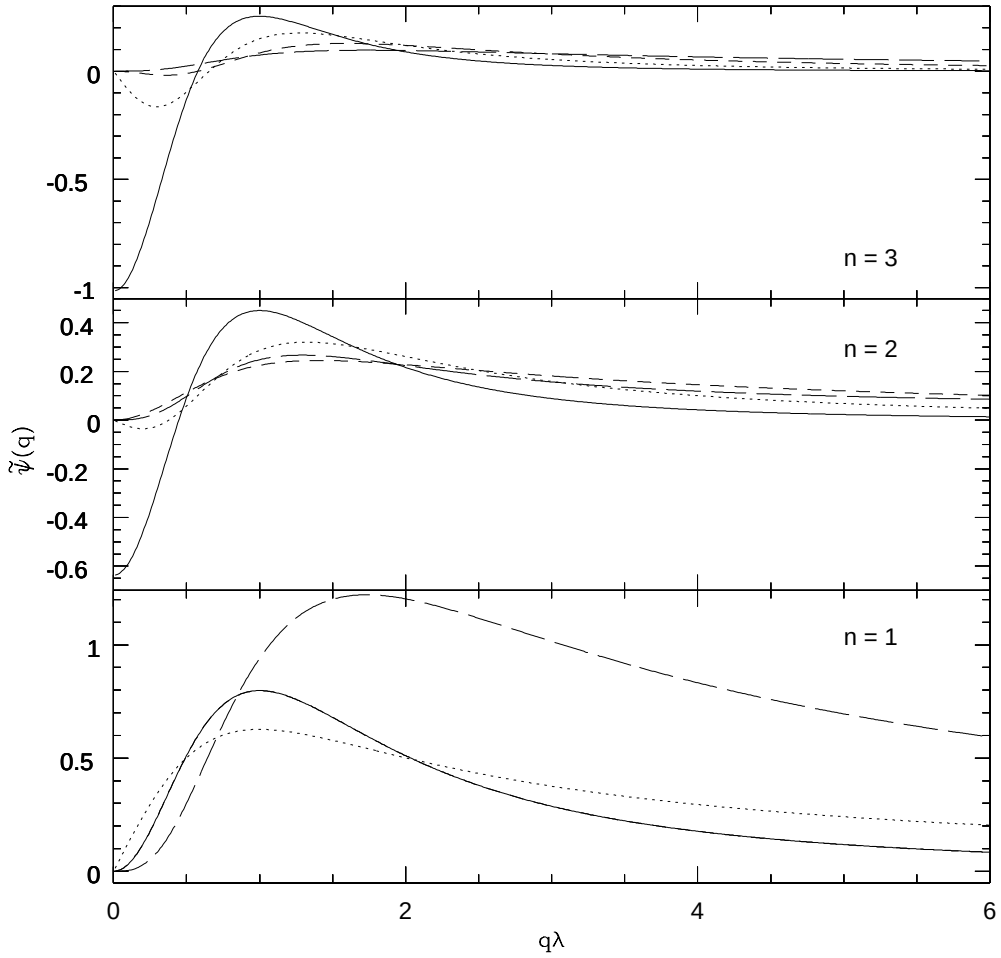

Figure 2.2: Scale-adaptive filters for an exponential source in a background $P(q) = Dq^{-\gamma}$ for $\gamma = 0$ (white noise, solid line), $\gamma = 1$ ($1/f$ noise, dotted line), $\gamma = 2$ (short-dashed line) and $\gamma = 3$ (large-dashed line). One, two and three-dimensional cases are represented. There is a degeneration in the case $n = 3$, where the $\gamma = 0$ line overlaps with the $\gamma = 2$ one.

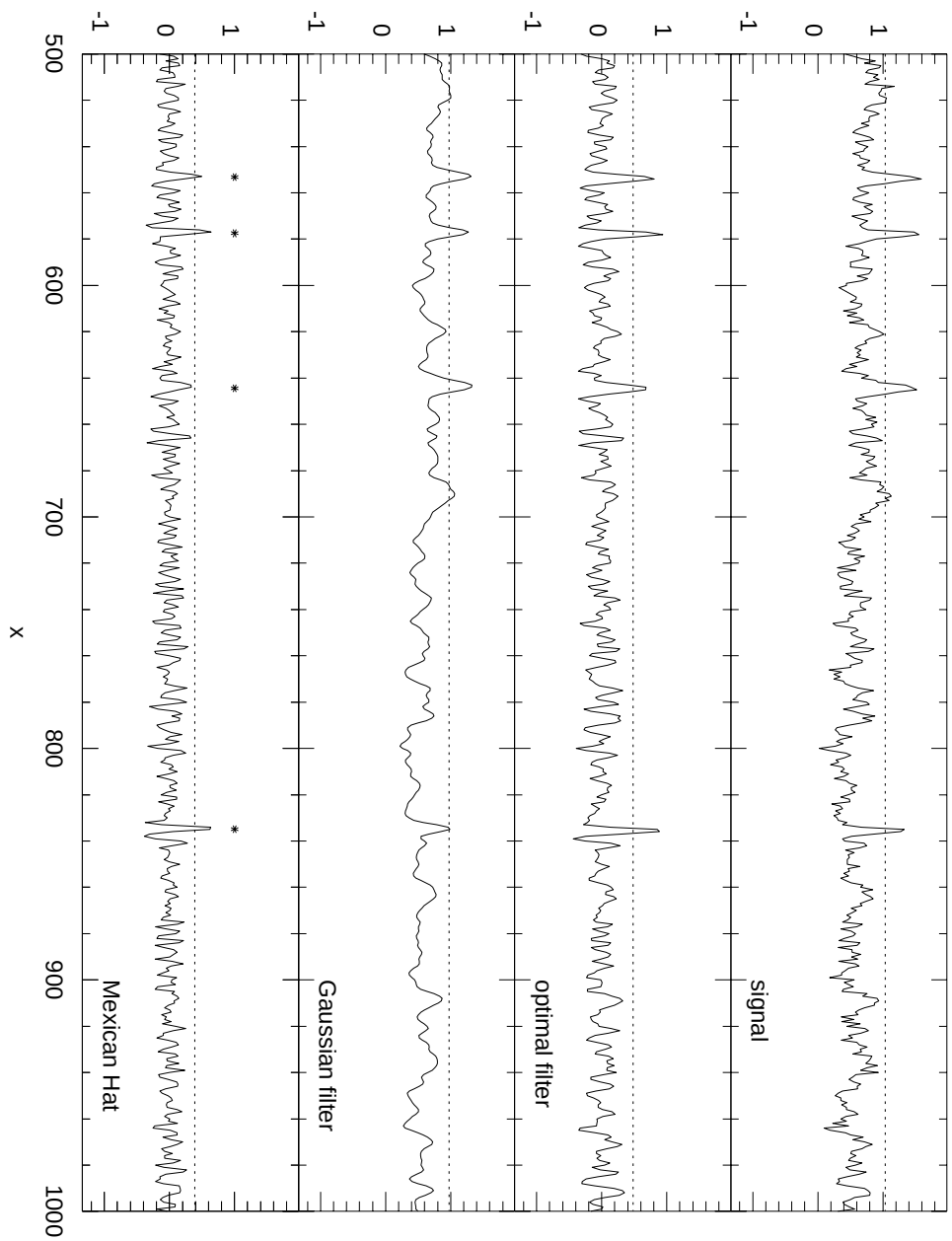

Figure 2.3: Performances of scale-adaptive filter, Gaussian filter and Mexican Hat wavelet. In the top panel a portion of a 32768-pixel one-dimensional simulation is shown. In the simulation there are present Gaussian sources of width $\theta = 3\theta_p$ and $1/f$ noise. The amplitude of the sources is 1 for all of them. The dispersion of the background is set to be 1/3. The second (third) panel from the top shows the map after filtering with the scale-adaptive filter (a Gaussian of width θ). The bottom panel shows the map after filtering with a Mexican Hat of width θ. The positions of the sources are marked with asterisks. In the three cases the 3σ level of the resultant map is shown with an horizontal dotted line.

Chapter 3

Scale-Adaptive Filters in 2D images

The scale-adaptive filters introduced in Chapter 2 (see eq. 2.10) have been developed for the n-dimensional case, being n any integer. The most interesting cases are $n = 1$ (spectra, time-ordered data, etc.), $n = 2$ (images) and $n = 3$ (volumes). The unidimensional case was chosen as a straightforward example. In most cases, however, astronomers have to deal with two-dimensional data sets, such as photographic plates, CCD images and an assorted set of detection devices across the whole electromagnetic spectrum range, including microwave, radio and X-ray maps. Examples of objects of astronomical interest present in these images include point-like sources (such as stars, QSOs and galaxy nuclei) and extended objects (such as galaxies, galaxy clusters, globular clusters, etc.) that in many cases can be well approximated by more or less symmetric profiles (Gaussian profiles for point sources seen through Gaussian beams, exponential disks in spiral galaxies, β-models for galaxy clusters, etc.). In this chapter an application of scale-adaptive filters to one of the most important cases above mentioned: the detection of galaxy clusters in two-dimensional images. In particular, we will consider the detection of the well-known β profiles of galaxy clusters, that are broadly used in X-ray and microwave Astronomy. By 'detection' we mean the determination of the position of the sources as well as the estimation of parameters such as the intensity at the central pixel (*amplitude*, hereafter) and the characteristic scale of each source. Given the radial profile of the cluster, knowing the amplitude and the scale is

equivalent to knowing the total flux at the observation frequency.

Given several images at different frequency channels, it is possible to use the knowledge about the frequency dependence of the different components as well as their statistical properties to separate them. Wiener filtering (WF, Tegmark and Efstathiou 1996; Bouchet et al. 1999) and Maximum Entropy Method (MEM, Hobson et al. 1998, 1999) are powerful tools for component separation based on this idea. When the frequency dependence of the sources is not known the previous methods are inefficient. Moreover, if only a single image is provided, the knowledge of the ν-dependence of the different components is of no use. In that case, only spatial properties (such as characteristic scale, profile, structure, distribution and other statistical properties) might be used to perform the separation. In the case we are going to consider in this work, we intend to detect localized sources (that is, with a 'small' characteristic scale) with spherical symmetry and a given radial profile. All the other components in the image will be considered as a background ('noise') that should be removed in order to optimize the detection of the signal.

Wavelet formalism is well suited to deal with localized signals. The localized bases used in wavelet analysis allow one to obtain a precise representation of local objects both in spatial and frequency domains. In the context of signal processing, a wavelet can be considered as a kind of band-pass filter, capable of select a finite range of frequencies (scales) of an image. The Mexican Hat wavelet (MHW) has been successfully applied to simulated X-ray images (Damiani et al. 1997, Valtchanov et al. 2001, Freeman et al. 2002) in order to detect X-ray sources, as well as to detect and extract point sources from simulated microwave maps (Cayón et al. 2000, Vielva et al. 2001a and 2001b).

As noted in Chapter 2, we may wonder if the Mexican Hat wavelet is the best possible choice in every case or if, on the contrary, there are different families of functions, wavelets or not, which are better suited to each particular case. It is clear that such a filter should take into account the shape (profile) of the source, its characteristic scale and the statistical properties of the background in which is embedded. The MHW is nearly optimal to detect point sources convolved with a Gaussian beam because of the relation between the MHW and the Laplacian of the Gaussian. If the source profile is other than Gaussian, however, the optimal filter would be different. In this chapter optimal scale-adaptive filters will be applied to

the separation and detection of compact sources with a *multiquadric* profile.

3.1 Scale-Adaptive filter for multiquadric profiles

In some astrophysical/cosmological applications the source is modeled by a multiquadric, i. e. the profile is given by

$$\tau(x) = \frac{1}{\left[1 + (\frac{x}{r_c})^2\right]^{\lambda}}, \quad \lambda \geq 1/2. \tag{3.1}$$

Typical examples are, for the 2D case, the emissions in the microwave and X-ray bands with $\lambda = \frac{3\beta-1}{2}, \frac{6\beta-1}{2}$, respectively, for a β-profile for the electron number density $n_e(r) \propto [1 + (r/r_c)^2]^{-\frac{3}{2}\beta}$. Assuming the standard value $\beta = 2/3$ one trivially obtains $\lambda = 1/2, 3/2$ for the microwave and X-ray emissions, respectively.

Assuming a scale-free power spectrum $P(q) = Dq^{-\gamma}$, the equations (2.10) and (2.11) lead to the filter:

$$\tilde{\psi}(q) = \frac{1}{\alpha a'} \frac{\Gamma(\lambda)}{2^{1-\lambda}(\gamma + n)} (qr_c)^{\gamma+\lambda-\frac{n}{2}} \left[P\, K_{\lambda-\frac{n}{2}}(qr_c) + Q\, qr_c K_{1+\lambda-\frac{n}{2}}(qr_c) \right], \tag{3.2}$$

$$P \equiv 2\gamma - (n-\gamma)(\gamma + 2\lambda), \quad Q \equiv 2(n-\gamma)\frac{\gamma + 2\lambda + 1}{\gamma + 4\lambda - n}, \tag{3.3}$$

$$a' \equiv \frac{2^{\gamma+2\lambda-3}}{\Gamma(\gamma + 2\lambda)} \Gamma\left(\frac{\gamma + n}{2}\right) \Gamma^2\left(\frac{\gamma + 2\lambda}{2}\right) \Gamma\left(\frac{\gamma + 4\lambda - n}{2}\right). \tag{3.4}$$

In the previous equations Γ denotes the Gamma function whereas K denotes the Bessel K function. Table 3.1 gives the analytical form of the scale-adaptive filter for $\lambda = 1/2, 3/2$ and different values of the spectral index $\gamma = 0, 1, 2, 3$ on Fourier space and real space. The 2D scale-adaptive filters for the same parameters are shown in the left side of figure 3.1.

3.2 Other filters

For comparison with the scale-adaptive filters, we shall briefly review other two filters that have been extensively used in the literature: the Mexican Hat wavelet

λ	γ	$\tilde{\Psi}_a(q)$	$\Psi_a(x)$
1/2	0	$\frac{2}{\pi}e^{-y}$	$\frac{2}{\pi r_c^2}\frac{1}{(1+x^2)^{3/2}}$
1/2	1	$\frac{2}{\pi}ye^{-y}$	$\frac{2}{\pi r_c^2}\frac{2-x^2}{(1+x^2)^{5/2}}$
1/2	2	$\frac{2}{\pi}ye^{-y}$	$\frac{2}{\pi r_c^2}\frac{2-x^2}{(1+x^2)^{5/2}}$
1/2	3	$\frac{2}{\pi}(3-y)y^2e^{-y}$	$\frac{4}{\pi r_c^2}\frac{-12x^4+21x^2-2}{(1+x^2)^{9/2}}$
3/2	0	$\frac{2}{\pi}(2y-1)e^{-y}$	$\frac{6}{\pi r_c^2}\frac{1-x^2}{(1+x^2)^{5/2}}$
3/2	1	$\frac{4}{3\pi}y^2e^{-y}$	$\frac{4}{3\pi r_c^2}\frac{2-3x^2}{(1+x^2)^{7/2}}$
3/2	2	$\frac{4}{3\pi}y^2e^{-y}$	$\frac{4}{3\pi r_c^2}\frac{2-3x^2}{(1+x^2)^{7/2}}$
3/2	3	$\frac{4}{15\pi}(5-y)y^3e^{-y}$	$\frac{4}{3\pi r_c^2}\frac{9x^6-84x^4+88x^2-2}{(1+x^2)^{11/2}}$

Table 3.1: Optimal scale-adaptive filters associated to the $\beta = 2/3$ profile for the cases $\tau(r) = \frac{1}{(1+(r/rc)^2)^\lambda}$, $\lambda = 1/2$ (microwave) and $\lambda = 3/2$ (X-ray). Col. (1): λ . Col. (2): Background (noise) exponent γ. Col (3): Filter in Fourier space ($y = qr_c$). Col. (4): Filter in real space ($x = r/r_c$).

and the 'matched' filter (MF). Both of them have already been introduced in previous chapters of this thesis; they have been used in 2D problems even more that in one-dimensional data sets. Here the main facts concerning astronomical images and these filters are summarized.

3.2.1 The Mexican Hat wavelet

The well-known n-dimensional Mexican Hat wavelet is defined by

$$\Psi(\vec{x}; R, \vec{b}) = \frac{1}{R^n}\psi(\frac{|\vec{x}-\vec{b}|}{R}), \quad \psi(x) \propto (n-x^2)e^{-x^2/2} \tag{3.5}$$

$$\psi(q) \propto q^2 e^{-\frac{1}{2}q^2}. \tag{3.6}$$

This type of wavelet has been extensively used for point source detection. Optical images of galaxy fields have been analyzed to detect voids and high-density structures in the first CfA redshift survey slice (Slezak et al. 1993). Microwave images have been analyzed (Cayón et al. 2000; Vielva et al. 2001a) and combined with the maximum entropy method (Vielva et al. 2001b) to obtain catalogs of point sources from simulated maps at different frequencies that will be observed by the future Planck mission. On the other hand, the MHW has also been used to detect

X-ray sources (Damiani et al. 1997) and presently for the on-going XMM-Newton mission (Valtchanov et al. 2001) and Chandra (Freeman et al. 2001).

3.2.2 The matched filter

If one removes condition (3) defining the scale-adaptive filter in § 2.2 another type of filter can be found after minimization of the variance:

$$\tilde{\psi}_m(q) = \frac{1}{\alpha a}\frac{\tau(q)}{P(q)}. \tag{3.7}$$

This is usually called a *matched* filter (see § 1.2.4 and Appendix A). In general, the *matched* and *scale-adaptive* filters are different. In the former case, one obtains a slightly larger gain (although for $\gamma \geq 0.5$ the gain is the same from the practical point of view, i.e. the relative difference is less than 20%) but sources must be identified 'a posteriori' with an extra criterion whereas the scale-adaptive filter allows one to get the sources in a straightforward manner. Thus, from the methodological point of view and reliability (detection of spurious sources) is better to use the scale-adaptive filter unless the image is completely dominated by white noise.

Matched filters have been used recently to detect clusters of galaxies from optical imaging data (Postman et al. 1996; Kawasaki et al. 1998). In this approach the method uses galaxy positions, magnitudes and photometric/spectroscopic redshifts if available to find clusters and determine their redshift.

For the case of a source profile given by eq. (3.1) and a scale-free power spectrum given by $P(q) \propto q^{-\gamma}$, the previous formula (3.7) leads to the following matched filter:

$$\tilde{\psi}_m(q) = \frac{1}{\alpha a'}\frac{\Gamma(\lambda)}{2^{1-\lambda}}(qr_c)^{\gamma+\lambda-\frac{n}{2}}K_{\lambda-\frac{n}{2}}(qr_c), \tag{3.8}$$

with a' given by eq. (3.4). For $\lambda = 1/2$ the matched filter is not defined for white noise (i.e. $\gamma = 0$), whereas for other values of γ is given by table 3.2 in Fourier space and real space, respectively. In the case of white noise, we can use the modified profile

$$\tau(x) = N\left[\frac{1}{\sqrt{r_c^2+x^2}} - \frac{1}{\sqrt{r_v^2+x^2}}\right],$$

λ	γ	$\tilde{\Psi}_o(q)$	$\Psi_o(x)$
1/2	0	$B\frac{e^{-qr_c}-e^{-qr_v}}{q}$	$\frac{B}{N}\tau(x)$
1/2	1	$\frac{1}{\pi}e^{-qr_c}$	$\frac{1}{\pi r_c^2}\frac{1}{(1+x^2)^{3/2}}$
1/2	2	$\frac{2}{\pi}(qr_c)e^{-qr_c}$	$\frac{2}{\pi r_c^2}\frac{2-x^2}{(1+x^2)^{5/2}}$
1/2	3	$\frac{2}{\pi}(qr_c)^2e^{-qr_c}$	$\frac{2}{\pi r_c^2}\frac{2-3x^2}{(1+x^2)^{7/2}}$
3/2	0	$\frac{2}{\pi}e^{-qr_c}$	$\frac{2}{\pi r_c^2}\frac{1}{(1+x^2)^{3/2}}$
3/2	1	$\frac{2}{\pi}(qr_c)e^{-qr_c}$	$\frac{2}{\pi r_c^2}\frac{2-x^2}{(1+x^2)^{3/2}}$
3/2	2	$\frac{4}{3\pi}(qr_c)^2e^{-qr_c}$	$\frac{4}{3\pi r_c^2}\frac{2-3x^2}{(1+x^2)^{3/2}}$
3/2	3	$\frac{2}{3\pi}(qr_c)^3e^{-qr_c}$	$\frac{2}{\pi r_c^2}\frac{3x^4-24x^2+8}{(1+x^2)^{3/2}}$

Table 3.2: Matched filters associated to the $\beta = 2/3$ profile for the cases $\tau(r) = \frac{1}{(1+(r/rc)^2)^\lambda}$, $\lambda = 1/2$ (microwave) and $\lambda = 3/2$ (X-ray). Col. (1): λ . Col. (2): Background (noise) exponent γ. Col (3): Filter in Fourier space ($y = qr_c$). Col. (4): Filter in real space ($x = r/r_c$). For the case $\lambda = 1/2$ and $\gamma = 0$ the modified profile (3.9) has been used.

$$N = \frac{r_c}{1 - r_c/r_v}, \quad B = \frac{1}{4\pi N \ln\left[\frac{r_c+r_v}{2\sqrt{r_c r_v}}\right]}, \tag{3.9}$$

where r_v is a cut-off radius. The behavior of this profile is: $\tau(0) = 1$ and $\tau(x) \propto x^{-3}$ for $x \gg r_v$.

The matched filters appear in the right side of figure 3.1. Note that the scale-adaptive filters reach their peak around $qr_c \sim 1$, whereas there is more dispersion in the position of the peak of the matched filters. This means that scale-adaptive filters put more emphasis around the characteristic scale of the sources. We remark that for $\gamma = n$ the scale-adaptive filter and the matched filter coincide.

3.3 Numerical simulations: application and results

In order to show the performance of the scale-adaptive filters we have simulated realizations of multiquadric profiles embedded in backgrounds of the type $P(q) = Dq^{-\gamma}$. For the sake of simplicity we simulate ideal multiquadric profiles, not taking into account the effect of the detector beam. However, the method can

be generalized to include this effect by modifying the input source profile. We also consider only integer values of γ, so we can directly use the filters from table 3.1. In a more realistic case, where the power spectrum does not follow such a simple law, it could be directly estimated from the data. Then, the filter is numerically calculated.

3.3.1 Microwave emission and the SZ-effect of clusters

One of the most promising applications of scale-adaptive filters is the detection and extraction of the emission of galaxy clusters due to the Sunyaev-Zel'dovich (SZ) effect in microwave maps. Maps of the Cosmic Microwave Background (CMB) contain contributions from a variety of foregrounds (the SZ effect among them) and different types of noise. If we approximate the power spectrum of a typical CMB map by the law $P(q) = Dq^{-\gamma}$, the effective index γ ranges between values near 0 (for regions dominated by white noise) to ~ 3 (for regions dominated by dust emission). As an example, we first simulated a 512×512 pixel field containing 100 randomly distributed 'clusters' with a multiquadric profile ($\lambda = 1/2$). All the 'clusters' have the same scale ($r_c = 1.0$ pixel) and amplitudes distributed between 0.1 and 1 (in arbitrary units). With a convenient rescaling of the amplitude, this could simulate, for example, a $12^\circ.8 \times 12^\circ.8$ field of the sky filled with clusters of several arcmin of extent. The simulated clusters are shown in the bottom panel of figure 3.2. A $P(q) = Dq^{-3}$ background was added so that the peaks of sources are *on average* at the 2σ level of the final map. The map containing the sources and the background ('noise') is shown in the center of figure 3.2. In the following, we will call this simulation 'Simulation I'.

Following table 3.1, the scale-adaptive filter for the case $\lambda = 1/2$, $\gamma = 3$ is $\tilde{\Psi}_o(q) = 4(3 - y)y^2e^{-y}/3\pi$, $y = qr_c$. The result of applying the scale-adaptive filter on the simulated field is shown in the top panel of figure 3.2. For the sake of clarity, only the pixels above the 3σ level have been plotted. After filtering, the detection is performed by looking for peaks above a certain threshold. Knowing the initial position and amplitude of the simulated sources, it is possible to determine quantities such as the mean error in the estimation of the source parameters as well as the reliability of the filter (that is, the probability of detecting spurious

'sources'). The results are summarized in table 3.3. The first column in table 3.3 shows the σ threshold used for the detection. The second column shows the number of true detections found. The third column indicates the number of spurious 'sources' detected. The fact that above 4σ we find more spurious detections than above the 3σ threshold may seem surprising. This can be easily explained taking into account that our detection method looks for sets of connected pixels over a given threshold. Therefore, a 'lump' with two peaks may be seen as a single source if the detection threshold is low enough. A higher detection threshold allows to split the 'lump' into separate peaks. The fourth column in table 3.3 indicates the mean error in the determination of the position of the source in pixel units. As can be seen, all the detected sources were correctly located. The mean relative bias given in the fifth column of table 3.3 is defined as $\bar{b}_A(\%) = 100 \times \frac{1}{N} \sum (A_i - A_{o_i})/A_{o_i}$, where A_{o_i} and A_i are the original and estimated amplitudes of the sources respectively and N is the number of considered sources. The mean relative error given in the sixth column is defined as $\bar{e}_A = 100 \times \frac{1}{N} \sum \mid A_i - A_{o_i} \mid /A_{o_i}$. The seventh column gives the mean gain, as defined in eq. (2.20). The most striking fact in this case is the high gain (~ 4.4) that is found in all the cases. This is due to the fact that the scale-adaptive filter (which acts as a band-pass filter) is very efficient in removing the large scale structure that characterizes the q^{-3} background. The individual gain for each detection as a function of the true amplitude of the source is shown in the bottom panel of figure 3.3. We see that there is a certain dispersion around the average value (represented in figure 3.3 as an horizontal dotted line) but, in general, the gain is nearly independent of the original amplitude of the source. Only for the brightest and the faintest sources this independence is lost. Bright sources tend to have lower gains than the average and vice versa. This is related to what is seen in the top panel of figure 3.3, which shows the estimated amplitudes of the detected sources versus the original amplitudes. Between the original and estimated amplitudes there is a strong linear regression but with a small positive deviation from the $y = x$ law. This corresponds to the $\sim 7\%$ relative bias that is given in table 3.3. Again, both bright and faint sources deviate from this law.

Figure 3.3 can be explained considering the detection strategy we use. We look for peaks above a certain threshold, that is, for local maxima in the filtered image. Therefore the detection method is biased to give higher amplitudes (it selects the

σ	detected	spurious	mean offset	b_A	$\bar{e}_A$	$\bar{g}$
3.0	80	14	0.0	7.7	10.7	4.4
4.0	71	17	0.0	6.6	9.5	4.4
5.0	59	12	0.0	7.2	9.7	4.4

Table 3.3: Detections in Simulation I. Col. (1): σ detection level. Col. (2): Number of true detections above the σ threshold. Col (3): Number of spurious sources. Col. (4): Mean offset in the position of the source (pixels). Col. (5): Mean relative bias in the determination of the amplitude (%), defined as the average of $100 \times \frac{A-A_o}{A_o}$. Col. (6): Mean relative error in the determination of the amplitude (%), defined as the average of $100 \times \frac{|A-A_o|}{A_o}$. Col. (7): Mean gain (as defined in eq. 2.20)

pixels in which the contribution of the source plus the residual noise is maximum). This effect has more relevance when the source is faint and less when the source is bright. Indeed, the gain of the detected faint point sources is larger than average, since otherwise they would be below the detection level. As A_o increases in the top panel of figure 3.3, the bias decreases and eventually drops to negative values. This suggests the existence of a systematic effect that gives a negative bias and that is compensated by the positive bias produced by the detection method when the sources are weak enough. Such negative bias has been reported in relation with scale-adaptive filters in Sanz et al 2000 and Herranz et al. 2001a, and is due to the fact that the normalization given by condition (1) of scale-adaptive filters is calculated for an ideal profile on an infinite, continuous field (the limits of the integral are 0 and ∞), while the real images are finite and discrete (pixelized).

In spite of these small systematics, the performance of the optimal filter is very good in the sense that with a simple application of a filter and a thresholding detection scheme we are able to recover a very significant number of sources and estimate their amplitudes with errors not larger than $\sim 10\%$. The reliability of the method is such that over a 3σ threshold there are a 15% of spurious detections.

In a more realistic case the scale r_c of the sources will not be known *a priori*. For example, in a CMB map, clusters of different scales will be present, going from almost point-like sources to large structures that extend across several pixels. Under these conditions, it does not seem clear which one should be the scale of the

scale-adaptive filter. Fortunately, condition (3) of the optimal scale-adaptive filters implies that the coefficient at the position of the source and at the 'right' scale is a maximum. Therefore, the strategy to follow is to filter the image with a set of N filters with different 'tentative' scales r_{c_i}, $i = 1, ..., N$. For a given detection, the maximum value among the coefficients at the position of the source of the different filtered maps will correspond to the r_{c_i} closer to the r_c of the cluster.

To test this point we performed a new simulation (Simulation II, hereafter) with 100 'clusters' with r_c distributed between 0.5 and 2.0 pixels and amplitudes between 0.1 and 1.0 (in arbitrary units). The noise is similar to that of Simulation I. The simulated map was filtered with scale-adaptive filters with the parameter $r_c =$ 0.25, 0.50, 0.75, 1.00, 1.25, 1.50, 1.75, 2.00, 2.25 and 2.50 pixels. In figure 3.4 are plotted the original clusters (bottom panel), the simulated map with noise (central panel) and the map filtered with the scale-adaptive filter $r_c = 1.50$ (top panel). The sources are detected over each filtered map by selecting peaks above a certain threshold. If one of such peaks is detected at the same position on different filtered maps, it is unlikely that it corresponds to pixel-scale noise (although it could be a fluctuation of the noise at a scale of the order of the source scale). Looking for detections in several filtered maps will help to reduce the number of spurious sources in the output. For every detection that is present in several filtered maps we look for the maximum of the coefficients at the central position. In particular, we only consider sources that appear at least at 5 scales. This number was chosen because with 5 consecutive scales is possible to 'cover' the pixel size in the scale space and because it gives a good compromise between the number of detections and the number of spurious detections, as will be seen through this section. The maximum gives both the scale of the source and its amplitude.

