Review Article

Compact Source Detection in Multichannel Microwave Surveys: From SZ Clusters to Polarized Sources

Diego Herranz,[1, 2] Francisco Argüeso,[3] and Pedro Carvalho[2]

[1] *Instituto de Física de Cantabria, CSIC-UC, Avenue los Castros s/n, 39005 Santander, Spain*
[2] *Astrophysics Group, Cavendish Laboratory, JJ Thomson Avenue, Cambridge CB3 0HE, UK*
[3] *Departamento de Matemáticas, Universidad de Oviedo, Calvo Sotelo s/n, 33007 Oviedo, Spain*

Correspondence should be addressed to Diego Herranz, herranz@ifca.unican.es

We describe the state-of-the art status of multifrequency detection techniques for compact sources in microwave astronomy. From the simplest cases where the spectral behaviour is well known (i.e., thermal SZ clusters) to the more complex cases where there is little a priori information (i.e., polarized radio sources) we will review the main advances and the most recent results in the detection problem.

1. Introduction

Extragalactic foregrounds play a crucial role in microwave astronomy, not only by their effect as contaminants of the Cosmic Microwave Background (CMB) but also by their own right as cosmological probes. Galaxies and galaxy clusters, if not properly identified and taken into account, can seriously affect the measurement of the CMB anisotropies angular power spectrum in temperature [1–3] and polarization [4, 5], CMB non-Gaussianity tests [6–12], and even the performance of component separation methods used for the study of Galactic foregrounds (on the other hand, the opposite is also true: foregrounds can affect the performance of compact source detection algorithms. In general, compact source detection algorithms find it easier to deal with diffuse foregrounds than diffuse component separation techniques to deal with compact sources, so the typical CMB analysis pipeline includes the detection of compact sources as a previous step to diffuse component separation) [13, 14]. On the other hand, galaxy and galaxy cluster surveys in the submm regime of the electromagnetic spectrum are powerful tools for cosmology [15–17]. This is the motivation of the considerable number of works on extragalactic foreground detection that have appeared in the literature on recent years.

As opposed to Galactic foregrounds, that are typically extended as diffuse clouds over large areas of the sky, individual extragalactic objects appear as compact blobs that subtend very small angular scales. For this reason, both galaxies and galaxy clusters are often referred to as *compact sources* and their detection/separation is typically treated as a problem apart from the one posed by the separation of Galactic diffuse components.

It is precisely the compactness of extragalactic sources that makes it possible to detect them against the fluctuations of the diffuse components (CMB included) in single-frequency (channel) settings. Most of the detection methods that have been proposed in the literature make use of this *scale diversity*. The well-known SExtractor package [18], for example, is particularly good at estimating and then subtracting the background at coarse scales and then detecting compact sources by looking for small sets of connected pixels above a given threshold. For the same reason, techniques based on scale-selecting devices such as band-pass filters [19–24] and wavelets [25, 26] have proven to be very useful. More sophisticated detection Bayesian algorithms [27–31] make also use at some point of the scale diversity of compact sources versus diffuse foregrounds. A recent review on methods for the detection of compact sources in microwave images can be found in [32].

The majority of current CMB experiments, such as the Wilkinson Microwave Anisotropy Probe (WMAP) [33] and *Planck* [34], can observe the sky in several frequency bands simultaneously. Some of them are also capable of measuring not only the intensity of the microwave radiation, but also its polarization. This means that extragalactic compact sources can be observed in several different images or *channels*. Multichannel, (In this work we wil use the term *multichannel detection* instead of the more common *multifrequency* term because it can accommodate both the classic problem of detecting a source using different frequency maps and the detection of a source in polarization using images of the Stokes' parameters *I*, *Q*, *U*. As we will see, both problems are formally equivalent and therefore we prefer the more general terminology), information can be used to improve the chances of detecting compact sources. In this paper we will review the state of the art multichannel detection techniques for compact sources in microwave astronomy.

Multichannel (multi-frequency) astronomy is almost as old as the telescope itself: we have been doing it since the first colour filters became available. During the XXth and XXIst centuries astronomy has expanded its range of operation into the radio, microwave, infrared, ultraviolet, X-ray, and gamma regimes of the electromagnetic spectrum: a tremendous amount of information we are only beginning to piece together. The most basic multichannel consideration we can do is given that we have observed some object in the optical, let us see how does it look like in, for example, the infrared. Catalogue cross-correlation and band merging, follow-ups and the construction of spectral energy distribution (SED) curves are fundamental parts of this multichannel quest for knowledge. But we are *not* going to review this.

A followup, looking at a certain position of the sky where a particular source has been already detected in a different region of the electromagnetic spectrum, can only work if we have a clear detection of the source in the first place. Sometimes, however, the sources are just too faint to be detected with enough significance in any of the available channels. But maybe if we could join together all the channels we would be able to detect the source. Even in the cases where the source is clearly detected in one channel, it may be too weak to be observed in other channels. Once again, one can ask if there is any possibility to use other information apart from just the source position from the "good" channels in order to enhance the source in the "bad" ones. Or we can more generally ask if there is a better way to combine all the channels to get better SEDs than just going channel by channel separately. This is the aim of our paper.

The key for multichannel source separation is *spectral diversity*: we hope the signal of interest to scale from one image to other in a different way than the other components (background). Assuming a linear mixture model in which *N* different, unknown sources are added with a set of channel-dependent weights to produce *M* observed channels ($M \geq N$) and making certain assumptions about the statistical properties of the sources, for example statistical independence, or non-Gaussianity, the Blind Source Separation (BSS) problem is solvable by means of a number of

statistical signal processing techniques such as the Maximum Entropy Method [35], Independent Component Analysis [36], Correlated Component Analysis [37, 38], or wavelet-based Internal Template Fitting [39], just to mention a few of them. For a more detailed review of BSS methods applied to microwave experiments, see [14].

Unfortunately for our purposes the above mentioned BSS techniques are not well suited for the detection of extragalactic compact sources, except for the particular case of galaxy clusters observed through the thermal Sunyaev-Zel'dovich (tSZ) effect. Individual galaxies leave their imprint on the microwave sky through an enormous variety of astrophysical mechanisms, from radio active lobes to dust thermal emission, so that, strictly speaking, each individual galaxy has its own unique spectral behaviour. In the linear mixing model this translates into $N \gg M$, and the BSS problem is sorely under determined. New methods, specifically tailored for compact sources, become necessary. Even in the case of tSZ clusters, where all the sources share the same spectral behaviour, it may be advisable to apply other techniques different from the abovementioned BSS methods. The tSZ is subdominant at all frequencies, making it very difficult for BSS techniques to detect but the brightest clusters in the sky. The multichannel detection methods we are going to describe in this paper make use, up to different extent, of both *scale* and spectral diversities in order to optimize the detectability of compact sources.

In this paper we will review the most widely used methods for detecting extragalactic compact sources using multichannel data. We will first formalize the problem in Section 2. Then we will review the different approaches currently used in microwave astronomy. We choose the order in which we introduce the methods on the basis of the amount of information that is used to achieve detection: we will start in Section 3 with traditional stacking and band-merging techniques that make a minimal use of multichannel information and then proceed in Section 4 to a linear filtering scheme that takes into account the correlation of the background among different channels. In Section 5 we will discuss how to filter the data when both the correlation of the background among channels and the spectral behaviour of the sources are known. We will see how the second condition can be relaxed. Then we will discuss the more general theory of Bayesian detection in Section 6. Finally, we will devote Section 7 to the particular case of polarization data.

2. Compact Sources in Multichannel Observations

Let us consider a set of *N* two-dimensional images (channels) in which there is an unknown number of compact sources embedded in a mixture of instrumental noise and other astrophysical components. Without loss of generality, let us consider the case of a single source with a certain typical angular scale *R* and located at the origin of the coordinates. Our data model is

$$D_j(\vec{x}) = s_j(\vec{x}; R) + n_j(\vec{x}), \tag{1}$$

where the subscript $j = 1,\ldots,N$ denotes the index of the channel: it may refer to a given frequency (i.e., 30 GHz, 44 GHz), to a polarization channel (i.e., the Stokes' parameters I, Q, U) or any other image indexing we can consider. The scale R of the sources can vary from one object to other (as it happens for galaxy clusters) or be the same for all the objects belonging to a given class (i.e., radio sources observed by low angular resolution experiments such as WMAP). It is common to factorize the source term as the product of a channel-dependent *amplitude* or *intensity* and a spatial profile:

$$s_j(\vec{x}) = A_j \times \tau_j(\vec{x}; R). \qquad (2)$$

The spatial profile, in turn, includes the possible effects of any channel-dependent point spread function (beam):

$$\tau_j(\vec{x}; R) = b_j(\vec{x}) \otimes \tau_j^0(\vec{x}; R). \qquad (3)$$

The above formula can be expressed more easily in Fourier space:

$$\tau_j(\vec{k}; R) = b_j(\vec{k}) \times \tau_j^0(\vec{k}; R), \qquad (4)$$

where for simplicity we have kept the same symbol for the functions in real and in Fourier space (the argument $\vec{x}$ or $\vec{k}$ indicates clearly enough in which space we are). We choose the normalization of the profile so that

$$\tau_j^0(\vec{0}; R) = 1. \qquad (5)$$

Finally, for simplicity we will consider symmetric beams and profiles, $\tau_j(\vec{x}; R) = \tau_j(|\vec{x}|; R) = \tau_j(x; R)$. This assumption is not strictly necessary and can be relaxed on demand but it greatly simplifies calculations and is quite reasonable in most applications.

