Universidad de Cantabria

Departamento de Física Moderna

CSIC - Universidad de Cantabria

Instituto de Física de Cantabria

Detection of Point Sources in Maps of the Cosmic Microwave Background Radiation by means of Optimal Filters

A dissertation submited in partial of the requirements

for the degree of Doctor of Philosophy in Physics

by

Marcos López-Caniego Alcarria

2006

CHAPTER 3

Biparametric Scale Adaptive Filter: 1D approach

This chapter considers the problem of compact source detection on a Gaussian background. We make a one-dimensional treatment. Two relevant aspects of this problem are considered: the design of the detector and the filtering of the data. Our detection scheme is based on local maxima and it takes into account not only the amplitude but also the curvature of the maxima. A Bayesian Neyman-Pearson test is used to define the region of acceptance, that is given by a sufficient linear detector that is independent on the amplitude distribution of the sources. We study how detection can be enhanced by means of linear filters with a scaling parameter and compare some of them that have been proposed in the literature (the Mexican Hat wavelet, the matched and the scale-adaptive filters). We introduce a new filter, that depends on two free parameters (biparametric scale-adaptive filter). The value of these two parameters can be determined, given the a priori pdf of the amplitudes of the sources, such that the filter optimizes the performance of the detector in the sense that it gives the maximum number of real detections once fixed the number density of spurious sources. The new filter includes as particular cases the standard matched filter, the scale-adaptive filter and the Mexican Hat wavelet. Then, by construction, the biparametric scale adaptive filter outperforms all the other considered filters (including the standard matched filter). The combination of a detection scheme that includes information on the curvature and a flexible filter that incorporates two free parameters (one of them a scaling) improves significantly the number of detections in some interesting cases. In particular, for the case of weak sources embedded in white noise the improvement with respect to the

standard matched filter is of the order of 40%. Finally, an estimation of the amplitude of the source (most probable value) is introduced and it is proven that such an estimator is unbiased and it has maximum efficiency. We perform numerical simulations to test these theoretical ideas in a practical example and conclude that the results of the simulations agree with the analytical ones, see López-Caniego et al. [87].

3.1 Introduction

The detection, identification and removal of the extragalactic point sources (EPS) is fundamental for the study of the Cosmic Microwave Background Radiation (CMB) data (Franceschini et al. [47], Toffolatti et al. [141], De Zotti et al. [31]). In particular, the contribution of EPS is expected to be very relevant at the lower and higher frequency channels of the future ESA Planck Mission [96, 115]).

The heterogeneous nature of the EPS that appear in CMB maps as well as their unknown spatial distribution make difficult to separate them from the other physical components (CMB, Galactic dust, synchrotron, etc) by means of statistical component separation methods. Techniques based on the use of linear filters, however, are well-suited for the task of detecting compact spikes on a background. Several techniques based on different linear filters have been proposed in the literature: the Mexican Hat Wavelet (MHW,Cayón et al. [19], Vielva et al. [144, 145]), the classic *matched* filter (MF, Tegmark & de Oliveira-Costa [139]), the Adaptive Top Hat Filter (Chiang et al. [20]) and the scale-adaptive filter (SAF, Sanz, Herranz, & Martínez-Gónzalez [124]). A certain deal of controversy has appeared about which one, if any, of the previous filters is *optimal* for the detection of point sources in CMB data.

In order to answer that question it is necessary to consider first a more fundamental issue, the concept of *detection* itself. The detection process can be posed as follows: given an observation, the problem is to *decide* whether or not a certain signal was present at the input of the receiver. The decision is not obvious since the observation is corrupted by a random process that we call 'noise' or 'background'.

Formally, the *decision* is performed by choosing between two complementary hypotheses: that the observed data is originated by the background alone (*null hypothesis*), and the hypothesis that the observation corresponds to a combination of the background and the signal. To decide, the detector should use the totality of the available information in terms of the probabilities of both hypotheses given the data. The *decision device* separates the space $\mathcal{R}$ of all possible observations in two disjoint subspaces, $\mathcal{R}_*$ and $\mathcal{R}_-$, so that if an observation $y \in \mathcal{R}_-$ the null hypothesis is accepted, and if $y \in \mathcal{R}_*$

the null hypothesis is rejected, that is, a source is 'detected'. Hence, we will call any generic decision device of this type a *detector*.

Any detector can produce two kinds of errors: on the one hand, it can produce a *false alarm* or *spurious detection* when an observation in which no source was present is assigned to the subspace $\mathcal{R}_*$. The probability of this kind of error depends on the statistical properties of the background and the choice of the detector. On the other hand, a signal that is present in the observation can be missed by the detector (i.e. the observation is wrongly assigned to the subspace $\mathcal{R}_-$). This error is often referred to as *false dismissal*. The probability of false dismissal depends on the statistical properties of the background, the choice of the detector and the properties of the signal (for example, its intensity). In general, it is not possible to decrease the incidence of both types of error at the same time: one of them can be reduced at expense of increasing the other. The goodness of a given detector must be established by taking into account the balance between these two types of error.

The simplest example of detector, and one that has been exhaustively used in Astronomy, is *thresholding*. Thresholding considers that the space $\mathcal{R}$ of observations consists of all the possible values of the measured intensity ξ (in the case of an astronomical image) and subdivides this space into two simple regions $\mathcal{R}_- \equiv \{\xi \in \mathcal{R} : \xi < \xi_*\}$ and $\mathcal{R}_* \equiv \{\xi \in \mathcal{R} : \xi \geq \xi_*\}$. The value ξ_* is an arbitrarily chosen *threshold* that is often expressed as a number of times the standard deviation of the background, $\xi_* = \nu_* \sigma_0$. Thresholding works on the assumption that the probability of finding a value of ξ due to the background decreases as the value of ξ increases. In the case of a Gaussian background, this assumption has a very precise meaning and it allow us to control straightforwardly the probability of occurrence of spurious detections simply by setting a large enough threshold. However, this may lead to a very high probability of false dismissals.

Unfortunately, in many cases the sources are very faint and this makes very difficult to detect them: a high threshold means that the number of detections will be very small. Here is where *filtering* enters in scene. The role of filtering is to transform the data in such a way that a detector can perform better than before filtering. For example, a filter can be designed to reduce the fluctuations of the background so that we can safely use lower detection thresholds and, hopefully, increase the number of detections without increasing the number of spurious detections. We remark that *a filter is not a detector*: the decision device we call 'detector' can be applied after the application of any imaginable filter, or even no filter at all, while the use of any filter without a posterior detection criterion means nothing. However, the two different steps in the

process (filtering and detection) are not independent. In the thresholding example, the use of a filter that cancels most of the fluctuations in the background allows us to change the detection threshold from its original high value to a lower one. Given an adopted detector and a background, it is licit to ask which is the filter that creates the most favorable conditions in the filtered background for the detector to perform. In other words, the 'optimality' of a filter for detection depends on the type of detector chosen which, in turn, depends on the specific *goal* that the observer has in mind: in certain cases the observer will accept a relatively large number of spurious detections in order to have a large number of true targets, whereas in other cases it could be more important to be certain that the detections are all of them reliable, and so on.

For example, let us consider that we have chosen thresholding as our detection device. In that case, it is well known that the optimal linear filter is the matched filter. It produces the maximum amplification of the signal with respect to the background fluctuations, so that the threshold for a given probability of spurious detections is minimum, allowing the thresholding detector to find more sources than would be detected if we filtered with any other linear filter. A sub-optimal approach is to select *a priori* a filter and adapt its scale in order to produce a maximum amplification in a given background. Such is the case of the Mexican Hat Wavelet at the optimal scale (MHO, Vielva et al. [144, 145]) and the Adaptive Top Hat Filter (ATHF, Chiang et al. [20]).

Thresholding has a number of advantages, among them the facts that it is straightforward, it has an obvious meaning in the case of Gaussian backgrounds, and it has been successfully used for many years in many fields of science. However, it does not use all the available information contained in the data to perform decisions. Therefore, one can ask the following question: is it possible to devise a detector that uses additional information apart from mere intensities and that produces better results than thresholding? And, if so, which is the filter that optimizes the performance of such a detector?

Let us focus on the case of one-dimensional data (such as stellar spectra or time-ordered scannings of the sky in CMB experiments) and linear filters. Data in a one-dimensional array is entirely described by two quantities, namely the position in the array (corresponding to the spatial or temporal coordinate, for simplicity we will refer it as *spatial* information) and the value (intensity) at each position. Thresholding uses only the intensity distribution to make the decision. Clearly, the inclusion of spatial information in a detector should be useful. For example, it could help to distinguish the sources from fluctuations in the background with similar scale but a different shape. A full description of this 'spatial information' should include the probability distribution of events (both due to background and sources) in space, with all its infinite moments.

We will somewhat relax this demand of information assuming that the background is homogeneous and isotropic, and asking at each point for some information about the shape of the sources (e. g., the curvature of the peaks) and the autocorrelation of the background.

In fact, even a simple filtering-and-thresholding scheme uses implicitly some degree of spatial information. Both the MHO and the ATHF adapt to the scale at which the contrast between sources and background produces the maximum amplification. The MF includes as well the information on the profile of the sources in order to amplify the structures whose shape correlates with the shape of the sources. SAF goes a step further in constraining additionally the scale of the filter. Moreover, in most cases the detection is performed not in all the points of the data but only in the peaks, that is, *in those points where the curvature is positive*.

An example of detector that uses more information than simple thresholding is given by the Neyman-Pearson decision rule:

$$L(x_i) = \frac{p(x_i|H_1)}{p(x_i|H_0)} \geq L_* \tag{3.1.1}$$

where $L(x_i)$ is called the likelihood ratio, $p(x_i|H_0)$ is the probability density function (pdf) associated to the null hypothesis (i.e. there is no source) whereas $p(x_i|H_1)$ is the pdf corresponding to the alternative hypothesis (i.e. there is a source). x_i are a set of variables which are measured from the data (e.g. amplitude, curvature). L_* is an arbitrary constant, which defines the region of acceptance, and must be fixed using some criterion.

In a recent work, Barreiro et al. [5] introduce a novel technique for the detection of sources based on the study of the number density of maxima for the case of a Gaussian background in the presence or absence of a source. In order to define the region of acceptance the Neyman-Pearson decision rule is used with pdf's associated to the previous number densities and using the information of both the intensities ξ and the curvatures κ of the peaks in a data set. In addition, L_* is fixed by maximising the *significance*, which is the weighted difference between the probabilities of having and not having a source. In that work the performances of several filters (SAF, MF and MHW) is compared in terms of their *reliability*, defined as the ratio between the number density of true detections over the number density of spurious detections. They find that, on the basis of this quantity, the choice of the optimal filter depends on the statistical properties of the background. For the case of backgrounds that can be described with a power spectrum of the form $P(q) \propto q^{-\gamma}$, the SAF outperforms the other two filters for the case $1 < \gamma \leq 1.6$, whereas in the range $0 \leq \gamma \leq 1$ the MF is the most reliable.

The MHW is the most reliable filter in this sense when $\gamma > 1.6$. However, the criterion chosen to fix L_* based on the significance does not necessarily leads to the optimal reliability. Therefore, if we are considering the reliability as the main criterion to compare filters, a different criterion for L_*, closely related to the number densities, must be used.

In a posterior work, Vio, Andreani, & Wamsteker [148], following the previous work, adopt the same Neyman-Pearson decision rule, based on the pdf's of maxima of the background and background plus source, to define the region of acceptance. However, they propose to find L_* by fixing the number density of spurious detections and compare the number density of true detections for all the filters. In this case, the MF outperforms the other two filters.

In the present work, our goal will also be to find an optimal filter that gives a maximum number density of detections fixing a certain number density of spurious sources. In order to define the detector, we will use a decision rule based on the Neyman-Pearson test. We will consider some standard filters (MF, SAF and MH) introduced in the literature as well as a new filter that we call the Biparametric Scale Adaptive Filter (BSAF). In all the filters appears in a natural way the scale of the source. We will modify such a scale introducing an extra parameter. In fact, it has been shown by López-Caniego et al. [86] that the standard Matched Filter can be improved under certain conditions by filtering at a different scale from that of the source. The performance of the BSAF will be compared with the other filters.

The overview of this chapter is as follows: In section 2, we introduce two useful quantities: number of maxima in a Gaussian background in the absence and presence of a local source. In section 3, we introduce the detection problem and define the region of acceptance. In section 4, we introduce an estimator of the amplitude of the source that is proven to be unbiased and maximum efficient. In section 5 and 6, we obtain different analytical and numerical results regarding weak point sources and scale-free background spectra and compare the performance of the new filter with other used in the literature. We include a subsection dedicated to study the robustness of the method. In section 7, we describe the simulations performed to test some theoretical aspects and give the main results and finally, in section 8, we summarize the conclusions and applications of this work. Appendix A is a sketch to obtain a sufficient linear detector whereas we obtain the linear unbiased and maximum efficient estimator in appendix B.

