Biparametric adaptive filter: detection of compact sources in complex microwave backgrounds

M. López-Caniego* and P. Vielva

Instituto de Física de Cantabria (CSIC-Universidad de Cantabria), Avda. de Los Castros S/N, 39005 Santander, Spain

ABSTRACT

In this paper we consider the detection of compact sources in maps of the cosmic microwave background radiation following the philosophy behind the Mexican hat wavelet family (MHW*n*) of linear filters. We present a new analytical filter, the biparametric adaptive filter (BAF), that is able to adapt itself to the statistical properties of the background as well as to the profile of the compact sources, maximizing the amplification and improving the detection process. We have tested the performance of this filter using realistic simulations of the microwave sky between 30 and 857 GHz as observed by the *Planck* satellite, where complex backgrounds can be found. We demonstrate that doing a local analysis on flat patches allows one to find a combination of the optimal *scale* of the filter *R* and the *index* of the filter *g* that will produce a global maximum in the amplification, enhancing the signal-to-noise ratio (SNR) of the detected sources in the filtered map and improving the total number of detections above a threshold. We conclude that the new filter is able to improve the overall performance of the MHW2, increasing the SNR of the detections and, therefore, the number of detections above a 5σ threshold. The improvement of the new filter in terms of SNR is particularly important in the vicinity of the Galactic plane and in the presence of strong Galactic emission. Finally, we compare the sources detected by each method and find that the new filter is able to detect more new sources than the MHW2 at all frequencies and in clean regions of the sky. The BAF is also less affected by spurious detections, associated with compact structures in the vicinity of the Galactic plane.

Key words: filters.

1 INTRODUCTION

The emission of extragalactic point sources at microwave frequencies is known to be one of the most critical contaminants of the cosmic microwave background (CMB) anisotropies (de Zotti et al. 2005; Toffolatti et al. 2005). This foreground emission is a strong source of bias in the estimation of the CMB temperature and polarization angular power spectra. Even at the frequency range of 60–90 GHz (which covers an optimal window for observing the CMB), this effect is important. In particular, point sources contaminate the power spectra for multipoles $\ell \geq 800$ (Tegmark 1997; Tucci et al. 2005), which is translated into a bias when determining the cosmological parameters from these quantities. In particular, the scalar spectral index n_s and the optical depth τ are two of the most biased parameters (Rubiño-Martín & Barreiro, private communication). In addition, the foreground emission of point sources also introduces a high level of non-Gaussianity in the CMB anisotropies (e.g.

Argüeso, González-Nuevo & Toffolatti 2003). This non-Gaussian signal is a very important confusion noise when someone is probing non-standard models of structure formation, which actually predict a certain degree of non-Gaussianity on the CMB anisotropies, typically at a similar or lower level than the Gaussian deviation caused by the point sources (e.g. Komatsu et al. 2003; Curto, Martínez-González & Barreiro 2009). In addition to the *role* of the extragalactic point sources as CMB contaminants, characterizing their properties at the microwave frequencies is a very important field per se. For example, studying the number of objects per flux interval provides useful information to understand the history of galaxy evolution (e.g. González-Nuevo et al. 2008; Tucci et al. 2011).

A large number of works have been presented in the literature to mitigate the impact of extragalactic point sources on the science that can be extracted from the analysis of the CMB anisotropies. Two global approaches are usually followed, being, in fact, complementary. On the one hand, detection algorithms are applied to the CMB images for removing or masking the emission due to the brightest point sources. On the other hand, statistical modelling of the background of the residual point sources is adopted. This modelling

*E-mail: caniego@ifca.unican.es

is followed in the determination of the cosmological parameters as well as when exploring the compatibility of the CMB signal with a given non-Gaussian scenario. This work lies in the former category: the detection of the brightest point sources. The literature on this field is quite large, not only in relation to the development of specific tools for the CMB problem but also to the adaptation of techniques originally developed in radio (e.g. Högbom 1974) and optical (e.g. Bertin & Arnouts 1996) astronomy. We refer to Herranz & Vielva (2010) for a complete and recent tutorial on the detection of compact sources in CMB maps.

Some of the most popular tools to perform the detection and flux estimation of point sources are wavelets. Wavelets represent the simplest case of point source detection, since it is just based on a thresholding criterion, without any prior knowledge on the statistical properties or the background (as the matched filter; e.g. Tegmark & de Oliveira-Costa 1998; Argüeso et al. 2009; López-Caniego et al. 2009) or, even more, on the statistical properties of the background and the signal (as the Bayesian methods; e.g. López-Caniego et al. 2005; Carvalho, Rocha & Hobson 2009; Argüeso et al. 2011). For CMB experiments with a point spread function (PSF) well defined by a Gaussian function, the Mexican hat wavelet (MHW; built as function proportional to the Laplacian of a Gaussian kernel) has proved to be a very good tool (e.g. Cayón et al. 2000). In addition, the wavelet scale R can be optimized to provide a larger number of detections, depending on the statistical properties of the image (Vielva et al. 2001). Further improvement can be achieved by applying subsequently the Laplacian operator to the Gaussian kernel. This produces a series of wavelets known as the Mexican hat wavelet family (MHWn; González-Nuevo et al. 2006, where MHW1 is the standard MHW). The degree of derivation n is a reflection of the statistical properties of the background. In particular, low values of n define a filter that has most of the power at low frequencies (or equivalently, in the context of the CMB, at large scales). As the value of n increases, the power of the filter moves towards smaller scales. Therefore, if an image is dominated by large-scale structures in the background, then one should expect a better point source detection efficiency for a higher value of n than for a smaller one.

In López-Caniego et al. (2006), using simulations based on the *Planck* Reference Sky Model available at that time, it was shown that, on average, the MHW2 provides a larger number of point source detections than the MHW and the matched filter (MF), and a better flux estimation (see López-Caniego et al. 2006 for further details of the simulations, and Tauber et al. 2010 for a description of the *Planck* mission). As a matter of fact, the MHW2 has been successfully applied to the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) maps (López-Caniego et al. 2007), providing a larger number of detections than the matched filter used by *WMAP* (Hinshaw et al. 2007; Wright et al. 2009) as well as other methods (Chen & Wright 2008, 2009).

In this work we generalize the idea behind the MHWn, by defining a biparametric filter, $\psi_g(R)$, where the parameter related to the background fluctuations g is allowed to vary in a continuous way. Therefore, the determination of the g and R parameters is done jointly, attending to the statistical properties of the background. The paper is organized as follows. In Section 2 the biparametric filter is defined. We also discuss under which circumstances the new filter defaults to other known filters. The performance of the method is illustrated in Section 3 by analysing in detail some simulated microwave images. A statistical analysis on how the biparametric filter behaves as a function of the frequency and the Galactic latitude is presented. Conclusions are given in Section 4.

2 METHODOLOGY

As mentioned above, considering the statistical properties of the background in the vicinity of a compact object is a key step in the design of a filter to be used for the detection of point sources. This issue can be as simple as looking for the optimal *scale* that provides the maximum amplification for wavelets or as sophisticated as providing a complete likelihood and prior functions, for instance, when applying Bayesian approaches. In all-sky CMB experiments such as *WMAP* (Bennett et al. 2003) or *Planck* (Tauber et al. 2010) one can deal with very different backgrounds. For instance, the properties of the background have a typical dependence on the Galactic latitude: the major emission corresponds to the Galactic components
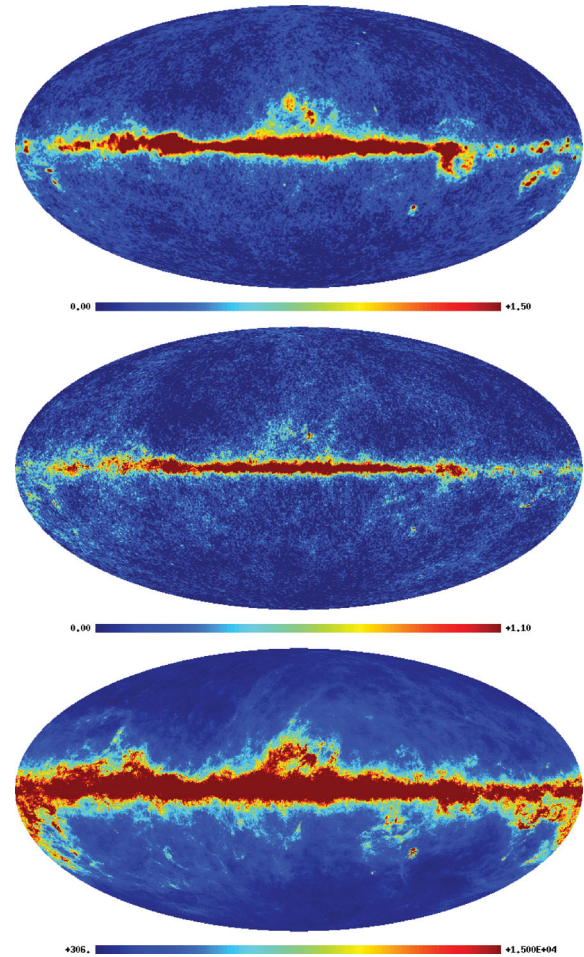

Figure 1. From top to bottom, a combination of noise, CMB, Galactic diffuse emission and compact source emission maps at 30, 143 and 857 GHz simulated with the pre-launch *Planck* Sky Model.