The term $n_j(\vec{x})$ in (1) is the generalized noise in the jth channel, containing not only instrumental noise, but also CMB and all the other astrophysical components apart from the compact sources. Let us suppose that the noise has zero mean and that it can be characterized by its cross-power spectrum:

$$\left\langle n_j(\vec{k}) n_l^*(\vec{k'}) \right\rangle = P_{jl}(k)\delta^2(\vec{k} - \vec{k'}). \qquad (6)$$

In other words, we work under the assumption that the properties of the noise can be sufficiently described by its second-order statistics. Please note that the homogeneity of the background is a necessary condition for the power spectrum to be a full second-order description of the background; this condition is not globally met in real astronomical images (where the Galactic foregrounds, for example, are strongly non homogeneous), but we can always take patches small enough to satisfy local homogeneity (at least on a first approximation).

2.1. Notation. In this paper we will use the notation $\vec{x}$ to indicate coordinate vectors (both in real and Fourier spaces) and the notation $\mathbf{x}$ to indicate vectors whose components are labelled with the indexes of the different channels. According to this notation, we can write (1)–(6) as

$$\mathbf{D}(\vec{x}) = \mathbf{s}(x; R) + \mathbf{n}(\vec{x}),$$
$$\mathbf{s}(x) = \mathbf{A}\boldsymbol{\tau}(x; R),$$
$$\boldsymbol{\tau}(k; R) = \mathbf{b}(k)\boldsymbol{\tau^0}(k; R), \qquad (7)$$
$$\left\langle \mathbf{n}^t(\vec{k})\mathbf{n}(\vec{k'}) \right\rangle = \mathbf{P}(k)\delta^2(\vec{k} - \vec{k'}).$$

It will be useful in some cases to expand the vector of intensities as

$$\mathbf{A} = I_*\mathbf{f}, \qquad (8)$$

where I_* can be interpreted as a channel-independent intensity and $\mathbf{f}$ is a vector containing the spectral behaviour of the sources across the different channels. For example, when we consider the tSZ I_* can be associated with the cluster Compton-y parameter and $\mathbf{f}$ takes the well-known form, in thermodynamic temperature units,

$$f_{\hat{\nu}} \propto \hat{\nu}\frac{e^{\hat{\nu}} + 1}{e^{\hat{\nu}} - 1} - 4, \qquad (9)$$

where $\hat{\nu} = h\nu/kT$, ν is the frequency of observation, k is the Boltzmann constant, and T is the temperature of the CMB. Other example is provided by radio sources whose spectral behaviour can be approximated by a power law

$$f_\nu \propto \left(\frac{\nu}{\nu_0}\right)^{-\alpha}, \qquad (10)$$

where ν_0 is some frequency of reference and α is known as the *spectral index*. In this case, I_* can be interpreted as the source flux density at the reference frequency ν_0. Please note that analytic spectral laws such as (9) or (10) are not always available, or even convenient.

A list of symbols used in this paper can be found at the end of the paper.

3. Basic Multichannel Detection

Probably, the simplest imaginable approach to the multichannel problem consists in trying somehow to transform it to a much simpler single-channel problem. The theory and applications of single-channel source detection are well studied in the literature; we will assume in this work that the reader is already familiar with the topic. A short, recent review can be found in [32] and references therein.

Let us start with the simplest, and most frequent, case of multichannel observation we can conceive. Imagine we have just two independent observations D_1 and D_2 of the same source, taken at the same frequency and with the same instrumental characteristics (beam size, noise level, etc.). A perfect example is any pair of identical radiometers in an experiment such as *Planck*. Let us also assume that the two observations are simultaneous or at least that the source under study has not experienced variability in the time passed between observations. Then, it is well known

that if the noises of the two observations are Gaussian with rms $\sigma_1 = \sigma_2 = \sigma$ and independently distributed, the linear combination $\tilde{D} = (D_1 + D_2)/2$ has a lower rms noise $\tilde{\sigma} = \sigma/\sqrt{2}$. If the different channels have different noise levels but the noises are still uncorrelated, another well-known result is that the noise of the linear combination of the channels can be minimized by inverse noise variance weighting of the individual channels. Once the multichannel data has been combined (or *projected*) into a single channel (or *plane*) it is straightforward to apply threshold detection, filtering plus thresholding, or any other of the techniques described in [32]. With the appropriate modifications, the weighting internal linear combination scheme can be extended to correlated noise [40]. The problem with this approach is that if the source has not the same intensity in all the channels (or if its spectral behaviour is not perfectly known) it is impossible to relate the intensity observed in the combined image with the true intensity of the source.

Other merging schemes can be advisable in some other circumstances. Imagine that some of the components of the noise in (1) are constant across all the channels, the best example is the CMB itself in images expressed in thermodynamic temperature units. Then an obvious way to reduce the noise is to subtract pairs of channels. As an example, [41, 42] used combinations of the WMAP W and V bands in order to produce a CMB-free map to better detect the elusive radio galaxies in the WMAP 5-year data. The problem with this approach is that if a source has by chance a spectrum flat enough, it may be also cancelled by the subtraction.

4. Matrix Multifilters

A recurrent problem of channel linear combination is that accurate photometry, that is, the determination of I_*, is not possible unless the exact values of all the components of the spectral behaviour **f** are known. But the spectral behaviour is not known in most cases. General laws such as (10) are not reliable when the range of frequencies involved is wide enough, and even if they are there is the additional problem that different sources in a given patch of the sky will have different spectral indexes. The classic solution in these cases, if one wants to persist using linear combination nevertheless, is to go back to the individual channels and extract photometric measurements on them.

In order to overcome these problems authors in [43, 44] proposed a method that combines the signal-to-noise boosting capacity of linear filters with a particular merging scheme that makes totally irrelevant the spectral behaviour **f**. The method is called *matrix multifilters* (MTXFs) and its foundations can be summarized as follows.

Let $\Psi_{ij}(\vec{x})$ be a set (matrix) of $N \times N$ linear filters, and let us define the set of quantities:

$$w_i(\vec{x}) = \sum_j \int d\vec{y}\, \Psi_{ij}(\vec{x} - \vec{y}) D_j(\vec{y}). \tag{11}$$

The quantity above is the sum of a set of linear filterings of the individual channels. We intend to use the combined filtered images $w_i(\vec{x})$ as estimators of the source amplitudes A_i. For that purpose, we require that the filters Ψ_{ij} satisfy the conditions below.

  (i) The combined filtered image at the position of the source $w_i(\vec{0})$ is an *unbiased* estimator of A_i.

 (ii) The combined filtered image at the position of the source $w_i(\vec{0})$ is an *efficient* estimator of A_i, that is, the variance of the estimator is minimum.

Looking at the structure of (11) it is easy to verify that a sufficient condition in order to guarantee the first condition above is

$$\int d\vec{y}\, \Psi_{ij}(\vec{x} - \vec{y}) \tau_j(\vec{y}) = \delta_{ij}(\vec{x}), \tag{12}$$

that is, the filter functions are orthonormal to the source profile functions. This guarantees statistical unbiasedness independently of whatever values **f** may take. The solution to the problem given the two conditions above and the constraints (12) is, in Fourier space:

$$\mathbf{\Psi}^* = \mathbf{\Lambda} \mathbf{P}^{-1}, \tag{13}$$

where $*$ denotes complex conjugation and

$$\Lambda_{lm} = \lambda_{lm} \tau_m,$$
$$\boldsymbol{\lambda} = \mathbf{H}^{-1}, \tag{14}$$
$$H_{lm} = \int d\vec{k}\, \tau_l(\vec{k}) P_{lm}^{-1}(k) \tau_m^*(\vec{k}).$$

In the case where **P** is diagonal (no noise correlation among channels), it can be shown that the matrix multifilters default to a simple matched filter applied individually to each channel.