3.2 Background peaks and compact sources

3.2.1 The background

Let us assume a 1D background (e. g. one-dimensional scan on the celestial sphere or time ordered data set) represented by a Gaussian random field $\xi(x)$ with average value $\langle\xi(x)\rangle = 0$ and power spectrum $P(q)$, $q \equiv |Q|$: $\langle\xi(Q)\xi^*(Q')\rangle = P(q)\delta_D(Q-Q')$, where $\xi(Q)$ is the Fourier transform of $\xi(x)$ and δ_D is the 1D Dirac distribution. The distribution of maxima was studied by Rice (1954) in a pioneering article. The expected number density of maxima per intervals $(x, x+dx)$, $(\nu, \nu+d\nu)$ and $(\kappa, \kappa+d\kappa)$ is given by

$$n_b(\nu,\kappa) = \frac{n_b\,\kappa}{\sqrt{2\pi(1-\rho^2)}} e^{-\frac{\nu^2+\kappa^2-2\rho\nu\kappa}{2(1-\rho^2)}}, \tag{3.2.1}$$

being n_b the expected total number density of maxima (i.e. number of maxima per unit interval dx)

$$n_b \equiv \frac{1}{2\pi\theta_m}, \quad \nu \equiv \frac{\xi}{\sigma_0}, \quad \kappa \equiv \frac{-\xi''}{\sigma_2}, \tag{3.2.2}$$

$$\theta_m \equiv \frac{\sigma_1}{\sigma_2}, \quad \rho \equiv \frac{\sigma_1^2}{\sigma_0\sigma_2} = \frac{\theta_m}{\theta_c}, \quad \theta_c \equiv \frac{\sigma_0}{\sigma_1},$$

where $\nu \in (-\infty, \infty)$ and $\kappa \in [0, \infty)$ represent the normalized field and curvature, respectively. σ_n^2 is the moment of order $2n$ associated to the field. θ_c, θ_m are the coherence scale of the field and maxima, respectively. As an example, figure 3.1 shows the values of the ratio $n_b(\nu,\kappa)/n_b$ for the case $\rho = 0.7$ (a typical value for the backgrounds we are considering). In this case, the expected density of maxima has a peak around $\nu \simeq 0.8$ and $\kappa \simeq 1.1$, that is, most of the peaks appear at a relatively low threshold and curvature, and the density of peaks decreases quickly for extreme values of ν and κ.

If the original field is linear-filtered with a circularly-symmetric filter $\Psi(x; R, b)$, dependent on 2 parameters (R defines a scaling whereas b defines a translation)

$$\Psi(x; R, b) = \frac{1}{R}\psi\left(\frac{|x-b|}{R}\right), \tag{3.2.3}$$

we define the filtered field as

$$w(R, b) = \int dx\, \xi(x)\Psi(x; R, b). \tag{3.2.4}$$

Then, the moment of order $2n$ of the linearly-filtered field is

$$\sigma_n^2 \equiv 2\int_0^\infty dq\, q^{2n} P(q)\psi^2(Rq), \tag{3.2.5}$$

being $P(q)$ the power spectrum of the unfiltered field and $\psi(Rq)$ the Fourier transform of the circularly-symmetric linear filter.

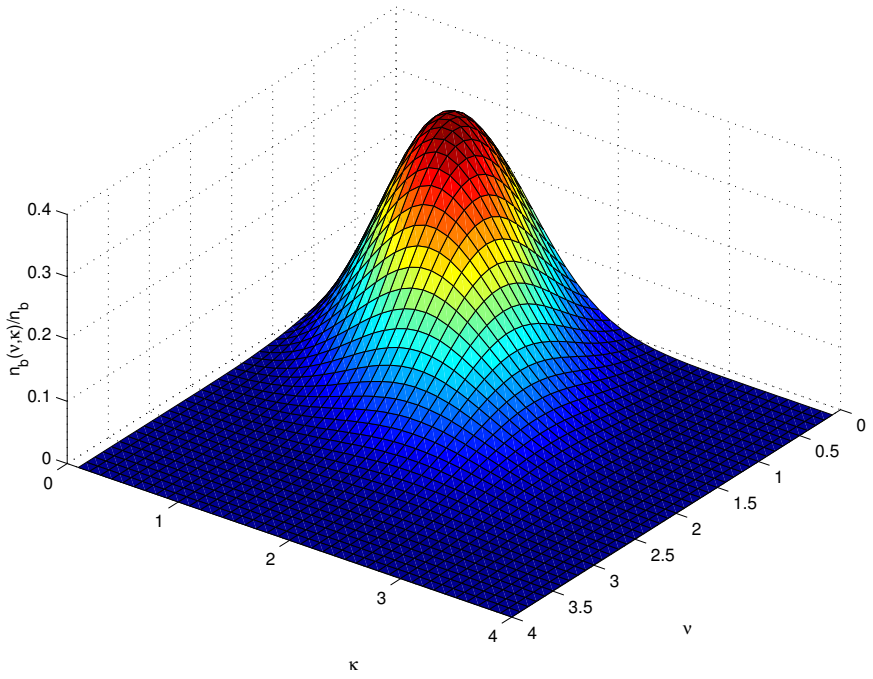

Figure 3.1 Value of $n_b(\nu,\kappa)/n_b$ for $\rho = 0.7$.

3.2.2 The presence of a local source

Now, let us consider a Gaussian source (i.e. profile given by $\tau(x) = \exp(-x^2/2R^2)$, where R is the beam width) embedded in the previous background. Then, the expected number density of maxima per intervals $(x, x+dx)$, $(\nu, \nu+d\nu)$ and $(\kappa, \kappa+d\kappa)$, given a source of amplitude A in such spatial interval, is given by Barreiro et al. [5]

$$n(\nu,\kappa|\nu_s) = \frac{n_b\,\kappa}{\sqrt{2\pi(1-\rho^2)}}\times$$

$$e^{-\frac{(\nu-\nu_s)^2+(\kappa-\kappa_s)^2-2\rho(\nu-\nu_s)(\kappa-\kappa_s)}{2(1-\rho^2)}}, \tag{3.2.6}$$

where $\nu \in (-\infty,\infty)$ and $\kappa \in [0,\infty)$, $\nu_s = A/\sigma_0$ is the normalized amplitude of the source and $\kappa_s = -A\tau''_\psi/\sigma_2$ is the normalized curvature of the filtered source. The last expression can be obtained as

$$\kappa_s = \nu_s y_s,\ y_s \equiv -\frac{\theta_m^2\tau''_\psi}{\rho},\ -\tau''_\psi = 2\int_0^\infty dq\, q^2\tau(q)\psi(Rq). \tag{3.2.7}$$

We consider that the filter is normalized such that the amplitude of the source is the same after linear filtering: $\int dx\,\tau(x)\Psi(x;R,b) = 1$.

3.3 The detection problem

We want to make a decision between filters based on *detection*. To make such a decision, we will focus on the following two fundamental quantities: a) the number of spurious sources which emerge after the filtering and detection processes and b) the number of real sources detected. As seen in the previous section, these quantities are properties of the Gaussian field and source that can be calculated through equations (3.2.1) and (3.2.6). As we will see, the previous properties are not only related to the SNR gained in the filtering process but depend on the filtered momenta to 4th-order (in the 1D case), i.e. the amplification and the normalized curvature.

3.3.1 The region of acceptance

Let us consider a local peak in the 1D data set characterised by the normalized amplitude and curvature (ν_s, κ_s). Let $H_0 : n.d.f.\ n_b(\nu, \kappa)$ represent the *null* hypothesis, i.e. the local number density of background maxima, and $H_1 : n.d.f.\ n(\nu, \kappa)$ represents the *alternative* hypothesis, i.e. the local number density of maxima when there is a compact source.

$$n(\nu, \kappa) = \int_0^\infty d\nu_s p(\nu_s) n(\nu, \kappa | \nu_s). \tag{3.3.1}$$

We have assumed a Bayesian approach: at a concrete pixel we get the number of source detections weighting with the *a priori* probability $p(\nu_s)$, if the sources are spatially distributed as a Poissonian process.

Given the data (ν, κ), we can associate to any region $\mathcal{R}_*(\nu, \kappa)$ two number densities n_b^* and n^*

$$n_b^* = \int_{\mathcal{R}_*} d\nu\, d\kappa\, n_b(\nu, \kappa), \tag{3.3.2}$$

$$n^* = \int_{\mathcal{R}_*} d\nu\, d\kappa\, n(\nu, \kappa). \tag{3.3.3}$$

Then, n_b^* is the number density of spurious sources, i.e. due to the background, expected inside the region $\mathcal{R}_*(\nu, \kappa)$, whereas n^* is the number density of maxima expected in the same region of the (ν, κ) space in the presence of a local source. Hereinafter, we will call it the *number density of detections*.

$\mathcal{R}_*$ is called the *acceptance* region. We remark that in order to get the true number of real source detections such a number must be multiplied by the probability to have a source in a pixel in the original data set.

We will assume a Bayesian Neyman-Pearson (BNP) decision rule using number densities instead of probabilities: the acceptance region $\mathcal{R}_*$ is given by the highest number density of detections n^*, for a given number density of spurious detections n_b^*. Such a region is (criterion for detection)

$$L(\nu,\kappa) \equiv \frac{n(\nu,\kappa)}{n_b(\nu,\kappa)} \geq L_*, \tag{3.3.4}$$

where L_* is a constant and $L(\nu,\kappa)$ defines the *d*etector. Therefore, the decision rule is expressed by the likelihood ratio: if $L \geq L_*$ we decide that the signal is present, whereas if $L < L_*$ we decide that the signal is absent. Note that the previous region is equivalent to the one defined by the usual Neyman-Pearson test in terms of probabilities

$$\frac{p(\nu,\kappa)}{p_b(\nu,\kappa)} \geq L'_* \tag{3.3.5}$$

where $p_b(\nu,\kappa)$, $p(\nu,\kappa)$ are the pdf's associated to the number densities given by equations (3.2.1) and (3.3.1) and, in order to compare different filters, the constant L'_* must be found by fixing the number density of spurious sources in the region of acceptance n_b^* instead of the "false alarm" probability $\alpha \equiv n_b^*/n_b$.

It can be proved that the previous region of acceptance $\mathcal{R}_*$ is equivalent to the sufficient linear detector (see Appendix A)

$$\mathcal{R}_* : \varphi(\nu,\kappa) \geq \varphi_*, \tag{3.3.6}$$

where φ_* is a constant and φ is given by

$$\varphi(\nu,\kappa) \equiv \frac{1-\rho y_s}{1-\rho^2}\nu + \frac{y_s-\rho}{1-\rho^2}\kappa \tag{3.3.7}$$

We remark that the assumed criterion for detection leads to a *linear* detector φ (i.e. linear dependence on the threshold ν and curvature κ). Moreover, this detector is independent of the pdf of the source amplitudes.

3.3.2 Spurious sources and real detections

Once obtained the region of acceptance $\mathcal{R}_*$ in the previous subsection, one can calculate the number density of spurious sources and the number density of detections as given by equation (3.3.2)

$$\begin{aligned} n_b^* &= \frac{n_b}{2}\left[\operatorname{erfc}\left(\frac{\varphi_*\sqrt{1-\rho^2}}{\sqrt{2}(1-\rho y_s)}\right)\right. \\ &\left. +\sqrt{2}My_s\mathrm{e}^{-M^2\varphi_*^2}\operatorname{erfc}\left(-\frac{\sqrt{1-\rho^2}}{1-\rho y_s}y_sM\varphi_*\right)\right], \end{aligned} \tag{3.3.8}$$

$$M \equiv \sqrt{\frac{1-\rho^2}{2(1-2\rho y_s + y_s^2)}},$$

$$\begin{aligned} n^* &= \frac{n_b}{\sqrt{2\pi}} \frac{1-\rho y_s}{(\mu + y_s^2)\sqrt{1-\rho^2}} \times \\ &\int_{\varphi_*}^{\infty} d\varphi I(\varphi)[1+B(z)] e^{-\frac{(1-\rho^2)\varphi^2}{2(1-\rho y_s)^2}}, \end{aligned}$$

where

$$z = \frac{y_s \varphi}{1-\rho y_s} \sqrt{\frac{1-\rho^2}{2(\mu + y_s^2)}},$$

$$B(z) = \sqrt{\pi} z e^{z^2} \operatorname{erfc}(-z),$$

$$\mu \equiv \frac{(1-\rho y_s)^2}{1-\rho^2},$$

$$I(\varphi) = \int_0^{\infty} d\nu_s \, p(\nu_s) e^{\nu_s \varphi - \frac{1}{2}\nu_s^2(\mu + y_s^2)}. \tag{3.3.9}$$

Then, one can invert the equation for the number of spurious to get $\varphi_* = \varphi_*(\frac{n_b^*}{n_b}; \rho, y_s)$ that allows to rewrite the equation for the number of detections as $n^* = g(n_b^*; \theta_m, \rho, y_s)$, that is a generalized "receiving operating curve"(R.O.C. in the signal processing jargon).