Table 1. The Galactic coordinates of the centres of the three regions that we have selected visually in the simulated maps at 30, 143 and 857 GHz in order to have increasing background complexities. We also show the optimal value of the *index* of the filter g that we have found in each case.

Region	GLAT	GLON	g_{30}	g_{143}	g_{857}
1	+60	70	6.66	4.33	1.83
2	−36	290	6.66	6.50	4.83
3	+01	123	5.66	6.83	6.66

(synchrotron, free–free, dust) near the Galactic plane, whereas the CMB and the background of extragalactic sources are, typically, more important at intermediate and high latitudes. In addition, depending on the observational frequency, radio emissions (Galactic and extragalactic) are more important in the lower part of the microwave frequency range (i.e. from 10 to 100 GHz), whereas thermal dust and the cosmic infrared background are the major contributors to highest frequencies (i.e. from 300 to 1000 GHz). However, at intermediate frequencies the CMB is, overall, the most important background. Each one of these emissions has its own specificities and, therefore, it is suboptimal to use the same filter at all frequencies and positions in the sky.

In this paper we explore the detection of objects in different types of backgrounds commonly found in CMB experiments. These objects are, in general, point-like sources that have been convolved with the beam PSF of the instrument used for the observation. They can be described as a signal $s(\boldsymbol{x}) = A\tau(\boldsymbol{x})$, where A is the intrinsic flux density of the object, $\tau(\boldsymbol{x})$ is the PSF of the beam and $\boldsymbol{x}$ is a unit vector in the sky. Although the profile $\tau(\boldsymbol{x})$ can be described

by any function, the formalism adopted in this paper assumes that it is given by an isotropic PSF, i.e. $\tau(\boldsymbol{x}) \equiv \tau(|\boldsymbol{x}|) = \tau(x)$. We adopt this simplification because for most of the CMB experiments an effective isotropic window function is usually defined, even when the PSF has a certain degree of anisotropy. In any case, of course, the filter proposed in this paper can be easily generalized for non-circular beam profiles.

2.1 The biparametric adaptive filter

Following the idea behind the MHWn, in this paper we propose a filter that has two free parameters that will allow us to better incorporate into the filter the statistical properties of the background. As for the case of wavelets, one of the parameters of the filter is the *scale R*. The filter *scale* is associated with a compression/expansion of the typical scale of the PSF and provides the size of the filtering kernel. In addition, we incorporate the *index* of the filter g that can be seen as a generalization of the role played by the order of the Laplacian operator used in the definition of the MHWn (e.g.
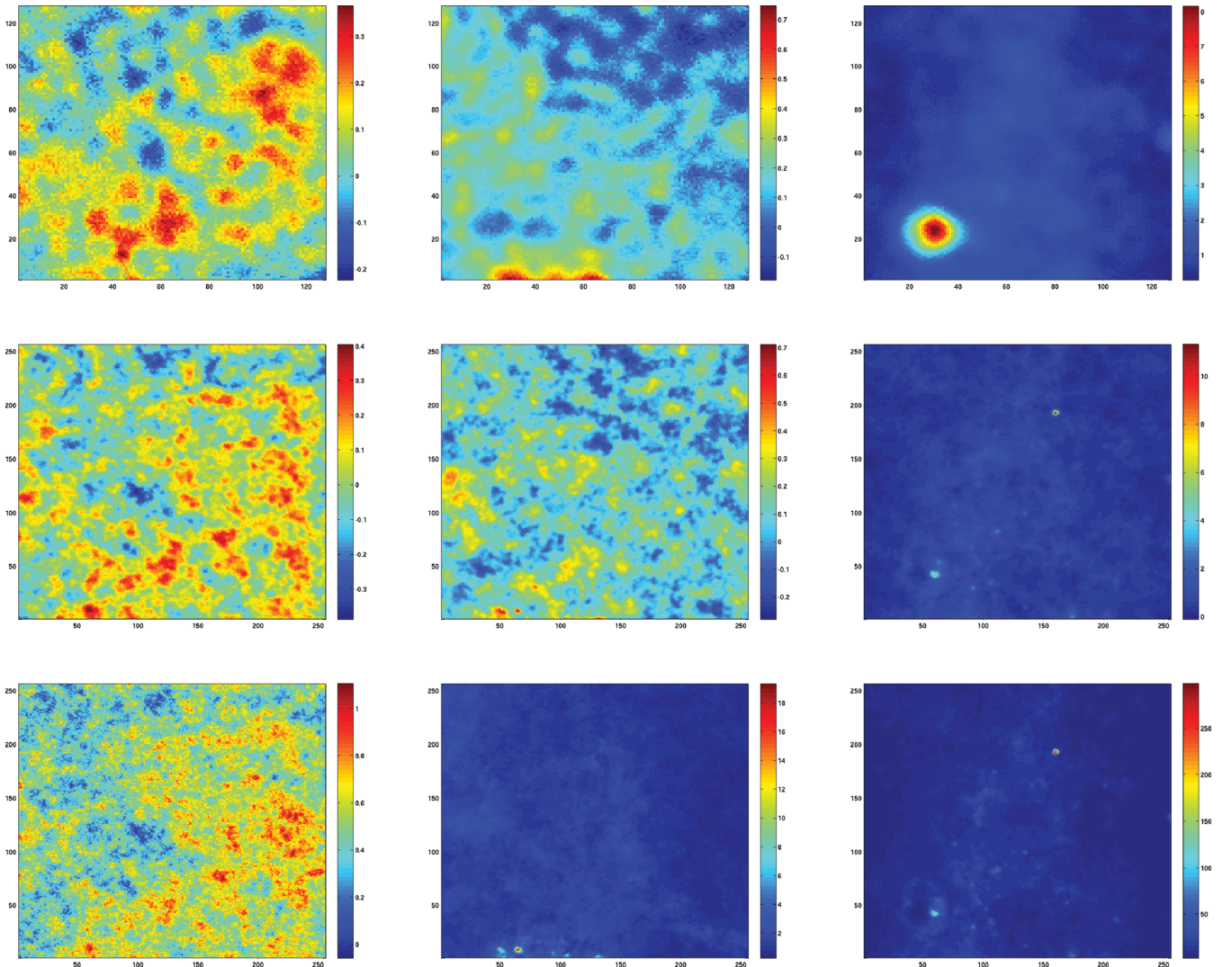

Figure 2. Selected regions with increasing background complexities extracted from the 30, 143 and 857 GHz simulated maps. At 30 GHz these patches have 128×128 pixels in size, where the pixel size is 3.43 arcmin. At 143 and 857 GHz these patches have 256×256 pixels in size, where the pixel size is 1.71 arcmin. The upper panels correspond to the 30 GHz regions, the middle panels to the 143 GHz regions and the lower panels to the 857 GHz regions. The left-, centre and right-hand panels correspond to regions 1, 2 and 3, respectively. The coordinates of the centre of these three regions can be found in Table 1.
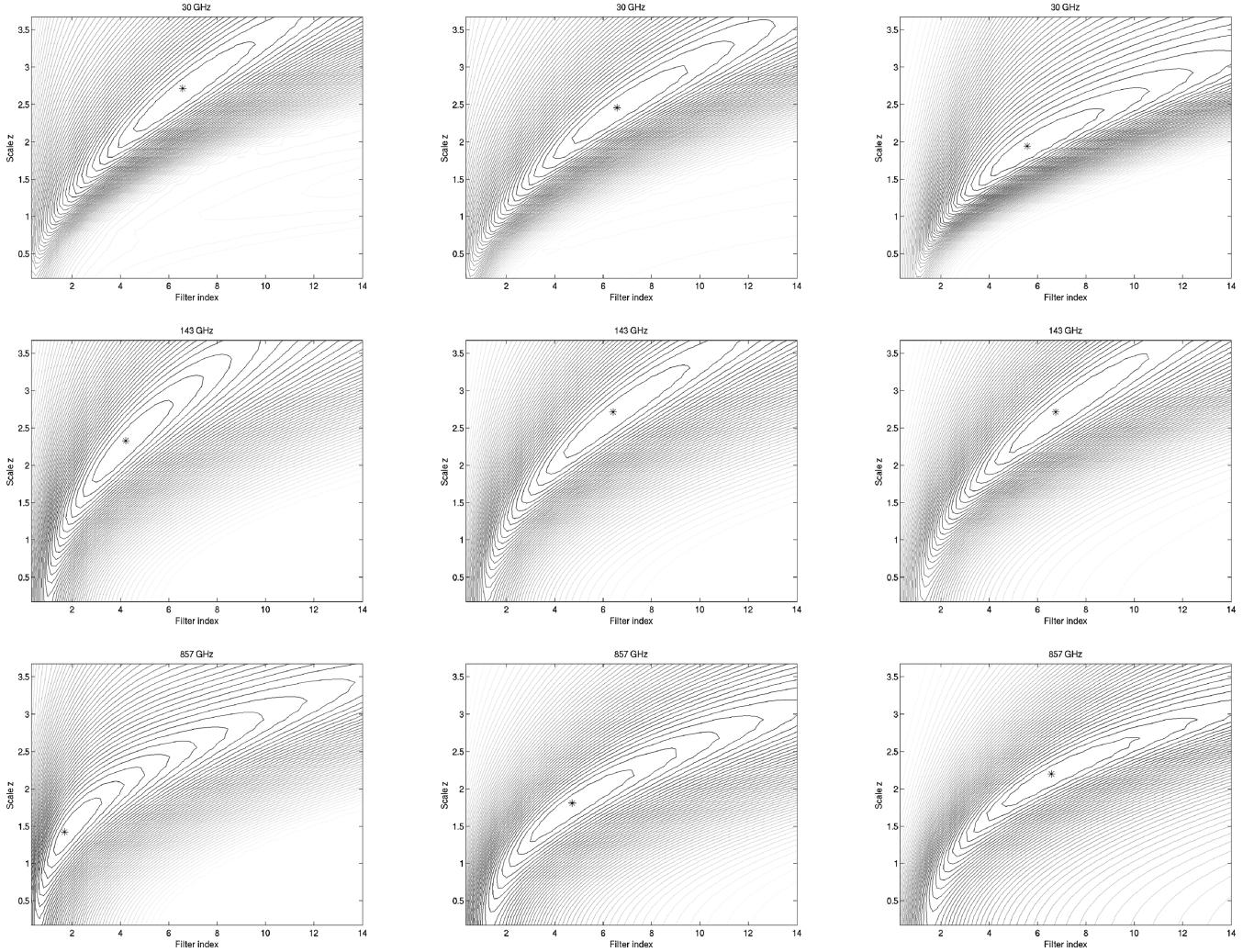
© 2012 The Authors, MNRAS **421,** 2139–2154
Monthly Notices of the Royal Astronomical Society © 2012 RAS