The MTXF do not project N channels into a single effective plane where to detect the sources. They project instead N channels into N new planes where the sources can be detected and their N amplitudes $A_i (i = 1, \ldots, N)$ can be estimated separately. The difference with the single-channel approach is that each one of the N new planes w_i is constructed by combination of the original N channels in such a way the noise is cancelled more effectively: the MTXFs use the multichannel cross-correlation of noise but do not use at all any information about the spectral behaviour of the sources. It can be described as a "semi-multichannel approach", in the sense that it uses only half of the available information. When the noise cross-correlation among channels is zero, the method defaults to standard single-channel-matched filtering. But when the cross-correlation is not null, the MTXF give better signal-to-noise boosts than the standard single-channel matched filter. Figure 1 shows the comparison between the single-channel matched filter and the MTXF for a simulation of the *Planck* 30, 44, 70 and 100 GHz channels. Note that for the 44 and, 70 GHz channels the output matrix-filtered maps look far cleaner than their matched filtered equivalents. For the 30 GHz channel, the distinction is
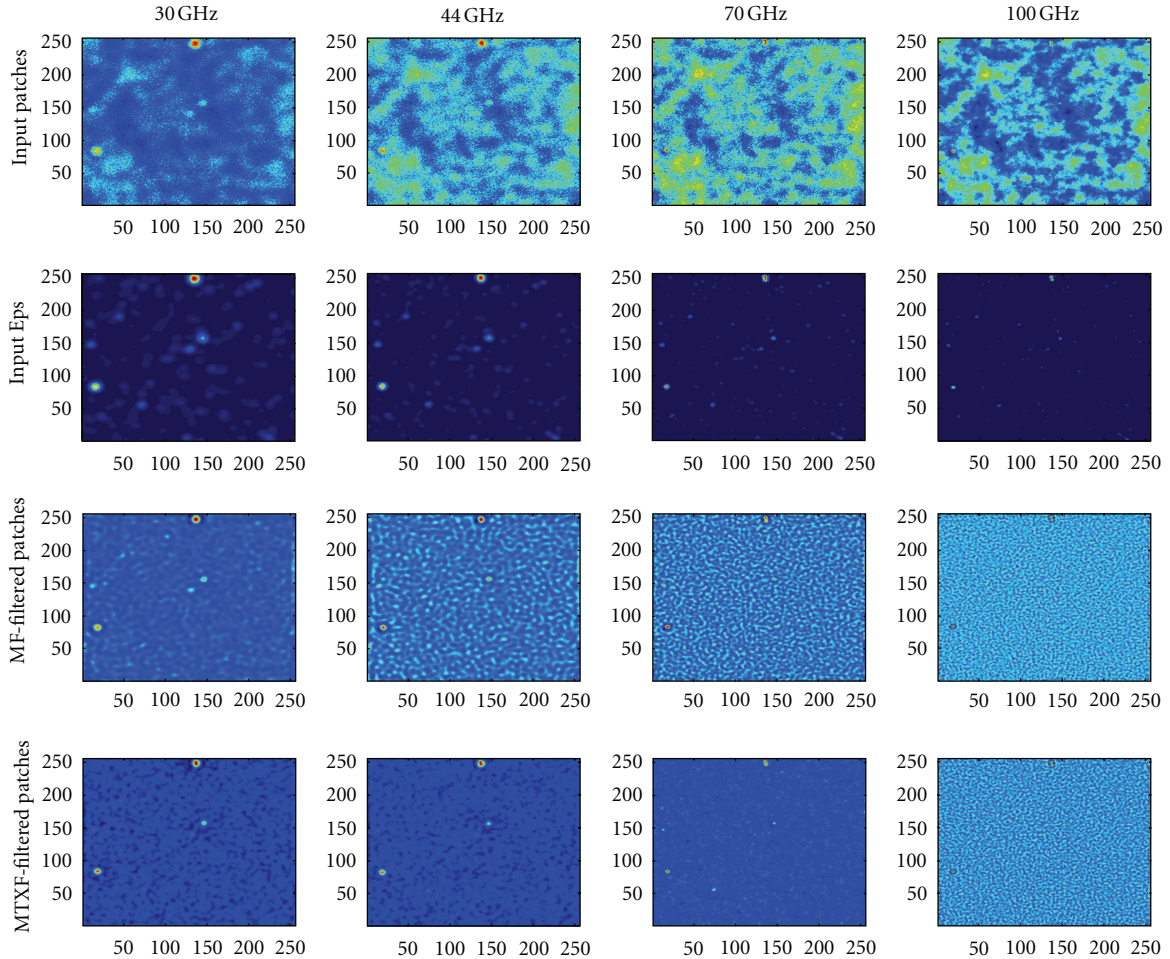
FIGURE 1: Comparison between single-channel-matched filtering and multichannel MTXF filtering for a four-channel simulation of a patch of the sky as observed by the *Planck* 30, 44, 70 and 100 GHz detectors. The top row of images show the simulated patches. The second row shows just the point sources present in the simulations. The third row shows the patches after being filtered with single-channel-matched filters. The fourth row shows the patches filtered with the corresponding MTXF. The *x*-and *y*-axes are in pixel units (pixel size = 3.435 arcmin).

not so clear, but some improvement can be appreciated nevertheless. Finally, for the 100 GHz channel both filtered images look practically identical. The gain factors obtained for these images with the MTXF are [2.9, 3.8, 3.5, 2.8] for the [30, 44, 70, and 100] GHz channels. The signal-to-noise gain ratio between the MTXF and the matched filter is GMTXF/GMF = [1.38, 1.52, 1.49, and 1.00] for the [30, 44, 70, 100] GHz channels. Note that in one of the channels (100 GHz) the signal-to-noise gain is equal to one, although this is not always the case, the MTXFs tend to default to the standard-matched filter for some of the channels when the multichannel mixing does not add information in a constructive way: this is often (but not always) the case for the *Planck* 100 GHz when combined with the lower frequency channels that have worse angular resolution and higher noise levels. This cannot be taken as a general rule, because the conditions change from one region of the sky to another. Only in *the worst* case the 100 GHz-MTXF-filtered image is as good as the single-channel-MF-filtered image. Regarding the number of true

and spurious detections produced by both methods Figure 2 shows the receiver operating characteristic (ROC) curves for the MTXF and the MF in the four considered channels; a clear improvement can be appreciated at 30, 44, and 70 GHz, whereas both methods work similarly for thr 100 GHz case. This serves as an indication of how MTXF can produce results better or at least equal to single-channel-matched filters without making any use of the sources spectral diversity.

5. Matched Multifilter

Multichannel detection techniques reach the summit of their power when they fully exploit the spectral diversity of compact sources with respect to the background. he fully multichannel compact source detection techniques that are being applied in the context of CMB astronomy lie in two main groups: linear filtering methods based on the so-called matched multi-filters plus some thresholding
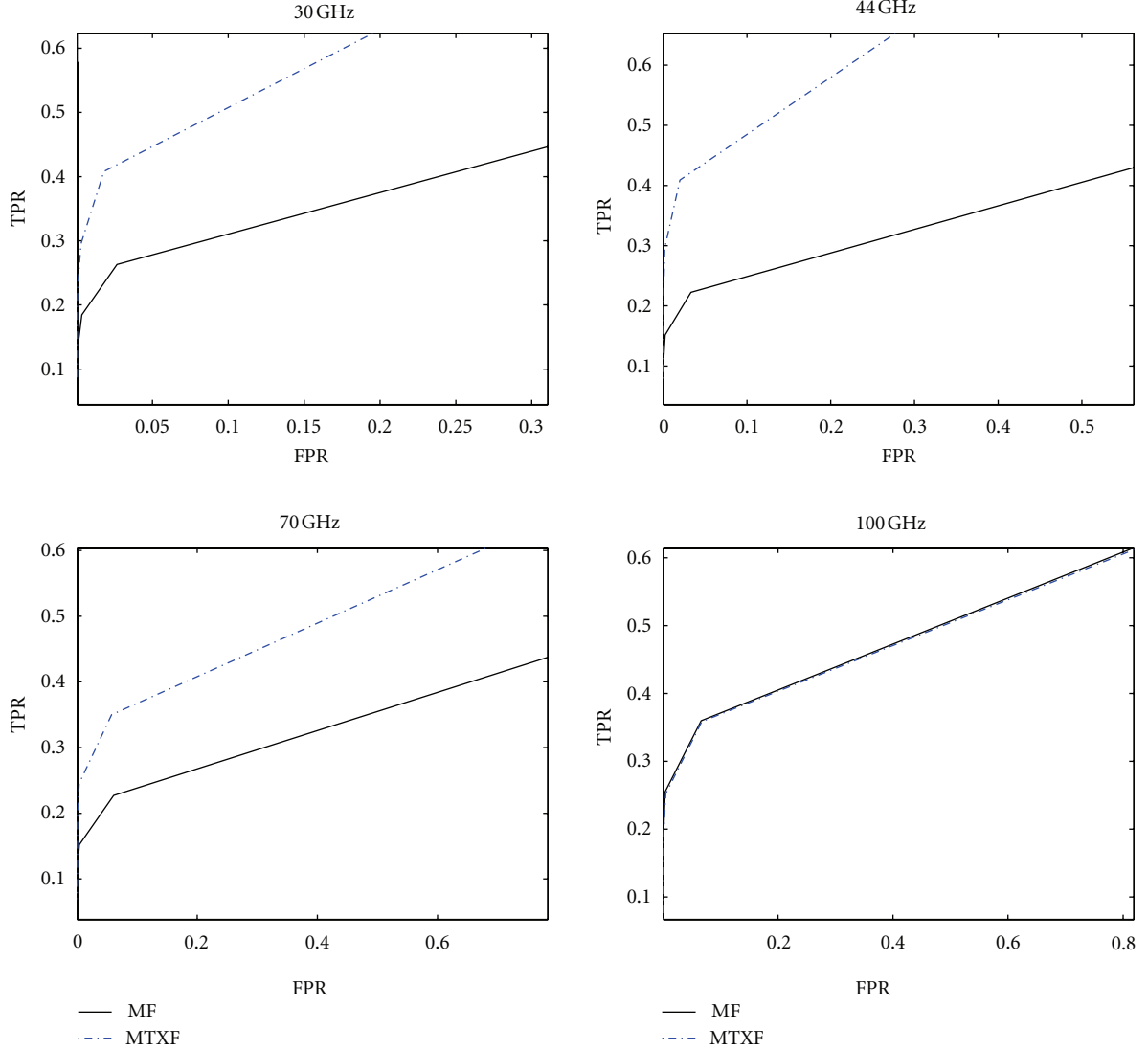
FIGURE 2: ROC curves for the filtering schemes and the four channels, considered. Solid line: MF. dot-dashed line: matrix filters. Axis labels: TPR stands for true positives ratio (associated to the completeness of a catalogue), FPR stands for false positives ratio (associated to the purity of the catalogue, a low FPR indicates a high purity).

detection criteria and Bayesian methods, to be discussed in the next section.