3.4 The estimation of the amplitude of the source

The signal has an unknown parameter, the amplitude A, that has to be estimated from the data (ν, κ). We shall assume that the most probable value of the distribution $n(\nu, \kappa | \nu_s)$ gives an estimation of the amplitude of the source (criterion for amplitude estimation). The result $\hat{\nu}_s$ is given by the equation

$$\hat{\nu}_s = \frac{\varphi(\nu, \kappa)}{y_s^2 + \mu}, \tag{3.4.1}$$

where the function φ is given by equation (3.3.7). One can prove that such an estimation gives a linear estimator that is unbiased and maximum efficient, i.e.

$$\langle \hat{\nu}_s \rangle = \nu, \quad \sigma_{\hat{\nu}_s}^2 = \frac{1}{y_s^2 + \mu}, \tag{3.4.2}$$

where $\langle\rangle$ denotes average value over realizations (see Appendix B). A confidence level for the estimation of the amplitude of the source can also be obtained from the *a posteriori* p.d.f. $p(\nu_s|\nu,\kappa) \propto n(\nu,\kappa|\nu_s)$.

3.5 Analytical results

3.5.1 Filters

We will consider as an application the detection of compact sources characterised by a Gaussian profile $\tau(x) = \exp(-x^2/2R^2)$, and Fourier transform $\tau = R\exp(-(qR)^2/2)$ though the extension to other profiles will be considered in the future. Such a profile is physically and astronomically interesting because represents the convolution of a point source (Dirac δ distribution) with a Gaussian beam.

The source profile above includes a "natural scale" R that characterises the source. This is a fundamental scale that will appear in all the filters we will consider here. Practically by definition, the standard MF and SAF operate on this scale, as well as the canonical MHW at the scale of the source. However, it has been shown that changing the scale at which the MHW and the MF filters the image can improve its performance in terms of detection (López-Caniego et al. [86], Vielva et al. [144, 145]). Following this idea, we will in the following introduce another degree of freedom in all the filters that allows us to change their scale in a continuous way (much similar to the scaling of a continuous wavelet). This degree of freedom is obtained by multiplying the scale R by a new parameter $\alpha > 0$. We will show that with this new parameter the improvement in the results is significant.

The Scale-adaptive Filter (SAF)

The idea of a scale-adaptive filter (or optimal pseudo-filter) has been recently introduced by the authors [124]. By introducing a circularly-symmetric filter, $\Psi(x;R,b)$, we are going to express the conditions in order to obtain a scale-adaptive filter for the detection of the source $s(x)$ at the origin taking into account the fact that the source is characterised by a single scale R_o. The following conditions are assumed: (1) $\langle w(R_o,0)\rangle = s(0) \equiv A$, i.e. $w(R_o,0)$ is an *unbiased* estimator of the amplitude of the source; (2) the variance of $w(R,b)$ has a minimum at the scale R_o, i.e. it is an *efficient* estimator; (3) $w(R,b)$ has a maximum at $(R_o,0)$. Then, the filter satisfying these conditions is given by [124].

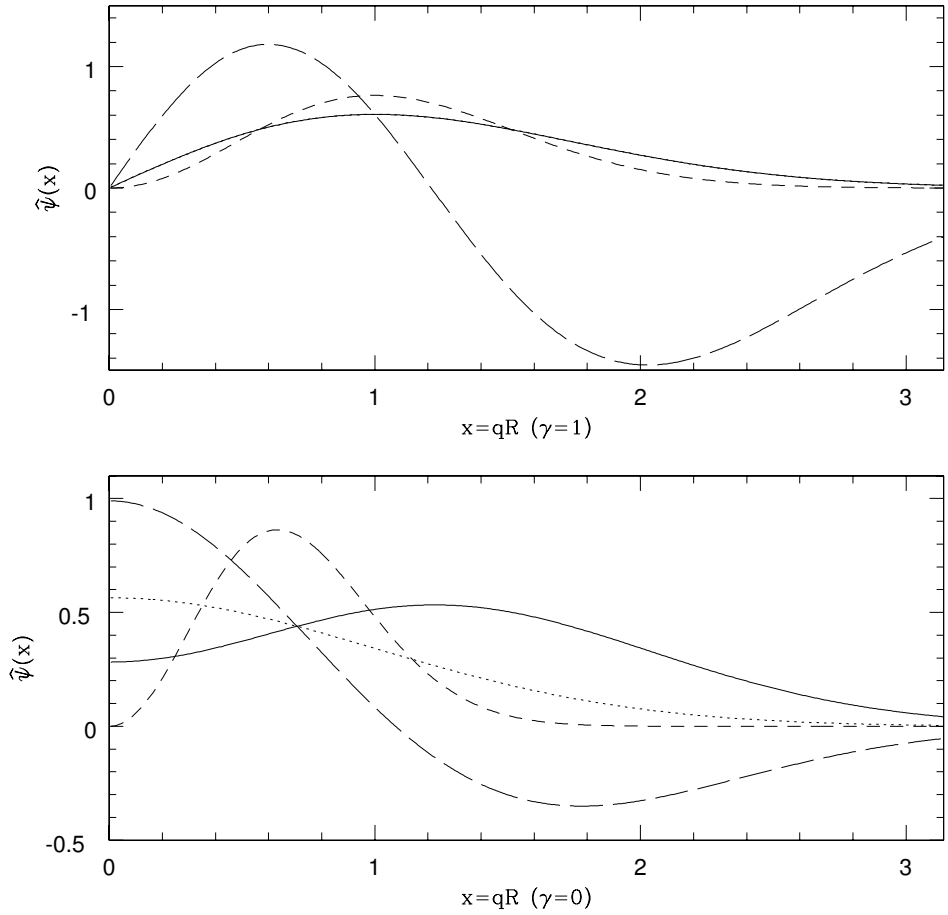

Figure 3.2 Different filters for the values of $\gamma = 0$ (lower panel) and $\gamma = 1$ (upper panel). The filters represented in all cases are: the SAF (solid line), MF (dotted line), MH (short dashed line) are shown for α = 1 and the BSAF (long dashed line) are given for $(\alpha, c) = (0.3, -0.86)$ for $\gamma = 0$ and $(\alpha, c) = (0.4, -0.68)$ for $\gamma = 1$. Note that for $\gamma = 1$ the SAF and the MF coincide

$$\psi_{SAF} = \frac{1}{ac - b^2}\frac{\tau(q)}{P(q)}\left[b + c - (a+b)\frac{dln\tau}{dlnq}\right],$$

$$a \equiv \int dq\,\frac{\tau^2}{P},\; b \equiv \int dq\,q\frac{\tau}{P}\frac{d\tau}{dq},\; c \equiv \int dq\,q^2\frac{1}{P}\left(\frac{d\tau}{dq}\right)^2, \tag{3.5.1}$$

Assuming a scale-free power spectrum, $P(q) \propto q^{-\gamma}$, and a Gaussian profile for the source, the previous set of equations lead to the filter

$$\tilde{\psi}_{SAF} = N(\alpha)x^{\gamma}e^{-\frac{1}{2}x^2}\left[1 + \frac{t}{m^2}x^2\right],\;\; x \equiv q\alpha R,$$

$$m \equiv \frac{1+\gamma}{2},\;\; t \equiv \frac{1-\gamma}{2},\;\; \Delta \equiv \frac{2\alpha^2}{1+\alpha^2},$$

$$N(\alpha) = \frac{\alpha}{\Delta^m\Gamma(m)}\frac{1}{1+\frac{t}{m}\Delta}, \tag{3.5.2}$$

where we have modified the scale as αR.

In this case the filter parameters θ_m, ρ and the curvature of the source y_s are given by

$$\frac{\theta_m}{\alpha R} = \sqrt{\frac{1 + \frac{t^2}{m} + \frac{2t}{m^2}}{(1+m)\left(1 + \frac{t^2}{m} + \frac{2t(2+t)}{m^2}\right)}}$$

$$\rho = \sqrt{\frac{m}{1+m}}\frac{1 + \frac{t^2}{m} + \frac{2t}{m^2}}{\sqrt{\left(1 + \frac{t^2}{m}\right)\left(1 + \frac{t^2}{m} + \frac{2t(2+t)}{m^2}\right)}},$$

$$y_s = H\sqrt{\frac{1 + \frac{t^2}{m}}{m\,(1+m)\left(1 + \frac{t^2}{m} + \frac{2t(2+t)}{m^2}\right)}},$$

$$H \equiv \Delta\frac{1+\frac{t}{m}}{1+\frac{t}{m}\Delta}\frac{m^2 + t(1+m)\Delta}{m^2 + t(1+m)} \tag{3.5.3}$$

Figure 3.2 shows the SAF for different values of the spectral index.

The Matched Filter (MF)

If one removes condition (3) defining the SAF in the previous subsection, it is not difficult to find another type of filter after minimization of the variance (condition (2)) with the constraint (1)

$$\psi_{MF} = \frac{1}{2a}\frac{\tau(q)}{P(q)}. \tag{3.5.4}$$

This will be called *matched* filter as is usual in the literature. Note that in general the matched and adaptive filters are different.

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

$$\tilde{\psi}_{MF} = N(\alpha)x^{\gamma}e^{-\frac{1}{2}x^2}, \quad x \equiv q\alpha R, \tag{3.5.5}$$

$$N(\alpha) = \frac{\alpha}{\Delta^m \Gamma(m)}$$

where m is given by equation (3.5.2).

Figure 3.2 shows the MF for the case $\alpha = 1$ (standard MF) and different values of the spectral index $\gamma = 0, 1$. We remark that for $\gamma = 1$ the scale-adaptive filter and the matched filter coincide, and for $\gamma = 2$ (not shown in the figure), the matched filter and the Mexican Hat wavelet are equal.

For the MF the parameters θ_c, θ_m, ρ and the curvature of the source y_s are given by

$$\frac{\theta_m}{\alpha R} = \frac{1}{\sqrt{1+m}}, \quad \rho = \sqrt{\frac{m}{1+m}}, \quad y_s = \rho\Delta \tag{3.5.6}$$

We remark that the linear detector $\varphi(\nu, \kappa)$ is reduced to $\varphi = \nu$ for the standard Matched Filter ($\alpha = 1$). i.e. curvature does not affect the region of acceptance for such a filter.

The Mexican Hat Wavelet (MH)

The MH is defined to be proportional to the Laplacian of the Gaussian function: $\psi_{MH}(x) \propto (1 - x^2)e^{-x^2/2}$. Thus, in Fourier space

$$\psi_{MH}(x) = \frac{2}{\sqrt{\pi}}x^2 e^{-\frac{1}{2}x^2}, x \equiv qR. \tag{3.5.7}$$

In this case the filter parameters θ_m, θ_c, ρ and the curvature of the source y_s are given by

$$\frac{\theta_m}{R} = \frac{1}{\sqrt{3+t}}, \ \rho = \sqrt{\frac{2+t}{3+t}}, \ y_s = \frac{3/2}{\sqrt{(2+t)(3+t)}}. \tag{3.5.8}$$

We comment that the generalization of this type of wavelet for two dimensions has been extensively used for point source detection in 2D images [19, 144]. As in the previous cases, the MH is modified in the form

$$\tilde{\psi}_{MH} = N(\alpha)x^2 e^{-\frac{1}{2}x^2}, \quad x \equiv q\alpha R, \tag{3.5.9}$$

$$N(\alpha) = \frac{2\alpha}{\sqrt{\pi}\Delta^{3/2}}$$

For the MH the parameters θ_c, θ_m, ρ and the curvature of the source y_s are given by

$$\frac{\theta_m}{\alpha R} = \frac{1}{\sqrt{3+t}}, \ \rho = \sqrt{\frac{2+t}{3+t}}, \ y_s = \frac{3\Delta/2}{\sqrt{(2+t)(3+t)}}. \tag{3.5.10}$$

Figure 3.2 shows the MH for different values of the spectral index.

The Biparametric Scale Adaptive Filter (BSAF)

If one removes condition (3) defining the SAF in subsection 5.1.1 and introduces the condition (3), $w(R_o, b)$ has a maximum in the filtered image at $b = 0$, i.e. $\langle w''(R_o, 0)\rangle < 0$, it is not difficult to find another type of filter

$$\psi \propto \frac{\tau(q)}{P(q)}(1 + c(qR)^2), \tag{3.5.11}$$

where c is an arbitrary constant that can be related to the curvature of the maximum. We remark that the constrain $\langle w\prime(R_o, 0)\rangle = 0$ is automatically satisfied for any circularly-symmetric filter if the source profile has a maximum at the origin.

For the case of a scale-free spectrum, the filter is given by the parametrized equation

$$\tilde{\psi}_{BSAF} = \frac{\alpha}{2J_\gamma} x^\gamma e^{-\frac{1}{2}x^2}(1 + cx^2), \ \ x \equiv q\alpha R. \tag{3.5.12}$$

where we have modified the scale as αR. Hereinafter, we will call this new filter containing two arbitrary parameters, $\alpha > 0$ and c, the biparametric scale-adaptive filter (BSAF).