Figure 3. Point source amplification as a function of the *index g* versus the *scale z = R/σ*ᵦ of the filter for a selection of regions. The upper, middle and lower panels correspond to the 30, 143 and 857 GHz maps, respectively. These panels correspond, from left to right, to the regions 1, 2 and 3. Every contour line represents 1/25th of the maximum amplification.

González-Nuevo et al. 2006).[1] It is related to the filter location in Fourier space or, conversely, with the filter oscillations in the real space. As it will be discussed in the next subsection, this parameter is somehow associated with the shape of the angular power spectrum of the background. The filter definition is such that it is compensated (e.g. Cayón et al. 2000) and, therefore, for the case adopted in this paper of 2D Euclidean images, it behaves as a wavelet. The only case in which there is no compensation and, therefore, the filter is not a wavelet, is for the particular case in which the background is defined by a white noise field. In this particular situation, the filter defaults to a Gaussian kernel with $g = 0$ (see next subsection).

Let us denote the biparametric adaptive filter (BAF; defined in terms of the two parameters R and g) by $\Psi(\boldsymbol{x}; R, g, \boldsymbol{b})$. It denotes the value of the filter centred at $\boldsymbol{b}$, in the position $\boldsymbol{x}$. Following the philosophy behind the wavelets, this filter can be seen as the scaling version of a translated mother filter:

$$\Psi(\boldsymbol{x}; R, g, \boldsymbol{b}) \equiv \frac{1}{R^2} \bar{\psi}_g \left(\frac{|\boldsymbol{x} - \boldsymbol{b}|}{R} \right). \tag{1}$$

---

[1] Notice that, whereas the order n of the Laplacian operator is a natural number, the *index g* is defined as a real number ≥ 0.

This mother filter $\bar{\psi}_g$ defines the BAF and, in Fourier space, is given by

$$\hat{\psi}_g (qR) = \frac{1}{\pi} \frac{1}{\Gamma ((2 + g)/2)} (qR)^g \, \tau(qR). \tag{2}$$

In particular, and following the convention in Sanz, Herranz & Martínez-Gónzalez (2001), convolving such a filter with a 2D image $f(\boldsymbol{x})$ we obtain the filter coefficients $\omega_g(R, \boldsymbol{b})$ defined as

$$\omega_g (R, \boldsymbol{b}) = \int \mathrm{d}\boldsymbol{x} \, f(\boldsymbol{x}) \Psi(\boldsymbol{x}; R, g, \boldsymbol{b}). \tag{3}$$

A very common situation in astronomy is to have a PSF defined by a Gaussian function $\tau(\boldsymbol{x}) = (1/2\pi\sigma_b^2)\mathrm{e}^{-(1/2)(\boldsymbol{x}/\sigma_b)^2}$, where σ_b is the Gaussian beam dispersion. For this particular case, the BAF is given by

$$\hat{\psi}_g (qR) = \frac{1}{\pi} \frac{1}{\Gamma ((2 + g)/2)} (qR)^g \, \mathrm{e}^{-(1/2)(qR)^2}, \tag{4}$$

and the filter coefficients at the position of a source with a profile $I_0\tau(qR)$, where I_0 is the amplitude of the source, is given by

$$\omega_g (R) = \frac{I_0 2^{(g+2)/2} z^g}{\left(1 + z^2\right)^{(g+2)/2}}, \tag{5}$$

where $z \equiv R/\sigma_b$. As it was introduced by Vielva et al. (2001), the filter parameters are determined by imposing a maximum amplification $\mathcal{A}$ of the point source amplitude in the filter coefficients map:

$$\mathcal{A} \equiv \frac{\omega_g\,(R)\,/\sigma_\omega}{I_0/\sigma}, \tag{6}$$

where σ is the dispersion of the image in the real space, I_0 is the amplitude of the source and σ_ω is the dispersion of the filter coefficients at the *scale R* and *index g* that can be defined as

$$\sigma_\omega^2\,(R, g) \equiv 2\pi \int_0^\infty \mathrm{d}q\; q\, P\,(q)\, \hat{\psi}_g^2(Rq), \tag{7}$$

with $P(q)$ being the angular power spectrum of the analysed image.

2.2 The BAF and other filtering kernels

The BAF introduced in equation (2) defaults, under certain conditions, to other well-known filter kernels extensively used in literature. In this sense, the BAF can be considered as a generalization of these filters. The most obvious generalization is for the MHW*n* (e.g. González-Nuevo et al. 2006, $\hat{\psi}_n$). As previously mentioned,

when the profile of the point sources is described by a Gaussian function, the BAF is given by equation (4). Both filters are related by $\hat{\psi}_g = (2^n/\pi)\hat{\psi}_n$, where $g = 2n$ and $n = 1, 2, 3, \ldots$.

Another interesting situation occurs when the background can be described by an angular power spectrum following a power law $P(q) = Cq^{-\gamma}$, for $q > 0$, and C being a normalization constant. In this case, it is easy to show, from equation (7), that the variance of the filter coefficients is given by

$$\sigma_\omega^2\,(R, g) = C\frac{\sigma_b^{\gamma-2}}{\pi}\frac{\Gamma\,((2g + 2 - \gamma)/2)}{[\Gamma\,((2 + g)/2)]^2}z^{\gamma-2}. \tag{8}$$

From equation (6), one can demonstrate that the maximum amplification occurs when $R \equiv \sigma_b$ and $g \equiv \gamma$, i.e. the BAF defaults to the matched filter.

3 THE SIMULATIONS

In order to test the performance of the new filter we will use realistic simulations of the microwave sky at 30, 44, 70, 100, 143, 217, 353, 545 and 857 GHz. These simulations have been generated with the pre-launch *Planck* Sky Model (Leach et al. 2008; Delabrouille et al., in preparation), a software package developed within the
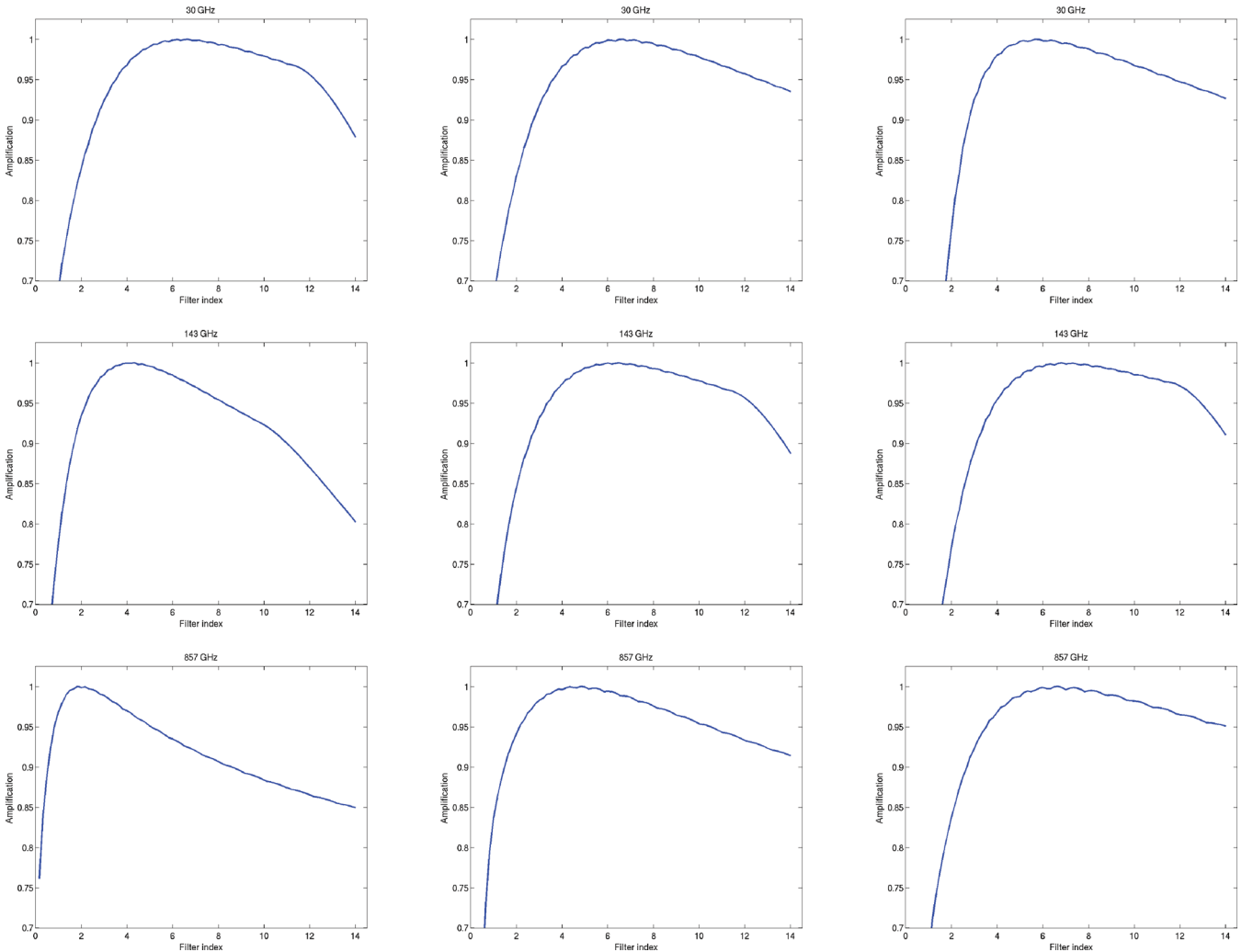