5.1. Standard-Matched Multifilters. Let us assume that the spectral behaviour f of the sources is known *a priori*. An example is the well known thermal Sunyaev-Zel'dovich effect which, ignoring relativistic effects, has a spectral behaviour given by (9). This opens interesting possibilities. For example, if $\mathbf{f}$ is known, it is straightforward to stack the different channels with an optimal weighting designed to minimize the effects of the background while keeping intact the intensity of the sources [45]:

$$\tilde{D}(\vec{x}) = \sum_i w_i D_i(\vec{x}) = \mathbf{w}^t \mathbf{D}(\vec{x}), \qquad (15)$$

such that if there is a source of intensity I_* located at the origin, then

$$\left\langle \tilde{D}(\vec{0}) \right\rangle = I_*. \qquad (16)$$

The solution to this kind of internal linear combination of channels is given by the generalized eigenvalue problem [45]

$$(\mathbf{G} - \lambda \mathbf{M})\mathbf{w} = 0, \qquad (17)$$

where the elements of the matrices $\mathbf{G}$ and $\mathbf{M}$ are given by

$$G_{ij} = f_i \tau_i(0) \tau_j(0) f_j,$$
$$M_{ij} = \left\langle n_i(\vec{x}) n_j(\vec{x}) \right\rangle. \qquad (18)$$

It is evident that $\mathbf{M}$ is a measure of the cross-correlation of the noise among the different channels. The combination (15)

using the weights obtained through (17) gives the optimal internal linear combination (ILC) map for the detection of sources with the spectral behaviour **f**. The ILC map can be further processed using a single-channel-matched filter as suggested by [45]. In an independent work, [40] combined simulated multiwavelength maps in order to increase the average signal-to-noise ratio of galaxies assuming that their spectral behaviour could be modelled by expressions such as (10) with a known spectral index.

In the previous ILC method, however, the separation between linear combination and filtering seems somewhat artificial. The process of filtering and projecting into a single effective plane can be achieved in a single step by means of the so-called *matched multifilters* [45–48]. Let us define the effective filtered plane as a sum of optimally filtered images

$$\hat{D}(\vec{x}; S) = \sum_i \int d\vec{y} D_i(\vec{y}) \psi_i(\vec{x} - \vec{y}; S), \qquad (19)$$

where the ψ_i are N linear filters depending on a certain scale parameter S. The meaning of this scale parameter will become evident shortly. If we impose to the effective filtered plane $\hat{D}$ the usual unbiasedness and efficiency conditions, that is

(i) at the position of the sources, $\hat{D}$ is an unbiased estimator of I_*;

(ii) the variance of $\hat{D}$ is minimum.

The solution, as proven in [45], is given in Fourier space by

$$\boldsymbol{\psi}\left(\vec{k}; S\right) = \alpha(S) \mathbf{P}^{-1}\left(\vec{k}\right) \mathbf{F}\left(\vec{k}; S\right), \qquad (20)$$

$$\alpha^{-1}(S) = \int d\vec{k} \mathbf{F}^t\left(\vec{k}; S\right) \mathbf{P}^{-1}\left(\vec{k}\right) \mathbf{F}\left(\vec{k}; S\right), \qquad (21)$$

where the vector **F** has components $F_i(\vec{k}; S) = f_i \tau_i(\vec{k}; S)$. We show explicitly the dependence on S for a reason that will be clear immediately. The filters defined by (20) take the name of *matched multifilters* (MMF).

Just to show the power of the MMF, in Figures 3 and 4 we show the unfiltered channels of a typical simulated SZ observations and the corresponding filtered plane. Figure 3 shows a small part of a realistic sky simulation obtained with the *Planck* Sky Model (PSM) package [49] for the *Planck* HFI frequencies of 100 143, 217, 353, 545, and 857 GHz. There is a bright galaxy cluster in the center of the images, but it is very hard to spot it visually. On the other hand, the presence of the cluster is evident in the MMF-filtered image shown in Figure 4.

5.2. Objects with Different Size. Until now, we have not made reference to the scale R of the sources that appeared in (1)–(5). In Sections 3 and 4 we have implicitly assumed that all the sources shared the same basic profile τ_0. This is pretty well the case of galaxies in low-resolution experiments such as WMAP and *Planck*: the angular scale of these objects is typically far smaller than the instrument psf and therefore all of them can be safely considered as *point sources*, with observed profiles that are basically equal to the observing beam. In that case, we can effectively forget about the scale R and assume that is something intrinsic to the profile τ_0 that does not require to be made explicit.

But this cannot be the case for galaxy clusters: many of them are resolved objects even at the relatively low angular resolutions of *Planck* [50]. Detection algorithms must therefore provide not only a list of positions and intensities but also of sizes of clusters. Moreover, they should give accurate photometry (values of I_*) for all the range of possible object sizes. MMF can be adapted to do precisely this in a very simple manner.

Let us assume that all the compact sources in a given multichannel image have the same spatial profile except for a scale parameter R. For example we may consider the universal galaxy cluster pressure profile given by [51] with a different contrast radius R_{500} for each cluster in our sample. If the jth cluster has a true value of its scale parameter equal to R_j, then is it straightforward to prove that the filters (20) satisfy the conditions of unbiasedness and maximum efficiency if and only if their scale is $S = R_j$. Moreover, as a part of the demonstration it can be shown that the signal-to-noise boosting given by the filter for that particular cluster is maximum also if and only if $S = R_j$. This suggests a very simple detection/estimation algorithm:

(1) filter the multichannel image with a set of filters (20) with different scales S_i;

(2) select as cluster candidates the positions of the maxima of the filtered images above a certain threshold;

(3) for each cluster candidate j, obtain the curve showing the signal-to-noise boosting versus the scale S_i and find the location $S_{\max}(j)$ of that curvely;

(4) make $R_j = S_{\max}(j)$.

5.3. Unbiased Matched Multifilters. Galaxy clusters interact with CMB photons not only through the thermal SZ effect, but also through the so-called kinematic SZ effect due to the proper motion of the cluster. This is a perfect example of a case in which the same object produces two signals with identical spatial distribution but different spectral behaviours. The mixing of the two signals can affect negatively the determination of each of them. Normally the kinematic effect is at least one order of magnitude fainter than the tSZ and therefore the bias introduced by it in tSZ measurements can be neglected. But the contrary is not true: the tSZ can contaminate significantly the estimation of the kinematic effect. In order to avoid this authors in [46] proposed a modification of MMF specifically tailored to cancel out this bias.

Let us consider a case in which we wish to observe the kinematic SZ effect without being disturbed by tSZ. In thermodynamic units the vector $\mathbf{f} = [f_1, \ldots, f_N]$ of the thermal effect is given by (9). The spectral behaviour of the kinematic effect is, in the same units, flat: $[1, \ldots, 1]$. We look for a set of N filters $\Phi_i, \ldots, \Phi_N$ such that the estimation of the
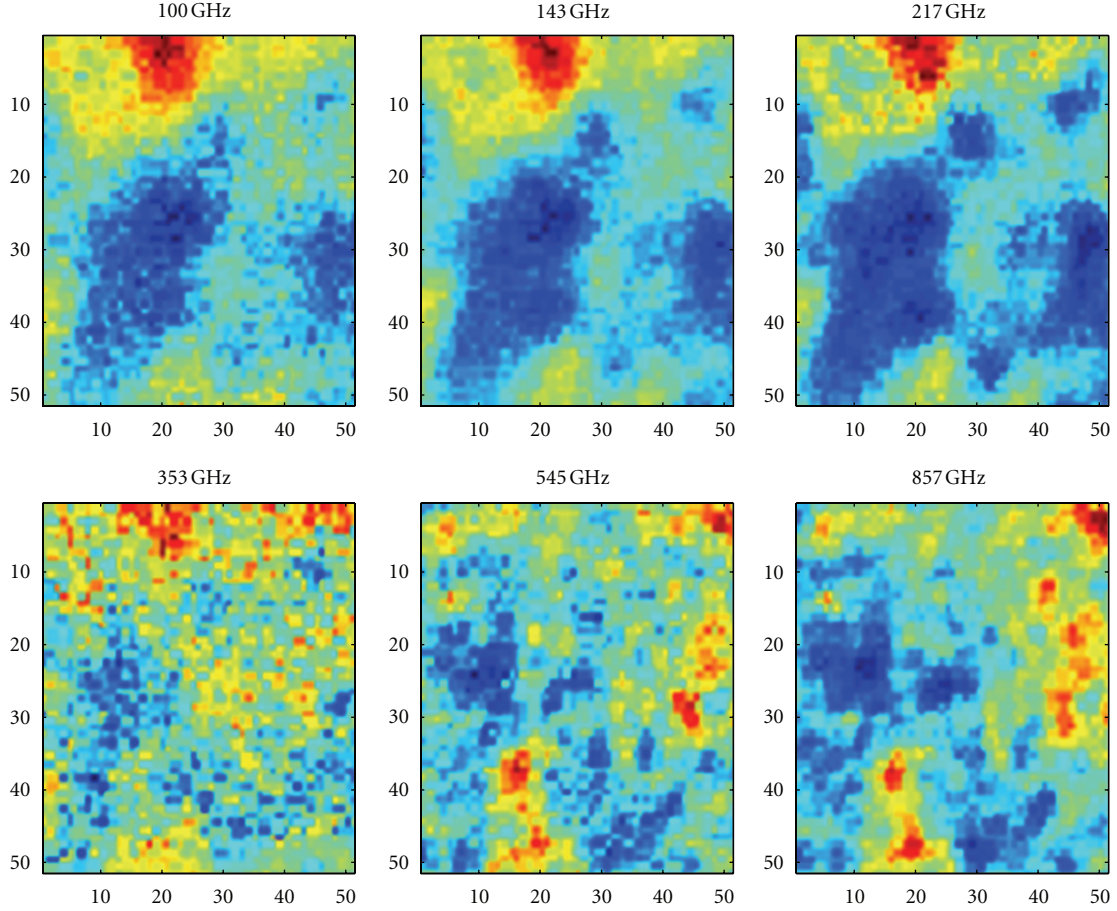
FIGURE 3: Simulated patches of the sky at 100, 143, 217, 353, 545 and 857 GHz. A bright galaxy cluster is in the centre of the patches, but it is almost invisible among the CMB and Galaxy fluctuations.