A calculation of the different moments leads to

$$\frac{\theta_m}{\alpha R} = \sqrt{\frac{G_{\gamma+2}}{G_{\gamma+4}}}, \ \rho = \frac{G_{\gamma+2}}{\sqrt{G_\gamma G_{\gamma+4}}}, \ y_s = \frac{J_{\gamma+2}}{J_\gamma}\sqrt{\frac{G_\gamma}{G_{\gamma+4}}}, \tag{3.5.13}$$

where m and t are defined in equation (3.5.2) and G_γ and J_γ are given by

$$G_\gamma \equiv \frac{1}{2}[1 + 2mc + m(m+1)c^2]\Gamma(m), \tag{3.5.14}$$

$$J_\gamma(\alpha) \equiv \frac{1}{2}[1 + mc\Delta]\Delta^m \Gamma(m). \tag{3.5.15}$$

Note that the BSAF contains all the other considered filters as particular cases: the MF is recovered for $c = 0$, when $c = t/m^2$ the BSAF defaults to the scale-adaptive filter and, finally, the MH wavelet is obtained in the two cases: $\gamma = 0$, $c \gg 1$ and $\gamma = 2$, $c = 0$.

3.5.2 A priori probability distributions $p(\nu_s)$

We will test two different p.d.f. $p(\nu_s)$: a uniform distribution in the interval $0 \leq \nu \leq \nu_c$ and a scale-free distribution with a lower and upper cut-off $\nu_i \leq \nu \leq \nu_f$. In particular, we will especially focus on values for the cut-off's that lead to distributions dominated by weak sources. It is in this regime where sophisticated detection methods are needed, since bright sources can be easily detected with simple techniques.

Uniform distribution

In this case,

$$p(\nu_s) = \frac{1}{\nu_c}, \quad \nu_s \in [0, \nu_c]. \tag{3.5.16}$$

This allows to obtain

$$I(\varphi) = \sqrt{\frac{\pi}{2}} \frac{e^{h^2}}{\nu_c \sqrt{y_s^2 + \mu}} \left[\text{erf}(h) + \text{erf}\left(\frac{\nu_c}{\sqrt{2}} \sqrt{y_s^2 + \mu} - h \right) \right],$$

$$h \equiv \frac{\varphi}{\sqrt{2(y_s^2 + \mu)}}. \tag{3.5.17}$$

We will consider a cut-off in the amplitude of the sources such that $\nu_c = 2$ after filtering with he standard MF. Note that this correspond to different thresholds for the rest of the filters.

Scale-free distribution with lower and upper cut-off

In this case,

$$p(\nu_s) = N \nu_s^{-\beta}, \quad \nu \in [\nu_i, \nu_f], \quad \beta \neq 1, \tag{3.5.18}$$

where the normalization constant N and $I(\varphi)$ are

$$N = \frac{\beta - 1}{\nu_i^{1-\beta}} \frac{1}{1 - \left(\frac{\nu_i}{\nu_f}\right)^{\beta - 1}} \tag{3.5.19}$$

$$I(\varphi) = N \int_{\nu_i}^{\nu_f} d\nu\, \nu^{-\beta} e^{\nu[\varphi - \frac{\nu}{2}(\mu + y_s)^2]}. \tag{3.5.20}$$

We will consider $\nu_i \simeq 0.5$ and $\nu_f \simeq 3$ after filtering with the standard MF and the corresponding thresholds for the other filters.

3.6 Numerical results

We explore the parameter space α, R and γ fixing the number density of spurious sources to levels of $n_b^* = 0.01 - 0.05$. We want to find the optimal filter in the sense of maximum number of detections. We use two different distributions of sources to show the robustness of the method, a Uniform distribution and a Scale-free distribution. In both cases, we obtain similar results: we find an optimal filter that significantly improves the standard MF, SAF and MH for different values of the spectral index in the interval $0 \leq \gamma \leq 1$. We have also studied the performance of the filters for $1 < \gamma \leq 2$ and have found no significant improvement of the BSAF with respect to the standard MF. Therefore, we do not include a detailed study of these cases in the results.

The BSAF has an additional degree of freedom, the parameter c, as it appears in equation (3.5.12). Note that the BSAF and the SAF are not the same filter. The parameter c in the BSAF can take any positive or negative value, while the coefficient t/m^2, for the SAF, is a known function of γ. We remark that, by construction, the BSAF always outperforms the other considered filters or, in the worst case, defaults to the best of them.

3.6.1 Uniform distribution

We use a uniform distribution of sources with amplitudes in the interval $A \in [0, 2]\sigma_0$, where σ_0 is the zero-order moment of the linearly-filtered map with the standard MF, i.e. $\nu \in [0, 2]$. It is clear that the corresponding upper limit for ν in the original signal is below 2, therefore, we are considering the detection of weak sources.

In figures 3.3 and 3.4 we plot n^*, the number density of detections, as a function of α for the cases $\gamma = 0$ and $\gamma = 0.5$ for the case $R = 3$, and $n_b^* = 0.05$, where R is given in pixel units.

For the case $\gamma = 0$, (shown in figure 3.3), n^* rapidly increases as α decreases for the BSAF as compared with the other filters. For the MF and BSAF we find a maximum at $\alpha \simeq 0.2$. This scale is a compromise between minimising the effect of the background

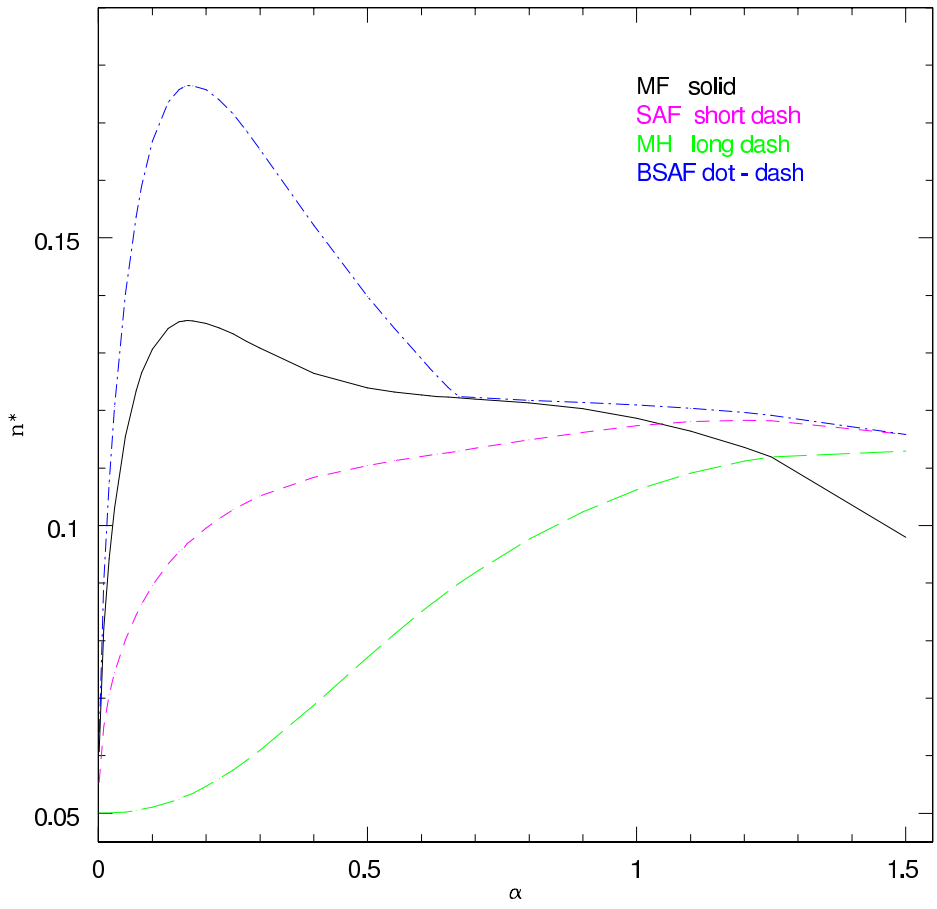

Figure 3.3 Uniform distribution. The expected number density of detections n^* as a function of the filter parameter α for $\gamma = 0$ for the BSAF (where c has been obtained by maximising the number of detections for each value of α), MF, SAF and MH filters. We consider the case $R = 3$, $n_b^* = 0.05$ and $\nu_c = 2$.

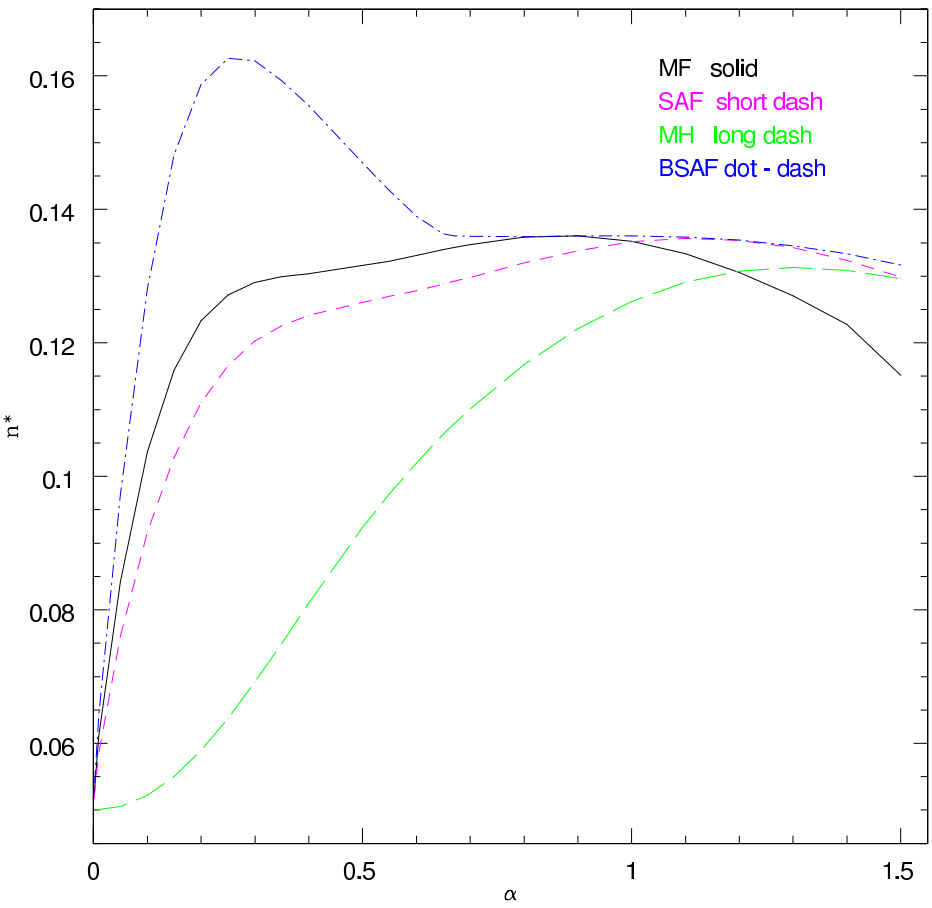

Figure 3.4 Uniform distribution. The expected number density of detections n^* as a function of the filter parameter α for $\gamma = 0.5$ for the BSAF, MF, SAF and MH filters. We consider the case $R = 3$, $n_b^* = 0.05$ and $\nu_c = 2$. As in the previous figure the parameter c of the BSAF has been determined by maximising the number of detections for each value of α.