The results are shown in table 3.4. The main difference between tables 3.4 and 3.3 is the lower gains that are obtained in Simulation II. This result is expected because the set of filters used only fit clusters with some specific values of r_c. For the rest of clusters the filters are only approximately optimal. Moreover, any error in the determination of r_c will lead to a wrong determination of the amplitude. The top panel of figure 3.5 shows the estimated r_c versus the original r_{c_o} in Simulation II. There is a significant dispersion around the $r_c = r_{c_o}$ line (represented by a dotted line in the top panel of figure 3.5). As shown in table 3.4, the mean error in the

σ	detected	spurious	mean offset	b_A	$\bar{e}_A$	b_{r_c}	$\bar{e}_{r_c}$	$\bar{g}$
3.0	57	0	0.0	11.5	14.0	-0.12	0.14	3.7
4.0	49	0	0.0	7.3	10.4	-0.13	0.14	3.5
5.0	38	0	0.0	6.1	8.5	-0.12	0.13	3.5

Table 3.4: Detections in Simulation II. Col. (1): σ detection level. Col. (2): Number of true detections above the σ threshold. Col (3): Number of spurious sources. Col. (4): Mean offset in the position of the source (pixels). Col. (5): Mean relative bias in the determination of the amplitude (%), defined as the average of $100 \times \frac{A_i - A_{o_i}}{A_{o_i}}$. Col. (6): Mean relative error in the determination of the amplitude (%), defined as the average of $100 \times \frac{|A_i - A_{o_i}|}{A_{o_i}}$. Col. (7): Mean bias in the determination of r_c (in pixel units), defined as the average of $\frac{r_{c_i} - r_{c_{o_i}}}{r_{c_{o_i}}}$. Col. (8): Mean absolute error in the determination of r_c (pin pixel units), defined as the average of $\frac{|r_{c_i} - r_{c_{o_i}}|}{r_{c_{o_i}}}$. Col. (9): Mean gain (as defined in eq. 2.20).

determination of r_c is ~ 0.15 pixels, with a similar bias towards estimating higher values of r_c than the real ones. If we increase the 'resolution' in r_c by increasing the number of filters the error is not significantly reduced. This indicates that it is not possible to determine scale parameters with a much better resolution than the pixel scale. In spite of this limitation, the errors in the determination of the amplitude remain reasonable ($\sim 10\%$). This indicates that the estimated amplitude does not vary significantly with r_c in the neighborhood of the real value of the core radius. Therefore, errors under the pixel scale in the estimation of r_c have little effect in the determination of the amplitude. Lower panel of figure 3.5 shows the estimation of the amplitudes in Simulation II. As in Simulation I, a small positive bias is found. The gain shows the same behavior with respect to the amplitude of the sources as in Simulation I (an asymptotic decrease of the gain with the amplitude of the clusters) and is almost independent of the size of the clusters. Small clusters tend to have slightly higher gains than large clusters. This could be due to the fact that small clusters are detected in maps filtered with small r_c parameter and therefore in these maps the pixel-scale residuals are worse removed than in maps filtered with a large r_c. Maps filtered with small r_c show more contribution from pixel-scale noise residuals and therefore the peak amplitude

is more likely to be overestimated. The number of spurious detections drops to 0 due to the fact that we discard those 'candidates' that are not detected in at least 5 filtered maps. Indeed a similar constraint is imposed in the MHW method used by Vielva et al. 2001a and 2001b, who performed a multiscale fit in order to estimate the amplitude of point sources as well as to discard spurious detections. Unfortunately, this constraint also reduces the number of true detections. This is an example of a well-known and unavoidable problem in detection theory: the cost of reducing spurious detections is to reduce the number of true detections and, conversely, relaxing the selection criteria to include more true sources leads to an increase of spurious detections.

3.3.2 X-ray emission and clusters

Other straightforward application of scale-adaptive filters is in the field of X-ray astronomy. X-ray emission from galaxy clusters roughly follows a multiquadric profile with $\lambda = 3/2$. Unfortunately, real X-ray images suffer from strongly non-isotropic point spread functions that distort cluster profiles to quite odd and asymmetric shapes. For this work, however, we will consider (as we did in the previous case of SZ emission) that instrumental effects have been somehow corrected and that we only have ideal clusters and noise. For Simulation III we distributed 100 ideal multiquadric profiles with $\lambda = 3/2$, r_c between 2.0 and 4.5 pixels and amplitudes between 0.1 and 1.0 (in arbitrary units). The dominating background in X-ray images is Poissonian shot-noise that, when the exposure time is long enough, can be approximated by white noise. We added white noise ($\gamma = 0$) to our simulation so that, given the amplitudes of the clusters (in their arbitrary units), the final noise level is roughly similar to the one of an XMM image of 95 ks of exposure time. The final signal to noise ratio is greater than in Simulations I and II. Figure 3.6 shows the simulated clusters (bottom panel), the complete simulation with noise (central panel) and the coefficient map that corresponds to the simulation filtered with an scale-adaptive filter with $r_c = 3.0$ pixels (top panel).

The detection and determination of the amplitude and the scale of the clusters were performed following the same steps than in Simulation II. The set of chosen 'core radii' was $r_c = 1.2, 1.3, 1.5, ..., 5.1$ pixels ($N = 40$ filters). To consider a

σ	detected	spurious	mean offset	b_A	$\bar{e}_A$	b_{r_c}	$\bar{e}_{r_c}$	$\bar{g}$
3.0	73	22	0.4	10.1	17.8	0.06	0.40	2.0
4.0	63	10	0.3	6.2	14.9	0.03	0.39	2.0
5.0	48	2	0.2	6.9	12.6	0.05	0.36	2.0

Table 3.5: Detections in Simulation III. Col. (1): σ detection level. Col. (2): Number of true detections above the σ threshold. Col (3): Number of spurious sources. Col. (4): Mean offset in the position of the source (pixels). Col. (5): Mean relative bias in the determination of the amplitude (%), defined as the average of $100 \times \frac{A_i - A_{o_i}}{A_{o_i}}$. Col. (6): Mean relative error in the determination of the amplitude (%), defined as the average of $100 \times \frac{|A_i - A_{o_i}|}{A_{o_i}}$. Col. (7): Mean bias in the determination of r_c (in pixel units), defined as the average of $\frac{r_{c_i} - r_{c_{o_i}}}{r_{c_{o_i}}}$. Col. (8): Mean absolute error in the determination of r_c (in pixel units), defined as the average of $\frac{|r_{c_i} - r_{c_{o_i}}|}{r_{c_{o_i}}}$. Col. (9): Mean gain (as defined in eq. 2.20).

detection as a source it must be present in at least ten of the filtered maps (we chose this number for the same reasons that were explained in Simulation II). The results are shown in table 3.5. Simulation III has two qualitative differences with respect to Simulations I and II. First of all, the cluster profile drops much faster in the X-ray case than in the SZ case. Second, the simulated background is white noise instead of $\propto q^{-3}$. This is not the optimal situation since the gain achieved by the scale-adaptive filter is not very high in comparison with other filters when the background is dominated by white noise. In particular, the gains obtained in Simulation III are only ~ 2 (table 3.5). Moreover, the fast decline of the cluster profile makes them much more compact (that is, more point-like), making more difficult to distinguish them from noise fluctuations. In other words, comparatively to a microwave image of the same pixel scale and in which there are clusters of the same r_c, it is harder to estimate parameters such as r_c because the signal is condensed in a few pixels. If, for example, r_c was of the order or less than half a pixel, the cluster would be almost indistinguishable from a point source. This explains why, in spite of having the r_c space quite densely sampled, the mean error in the determination of that parameter is almost of 0.5 pixels. There is no observed bias in the determination of r_c. Clusters in Simulation II had smaller core radii

than those in Simulation III, and therefore it was more likely to overestimate than to underestimate their size. In Simulation II, smaller values of r_c were chosen in order to avoid overlapping among the clusters (the compactness of X-ray clusters reduces the probability of overlapping). Besides, the $P(q) \propto q^{-3}$ noise fluctuations grow stronger with their scale, so large-scale residuals (that, when combined with the filtered profile, would lead to overestimation of the scale) were more likely in Simulation II that pixel-scale residuals (that, when combined with the filtered profile, would lead to underestimation of the scale). Under white noise conditions all scales of the background have the same power and therefore there is no reason for bias in any of the two directions, as is observed.

Figure 3.7 shows the performance of scale-adaptive filters in the determination of the amplitude (bottom panel) and r_c (top panel). The dispersion in the estimated r_c is comparatively much greater than the dispersion in the estimated A (lower panel of figure 3.7). The relative errors in the determination of the amplitude are also small ($\sim 15\%$, table 3.5). This indicates that the estimated amplitude is quite stable with respect to variations in r_c, that is, that curves A versus r_c are fairly flat indeed. The observed bias in the determination of the amplitude is around $5 - 10\%$, similar to that observed in Simulations I and II.

Finally, the compactness of the clusters embedded in white noise has the effect of increasing the number of spurious detections, specially for low detection thresholds. At the 5σ level, however, the number of spurious sources is less than 5% of the true detections. Therefore, the scale-adaptive filter is well suited to detect and extract multiquadric profiles even in the less favorable case of more compact clusters and a background dominated by white noise.

3.4 Conclusions

In this chapter the concept of scale-adaptive filter has been applied to a multiquadric profile characterized by two parameters: the core radius r_c and the the decay parameter λ, in order to obtain an unbiased estimation of the amplitude of the source. Explicit analytical formulas on Fourier space and simple analytical ones on real space for some source profiles and backgrounds have been obtained. In particular, the chapter focuses on the interesting cases of microwave and X-

ray emissions. A comparison with other standard filters is done. In particular, for $\gamma = n$ (e. g. 1D signal and $1/f$ noise, 2D image and $P(q) \propto q^{-2}$) the scale-adaptive and matched filters coincide.

Two-dimensional clusters with $\lambda = 1/2$, with different core radius r_c, embedded in a background $P(q) \propto q^{-3}$ that can mimic microwave emission from the cluster plus intrinsic microwave and foreground emissions and noise have been simulated. When the characteristic scale (the 'core radius' r_c) is a priori known it is possible to recover a great number of sources without any significant error in their position and with errors in the determination of the amplitude of $\sim 10\%$. About a 20% of spurious detections are also detected. This percentage of spurious sources could be reduced by introducing a multiscale analysis similar to the one used in the previous section (i.e. imposing that the sources appear in a certain number of maps filtered with different scales). However, this would also reduce the number of true detections. The mean gain in this simulated case is 4.4, meaning that over a 5σ detection threshold in the filtered maps we are able to detect sources that were only at 1.14σ in the original map.

When the characteristic scale is not a priori known, as will happen in a realistic case, the size of the clusters can be estimated by filtering the map with several filters at different scales and looking for the maximum among the coefficients at the position of the sources. Moreover, the multiscale analysis can be used to reject spurious detections (that typically appear only in one or a few of the filtered maps) by imposing that the source 'candidates' must be present in several filtered maps. Unfortunately, the price of removing spurious detections is always to reduce the number of detections. A simulation including clusters of different sizes has been performed. A multiscale analysis with ten filters of different scale has been applied to this simulation. Only sources appearing in five or more filtered maps were considered as true detections. Under these conditions, it was possible to recover 30% less sources than in the first simulation, but the number of spurious detections dropped to 0. The position of the sources was again recovered with no significant error. The mean error in the determination of the amplitude is $< 15\%$. Additionally, we were able to determine the scale parameter of the detected clusters with mean errors of ~ 0.15 pixels.

2D clusters with $\lambda = 3/2$, with different core radius r_c, embedded in a back-

ground $P(q)$ = constant (that can mimic X-ray emission from the cluster plus white noise) have been also simulated. The multiscale analysis allowed again to estimate the position, amplitude and scale of the sources with small errors. The results are slightly worse than in the microwave simulations because the X-ray clusters are more compact and because the background used in this simulation (white noise) is not the most favorable for the scale-adaptive filters. Scale-Adaptive filters take advantage of the fact that in most cases the power spectrum of the noise has a maximum at a scale that is different from the scale of the sources. In the white noise case the power spectrum is constant at all scales. In other words, the scale-adaptive filter produced high gains in the microwave simulations because the background showed strong large-scale features that were removed efficiently by the scale-adaptive filters. In the white noise case, the gain is only $g = 2.0$. In spite of that, the errors in the determination of the fundamental parameters remain small and the number of spurious detections over the 5σ detection threshold is less than 5% of the number of true sources detected.

Therefore, the scale-adaptive filter is well suited to detect and extract multiquadric profiles from images with a variety of backgrounds.

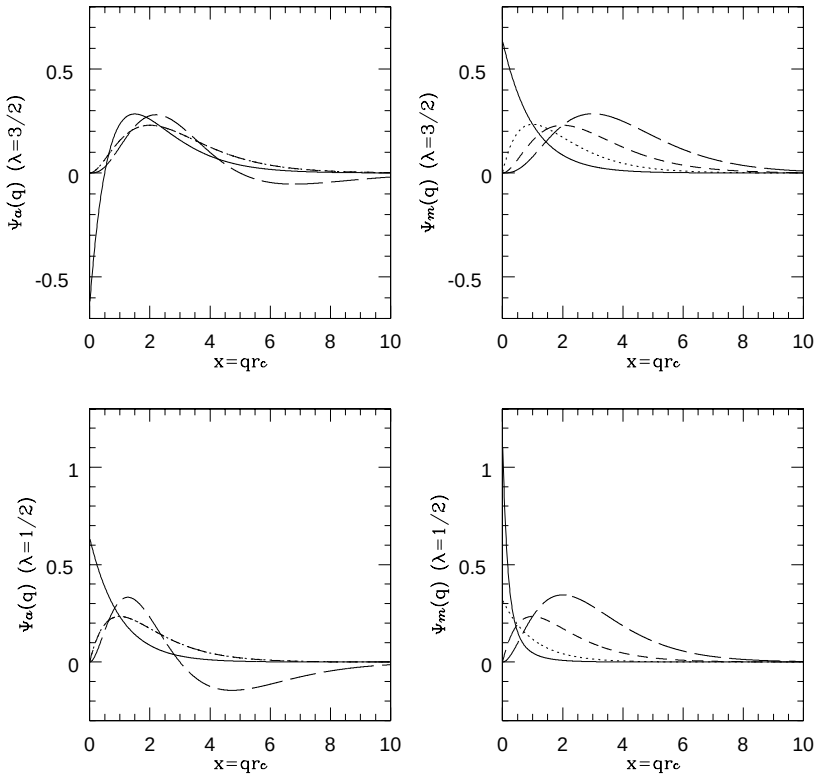

Figure 3.1: Optimal scale-adaptive filters and matched filters (in Fourier space) associated to the $\beta = 2/3$ profile for the cases $\tau(x) = \frac{1}{[1+(x/rc)^2]^\lambda}$, $\lambda = 1/2$ (microwave) and $\lambda = 3/2$ (X-ray) and a homogeneous and isotropic background with power spectrum $P(q) \propto q^{-\gamma}$. The two panels in the left of the figure show the optimal scale-adaptive filters for the cases $\gamma = 0$ (solid line), $\gamma = 1$ (dotted line), $\gamma = 2$ (short-dashed line) and $\gamma = 3$ (long-dashed line). There is a degeneracy for $\gamma = 1$ and $\gamma = 2$ in the two considered cases. The two panels in the right of the figure show the matched filters for the mentioned cases. For the matched filter and the case case $\lambda = 1/2$ and $\gamma = 0$ it has been used the modified profile (3.9).

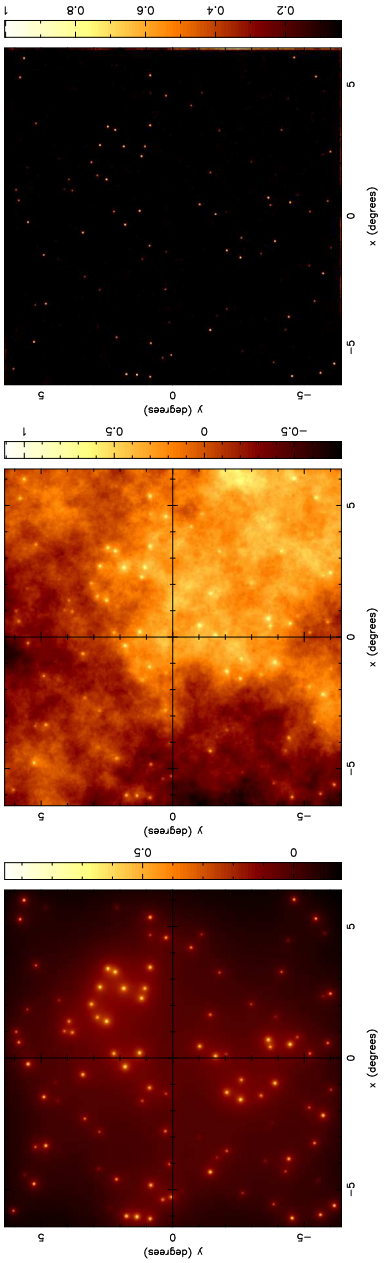

Figure 3.2: Simulation I. Bottom panel shows the simulated 'clusters'. Central panel shows the data to be analyzed ('clusters' plus q^{-3} noise). Top panel shows the coefficient map (after filtering). Only pixels above 3σ have been plotted in the top panel.

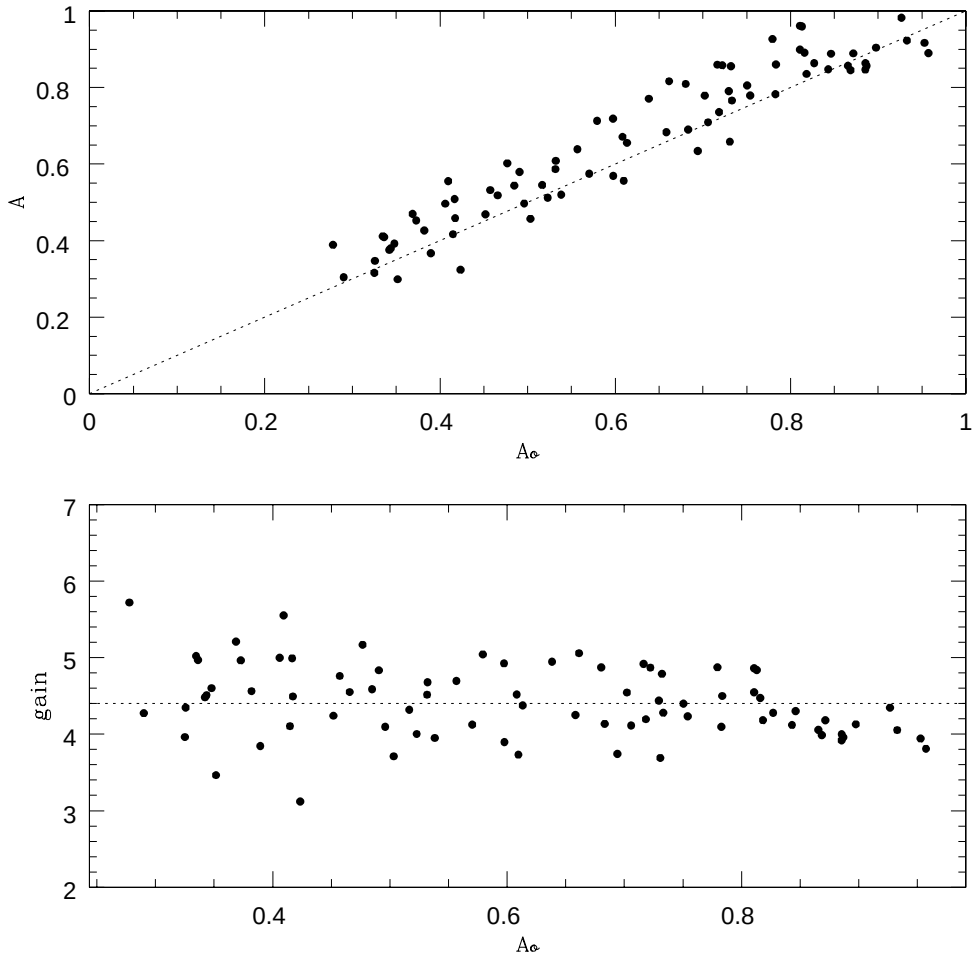

Figure 3.3: Results of the detection (over a 3σ threshold) in simulation 1. The top panel shows the estimated versus the original amplitudes. Points $A = A_o$ are given by a dotted line. The bottom panel shows the gain of the detected sources against the original amplitude. The mean gain is given by the horizontal dotted line.

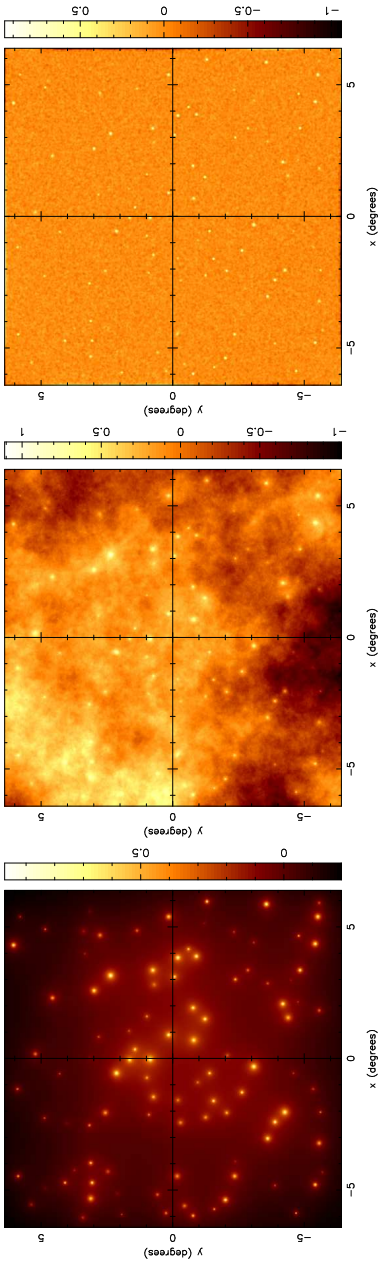

Figure 3.4: Simulation II. Bottom panel shows the simulated 'clusters'. Central panel shows the data to be analyzed ('clusters' plus q^{-3} noise). Top panel shows the coefficient map (after filtering with an scale-adaptive filter of $r_c = 1.50$).

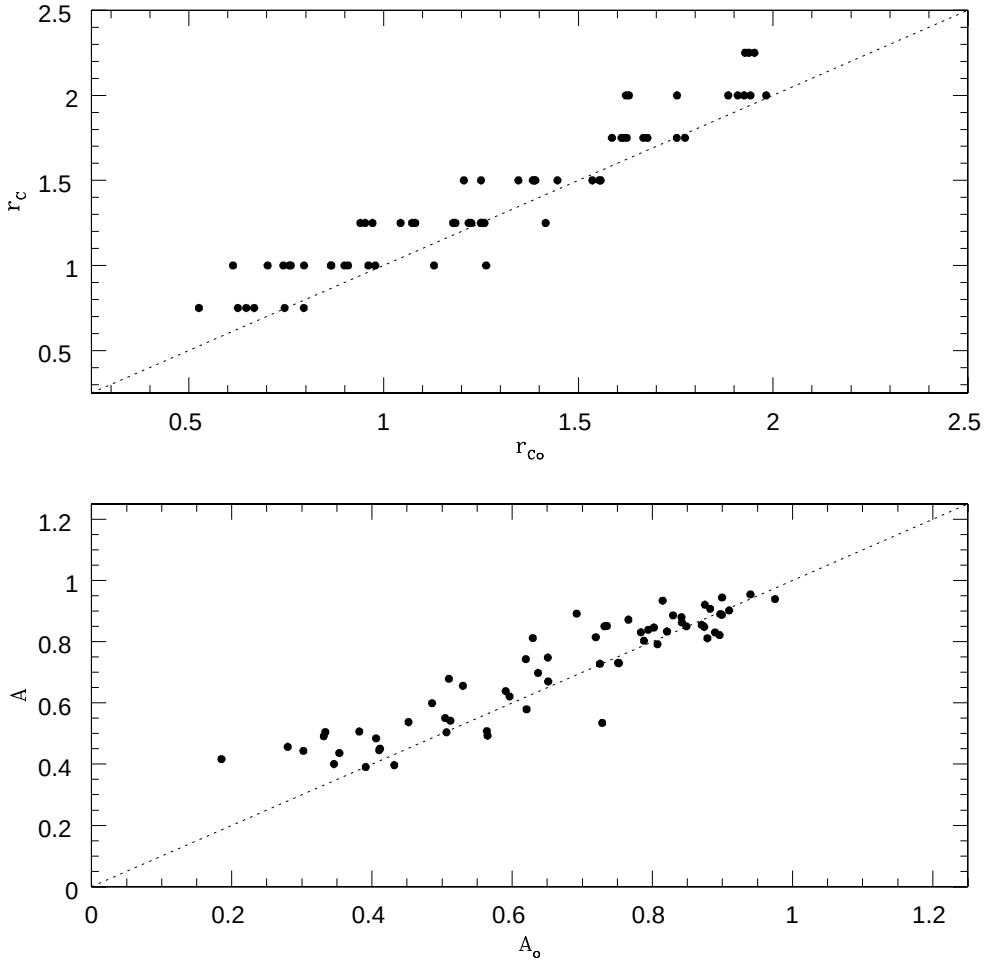

Figure 3.5: Results of the detection (over a 3σ threshold) in Simulation II. The top panel shows the estimated versus the original core radii. The dotted line corresponds to $r_c = r_{c_o}$. The bottom panel shows the estimated versus the original amplitudes. Points $A = A_o$ are given by a dotted line.

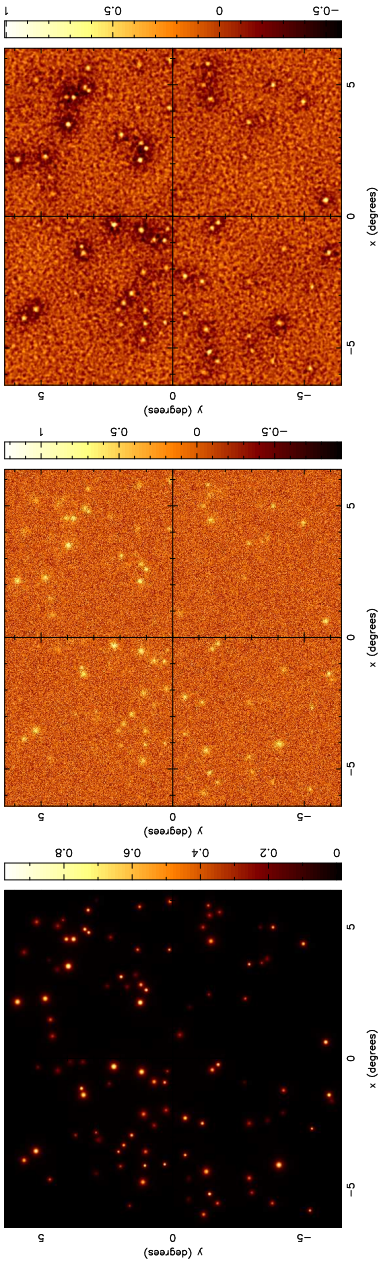

Figure 3.6: Simulation III. Bottom panel shows the simulated 'clusters'. Central panel shows the data to be analyzed ('clusters' plus q^{-3} noise). Top panel shows the coefficient map (after filtering with an scale-adaptive filter of $r_c = 3.50$).

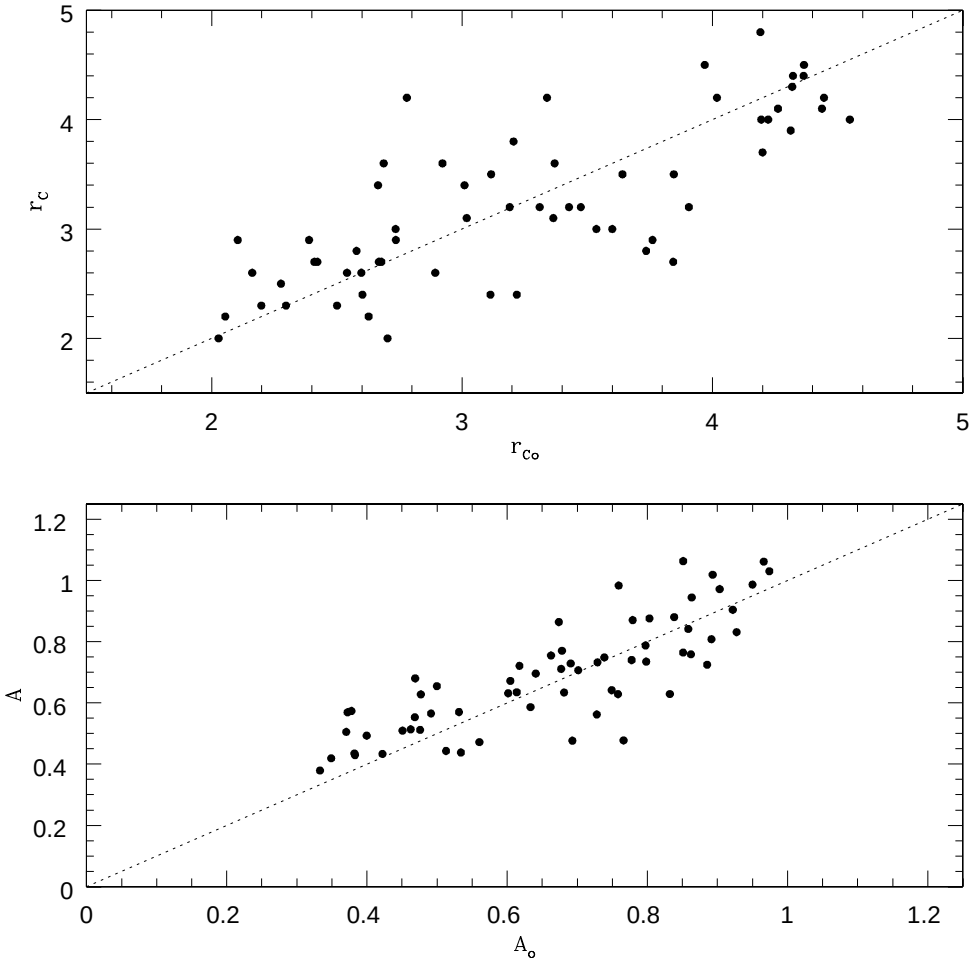

Figure 3.7: Results of the detection (over a 3σ threshold) in Simulation III. The top panel shows the estimated versus the original original core radii. The dotted line corresponds to $r_c = r_{c_o}$. The bottom panel shows the estimated versus the original amplitudes. Points $A = A_o$ are given by a dotted line.