Figure 4. Point source amplification as a function of the *index g*, conditioned on the optimal *scale R*. The upper, middle and lower panels correspond to the 30, 143 and 857 GHz maps, respectively. These panels correspond, from left to right, to the regions 1, 2 and 3. Note that the amplification has been normalized to the maximal amplification in each case.

Planck Collaboration that allows one to make simulations at the microwave frequencies of the CMB, Galactic diffuse emissions and compact sources. The Galactic interstellar emission adopted in this paper is described by a three-component model of the interstellar medium comprising free–free, synchrotron and dust emissions plus added small-scale fluctuations to reproduce the non-Gaussian nature of the interstellar emission. Free–free emission is based on the model of Dickinson, Davies & Davis (2003) assuming an electronic temperature of 7000 K, where the spatial structure of the emission is estimated using a Hα template corrected for dust extinction. Synchrotron emission is based on an extrapolation of the 408-MHz map of Haslam et al. (1982) from which an estimate of the free–free emission was removed. The thermal emission from interstellar dust is estimated using model 7 of Finkbeiner, Davis & Schlegel (1999). Point sources are modelled with two main categories: radio and infrared. Simulated radio sources are based on the NRAO VLA Sky Survey (NVSS) or Sydney University Molonglo Sky Survey (SUMSS; Condon et al. 1998; Mauch et al. 2003) and Green Bank 6-cm (GB6) or Parkes–MIT–NRAO (PMN) catalogues (Griffith et al. 1995; Gregory et al. 1996). Measured fluxes at 1 and/or 4.85 GHz are extrapolated to *Planck* frequencies assuming a dis-

tribution in flat and steep populations. Infrared sources are based on the *IRAS* catalogue (Beichman et al. 1988), and modelled as dusty galaxies. A detailed description of each component can be found in Leach et al. (2008). These simulations use the HEALPix pixelization scheme (Górski et al. 2005) with NSIDE = 1024 at 30, 44 and 70 GHz and NSIDE = 2048 for the rest. In Fig. 1 we show three of these simulations, at 30, 143 and 857 GHz.

3.1 Patch analysis

As it was mentioned above, the purpose of this work is to study the performance of the new filter in different backgrounds. For this reason, and to better illustrate the problem, we have studied three different regions of the sky at each of the following three frequencies: 30, 143 and 857 GHz. For each region, we have projected flat patches of 7.3×7.3 deg^2. At 30 GHz each patch has 128×128 pixels and a pixel size of 3.43 arcmin. At 143 and 857 GHz, each patch has 256×256 pixels and a pixel size of 1.71 arcmin. Each region has been selected by visual inspection to have increasing background complexities. In general, the statistical properties of the background in those patches with lower Galactic latitudes are more complex than
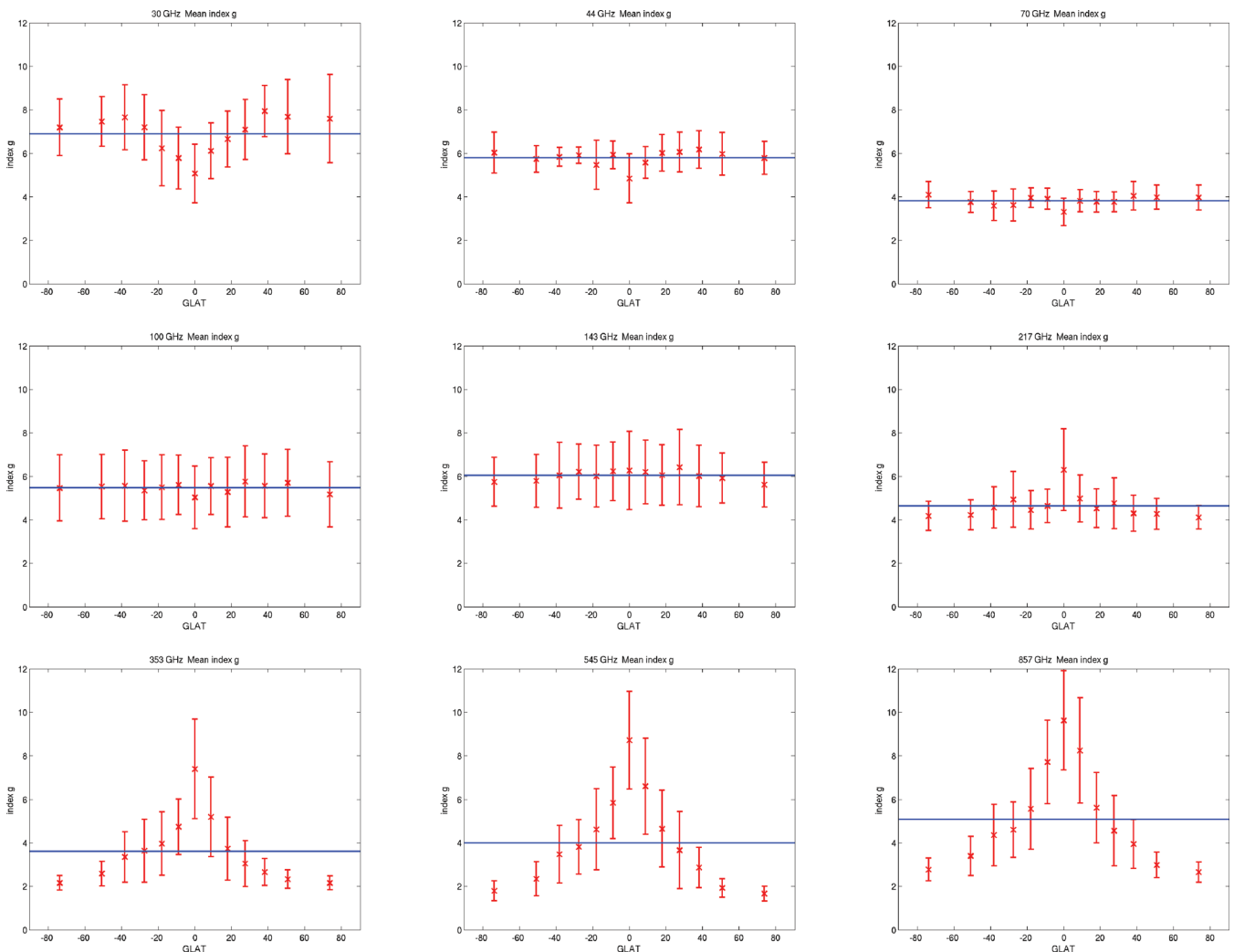
Figure 5. The mean filter *index g* per the galactic latitude bin and its dispersion are shown. The galactic latitude bins have been chosen to cover an equal area on the sky. In blue the mean *index g* across the sky is given.

those in the higher latitudes. The coordinates of the centres of the patches can be found in Table 1. In Fig. 2 one can see the nine selected regions.

In order to determine the filter characteristics that best adapts to the properties of each patch we look for the optimal *scale R* and *index g* of the filter that maximize the amplification of the sources present in the patch. An increase in the amplification will allow one to detect more sources above a certain threshold. In Table 1 one can see the values of *g* obtained for each patch and each frequency. Note that the optimal values of *g* and *R* vary from one region of the sky to another. For the 30 and 143 GHz cases, these variations are small even for different Galactic latitudes, particularly for the 143 GHz case, but at 857 GHz, where the Galactic dust and the far-infrared background dominate, the variations are large. This simple analysis already shows the importance of optimizing not only the *scale* but also the *index* of the filter.