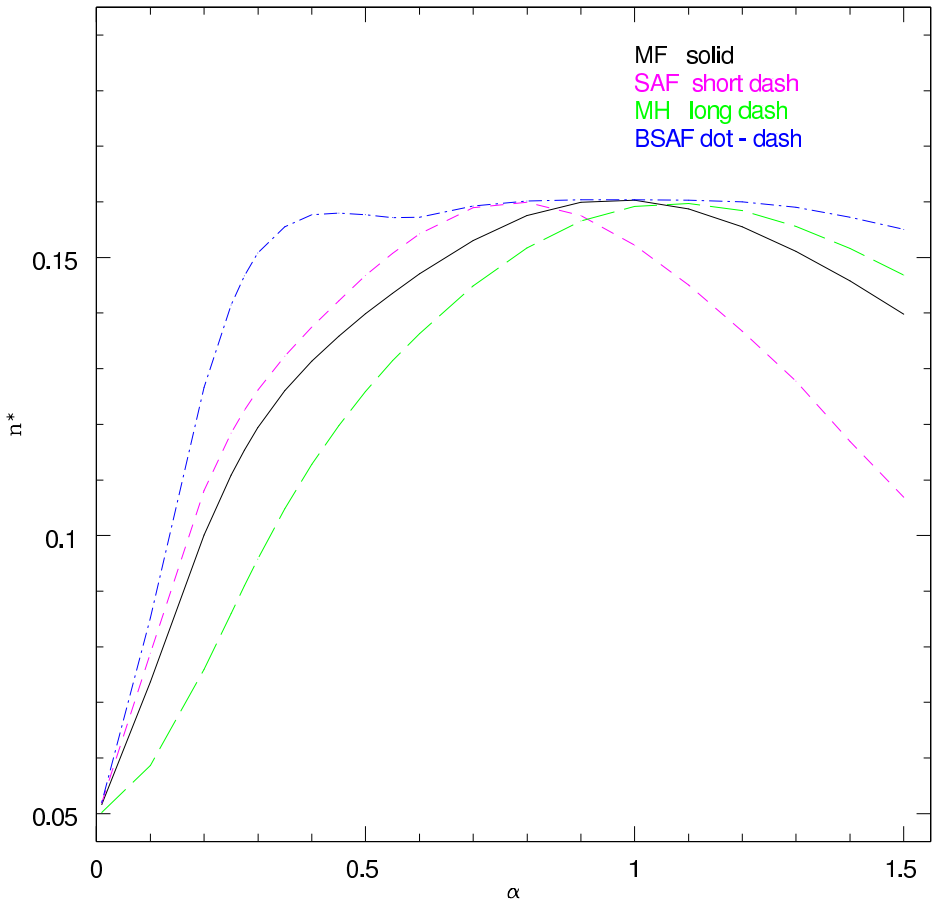

Figure 3.5 Uniform distribution. The expected number density of detections n^* as a function of the filter parameter α for $\gamma = 1.5$ for the BSAF, MF, SAF and MH filters. We consider the case $R = 3$, $n_b^* = 0.05$ and $\nu_c = 2$. As in the previous figure the parameter c of the BSAF has been determined by maximising the number of detections for each value of α.

R	n_b^*	γ	α	c	n_{BSAF}^*	n_{MF}^*	$RD[\%]$
2	0.01	0	0.4	-0.69	0.0860	0.0824	4.4
2	0.03	0	0.4	-0.68	0.1493	0.1311	13.9
		0.5	0.4	-0.59	0.1512	0.1474	2.5
2	0.05	0	0.4	-0.70	0.1900	0.1575	22
		0.5	0.4	-0.59	0.1935	0.1783	9
3	0.01	0	0.3	-0.86	0.0784	0.0658	19.1
3	0.03	0	0.3	-0.86	0.1282	0.1013	26.5
		0.5	0.3	-0.73	0.1242	0.1145	8.4
3	0.05	0	0.3	-0.86	0.1654	0.1186	39.4
		0.5	0.3	-0.75	0.1616	0.1352	19.5
		1	0.4	-0.58	0.1582	0.1487	6.3

Table 3.1 Uniform distribution. Number density of detections n^* for the standard MF($\alpha = 1$) and the BSAF with optimal values of c and α. RD means relative difference in number densities in percentage. $RD \equiv 100(-1 + n_{BSAF}^*/n_{MF}^*)$.

and amplifying the source. One can proven theoretically, that n^* converges to the value n_b^* when $\alpha \to 0$, as can be seen in the figures.

For completeness, we show the results for all the values of α. However, in practice, it does not make sense to filter at scales significantly smaller than the size of the pixel, which corresponds to $\alpha R \simeq 1$. Therefore, in the following, the results will be given taking into account only those values of $\alpha \gtrsim R^{-1}$. The improvement in n^* for $\alpha = 0.3$ is of the order 40% compared with the standard Matched Filter ($\alpha = 1$). If we compare with the MF at $\alpha = 0.3$, the improvement is of $\simeq 20\%$.

In figure 3.4, with $\gamma = 0.5$, the BSAF at $\alpha = 0.3$ again improves significantly the standard MF. In particular, the difference between both filters is of the order of 25%.

As γ increases, the improvement of the BSAF with respect to the standard MF decreases. In fact, for values of $1 < \gamma \leq 2$ they produce very similar results. As an example, we give the number of detections achieved for each filter for the case $\gamma = 1.5$, $n_b^* = 0.05$ and $R = 3$ in Fig. 3.5. It can be seen that the maximum number of detections is found approximately for the standard MF. However, we would like to point out that the SAF and MH wavelet at the optimal scale give the same number of detections as the standard MF. These results show the importance of filtering at scales αR instead of the usual scale of the source.

In table 3.1, we present some examples of the number density of detections for the

BSAF and for the standard MF ($\alpha = 1$). We calculate n^* for $R = 2$ and $R = 3$, with n_b^* in the interval 0.01 - 0.05 and values of $\gamma = 0, 0.5, 1$ (we give in the table only those cases where the BSAF improves at least a few per cent the standard MF). The values of α and c for the BSAF are found as the ones that maximise n^* in each case.

When looking at this table, it can be seen how the number density of detections increases when we increase n_b^*, as expected. Most important, the relative difference between the BSAF and the standard MF also increases when we increase R. In general, we find that if R is larger, the BSAF will behave better. But there is a limit to this, because the total number of spurious detections n_b, which is a function of R, is involved in the calculation of φ_*, and if n_b^* is too large, we can not solve for φ in the implicit equation (3.3.2).

In the previous case, the sources were very weak ($A \in [0, 2]\sigma_0$). For completeness, we have tested a uniform distribution with a mixture of weak and bright sources with amplitudes in the interval $A \in [0, 5]\sigma_0$, $R = 3$ pixels and $n_b^* = 0.05$. For $\gamma = 0$ we find that the optimal values of the parameters for the BSAF are $c = -0.79$ and $\alpha = 0.3$. The behaviour is similar to the one found in the weak sources case, although the improvement of the BSAF versus the MF is lower. In particular, the BSAF improves $\simeq$ 21% the standard MF ($\alpha = 1$), as compared to the 40% improvement for weak sources. It is interesting to note that for the same values of the parameters for the case with $A \in [0, 2]\sigma_0$, the optimal values of α and c are very similar in both cases, which is an indication of the robustness of the technique.

Figure 3.6 shows the relative difference in the number of detections, with respect to the standard MF, as a function of the spectral index γ for the different filters. At each point the optimal scale (and parameter c in the case of BSAF) has been used. We remark that for the interesting case of white noise a 40% of detections is gained with respect to the standard MF.

3.6.2 Scale-free distribution

For comparison purposes we use a Scale-free power-law distribution of sources with $\beta = 0.5$ and amplitudes in the interval $A \in [0.5, 3]\sigma_0$, where σ_0 is the zero-order moment obtained from the map filtering with the standard MF. It is clear that the corresponding upper limit for ν in the original signal is below 3, therefore, we are considering the detection of weak sources. We will show that the results are comparable to the ones obtained for the previous distribution.

In figures 3.7 and 3.8, we plot n^*, the number density of detections, as a function of

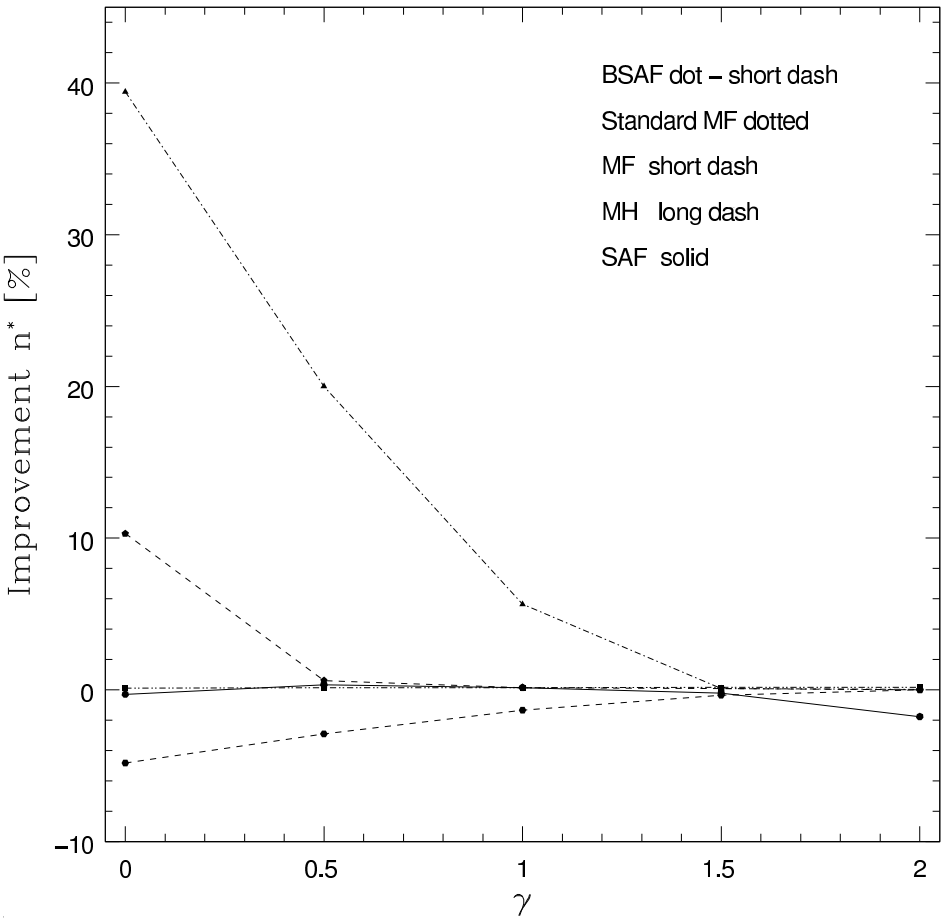

Figure 3.6 Uniform distribution. Relative difference in the number of detections, with respect to the standard MF, as a function of the spectral index γ.

α for the cases $\gamma = 0$ and $\gamma = 0.5$, where $R = 3$ and $n_b^* = 0.05$. In figure 3.7, n^* is significantly higher for the BSAF compared with the other filters at certain values of α. In this case, the improvement of the BSAF at $\alpha = 0.3$ compared with the standard MF is $\simeq 42\%$. If we compare with the MF at $\alpha = 0.3$, the improvement is $\simeq 33\%$. In figure 3.8, with $\gamma = 0.5$, an improvement of $\simeq 20\%$ is obtained for the BSAF at $\alpha = 0.3$ with respect to the standard MF.

As in the uniform distribution case, the BSAF gives very similar results to the MF in the range $1 < \gamma \leq 2$. In figure 3.9 we show the results for $\gamma = 1.5$, $n_b^* = 0.05$ and $R = 3$. Again we see that the optimal BSAF defaults to the standard MF.

In table 3.2, we show the number density of detections for the BSAF and for the standard MF ($\alpha = 1$) for $R = 2$ and $R = 3$, with n_b^* ranging from 0.01 to 0.05, and for values of $\gamma = 0, 0.5, 1$ (we only include the results for those cases where the relative difference between the BSAF and standard MF is at least a few per cent). We also give the optimal values of c and α where the BSAF performs better (taking into account the constraint αR of the order of 1 or greater).

Other values of $\beta \in [2.2, 2.5]$ have an intrinsic interest for astronomy, because they describe the distribution of compact sources in the sky at microwave wavelengths. We have explored these numbers (for $\gamma = 0$, $n_b^* = 0.05$, $R = 3$) and the improvement of the BSAF (with optimal values of $\alpha = 0.3$ and $c = -0.86$) versus the standard MF is still significant and of the order $\simeq 35\%$.

To test the robustness of the method, we have also considered a scale-free distribution of bright sources with $A \in [0.5, 5]\sigma_0$, for $\beta = 0.5$ and the same parameters as in the previous paragraph. In this case, the improvement is of the order of 25%, with optimal parameters for the BSAF of $\alpha = 0.3$ and $c = -0.76$. Moreover, we have also tested the performance of the filters for a scale-free distribution of bright sources with $A \in [3, 5]\sigma_0$, where σ_0 is the dispersion of the standard MF for the same case as before ($\gamma = 0$, $n_b^* = 0.05$, $R = 3$). We explored the parameter space of (c, α), looking for the best filter regarding detection. We find that, for this distribution, the optimal parameters for the BSAF are $c = 0$ and $\alpha = 1$, that correspond to the standard MF. Therefore, we conclude that for this case the BSAF defaults to the standard MF.

We would like to point out that for all the amplitude pdf's that we have considered (except for the one dominated by bright sources), the optimal parameters for the BSAF are similar for a given set of γ, R and n_b^*. This suggests that possible uncertainties in the knowledge of the pdf of the amplitude of the sources (provided we are not dominated by bright sources) will only marginally affect the estimation of the filter parameters and therefore the detections.