Chapter 4

Point source detection from Planck TOD using Scale-Adaptive Filters

Chapters 2 and 3 have introduced the concept of scale-adaptive filter and tested it on simplistic one and two-dimensional simulations. These simulations were simplistic in the sense that the simulated noises were realizations of pure $P(q) \propto q^{-\gamma}$ spectra; in that case, it was easy to obtain the analytic expression of the filters. Unfortunately, real life is quite more complicated than that. Except for the case of white noise ($P(q) = \text{constant}$), the backgrounds that usually appear in astronomical data sets can hardly be approximated by a single power law. The good news is that, if an estimation of the power spectrum $P(q)$ is available, any computer can easily calculate integrals (2.11). The resulting filter is in a certain way *semi-analytic*, because the profile $\tau(q)$ in eqs. (2.10) and (2.11) is a known mathematical function and the integrals a, b and c are numerically calculated.

This chapter presents an application of scale-adaptive filters to a fully realistic problem related to CMB science: the detection of point sources in time ordered CMB data sets such as the ones the future Planck mission will produce.

4.1 Introduction

One of the most thrilling challenges in the study of the Cosmic Microwave Background (CMB) is to deal with the problem of separating the cosmological signal

from the different foregrounds and noises that appear in CMB experiments. This problem will be specially relevant in future high-resolution experiments such as MAP (Bennett et al. 1996) and Planck (Mandolesi et al. 1998, Puget et al. 1998). From the point of view of determining constraints on fundamental cosmological parameters, the different foregrounds (synchrotron emission, galactic dust, free-free radiation, thermal and kinetic Sunyaev-Zel'dovich (SZ) emission from galaxy clusters and extragalactic point sources) are considered as contaminants and therefore must be removed together with the noise in order to extract the cosmological signal. Besides, knowledge about each of these 'contaminants' has a great scientific relevance on itself. Therefore, it is of exceptional importance to be provided with good techniques of denoising and foreground separation.

Extragalactic point sources are one of the most difficult foreground to separate in CMB maps. This is so because their spatial distribution and abundance remain uncertain and their frequency dependence is not well known. Besides, they can show temporal and even spectral variability. Therefore, methods that take advantage of the different statistical properties of the CMB and the foregrounds (e.g. different angular power spectra), as well as the knowledge of their distinctive frequency dependence, fail in detecting point sources. Such is the case of Wiener filter (Tegmark and Efstathiou 1996, Bouchet and Gispert 1999) and on Maximum Entropy Methods (MEM, Hobson et al. 1998, 1999; Stolyarov et al. 2001). Extragalactic point sources should be removed from the maps before any analysis should be performed. Maximum Entropy Methods can produce a catalogue of point sources as a residual (noise) from the separation process. However, only the faintest sources are recovered, and the brightest ones can still be observed in the residuals. The Independent Component Analysis technique (Baccigalupi et al 2000, Maino et al 2001) has also been applied to this problem with promising results: it has the advantage that, unlike the previous methods, it does not need any prior knowledge of the components to be separated, but its weakest point is actually the separation of point sources.

As noted in chapters 2 and 3, filtering techniques have been successful in both denoising and extracting the brightest point sources from CMB maps. All the point sources in a CMB map will have the same profile and size (the convolution of a δ-Dirac source with the antenna beam) and the filtering process should take

advantage of this fact using a filter optimal for that particular profile and size (scale). This leads to the idea of scale-adaptive filters.

Scale-adaptive filters are unbiased, give a maximum at the position and scale of the source and give the minimum variance of filter coefficients (i.e. are efficient estimators of the amplitude of the sources). They *adapt* themselves to the characteristics of the signal and the noise.

Although two-dimensional maps are the most useful and extended form to show CMB data, they are only available after an exhaustive process of analysis and reduction of the raw data from CMB experiments. These experiments scan different patches of the sky in a sequence producing a unidimensional set of time ordered data (TOD). TODs suffer from lower signal-to-noise ratios than final two-dimensional maps, but on the other hand are less likely to show artifacts coming from the data reduction, such as pixel-to-pixel noise correlations. Moreover, TODs can be analyzed in (almost) real time during the observations in order to produce early (preliminary) catalogs of sources.

4.2 One-dimensional scale-adaptive filter

TODs can be considered as a particular case of a one-dimensional image (a spectrum is another interesting case). The image data values can be expressed as

$$y(t) = s(t) + n(t), \tag{4.1}$$

where t is the time, $s(t)$ represents a symmetric source and $n(t)$ is a homogeneous and isotropic background with mean value $\langle n(t) \rangle = 0$ and characterized by the power spectrum $P(q)$ (q is the absolute value of the 'wave vector' associated to t). If $A = s(0)$ is the amplitude of the source we can introduce the profile $\tau(t)$ as $s(t) = A\tau(t)$.

Following Chapter 2 and Appendix A, it is possible to derive the form of the *scale-adaptive filter* for a given source profile, that is, the filter that satisfies the three conditions:

1. There exists a scale R_o such that at the point source position ($b = 0$) $<$ $w(R, 0) >$ has a maximum at that scale.

2. $< w(R_o, 0) >= s(0)$, i.e. $w(R_o, 0)$ is an unbiased estimator of the amplitude of the source.

3. The variance of $w(R, b)$ has a minimum at the scale R_o, i.e. we have an efficient estimator.

Given these three conditions, the solution (scale-adaptive filter) is found to be:

$$\tilde{\psi}(q) \equiv \psi(R_o q) = \frac{\tau(q)}{2P(q)\Delta}\left[b + c - (a+b)\frac{d\ln\tau}{d\ln q}\right] \tag{4.2}$$

$$a \equiv \int dq \frac{\tau^2}{P}, \quad b \equiv \int dq \frac{\tau}{P}\frac{d\tau}{d\ln q},$$
$$c \equiv \int dq \frac{1}{P}\left[\frac{d\tau}{d\ln q}\right]^2, \quad \Delta = ac - b^2 \tag{4.3}$$

The limits of the integrals go from 0 to ∞. Equations (4.2) and (4.3) are the one-dimensional form of equations (2.10) and (2.11).

In Chapter 2 analytic expressions for Gaussian sources and backgrounds of the type $P(q) \propto q^{-\gamma}$ were derived. In a more realistic case, the background can not be modelled in such a simple way, and integrals in (4.3) should be numerically estimated. When dealing with real data (or realistic simulations such as those used in this work) we must perform the following steps: first, determine the power spectrum of the background directly from the image. Second, evaluate integrals (4.3). Third, build the scale-adaptive filter (4.2) and make the convolution with the data. Finally, we can proceed to detect the sources, for example looking for peaks above a certain σ-level in the coefficient (filtered) image.

When determining the power spectrum of an image we obtain the power spectrum of both the background and the sources together. In the following we consider that the contribution of the point sources to the total power spectrum is negligible. This is a reasonable assumption in a realistic case, specially at medium and high wavelengths where the emission of IR and radio sources is weak. Another problem related to power spectrum determination is the variance of the power spectrum estimator. The variance of one-dimensional power spectrum estimators tends to be larger than that of two-dimensional cases because the number of sample points

is usually smaller. A possible solution to this problem is to estimate low resolution power spectra. However, a large amount of information is lost at scales that are crucial for the determination of the scale-adaptive filters. For this work we chose to average the estimated power spectra of contiguous rings of the TOD (corresponding to regions of the sky separated by only a few arcminutes and therefore possessing very similar underlying power spectra).

A typical profile of the scale-adaptive filter (in Fourier space) for a section of the simulated TOD used in this chapter is shown in figure 4.1. The profile (solid line) is irregular due to the roughness of the estimated power spectrum. These irregularities reflect the particularities of the data and define the scales where the scale-adaptive filter is more or less efficient. For comparison, a Mexican Hat Wavelet (dashed line) and a Gaussian (dotted line), both of them with a width equal to the width of the source, are also shown.

4.3 Data

For this work we chose the 30 GHz LFI28 channel of Planck because of the relatively small size of its resulting TOD. Higher frequencies, in spite of being expected to show more contribution from point sources due to their higher resolution and the contribution of IR sources at $\nu \geq 300$ GHz, lead to huge TOD sets and require many hours of computation to be simulated and analysed on an average computer and will be analyzed in a future work.

The data simulates a 6 months run of the 30 GHz LFI28 channel of Planck, covering the whole sky except for two circles of 1.82° around the ecliptic poles. The sky simulation consist of the addition (in flux) of a CMB simulation and three templates, each for the Galactic synchrotron radiation, Galactic dust and thermal SZ. This sky is then combined with a map of point sources, generated from a catalog performed following the model of Toffolatti et al. (1998) and then 'observed' by the Planck Pipeline Simulator for one of the LFI 30 GHz radiometers; resulting in a TOD for the observed sky (CMB, foregrounds and point sources) plus instrumental noise. The data are ordered in circular rings centered on points situated on the Ecliptic. The angle between the pointing axis and the rotation axis for the LFI28 instrument is 88.18°. The maximum separation between two

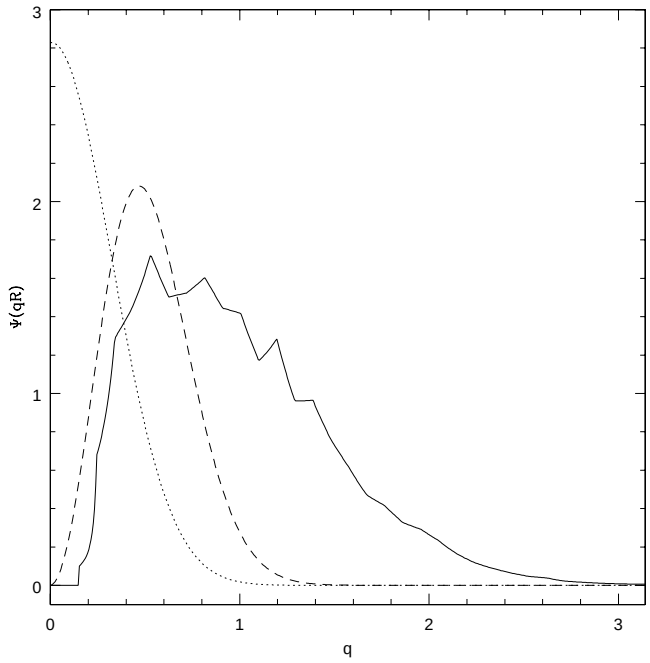

Figure 4.1: Several filters in Fourier space. The scale-adaptive filter particularised for an individual ring of the TOD simulations is represented by a solid line. The dashed line shows a Mexican Hat Wavelet of width equal to the width of the (Gaussian) source. The dotted line represents a Gaussian of width equal to the width of the source. All filters are normalized to give the true amplitude of the source after convolution.

consecutive rings is $2.5'$ (at the intersection with the Ecliptic plane). Each ring results from the average of 60 revolutions of the detector around the rotation axis, corresponding to one hour of integration time, and contains 1950 measures of antenna temperature. There are a total of 4383 of such rings, leading to 8546850 measures of temperature. The antenna has a FWHM of $33'$ and its response slightly

differs from a circular Gaussian one.

The simulation contains CMB emission, different extended foregrounds (Galactic synchrotron, dust and free-free, thermal and kinetic SZ emission from clusters, etc), point sources and instrumental noise. Both white noise and $1/f$ noise are present. The knee frequency f_ν is set to be < 20 mHz (less than the frequency of rotation). In the lower panel of figure 4.2 it is shown a segment of one of the rings of the simulation. There is a bright point source near pixel 400 (in fact, it is the brightest source in the simulation). Apart from this extraordinarily bright source, the features that dominate the image are Galactic emission (the large peaks around pixels 150 and 940) and noise.

4.4 Data analysis

The complete set of simulated TOD was filtered using the scale-adaptive filter described in § 4.2. Each individual ring was filtered separately. The power spectrum that appears in equations (4.2) and (4.3) was obtained by averaging the estimated power spectra of twenty-one consecutive rings (the ring that is being filtered, the ten previous rings and the ten subsequent rings).

In order to detect the sources from the filtered image one can set a certain threshold over the dispersion of the filtered ring and look for connected regions (peaks) above that threshold. That would be the most direct detection method if the different rings were independent of their neighbours. This is not the case, since adjacent rings scan very close regions and each source is expected to appear in more than one ring. The use of information coming from neighbouring rings allows to increase the effective signal to noise ratio of the detections and to discard spurious 'sources' due to noise fluctuations. The most straightforward way to detect sources is then to perform a kind of two-dimensional thresholding, looking for connected pixels at equal latitude that appear *in several adjacent rings*. In a first approximation, the position of the source will be the position of the maximum of that region of connected pixels. We will show that this approximation is good enough for the purpose of locating the sources with an error comparable to 1/3 of the antenna FWHM.

Two different regimes of 'noise' have to be removed in order to optimize the

detection of the sources. The large scale features due to Galaxy foregrounds as well as CMB fluctuations are strongly correlated between a given ring of data and its neighbours. On the other hand, the small scale noise, dominated by white instrumental noise, is expected to be nearly independent from one ring to another. This suggests a further step in the idea of combining information from nearby rings to increase the signal to noise ratio of the sources. By averaging each ring data (before filtering) we can construct a 'synthetic TOD' in which the large scale fluctuations are almost the same than in the original TOD but the noise at the scale of the pixel has been greatly diminished. The $1/f$ will still be present in the averaged TOD. The point sources of the original TOD are replaced in the averaged TOD by Gaussians of amplitude

$$A_a = \frac{1}{N_a}\sum_{i=1}^{N_a} A e^{-\frac{|\vec{x_o}-\vec{x_i}|^2}{2\sigma_b^2}} \tag{4.4}$$

where A is the true amplitude of the source, N_a is the number of rings that are averaged, σ_b is the beam width and the distance $|\vec{x_o}-\vec{x_i}|^2$ is the geodesic (spherical) angular distance between the pixel corresponding to the position of the source and the pixel that is being added to calculate the average. This is true when the beam is a perfect Gaussian because in that case the average is a weighted sum of Gaussians of equal width, that is, a new Gaussian of the same width and amplitude given by eq. (4.4). Therefore, we can filter the rather denoised, 'synthetic TOD', instead of the original TOD and correct the amplitude of the sources that we will detect using eq. (4.4) in order to recover an estimate $\hat{A}$ of the true amplitude A.

4.5 Results

In order to determine the goodness of the averaging of close rings we performed two different filterings of the TOD. In the first case, the raw data of the TOD were filtered ring by ring as described in § 4.4. In the second case, a 'synthetic TOD' in which each individual ring was the average of the original ring (at that position) with the twelve neighbours was constructed. The number of rings averaged is such that at the maximum ring spatial separation region on the sky (the ecliptic Equator) the separation between the two most separated rings is approximately equal

to the FWHM of the beam. After filtering, the detection and extraction was performed looking for sets of 5 or more connected pixels (in the 2-dimensional sense). A lesser number of connected pixels required would lead to a much higher number of spurious detections. To check if the detections correspond to real sources we compare with the catalogue of the 1000 brightest sources present in the simulations. This catalogue is complete and its flux limit is clearly lower than the expected flux limit of the detected sources.

In the lower panel of figure 4.3 the number of detections (defined as the number of peaks encountered above a certain threshold that correspond to a real source in the reference catalog) above several σ thresholds is shown for the case where the data have been averaged (filled circles and solid line) and the case where they have not (open circles and dashed line). The remaining quantities of interest (such as estimation of the amplitudes, fluxes, etc.) concerning the detection/extraction of sources over the averaged and then filtered TOD are shown in table 4.1 and will be discussed later. As expected, the number of sources detected when analysing the synthetic TOD is 4 to 5 times higher to the number of sources detected over the original, filtered TOD for every σ level.

Apart from having the greatest number of detections, it is also of great importance to reduce as much as possible the number of spurious detections. In the lower panel of figure 4.4 (labeled as 'simple detection') the ratio between spurious and 'true' detections is represented for the case where the data have been averaged (filled circles and solid line) and the case where they have not (open circles and dashed line). The ratio is lower for the case in which the data have not been averaged. It is evident that a kind of compromise has to be reached between gain and reliability. If we arbitrarily set a maximum proportion of spurious sources versus true ones, say a 10% (represented in fig. 4.4 with an horizontal dashed line), we can determine the minimum σ level that satisfies this condition. In this example, for the case of non-averaged rings, we can reach the 3σ and find 80 sources (to a minimum flux of 4.33 Jy) with a 7.5% of spurious detections. For the case of averaged rings, we must go to the 4σ level, where we find 224 sources (to a minimum flux of 0.885 Jy) with a 9.8% of spurious detections. We conclude that the averaging of neighbouring rings is a valid strategy to reduce pixel-scale noise. In the following, all the results will refer to filtering of averaged rings.

σ	detected sources	spurious sources	mean offset ($'$)	m.r.a.e. (%)	$< bias >$ (%)	min. flux (Jy)	compl. flux (Jy)
			Scale-Adaptive filter				
2.5	549	1368	14.08	24.22	3.15	0.542	4.337
3.0	403	296	13.56	24.32	1.31	0.644	4.337
3.5	351	54	12.73	21.75	-4.91	0.760	4.337
4.0	224	22	12.49	20.98	-7.69	0.885	4.337
5.0	150	9	12.20	19.82	-14.85	1.065	4.337
5.5	124	5	12.03	20.38	-15.54	1.197	4.337
			Gaussian filter				
2.5	11	11	18.88	538.4	537.0	6.511	17.070
3.0	7	6	17.84	746.4	744.2	8.787	10.070
3.5	7	13	17.50	898.1	898.1	7.908	18.866
4.0	4	11	17.24	1281.7	1281.7	10.340	18.866
5.0	4	12	13.09	1985.3	1985.3	13.926	18.866
5.5	4	15	13.09	1985.3	1985.3	13.926	18.866
			Mexican Hat Wavelet				
2.5	473	413	12.83	21.46	4.27	0.531	4.337
3.0	361	130	12.66	20.76	2.52	0.693	4.337
3.5	270	63	12.67	20.46	0.45	0.761	4.337
4.0	196	59	12.40	18.47	-4.22	0.844	4.337
5.0	139	49	12.50	19.47	-4.96	0.957	4.337
5.5	118	44	12.41	19.83	-5.53	1.099	4.337

Table 4.1: Detections at 30GHz with the Scale-Adaptive filter, compared to a Mexican Hat Wavelet (MHW) and a Gaussian window with FWHM of $33'$. Col. (1): σ detection level. Col. (2): Number of sources found. Col. (3): Number of spurious detections. Col. (4): Mean position offset. Col. (5): Mean relative absolute error of the amplitude (defined as $r.a.e. = \langle |A_0 - A_e|/A_0 \rangle$, where A_e is the estimated amplitude). Col. (6): Mean bias in the amplitude. Col. (7): Minimum reached flux. Col. (8): Flux over which the catalog of detections is complete.

The number of detections and spurious sources found with the scale-adaptive

filter applied to a synthetic TOD is shown in table 4.1. The determination of the position of the source, the mean relative absolute error in the determination of the amplitude (defined as $m.r.a.e. = \langle |A_0 - A_e|/A_0 \rangle$, where A_e is the estimated amplitude and A_0 is the real amplitude), the mean bias (defined in the same way as the m.r.a.e, but without the absolute value), the minimum flux reached and the completeness flux are also included in table 4.1. In each case the mean error in the position of the sources is comparable with the size of the 'pixel' of the TOD, $11'$. The determination of the amplitude using eq. (4.4) has relative errors ranging from 24.32% at 3σ threshold to 19.82% at 5.0σ. The error decreases as the detection threshold increases. This indicates that the estimation of the amplitude of weak sources is less accurate than the estimation of the amplitude of bright ones. Also, the determination of the amplitude is biased to higher values at low σ levels and to lower values (negative bias) at high σ levels. The positive bias for weak sources arises due to the peak finding algorithm: it preferably finds the maxima in pixels where the noise contribution is positive. The negative bias that appears at high σ levels is systematic and is due to the fact that the filter is designed to detect 'perfect' Gaussian sources whereas real data are non ideal. This systematic bias is never greater than 16%. At intermediate thresholds, both effects tend to cancel and, *on average*, the amplitude estimates are unbiased.

To compare with other 'classic' filters we repeated the process using a Gaussian filter and a Mexican Hat Wavelet (MHW, hereafter), both of them with a width equal to the beam width of the antenna ($33'$). The normalization of both filters was chosen so that the coefficient at the position of the source is equal to the amplitude of the source (that is, the filtering process does not change the amplitude of the sources). The scale-adaptive filter automatically satisfies this condition (condition 2 for an scale-adaptive filter). While the MHW and the scale-adaptive filters are both band-pass filters, the Gaussian is a low-pass filter, so the comparison with the Gaussian is a bit unfair: the Gaussian is expected to perform significantly worse than the other two filters. The number of detections above several thresholds for the two 'classic' filters together with the scale-adaptive filter are shown in table 4.1. The number of detections is similar for the MHW and the scale-adaptive filter, yet are slightly higher for the scale-adaptive filter. The lowest number of detections corresponds to the Gaussian filter. In the lower panel of figure 4.5 the number of

detections with the three different filters is compared. Detections with the scale-adaptive filter are shown with open circles and solid line. The open boxes and dashed line corresponds to MHW detections and the triangles and dashed line corresponds to Gaussian filter detections. The ratio between spurious and true detections for the three filters is shown in the top panel of figure 4.5. Except for the 2.5σ level, the Gaussian filter produces the worst ratio. The scale-adaptive filter gives spurious to detected (e/d) ratios that quickly decline with increasing σ thresholds. The e/d ratio for the MHW remains almost constant with σ in the considered cases and clearly exceeds the ratio obtained with the scale-adaptive filter. The m.r.a.e. and the bias in the determination of the amplitude are huge in the case of the Gaussian filter. Both have very similar values. That means that the main source of error is systematic (the filter is biased). Considering the m.r.a.e, the MHW seems to give amplitude estimates a few percent better than the scale-adaptive filter. Flux limits are similar in the MHW and scale-adaptive filter cases. The Gaussian filter leads to higher inferior flux and completeness limits.

Figure 4.2 provides a useful insight into what is happening with the different filters. The Gaussian filter smoothes the image, removing very efficiently the small scale noise but allowing the large structures (the Galaxy and others extended fluctuations) to remain in the image. The naive σ thresholding counts these bright, large structures as sources, leading to a big relative number of spurious detections. Besides, the sources that by chance lie on 'valleys' of the background can not be detected. On the other hand, sources that lie on areas of positive background are enhanced and can be more easily detected. This can explain the large and positive bias in the detections with the Gaussian filter.

Both MHW and scale-adaptive filter are better prepared to deal with this problem than the Gaussian window. Their profiles in Fourier space drop to zero at low frequencies and thus they are efficient in removing large scale structures. Images filtered with the MHW and the optimal scale-adaptive filter in figure 4.2 are similar. A visual inspection reveals that MHW smoothes better the high-frequency fluctuations. The scale-adaptive filter is more efficient in removing medium and large structures. The fluctuations around pixel 150 (corresponding to one of the two observations of the Galaxy in the ring) are better removed with the optimal scale-adaptive filter than with the MHW. This is more apparent in figure 4.1 where

the profiles of the different filters in Fourier space clearly indicate the faster drop of the scale-adaptive filter at low frequencies (large scales) and its slower drop at high frequencies (small scales). The MHW can have the same problem as the Gaussian filter. However, the probability of this problem should be smaller due to its better efficiency in removing the Galaxy and other large-scale structures. This effect explains the higher e/d ratio and the trend to positive bias.

The conclusion is that the scale-adaptive filter detects point sources better than the MHW and the Gaussian window. The number of detections is comparable to the number of detections with the MHW and clearly higher than with the Gaussian window. The relative number of spurious detections with the scale-adaptive filter is lower, except for very low detection thresholds, than with the MHW and the Gaussian window. Over 4σ the contamination of spurious detections is lower than 10%. At this level (4σ) the number of expected sources in all the sky is of a few hundreds (224 in our simulation) above fluxes of around 0.9 Jy.

Some tests can be performed in order to discover if the number of spurious sources, the most unpleasant effect of the filtering and detection process, can be reduced. Where do these spurious detections come from? One possibility is that the peak finding algorithm is detecting the Galaxy or other large-scale features. Such structure will appear in several adjacent rings, as sources do, and therefore a possibility of confusion exists. To test this potential source of contamination, we repeated the analysis excluding a band of 5° centered on the Galactic plane (corresponding to a 4.36% of the sky area). The top panels of figures 4.3 and 4.4 show the number of detections and the ratio of spurious/true detections, respectively, for the scale-adaptive filter in the cases where the rings have been averaged (as explained before, filled circles) and where they have not (open circles). In the first section of table 4.2 are the results for the case of averaged rings (the tabulated quantities are the same than in table 4.1) are shown. The decrease in the number of detections corresponds to the one expected for a uniform distribution of sources in the sky (around 5%). This indicates that the density of detections around the Galactic plane is not substantially different from the density of detections in other regions, less 'contaminated', of the sky. This can be seen in figure 4.6, where the 4σ detections have been represented in Galactic coordinates (the Galactic Plane being represented by a dashed line).

σ	detected sources	spurious sources	mean offset ($'$)	m.r.a.e. (%)	$< bias >$ (%)	min. flux (Jy)	compl. flux (Jy)
			Scale-Adaptive filter, excluding the Galaxy				
2.5	520	1272	14.06	23.79	2.67	0.542	13.722
3.0	379	257	13.51	23.69	0.58	0.644	13.722
3.5	280	34	12.67	20.87	-6.04	0.760	13.722
4.0	212	8	12.45	20.05	-8.97	0.885	13.722
5.0	142	1	12.16	18.73	-16.25	1.066	13.722
5.5	118	1	11.94	18.97	-17.22	1.197	13.722
			Mexican Hat Wavelet, excluding the Galaxy				
2.5	454	347	12.76	20.21	2.79	0.531	13.72
3.0	342	87	12.54	19.04	0.42	0.693	13.72
3.5	256	25	12.49	18.28	-2.37	0.806	13.72
4.0	187	14	12.15	15.71	-7.65	0.844	13.72
5.0	129	7	12.18	15.07	-10.64	0.957	13.72
5.5	108	7	12.01	14.60	-12.37	1.099	13.72

Table 4.2: Different tests for the detections at 30 GHz. Columns have the same meaning than in table 4.1. Two different cases are tabulated: detections with the scale-adaptive filter when a 5° band around the Galactic plane is excluded from the analysis and detections with a $33'$ MHW when a 5° band around the Galactic plane is excluded from the analysis.

The ratio between spurious and true detections remains almost untouched for low σ. For higher σ levels, the proportion of spurious detections that are due to the Galactic plane increases dramatically (comparing tables 4.1 and 4.2 we see that this proportion increases from a 7% at 2.5σ to a 89% at 5σ). This indicates that the contamination by the Galaxy dominates at high signal to noise ratios, whereas low-intensity contamination is dominated by noise fluctuations. In all cases, the number of peaks that correspond to the Galaxy is much smaller than the number of detected sources, thus implying that the scale-adaptive filter deals efficiently with large scale backgrounds.

We stated before that the higher number of spurious detections of MHW could be due to its non-optimal performance at large scales. To further test this hy-

pothesis we repeated the test for the MHW, now considering only the peaks found outside a 5° band centered in the Galactic plane, as we did with the scale-adaptive filter before. The results are shown in the last section of table 4.2. Most of the spurious sources (specially at high σ) lied near the Galactic plane, as we expected. And yet the remaining number of spurious sources is still greater than in the equivalent scale-adaptive filter case. The number of detections and the flux limits remain similar to the scale-adaptive filter case. The m.r.a.e. is also similar in the two cases. Finally, the mean bias in the determination of the amplitude is negative for high σ thresholds, as happens with the optimal scale-adaptive filter. Note that the MHW used in this work has the same scale as the antenna. In fact, that scale is not the optimal for detection (Vielva et al. 2001a). The optimal scale of the MHW for a particular case has to be determined from the power spectrum of the data. This, in a certain way, mimics the determination of the scale R_0 *that is automatically included* in the scale-adaptive filter via eqs. (4.2) and (4.3). The MHW at its optimal scales resembles the shape of the scale-adaptive filter in Fourier space (in the sense that its maximum is located near the maximum of the scale-adaptive filter) and thus the effectiveness of both filters should be similar.

In order to further decrease the number of spurious detections, we could take advantage of the fact that, in many realistic cases, due to the sky coverage of the experiment and its scanning strategy, many positions of the sky can be measured more than once (that is, at different epochs). For example, a source of 0° latitude will be detected once when the center of the ring is located on longitude $\phi = \phi_{source} - R$, being R the radius of the ring, and once again several months later, when the center of the ring is located on longitude $\phi = \phi_{source} + R$. When this occurs, it would be possible to almost duplicate the amount of information in some areas of the sky and therefore improve both the sensitivity and the reliability of the detection. However, such a refined detection can not be done when the data do not cover the whole sky, and therefore is useless for the construction of an early catalogue of sources during the mission flight.