To further illustrate this exercise, in Fig. 3 we show, for the selected regions, the amplification as a function of the *scale R* and *index g* of the filter. Since the behaviour of the amplification as a function of the optimal *scale R* has been already studied in previous works (e.g. Vielva et al. 2001; López-Caniego et al. 2006), we will concentrate on the properties of the *index g*. In this respect, in

Fig. 4 we show the results of the same analysis but representing the amplification as function of the *index g*, conditioned on the optimal *scale R*. As we mentioned above, for each region there is a combination of *R* and *g* that produces a single maximum in the amplification. In particular, if we look to the upper left-hand panel of Fig. 4 corresponding to the region 1 at 30 GHz, the maximum in the amplification is reached for $g = 6.66$. If we use, instead, the standard MHW at the optimal *scale*, $g = 2$, the amplification is only 82 per cent of that of the maximum. If we compare the MHW2 at the optimal *scale*, $g = 4$, the amplification is ∼97 per cent. If we look at one of the regions for the 143 GHz case, for example the middle central panel of Fig. 4, the maximum appears at $g = 6.5$. In this case we can see a similar behaviour for the MHW at the optimal *scale* as before, the amplification is ∼83 per cent of that of the maximum and for the MHW2 at the optimal *scale* the amplification is ∼97 per cent. More interesting is the case of 857 GHz, where the maximum in *g* changes significantly from one region of the sky to another. In this frequency, the maximum of the lower left-hand panel of Fig. 4 is at $g = 1.83$ and the MHW and MHW2 at the optimal *scale* produce an amplification of ∼95 and ∼100 per cent, respectively. Then, the maximum in the lower central panel is at $g = 4.33$, and the MHW and MHW2 at the optimal *scale* produce an amplification of
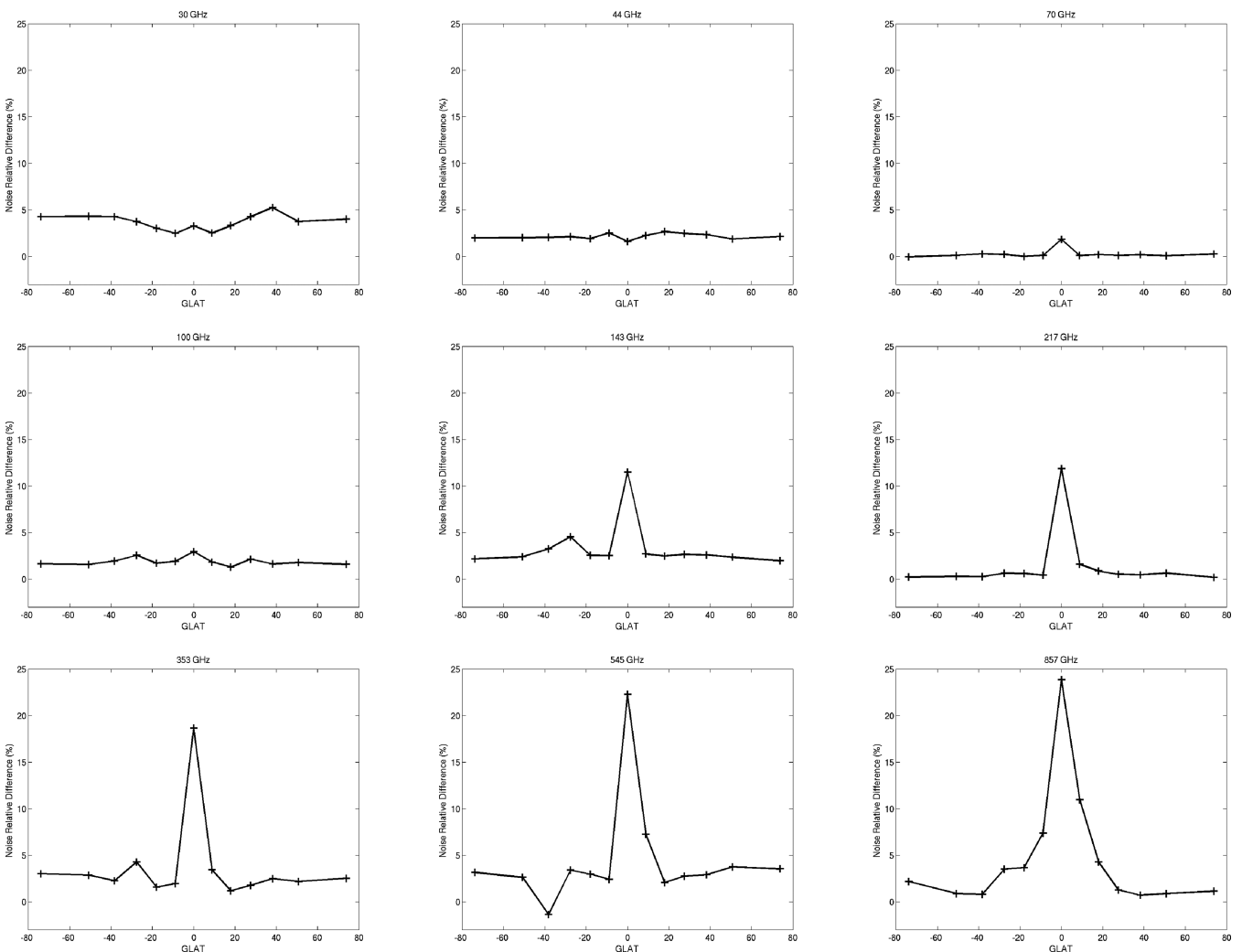

Figure 6. The average relative improvement (of BAF with respect to MHW2) in σ_ω is given. The BAF is able to improve the estimation of the noise up to 25 per cent in particularly complex regions, as in those areas dominated by Galactic emission.

© 2012 The Authors, MNRAS **421,** 2139–2154
Monthly Notices of the Royal Astronomical Society © 2012 RAS

~94 and ~100 per cent, respectively. Finally, the maximum in the lower right-hand panel is found at $g = 6.66$, and the MHW and MHW2 at the optimal *scale* produce an amplification of ~82 and ~97 per cent, respectively.

3.2 All-sky analysis

In the previous section we have studied the performance of the BAF optimizing the two parameters, the *scale* of the filter R and the *index*

of the filter g, that produce a maximum in the amplification of the filtered map as compared with the unfiltered one. This study was done looking at three visually selected regions in the sky for each of the three considered frequencies. The results for this analysis with only nine regions already show that choosing the right *index* g and *scale* of the filter R can increase the performance of the filter significantly. In this section we want to improve the statistics of the analysis by increasing the percentage of sky and frequency coverage of the study. For this purpose, we use nine full-sky simulations
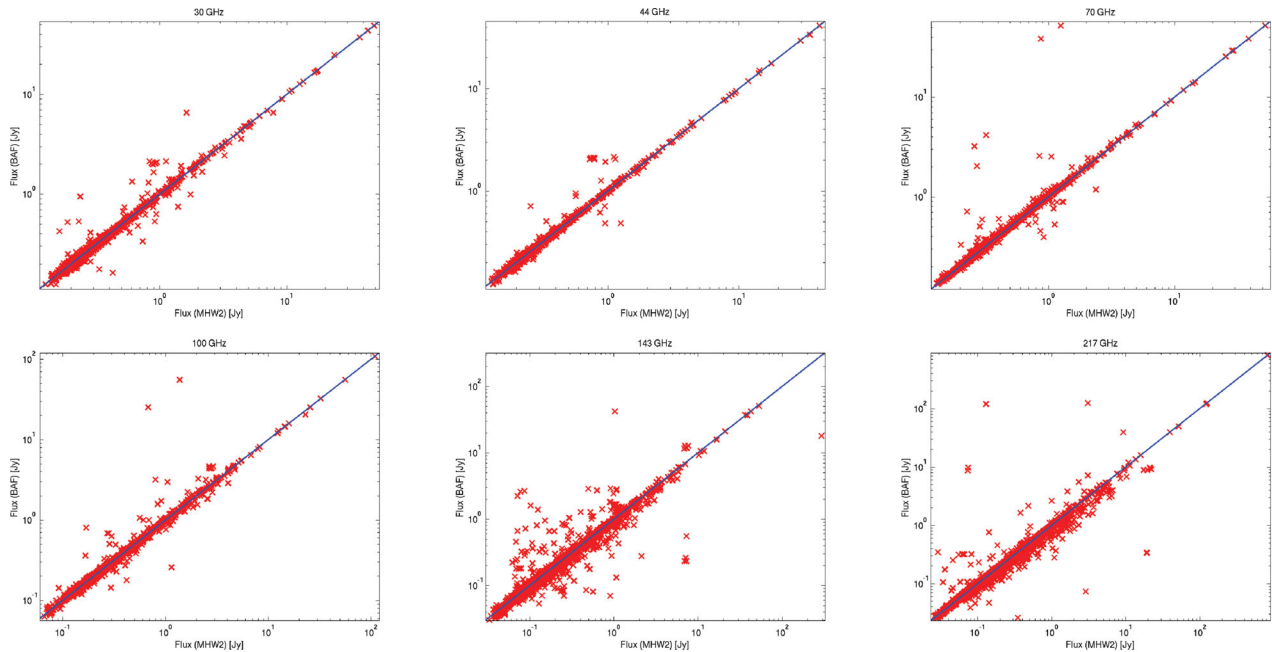