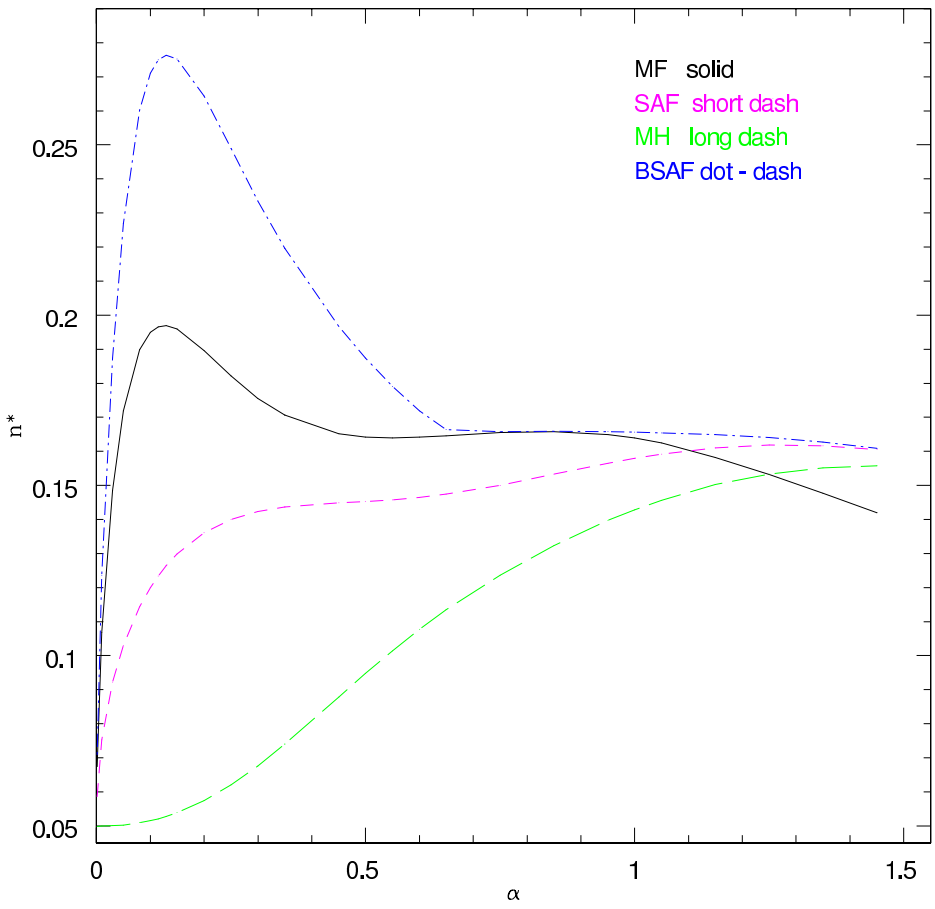

Figure 3.7 Scale-free distribution. The expected number density of detections n^* as a function of the filter parameter α for $\gamma = 0$ for the BSAF (using the optimal values of c), MF, SAF and MH filters. We consider the case $R = 3$, $n_b^* = 0.05$, $\nu_i = 0.5$, $\nu_f = 3$ and $\beta = 0.5$.

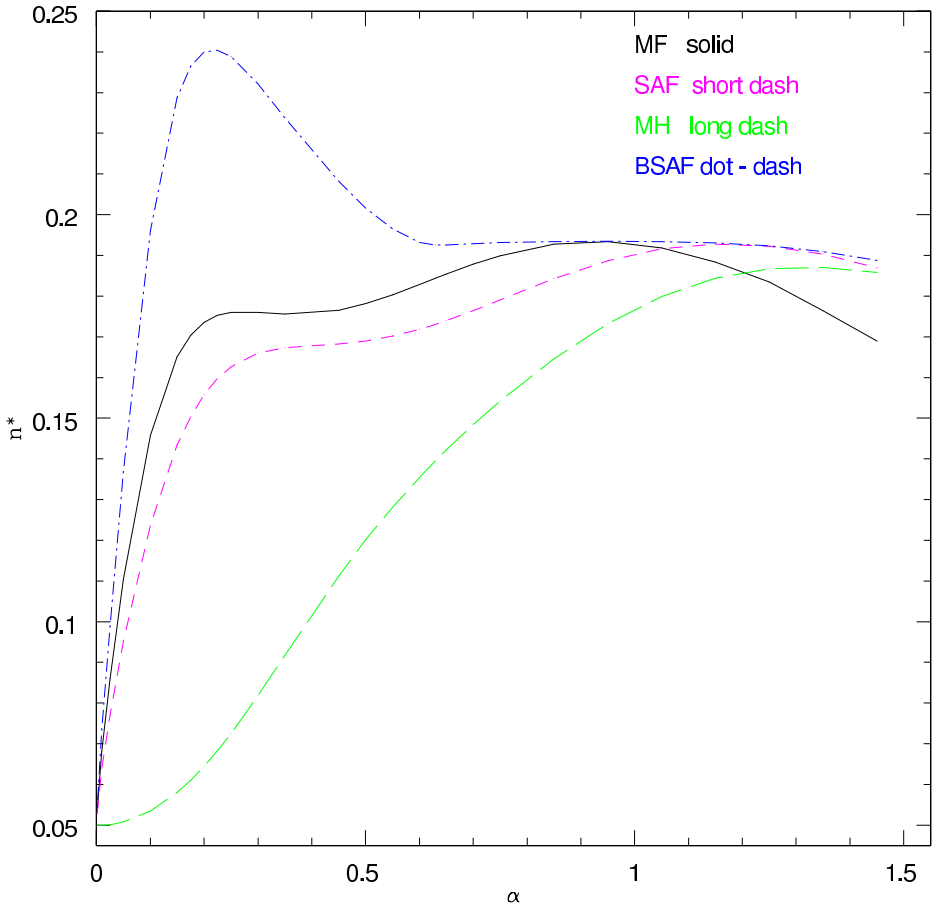

Figure 3.8 Scale-free distribution. The expected number density of detections n^* as a function of the filter parameter α for $\gamma = 0.5$ for the BSAF (using the optimal values of c), MF, SAF and MH filters. We consider the case $R = 3$, $n_b^* = 0.05$, $\nu_i = 0.5$, $\nu_f = 3$ and $\beta = 0.5$.

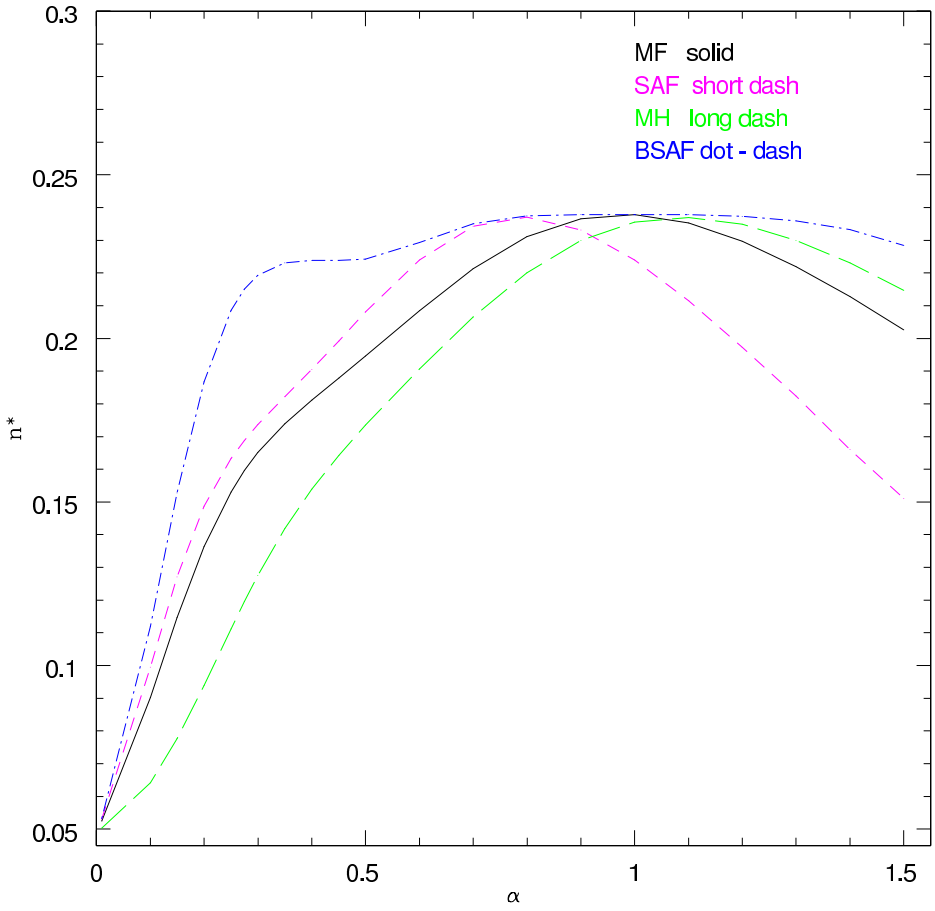

Figure 3.9 Scale-free distribution. The expected number density of detections n^* as a function of the filter parameter α for $\gamma = 1.5$ for the BSAF (using the optimal values of c), MF, SAF and MH filters. We consider the case $R = 3$, $n_b^* = 0.05$, $\nu_i = 0.5$, $\nu_f = 3$ and $\beta = 0.5$.

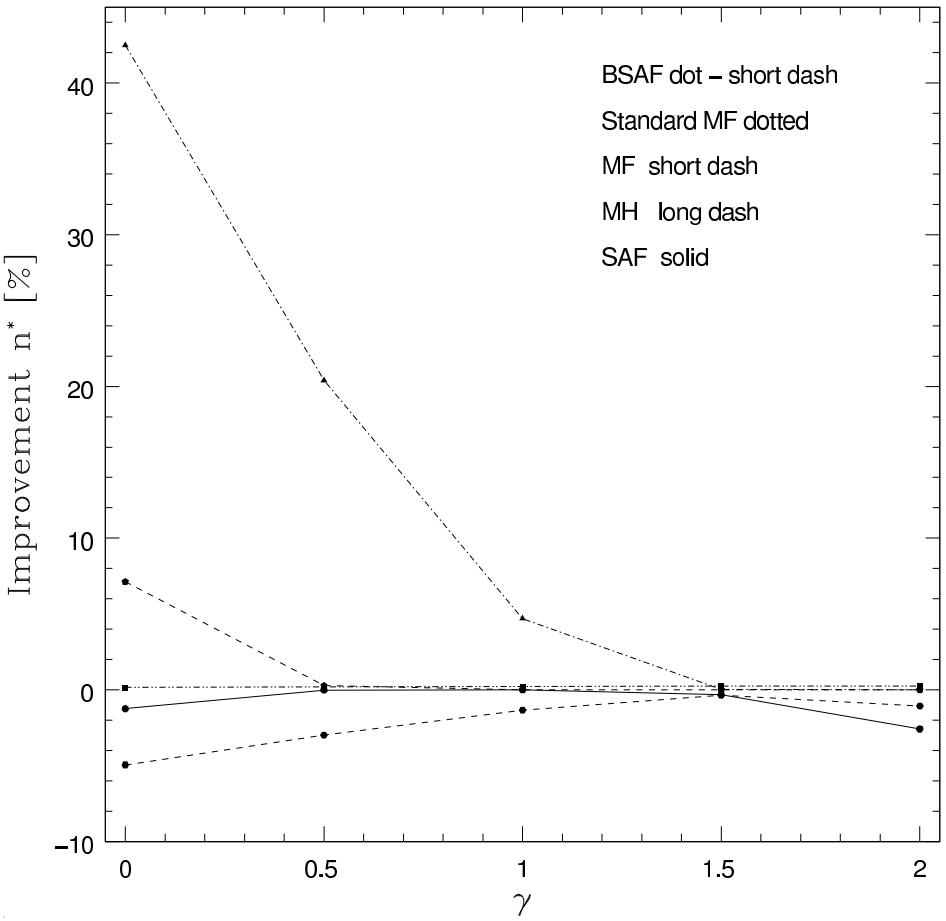

Figure 3.10 Scale-free distribution. Relative difference in the number of detections, with respect to the standard MF, as a function of the spectral index γ.

R	n_b^*	γ	α	c	n_{BSAF}^*	n_{MF}^*	$RD[\%]$
2	0.01	0	0.4	-0.66	0.1659	0.1590	8.2
2	0.03	0	0.4	-0.68	0.2376	0.2089	23.2
		0.5	0.4	-0.56	0.2451	0.2432	28
2	0.05	0	0.4	-0.68	0.2772	0.2311	32.8
		0.5	0.4	-0.57	0.2873	0.2705	11.6
3	0.01	0	0.3	-0.83	0.1336	0.1180	13.2
3	0.03	0	0.3	-0.83	0.1975	0.1512	30.6
		0.5	0.3	-0.71	0.1937	0.1767	9.6
3	0.05	0	0.3	-0.81	0.2335	0.1639	42.5
		0.5	0.3	-0.70	0.2321	0.1928	20.4
		1	0.3	-0.62	0.2271	0.2169	4.7

Table 3.2 Scale-free distribution. Number density of detections n^* for the standard MF(α=1) and the BSAF with optimal values of c and α. $RD \equiv 100(-1 + n_{BSAF}^*/n_{MF}^*)$.