4.6 Conclusions

In this chapter a sequence of time ordered data such as the one that the future Planck 30 GHz LFI28 channel will produce after the first 6 months of flight has been simulated and analysed using scale-adaptive filters. The data include all the main foregrounds as well as CMB fluctuations, point sources and instrumental noise. The resulting TOD has been ring-averaged in order to remove pixel-scale noise and then filtered with an scale-adaptive filter that includes the spectral properties of the data in order to maximize the detection of sources of a particular shape (Gaussian) and scale (the scale of the antenna). The optimal scale-adaptive filter was designed to produce an unbiased, efficient estimator of the amplitude of the sources at their position and to give a maximum of detections at the characteristic scale of the sources. The detection of the sources was performed by thresholding the filtered TOD and looking for connected sets of peaks belonging to adjacent rings. At a 4σ detection level 224 sources over a flux of 0.88 Jy are detected with a mean relative error (in absolute value) of 20.98% and a systematic bias of -7.69%. The position of the sources in the sky is determined with errors inferior to the size of the antenna. The catalogue of detected sources is complete at fluxes ≥ 4.337 Jy. The number of spurious detections is 22.

The performance of the scale-adaptive filter has been compared with the performances under the same conditions of a Gaussian window and a Mexican Hat Wavelet (MHW) of width equal to the beam width. The number of sources detected with the scale-adaptive filter is comparable to the number of sources detected with the MHW and much higher than the number of sources detected with the Gaussian. However, the scale-adaptive filter finds a significantly lower number of spurious sources than the other two filters. This is due to the fact that the scale-adaptive filter removes better the large scale fluctuations (e.g. the Galaxy) than the other two filters. To test this hypothesis a 5° band around the Galactic Plane was removed and the analysis was repeated with the scale-adaptive filter and the MHW. This test shows that most of the spurious detections that were found with the MHW were located on the Galactic Plane, whereas spurious detections that were found with the scale-adaptive filter are uniformly distributed in the sky.

The mean absolute error in the determination of amplitudes is slightly higher in

the case of the optimal scale-adaptive filter than in the MHW case. The estimation of the amplitudes with the Gaussian window suffer from very large errors. In most cases, the Gaussian detects only the Galaxy.

The number of spurious detections can be reduced by means of further analysis after the filtering. For example, when the sky coverage of the scan is big enough some areas of the sky can be observed twice or more times, increasing the signal to noise ratio of the sources that lie in such areas. This point is out of the scope of the present work, in which we only present the filter as a first step in the data reduction.

In conclusion, the scale-adaptive filter is an efficient, unbiased and reliable tool for the detection and extraction of punctual sources from TOD. A few hundred of sources over 1 Jy will be detected in the 30 GHz channels of the future Planck mission with $\leq 10\%$ of spurious detections. The catalogue will be complete at fluxes ≥ 4 Jy. One possible application of this technique would be the elaboration of early catalogues of sources. A greater number of detections is expected at higher frequencies. The simulation and analysis of such frequencies will be performed in a future work.

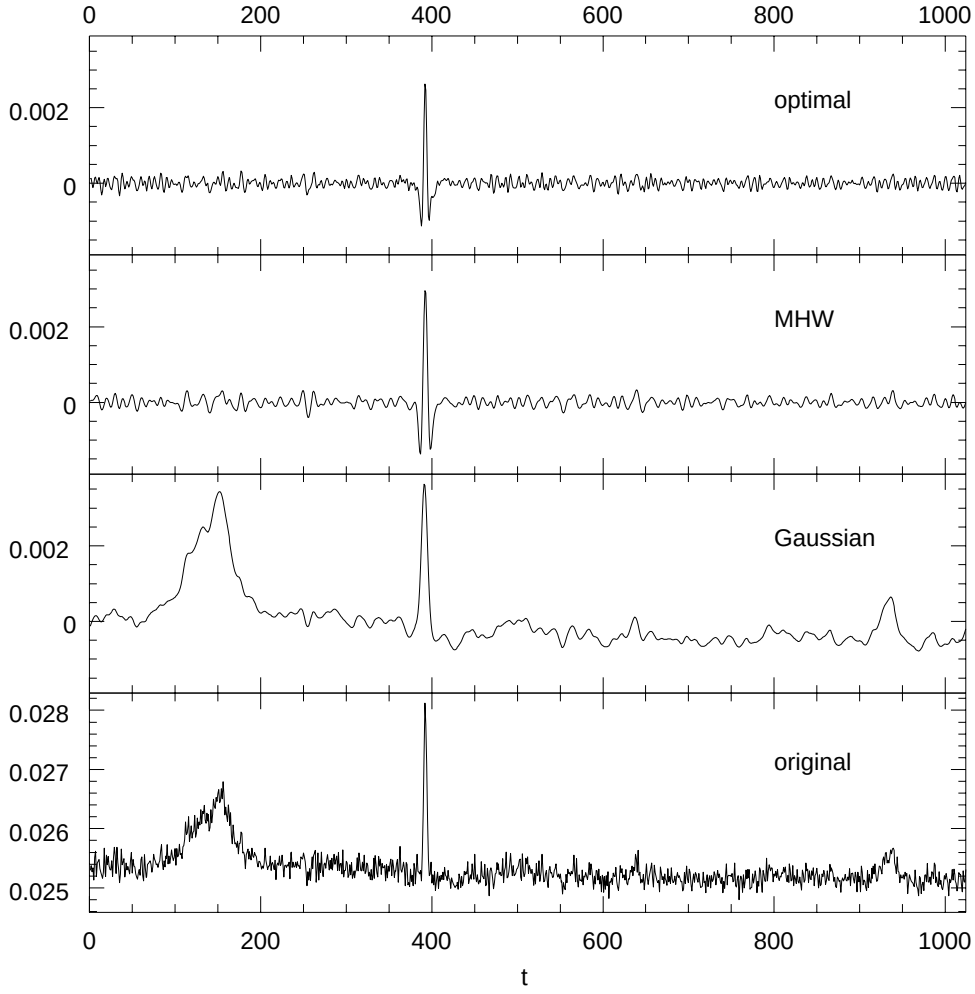

Figure 4.2: A 1024-pixel section of one of the rings before and after filtering. The lowest panel shows the section before any filtering. The second panel shows the section after filtering with a Gaussian of FWHM 33′ (equal to the one of the source). The third panel shows the section after filtering with a Mexican Hat Wavelet of width equal to the width of the source. The top panel shows the section after filtering with the (optimal) scale-adaptive filter.

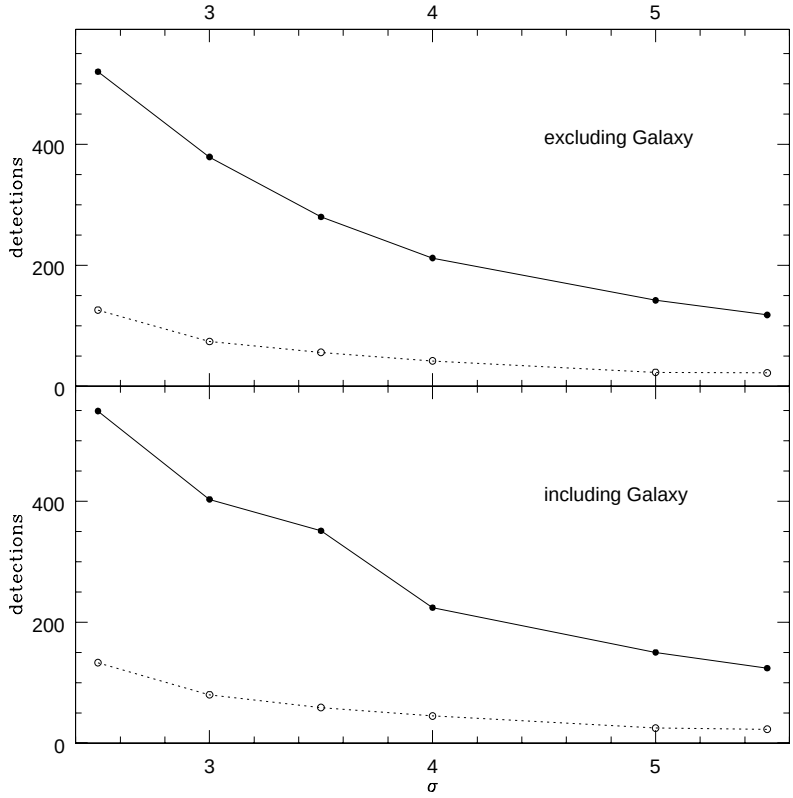

Figure 4.3: 30 GHz detections with the scale-adaptive filter as a function of the σ detection threshold. The open circles and dashed lines represent detections over the filtered TOD. The filled circles and solid line represent detections over a filtered synthetic TOD in which each ring is the result of the average of 9 rings of the original TOD. The synthetic TOD has the same number of rings than the original one. Two cases have been represented: simple detection over a certain threshold (lower panel) and simple detection excluding a 5° band around the Galactic plane (top panel).

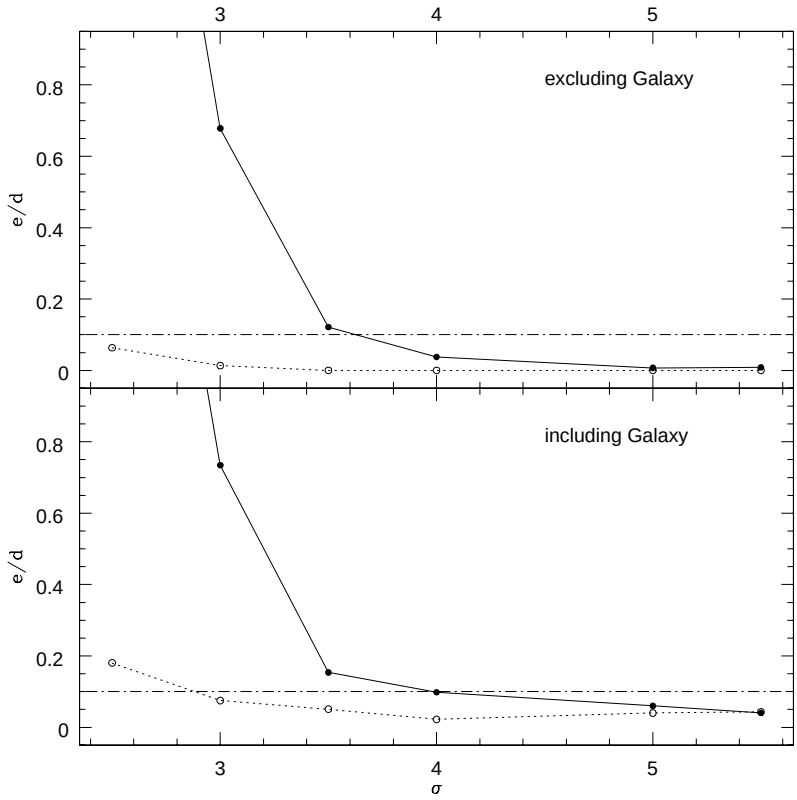

Figure 4.4: 30 GHz ratio between the number of spurious sources and the number of true detections with the scale-adaptive filter as a function of the σ detection threshold. The open circles and dashed lines represent detections over the filtered TOD. The filled circles and solid line represent detections over a filtered synthetic TOD in which each ring is the result of the average of 9 rings of the original TOD. The synthetic TOD has the same number of rings than the original one. Two cases have been represented: simple detection over a certain threshold (lower panel) and simple detection excluding a 5° band around the Galactic plane (top panel).

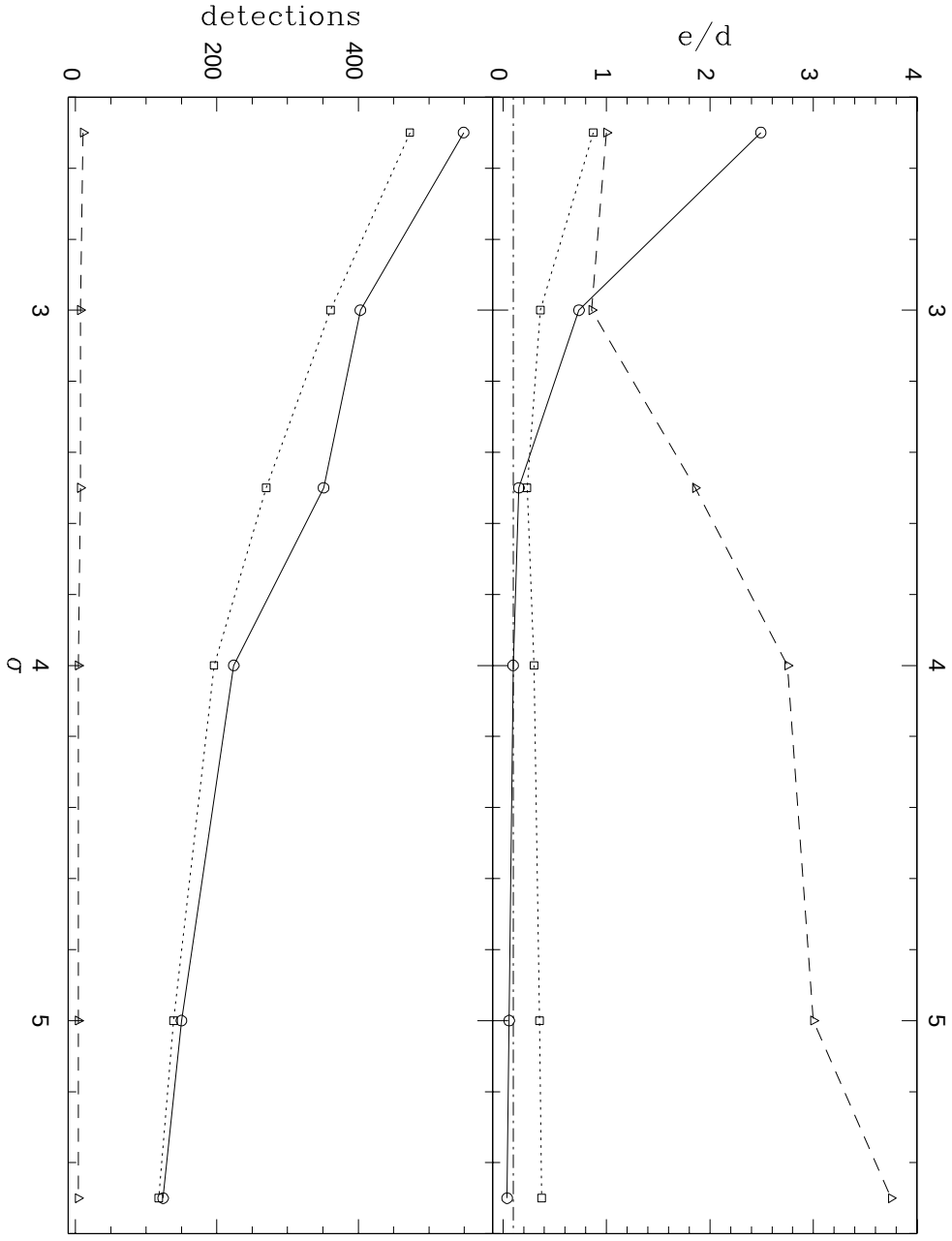

Figure 4.5: Comparison of the performances of three different filters over the 30 GHz TOD. Open circles and solid line refer to scale-adaptive filter results. Boxes and dashed line correspond to Mexican Hat Wavelet results. Triangles and dashed line correspond to Gaussian filter results. Both the MHW and Gaussian filter have a width of $33'$. The lower panel shows the number of detections above the different σ thresholds. The upper panel shows the ratio between spurious and true detections for the same σ thresholds.

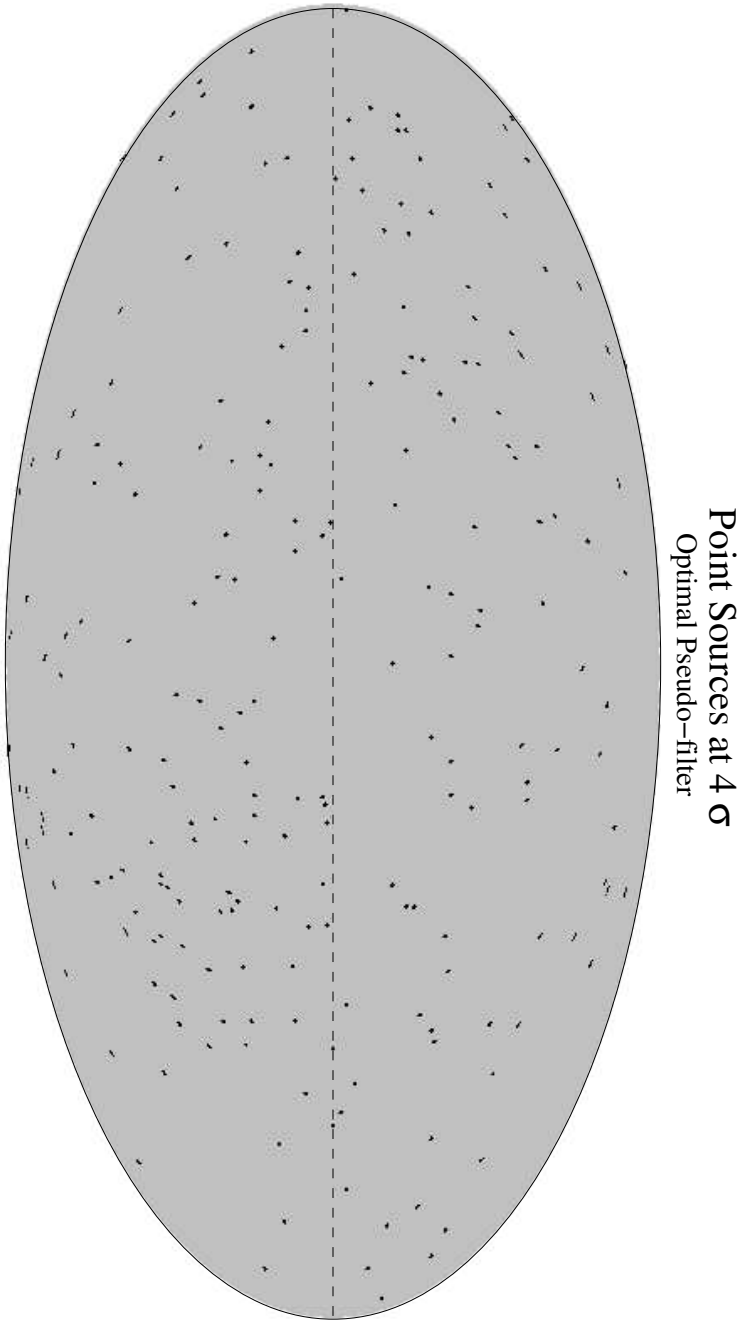

Figure 4.6: Positions of the sources detected with the scale-adaptive filter at a 4σ detection level. The sky is projected in Galactic coordinates. The dashed line represents the Galactic Equator. The scale-adaptive filter is able to detect sources even in the highly contaminated region around the Galactic Plane.

Chapter 5

Detection of clusters in Planck data

Another interesting application of scale-adaptive filters is the detection of galaxy clusters in CMB multi-channel data. The intensity of the CMB towards the line of sight of a cluster is distorted due to the Sunyaev-Zel'dovich effect (SZE), as can be seen in eqs. (1.13) and (1.14). The SZE has a characteristic frequency dependence (see figure 1.3) that can be used to improve the detection of clusters. The future Planck mission will allow to measure the whole sky in ten different frequency channels. Therefore, it is expected to provide a very good chance of detecting thousands of galaxy clusters. In this chapter, two different generalisations of the scale-adaptive and matched filters will be presented that optimize the detection of signals with a known frequency dependence in multifrequency data.

5.1 Introduction

The detection and characterisation of the Sunyaev-Zel'dovich effect (SZE) is one of the most promising fields of interest in millimetric and sub-millimetric Astronomy. The Cosmic Microwave Background (CMB) radiation is distorted in the direction of galaxy clusters due to the hot intracluster plasma. This distortion, among other contaminants, has to be corrected in order to study the anisotropies of the CMB. Besides, the characteristic signature of the SZE can be used for the study of the intracluster medium, its origin and its evolution. As the SZE does not depend on

the redshift of the clusters, it can also be used to detect clusters that otherwise would be undetectable. For example, the future ESA Planck satellite will produce a full-sky catalogue containing several tens of thousands of galaxy clusters; therefore, it is crucial to have a robust and efficient method to detect and extract the SZE.

Component separation of CMB images has been thoroughly studied and discussed in the literature. Regarding SZE detection, methods such as Wiener filtering (WF, Tegmark & Efstathiou 1996, Bouchet et al. 1997) and maximum entropy method (MEM, Hobson et al. 1998, 1999) have proven to be very powerful. In general, a component separation method that uses all the available information will be more powerful than other that does not assume any prior knowledge about the data. MEM, for instance, produces astounding results in the separation when the power spectrum and the frequency dependence of the components are well known. Unfortunately, if the assumptions about the data are wrong, errors may arise that would affect the separation of several (or all) components. This is particularly dangerous in methods that perform the separation of all the components simultaneously (WF, MEM, FastICA). The opposite approach is to use a robust method that makes as few assumptions as possible about the data. In this case the separation is less powerful due to the loss of valuable information that could be helpful to distinguish among the different components of the image. An example of a method that makes almost no assumptions is the non-parametric Bayesian SZE detection method given by Diego et al. (2001b), in which only the well-known frequency dependences of the SZE and the CMB are needed. An example of a method that can make use all the available information is MEM. Depending on the circumstances of the analysis, a compromise between these two poles must be reached.

Filtering techniques, such as Mexican Hat wavelets (MHW), matched filters (MF) and scale-adaptive filters (SAF), are single component separation methods that use some of the available information (i.e. the shape of the component that is going to be detected and the power spectrum of the combination of the rest of components) while reducing the errors that may appear due to error propagation in methods that separate all components simultaneously. MHW assumes a specific shape (a Gaussian) for the component to be separated ('sources') and, given the power spectrum of the background (that can be directly estimated from

the data), finds the optimal scale of the filter. MF generalises this idea by introducing more general profiles. Adaptive filters put additional emphasis on the characteristic scale of the sources in order to further reduce the number of spurious detections. Although usually the spherical symmetry of the sources is assumed, it is not a requirement of the method and the filters can be easily generalised to detect asymmetric features.

This chapter will discuss the generalisation of filtering techniques, in particular MF and AF, to the case of multiple images corresponding to different frequency channels. Multifrequency information can be used both to increase the signal of the sources and to reduce the contribution of the background ('noise'). This information can be used prior to the filtering of the images (by using the correlations between the different channels to find an optimal combination of channels that maximises the signal to noise ratio of SZE clusters) or be included directly in the construction of the filters.

5.2 Formalism

There are two different approaches we can follow to include the frequency dependence of a signal in the filtering of multi-channel data. On one hand, we can filter each channel separately, but carefully taking into account the cross-correlation between the different channels and the frequency dependence of the signal in order to obtain an output set of filtered maps that, added to each other, must be optimal for the detection. This philosophy gives birth to the *multifilter method*. On the other hand, one can use the information about correlations and frequency dependence before filtering in order to find the optimal combination of channels that maximizes the signal to noise ratio of the sources, and then use a filter on the optimally combined map. This leads to the design of a *single filter method*.

5.2.1 The multifilter method

Let us consider a set of 2-dimensional images (maps) with data values defined by

$$y_\nu(\vec{x}) = f_\nu s_\nu(x) + n_\nu(\vec{x}), \quad x \equiv |\vec{x}|, \quad \nu = 1, ..., N, \tag{5.1}$$

where N is the number of maps (or number of frequencies) and f_ν is the frequency dependence of the SZE. We choose a fiducial frequency ν_f such that $f_{\nu_f} = 1$, therefore A is the amplitude of the cluster at frequency ν_f. $\vec{x}$ is the spatial coordinate and $s_\nu(x) \equiv A\tau_\nu(x)$ represents a cluster with spherical symmetry placed at the origin with amplitude A and profile $\tau_\nu(x)$. If one assumes a β-profile for the electron number density $n_e(r) \propto \left[1 + (x/r_c)^2\right]^{-\frac{3}{2}\beta}$ and assuming the standard value $\beta = 2/3$, one trivially obtains for the microwave emission through the SZ-effect

$$\tau(x) = \frac{1}{\left[1 + (\frac{x}{r_c})^2\right]^{1/2}}. \tag{5.2}$$

This is the multiquadric profile presented in Chapter 3. A convolution with the antenna beam at the frequency ν, assumed here to be a 2D-Gaussian with dispersion θ_ν, gives the convolved profile $\tau_\nu(x)$.

The background $n_\nu(\vec{x})$ is modeled by a homogeneous and isotropic random field with average value $\langle n_\nu(\vec{x})\rangle = 0$ and cross-power spectrum $P_{\nu_1\nu_2}(q), q \equiv |\vec{q}|$: $\langle n_{\nu_1}(\vec{q}) n^*_{\nu_2}(\vec{q'})\rangle = P_{\nu_1\nu_2}(q)\, \delta^2_D(\vec{q} - \vec{q'})$, $n_\nu(\vec{q})$ is the Fourier transform of $n(\vec{x})$ and δ^2_D is the 2D Dirac distribution.

Let us for one moment forget the shape given by eq. (5.2) and consider a generic profile $\tau(x)$. The results in following sections are general for any τ.

Scale-Adaptive multifilter (SAMF)

The idea of an optimal scale-adaptive filter has been introduced and developed in previous chapters. By introducing a spherically-symmetric filter, $\Psi_\nu(\vec{x}; R_\nu, \vec{b})$, dependent on 4 parameters (R_ν defines a scaling whereas $\vec{b}$ defines a translation) for any of the frequencies

$$\Psi_\nu(\vec{x}; R_\nu, \vec{b}) = \frac{1}{R^2_\nu} \psi_\nu \left(\frac{|\vec{x} - \vec{b}|}{R_\nu} \right), \tag{5.3}$$

we define the filtered field as

$$w_\nu(R_\nu, \vec{b}) = \int d\vec{x}\, y_\nu(\vec{x}) \Psi_\nu(\vec{x}; R_\nu, \vec{b}), \tag{5.4}$$

and the total coefficient as

$$w(R_1, ..., R_n, \vec{b}) \equiv \sum_\nu w_\nu(R_\nu, \vec{b}). \tag{5.5}$$

The previous convolution can be written as a product in Fourier space, in the form

$$w_\nu(R_\nu, \vec{b}) = \int d\vec{q} e^{-i\vec{q}\cdot\vec{b}} y_\nu(\vec{q})\psi_\nu(R_\nu q), \tag{5.6}$$

where $y_\nu(\vec{q})$ and $\psi_\nu(q)$ are the Fourier transforms of $y_\nu(\vec{x})$ and $\psi_\nu(\vec{x})$, respectively. A simple calculation gives the average at the origin $\vec{b} = 0$, $\langle w_\nu(R_\nu, \vec{0})\rangle$ and the variance, $\sigma_w^2(R_1, ..., R_n) \equiv \langle w^2(R_1, ..., R_n, \vec{b})\rangle - \langle w(R_1, ..., R_n, \vec{b})\rangle^2$, of the filtered field

$$\langle w_\nu(R_\nu, \vec{0})\rangle = 2\pi A\, f_\nu \int dq\, q\tau_\nu(q)\psi_\nu(R_\nu q), \tag{5.7}$$

$$\sigma_w^2(R_1, ..., R_n) = 2\pi \sum_{\nu_1,\nu_2} \int dq\, q\, P_{\nu_1\nu_2}(q)\psi_{\nu_1}(R_{\nu_1} q)\psi_{\nu_2}(R_{\nu_2} q). \tag{5.8}$$

where the limits in the integrals go from 0 to ∞.

Now, we are going to express the conditions in order to obtain a scale-adaptive multifilter (SAMF) for the detection of the cluster at the origin, taking into account that the source has a bell shape with a single characteristic scale $R_{o\nu}$ in each map. The following conditions are assumed:

1. $\langle w(R_{o1}, ..., R_{on}, \vec{0})\rangle = s_{\nu_f}(0) \equiv A$, i. e. $w(R_{o1}, ..., R_{on}, \vec{0})$ is an *unbiased* estimator of the amplitude of the source.

2. The variance of $w(R_1, ..., R_n, \vec{b})$ has a minimum at the scales $R_{o1}, ..., R_{on}$, i. e. it is an *efficient* estimator.

3. $w_\nu(R_\nu, \vec{b})$ has a maximum at $(R_{o\nu}, \vec{0})$.

Then, the multifilter satisfying these conditions is given by the matrix equation

$$\tilde{\psi}(q) = \mathbf{P}^{-1}(\alpha\mathbf{F} + \mathbf{G}), \tag{5.9}$$

where we we have introduced the following column vectors: $\tilde{\psi}(q) = (\psi_\nu(R_\nu q))$, $\mathbf{F} = (f_\nu\tau_\nu)$, $\mathbf{G} = (\mu_\nu\beta_\nu)$, $\mu_\nu \equiv f_\nu\tau_\nu\left[2 + \frac{dln\tau_\nu}{dlnq}\right]$, and $\mathbf{P}^{-1}$ is the inverse matrix of $\mathbf{P} \equiv (P_{\nu_1\nu_2}(q))$,

$$\alpha = A^{-1}{}_0^0, \quad \beta^\nu = A^{-1}{}_0^\nu, \tag{5.10}$$

where $\mathbf{A}$ is the $(1+n)\times(1+n)$ matrix whose elements are given by

$$A_0^0 \equiv \int d\vec{q}\, \mathbf{F}^t\mathbf{P}^{-1}\mathbf{F}, \quad A_\nu^0 \equiv \int d\vec{q}\, \mu_\nu(\mathbf{F}^t\mathbf{P}^{-1})_\nu, \tag{5.11}$$

$$A_0^\nu \equiv \int d\vec{q}\,\mu^\nu (\mathbf{P}^{-1}\mathbf{F})^\nu, \quad A_{\nu'}^\nu \equiv \int d\vec{q}\,\mu^\nu \mu_{\nu'}\, P^{-1}{}_{\nu'}^{\nu}\,. \tag{5.12}$$

This will be called *scale − adaptive* multifilter extending the concept considered in previous chapter for a single map. Equations (5.9) and (5.11)-(5.12) are indeed the generalisation of eq. (2.10) and (2.11), respectively. The derivation of these formulas is qualitatively equivalent to the derivation of the scale-adaptive filter in Appendix A, though more laborious: condition (3) gives N constraints instead of one and the minimisation problem involves $(N+1)$ Lagrangian multipliers, leading to the resolution of $(N+1)$ equations with $(N+1)$ unknown quantities that, in matrix form, reduces to eq. (5.9).