Figure 7. Comparison of the estimated flux density of the common sources detected by both methods for channels 30–217 GHz.
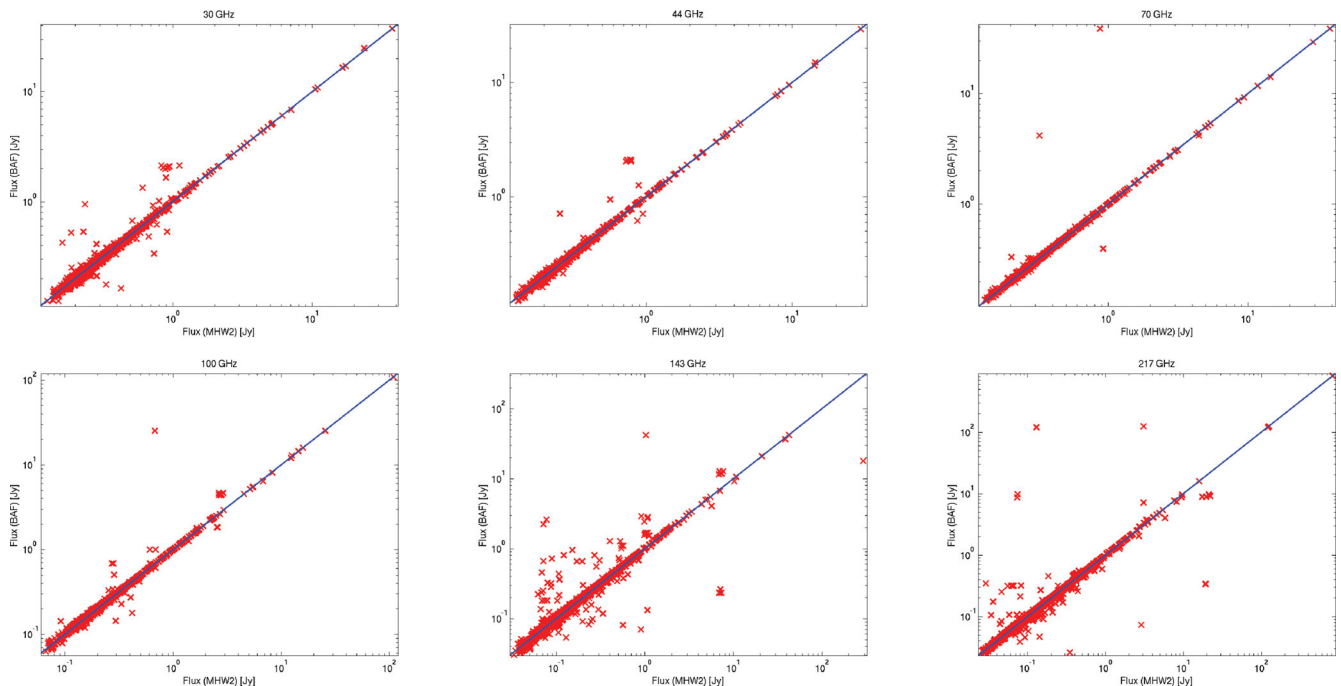

Figure 8. As in the previous figure, but applying a $\pm 3°$ three Galactic cuts to show that the small bias found in the 143 and 217 GHz panels rapidly disappears when we exclude from the figure sources very close to the Galactic plane.

between 30 and 857 GHz, dividing each of the nine simulations into 1344 flat patches that effectively overlap to cover the 100 per cent of the maps. Then we apply our maximization techniques to obtain the pair of optimal values R and g that define the filter, taking into account the local statistics of the background in each particular patch. In addition, we make considerations in terms of Galactic latitudes since we know that the complexity of the background increases when one gets closer to the Galactic plane. In Fig. 5 one can see the results of this analysis with respect to the *index g* for the nine frequency bands. In this figure we show the average filter *index* and its dispersion per Galactic latitude bin, where the bins have been chosen to cover the same area in the sky.

For the 30 GHz case, the mean of the average filter *index* is ~ 7, except for the Galactic region between [−20 and 20], where the synchrotron radiation dominates and the *index g* decreases a bit (see upper left-hand panel of Fig. 5). Then, the behaviour of the filter *index* between 44 and 143 GHz is fairly flat, although it varies between 4 and 6. It is important to note that the behaviour of g starts to change again at 217 GHz and becomes more and more steep when we get closer to 857 GHz, extending each time to higher Galactic latitudes. This is a clear correlation between the

contribution of the dust emission and the far-infrared background to the background signal in the patches. It reflects how the filter is adapting to background complexity. This effect can be clearly seen in the lower panels of Fig. 5.

In order to do a qualitative comparison between the MHW2 and the BAF, we cross-match the catalogues of objects detected with each technique at each frequency and use the common objects to build several sets of figures.

First, in Fig. 6 we represent the relative difference between the noise level estimated for one method and the other as a function of Galactic latitude and frequency. This allows us to see the overall behaviour of the noise estimated with the new filter compared with the MHW2. In the upper panels one can see that the estimation of the noise in the patches filtered with the BAF is a few per cent smaller than that of the MHW2. When we increase the frequency one starts to see a change in the behaviour, as expected, and between 143 and 857 GHz the estimation of the noise of the BAF is up to a 25 per cent smaller than that of the MHW2, but only in the vicinity of the Galactic plane. Note that even though the BAF, by definition, includes the MHW2 and its performance should always be equal or better than that of the MHW2, in the 545-GHz panel
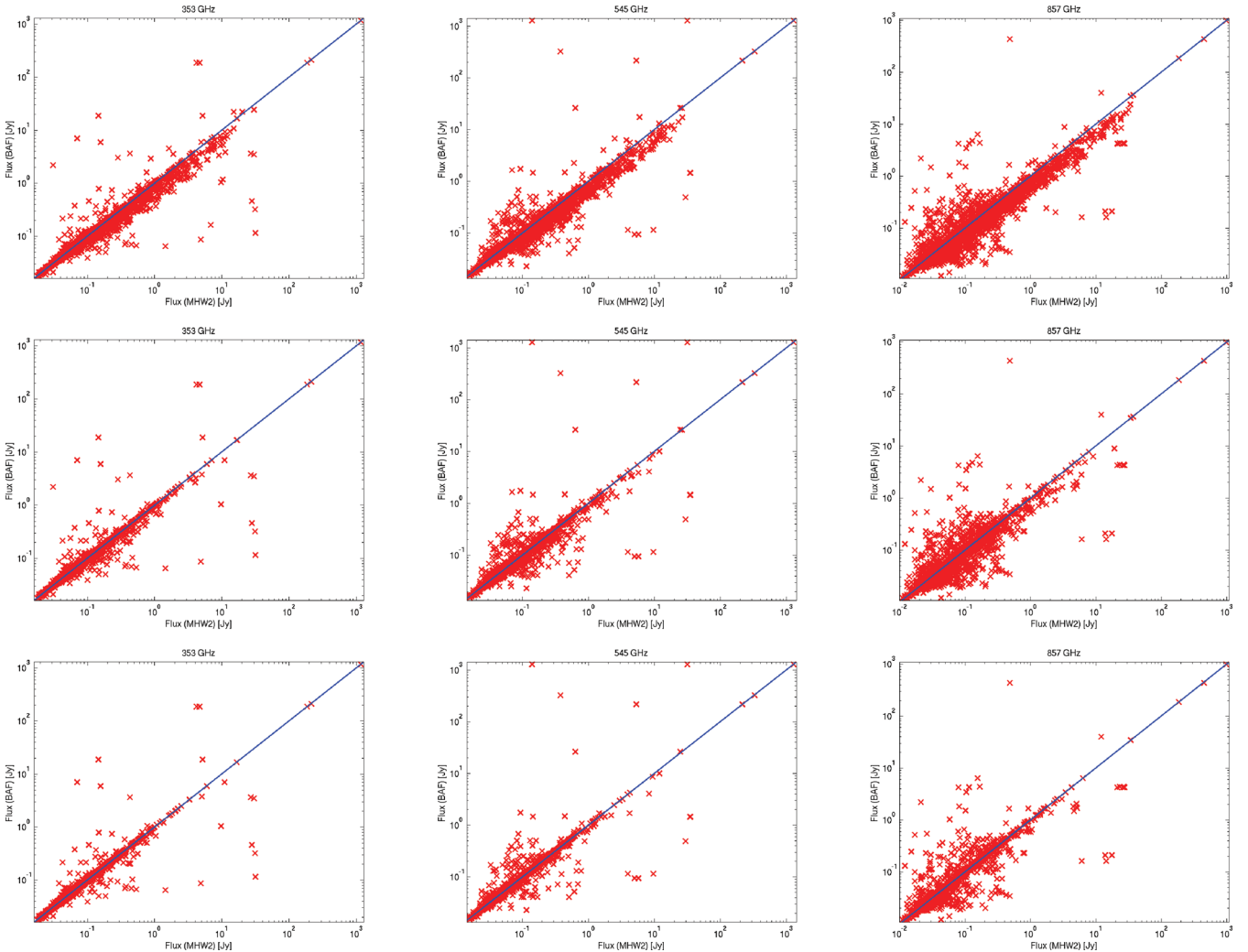

Figure 9. Comparison of the estimated flux density of the common sources detected by both filters for channels 353, 545 and 857 (left-, centre and right-hand panels) and applying three Galactic cuts (none, $\pm 3°$ and $\pm 15°$) to show that the bias found in the figures rapidly disappears when we exclude from the figure sources very close to the Galactic plane.