Figure 3.10 shows the relative difference in the number of detections, with respect to the standard MF, as a function of the spectral index γ for the different filters. At each point the optimal scale (and parameter c in the case of BSAF) has been used. We remark that for the interesting case of white noise more than a 40% of detections is gained with respect to the standard MF.

3.6.3 On the robustness of the filters

The filters considered here depend on a number of parameters (α in the case of SAF, MF and MH and α and c in the case of BSAF) that must be determined in order to get the maximum number of detections for a fixed number of spurious detections. While for a given filter the region of acceptance is explicitly independent of the source distribution, the methodology presented here for the estimation of the optimal filter parameters depends on some assumed parameters of the source distribution (namely β, ν_i and ν_f) and the noise power spectrum (γ). A full study of the robustness of the method for all the filters is out of the scope of this work. However, we have considered some interesting cases as tests of the robustness of the method.

In order to ascertain to what extent the uncertainties in the β parameter of the source distribution affects the determination of the optimal filter parameters, we repeated our calculations using wrong assumptions on its value. An interesting case corresponds

to assume that the source distribution is uniform when it is scale-free and vice versa. Tables 3.1 and 3.2 show that both uniform and scale-free distributions lead to similar values of the optimal α and c parameters. If we use the optimal (α, c) that were obtained for the uniform distribution in the scale-free distribution with $\beta = 0.5$ the differences in the number of detections are very small (lower than 0.1%). The same happens in the opposite case.

Regarding the cut-offs, we observed that the qualitative behaviour of the filters is preserved even when the lower-limit is changed. The change of the upper-limit is not interesting since bright sources are always detected by all the filters. The shape of the $\alpha - n^*$ curves is unchanged. The position of the maxima does not change significantly, whereas the relative performances of the filters varies smoothly with the amplitude of the sources, and therefore changing the cut-offs leads only to a global change in the number of detections (amplitude of the curves) for all the filters.

We conclude that the uncertainty in the knowledge of the source distribution is not a critical issue in the cases we considered.

A more delicate issue is the one that appears when one assumes a γ parameter that is very different from the true one. In that case the shape of the filters changes dramatically (except in the case of the MH whose shape is independent of γ) and this may lead to wrong results. However, note that there are very well established techniques to estimate the power spectrum. Albeit in this academic case we consider power law-type backgrounds, it is straightforward to apply the method to any kind of power spectrum that can be present in the data.

As an example, we considered a case where the background corresponds to a true value $\gamma_t = 0.5$ whereas the filters have been constructed with a wrong $\gamma_w = 0.6$, that is, a 20% error in the determination of γ. The resulting $\alpha - n_*$ curves can be observed in figure 3.11.

The behaviour of the BSAF, the MF and the SAF is qualitatively similar to the case where the noise power spectrum is perfectly known, but the performance of the three filters is poorer. The MH curve is identical to the ideal case since the shape of the filter does not depend on γ. The BSAF still outperforms all the other filters, although the improvement in the number of detections with respect to the MF slightly decreases.

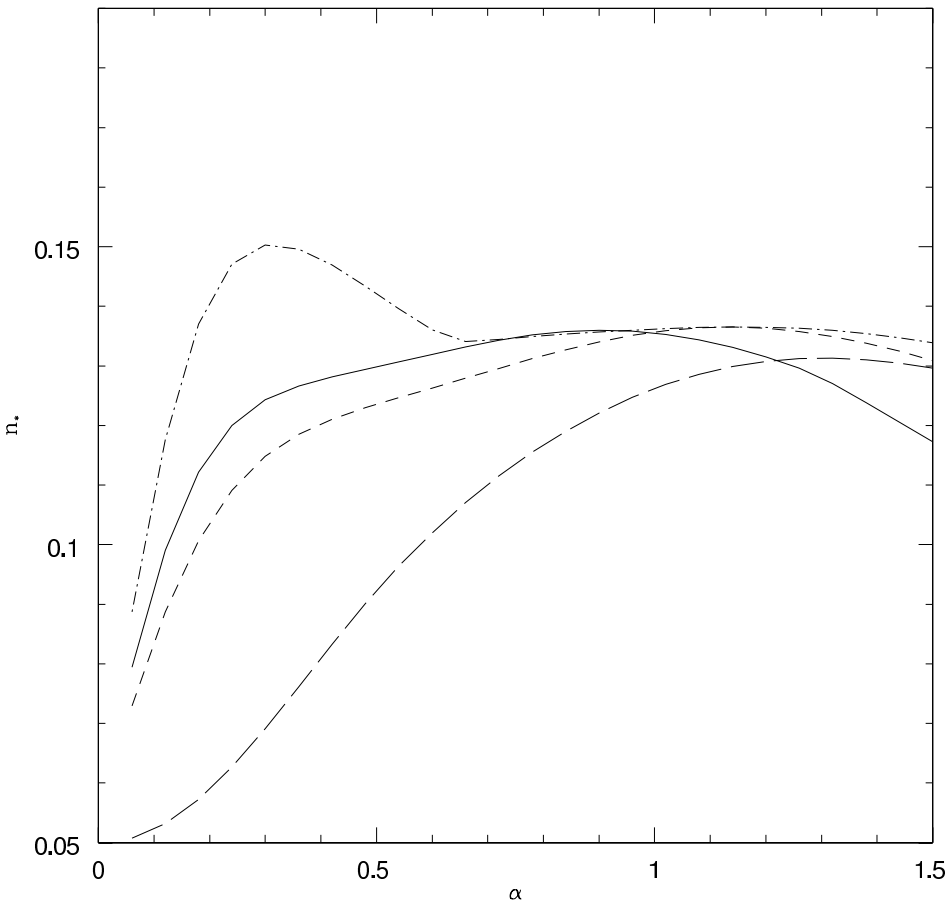

Figure 3.11 Filter performances when the γ parameter of the background is poorly known. A uniform distribution is considered. The true value of the background power spectrum index is $\gamma_t = 0.5$ whereas the filters hasve been constructed with a wrong $\gamma_w = 0.6$. Note how the figure compares with figure 3.4.

3.7 Simulations

3.7.1 The simulations

To further test these ideas, we generate a set of one-dimensional images containing a Gaussian background characterised by a white noise power spectrum (γ=0) to see how our theoretical framework works in a practical example. First, we study the properties of the maxima of the background alone, and second we add point sources and study the performance of the BNP detector after filtering with the different filters considered in the previous sections.

The different simulations are performed as follows. The images contain a number $N = 4096$ pixels, big enough so that the addition of a single source does not modify significantly the dispersion of the images. The background is generated with dispersion $\sigma_0 = 1$ (in arbitrary units); the sources that we consider for this example had a characteristic scale $R = 3$ pixels. The images filtered with a standard MF for this scale ($R = 3$, $\alpha = 1$) have dispersion $\sigma_0^{\rm MF} \simeq 0.43$. Since we are interested in the detection of weak sources, we add point sources distributed in intensity following a uniform distribution in the interval $\nu \in [0, 2]$ with respect to the standard MF. In terms of amplitudes, this corresponds to a distribution $A \in [0, 0.86]$, where $A = \nu\sigma_0^{\rm MF}$.

Given any simulation, it is possible to obtain the momenta σ_0, σ_1 and σ_2 directly from the image and, therefore, the quantities ρ, y_s needed to know the value of the linear detector φ. For every maximum in the image, it is possible as well to measure directly its amplitude A and curvature κ. The curvature is easily obtained by Fourier transforming the image, multiplying by q^2 and going back to real space. This gives the value of $-\xi''$ at each point and κ is obtained dividing by σ_2. For each considered filter, it is in principle possible to calculate the critical value φ_* that defines the acceptance region using eq. (16) just as we did in the previous sections, and hence to proceed with the BNP decision rule. Instead, since we are dealing with simulations of real data we will follow a fully empirical approach.

The argument goes as follows. We fix the number density of spurious detections, i.e., the number of maxima of the background that could be classified as "sources" by our detection criterion. Then we simulate a set of images corresponding only to background and filter them with the filter under study. We focus on the background maxima and try to determine the value of φ_* that makes the BNP rule to produce the specified number of spurious detections. For each filter, we analyse, let us say, 50000 noise realisations and focus on what happens in a certain pixel (we choose the central pixel of the simulation in order to avoid border effects). For every realisation, we check

if there is a maximum at this position or not. If a maximum is present, the value of φ is calculated. All the values of φ obtained in that way are sorted into descending numbers (large to small). The value of φ_* is then given by the φ corresponding to the r-th element ($r = n_b^* N_{tot}$) of the sorted list (that is, φ_* is the value of φ so that there are $n_b^* N_{tot}$ background maxima with $\varphi \geq \varphi_*$). For this example, we considered $n_b^* = 0.05$ and therefore $n_b^* N_{tot} = 2500$.

After the estimation of φ_*, we add a source with a Gaussian profile of dispersion $R = 3$ pixels at the central position (pixel $N/2 + 1$) of the background image and we filter with the considered filters respectively. All the relevant quantities needed for the detection are estimated directly from the images. We proceed to apply the detector to any maxima where the existence of a source is suspected at the pixel we are considering. Finally, n^*, for each of the filters, will be the number of sources with an estimated $\varphi \geq \varphi_*$ divided by the total number of realizations N_{tot}.

This procedure to obtain φ_* requires some cpu time, but the results, shown in figure (3.12), reveal the good agreement between the theoretical expected values and the simulations.

If we use the theoretical value of φ_*, instead of obtaining it from the simulations, we can reduce the computation time, but the improvement of the BSAF versus the MF may be reduced by a factor some times as large as 2. This can be explained by the fact that the effect of the pixel is not considered in the theoretical calculations. This affects specially the BSAF constructed with negative values of c, because they are more extense in Fourier space and therefore remove less power at small scales.

We have theoretically tested different values of n_b^*, ν_c, γ, R and α. We have performed numerical simulations for the most favorable case, assuming a uniform distribution with $\nu_c = 2$, where the BSAF outperforms the standard MF. This corresponds to the case $n_b^* = 0.05$, $R = 3$ and $\gamma = 0$. In the figure 3.12 we present the results from the simulations for this case and the comparison with the theoretical calculations. For the sake of simplicity, we give the results only for the BSAF and the MF since the other two filters (SAF and MH) perform worse in the considered case.

The lines in this plot show the theoretical results for each filter, the triangles the result from the simulations for the BSAF and the squares the results for the MF. We concentrate on the BSAF, which corresponds to the dot-dash line. As we mentioned in previous sections, the BSAF significantly improves the standard MF for $\gamma = 0$. We have done five simulations, each one of them generated with $\simeq 50000 - 60000$ realizations, 5000 of them containing a maximum of the background in the central pixel. We have obtained φ_* and used it to estimate the number of detections. The simulations follow

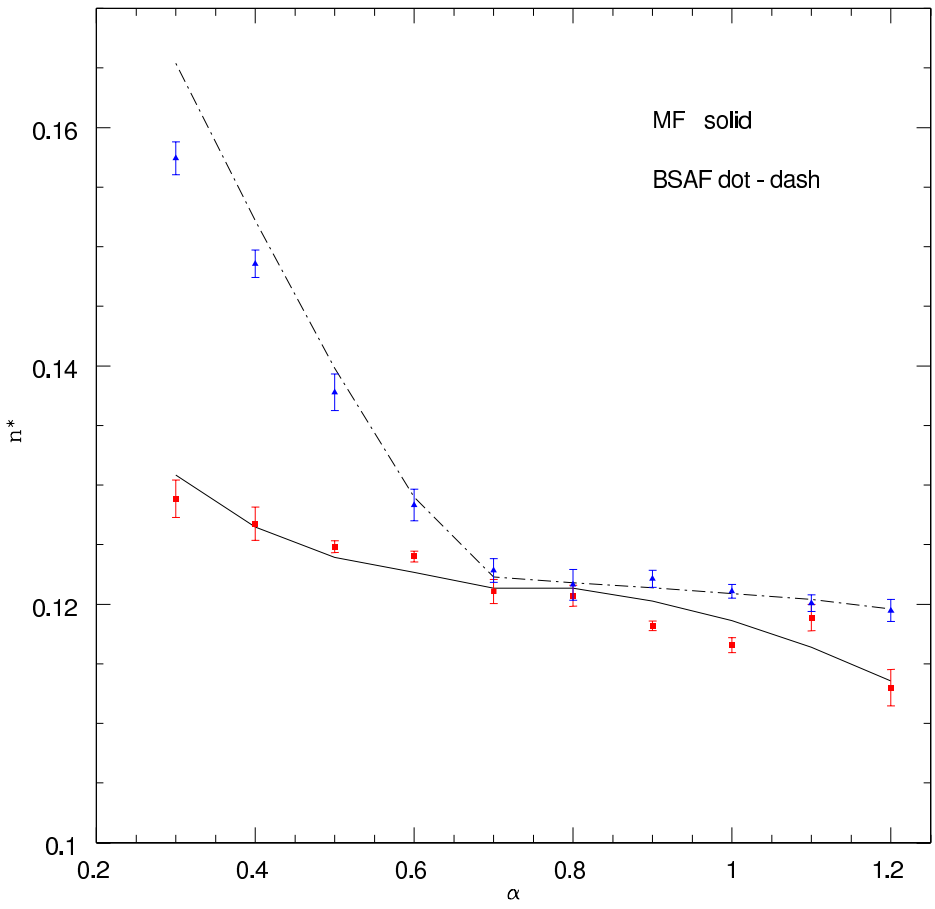

Figure 3.12 Expected number density of detections versus the filter parameter α for the BSAF and MF. The solid and dot-dash lines represent the theoretical number density n^* and the squares and triangle are the results from the simulations for $\nu_c = 2$, $n_b^* = 0.05$ R=3 and $\gamma = 0$, for the MF and the BSAF respectively.

the theoretical results well. In the region where $\alpha \simeq 0.3$, there is a small deviation from theory which we believe is related to the fact that we are close to the scale of the pixel, but still, significantly close to the expected theoretical value.