On the other hand, the variance is given by

$$\sigma_w^2 = \int d\vec{q}\,\tilde{\psi}^t \mathbf{P}\tilde{\psi} = (A_0^0)^{-1}\left[1 + 2\int d\vec{q}\,\mathbf{F}^t\mathbf{P}^{-1}\mathbf{G}\right] + \int d\vec{q}\,\mathbf{G}^t\mathbf{P}^{-1}\mathbf{G}. \tag{5.13}$$

A particular case is when $\mathbf{P}$ is a diagonal matrix, i. e. there is no cross-correlation between the backgrounds in the different maps $P_{\nu\nu'} = \delta_{\nu\nu'}P_\nu(q)$. In this case the multifilter is given by

$$\tilde{\psi}_\nu(q) = \frac{f_\nu \tau_\nu}{P_\nu}\frac{1}{a - \sum_\nu f_\nu^2 \frac{b_\nu^2}{H_\nu}}\left[1 - \left(2 + \frac{dln\tau_\nu}{dlnq}\right)\frac{b_\nu}{H_\nu}\right], \tag{5.14}$$

$$a = \sum_\nu f_\nu^2 \int d\vec{q}\,\frac{\tau_\nu^2}{P_\nu}, \quad b_\nu = f_\nu^2 \int d\vec{q}\,\frac{\tau_\nu^2}{P_\nu}\left(2 + \frac{dln\tau_\nu}{dlnq}\right)$$

$$H_\nu = f_\nu^2 \int d\vec{q}\,\frac{\tau_\nu^2}{P_\nu}\left(2 + \frac{dln\tau_\nu}{dlnq}\right)^2. \tag{5.15}$$

If, additionally, one assumes that the backgrounds are white noise, i. e. P_ν are constants and the resolution is the same in all maps, originating the convolved profile $\tau(r_c, \theta)$, then

$$f_\nu \tilde{\psi}_\nu(q) = \frac{\tau}{N}\frac{f_\nu^2 P_\nu^{-1}}{\sum_\nu f_\nu^2 P_\nu^{-1}}\left[1 - \frac{b}{H}(2 + \frac{dln\tau}{dlnq})\right], \tag{5.16}$$

$$N \equiv \int d\vec{q}\,\tau^2 - \frac{b^2}{H}, \quad b = \int d\vec{q}\,\tau^2\left(2 + \frac{dln\tau}{dlnq}\right),$$

$$H = \int d\vec{q}\,\tau^2\left(2 + \frac{dln\tau}{dlnq}\right)^2. \tag{5.17}$$

Matched multifilter (MMF)

If we do not assume condition (3) defining the SAMF in the previous subsection, it is not difficult to find another type of multifilter after minimization of the variance (i.e. condition (2) with the constraint (1)). Then, the multifilter satisfying such conditions is given by the matrix equation

$$\tilde{\psi}(q) = \alpha \mathbf{P}^{-1}\mathbf{F}, \quad \alpha^{-1} = \int d\vec{q}\mathbf{F}^t\mathbf{P}^{-1}\mathbf{F}, \tag{5.18}$$

where $\mathbf{F}$ is the column vector $\mathbf{F} = (f_\nu\tau_\nu)$ and $\mathbf{P}^{-1}$ is the inverse matrix of the cross-spectrum one P. This will be called *matched* multifilter extending the concept usually considered for a single map. On the other hand, the variance is given by

$$\sigma_w^2 = \int d\vec{q}\tilde{\psi}^t\mathbf{P}\tilde{\psi} = \alpha. \tag{5.19}$$

A particular case is when P is a diagonal matrix, i.e. there is no cross-correlation between the backgrounds in the different maps $P_{\nu\nu'} = \delta_{\nu\nu'}P_\nu(q)$. In this case the multifilter is given by

$$\tilde{\psi}_\nu(q) = \alpha f_\nu\tau_\nu P_\nu^{-1}, \quad \alpha^{-1} = \sum_\nu f_\nu^2 \int d\vec{q}\frac{\tau_\nu^2}{P_\nu}. \tag{5.20}$$

If, additionally, one assumes that the backgrounds are white noise, i.e. P_ν are constants and the resolution is the same in all maps, originating the convolved profile $\tau(r_c, \theta)$, then

$$f_\nu\tilde{\psi}_\nu(q) = \frac{\tau}{N}\frac{f_\nu^2 P_\nu^{-1}}{\sum_\nu f_\nu^2 P_\nu^{-1}}, \quad N \equiv \int d\vec{q}\tau^2. \tag{5.21}$$

5.2.2 The Single Filter method

The idea is to build up a template (optimal map) combination of the original ones (see eq. (5.1)) with a set of weights c_ν:

$$y(\vec{x}) = At(x) + \epsilon(\vec{x}), \tag{5.22}$$

$$y(\vec{x}) \equiv \sum_\nu c_\nu y_\nu(\vec{x}), \;\; t(x) \equiv \sum_\nu c_\nu f_\nu\tau_\nu(x), \;\; \epsilon(\vec{x}) \equiv \sum_\nu c_\nu n_\nu(\vec{x}). \tag{5.23}$$

The optimal values of the weights are found assuming the criterion: maximisation of the quotient $Q \equiv At(\vec{0})/\sigma_\epsilon$, σ_ϵ is the dispersion of the noise field $\epsilon(\vec{x})$. The last quantity is given by

$$\sigma_\epsilon^2 = \mathbf{c}^t\mathbf{M}\mathbf{c}, \quad \mathbf{c} = (c_\nu), \quad \mathbf{M} = (M_{\nu\nu'}), \quad M_{\nu\nu'} = \langle n_\nu(\vec{x}) n_{\nu'}(\vec{x})\rangle. \tag{5.24}$$

The previous problem is equivalent to maximize

$$Q^2 = A^2 \frac{\mathbf{c}^t\mathbf{G}\mathbf{c}}{\mathbf{c}^t\mathbf{M}\mathbf{c}}, \quad \mathbf{G} = (G_{\nu\nu'}), \quad G_{\nu\nu'} = f_\nu \tau_\nu(0) f_{\nu'} \tau_{\nu'}(0). \tag{5.25}$$

The solution is easily found to be the eigenvector $\mathbf{c}$ associated to the largest eigenvalue λ of the generalized eigenvalue problem

$$(\mathbf{G} - \lambda\mathbf{M})\mathbf{c} = 0. \tag{5.26}$$

Scale-Adaptive single filter (SAF)

Once we have obtained the template $t(x)$ given by eq. (5.23) with the optimal weights $\mathbf{c}$ given by eq. (5.26), we can apply a scale-adaptive filter to the single image given by eq. (5.22). The only difference to the approach given in Chapter 2 is that the profiles of the sources present in the combined map are more complicated (namely, the superposition of individual profiles which in many cases are convolved with different resolution beams).

Matched single filter (MSF)

Once we have obtained the template $t(x)$ given by eq. (5.23) with the optimal weights $\mathbf{c}$ given by eq. (5.26), we can apply a matched filter to the single image given by eq. (5.22).

$$\tilde{\psi}(q) = \frac{1}{2\pi a}\frac{t(q)}{P(q)}, \quad a = \int dq\, q \frac{t^2(q)}{P(q)},$$

$$P(q) = \mathbf{c}^t\mathbf{N}\mathbf{c}, \quad \mathbf{N} = (P_{\nu\nu'}(q)), \tag{5.27}$$

where $t(q)$ and $P(q)$ are the Fourier transform of $t(x)$ and the power spectrum of the noise ϵ, respectively. This last formula applies also to the case of the scale-adaptive single filter.

Frequency (GHz)	FWHM (arcmin)	Pixel size (arcmin)	Fractional bandwidth ($\Delta\nu/\nu$)	σ_{noise} (μK)
30	33.0	6.0	0.20	4.4
44	23.0	6.0	0.20	6.5
70	14.0	3.0	0.20	9.8
100 (LFI)	10.0	3.0	0.20	11.7
100 (HFI)	10.7	3.0	0.25	4.6
143	8.0	1.5	0.25	5.5
217	5.5	1.5	0.25	11.7
353	5.0	1.5	0.25	39.3
545	5.0	1.5	0.25	400.7
857	5.0	1.5	0.25	18182

Table 5.1: Technical characteristics of the 10 simulated Planck channels. Column two lists the FWHM assuming a Gaussian beam. Column three shows the pixel size in arcmin. Column four lists the fractional bandwidths of each channel. In column five the instrumental noise is ΔT (μK) per resolution element for 12 months of observation.

5.3 Simulations

Realistic Planck simulations were performed in order to test the ideas presented in § 5.2. The simulations mimic the main features of the future Planck mission. They are realistic in the sense that they include the latest available information about the different components (CMB, Galactic components, SZ effect and extra-galactic point sources) and that they reproduce the technical specifications of the different Planck channels (pixel sizes, antenna beams and noise levels). Table 5.1 shows the principal technical features of the simulated maps. We assume that the components can be factorized as a spatial template at 300 GHz with a known frequency dependence. The simulations were performed in patches of the sky of $12.8° \times 12.8°$. However, the method can be extended to all the sphere.

The CMB simulation was generated using the C_l's provided by the CMBFAST code (Seljak & Zaldarriaga, 1996) for a spatially-flat ΛCDM Universe with $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$ (Gaussian realization).

The simulations include four different Galactic emissions: thermal dust, spin-

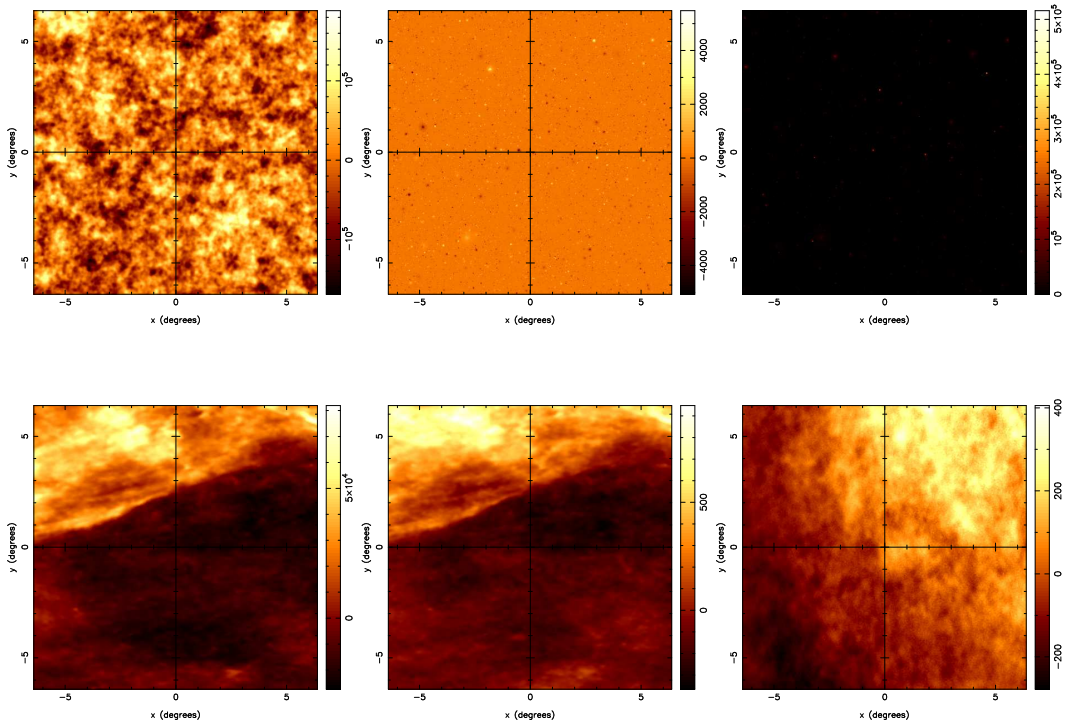

Figure 5.1: Components present in the simulation at 300 GHz. From left to right and from top to bottom the components are: CMB, kinetic SZ, thermal SZ, Galactic dust, free-free and synchrotron. The units of the maps are Jy/sr.

ning dust, free-free and synchrotron. Thermal dust was simulated using the template given by Finkbeiner et al. (1999). This model assumes that dust emission is due to two grey-bodies: a hot one with a dust temperature $T_D^{hot} \simeq 16.2K$ and a emissivity $\alpha^{hot} \simeq 2.70$, and a cold one with $T_D^{cold} \simeq 9.4K$ and $\alpha^{cold} \simeq 1.67$. These quantities are mean values.

For the free-free template we used one correlated with the dust emission in the manner proposed by Bouchet et al. (1996). The frequency dependence of the free-free emission is assumed to vary a $I_\nu \propto \nu^{-0.16}$, and is normalised to give an rms temperature fluctuation of $6.2\mu K$ at 53 GHZ.

The synchrotron spatial template has been produced using the all sky map

of Fosalba & Girardino [1], which is an extrapolation of the 408 MHz radio map of Haslam et al. (1982), from the original 1 deg resolution to a resolution of 5 arcmin. The small-scale structure has been extrapolated to 1.5 arcmin following a power-law power spectrum with an exponent of -3. The frequency dependence is assumed to be $I_\nu \propto \nu^{-0.9}$ and is normalized to the Haslam 408 MHz map. We include in our simulations the information on the changes of the spectral index as a function of electron density in the Galaxy. This template has been done combining the Haslam map with the Reich & Reich 1986 map at 1420 MHz and the Jonas et al. 1998 map at 2326 MHz, and can be found in the previous FTP site.

We have also taken into account the possible Galactic emission due to spinning dust grains, proposed by Draine & Lazarian 1998. This component could be important at the lowest frequencies of Planck (30 and 44 GHz) in the outskirts of the Galactic plane.

The extragalactic point source simulations follow the same cosmological model than the CMB simulations and correspond to the model of Toffolati et al. (1998). The thermal SZ effect was made for the same cosmological model. The cluster population was modelled following the Press-Schechter formalism (Press & Schechter 1974) with a Poissonian distribution in θ and ϕ. The simulated cluster population fits well all the available X-ray and optical cluster data sets (see Diego et al. 2001a for a discussion). The different components used for the simulation are shown at 300 GHz in figure 5.1. Figure 5.2 shows the simulated channels taking into account all the components and the antenna beam effect.

The simulation described above is the same used by Diego et al. (2001b). We have chosen this particular simulation in order to compare results with that work under the same conditions.

5.4 Results and discussion

5.4.1 Testing the methods

Before applying the methods presented in § 5.2 to the simulations described in § 5.3, let us illustrate how they work in a simplified case. Section § 5.2 introduced

[1] ftp://astro.estec.esa.nl/pub/synchrotron

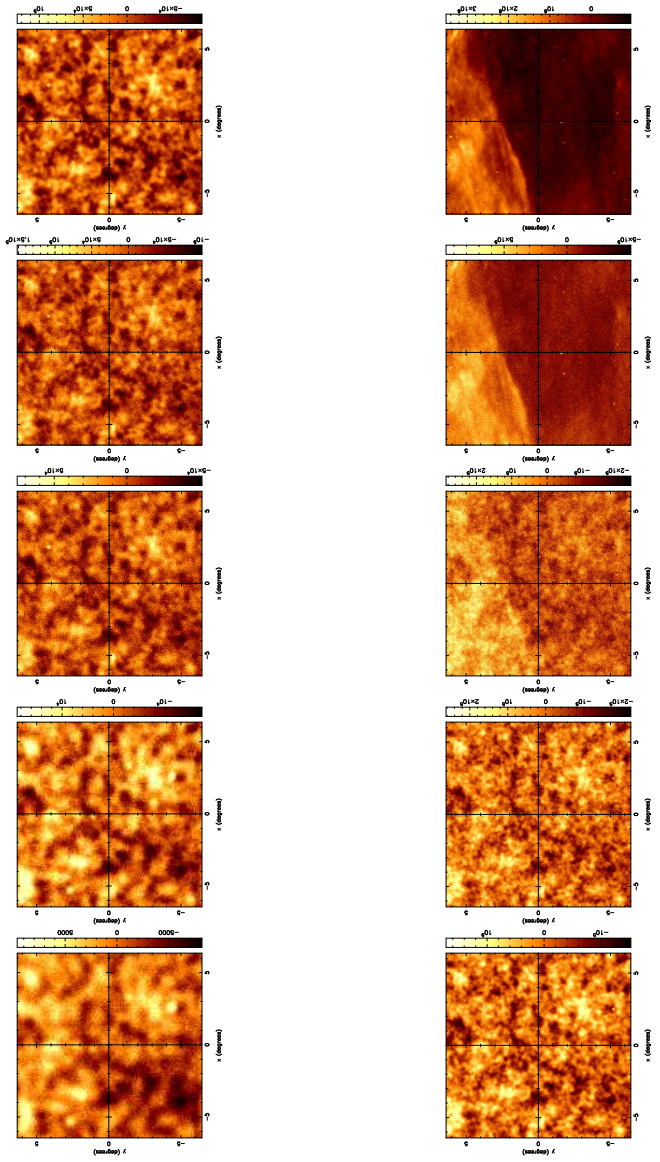

Figure 5.2: Simulated Planck channels. Each map corresponds to the same 12.8° × 12.8° area of the sky at the frequency of the channel. 30, 30, 44, 70, 100 (LFI), 100 (HFI), 143, 217, 353, 545 and 857 GHz channels are represented. The units of the maps are Jy/sr.

two different methods that include the multi-frequency information (optimal combination and multifiltering) and that can be implemented for two different kind

of filters (matched filters and adaptive filters). This makes four different possible combinations to filter the data; each way will have advantages and shortcomings with respect to the others.

From the methodological point of view, the main difference between the single filter method and the multifilter method is that the first one uses the multi-frequency information to give the *optimal starting point* for the filter, while multifilters produce the *optimal ending point* after filtering. In that sense, multifilters are more powerful than a single filter applied to an optimally combined map. The cost of this higher efficiency is an increase of the complexity of the filters and therefore of the computational time required to perform the data analysis. For example, to filter 10 frequency channels the multifilter method will need ~ 10 times more than the single filter method (because the last one only filters once).

Regarding the type of filter, matched filters give higher gains while adaptive filters reduce the number of errors in the detections (spurious sources). This is due to the fact that the adaptive filter definition of 'optimal filter' includes a constraint about the scale of the sources that is not present in matched filter design. This constraint characterises with more precision the sources but also restricts the minimisation of the variance in the filtered maps (that is, the final gain).

To give an example of the previous ideas, we performed a '10 frequencies' simple simulation. The simulation includes the same foregrounds and technical features than the simulations described in § 5.3, except for the SZ effect. Instead of the realistic SZ clusters we simulated 200 clusters with the same size ($r_c = 0.5$ pixels) and amplitudes uniformly distributed between 0 and the maximum amplitude of the clusters belonging to the realistic simulation. The simulated cluster have the frequency dependence of the SZ effect and the radial profile

$$\tau(x) = N\left(\frac{1}{\sqrt{r_c^2 + x^2}} - \frac{1}{\sqrt{r_v^2 + x^2}}\right) \tag{5.28}$$

where $N = r_v r_c/(r_v + r_c)$. Here r_c denotes the core radius and r_v is a 'cut-off scale' that can be interpreted as the virial radius of the cluster. The profile given by eq. (5.28) is a modified multiquadric profile that behaves as a β model for $x << r_v$ and decays quickly for $x >> r_v$.

The simulated maps were filtered using the four possible combinations of filters and methods presented in this work. The single filter method took about 1 minute of processing in a 700 MHz PC. The multifilter method took about 10 minutes of processing in the same computer. Processing times were slightly higher for the case of adaptive filters than in the case of matched filters. To compute the filters and the combinations of the maps the low resolution channels were re-binned to the resolution of the highest resolution channel (1.5 arcmin). Figure 5.3 shows the filters (in Fourier space) employed in the multifilter method. The adaptive filters are represented by solid lines, whereas matched filters are represented by dotted lines. Filters, specially those used in low frequency channels, are quite complicate and show several peaks at different wave numbers q that correspond to the different scales the filters are trying to identify on the images. Conversely, the filters used in the single filter method are rather simple. They are represented in figure 5.4. The differences between matched and adaptive filters in the single filter method are small. This indicates that there will be few differences in the results of both filters. In the case of the multifilters the differences are greater, in particular for the 217 GHz channel.

The results of the test are shown in table 5.2. After filtering, the sources were detected by looking for peaks above a 4σ threshold and then compared with a catalogue of the original simulated clusters. The number of detections is higher in the multifilter case. this is not surprising since by definition the multifilter method is more powerful than the single filter method. If we detect over a higher threshold the difference between the two methods increase. for example, over 5σ the matched multifilters produce 90 detections whereas the matched filter in the single filter method gives only 78 (a 15 percent less).

A detected peak is considered a spurious detection if the distance to the closest object in the original catalogue is greater than 1.5 pixels. The number of spurious detections is strikingly low in all cases. The adaptive filter seems to work better than the matched filter (0 spurious detections instead of 1 in the single filter method). Of all the components present in the simulation, point sources are the most likely to produce spurious detections due to the similar scale of sources and SZ clusters. However, the frequency dependence of the SZ effect greatly reduces the probability of contamination.

METHOD	number of detections	number of spurious	mean offset (pixels)	$\bar{b}_A(\%)$	$\bar{e}_A$ (%)
Single Filter/matched	110	1	0.41	33.3	33.3
Single Filter/adaptive	113	0	0.38	32.5	32.5
Multifilter/matched	116	1	0.42	29.8	29.9
Multifilter/adaptive	109	1	0.57	34.4	34.5

Table 5.2: Results of the test for 10 frequency channels and clusters of equal size ($r_c = 0.5$ pixels). The detection was performed over the 4σ threshold in all cases. The filtering method is listed in the first column. Second column shows the number of true detections found over the 4σ detection threshold. The third column indicates the number of spurious sources detected in each case. The fourth column lists the mean error in the determination of the position of the detected sources. Column five lists the mean relative bias in the determination of the amplitude, defined as $\bar{b}_A = 100\langle(A - A_0)/A_0\rangle$. Column six shows the mean relative error in the determination of the amplitude, defined as $\bar{e}_A = 100\langle|A - A_0|/A_0\rangle$.

The position of the sources was determined with errors below the pixel size (1.5 arcmin). The error in the determination is large ($\sim 35\%$). Table 5.2 shows that this error is due to a systematic bias. To explain this bias let us consider the case of matched filter. The filter normalization is given by integral $a = \int_0^\infty dqq\tau^2/P$, being τ the source profile in Fourier space and P the power spectrum of the background. However, when we analyse a pixelised patch of the sky we do not have information about the power spectrum in all the wave numbers q. A pixellised image is limited by a minimum value q_{min} and a maximum q_{max}. Therefore, the normalization we calculate is $a' = \int_{q_{min}}^{q_{max}} dqq\tau^2/P$. For the case of images with the same size and pixel scale than in our simulations and a multiquadric profile τ with $r_c = 0.5$ pixels, the normalisations can differ from a 10% for spectral indexes of the background $\gamma = 1$ to a 60% for spectral indexes of the background $\gamma = 3$. The case for adaptive filters is more complicated but qualitatively similar. This bias is independent of the source flux and can be calibrated using simulations.

Figure 5.5 shows the contribution of each channel to the source amplitude estimation for each method. The channels with larger contribution are the 143 GHz and the 353 GHz. This is not surprising since they are the channels with more SZ contribution. The 100 GHz HFI channel has more contribution than the 100 GHz

LFI because of its better signal to noise ratio. The single filter method puts more emphasis in the 100 GHz and less in the 143 GHz channels than the multifilter method. Contributions from the 30 GHz, 217 GHz and 857 GHz channels are negligible. However, that does not mean that these channels do not contribute to the filter construction. For example, if we repeat the analysis with only the 5 channels with more contribution (70, 100 LFI, 100 HFI, 143 and 353 GHz), the number of detections reduces in $\sim 10\%$ and the number of spurious detections raise to 6 (at 4σ detection threshold) for the single filter method case (for the two filters) and to 9 (adaptive filter) and 5 (matched filter) for the multifilter method. Even the 217 GHz channel is important; repeating the analysis with all channels except for the 217 GHz increases the error in the determination of the amplitudes and also increases the number of spurious detections at low σ thresholds.

We can summarise the conclusions of the test as follows:

- Single Filter method is faster than the multifilter method.
- Multifilter method is more powerful in the detection and estimation of cluster parameters.
- Multifrequency information reduces the number of spurious detections. Therefore, it is not critical to use an adaptive filter to that end. The matched filter allows to detect more sources.
- Some channels contribute more than others to the analysis, but all of them carry valuable information; the analysis should include all the available information.

5.4.2 Results for realistic simulations

Taking into account the insights provided by the test presented in the last subsection, we are now prepared to confront the analysis of realistic simulations. The main difference with respect to the previously performed test is that in the realistic case clusters have different, non a priori known sizes. In Chapter 3 a way to deal with this problem was proposed:

- We choose a trial core radius r_c and construct the correspondent filters.

- We filter the data with the filters, *varying the scales* R_ν *of eq.* (5.3). This is performed by substituting $\tilde{\psi}(q)$ by $\tilde{\psi}(qx)$, where $x = R'_\nu/R_\nu$ is simply a dilation factor.

- In the case of scale-adaptive filters, condition (3) (see § 5.2.1) imply that, if the r_c tested corresponds with the true core radius of the cluster, the coefficients will be maximum when $x = 1$. If this is not true, the trial r_c is discarded.

- In the case of matched filters, condition (3) no longer exits. However, a similar criterion can be used: the best performance of the filter will happen when the scale of the filter and the scale of the clusters is the same. Therefore, if the maximum of the signal to noise ratio of the coefficients is not maximum for $x = 1$, that means that the chosen radius r_c is not the best choice.

- We repeat the process with as many different values of r_c as we want.

In Chapter 3 the previous method has been successfully used on single frequency maps containing simulated multiquadric profiles and different kind of backgrounds.

Following the results of the test in the last subsection, we choose a matched multifilter to perform the analysis. After applying our method to the simulations we detect the clusters by looking for sets of connected pixels above a certain threshold. At the 3σ level (regions with 5 or more connected pixels) we are able to detect 63 cluster candidates. Of these 63 candidates 62 correspond with real clusters. The spurious detection appear in one of the borders of the image and therefore can be considered as a border effect. If we lower to 2σ we detect 284 candidates. Of these 284, 257 were real clusters and 27 were spurious detections. The mean error in the determination of the position of the clusters was ~ 1 pixel.

With the multi-scale analysis we were able to determine the core radii of the clusters with a mean absolute error of 0.30 pixels. The mean bias in the determination of the core radii was -0.15 pixels. Since pixelisation effects are expected to corrupt structures whose typical scale of variation is much smaller than the pixel size, all clusters with a core radius small enough will not follow the multiquadric profile. Most of the clusters have very small core radii and therefore can be considered as point-like sources. To detect these clusters, a Gaussian profile

(corresponding to the beam at each channel) was assumed instead of the multiquadric given by eq. (5.28). The separation between clusters that are considered as point sources and extended sources was set in $r_c = 0.4$ pixel because if r_c is below this limit the FWHM of the multiquadric profile is of the order of a pixel. All the clusters that were detected as point sources were considered to have $r_c = 0.1$ pixel. The bias in the determination of the core radius above mentioned is due to the asymmetry introduced when assigning $r_c = 0.1$ pixel to all clusters whose core radius is less than 0.4 pixel. Figure 5.6 show the resulting maps after filtering with different scales (point source in the left, $r_c = 1.0$ pixel in the middle and $r_c = 2.5$ pixel in the right side). Small clusters stand out mainly in the left panel, while large clusters are emphasized in the right panel.

Figure 5.7 shows the detected amplitudes of the 80 brightest detected clusters *vs.* the true amplitudes. The brightest sources in the graph were recovered with relative errors of $\sim 30\%$. The error goes worse as the true amplitude decreases. For the weakest clusters, the detected amplitude tends to a horizontal line. This bias is due to two factors: first, only the faint clusters that are overamplificated are able to reach the detection threshold; second, the detection is performed looking for peaks in the data and, hence, those clusters that are by chance enhanced by residual fluctuations of the background (noise) are more likely detected.

The catalogue of detected sources was complete over 0.17 Jy (at 300 GHz). Below this limit the completeness decreases gradually. There are a few very faint clusters that are detected with this method (fluxes of few tens of mJy), but the practical detection limit is $40 - 50$ mJy (at 300 GHz). A number of 62 detections in our small sky patch of $(12.8^\circ)^2$ means that we expect ≈ 15000 detections in all sky.

The performance of the other filtering techniques proposed in this work was compared to the performance of the matched multifilter. For the comparison, only the point-like clusters (that dominate the number counts) were considered. The scale-adaptive multifilter, as seen in § 5.4.1, performs worse than the matched multifilter because their main advantage, that is the removal of spurious sources, has already been accomplished by the multi-frequency information included in the filters. Since the clusters we are trying to detect are very faint, what counts here is the gain. At 3σ, the scale-adaptive multifilter detects a 40 percent less

clusters than the matched multifilter. Single filters are slightly less sensitive than multifilters, producing a 20 percent less detections at 3σ than with the matched multifilter in the case of both matched and scale-adaptive filters. In the case of single filters, matched and scale-adaptive perform very similarly (as was expected from figure 5.4, where both filters look alike).