of Fig. 6 we see that for one of the bins, the relative difference is negative. This is not a problem of the filter but a rare artefact of the implementation of the algorithm that occurs due to the fact that we divide the sky into overlapping patches and, sometimes, the same source is detected with the same technique in two adjacent patches but with different signal-to-noise ratio (SNR). Since we keep the ones with highest SNR, it could happen that a source is detected with a higher SNR with the BAF in one patch and with the MHW2 in the adjacent one. In this case we cannot guarantee that

the noise estimation of the BAF is equal or higher than the MHW2 because they are in fact looking at slightly different regions of the sky. Secondly, with respect to the flux density estimation of the sources obtained with the BAF or the MHW2 techniques, in Fig. 7 one can see that between 30 and 217 GHz the fluxes obtained with both methods follow a clear one-to-one line. At 143 GHz one starts to see an increase in the dispersion and a small bias that is more obvious at higher frequencies. These differences mitigate when all the objects that lie within a $\pm 3°$ cut in latitude are excluded. These
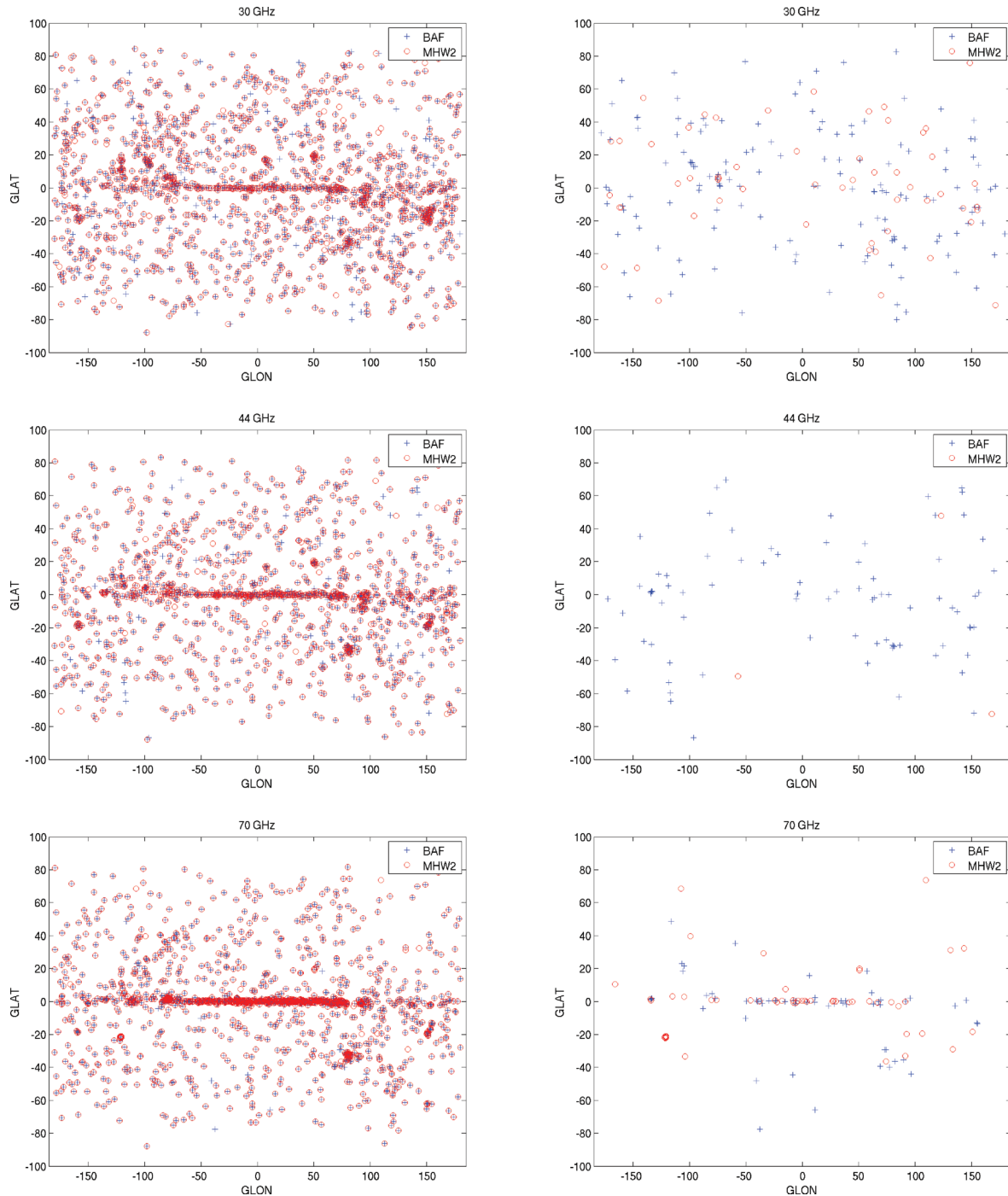

Figure 10. Position of the detected sources at 30 GHz (upper panels), 44 GHz (medium panels) and 70 GHz (lower panels) for the two considered techniques, MHW2 and BAF. In the left-hand panels we show the position in the sky of all the objects detected above SNR > 5. In the right-hand panels we show only those objects SNR > 5 that were detected by one method and not by the other, and vice versa.
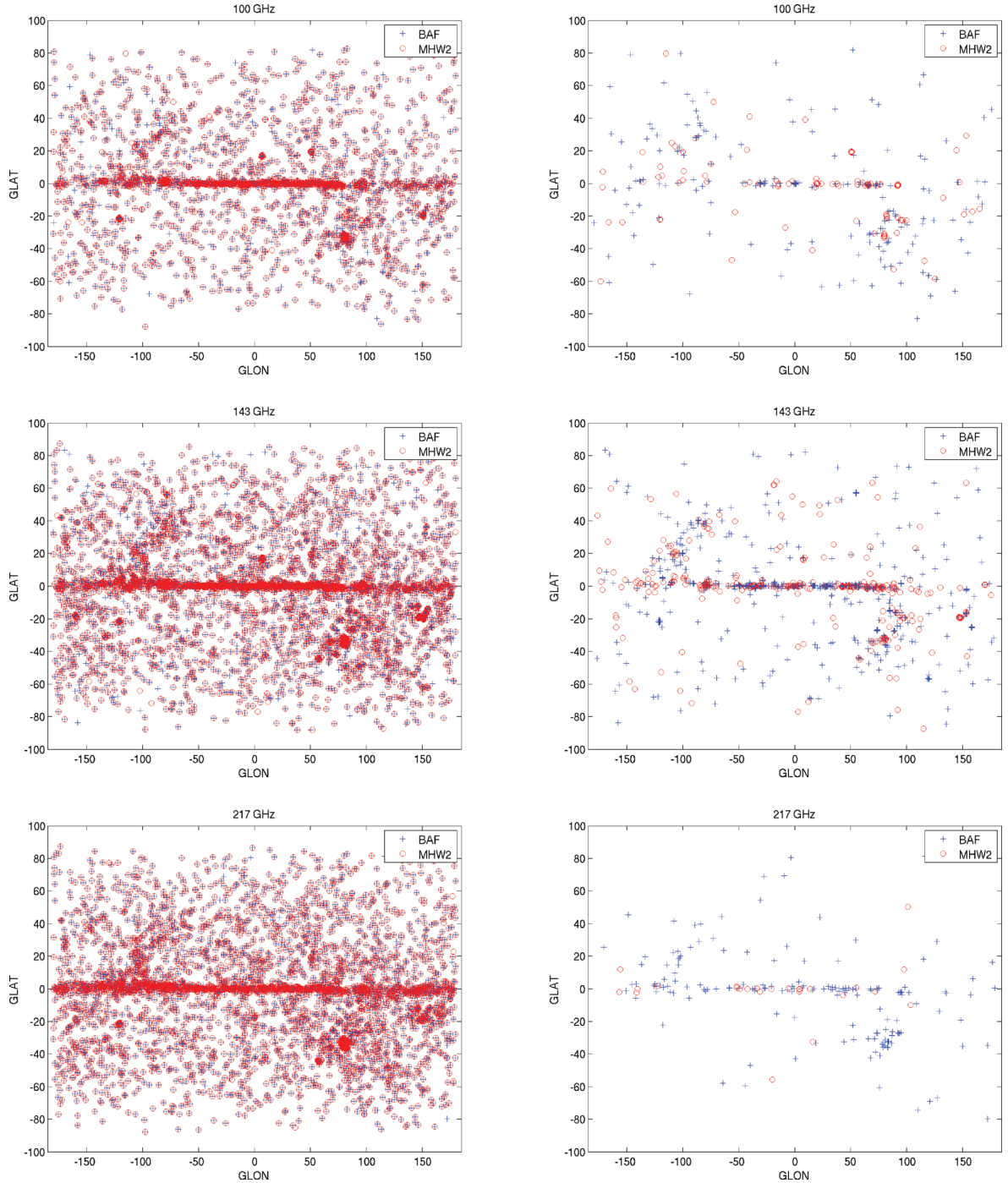
Figure 11. Position of the detected sources between 100 and 217 GHz for the two considered techniques, MHW2 and BAF. In the left-hand panels we show the position in the sky of all the objects detected above SNR > 5. In the right-hand panels we show only those objects SNR > 5 that were detected by one method and not by the other, and vice versa.

plots can be seen in Fig. 8, where most of the scatter and bias has now disappeared.

To further investigate this correlation between biased fluxes and extended Galactic regions at high frequencies, we have plotted the flux of the MHW2 versus that of the BAF for 353, 545 and 857 GHz using an additional Galactic cut of $\pm15°$. In Fig. 9 we show the fluxes without applying any Galactic cut (upper panels), applying a $\pm3°$ cut (middle panels) and applying a $\pm15°$ cut (lower panels). Again,

the scatter and bias found in the upper panels decreases significantly when increasing the Galactic cut, as we expected.