In order to test how the finite pixel size affects the previous results, we repeated the theoretical calculations introducing a cut-off in our integrals at $q_{max} = \pi$, corresponding to the limiting sampling frequency (inverse pixel units) in Fourier space. The results for the considered source scale (R=3 pixel) and different noise power spectra ($\gamma = 0$, 0.5, 1.5), for $n_b^* = 0.05$, show that the cut-off does not significantly change the predicted number of detections for α values ≥ 0.3. The differences appear at the 1% level, decreasing the number of detections as it was expected and approaching to the simulated values. This leaves a 5% of discrepancy between the theoretical and the simulated value for the BSAF and $\alpha = 0.3$ (pixel scale) that we believe is due to further discretization effects, affecting for example the estimation of curvatures. This discrepancy level is lower for the MF case.

3.7.2 The estimation of the amplitude of the source

We can estimate the amplitude of a source using the unbiased and maximum efficient estimator from equation (3.4.1) and then compare it with the amplitude that we have randomly generated. In figure 3.13, we plot the real amplitude versus the estimated one for the BSAF (top panels) and the MF (bottom panels).

In the left panels, we have simulated a uniform distribution of sources $A \in [0, 2]\sigma_0$. The parameters used for these simulations are $n_b^* = 0.05$, $\gamma = 0$. The points and the error bars are calculated as the average and the dispersion of the detected sources that fall in each of the amplitude bins from a total of $\simeq 10000$ detected sources. We find a similar positive bias in the determination of the amplitude for the BSAF ($\alpha = 0.3$, $c = -0.86$) and MF ($\alpha = 0.3$). However, the error bars corresponding to the BSAF are slightly smaller than those of the MF.

In the right panels, we give the results for a uniform distribution of sources with $A \in [0, 5]\sigma_0$. We find that the BSAF ($\alpha = 0.3$, $c = -0.79$) is unbiased for bright sources (see figure (3.13)), as compared to the MF ($\alpha = 1$). The BSAF, with $\alpha = 0.3$, is optimal regarding both detection and estimation, whereas the standard MF is biased regarding estimation for this distribution. As before, the points and the error bars are calculated as the average and the dispersion of the detected sources that fall in each of the amplitude bins from a total of $\simeq 20000$ detected sources. The parameters used for these simulations are $n_b^* = 0.05$ and $\gamma = 0$.

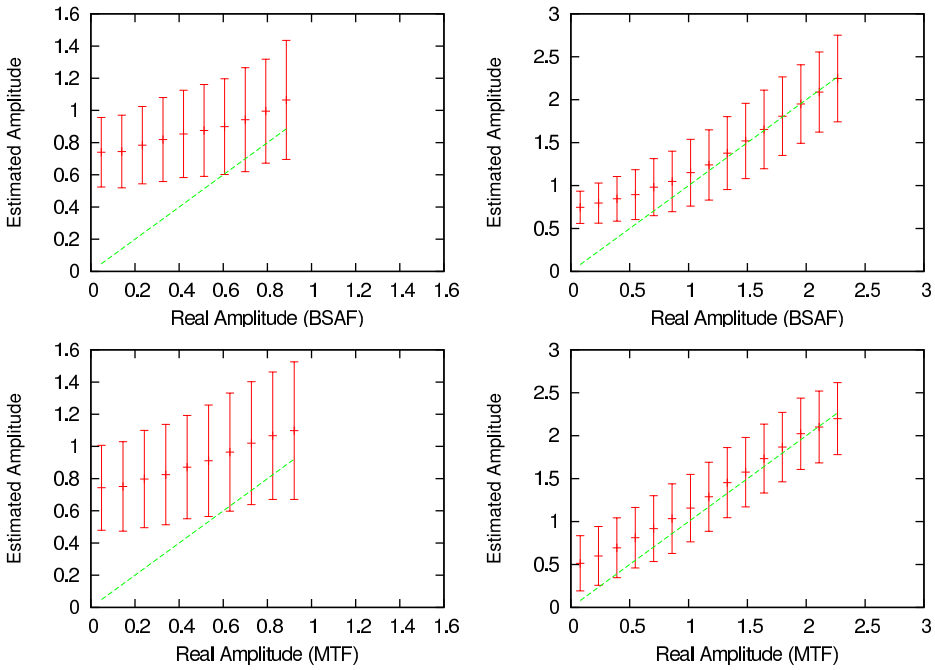

Figure 3.13 Real amplitudes versus Estimated amplitudes. For all four cases, the simulations have been done using a uniform distribution with cut-off's $A \in [0,2]\sigma_0$ and $A \in [0,5]\sigma_0$ for the left and right panels, respectively. We have also assumed $n_b^* = 0.05$ and $\gamma = 0$.

The fact that sources with small amplitudes are significantly affected by a positive bias can be explained taking into account that these sources are more easily detected if they lie over a positive contribution of the background. This contributes systematically to the overestimation of the amplitude. We would like to point out that this estimator produces appreciably better results than a naive estimation using directly the measured values at the maxima.

3.8 Conclusions

Nowadays, the detection of compact sources on a background is a relevant problem in many fields of science. A number of detection techniques use linear filters and thresholding-based detectors. Our approach to the problem of detector design is different. We use a Bayesian Neyman-Pearson rule that takes into account *a priori* information of the distribution of sources and the number density of maxima to define the acceptance region.

Our approach is based on maxima and includes information of both the curvature and the amplitude. The standard methods use the amplification of the sources produced by filtering. In our case, we take adventage not only of the amplification but also of the spatial information: the curvature of the background is different from that of the sources, and we use this to improve our detection rule. This information is contained in the momenta of the filtered field up to fourth order. The background is modelled by an homogeneous and isotropic Gaussian random field, characterized by a scale-free power spectrum $P(q) \propto q^{-\gamma}$, $\gamma \geq 0$.

We design a new filter that we call BSAF in such a way that the use of our improved detection rule based on amplification and curvature on the filtered field will increase the number of detections for a fixed number of spurious sources. We generalize the functional form of this filter, as well as other standard filters, and introduce another degree of freedom, α, that allows us to filter at any scale, including that of the source R. We have shown the benefits of filtering at scales smaller than R, which significantly improves the number of detections.

As an example, we have considered two different distributions of sources. A uniform distribution in the interval $\nu \in [0, 2]$ in the filtered field, and a scale-free power law distribution in the interval $\nu \in [0.5, 3]$, i.e. weak sources. The BSAF has proven to be significantly better than the standard MF, the SAF and MH wavelet in certain cases. In particular, the improvement in the number of detections of the BSAF at $\alpha = 0.3$ with respect to the standard MF is $\simeq 40\%$ for $\gamma = 0$, $n_b^* = 0.05$ and $R = 3$. We have also

tested the performance of the filters for a mixture of weak, intermediate and bright sources. For a uniform distribution with $A \in [0,5]\sigma_0$ and for a scale-free distribution with $A \in [0.5,5]\sigma_0$, the BSAF also improves the MF. However, for a scale-free distribution with $A \in [3,5]\sigma_0$, i.e., dominated by bright sources, we find that the optimal BSAF defaults to the standard MF, which gives the maximum number of detections in this case.

By construction, the BSAF gives in any case the best performance among the considered filters: the SAF and the MF are particular cases of the BSAF and the strategy we follow, i.e. maximization of the detections, guarantees that the parameters of the BSAF will default to the best possible filter in each case. Therefore, the number density of detections obtained with the BSAF will be at least equal to the best of the other three filters, and in some cases superior. In some other cases, however, the gain is small and it is justified to use an analitically simpler filter. We also remark that the BSAF is equivalent to the MH in some particular cases.

Our results suggest that, from the practical point of view, one could use the BSAF when $0 \lesssim \gamma \lesssim 1$ since, in this range, clearly improves the number of detections with respect to the other filters. However, for $\gamma \gtrsim 1.0$ the usage of the MH is justified due to its robustness in its functional form and it gives approximately the number of detections obtained either with the BSAF or MF.

For all the studied cases of sources distributions (except for the one dominated by bright sources) and fixing the values of γ, n_b^* and R, we find that the optimal parameters of the BSAF are only weakly dependent on the distribution of the sources.

We have done some simple tests in order to study the robustness of the method when the knowledge about the source pdf and/or the background spectral index is not perfect. We find that the values of the optimal filter parameters vary slightly when we assume that the source distribution is uniform when it is scale-free and vice versa. The uncertainties in the cut-off values of the source pdf slightly affect the number of detections, but in a similar way for all the filters, and therefore the relative behaviour of the filters do not change. Errors in the estimation of the spectral index γ reduces the efectiveness of the BSAF, but it still outperforms the other filters. All of this indicates that our detection scheme is robust against uncertainties in the knowledge of the distribution of the sources and spectral index. An extensive study of the robustness of filters is out of the scope of this work.

To test the validity of our results in a practical example, we have tested our ideas with simulations for the uniform distribution (using $n_b^* = 0.05$, $R = 3$, $\gamma = 0$) and find that the results follow the expected theoretical values.

Regarding source estimation, we propose a linear estimator which is unbiased and of maximum efficiency, that we have also tested with simulations.

The ideas presented in this work can be generalized: application to other profiles (e.g. multiquadrics, exponential) and non-Gaussian backgrounds is physically and astronomically interesting. The extension to include several images (multi-frequency) is relevant. The generalization to two-dimensional data sets (flat maps and the sphere) and nD images is also very interesting. Finally the application of our method to other fields is without any doubt. We are currently doing research in some of these topics.

3.9 Appendix A

The ratio $L(\nu,\kappa|\nu_s) \equiv n(\nu,\kappa|\nu_s)/n_b(\nu,\kappa)$ can be explicitly written as

$$L(\nu,\kappa|\nu_s) = e^{\varphi\nu_s - \frac{1}{2}(\mu+y_s^2)\nu_s^2}, \tag{3.9.1}$$

and taking into account the Bayesian criterion for detection

$$Ł(\nu,\kappa) \equiv \int_0^\infty d\nu_s\, p(\nu_s) L(\nu,\kappa|\nu_s) \geq L_*, \tag{3.9.2}$$

where L_* is a constant. By differentiating the previous equation with respect to φ

$$\frac{\partial L}{\partial \varphi} = \int_0^\infty d\nu_s\, p(\nu_s)\nu_s e^{\varphi\hat{\nu}_s - \frac{1}{2}(\mu+y_s^2)\hat{\nu}_s^2} \geq 0. \tag{3.9.3}$$

Therefore, $Ł(\nu,\kappa) \geq L_*$ is equivalent to $\varphi \geq \varphi_*$, where φ_* is a constant, i.e. $\varphi(\nu,\kappa)$ given by equation (3.3.7) is a sufficient linear detector.

3.10 Appendix B

Let us assume a linear estimator combination of the normalized amplitude ν and normalized curvature κ with the constrain

$$\hat{\nu}_s = A\nu + B\kappa. \tag{3.10.1}$$

If the estimator is unbiased, i.e. $\langle\hat{\nu}_s\rangle = \nu_s$, taking into account that $\langle\nu\rangle = \nu_s$ and $\langle\kappa\rangle = \nu_s y_s$, we obtain the constrain

$$A + By_s = 1. \tag{3.10.2}$$

On the other hand, the variance is given by

$$\sigma_{\hat{\nu}_s}^2 = A^2 + B^2 + 2\rho AB, \tag{3.10.3}$$

taking into account that $\sigma_\nu^2 = \sigma_\kappa^2 = 1$, $\langle \nu\kappa \rangle = \rho + y_s \nu_s^2$. By minimizing the previous expression with the constrain (3.10.2), one obtains

$$A = \frac{1}{y_s^2 + \mu} \frac{1 - \rho y_s}{1 - \rho^2}, \quad B = \frac{1}{y_s^2 + \mu} \frac{y_s - \rho}{1 - \rho^2}, \tag{3.10.4}$$

i.e. $\hat{\nu}_s = \varphi / (y_s^2 + \mu)$.

.