5.5 Discussion

The results obtained with the matched multifilter are comparable at 3σ and better at 2σ than those obtained by Diego et al. (2001b). To perform their method it is necessary to previously perform a careful cleaning of the maps. This cleaning include point source removal with a MHW technique (Cayón et al. 2000, Vielva el al. 2001a), dust substraction using the 853 GHz map and CMB substraction in Fourier space up to a certain wavenumber limit using the 217 GHz map. In order to compare, we performed the same cleaning method and then applied the matched multifilter. We obtained a 7% more detections at 3σ detection threshold than before. We also detected 7% more clusters and the number of spurious detections decreased in a 12% at the 2σ level. The improvement is small taking into account all the cleaning work that is necessary to obtain it. Moreover, it indicates that the multifilter scheme is powerful and robust enough to successfully deal with contaminants such as point sources, dust and the CMB signal.

The filtering techniques described in this work have proven to be useful tools for the detection of signals that are present in multi-frequency data and that are known to have a prescribed frequency dependence. The only other assumption about the signals is their profile; any other quantities of interest are directly calculated from the data. We have assumed in this work that the sources have circular symmetry, but the method can be extended to deal with asymmetric structures. Future work will take this point into account. We have applied these techniques to the problem posed by the presence of Sunyaev-Zel'dovich effect in CMB maps. The best results are obtained with matched multifilters. The detection level of the method is comparable to other single-component separation methods, but the number of assumptions and previous data processing is reduced to a minimum. Due to the weakness of the signals that we tried to detect, the determination

of the amplitude of the clusters is biased; this is the strongest limitation of the method. Therefore, the method should be complemented with a posteriori aids (for example data fitting, etc.) in order to improve the amplitude estimation of weak clusters. Finally, the filters here presented can be used in other fields of Astronomy, such as X-ray observations, large-scale structure detection, etc.

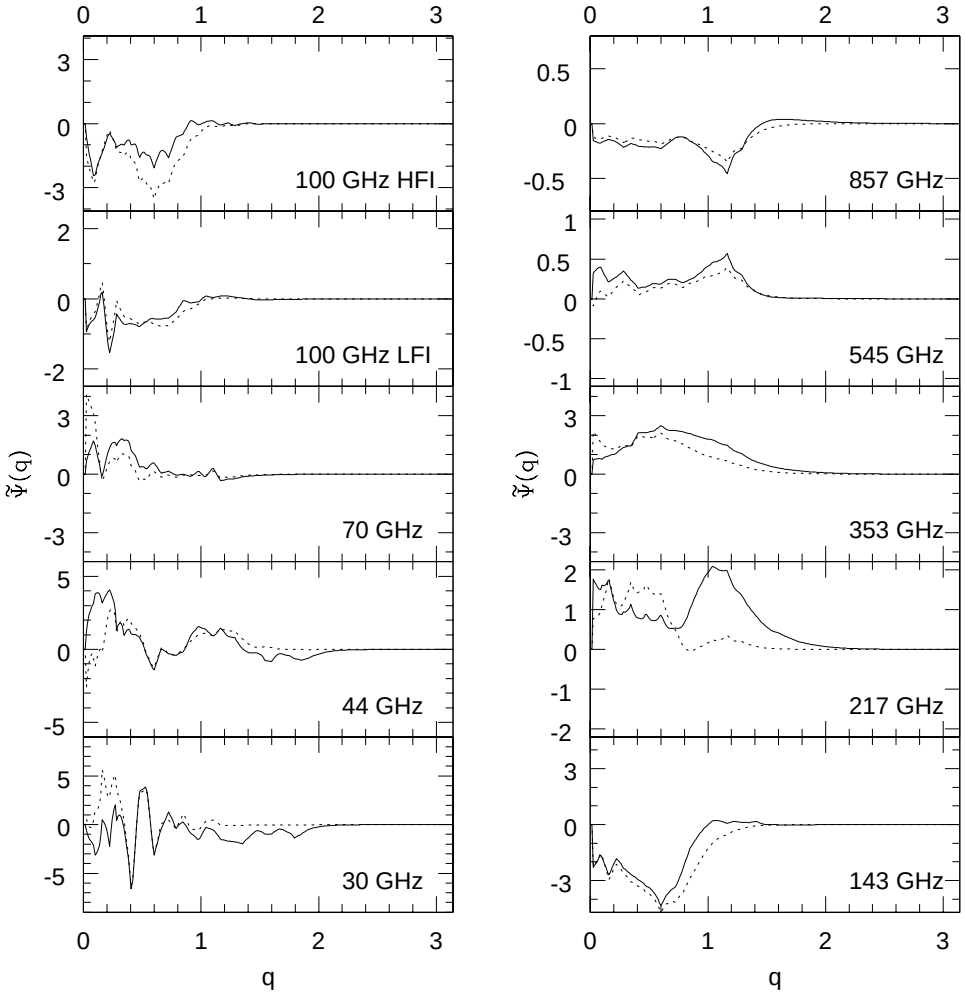

Figure 5.3: Filters used to test the multifilter method. Adaptive filter (*solid line*) and matched filter (*dotted line*) are represented in Fourier space for each frequency channel.

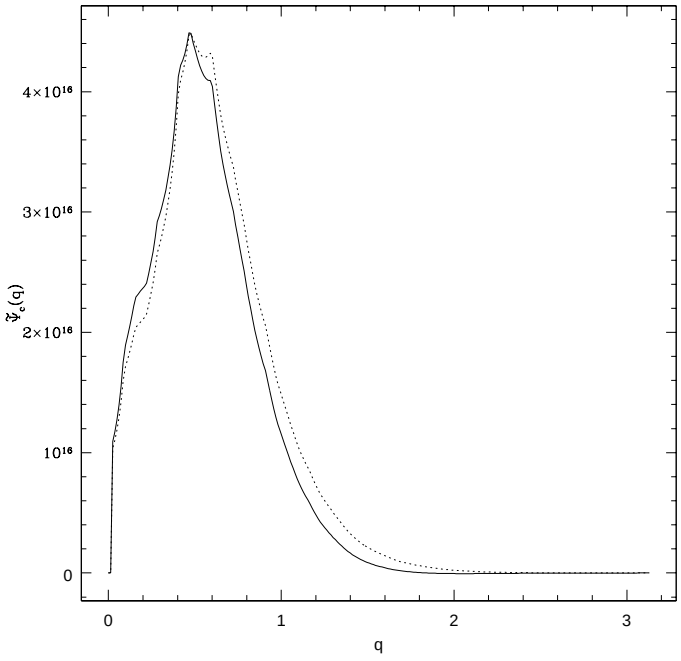

Figure 5.4: Filters used to test the single filter method. The adaptive filter (*solid line*) and the matched filter (*dotted line*) are represented in Fourier space.

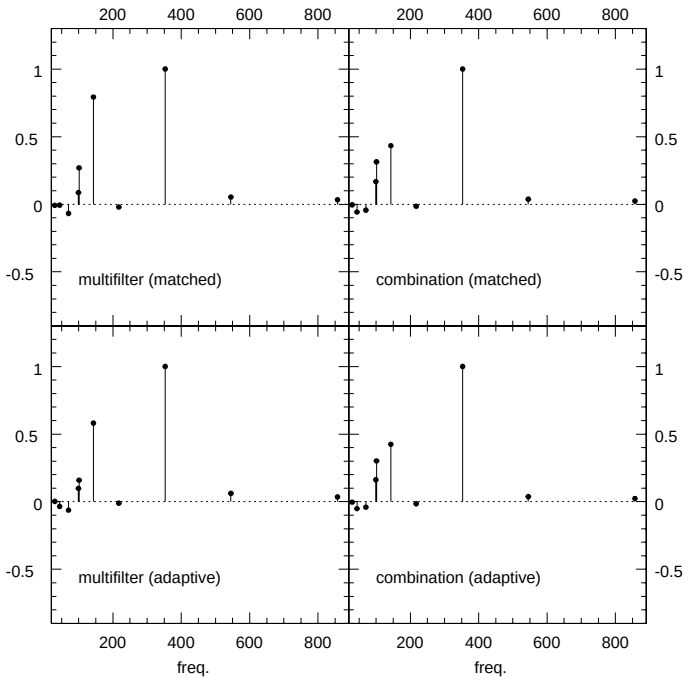

Figure 5.5: Contributions of the different channels to the source amplitude estimation. The contributions are normalised so that the channel with maximum contribution has contribution 1.

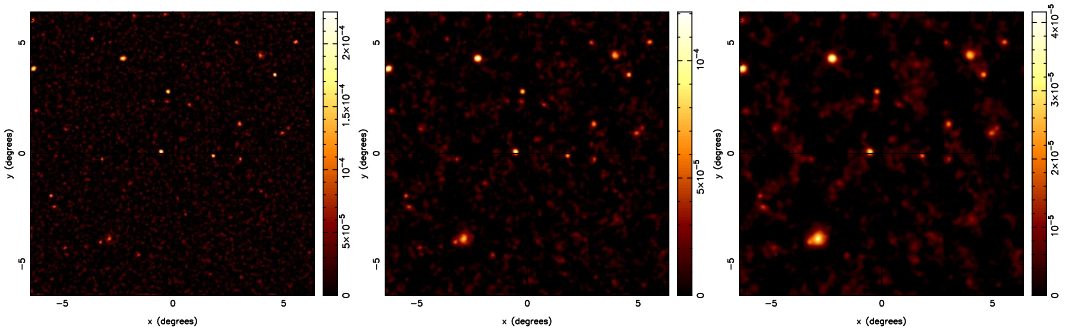

Figure 5.6: Multifiltered data at several scales. Left panel shows the output after filtering with a matched multifilter that considers that clusters are point-like. Center panel shows the output after filtering with a matched multifilter for a modified multiquadric profile with $r_c = 1.0$ pixel. Right panel shows the output after filtering with a matched multifilter for a modified multiquadric profile with $r_c = 2.5$ pixel. Note how different structures are enhanced at different scales. The units of the maps are in $\Delta T/T$.

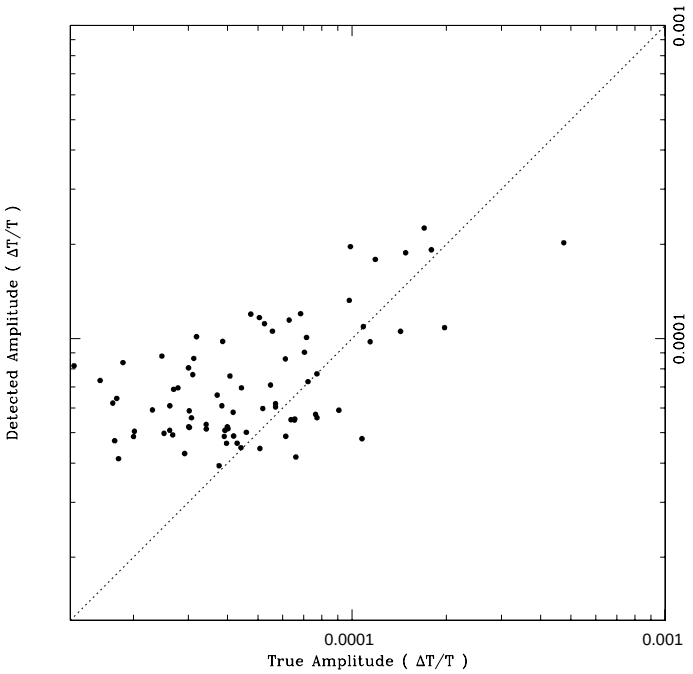

Figure 5.7: Amplitudes of the 80 brightest detected clusters *vs.* the true amplitudes. The dotted line corresponds to the perfect situation True Amplitude = Detected Amplitude.

Chapter 6

Conclusions and future work

The work presented in this thesis deals with the application of filtering techniques to component separation of CMB data. The relevance of CMB studies has been pointed out in the introduction of this thesis; it is necessary to develop efficient and robust component separation techniques in order to identify the different physical processes that are present at microwave wavelengths.

Filtering techniques are well suited to isolate (that is, detect) signals that have a known characteristic behaviour in Fourier space. Examples of such signals are extragalactic point sources and galaxy clusters in CMB maps. The initial question of this work was, 'is there any optimal way to filter a given data set in order to recover certain signals?'

Chapter 2 introduced the use of scale-adaptive filters that optimize the detection/extraction of sources on a background. As a first approach, it is assumed that such sources are described by a spherical (central) profile and that the background is represented by a homogeneous & isotropic random field. The scale-adaptive filters are derived so that

- They are *unbiased* estimators of the amplitude of the sources.
- They are *efficient* in the sense that they minimize the variance of the filtered data.
- They give a maximum at a certain scale R_o related to the characteristic scale of the sources. This last condition leads straightforwardly to

a reduction of the probability of false detections (improvement of the *reliability*).

A fully n-dimensional treatment of scale-adaptive filters is done, making emphasis in astrophysical applications for spectra, images and volumes, for the cases of exponential and Gaussian source profiles and scale-free power spectra to represent the background. The scale-adaptive filters were compared with the Gaussian filter and the Mexican Hat Wavelet, surpassing them both in number of detections and reliability.

Chapter 3 deals with the detection of more general profiles in two-dimensional images. The profile of a galaxy cluster at different wavelengths, for example, can be approximated by functions belonging to the *multiquadric* family, i. e. $\tau(x) = \left[1 + (x/r_c)^2\right]^{-\lambda}, \lambda \geq \frac{1}{2}, x \equiv |\vec{x}|$. This chapter included simple simulations of 'clusters' at microwave and X-ray frequencies embedded in different noises. The scale-adaptive filters are able to detect clusters with high reliability. Besides, it is shown that the scale-adaptive filters can be used to determine not only the amplitude of the sources but also their size, using the third condition mentioned above.

Chapter 4 apply scale-adaptive filters to optimize the detection and extraction of point sources from a one-dimensional array of time-ordered data such as the one that will be produced by the future 30 GHz LFI28 channel of the ESA Planck mission. This simulation is fully realistic and include all the foregrounds and instrumental effects to be present in Planck channels. The scale-adaptive filter produced the following results:

- At a 4σ detection level 224 sources over a flux of 0.88 Jy are detected with a mean relative error (in absolute value) of 21% and a systematic bias of -7.7%.
- The position of the sources in the sky is determined with errors inferior to the size of the pixel.
- The catalogue of detected sources is complete at fluxes ≥ 4.3 Jy.

- The number of spurious detections is less than a 10% of the true detections (*high reliability*).

The results were compared with the ones obtained by filtering with a Gaussian filter and a Mexican Hat Wavelet of width equal to the scale of the sources. The adaptive filter outperforms the other filters in all considered cases. We conclude that optimal adaptive filters are well suited to detect and extract sources with a given profile embedded in a background of known statistical properties. In the Planck case, they could be useful to obtain a real-time preliminary catalogue of extragalactic sources, which would have a great scientific interest, e. g. for follow-up observations.

Chapter 5 introduce multifilters and single filters that optimize the detection of clusters on microwave maps. The Sunyaev-Zel'dovich effect generates source profiles that can be approximately represented by multiquadric profiles. The filters are applied to small patches (corresponding to 10 frequency channels) of the sky such as the ones that will produce the future ESA Planck mission. The method presented in this chapter is a generalization of the scale-adaptive filters (see chapter 2) that includes all the cross-correlation between maps as well as the frequency dependence of the Sunyaev-Zel'dovich effect. In a square patch of the sky of $12.8^\circ \times 12.8^\circ$ the matched multifilter detects 62 clusters. That means a number of detections of ≈ 15000 in whole sky. The catalogue will be complete over fluxes $S > 170$ mJy (at 300 GHz). The method here presented only assumes the profile of the cluster, allowing to determine both the amplitude (flux) and the scale (core radius) of the clusters in a robust manner.

Besides the applications described in this thesis, scale-adaptive filters may be used in many other fields of Astronomy and science in general. Examples of such applications are the analysis of spectra, detection and extraction of large-scale structure, detection of X-ray clusters, etc. The scale-adaptive filters will be generalised in the future for non-symmetric profiles, thus allowing to detect substructure in galaxy clusters or to include more realistic beam profiles in the calculations.

One of the most promising work lines for the future will be the generalisation of

the treatment presented in Chapter 4 to include information from different Planck instruments in the TOD. For example, the final 30 GHz map will have contributions from several radiometers. In Chapter 4 only one of them was taken into account. By taking all of them into account it is expected to increase the SNR of the sources. Besides, TOD can be used to detect Sunyaev-Zel'dovich clusters in a similar way to described in Chapter 5. In this case, not only different instruments but also different frequency channels must be taken into account.

Another possibility is the combination of scale-adaptive filters with other component separation techniques. In Vielva *et al.* (2001b) the MHW was combined with MEM in order to clean up the maps from bright sources as a previous step to maximum entropy analysis. Similarly, SAFs can be combined with other techniques to improve the final results.

Appendix A

Derivation of the formula for the Scale-Adaptive Filters

Let us consider a set of n-dimensional data:

$$y(\vec{x}) = s(x) + n(\vec{x}), \tag{A.1}$$

where $\vec{x} = (x_1, x_2, \cdots, x_n)$ is a position vector and $x = |\vec{x}|$. Therefore, $s(x)$ represents a source with central symmetry whose *profile* $\tau(x)$ can be defined as $\tau(x) = s(x)/A$, $A \equiv s(0)$. The background $n(\vec{x})$ is assumed to be statistically homogeneous and isotropic random field with mean value $\langle n(\vec{x})\rangle = 0$ and characterized by the power spectrum $P(q)$, $q \equiv |\vec{q}|$:

$$\langle n(\vec{q}) n^*(\vec{q}\,')\rangle = P(q)\delta^n(\vec{q} - \vec{q}\,'), \tag{A.2}$$

where $n(\vec{q})$ is the n-dimensional Fourier transform, symbol n^* represents the complex conjugate of n, $\vec{q}$ is the wave vector and δ^n is the n-dimensional Dirac distribution. Throughout all this work the Fourier transform convention is the same as in eq. (1.21),

$$\begin{aligned} f(\vec{q}) &= \frac{1}{(2\pi)^{n/2}} \int d\vec{x} f(\vec{x}) e^{i\vec{q}\vec{x}}, \\ f(\vec{x}) &= \frac{1}{(2\pi)^{n/2}} \int d\vec{q} f(\vec{q}) e^{-i\vec{q}\vec{x}}. \end{aligned} \tag{A.3}$$

Note that, with convention (A.3), any convolution can be expressed as

$$f \otimes g(\vec{b}) = \int d\vec{x} f(\vec{x}) g(\vec{b} - \vec{x}) =$$

$$\frac{1}{(2\pi)^n}\int\int d\vec{q}\,d\vec{q}\,'f(\vec{q})g(\vec{q}\,')e^{-i\vec{q}\,'\vec{b}}\int d\vec{x}e^{-i\vec{x}(\vec{q}-\vec{q}\,')} =$$
$$\frac{1}{(2\pi)^n}\int\int d\vec{q}\,d\vec{q}\,'f(\vec{q})g(\vec{q}\,')e^{-i\vec{q}\,'\vec{b}}(2\pi)^n\delta^n(\vec{q}-\vec{q}\,') =$$
$$\int d\vec{q}f(\vec{q})g(\vec{q})e^{-i\vec{q}\vec{b}}. \tag{A.4}$$

This justifies the result in eq. (1.23). Let us introduce a spherical (centrally symmetric) filter, $\Psi(\vec{x};\vec{b},R)$, dependent on $n+1$ parameters:

$$\Psi(\vec{x};R,\vec{b}) = \frac{1}{R^n}\psi\left(\frac{|\vec{x}-\vec{b}|}{R}\right), \tag{A.5}$$

where $\vec{b}$ defines a translation whereas R defines a scaling. The filtered field $w(R,\vec{b})$ is

$$w(R,\vec{b}) = \int d\vec{x}\,y(\vec{x})\Psi(\vec{x};\vec{b},R). \tag{A.6}$$

Using the same reasoning that in eq. (A.4), the filtered field can be expressed as a product in Fourier space:

$$w(R,\vec{b}) = \int d\vec{q}\,e^{-i\vec{q}\vec{b}}y(\vec{q})\psi(Rq). \tag{A.7}$$

A simple calculation -taking into account eqs. (A.1) and (A.2)- gives the average at the origin $\vec{b}=0$, $\langle w(R,\vec{0})\rangle$, and the variance, $\sigma_w^2(R) = \langle w^2(R,\vec{b})\rangle - \langle w(R,\vec{b})\rangle^2$, of the filtered field:

$$\langle w(R,\vec{0})\rangle = \alpha\int dq\,q^{n-1}\,s(q)\psi(Rq),$$
$$\sigma_w^2(R) = \alpha\int dq\,q^{n-1}\,P(q)\psi^2(Rq), \tag{A.8}$$

where $q=|\vec{q}|$, $\alpha = 2\pi^{n/2}\Gamma^{-1}(n/2)$ and the limits in the integrals go from 0 to ∞. The conditions that the filter (A.5) must satisfy are:

$$\exists\ R_o \text{ so that } \left.\frac{d\langle w(R,\vec{0})\rangle}{dR}\right|_{R=R_o} = 0 \tag{A.9}$$

$$\langle w(R_o,\vec{0})\rangle = s(0) = A \tag{A.10}$$

$$\sigma_w^2(R_o) \text{ is minimum with respect to } \psi \tag{A.11}$$

Condition (A.9) means that ψ is a *optimal scale estimator*. It gives a constraint on ψ given by

$$\frac{d\langle w(R,\vec{0})\rangle}{dR} = \frac{d}{dR}\left[A\alpha \int dq q^{n-1}\tau(q)\psi(qR)\right]. \tag{A.12}$$

Changing variables $x \equiv qR$ we have

$$\frac{d\langle w(R,\vec{0})\rangle}{dR} = A\alpha \frac{d}{dR}\left[\frac{1}{R^n}\int dx x^{n-1}\tau\left(\frac{x}{R}\right)\psi(x)\right] =$$
$$A\alpha\left[-n\frac{1}{R^{n+1}}\int dx x^{n-1}\tau\left(\frac{x}{R}\right)\psi(x) + \frac{1}{R^n}\int dx x^{n-1}\psi(x)\frac{d\tau(x/R)}{dR}\right] \tag{A.13}$$

but $\frac{d\tau(x/R)}{dR} = \frac{d\tau(u)}{dR} = \frac{d\tau(u)}{du}\frac{du}{dR} = -\frac{x}{R^2}\frac{d\tau(u)}{du}$, where we have changed $u \equiv x/R$. Now, note that reversing the variable change $q = x/R$ we have that $u = q$ and then

$$\frac{d\langle w(R,\vec{0})\rangle}{dR} = -\frac{A\alpha}{R}\int dq q^{n-1}\tau(q)\psi(qR)\left[n + \frac{q}{\tau}\frac{d\tau}{dq}\right] =$$
$$-\frac{A\alpha}{R}\int dq q^{n-1}\tau(q)\psi(qR)\left[n + \frac{d\ln\tau}{d\ln q}\right] = 0. \tag{A.14}$$

So we have a first constraint:

$$\int dq q^{n-1}\tau(q)\psi(qR_o)\left(n + \frac{d\ln\tau}{d\ln q}\right) = 0. \tag{A.15}$$

The condition (A.10) means that the filter is an *unbiased* estimator of the amplitude of the sources and straightforwardly gives the constraint:

$$\int dq\, q^{n-1}\,\tau(q)\psi(qR_o) = \frac{1}{\alpha}. \tag{A.16}$$

To include condition (A.11), that reflects the fact that the filter is an *efficient* estimator of the amplitude of the sources, we introduce a couple of Lagrangian multipliers $\bar{\lambda}$ and $\bar{\mu}$ and define

$$\mathcal{L}(\psi) = \sigma_w^2(\psi) + \bar{\lambda}\left[\int dq\, q^{n-1}\,\tau(q)\psi(qR_o) - \frac{1}{\alpha}\right]$$
$$+\bar{\mu}\left[\int dq q^{n-1}\tau(q)\psi(qR_o)\left(n + \frac{d\ln\tau}{d\ln q}\right)\right]. \tag{A.17}$$

Taking into account the expression (A.8) we find the extrema of $\mathcal{L}$ by taking variations with respect to ψ and making the result equal to zero. For the sake of

simplifying the notation, we will omit the q dependence of $P(q)$, $\tau(q)$, $\psi(qR_o)$, etc.

$$\frac{\mathcal{L}(\psi+\delta\psi)}{\alpha}-\frac{\mathcal{L}(\psi)}{\alpha}=2\int dqq^{n-1}P\psi\delta\psi^*+$$
$$\lambda\int dqq^{n-1}\tau\delta\psi^*+\mu\int dqq^{n-1}\tau\delta\psi^*\left(n+\frac{d\ln\tau}{d\ln q}\right), \tag{A.18}$$

where $\lambda=\bar{\lambda}/\alpha$, $\mu=\bar{\mu}/\alpha$. The right hand side of the previous equation is zero iff

$$0=2P\psi+\lambda\tau+\mu\tau\left(n+\frac{d\ln\tau}{d\ln q}\right). \tag{A.19}$$

Therefore,

$$\psi(qR_o)=\frac{\tau(q)}{\alpha P(q)}\left[\lambda+\mu\left(n+\frac{d\ln\tau}{d\ln q}\right)\right], \tag{A.20}$$

where the multipliers have been redefined for convenience. Now, to obtain the values of λ and μ we go back to constraints (A.15) and (A.16). The first of them gives the equation

$$a\lambda+(na+b)\mu=1, \tag{A.21}$$

where as the second constraint gives

$$(na+b)\lambda+(n^2a+nab+c)\mu=0, \tag{A.22}$$

and where a, b and c are the integrals:

$$\begin{aligned} a &\equiv \int dq\, q^{n-1}\,\tfrac{\tau^2}{P}, \\ b &\equiv \int dq\, q^{n-1}\,\tfrac{\tau}{P}\tfrac{d\tau}{d\ln q}, \\ c &\equiv \int dq\, q^{n-1}\,\tfrac{1}{P}\left[\tfrac{d\tau}{d\ln q}\right]^2. \end{aligned} \tag{A.23}$$

The solution to the system (A.21)-(A.22) is:

$$\lambda=\frac{n^2a+2nb+c}{\Delta},\quad \mu=-\frac{na+b}{\Delta},\quad \Delta=ac-b^2. \tag{A.24}$$

Substituting (A.24) in (A.20) we reach the formula for the scale-adaptive filter:

$$\tilde{\psi}(q)\equiv\psi(R_oq)=\frac{1}{\alpha}\frac{\tau(q)}{P(q)}\frac{1}{\Delta}\left[nb+c-(na+b)\frac{d\,ln\tau}{d\,lnq}\right] \tag{A.25}$$

If instead of using the three conditions (A.9)-(A.11) we only use the two last ones, the only constraint to the minimization is (A.15) and in eq. (A.17) only

the multiplier $\bar{\lambda}$ appear. Then the solution does not depend on the logarithmic derivative of τ and the filter that is obtained is:

$$\psi(R_o q) = \frac{1}{\alpha} \frac{\tau(q)}{aP(q)}, \tag{A.26}$$

where a is the same integral that in eq. (A.23). This kind of filter is called *matched filter*.

Appendix B

Resumen de la tesis en castellano

B.1 Introducción

B.1.1 La radiación cósmica del fondo de microondas

La Radiación Cósmica del Fondo de Microondas (RCFM) es, como su propio nombre indica, una radiación muy débil y uniforme que continuamente llega a la Tierra desde todas las direcciones. Las dos características más notorias de la RCFM son que su espectro electromagnético corresponde exactamente con el de un cuerpo negro cuya temperatura es $T_o = 2.73K$ y que su intensidad es extraordinariamente homogénea e isótropa. Hasta donde se sabe, sólo procesos en los que haya equilibrio termodinámico son capaces de producir espectros de cuerpo negro, lo cual sugiere que la RCFM tuvo un origen térmico. Por otra parte, la homogeneidad e isotropía de la RCFM implican que los procesos que la originaron afectaron a todo el Universo observado. Conjuntamente, estas dos propiedades pdrían ser explicadas fácilmente si el Universo hubiera atravesado en algún momento de su pasado una fase de equilibrio termodinámico en la cual las condiciones de densidad y temperatura alcanzaran valores muy elevados. Esto es precisamente lo que predicen los modelos cosmológicos basados en la 'Gran Explosión' (*'Big Bang'*). De hecho, la RCFM ya había sido predicha dentro del marco de la nucleosíntesis primordial (Gamow 1948a,b, Alpher & Herman 1948) años antes de su descubrimiento por parte de Penzias y Wilson en 1964 (Penzias & Wilson 1965).

Origen de la RCFM

Inmediatamente después de la Gran Explosión las condiciones de temperatura, densidad y presión en el Universo eran extremas. En esas condiciones, la materia se encontraba totalmente ionizada y electrones, bariones y fotones formaban un 'fluido' que se mantenía en equilibrio termodinámico gracias a las continuas interacciones entre las partículas que lo formaban. A medida que el Universo se expande, la temperatura y la presión disminuyen, de modo que, después de $\sim 3 \times 10^5$ años ($z \sim 1000$), la temperatura había bajado a $T \sim 3000K$. Por debajo de esta temperatura, los electrones y los protones se combinaron para formar átomos neutros, dejando libertad a los fotones para propagarse. El Universo se volvió *transparente*. Tras este momento, llamado *época de la recombinación* o del *desacoplo*, materia y radiación siguieron caminos independientes. La inmensa mayoría de aquellos fotones jamás volvieron a interaccionar con la materia, manteniendo así el espectro de cuerpo negro. La temperatura de la RCFM ha disminuido desde la época de recombinación de acuerdo con $T(z) = (1+z)T_o$, siendo $T_o = 2.73K$ la temperatura actual de la RCFM. Por lo tanto, la RCFM puede considerarse como la imagen más antigua del Universo, ya que previamente al desacoplo el Universo era totalmente opaco.