kinematic effect is not affected by the thermal. A sufficient condition for this is

$$\int d\vec{k}\,\boldsymbol{\tau}^t\boldsymbol{\Phi} = 1,$$

$$\int d\vec{k}\,\mathbf{F}^t\boldsymbol{\Phi} = 0, \tag{22}$$

which can be interpreted as a kind of orthogonality with respect to the spectral behaviour laws of the components. The solution, when we add the maximum efficiency constraint, looks similar to the MMF:

$$\boldsymbol{\Phi} = \frac{1}{\Delta}\mathbf{P}^{-1}(-\beta\mathbf{F} + \alpha\boldsymbol{\tau}). \tag{23}$$

In this equation the constants α, β and Δ are given by

$$\alpha = \int d\vec{k}\,\mathbf{F}^t\mathbf{P}^{-1}\mathbf{F},$$

$$\beta = \int d\vec{k}\,\boldsymbol{\tau}^t\mathbf{P}^{-1}\mathbf{F},$$

$$\gamma = \int d\vec{k}\,\boldsymbol{\tau}^t\mathbf{P}^{-1}\boldsymbol{\tau}, \tag{24}$$

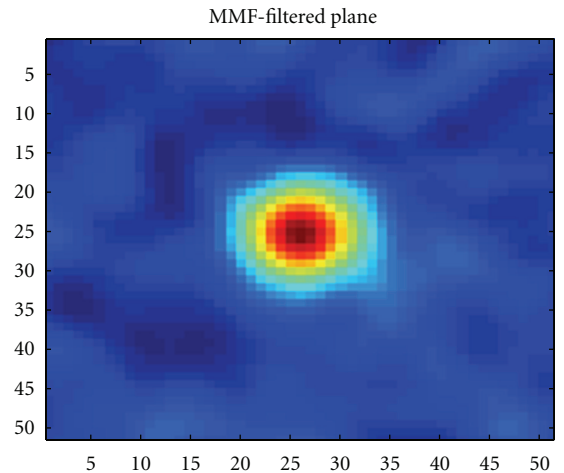$$\Delta = \alpha\gamma - \beta^2.$$

FIGURE 4: The same area of the simulated sky as in Figure 3, after filtering with MMF. Now the galaxy cluster is clearly visible.

Similarly, it is possible, albeit less interesting, to design filters that give the thermal SZ effect while cancelling the bias introduced by the kinematic effect. Both families of filters are called *unbiased matched multifilters* (UMMF).

5.4. Unknown Spectral Behaviour. The prior knowledge of $\mathbf{f}$ is the fundamental key of the success of MMF and the reason why they are able to get high signal-to-noise ratio boosts with respect to the individual channels. However, as we mentioned before this prior knowledge is only certain when considering the SZ effect (and while ignoring relativistic effects). Extragalactic radio and infrared sources are not that simple to model. One possible solution is to group sources in broad families, radio flat, radio steep, dusty galaxies of a certain type, and so forth, and to define average spectral laws, such as (10), for each family. This approach will not be optimal for individual sources and will certainly introduce biases in the photometry but at least is expected to improve the number of detections with respect to single-channel detection.

Fortunately MMF allow us to solve the conundrum in a very elegant way by means of an adaptive filtering scheme similar to the one described in Section 5.2.

Now imagine that $\mathbf{f}$ describes the true (unknown) spectral behaviour of a source and that $\mathbf{g} = [g_i]$, $i = 1, \ldots, N$ is a new vector of equal size as $\mathbf{f}$ but whose elements can take any possible value. We can define the MMF for vector $\mathbf{g}$:

$$\boldsymbol{\psi}_g = \alpha_g \mathbf{P}^{-1} \mathbf{G},$$

$$\alpha_g^{-1} = \int d\vec{k}\, \mathbf{G}^t \mathbf{P}^{-1} \mathbf{G}, \qquad (25)$$

$$G_i = g_i \tau_i,$$

where all the definitions are in Fourier space and we have not written down the explicit dependence on $\vec{k}$ for simplicity. If we apply the filter (25) to a source with true spectral behaviour $\mathbf{f}$ and true intensity I_* we will get

$$I_g = \widehat{D}\left(\vec{0}\right) = I_* \alpha_g \int d\vec{k}\, \mathbf{G}^t \mathbf{P}^{-1} \mathbf{F}, \qquad (26)$$

so $I_g \neq I_*$ unless $\mathbf{g} = \mathbf{f}$. This is no help since we do not know the true value I_*. But it can be shown [52] that the quantity

$$\mathrm{SNR}_g = \frac{I_g}{\sigma_g}, \qquad (27)$$

where σ_g^2 is the variance of the image filtered with (25), is maximum when $\mathbf{g} = \mathbf{f}$. SNR_g is obviously the signal-to-noise ratio of the source after filtering. This allows us to reproduce the same kind of algorithm that was used in Section 5.2 to detect the sources and estimate simultaneously their position, intensity, and spectral behaviour (instead of their scale). The only difference here is that while in Section 5.2 there was only one parameter per cluster to be determined (the scale), here we have $N - 1$ parameters per cluster (components of $\mathbf{f}$) to determine. The method has been recently applied to the two highest frequency channels of WMAP [53], leading to a new catalogue of 157 objects detected at the 5σ level at 94 GHz, which is a substantial improvement over the WMAP Five-Band Point Source Catalogue.

6. Bayesian Multichannel Detection and Estimation

All the multichannel detection methods above have a common two-step methodological approach: in a first moment the data are somehow preprocessed and then objects are detected by means of some criterion, typically thresholding, that hopefully separates them from the background noise. For the preprocessing step, one chooses beforehand a given tool (e.g., linear combinations of channels or a given kind of filter/wavelet) and adjusts a small number of parameters (e.g., the weights of the linear combination, the scale of the wavelet, etc.) according to some optimality criterion (e.g., unbiasedness, maximum efficiency). For the detection step, an arbitrary threshold is chosen with the hope to reach a compromise between minimizing the number of spurious detections and maximizing the number of true detections. A third step for the estimation of a number of parameters of the sources (e.g., intensity, size...) can be attempted after detection or, in some cases, simultaneously to it, (even if *detection* and *estimation* can be done at the same time under certain circumstances, it must be noted that in statistics they are fundamentally different concepts), (as it occurred with MMF). Errors to the estimated parameters are calculated by means of some prescription such as Fisher analysis. All these steps are somewhat arbitrary and are focused only in the statistical properties of the background (and the deterministic spectral behaviour of the sources, in the case of multichannel detection). This approach leaves out all probabilistic (and potentially useful) information about how many sources are expected to be found above a given flux limit, how are they distributed as a function of intensity, how many classes of sources are there and in which proportion are they present in the data, and so forth.

The Bayesian system of inference is the only one that provides a consistent extension of deductive logic to a broader class of degrees-of-belief by mapping them into the real interval $[0, 1]$ [54]. Bayesian inference provides a logically consistent way of tackling the detection problem as a part of the *decision theory*, while incorporating a probabilistic description of the sources as *a priori* information in a natural way. The Bayesian framework also provides a sensible detection criterion through the Bayesian posterior odds ratio [54, 55]. In estimation, Bayesian methods give a full description of the a posteriori distribution of the parameters given the current data, thus allowing us to obtain expectation values, confidence level contours, and any other statistics of interest.

A detailed description of multichannel Bayesian detection of compact sources is given in [55]. Here we will just summarize the main theoretical aspects of the problem. Let us start with Bayes' theorem:

$$Pr(\boldsymbol{\Theta} \mid \mathbf{D}, H) = \frac{Pr(\mathbf{D} \mid \boldsymbol{\Theta}, H) Pr(\boldsymbol{\Theta} \mid H)}{Pr(\mathbf{D} \mid H)}, \qquad (28)$$

where $\mathbf{D}$ is the vector of the observations as in (1) and is the vector $\boldsymbol{\Theta}$ contains all the relevant parameters (positions, intensities, sizes, etc.) to the detection problem and H is the underlining hypothesis. In the usual Bayesian terminology,

$Pr(\Theta \mid \mathbf{D}, H)$ is the *posterior* probability distribution of the parameters, $Pr(\mathbf{D} \mid \Theta, H) \equiv \mathcal{L}(\Theta)$ is the *Likelihood*, $Pr(\Theta \mid H) \equiv Pr(\Theta)$ is the *prior,* and $Pr(\mathbf{D} \mid H) \equiv \mathcal{Z}$ is the *Bayesian evidence.* The most simple detection scenario can be described as a decision between two incompatible hypothesis H_1 (there is a source) and H_0 (there is not a source). In this case, it can be shown [54, 55] that the decision criterion that minimizes the expected loss (i.e., that optimizes the balance between the number of true and false detections) is given by the following condition on the Bayesian odds ratio:

$$\ln\left[\frac{Pr(H_1 \mid \mathbf{D})}{Pr(H_0 \mid \mathbf{D})}\right] = \ln\left[\frac{\mathcal{Z}_1}{\mathcal{Z}_0}\frac{Pr(H_1)}{Pr(H_0)}\right] \underset{H_0}{\overset{H_1}{\gtrless}} \xi, \qquad (29)$$

where ξ is a threshold that is fixed for any given loss function. The choice of the loss function depends on the particular setting of the experiment; a common custom in microwave astronomy is to give identical weight to a spurious detection than to a missing true detection.