Observables

La isotropía de la RCFM no es totalmente perfecta. En ella aparecen pequeñas irregularidades en intensidad cuya magnitud típica es de una cienmilésima parte de la intensidad media de la RCFM. Dichas irregularidades reciben el nombre de *anisotropías* de la RCFM. Las anisotropías de la temperatura de la RCFM se describen por un campo aleatorio 2-dimensional $\frac{\Delta T}{T}(\vec{n}) \equiv \frac{T(\vec{n})-T_o}{T_o}$, donde $\vec{n}$ es un vector unitario sobre la esfera. Este campo suele escribirse como una expansión en armónicos esféricos:

$$\frac{\Delta T}{T}(\vec{n}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\vec{n}) \,, \tag{B.1}$$

En dicha expansión, valores bajos de ℓ corresponden a escalas angulares grandes, mientras que valores altos reflejan las anisotropías a pequeña escala angular. Los coeficientes a_{lm} son variables aleatorias independientes de media cero. Si las fluc-

tuaciones de temperatura son isotrópas en sentido estadístico, la varianza de los a_{lm} es independiente de m:

$$< a_{\ell m} a^*_{\ell' m'} >= C_\ell \delta_{\ell\ell'} \delta_{mm'} \,, \tag{B.2}$$

donde los promedios se definen sobre conjuntos estadísticos. El conjunto de valores C_ℓ constituye el *espectro de potencias angular* o, simplemente, el espectro de potencias de la radiación.

Si las fluctuaciones de la RCFM son gaussianas, el espectro de potencias angular describe por completo el campo de temperaturas. El espectro de potencias se suele representar como la potencia por intervalo logarítmico del número de onda:

$$\Delta^2_T = \frac{\ell(\ell+1)}{2\pi} C_\ell T^2 \,. \tag{B.3}$$

Los C_ℓ pueden ser calulados con gran exactitud en función de los parámetros cosmológicos (Seljak & Zaldarriaga 1996, Hu *et al.* 1998), de lo que se deduce la importancia de su estudio a la hora de intentrar determinar los valores de dichos parámetros. El comportamiento del espectro de potencias angular en función de ℓ refleja de manera muy sensible los fenómenos físicos fundamentales del Universo primitivo, por lo que resulta de vital importancia su estudio e interpretación. Aun en el caso de que las fluctuaciones de la RCFM no sean gaussianas, el espectro de potencias angular de la RCFM sigue siendo una herramienta fundamental en Cosmología.

El conocimiento del espectro de potencias de la RCFM está limitado por varias razones:

- En primer lugar, porque un observador real siempre estará limitado a observar un único Universo, con un único conjunto de $a_{\ell m}$, de tal manera que para ℓ bajos el número máximo posible de $a_{\ell m}$ observables es muy pequeño. Esto nos lleva a una varianza inevitable de orden $\frac{2}{2\ell+1} C^2_\ell$. Esta limitación es fundamental e insoslayable, y recibe el nombre de *varianza cósmica*.

- Otro efecto que reduce nuestra habilidad para medir con precisión el espectro de potencias de la radiación es la *varianza muestral*, la cual es debida a un cubrimiento parcial del cielo. Esta limitación también es inevitable, ya que

aunque un experimento dado sea capaz de abarcar todo el cielo siempre habrá regiones altamente contaminadas que deban descartarse del análisis.

- Por último, están las propias limitaciones instrumentales de los experimentos de medición de la RCFM.

En la figura 1.1 se muestra el estado actual de los experimentos de RCFM. Estos experimentos han permitido imponer fuertes ligaduras en el modelo cosmológico que actualmente se maneja, y en un futuro próximo se espera que permitan determinar los parámetros cosmológicos fundamentales con errores del orden de unos pocos tantos por ciento.

Tipos de anisotropías de la RCFM

Las anisotropías de la RCFM pueden clasificarse como *primarias*, si el momento de su origen es previo o igual al del desacoplo, y *secundarias*, si es posterior. Durrer (2001) y Hu & Dodelson (2001) ofrecen excelentes revisiones sobre el tema.

Anisotropías primarias. Su origen último son las fluctuaciones en la densidad del fluido materia-radiación en la época del desacoplo. Dichas fluctuaciones sirven también como semillas para la formación de las estructuras que actualmente se observan en el Universo. Se suelen distinguir dos tipos de fluctuaciones: *adiabáticas* y de *isocurvatura*. Las fluctuaciones adiabáticas se caracterizan por una fluctuación nula de la entropía específica asociada a cada componente:

$$\delta(\frac{n_b}{n_\gamma}) = \delta(\frac{n_X}{n_\gamma}) = 0\,, \tag{B.4}$$

donde los subíndices γ, b y X hacen referencia a fotones, bariones y materia no bariónica, respectivamente. Esta ecuación implica la siguiente relación en el desacoplo:

$$\delta_\gamma = \frac{4}{3}\delta_b = \frac{4}{3}\delta_X. \tag{B.5}$$

Las fluctuaciones de isocurvatura se caracterizan por una fluctuación nula de la energía total en cada punto, esto es, $\delta(\rho_\gamma + \rho_b + \rho_X) = 0$, con lo que la curvatura del espacio se mantiene constante. En adición, se suele asumir que

la entropía por barión permanece constante, $\delta(n_b/n_\gamma) = 0$. Eso conduce a

$$\delta_\gamma = -\frac{4\rho_X}{3\rho_b + 4\rho_\gamma}\delta_X. \tag{B.6}$$

Los modelos de inflación (Guth 1982, Linde 1982, 1983, Albrecht & Steindhart 1982) consideran que las fluctuaciones se deben a fluctuaciones cuánticas cuya escala fue enormemente amplificada por un periodo de inflación cósmica. Estos modelos tienden a favorecer fluctuaciones adiabáticas (Kolb & Turner 1990) de tipo gaussiano, aunque existen modelos inflacionarios no estándard que predicen fluctuaciones de isocurvatura o no gaussianas. Modelos no inflacionarios, basados en defectos topológicos (por ejemplo cuerdas cósmicas), predicen por lo general fluctuaciones no gaussianas (para una revisión sobre el tema, véase Vilenkin & Shellard 1994). Además de a las fluctuaciones de densidad de fotones, las anisotropías primarias se deben también al corrimiento al rojo sufrido por los fotones mientras escapan de los pozos de potencial existentes en el momento del desacoplo (*efecto Sachs-Wolfe*). Por último, aparece un término debido al efecto Doppler causado por las velocidades peculiares de los últimos emisores de fotones:

$$\frac{\Delta T}{T}(\vec{n}) \approx \frac{1}{4}\delta_{\gamma d} + \frac{1}{3}\phi_d - \vec{n}\vec{v}_d, \tag{B.7}$$

donde $\vec{n}$ es la dirección de observación y el subíndice d indica cantidades tomadas en el tiempo de desacoplo (se toman unidades $c = 8\pi G = 1$). El primer término representa las fluctuaciones en la densidad de fotones, el segundo recoge el efecto Sachs-Wolfe y el tercero es el término Doppler arriba mencionado.

Anisotropías secundarias. Se deben fundamentalmente a dos tipos de fenómenos: *gravitatorios* y de *reionización*.

- *Efectos gravitatorios*: Uno de ellos es el llamado efecto *Sachs-Wolfe integrado* (ISW). Cuando un fotón cae y escapa de un pozo de potencial constante en el tiempo, el cambio neto en la energía del fotón es cero. Sin embargo, si la profundidad del pozo de potencial varía mientras el fotón lo cruza, la energía ganada al caer ya no se cancela con la perdida

al escapar. La amplitud del efecto ISW está dado por una integral a lo largo del camino del fotón (Martínez-González *et al.* 1990):

$$\frac{\Delta T}{T} = \int \frac{\partial \phi}{\partial t}(\vec{r}, t) dt \,. \tag{B.8}$$

Hay contribuciones de este efecto en los siguientes casos: el universo no está completamente dominado por la materia al tiempo del desacoplo (efecto ISW temprano), la constante cosmológica $\Lambda \neq 0$ o $\Omega + \Lambda \neq 0$ (efecto ISW tardío) o el régimen lineal no es válido (efecto Rees-Sciama). Por otra parte, el efecto *lente gravitatoria* modifica la trayectoria de los fotones sin cambiar su energía, lo que produce una distorsión de la imagen de la superficie de desacoplo. Finalmente, las ondas gravitatorias también pueden generar anisotropías secundarias en la RCFM.

- *Fenómenos de reionización*: Si en algún momento de su historia el Universo ha sufrido de nuevo una fase de fuerte ionización, una fracción de los fotones de la RCFM habrá interaccionado con electrones libres, borrando parte de las anisotropías primarias. La importancia de este borrado dependerá del grado de reionización alcanzado y del tamaño de la región del Universo (posiblemente todo él) que haya sufrido dicha reionización. Un efecto similar es la distorsión de la energía de los fotones de la RCFM debida a dispersión Compton en el gas caliente que se encuentra en el interior de los cúmulos de galaxias (*efectos Suyaev-Zel'dovich térmico y cinemático*).

Emisiones contaminantes

El cielo de microondas contiene no sólo la señal de la RCFM sino también la emisión debida a otros objetos astronómicos, que desde el punto de vista del análisis de la RCFM son considerados *contaminantes*. Además de estos contaminantes, los experimentos que detectan la RCFM presentan de forma inevitable los datos 'corrompidos' con cierta dosis de ruido debido a la respuesta de los instrumentos. Este ruido debe ser eliminado en la medida de lo posible si queremos analizar correctamente la RCFM. Los contaminantes que se deben a procesos astronómicos, sin embargo, requieren una consideración especial. Los principales contaminantes presentes en las longitudes de onda de los experimentos de microondas son:

Radiación sincrotrón. Es debida a electrones relativistas acelerados en el seno de campos magnéticos y domina la emisión galáctica a frecuencias bajas $\nu \lesssim 20$ GHz. La dependencia espectral de este contaminante es aproximadamente descrita como $I \propto \nu^{-\beta_{syn}}$ con $\beta_{syn} \sim 0.9$.

Radiación de frenado . Es la radiación emitida por electrones libres al ser acelerados por iones en el gas interestelar. Este contaminante es el menos conocido de las emisiones galácticas, ya que se cree que domina únicamente en un rango muy pequeño de frecuencias ($\simeq 25 - 75GHz$), donde además la emisión galáctica total es mínima. La dependencia espectral de la radiación de frenado sigue aproximadamente la ley $I \propto \nu^{-\beta_{ff}}$ con un índice espectral $\beta_{ff} \sim 0.16$

Emisión por parte del polvo . Los granos de polvo de nuestra galaxia son calentados por la radiación interestelar, absorbiendo fotones ópticos y ultravioletas y reemitiendo la energía en el infrarrojo lejano. La emisión total observada es la suma sobre la emisión de cada grano de polvo a lo largo de la línea de visión. Este contaminante domina la emisión galáctica a frecuencias $\nu \gtrsim 90$GHz y su dependencia espectral se suele modelizar como una ley de emisión de cuerpo negro modificada $I \propto B_\nu(T_d)\nu^\alpha$ con $\alpha \sim 2$ y $T_d \sim 18$K. La existencia de una segunda componente de polvo fría con una temperatura ~ 7K ha sido sugerida en la literatura para ajustar la emisión de polvo galáctica. Si esta componente fría existiese, podría dominar la emisión de polvo a bajas frecuencias. Sin embargo, todavía hay una gran incertidumbre sobre su existencia. Una posible componente anómala debida a granos de polvo en rotación sería detectable a frecuencias de 10 a 100 GHz.

Fuentes puntuales extragalácticas. Poblaciones de fuentes diferentes dominan por encima y por debajo de $\nu \sim 300$GHz. A frecuencias más bajas, las radio fuentes dan la contribución principal mientras que a frecuencias más altas dominan las fuentes del infrarrojo lejano. Estas poblaciones consisten principalmente de núcleos galácticos activos, blazars y cuásares radioemisores en el radio y de galaxias espirales inactivas en el infrarrojo lejano. El número de cuentas y la dependencia espectral de estas poblaciones están sujetos a gran

incertidumbre debido a la falta de catálogos de fuentes puntuales en el rango de frecuencias explorado por los experimentos de la RCFM, especialmente a las frecuencias de Planck (30-900 GHz). El espectro de potencias de las fuentes puntuales es, aproximadamente, el de un ruido blanco, es decir, $C_\ell =$ cte. para todas las escalas, siempre y cuando el agrupamiento intrínseco de las fuentes no se sea muy fuerte.

Cúmulos de galaxias. Debido al efecto Sunyaev-Zel'dovich, los fotones de la RCFM ven alterada su energía en la dirección de los cúmulos de galaxias. Aunque este efecto sea propiamente una anisotropía secundaria, es tratado aquí por dos razones: primero, porque su origen puede ser localizado en objetos astronómicos concretos, como es el caso de los contaminantes. Segundo, porque, al igual que las fuentes puntuales extragalácticas, los cúmulos de galaxias aparecen en los mapas de RCFM como objetos compactos y, por lo tanto, las técnicas de separación de componentes de las que trata esta tesis son muy adecuadas para detectar estos objetos. El efecto Sunyaev-Zel'dovich se debe a cambios en la dirección y le energía de los fotones debido a dispersión (*scattering*) Compton inversa:

$$\epsilon' = \frac{\epsilon}{1 + (1 - \cos\phi)\frac{\epsilon}{m_e c^2}}. \tag{B.9}$$

La anterior fórmula está calculada en el sistema de referencia de reposo del electrón y en ella ϵ y ϵ' son las energías del fotón antes y después de la dispersión, mientras que ϕ es el ángulo de deflexión del fotón. Múltiples dispersiones en el seno de los cúmulos de galaxias dan origen a una distorsión en la temperatura de la RCFM

$$\frac{\Delta T(x)}{T_0} = g(x) y_c, \tag{B.10}$$

donde donde $y_c = \frac{k_B T_e}{m_e c^2}$ es el *parámetro de Compton*, $x = h\nu / k_B T_{CMB}$ y $g(x) = (x \coth(x/2) - 4)$ es el *factor de forma espectral*. La cantidad τ es la *profundidad óptica* del gas. En términos de intensidades,

$$\Delta I(x) = I_0 f(x) y_c \,,\; f(x) = \frac{x^4 e^x}{(e^x - 1)^2} g(x). \tag{B.11}$$

Esta distorsión tiene una dependencia de la frecuencia muy característica que puede verse en la figura 1.3. Cuando $x = 3.81$ ($\nu = 217$ GHz), el cambio en la intensidad de la RCFM debido al efecto SZ es nulo. Las distorsiones de temperatura típicas debidas al efecto SZ son del orden de $\delta T/T \sim 10^{-4}$, esto es, diez veces mayores que las anisotropías intrínsecas de la RCFM. Además del efecto SZ descrito, que recibe el nombre de *efecto Sunyaev-Zel'dovich térmico*, existe otra variante debida al movimiento conjunto de los electrones del gas por causa del movimiento propio de los cúmulos. Este efecto, conocido como *efecto Sunyaev-Zel'dovich cinemático*, se distingue del térmico debido a que las distorsiones que produce no dependen de la frecuencia sino de la velocidad del cúmulo:

$$\left(\frac{\Delta T}{T}\right)_{kin} = -\frac{v_r}{c}\tau, \tag{B.12}$$

donde τ es la *profundidad óptica* del gas. El efecto SZ cinemático suele ser unas cien veces más débil que el SZ térmico. El efecto SZ es enormemente útil como herramienta cosmológica, ya que su intensidad no depende de la distancia a la que están los cúmulos que lo originan, por lo que permitirá detectar los primeros cúmulos que se formaron en el Universo. Por lo tanto, la detección del efecto SZ se convertirá en punto de referencia obligado para todo modelo de formación de estructura que se precie.

El problema de la separación de componentes

Como hemos visto, los 'contaminantes' que aparecen en la región de microondas del espectro son debidos a objetos astronómicos de la mayor importancia. Es por ello necesario no sólo eliminarlos de los mapas de RCFM para poder estudiar bien esta última, sino también conservarlos aparte para su posterior estudio e interpretación. Por lo tanto, antes de realizar cualquier análisis es necesario eliminar el ruido instrumental y lograr una adecuada *separación de las distintas componentes físicas que intervienen en los datos.*

Los diversos métodos de separación de componentes que se han propuesto aprovechan en mayor o menor medida las distintas propiedades estadísticas de los contaminantes y la RCFM para conseguir su separación. Los principales métodos existentes son: el filtro de Wiener (WF, Tegmark & Efstathiou 1996, Bouchet

et al. 1997), el método de máxima entropía (MEM, Hobson et al. 1998, 1999), análisis de componentes independientes, (FastICA, Maino et al. 2001), análisis mediante la wavelet 'Mexican Hat' (MHW, Cayón et al. 2000, Vielva el al. 2001a) y técnicas que utilizan filtros 'ajustados' (*matched filters*, en inglés) (MF, Tegmark & Oliveira-Costa 1998). Para una comparación de diferentes técnicas aplicadas a la RCFM, véase Tegmark (1997) y Jones (1998). Las técnicas de WF, MF y MHW serán tratadas en mayor profundidad a lo largo de este trabajo.

Cuando se conocen pobremente las propiedades estadísticas de alguna de las componentes, o si se quiere obtener un método *robusto* que haga el menor número posible de suposiciones acerca de dichas propiedades, las técnicas de *filtrado* sobresalen. Ejemplos de técnicas de filtrado son WF, MHW y MF. Este trabajo está dedicado al desarrollo de una nueva técnica de filtrado destinada a la detección de fuentes compactas en mapas de RCFM.

B.1.2 Filtros

Un filtro es una herramienta física o matemática que, aplicada a unos datos de entrada, los modifican de un cierto modo deseado. Dicho de una manera más formal, un filtro es un *operador*

$$L : f(t) \longrightarrow g(t) = Lf(t). \tag{B.13}$$

El filtro es *lineal* si la cantidad filtrada es un funcional lineal de las cantidades de entrada. El filtro es *invariante con respecto al tiempo* si, al producirse un retardo τ en los datos de entrada, los datos de salida también sufren ese mismo retardo, esto es, $g(t-\tau) = Lf(t-\tau)$. La mayor parte de los filtros que se suelen utilizar en muchos ámbitos de la ciencia y la ingeniería son lineales e invariantes con respecto al tiempo. La utilidad de estos filtros radica en que su acción puede expresarse como una convolución:

$$Lf(t) = \int_{-\infty}^{\infty} f(u)h(t-u)du = \int_{-\infty}^{\infty} h(u)f(t-u)du = h \otimes f(t), \tag{B.14}$$

donde $h(t)$ recibe el nombre de *impulso* del filtro. Esta convolución se puede escribir en el espacio de Fourier como un producto:

$$Lf(t) = g(t) = h \otimes f(t) = \int_{-\infty}^{\infty} \hat{h}(q)\hat{f}(q)e^{-iqt}dq, \tag{B.15}$$

siendo la transformada de Fourier $\hat{h}$ del impulso del filtro lo que se denomina *función de transferencia* del filtro. Estas dos últimas ecuaciones implican que los filtros lineales e invariantes con respecto al tiempo pueden ser interpretados como herramientas capaces de seleccionar determinadas frecuencias (modos) de una señal en el espacio de Fourier. Por lo tanto, los filtros son muy útiles a la hora de:

- Eliminar ruido presente en frecuencias (escalas) altas.
- Suavizar datos.
- Seleccionar ciertas bandas de frecuencia (escala).
- Eliminar componentes de una señal.
- Cancelar variaciones de baja frecuencia (escala) superpuestas a una señal.

El diseño de filtros adecuados a un determinado problema es, en cierta medida, un arte. En general, los filtros ideales (esto es, los que son cero fuera de un determinado rango de frecuencias y valen la unidad dentro de dicho rango) no son especialmente útiles debido a que generan efectos en forma de anillos y además las distintas señales que componen los datos reales rara vez se hayan en regiones disjuntas del espectro. Por lo tanto, se buscan filtros continuos que sean *óptimos* para el problema en cuestión.

Filtros óptimos

Habitualmente, el problema al que se aplican los filtros es alguna variante del que aparece en el siguiente diagrama:

$$\text{SEÑAL} + \text{RUIDO} \longrightarrow \text{FILTRO} \longrightarrow \text{SEÑAL} + \text{ruido reducido}$$

En el mejor de los casos, el 'ruido reducido' debe ser lo más cercano posible a cero. La forma de establecer dicha cercanía es a través de algún criterio estadístico. Según el criterio elegido, se obtiene un tipo de filtro u otro. Dos claros ejemplos son:

El filtro de Wiener Sea un conjunto de datos $\mathbf{d} = (d_1, d_2, ..., d_N)$ que son utilizados para estimar la señal subyacente $\mathbf{s} = (s_1, s_2, ..., s_M)$ y que pueden escribirse como:

$$\mathbf{d} = \mathbf{Rs} + \mathbf{n}, \tag{B.16}$$

donde $\mathbf{R}$ es una matriz $N \times M$ conocida y $\mathbf{n}$ es un vector que representa el ruido instrumental. Se busca un filtro $\mathbf{W}$ tal que, aplicado a los datos, produzca unas cantidades filtradas $\hat{\mathbf{s}}$ que sean lo más parecidas a la señal original *en el sentido de los mínimos cuadrados*, esto es, que $\langle|\mathbf{s} - \hat{\mathbf{s}}|^2\rangle$ sea mínimo. Esto nos lleva a un filtro (p.ej. Rybicki y Press 1992):

$$\mathbf{W} = \mathbf{SR}^t \left(\mathbf{RSR}^t + \mathbf{N}\right)^{-1}, \tag{B.17}$$

donde $\mathbf{N} =< nn^t >$ es la matriz de correlación del ruido.

Filtros 'ajustados' Si en vez de pedir que las cantidades filtradas sean lo más parecidas posible a la señal original en el sentido de los mínimos cuadrados se impone que la relación señal/ruido final sea máxima, esto es, que haya una *ganancia* máxima con respecto a la situación original a la hora de detectar una señal dada, se obtiene un filtro

$$\Phi(q) \propto \frac{s^*(q)}{P(q)}, \tag{B.18}$$

donde $s^*(q)$ es el conjugado del perfil de la señal en el espacio de Fourier y $P(q)$ es el espectro de potencias de los datos. Dicho filtro recibe el nombre de 'filtro ajustado' (*matched filter*). Por ejemplo, el filtro ajustado para el caso de una señal gaussiana inmersa en ruido blanco es precisamente un filtro gaussiano de idéntica anchura que la señal.

Filtros y wavelets

En los últimos años las técnicas de wavelets (*'ondículas'*) han revolucionado el mundo del análisis de imágenes y el procesado de señales en general. Las wavelets permiten obtener bases funcionales que poseen al mismo tiempo buena localización tanto en el espacio real como en el de Fourier, por lo que se convierten en herramientas muy adecuadas para tratar señales localizadas como las que se manejan

en esta tesis. Consideremos la *transformada continua de wavelet*. Sea la wavelet

$$\psi_{(R,b)}(t) = R^{-1/2}\psi\left(\frac{t-b}{R}\right), \tag{B.19}$$

entonces se define la transformada continua de wavelet como

$$Wf(R,b) = \int_{\infty}^{\infty} f(t)\psi_{(R,b)}(t)\,dt = f \otimes \bar{\psi}_R(b)\,, \tag{B.20}$$

donde

$$\bar{\psi}_R(t) = R^{-1/2}\psi\left(\frac{-t}{R}\right). \tag{B.21}$$

Esta transformación permite descomponer de manera continua una señal en elementos que contienen información acerca del comportamiento de dicha señal en distintas escalas. La transformada continua de wavelet puede interpretarse a la luz de las anteriores ecuaciones como una especie de filtrado en el que intervienen las escalas R. Esta idea va a ser muy importante en el desarrollo de esta tesis.

B.2 Filtros Adaptados a la Escala

Como se ha comentado ya, las técnicas de filtrado ofrecen una posible solución a los problemas de separación de componentes y eliminación de ruido instrumental en muchos campos de la Astrofísica. En particular, nos centraremos en cómo localizar señales compactas (*fuentes*, a partir de ahora) cuando éstas se hallan inmersas en un fondo de ruido que supondremos estadísiticamente homogéneo e isótropo y que viene caracterizado por su espectro de potencias. Para simplificar el estudio, supondremos que las fuentes tienen simetría esférica, lo cual es una muy buena aproximación en la mayoría de los casos. Casos típicos que uno puede encontrarse en Astrofísica son: el estudio de líneas de emisión o de absorción en espectros, el análisis de series de datos ordenados en el tiempo, la limpieza de imágenes bidimensionales, la detección y extracción de fuentes puntuales extragalácticas en mapas de RCFM, la detección de cúmulos en imágenes de rayos X o la detección de estructuras en conjuntos tridimensionales de datos.

En todos los casos anteriores, cabe preguntarse cuál será el filtro óptimo para conseguir el fin deseado. Por supuesto, la consecución de dicho objetivo dependerá de la definición de *'óptimo'* que se haga. En esta sección se introducirá un nuevo

tipo de filtro, el *filtro adaptado a la escala*, que es óptimo desde el punto de vista de la detección fiable de fuentes.

B.2.1 Obtención del filtro adaptado a la escala

Imaginemos un conjunto de datos n-dimensional que se puede representar como

$$y(\vec{x}) = s(x) + n(\vec{x}), \tag{B.22}$$

donde $\vec{x}$ son las coordenadas en cada punto de la imagen y $s(x) = A\tau(x)$ representa una señal con simetría esférica ($x \equiv |\vec{x}|$), de amplitud A y perfil τ ($\tau(0) = 1$). El ruido de fondo $n(\vec{x})$ es un campo aleatorio homogéneo e isótropo de media cero que se puede describir a través del espectro de potencias

$$\langle n(\vec{q})n^*(\vec{q}\,')\rangle = P(q)\delta^n(\vec{q} - \vec{q}\,'), \tag{B.23}$$

siendo $n(\vec{q})$ la trnasformada de Fourier de n. Sea un filtro también con simetría central

$$\Psi(\vec{x}; R, \vec{b}) = \frac{1}{R^n}\psi\left(\frac{|\vec{x} - \vec{b}|}{R}\right), \tag{B.24}$$

donde $\vec{b}$ define una translación mientras que R define una dilatación. El campo filtrado es entonces

$$w(R, \vec{b}) = \int d\vec{x}\, y(\vec{x})\Psi(\vec{x}; \vec{b}, R), \tag{B.25}$$

que se puede expresar como producto en el espacio de Fourier:

$$w(R, \vec{b}) = \int d\vec{q}\, e^{-i\vec{q}\vec{b}} y(\vec{q})\psi(Rq). \tag{B.26}$$

Las condiciones que se exigen al filtro para que sea óptimo son:

1. Existe una escala R_o tal que $\langle w(R, \vec{0})\rangle$ es máximo a dicha escala. Esta condición servirá para seleccionar las fuentes frente a posibles detecciones falsas indeseadas.

2. $\langle w(R_o, \vec{0})\rangle = s(0) \equiv A$, esto es, $w(R, \vec{0})$ es un estimador sin sesgo de la amplitud de las fuentes.

3. La varianza de $w(R, \vec{b})$ es mínima a la escala R_o, esto es, tenemos un estimador eficiente de la amplitud de las fuentes.

Estas tres condiciones conducen al filtro (ver Apéndice A):

$$\tilde{\psi}(q) \equiv \psi(R_o q) = \frac{1}{\alpha}\frac{\tau(q)}{P(q)}\frac{1}{\Delta}\left[nb + c - (na+b)\frac{d\,ln\tau}{d\,lnq}\right], \tag{B.27}$$

$$\begin{aligned} \Delta &= ac - b^2, \\ a &\equiv \int dq\, q^{n-1}\, \tfrac{\tau^2}{P}, \\ b &\equiv \int dq\, q^{n-1}\, \tfrac{\tau}{P}\tfrac{d\,\tau}{d\,lnq}, \\ c &\equiv \int dq\, q^{n-1}\, \tfrac{1}{P}\left[\tfrac{d\,\tau}{d\,lnq}\right]^2. \end{aligned} \tag{B.28}$$

Este filtro, desarrollado en el capítulo 2 por vez primera, recibe el nombre de *filtro adaptado a la escala* asociado a la fuente de perfil τ y al fondo con espectro de potencias P. El objetivo de este filtro es *maximizar la probabilidad de detección de las fuentes al tiempo que se mantiene baja la probabilidad de encontrar falsas detecciones*. Lo primero se consigue gracias a las condiciones (2) y (3), mientras que del segundo punto se encargará la condición (1). En efecto, el comportamiento del coeficiente $w(R,\vec{0})$ con la escala R es un poderoso elemento para discriminar las señales auténticas de las que no lo son.

Este filtro puede ser obtenido también en el espacio real:

$$\Psi(\vec{x}; R_o, \vec{0}) = \frac{1}{R_o^n}\psi\left(\frac{|\vec{x}|}{R_o}\right) = \frac{1}{\alpha\Delta}\left[(nb+c)F(x) - (na+b)G(x)\right], \tag{B.29}$$

donde F y G son las transformadas de Fourier inversas de $\frac{\tau}{P}$ y de $\frac{1}{P}\frac{d\tau}{dlnq}$, respectivamente. Se puede definir el *nivel de detección* como

$$\mathcal{D}_w \equiv \frac{\langle w(R,\vec{0})\rangle}{\sigma_w(R)}. \tag{B.30}$$

Entonces, el filtro producirá una *ganancia*:

$$g \equiv \frac{\mathcal{D}_w}{\mathcal{D}} = \frac{\sigma_b}{\sigma_w(R_o)}. \tag{B.31}$$

Varios casos de interés son cuando las fuentes tienen un perfil gaussiano (como es el caso de las fuentes puntuales extragalácticas detectadas mediante antenas de instrumentos de microondas) o cuando tienen perfil exponencial (como es el caso de los discos de galaxias espirales). En el capítulo 2 se estudian en detalle ambos casos, haciéndose hincapié en los particularmente interesantes casos en los que el ruido puede describirse por un espectro plano (ruido blanco) o por uno de tipo $1/f$.