6.1. Likelihood.
The form of the likelihood is determined by the statistical properties of the generalised noise (background sky emission plus instrumental noise) in each frequency channel. If the generalised noise is statistically homogeneous it is more convenient to work in Fourier space, since there are no correlations between the Fourier modes of the generalised noise. Please note that although Galactic emission is not homogeneous, one can always operate locally in patches small enough to make the homogeneity assumption valid (as a first approximation). It is also common to assume that both the background emission and instrumental noise are Gaussian random fields. This is a very accurate assumption for instrumental noise and the primordial CMB but more questionable for Galactic emission. Under the previous assumptions, the likelihood ratio between the hypotheses H_1 and H_0 can be written as

$$\ln\left[\frac{\mathcal{L}_{H_1}(\Theta)}{\mathcal{L}_{H_0}(\Theta)}\right] = \sum_{\vec{k}} \mathbf{D}^t\left(\vec{k}\right)\mathbf{P}^{-1}\left(\vec{k}\right)\mathbf{s}\left(\vec{k};\Theta\right)$$
$$- \frac{1}{2}\sum_{\vec{k}} \mathbf{s}^t\left(\vec{k};\Theta\right)\mathbf{P}^{-1}\left(\vec{k}\right)\mathbf{s}\left(\vec{k};\Theta\right). \qquad (30)$$

Note that the effect of the products $\mathbf{X}^t\mathbf{P}^{-1}\mathbf{Y}$, with $\mathbf{X}$ and $\mathbf{Y}$ generic vectors, is to project the multichannel data into one single equivalent channel (or plane), as it happened in the MMF and UMMF. The similarities are more profound. It is shown in [55] that when the source term $\mathbf{s}$ is written as the sum of N_s compact sources distributed across the image, the likelihood ratio (30) can be split into two parts:

(i) a sum of "auto-terms" each containing just the parameters corresponding to the ith source ($i = 1, \ldots, N_s$)

(ii) a sum of "cross-terms" with mixed parameters corresponding to the ith, jth sources ($i, j = 1, \ldots, N_s$).

The cross-term goes quickly to zero when the sources are well separated but must be taken into account when source blending is frequent (crowded fields). It is worth noting that maximising the likelihood ratio (30), in the absence of the cross-term (negligible source blending), with respect to the

source intensities $I_{*,j}$, leads precisely to the MMF described in Section 5. This is the multichannel generalization of the well-known result that matched filtering is the solution of the generalized maximum likelihood test (GLRT) for Gaussian noise and a single source [32, 56].

6.2. Priors.
If the data model provides a good description of the observed data and the signal-to-noise ratio is high, then the likelihood will be very strongly peaked around the true parameter values and the prior will have little or no influence on the posterior distribution. At the faint end of the source population, however, priors will inevitably play an important role. Moreover, since for most cases in astronomy the faint tail overwhelmingly dominates the population, the selection of the priors becomes important and has to be addressed very carefully. Physical (informative) priors are particularly useful when addressing the detection problem, whereas noninformative priors can be more adapted to the task of parameter estimation once the sources have been detected [55]. For multichannel detection, the list of priors to be considered must include the following.

(1) *Prior on the positions*: for extragalactic sources, it is reasonable to assume a uniform prior. Clustering can be a problem, particularly for dusty galaxies observed at $\nu \geq 300$ GHz.

(2) *Prior on the number of sources*: following the same rationale of local uniformity, that is no clustering, the probability of finding N_s objects (above a given flux limit) in a sky patch follows a Poisson distribution

$$\pi(N_s) = Pr(N_s \mid \lambda) = e^{-\lambda}\frac{\lambda^{N_s}}{N_s!}, \qquad (31)$$

where λ is the expected number of sources per patch.

(3) *Prior on intensity*: a usual informative prior on the distribution in intensity of extragalactic sources is a Generalized Cauchy Distribution such as in [31]. This provides a good model for the observed distribution of fluxes, fitting the de Zotti or Tucci models almost perfectly [2, 3]. Moreover, the distribution can be properly normalised as required for evidence evaluation. For the case of galaxy clusters, a power law distribution fits well the cluster populations assuming a standard WMAP best-fit ΛCDM cosmology [57] and the Jenkins mass function [58].

(4) *Prior on the sizes*: Point sources are best modelled by imposing the prior $\pi(R) = \delta(R)$ on the radius. For galaxy clusters, the fact that a significant number of them can be resolved must be taken into account. In [55] a truncated exponential law is found to fit the simulated catalogues very well.

(5) *Prior on the spectral parameters*: for galaxy clusters observed through the thermal Sunyaev-Zel'dovich effect, the spectral behaviour can be safely fixed to (9), if we ignore relativistic effects. For point sources, however, the fact that each source has a spectral behaviour that is different from all the others must

be accommodated. Either if we use power laws such as (10) for radio sources or grey body laws for dusty galaxies, there is at least one spectral parameter that changes from one source to another. Informative priors can be derived from the available source counts models [2, 3], or uniform priors can be used when physical models are not available.

(6) *Priors on the models*: if there is more than a kind of source in the data (e.g., we have radio sources, dusty galaxies, Galactic compact sources and galaxy clusters in the same image) then appropriate priors in the probability of the different hypotheses H_j should be included in the Bayesian framework. More importantly, the Bayesian framework allows us to explicitly consider the probability of background fluctuations above a certain level in order to optimize the decision between the hypotheses H_j and the null hypothesis H_0 (i.e., no source).

6.3. Bayesian Detection Methods in the Literature. Solving the Bayesian equations for object detection and parameter estimation typically implies sampling from a very complex posterior distribution with variable dimensionality (dependent on the number of objects). Thus the main problem that Bayesian methods need to overcome is the computational burden of evaluating integrals over the posterior and its marginals. This can be avoided in some cases by semianalytical Maximum A Posteriori evaluation [31], but if a full analysis of the posterior is required then sampling is unavoidable. Typical implementations include MonteCarlo Markov chain (MCMC) algorithms [27] and posterior refinements such as nested sampling [28, 59]. The authors in [29] implemented a simultaneous multiple minimization code based on Powell's direction set algorithm [60], that is generalized to the multichannel case in [55]. A full description of that algorithm and its catalogue making procedure out of the scope of this paper.

7. Polarized Sources

An interesting case of multichannel setting is the detection of extragalactic sources in polarization data. Mathematically the problem is not different from the general model given in Section 2, but the particularities of polarization observations have motivated the appearance of some specific techniques apart from the ones already discussed.

CMB polarization has been described as the next observational frontier of cosmology. In the last few years, some methods have been specifically developed to address the important problem of detecting compact sources in polarization data. In this section, we will give a brief review of these methods, which have been applied to CMB simulations as well as to real data [61–63].

Polarization of light is conveniently described in terms of the Stokes parameters I, Q, U, and V, (see [64] for an excellent review on CMB polarization). Q and U are the linear polarization parameters and V indicates the

circular polarization. Whereas Q, U, and V depend on the orientation of the receivers, the total polarization, defined as

$$P \equiv \sqrt{Q^2 + U^2 + V^2}, \qquad (32)$$

is invariant with respect to the relative orientation of the receivers and the direction of the incoming signal and therefore has a clear physical meaning. This quantity can be treated as the modulus of a vector. Thus, the methods presented in [61] have to do with the study of a set of images which contain signals whose individual intensities can be considered as components of a vector, but where the quantity of interest is the modulus. First, we will consider the case of linear polarization, $V \equiv 0$, given that the CMB is not circularly polarized in the standard cosmological models [64]. However, since some models predict a possible circular polarization [65], we will comment briefly on this case later.

If we have a compact source embedded in the data of Q and U, these can be expressed in the following way:

$$D_{Q,U}(\vec{x}) = A_{Q,U}\tau(\vec{x}) + n_{Q,U}(\vec{x}). \qquad (33)$$

with $A_{Q,U}$ the compact source intensity in Q and U, $\tau(\vec{x})$, the beam profile and $n_{Q,U}(\vec{x})$ the corresponding noise in both components. The P-map, $P(\vec{x}) \equiv (D_Q^2(\vec{x}) + D_U^2(\vec{x}))^{1/2}$, is characterised by a source at the centre of the image with amplitude $A \equiv (A_Q^2 + A_U^2)^{1/2}$ immersed in nonadditive noise which is correlated with the signal.

We assume the same beam profile for the images in Q and U, as well as a Gaussian independently distributed noise with zero mean and the same variance for Q and U. Given these typical conditions, the distribution of P if a source is present is the Rice distribution:

$$f(P \mid A) \ = \ \frac{P}{\sigma^2} e^{-(A^2+P^2)/2\sigma^2} I_0\left(A\frac{P}{\sigma^2}\right), \qquad (34)$$

where σ is the noise rms deviation and I_0 is the modified Bessel function of zero order. By using the Neyman-Pearson lemma [32], a filter is obtained which produces the maximum likelihood estimator of the amplitude, $\hat{A}$. In the case of a pixelated image, we can write

$$\hat{A}\sum_i \frac{\tau_i^2}{\sigma_i^2} = \sum_i y_i \frac{I_1(\hat{A}y_i)}{I_0(\hat{A}y_i)}, \qquad y_i \equiv \frac{P_i\tau_i}{\sigma_i^2}, \qquad (35)$$

where the indexes refer to the pixels and I_1 refers to the modified Bessel function of the first order. This filter is called the Neyman-Pearson filter (NPF). Alternatively, a matched filter can be applied to each image and then, with the two filtered images Q_{MF}, U_{MF}; a nonlinear fusion $P \equiv (Q_{MF}^2 + U_{MF}^2)^{1/2}$ can be made pixel by pixel. This filter is called the filtered fusion (FF).

Simulations with the *Planck* characteristics showed that the FF performed better than the NPF especially for low polarization fluxes. The FF outperformed the NPF when the power of the detections with a given significance and the flux and position estimation were compared.

If the circular polarization is taken into account, the NPF and the FF can be also calculated. Indeed, the filters can be

computed for the modulus of a vector with any number of components [63]. Simulations showed that in the case of circular polarization, the FF also produces better results than the NPF.

Taking into account the results with simulations, The authors in [62] applied the FF to the detection of polarized sources in the WMAP 5-year data. They detected, with a significance level greater than 99.99% in at least one WMAP channel, 22 objects, 5 of which were doubtful, since they did not have a plausible low frequency counterpart. These detections were a clear advance with respect to the 5 polarized sources listed by the WMAP team when they analysed the same data. The application of a filter targeted specifically for polarization detection had a clear influence on the improvement.

The application of these methods to Planck data can be expected in the future. It would also be interesting to combine them with a Bayesian approach, so that some previous information about polarization properties of the sources could be taken into account.

8. Final Remarks

CMB image processing is a relatively young area of research. Much work remains to be done. There is an incipient but resolved interest among CMB cosmologists to incorporate the newest ideas to solving the problem of compact sources in microwave images.

In this paper we have reviewed the status of the different algorithms and methods that attempt the detection of extra-galactic compact sources (galaxies and clusters of galaxies) using multichannel (i.e., obtained by means of more than one single detector) data. This should not be confounded with the traditional band-merging approach to catalogue making. We included the case of detection in polarized data because it is formally equivalent to the detection in multi-frequency experiments. The techniques we have reviewed include linear fusion of images, different multichannel filtering methods and Bayesian object detection. Although the area of research is young, all of these methods have already been applied to astronomical data sets with great success. In next years, however, we expect to see new interesting developments in the field.

One of the most promising ideas for future research is related to the notions of sparsity and l_p-approximations. For the particular case of point like objects, the idea of sparse dictionaries comes naturally [66]. However, the full application to multichannel CMB compact source detection has not been addressed yet. Wavelet techniques, that are very popular in single-frequency source detection, open another interesting possibility that has not been explored yet. The same applies to other space-scale representations such as Gabor and Wigner-Ville transforms. Undoubtedly, both Bayesian and multi-filtering techniques will see substantial improvements in the next few years. All these new developments will be really come in handy for the new generation of many channel cosmological surveys such as the upcoming J-PAS (http://www.j-pas.org/).

List of Symbols

$\vec{x}, \vec{y}$:	Position vector
$\vec{k}$:	Wave vector (Fourier)
N:	Number of channels
D_j:	Observed data at the jth channel
s_j:	Signal at the jth channel
n_j:	Noise at the jth channel
A_j:	Source amplitude at the jth channel
τ_j:	Source profile at the jth channel
τ^0:	Intrinsic source profile (before smearing by the beam)
b_j:	Beam point spread function at the jth channel
R:	Scale parameter
$\mathbf{P}$:	Cross-power spectrum matrix
$\mathbf{f}$:	Spectral behaviour of the source
$\boldsymbol{\Psi}, \boldsymbol{\Phi}$:	Matrix (or vector) of filters
SNR:	Signal-to-noise ratio
$\boldsymbol{\Theta}$:	Vector of parameters (position, size, amplitude…)
I, Q, U, V:	Stokes' polarization parameters
P:	Total polarization
I_0:	Modified Bessel function of zero order.

References

[1] L. Toffolatti, F. Argüeso Gómez, G. De Zotti et al., "Extra-galactic source counts and contributions to the anisotropies of the cosmic microwave background: predictions for the planck surveyor mission," *Monthly Notices of the Royal Astronomical Society*, vol. 297, no. 1, pp. 117–127, 1998.

[2] G. De Zotti, R. Ricci, D. Mesa et al., "Predictions for high-frequency radio surveys of extragalactic sources," *Astronomy and Astrophysics*, vol. 431, no. 3, pp. 893–903, 2005.

[3] M. Tucci, L. Toffolatti, G. de Zotti, and E. Martínez-González, "High frequency predictions for number counts and spectral properties of extragalactic radio sources. New evidence of

a break at mm wavelengths in spectra of bright blazar sources," *Astronomy & Astrophysics*, vol. 533, article A57, 21 pages, 2011.

[4] M. Tucci, E. Martínez-González, L. Toffolatti, J. González-Nuevo, and G. De Zotti, "Predictions on the high-frequency polarization properties of extragalactic radio sources and implications for polarization measurements of the cosmic microwave background," *Monthly Notices of the Royal Astronomical Society*, vol. 349, no. 4, pp. 1267–1277, 2004.

[5] M. Tucci, E. Martínez-González, P. Vielva, and J. Delabrouille, "Limits on the detectability of the cmb b-mode polarization imposed by foregrounds," *Monthly Notices of the Royal Astronomical Society*, vol. 360, no. 3, pp. 935–949, 2005.

[6] F. Argüeso, J. González-Nuevo, and L. Toffolatti, "Contributions of point extragalactic sources to the cosmic microwave background bispectrum," *Astrophysical Journal Letters*, vol. 598, no. 1, pp. 86–96, 2003.

[7] E. Komatsu, A. Kogut, M. R. Nolta et al., "First-year wilkinson microwave anisotropy probe (wmap) observations: tests of gaussianity," *Astrophysical Journal, Supplement*, vol. 148, no. 1, pp. 119–134, 2003.

[8] S. Smith, G. Rocha, A. Challinor et al., "Estimating the bispectrum of the very small array data," *Monthly Notices of the Royal Astronomical Society*, vol. 352, no. 3, pp. 887–902, 2004.

[9] R. B. Barreiro, E. Martínez-González, P. Vielva, and M. P. Hobson, "Effect of component separation on the temperature distribution of the cosmic microwave background," *Monthly Notices of the Royal Astronomical Society*, vol. 368, no. 1, pp. 226–246, 2006.

[10] F. Argüeso, J. L. Sanz, R. B. Barreiro, D. Herranz, and J. González-Nuevo, "Statistical analysis of undetected point sources in cosmic microwave background maps," *Monthly Notices of the Royal Astronomical Society*, vol. 373, no. 1, pp. 311–320, 2006.

[11] D. Babich and E. Pierpaoli, "Point source contamination in cmb non-gaussianity analyses," *Physical Review D*, vol. 77, no. 12, Article ID 123011, 2008.

[12] F. Lacasa, N. Aghanim, M. Kunz, and M. Frommert, "Characterization of the non-Gaussianity of radio and IR point sources at CMB frequencies," *Monthly Notices of the Royal Astronomical Society*, vol. 421, no. 3, pp. 1982–1995, 2012.

[13] P. Vielva, R. B. Barreiro, M. P. Hobson et al., "Combining maximum-entropy and the mexican hat wavelet to reconstruct the microwave sky," *Monthly Notices of the Royal Astronomical Society*, vol. 328, no. 1, pp. 1–16, 2001.

[14] S. M. Leach, J. F. Cardoso, C. Baccigalupi et al., "Component separation methods for the planck mission," *Astronomy and Astrophysics*, vol. 491, no. 2, pp. 597–615, 2008.

[15] J. E. Carlstrom, G. P. Holder, and E. D. Reese, "Cosmology with the sunyaev-zel'dovich effect," *Annual Review of Astronomy and Astrophysics*, vol. 40, pp. 643–680, 2002.

[16] G. De Zotti, M. Massardi, M. Negrello, and J. Wall, "Radio and millimeter continuum surveys and their astrophysical implications," *Astronomy and Astrophysics Review*, vol. 18, no. 1-2, pp. 1–65, 2010.

[17] S. Eales, L. Dunne, D. Clements et al., "The herschel atlas," *Publications of the Astronomical Society of the Pacific*, vol. 122, no. 891, pp. 499–515, 2010.

[18] E. Bertinl, "Sextractor: software for source extraction," *Astronomy and Astrophysics Supplement Series*, vol. 117, no. 2, pp. 393–404, 1996.

[19] M. Tegmark and A. De Oliveira-Costa, "Removing point sources from cosmic microwave background maps," *Astrophysical Journal Letters*, vol. 500, no. 2, pp. L83–L86, 1998.

[20] J. L. Sanz, D. Herranz, and E. Martínez-Gónzalez, "Optimal detection of sources on a homogeneous and isotropic background," *Astrophysical Journal Letters*, vol. 552, no. 2, pp. 484–492, 2001.

[21] D. Herranz, J. Gallegos, J. L. Sanz, and E. Martínez-González, "Point source detection and extraction from simulated planck time-ordered data using optimal adaptive filters," *Monthly Notices of the Royal Astronomical Society*, vol. 334, no. 3, pp. 533–541, 2002.

[22] D. Herranz, J. L. Sanz, R. B. Barreiro, and E. Martínez-González, "Scale-adaptive filters for the detection/separation of compact sources," *Astrophysical Journal Letters*, vol. 580, no. 1, pp. 610–625, 2002.

[23] M. López-Caniego, D. Herranz, R. B. Barreiro, and J. L. Sanz, "Filter design for the detection of compact sources based on the neyman-pearson detector," *Monthly Notices of the Royal Astronomical Society*, vol. 359, no. 3, pp. 993–1006, 2005.

[24] M. López-Caniego, D. Herranz, J. González-Nuevo et al., "Comparison of filters for the detection of point sources in planck simulations," *Monthly Notices of the Royal Astronomical Society*, vol. 370, no. 4, pp. 2047–2063, 2006.

[25] P. Vielva, E. Martínez-González, J. E. Gallegos, L. Toffolatti, and J. L. Sanz, "Point source detection using the spherical mexican hat wavelet on simulated all-sky planck maps," *Monthly Notices of the Royal Astronomical Society*, vol. 344, no. 1, pp. 89–104, 2003.