B.2.2 Simulaciones

Para poner a prueba el método propuesto, se han realizado simulaciones unidimensionales en las cuales se han colocado fuentes con perfil gaussiano inmersas en ruido $1/f$. Los filtros adaptados a la escala son comparados con otros filtros clásicos tales como la gaussiana y la wavelet 'Mexican Hat'. Nuestro resultado principal es que los filtros adaptados a la escala superan a dichos filtros en número de detecciones al mismo tiempo que reducen el número de falsas detecciones de manera muy apreciable. Las fuentes son encontradas con errores inferiores al 15% en la determinación de la amplitud. Estos resultados permiten la extracción de las fuentes en el espacio real, con lo que se consige la separación de componentes en lo que a las fuentes consideradas respecta.

B.3 Filtros adaptados a la escala aplicados a imágenes en 2D

Los filtros adaptados a la escala pueden ser construidos para perfiles más generales que los gaussianos y exponenciales considerados anteriormente. Por ejemplo, en Astrofísica es muy común encontrar perfiles de tipo

$$\tau(x) = \frac{1}{\left[1 + \left(\frac{x}{r_c}\right)^2\right]^{\lambda}}, \quad \lambda \geq 1/2. \tag{B.32}$$

Ejemplos típicos son la emisión en microondas y en rayos X de los cúmulos de galaxias, que se pueden aproximar bien mediante los ya conocidos 'modelos-β', que se relacionan fácilmente con la ecuación anterior a través de $\lambda = \frac{3\beta-1}{2}$ (microondas) y $\lambda = \frac{6\beta-1}{2}$ (rayos X), siendo la densidad en número de electrones en los cúmulos $n_e(r) \propto [1 + (r/r_c)^2]^{-\frac{3}{2}\beta}$. Tomando el valor estándar $\beta = 2/3$ se obtiene $\lambda = 1/2, 3/2$ para el caso de microondas y el de rayos X, respectivamente. El perfil (B.32) recibe el nombre de *multicuádrico*.

B.3.1 Filtro adaptado a la escala asociado a perfiles multicuádricos

Si asumimos un perfil como el dado en eq. (B.32) y un fondo de ruido cuyo espectro de potencias se pueda aproximar por $P(q) = Dq^{-\gamma}$, se obtiene un filtro adaptado

a la escala del tipo:

$$\tilde{\psi}(q) = \frac{1}{\alpha a'} \frac{\Gamma(\lambda)}{2^{1-\lambda}(\gamma+n)} (qr_c)^{\gamma+\lambda-\frac{n}{2}} \left[P\, K_{\lambda-\frac{n}{2}}(qr_c) + Q\, qr_c K_{1+\lambda-\frac{n}{2}}(qr_c) \right], \quad \text{(B.33)}$$

$$P \equiv 2\gamma - (n-\gamma)(\gamma+2\lambda), \quad Q \equiv 2(n-\gamma)\frac{\gamma+2\lambda+1}{\gamma+4\lambda-n}, \quad \text{(B.34)}$$

$$a' \equiv \frac{2^{\gamma+2\lambda-3}}{\Gamma(\gamma+2\lambda)} \Gamma\left(\frac{\gamma+n}{2}\right) \Gamma^2\left(\frac{\gamma+2\lambda}{2}\right) \Gamma\left(\frac{\gamma+4\lambda-n}{2}\right), \quad \text{(B.35)}$$

donde Γ es la función gamma y K es la función K de Bessel. La tabla 3.1 contiene la expresión de estos filtros en el espacio de Fourier y el espacio real para algunos casos de λ y γ.

B.3.2 Simulaciones y resultados

En el capítulo 3 se han realizado tres simulaciones distintas para poner a prueba el filtro (B.33). En el primer caso, se distribuyeron 100 cúmulos con $\lambda = 1/2$ e igual tamaño ($r_c = 1$ píxel) sobre un fondo $P(q) \propto q^{-3}$ de tal manera que la relación señal/ruido media de las fuentes fuese 2.0. El filtro adaptado a la escala correspondiente produjo en esas circunstancias ganancias medias $g = 4.4$, lo que quiere decir que fuentes situadas inicialmente a 2σ pasaron a estar por encima de 8σ a la hora de la detección. Las amplitudes de las fuentes se lograron determinar con errores medios del orden del 10%.

En la segunda simulación se distribuyeron 100 cúmulos de tipo $\lambda = 1/2$, variando ahora no sólo sus amplitudes sino también sus tamaños (0.5 píxel $< r_c <$ 2.0 píxel). De nuevo el fondo fue simulado con $P(q) \propto q^{-3}$ y señal/ruido inicial de 2.0. La condición (1) de los filtros adaptados a la escala permite determinar el tamaño de la fuentes mediante el siguiente procedimiento: varíese los valores de r_c y la escala del filtro hasta que se encuentre un máximo en $w(R, \vec{0})$. Cuando eso ocurra, el r_c para el que se haya encontrado ese máximo será el del cúmulo que se ha detectado. Aplicando este sencillo método, somos capaces de determinar los radios r_c con errores menores que 0.15 píxels. La ganancia media producida por los filtros fue en este caso $g = 3.5$, mientras que, gracias al propio criterio de escalas aplicado para determinar el radio, el número de detecciones falsas se redujo a cero.

Por último, se simularon 100 cúmulos con $\lambda = 3/2$ y radios r_c comprendidos entre 2.0 y 4.5 píxels, inmersos en ruido blanco. La simulación se hizo de tal

manera que se aproximara de forma muy cruda a una imagen típica como las que puede obtener el observatorio XMM con tiempos de exposición de 95 ks. En estas condiciones, la ganancia media es moderada ($g = 2.0$). Utilizando el método de variación de escalas descrito arriba, se determinaron los radios con errores medios menores que 0.4 píxeles, sin sesgos sistemáticos apreciables. Las amplitudes se determinaron con errores medios inferiores al 15%. La determinación de parámetros tales como r_c en cúmulos de rayos X es más difícil que en el caso de microondas debido a la caída más brusca del perfil de los primeros, que limita la información disponible en un área más pequeña de la imagen.

B.4 Detección de fuentes puntuales a partir de secuencias de datos ordenados en el tiempo (DOT) de Planck

La primera aplicación a simulaciones realistas de los filtros adaptados a la escala se ha realizado en el capítulo 4 de esta tesis. En él se ha analizado una simulación correspondiente a los seis primeros meses de datos tomados por el instrumento LFI28 de 30 GHz de la futura misión de la Agencia Europea del Espacio Planck. Planck cubrirá todo el cielo, tomando datos en diez canales diferentes de frecuencia, y permitirá medir las anisotropías de la RCFM con una sensibilidad y resolución sin precedentes. El canal de 30 GHz, aunque no es el más adecuado para detectar fuentes puntuales, es el primero que se simuló dentro del área de simulación de datos del Consorcio de Planck, además de ser el más manejable desde el punto de vista computacional debido a que su tamaño es menor que el de otros canales (con mayor resolución angular).

Aunque la forma usual de presentar los datos de RCFM es a través de mapas bidimensionales (o, cuando se extienden a todo el cielo, en representación esférica), los datos se adquieren en forma de secuencias ordenadas de medidas a lo largo del tiempo que luego se combinan para formar los mapas finales. El proceso de confección de los mapas es complicado y, salvo que se realice con el más extremo de los cuidados, fácilmente puede introducir artificios indeseados que corrompen la información contenida en los mapas. Es por tanto deseable realizar un análisis

paralelo de la secuencia de datos ordenados en el tiempo (DOT). Además, los DOT pueden ser analizados casi inmediatamente después de haber sido obtenidos, de forma que se puede generar un catálogo de fuentes en tiempo real que puede ser utilizado, por ejemplo, para calibración de los instrumentos del satélite.

La secuencia de datos simulada utilizada en este trabajo contiene 8546850 mediciones de temperatura, ordenadas en forma de 4383 anillos (correspondiente cada uno al promedio de sesenta rotaciones del satélite), cada uno de los cuales contiene 1950 datos. La antena considerada es realista y tiene una resolución de $33'$. Los datos incluyen RCFM y todos los principales contaminantes (sincrotrón, radiación de frenado, polvo, fuentes puntuales y ruido instrumental). Previamente al análisis, se promedió cada anillo con los doce anillos adyacentes para eliminar ruido blanco. Para el análisis se emplearon filtros adaptados a la escala, utilizando espectros de potencias obtenidos directamente de los datos. Por comparación, también se utilizaron un filtro gaussiano y una wavelet de tipo Mexican Hat para filtrar los datos.

De los tres filtros utilizados, el que mejores resultados produjo fue el filtro adaptado a la escala. En número de detecciones supera ampliamente al filtro gaussiano y ligeramente a la Mexican Hat. En cuanto al número de detecciones falsas, el filtro adaptado a la escala obtiene muchas menos que los otros dos filtros (menos de un 10% al nivel de detección 4σ, frente a un 274% en el caso del filtro gaussiano y un 30% en el caso de la Mexican Hat). Eso es debido a que el filtro adaptado a la escala elimina mejor las fluctuaciones debidas a elementos como la Galaxia, que no tienen la misma escala que las fuentes.

El filtro adaptado a la escala permite encontrar 224 fuentes reales a partir de los datos, más 22 detecciones falsas (menos de un 10% de las reales). El error medio en la determinación de la posición de las fuentes en el cielo es del orden de la resolución del experimento ($11'$). El error medio en la determinación de la amplitud de las fuentes es de 20%. El catálogo obtenido es completo a partir de los 4.3 Jy, y contiene fuentes con flujos de hasta 0.9 Jy. Se espera que el número de fuentes que se detecte en las frecuencias altas de Planck sea mayor, por lo que en el futuro se emprenderá el análisis de dichas frecuencias.

B.5 Detección de cúmulos usando filtros adaptados a la escala

El efecto Sunyaev-Zel'dovich (SZ) térmico plantea la interesante posibilidad de emplear varios canales de frecuencia a la vez para potenciar la detección de cúmulos en mapas de RCFM. Métodos de separación de componentes ya consagrados como son el filtro de Wiener y el método de máxima entropía aprovechan el conocimiento acerca de la dependencia frecuencial del SZ y del resto de componentes para lograr su separación. Sería interesante poder incorporar esta información a nuestras técnicas de filtrado. El capítulo 5 está dedicado a lograr ese fin. Para ello, encontramos dos métodos:

B.5.1 Método de filtrado múltiple

La idea consiste en construir un conjunto de filtros que tengan en cuenta la información acerca de la dependencia frecuencial de la señal y de las correlaciones entre los diferentes canales de tal modo que, tras aplicar a cada canal su filtro correspondiente, la suma de los coeficientes sea óptima para la detección. Imponiendo unas condiciones totalmente equivalentes a las tres que definen el filtro adaptado a la escala en una sola imagen se obtienen los *filtros múltiples adaptados a la escala*:

$$\tilde{\psi}(q) = \mathbf{P}^{-1}(\alpha\mathbf{F} + \mathbf{G}), \tag{B.36}$$

donde hemos introducido el vector columna $\tilde{\psi}(q) = (\psi_\nu(R_\nu q))$ y las matrices $\mathbf{P}^{-1}$ (que es la matriz inversa del espectro de potencias cruzado de los mapas $\mathbf{P} \equiv (P_{\nu_1\nu_2}(q))$), $\mathbf{F} = (f_\nu\tau_\nu)$ y $\mathbf{G} = (\mu_\nu\beta_\nu)$. Las cantidades que intervienen en esta última definición, junto con la constante α, vienen dados por

$$\mu_\nu \equiv f_\nu\tau_\nu\left[2 + \frac{dln\tau_\nu}{dlnq}\right], \tag{B.37}$$

$$\alpha = A^{-1}{}_0^0, \quad \beta^\nu = A^{-1}{}_0^\nu, \tag{B.38}$$

donde $\mathbf{A}$ es la matriz $(1+n)\times(1+n)$ coyos elementos vienen dados por

$$A_0^0 \equiv \int d\vec{q}\mathbf{F}^t\mathbf{P}^{-1}\mathbf{F}, \quad A_\nu^0 \equiv \int d\vec{q}\mu_\nu(\mathbf{F}^t\mathbf{P}^{-1})_\nu, \tag{B.39}$$

$$A_0^\nu \equiv \int d\vec{q}\,\mu^\nu (\mathbf{P}^{-1}\mathbf{F})^\nu, \quad A_{\nu'}^\nu \equiv \int d\vec{q}\,\mu^\nu \mu_{\nu'}\; P^{-1}{}_{\nu'}^{\nu}\,. \tag{B.40}$$

Estos filtros de aspecto tan complicado son la generalización natural del filtro adaptado a la escala para el problema con múltiples frecuencias.

B.5.2 Método de filtrado simple

Por otra parte, otra opción sería combinar los distintos canales de frecuencia de tal manera que la relación señal/ruido de los cúmulos en el mapa combinado fuese máxima. Dicha combinación tendría en cuenta la dependencia frecuencial de los cúmulos y también las correlaciones entre canales. Después de la combinación, el mapa resultante se podría filtrar normalmente como si fuera un problema parecido a los que se presentan en el capítulo 3. El problema se reduce a encontrar un conjunto de pesos $\mathbf{c} = (c_{\nu_1}, c_{\nu_2}, \cdots, c_{\nu_N})$ tales que, tras la combinación $y(\vec{x}) \equiv \sum_\nu c_\nu y_\nu(\vec{x})$, la relación señal/ruido del efecto SZ sea máxima. Se puede demostrar que estos pesos vienen dados por el problema de autovalores generalizado:

$$(\mathbf{G} - \lambda\mathbf{M})\mathbf{c} = 0, \tag{B.41}$$

donde $\mathbf{G} = (G_{\nu\nu'})$, $G_{\nu\nu'} = f_\nu \tau_\nu(0) f_{\nu'} \tau_{\nu'}(0)$ y $\mathbf{M} = (M_{\nu\nu'})$, $M_{\nu\nu'} = \langle n_\nu(\vec{x}) n_{\nu'}(\vec{x})\rangle$. El vector $\mathbf{c}$ deseado es el que corresponde al mayor de los autovalores λ

B.5.3 Simulación de prueba

En el capítulo 5 se consideran cuatro casos posibles: para cada uno de los dos métodos atrás mencionados, podemos construir filtros adaptados a la escala o *'matched filters'*. Para aclarar las potencialidades de cada una de las cuatro variantes posibles, hemos realizado una simulación de una región del cielo de $12.8° \times 12.8°$ vista a través de los diez canales de frecuencia de Planck en la que se incluyen de manera realista todos los contaminantes que se espera detectar. En dicha simulación se distribuyeron 200 cúmulos de perfil multicuádrico, todos del mismo tamaño para simplificar, con intensidades distribuidas uniformente entre 0 y el flujo del cúmulo más brillante (según simulaciones realistas) que se espera en dicha región del cielo. Los resultados sobre esta simulación de prueba se resumen en:

- El método de filtrado simple es del orden de diez veces más rápido en su ejecución que el método de filtrado múltiple.
- Sin embargo, el método de filtrado múltiple es más potente a la hora de la detección, permitiendo alcanzar flujos más bajos.
- La información acerca de la dependencia frecuencial del efecto SZ permite eliminar casi todas las detecciones falsas. En ese sentido, la principal ventaja de los filtros adaptados a la escala frente a los 'matched filters', que era precisamente su capacidad para eliminar detecciones falsas, carece prácticamente de importancia. Es sabido que los 'matched filters' producen por construcción ganacias superiores a cualquier otro filtro. Por lo tanto, en este problema en concreto es más recomendable usar simplemente 'matched filters' que filtros adaptados a la escala.
- Pese a que ciertos canales contribuyen mucho más al resultado final que otros, todos ellos aportan información valiosa al proceso.

B.5.4 Resultados utilizando una simulación realista

Si en vez de usar la simulación de prueba mencionada arriba se emplean simulaciones realistas de cúmulos (atendiendo a cuestiones como tamaños, distribución en número de cuentas según el flujo, etc) se obtienen los siguientes resultados:

- Sobre el nivel de detección 3σ se encuentran 62 cúmulos y ninguna detección falsa.
- Sobre el nivel de detección 2σ se encuentran 257 cúmulos y 27 detecciones falsas.
- La determinación del radio r_c de los cúmulos siguiendo el método de variación de escalas descrito en el capítulo 3 tiene un error máximo inferior a 0.3 píxeles. La mayor parte de las fuentes detectadas poseen son prácticamente puntuales.
- El flujo de completitud del catálogo de fuentes recuperado es de 170 mJy (a 300GHz).

- La determinación de la amplitud (flujo) de las fuentes tiene errores en torno al 30% para el caso de los cúmulos más brillantes. Para el caso de los cúmulos débiles, que dominan el número de cuentas, estos errores aumentan mucho.

Por tanto, el método demuestra ser potente en el sentido de la detección, pero no tanto a la hora de la determinación de flujos. Trabajos futuros se encaminarán a mejorar este aspecto, por ejemplo mediante ajustes en el espacio de coeficientes filtrados. En cualquier caso, este método permitirá encontrar cerca de 15000 (a 3σ) o 65000 (a 2σ) cúmulos en todo el cielo con un nivel muy bajo de detecciones falsas.

“ Here ends [...]. If it has passed from the high and the beautiful to darkness and ruin, that was of old the fate of Arda Marred; and if any change shall come and the Marring be amended, Manwë and Varda may know; but they have not revealed it, and it is not declared in the dooms of Mandos.”

J.R.R. Tolkien, *“The Silmarillion”*

Bibliography

[1] Albrecht A. & Steinhardt P.J., 1982, Phys.Rev.Lett., 48, 1220

[2] Alpher R.A. & Herman R.C., 1948, Nature, 162, 774

[3] Avelino P.P., Caldwell R.R. & Martins C.J.A.P., 1998, Ap&SS, 261, 319.

[4] Baccigalupi, C., Bedini, L., Burigana, C., De Zotti, G., Farusi, A., Maino, D., Maris, M., Perrotta, F., Salerno, E., Toffolatti, L., & Tonazzini, A., 2000, MNRAS, 318, 769.

[5] Bardeen J.M., Steinhardt P.J. & Turner M.S., 1983, Phys.Rev., D28, 679

[6] Barreiro, R.B., 1999, 'Reconstruction and characterization of cosmic microwave background anisotropy maps', PhD thesis, U. Cantabria.

[7] Barreiro, R.B. & Hobson, M.P., 2001, MNRAS, 327, 813.

[8] Bennett, C.L. *et al.* , 1996, ApJ, 464, L1. MAP homepage http://map.gsfc.nasa.gov

[9] Bertin E., Arnouts S., 1996, A&AS, 117, 393

[10] Birkinshaw M., 1999, Phys.Rept., 310, 97

[11] Bouchet F.R., Gispert R., Boulanger F., Puget J.L., 1997, in 'Microwave Background Anisotropies', Proc. XVIth Moriond Astrophysics Meeting, Bouchet F.R., Gispert R., Guiderdoni B. & Trân Thanh Vân J. eds., Editions Frontiéres, p.481.

[12] Bouchet, F.R., Gispert, R. & Puget, J.L., 1995, in 'Unveiling the Cosmic Infrared Background', AIP Conference Proceedings 348, Baltimore, Maryland, USA, E. Dawek, editor, p. 255-268.

[13] Bouchet, F.R., Gispert, R. & Puget, J.L., 1996 in Drew, E., ed., Proc. AIP Conf. 384, 'The mm/sub-mm foregrounds and future CMB missions', AIP Press, New York, p. 255.

[14] Bouchet, F. R. & Gispert, R. 1999, New Astronomy 4, 449.

[15] Bouchet, F.R., Prunet, S. & Sethi, S.K., 1999, MNRAS, 302, 663.

[16] Burrus, C.S., Gopinath, R.A. & Guo, H., 1998, 'Introduction to Wavelets and Wavelets Transforms', Prentice Hall, New Jersey.

[17] Cayón, L., Sanz, J.L., Barreiro, R.B., Martínez-González, E., Vielva, P., Toffolatti, L., Silk, J., Diego, J.M., Argüeso, F., 2000, MNRAS, 315, 757.

[18] Cayon, L., Sanz, J.L., Martínez-González, E., Banday, A.J., Argüeso, F., Gallegos, J.E., Gorski, K.M. & Hinshaw, G., 2001, MNRAS, 326, 1243.

[19] Damiani F., Maggio A., Micela G. & Sciortino S., 1997, ApJ, 483, 350

[20] Daubechies I., 1988, Comm. Pure Appl. Math., 41, 909

[21] de Oliveira-Costa A., Tegmark M., Gutiérrez C.M., Jones A.W., Davies R.D., Lasenby A.N., Rebolo R. & Watson R.A., 1999, ApJ, 527, L9.

[22] Dicke, R.H., Peebles, P.J.E., Roll, P.G. & Wilkinson, D.T., 1965, ApJ, 142, 414.

[23] Diego, J.M., 2000, 'Methods of Analysis of the Cosmic Microwave Background and the Sunyaev-Zel'dovich effect', PhD thesis, U. Cantabria.

[24] Diego, J.M., Martínez-González, E., Sanz, J.L., Cayón, L., & Silk, J., 2001a, MNRAS, 325, 1533.

[25] Diego, J.M., Vielva, P., Martínez-González, E., Silk, J. & Sanz, J.L., 2001b, MNRAS submitted, astro-ph/0110587.

[26] Draine B.T. & Lazarian A., 1998, ApJ, 494, L19

[27] Draine B.T. & Lee H.M., 1984, ApJ, 285, 89

[28] Efstathiou G. & Bond J.R., 1986, MNRAS, 218, 103

[29] Finkbeiner, D.P., Davis, M. & Schlegel, D.J., 1999, ApJ, 524, 867.

[30] Freeman, P.E., Kashyap, V., Rosner, R. & Lamb, D.Q., 2002, accepted for publication in Ap. J. Supp. (v. 138 Jan. 2002).

[31] Gamow, G. 1948a, Phys.Rev., 74, 505

[32] Gamow, G. 1948b, Nature, 162, 680

[33] Guiderdoni, B., 1999, in 'Microwave Foregrounds', eds. A. de Oliveira-Costa & M. Tegmark ASP, San Francisco, 173.

[34] Guth A.H., 1982, Phil. Trans. R. Soc., A307, 141

[35] Guth A.H. & Pi S.Y., 1982, Phys.Rev.Lett., 49,1110

[36] Halverson, N.W. *et al.* , 2001, ApJ in press, astro-ph/0104489.

[37] Harrison E.R., 1970, Phys.Rev., D1, 2726

[38] Haslam C.G.T., Klein U., Salter C.J., Stoffel H., Wilson W.E., Cleary M.N., Cooke D.J. & Thomasson P., 1982, Astr. Ap., 100, 209

[39] Haykin, S., 1996, 'Adaptive filter theory', Prentice-Hall, New York.

[40] Hawking S.W., 1982, Phys.Lett., 115B, 295

[41] Herranz, D., Gallegos, J., Sanz, J.L., Martínez-González, E., 2001a, submitted to MNRAS.

[42] Herranz, D., Sanz, J.L., Barreiro, R.B., Martínez-González, E., 2001b, submitted to ApJ.

[43] Hobson, M.P., Jones, A.W., Lasenby, A.N., Bouchet, F.R., 1998, MNRAS, 300, 1.

[44] Hobson, M.P., Barreiro, R.B., Toffolatti, L., Lasenby, A.N., Sanz, J.L., Jones, A.W., Bouchet, F.R., 1999, MNRAS, 306, 232.

[45] Hu, W., Seljak, U., White, M. & Zaldarriaga, M. 1998, Phys.Rev., D57, 3290

[46] Hu, W. & Dodelson, S., 2001, to appear in Annu. Rev. Astron. and Astrophys., astro-ph/0110414.

[47] Jeannerot R., 1996, Phys.Rev., D53, 5426

[48] Jonas, J.L., Baart, E.E. & Nicolson, G.D., 1998, MNRAS, 297, 977.

[49] Jones A.W., 1998, 'Application of novel analysis techniques to Cosmic Microwave Background Astronomy', PhD thesis, U.Cambridge.

[50] Kamionkowski M., Spergel D.N. & Sugiyama N., 1994, ApJ, 426, 57

[51] Kawasaki, W., Shimasaku, K., Doi, M. & Okamura, S., 1998, A&A, 130, 567

[52] Kolb E.B. & Turner M., 1990,'The Early Universe', Addison-Wesley, New York.

[53] Lee, A.T. *et al.* , 2001, ApJ in press, astro-ph/0104459.

[54] Linde A., 1982, Phys.Lett., 108B, 389

[55] Linde A., 1983, Phys.Lett., 129B, 177

[56] Linde A. & Riotto A., 1997, Phys.Rev. D56, 1841

[57] Mallat, S., 1998, 'A wavelet tour of signal processing', Academic Press, San Diego.

[58] Mandolesi, N. et al. 1998, proposal submitted to ESA for the Planck Low Frequency Instrument.

[59] Maino, D., Farusi, A., Baccigalupi, C., Perrota, F., Banday, A.J., Bedini, L., Burigana, C., De Zotti, G., Górsky, K.M. & Salerno, E., 2001, MNRAS submitted, astro-ph/0108362.

[60] Martínez-González E., Sanz J.L. & Silk J., 1990, ApJ, 355, L5

[61] Martínez-González E., Sanz J.L. & Cayón L., 1997, ApJ, 484,1

[62] Martínez-González E., Gallegos, J.E., Argüeso, F., Cayón, L. & Sanz, J.L., submitted to MNRAS, astro-ph/0111284.

[63] Mauskopf, P.D., *et al.* 2000, ApJL, 536, 59.

[64] Netterfield C.B. *et al.* , ApJ in press, astro-ph/0104460.

[65] Ogden R. T., 1997, 'Essential wavelets for statistical applications and data analysis', Birkhäuser, Boston

[66] Peebles P.J.E., 1999a, ApJ, 510, 523

[67] Peebles P.J.E., 1999b, ApJ, 510, 531

[68] Penzias A.A. & Wilson R.W., 1965, ApJ, 142, 419

[69] Press, W.H., Flannery, B.P., Teukolsky, S.A., & Vetterling, W.T., 1989, 'Numerical Recipes', Cambridge University Press, Cambridge.

[70] Press W.H. & Schechter, P., 1974, ApJ, 187, 425.

[71] Postman, M., Lubin, L. M., Gunn, J. E., Oke, J. B., Hoessel, J. G., Schneider, D. P. & Christensen, J. A., 1996, ApJ, 615, 111.

[72] Puget, J.L. *et al.* 1998, proposal submitted to ESA for the Planck High Frequency Instrument.

[73] Reach W.T. *et al.* , 1995, ApJ, 451, 188

[74] Reich P. & Reich W., 1986, A&A Suppl., 63, 205

[75] Rephaeli, Y., 2001, invited review, proceedings of the 9th Marcel Grossmann Meeting, astro-ph/0105192.

[76] Reynolds, R.J., Tufte, R.J., Haffner, S.L., Jaehning, K. & Pervival, J.W., 1998, Publ. Astron. Soc. Aust. 15, 14.

[77] Rybicki G.B. & Press W.H., 1992, ApJ, 398, 169

[78] Sachs R.K. & Wolfe A.M., 1967, ApJ, 147, 73

[79] Salopek D.S., 1992, Phys.Rev., D45, 1139

[80] Sanz J.L., Argüeso F., Cayón L., Martínez-González E., Barreiro R.B. & Toffolatti L. 1999a, submitted to MNRAS

[81] Sanz J.L., Barreiro R.B., Cayón L., Martínez-González E., Ruiz G.A., Díaz F.J., Argüeso F., Silk J., Toffolatti L., 1999b, submitted to A&A

[82] Sanz, J.L., Herranz, D. & Martínez-González, E., 2001, ApJ, 552, 484.

[83] Schlegel D., Finkbeiner D. & Davies M., 1998, ApJ, 500, 525

[84] Seljak U. & Zaldarriaga M., 1996, ApJ, 469, 437

[85] Silk J., 1968, ApJ, 151, 459

[86] Slezak E., De Lapparent V. & Bijaoui A., 1993, ApJ, 409, 517

[87] Smoot G.F., 1999, in 'Microwave Foregrounds', de Oliveira-Costa A. & Tegmark M. eds, ASP, San Francisco, 61.

[88] Smoot G.F. *et al.* , 1992, ApJ, 396, L1

[89] Starobinskii A.A., 1982, Phys.Lett., 117B, 175

[90] Stolyarov, V., Hobson, M.P., Ashdown, M.A.J. and Lasenby, A.N. 2001, submitted to MNRAS, astro-ph/0105432.

[91] Tegmark M., 1997, ApJ, 480, L87

[92] Tegmark, M. & Efstathiou, G., 1996, MNRAS, 281, 1297.

[93] Tegmark, M. & de Oliveira-Costa, A., 1998, MNRAS, 500, 83.

[94] Tenorio L., Jaffe A.H., Hanany S. & Lineweaver C.H., 1999, MNRAS, 310, 823.

[95] Toffolatti L., Argüeso Gómez F., De Zotti G., Mazzei P., Francheschini A., Danese L. & Burigana C., 1998, MNRAS, 297, 117 G

[96] Valtchanov, I., Pierre, M. & Gastaud, R., 2001, A&A, accepted

[97] Vielva, P., Martínez-González, E., Cay/'on, L., Diego, J.M., Sanz, J.L. & Toffolatti, L., 2001a, MNRAS, 326, 181.

[98] Vielva, P., Barreiro, R.B., Hobson, M.P., Martínez-González, E., Lasenby, A.N., Sanz, J.L. & Toffolatti, L., 2001b, MNRAS accepted, astro-ph/0105387.

[99] Vilenkin A. & Shellard E.P.S., 1994, 'Cosmic Strings and other Topological Defects', Cambridge University Press

[100] Wright E.L., 1991, ApJ, 381, 200

[101] Zaroubi, S., Hoffman, Y., Fisher, K.B. & Lahav, O., 1995, ApJ 449, 446.

[102] Zel'dovich Ya.B., 1972, MNRAS, 160, 1