[26] J. González-Nuevo, F. Argüeso, M. López-Caniego et al., "The mexican hat wavelet family: application to point-source detection in cosmic microwave background maps," *Monthly Notices of the Royal Astronomical Society*, vol. 369, no. 4, pp. 1603–1610, 2006.

[27] M. P. Hobson and C. McLachlan, "A bayesian approach to discrete object detection in astronomical data sets," *Monthly Notices of the Royal Astronomical Society*, vol. 338, no. 3, pp. 765–784, 2003.

[28] F. Feroz and M. P. Hobson, "Multimodal nested sampling: an efficient and robust alternative to markov chain monte carlo methods for astronomical data analyses," *Monthly Notices of the Royal Astronomical Society*, vol. 384, no. 2, pp. 449–463, 2008.

[29] P. Carvalho, G. Rocha, and M. P. Hobson, "A fast bayesian approach to discrete object detection in astronomical data sets—Powellsnakes I," *Monthly Notices of the Royal Astronomical Society*, vol. 393, no. 3, pp. 681–702, 2009.

[30] M. P. Hobson, G. Rocha, and R. S. Savage, *Bayesian Source Extraction*, chapter 7, Cambridge University Press, Cambridge, UK, 2009.

[31] F. Argüeso, E. Salerno, D. Herranz, J. L. Sanz, E. E. Kuruoğlu, and K. Kayabol, "A bayesian technique for the detection of point sources in cosmic microwave background maps," *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 1, pp. 410–417, 2011.

[32] D. Herranz and P. Vielva, "Cosmic microwave background images," *IEEE Signal Processing Magazine*, vol. 27, no. 1, pp. 67–75, 2010.

[33] C. L. Bennett, M. Bay, M. Halpern et al., "The microwave anisotropy probe mission," *Astrophysical Journal Letters*, vol. 583, no. 1, pp. 1–23, 2003.

[34] J. A. Tauber, N. Mandolesi, J. L. Puget et al., "Planck prelaunch status: the planck mission," *Astronomy and Astrophysics*, vol. 520, no. 1, Article ID aa12983-09, 2010.

[35] V. Stolyarov, M. P. Hobson, M. A. J. Ashdown, and A. N. Lasenby, "All-sky component separation for the planck

mission," *Monthly Notices of the Royal Astronomical Society*, vol. 336, no. 1, pp. 97–111, 2002.

[36] D. Maino, A. Farusi, C. Baccigalupi et al., "All-sky astrophysical component separation with fast independent component analysis (fastica)," *Monthly Notices of the Royal Astronomical Society*, vol. 334, no. 1, pp. 53–68, 2002.

[37] L. Bedini, D. Herranz, E. Salerno, C. Baccigalupi, E. E. Kuruoğlu, and A. Tonazzini, "Separation of correlated astrophysical sources using multiple-lag data covariance matrices," *EURASIP Journal on Applied Signal Processing*, vol. 2005, no. 15, pp. 2400–2412, 2005.

[38] A. Bonaldi, S. Ricciardi, S. Leach, F. Stivoli, C. Baccigalupi, and G. De Zotti, "Wmap 3-yr data with correlated component analysis: anomalous emission and impact of component separation on the cmb power spectrum," *Monthly Notices of the Royal Astronomical Society*, vol. 382, no. 4, pp. 1791–1803, 2007.

[39] F. K. Hansen, A. J. Banday, H. K. Eriksen, K. M. Górski, and P. B. Lilje, "Foreground subtraction of cosmic microwave background maps using WI-FIT (wavelet-based high-resolution fitting of internal templates)," *Astrophysical Journal*, vol. 648, no. 2, pp. 784–796, 2006.

[40] P. Naselsky, D. Novikov, and J. Silk, "A combined multifrequency map for point source subtraction," *Monthly Notices of the Royal Astronomical Society*, vol. 335, no. 3, pp. 550–554, 2002.

[41] X. Chen and E. L. Wright, "Extragalactic point-source search in wmap 61 and 94 Ghz data," *Astrophysical Journal Letters*, vol. 681, no. 2, pp. 747–755, 2008.

[42] E. L. Wright, X. Chen, N. Odegard et al., "Five-year wilkinson microwave anisotropy probe observations: source catalog," *Astrophysical Journal, Supplement Series*, vol. 180, no. 2, pp. 283–295, 2009.

[43] D. Herranz and J. L. Sanz, "Matrix filters for the detection of extragalactic point sources in cosmic microwave background images," *IEEE Journal on Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 727–734, 2008.

[44] D. Herranz, M. López-Caniego, J. L. Sanz, and J. González-Nuevo, "A novel multifrequency technique for the detection of point sources in cosmic microwave background maps," *Monthly Notices of the Royal Astronomical Society*, vol. 394, no. 1, pp. 510–520, 2009.

[45] D. Herranz, J. L. Sanz, M. P. Hobson et al., "Filtering techniques for the detection of sunyaev-zel'dovich clusters in multifrequency maps," *Monthly Notices of the Royal Astronomical Society*, vol. 336, no. 4, pp. 1057–1068, 2002.

[46] D. Herranz, J. L. Sanz, R. B. Barreiro, and M. López-Caniego, "The estimation of the sunyaev-zel'dovich effects with unbiased multifilters," *Monthly Notices of the Royal Astronomical Society*, vol. 356, no. 3, pp. 944–954, 2005.

[47] J. B. Melin, J. G. Bartlett, and J. Delabrouille, "Catalog extraction in sz cluster surveys: a matched filter approach," *Astronomy and Astrophysics*, vol. 459, no. 2, pp. 341–352, 2006.

[48] B. M. Schäfer, C. Pfrommer, M. Bartelmann, V. Springel, and L. Hernquist, "Detecting Sunyaev-Zel'dovich clusters with planck -I. construction of all-sky thermal and kinetic SZ maps," *Monthly Notices of the Royal Astronomical Society*, vol. 370, no. 3, pp. 1309–1323, 2006.

[49] Planck Collaboration, "The pre-launch planck sky model: a model of sky emission at submillimetre to centimetre wavelengths," in preparation.

[50] P. A. R. Ade, N. Aghanim, M. Arnaud et al., "Planck Early Results. VII. The Early Release Compact Source Catalog," *Astronomy & Astrophysics*, vol. 536, article A7, 26 pages, 2011.

[51] M. Arnaud, G. W. Pratt, R. Piffaretti, H. Böhringer, J. H. Croston, and E. Pointecouteau, "The universal galaxy cluster pressure profile from a representative sample of nearby systems (rexcess) and the Y_{SZ} - M_{500} relation," *Astronomy and Astrophysics*, vol. 517, no. 15, article no. A92, 2010.

[52] L. F. Lanz, D. Herranz, J. L. Sanz, J. González-Nuevo, and M. López-Caniego, "A multifrequency method based on the matched multifilter for the detection of point sources in cmb maps," *Monthly Notices of the Royal Astronomical Society*, vol. 403, no. 4, pp. 2120–2130, 2010.

[53] L. F. Lanz, D. Herranz, M. López-Caniego et al., "Extragalactic point source detection in WMAP 7-year data at 61 and 94 GHz," http://arxiv.org/abs/1110.6877.

[54] E. Jaynes and G. Bretthorst, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, UK, 2003.

[55] P. Carvalho, G. Rocha, M. Hobson, and A. Lasenby, "Powell-Snakes II—Multi-frequency detection," in preparation.

[56] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*, Prentice Hall, Upper Saddle River, NJ, USA, 1998.

[57] E. Komatsu, K. M. Smith, J. Dunkley et al., "Seven-year wilkinson microwave anisotropy probe (WMAP) observations: cosmological interpretation," *Astrophysical Journal, Supplement Series*, vol. 192, no. 2, article 18, 2011.

[58] A. Jenkins, C. S. Frenk, S. D. M. White et al., "The mass function of dark matter haloes," *Monthly Notices of the Royal Astronomical Society*, vol. 321, no. 2, pp. 372–384, 2001.

[59] F. Feroz, M. P. Hobson, and M. Bridges, "Multinest: an efficient and robust bayesian inference tool for cosmology and particle physics," *Monthly Notices of the Royal Astronomical Society*, vol. 398, no. 4, pp. 1601–1614, 2009.

[60] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran: The Art of Scientific Computing*, W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, Eds., Cambridge University Press, Cambridge, UK, 1992.

[61] F. Argüeso, J. L. Sanz, D. Herranz, M. López-Caniego, and J. González-Nuevo, "Detection/estimation of the modulus of a vector. Application to point-source detection in polarization data," *Monthly Notices of the Royal Astronomical Society*, vol. 395, no. 2, pp. 649–656, 2009.

[62] M. López-Caniego, M. Massardi, J. Gonzlez-Nuevo et al., "Polarization of the wmap point sources," *Astrophysical Journal Letters*, vol. 705, no. 1, pp. 868–876, 2009.

[63] F. Argüeso, J. Luis Sanz, and D. Herranz, "Filter design for the detection/estimation of the modulus of a vector: application to polarization data," *Signal Processing*, vol. 91, no. 7, pp. 1527–1534, 2011.

[64] M. Kamionkowski, A. Kosowsky, and A. Stebbins, "Statistics of cosmic microwave background polarization," *Physical Review D*, vol. 55, no. 12, pp. 7368–7388, 1997.

[65] A. Cooray, A. Melchiorri, and J. Silk, "Is the cosmic microwave background circularly polarized?" *Physics Letters B*, vol. 554, no. 1-2, pp. 1–6, 2003.

[66] F. Martinelli and J. Sanz, "Sparse Representations versus the Matched Filter," in *Signal Processing with Adaptive Sparse Structured Representations (SPARS '09)*, R. Gribonval, Ed., Inria Rennes-Bretagne Atlantique, Saint Malo, France, 2009.