This result is telling us that both filters produce unbiased flux density estimations outside the Galactic plane and is highlighting a problem in its vicinity. We suspect that the problem may be related to any or both of the following effects. First, the fact that the extremely bright and spatially variable background signal is contributing to the recovered flux density of the point sources. In

other words, the filters are not able to fully remove the contribution from the background. Secondly, the filters are not able to recover unbiased flux density estimations of the extended sources present in the Galactic plane, because they have not been designed for this purpose and their response to these objects can be different. In order to test these ideas, we have performed additional simulations injecting 200 point sources with the same flux density and spatially distributed along the Galactic plane of the 545 GHz simulated map. Then we have attempted to estimate their flux densities using the BAF and MHW2 techniques and obtain that we are able to recover them with differences smaller than 1 per cent, i.e. much lower than the ones in Fig. 9. Therefore, we are confident that the reason for the bias seen at 217 GHz and above is due to the uneven performance of both filters when dealing with extended objects, a situation for which they have not been designed.

Moreover, to illustrate the performance of the new filter we will do a simple comparison with the MHW2 in terms of the number of detections above SNR > 5 obtained by each method when applied
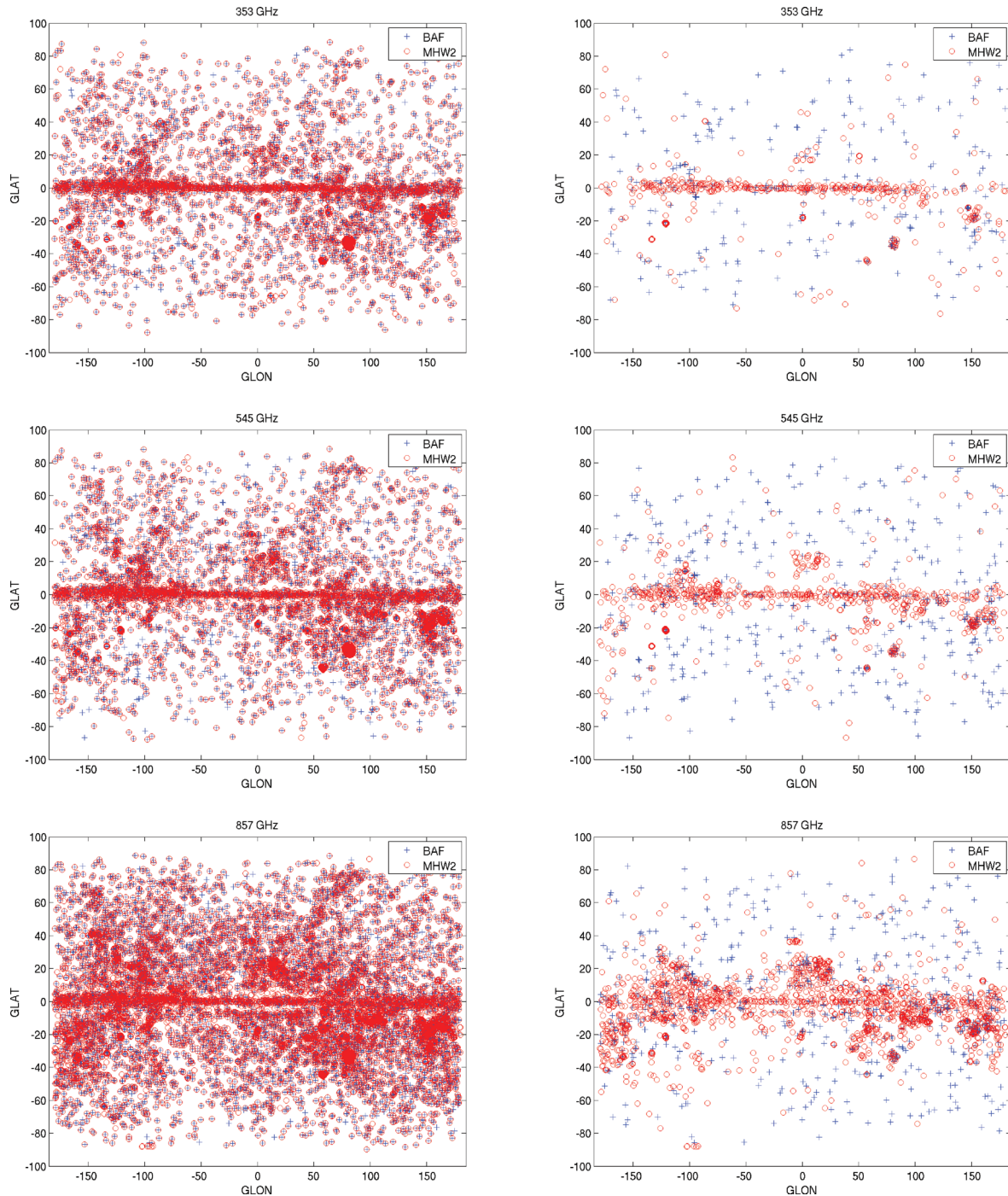
Figure 12. Position of the detected sources between 353 and 857 GHz for the two considered techniques, MHW2 and BAF. In the left-hand panels we show the position in the sky of all the objects detected above SNR > 5. In the right-hand panels we show only those objects SNR > 5 that were detected by one method and not by the other, and vice versa.

to the nine simulated maps. In the left- and right-hand panels of Figs 10–12 we show the positions in the sky of the sources detected by the BAF (left-hand panels) and those detected by the BAF or the MHW2 only (right-hand panels). In these figures one can see that the number of objects detected in the vicinity of the Galactic plane, as well as other complex regions such as the Large Magellanic Cloud (LMC), Orion, etc., is large and experience tell us that a fraction of those detections is not true sources but bright compact Galactic structure that looks like sources. Starting at 217 GHz, one can notice that the number of detections increases rapidly due to the change in the dominant population of sources, from non-thermal radio sources to thermal infrared sources. In addition, one can see that the BAF seems to detect more new sources all over the sky as opposed to the MHW2 that concentrates most of its new detections in the vicinity of the Galactic plane, many of which tend to be spurious detections caused by filaments and extended structures in the galaxy. This is an important result because it shows the potential of the BAF, a filter that better adapts itself to the local properties of the background, removing part of the noise and large-scale emission more effectively. This will imply a reduction in the number of detections in complex regions (that are likely to be spurious) while increasing the number of real detections in cleaner regions of the sky due to the improved estimation of the noise.

In Table 2, we show the number of detections with a SNR > 5 for both techniques when applied to the same regions of the sky, in one case covering the whole sky and in the other case applying a large Galactic cut of $\pm 30°$. One can see that the BAF detects, in general, more sources, especially between 30 and 143 GHz. In some channels the improvement is almost inexistent, for example at 70 GHz. Here, due to the properties of the background, the optimal value of g is essentially 4, which corresponds to the MHW2. On the contrary, in the upper frequency channels (353, 545 and 857 GHz) the number of SNR > 5 sources produced by the MHW2 is a bit larger than those of the BAF. Note that these results apply when considering the detection above SNR > 5 in the whole sky, including the Galactic plane where we know that the MHW2 is detecting more sources likely to be Galactic emission rather than extragalactic compact sources. If one looks at the second case, where a Galactic cut has been applied, the BAF detects more sources than the MHW2 in all cases except for one, 217 GHz, where the difference is very small. Note that these numbers are in agreement to what it was mentioned above. If we inspect the right-hand panels of Figs 10–12 one can see that the MHW2 tends to detect more objects in complex regions than the new filter, many of which are most likely not true sources.

Table 2. The number of detections above SNR > 5 that we have found applying the MHW2 and the BAF techniques at the nine simulated maps. In addition we show the number of detections above SNR > 5 and Galactic cut of $\pm 30°$ to have an idea of the effect of the Galactic emissions when doing this kind of simple comparisons.

ν	BAF	MHW2	BAF$_{30}$	MHW2$_{30}$
30	1400	1298	598	545
44	1082	1037	434	412
70	1175	1172	401	398
100	1975	1889	752	695
143	4001	3608	1531	1440
217	3753	3856	1557	1570
353	3495	3563	1313	1247
545	4325	4461	1582	1440
857	8460	9137	3429	3382

Table 3. We present the total number of detections above SNR > 5 obtained by one method and not by the other, as well as the number of detections inside and outside a mask that we have defined for each frequency and that includes the 15 per cent of the brightest pixels in the maps between 30 and 217 GHz and a more conservative 25 per cent of the brightest pixels between 353 and 857 GHz. We use these masks to define what can be considered as a complex region and what is not.

ν	MHW2			BAF		
(GHz)	All sky	Out	In	All sky	Out	In
30	55	33	22	146	111	35
44	3	3	0	84	49	35
70	54	23	31	60	18	42
100	99	45	54	201	125	76
143	287	104	183	731	279	452
217	25	6	19	167	88	79
353	448	92	356	312	175	137
545	720	138	582	530	323	207
857	1614	353	1261	882	391	491

This is telling us that the new filter is removing the background more efficiently which implies not only an improved estimation of the SNR but also a decrease in the number of spurious detections.

In addition, we have compared the catalogues of detections obtained with the BAF and the MHW2, searching for those objects detected only by one method and not by the other. We find that the BAF provides a higher number of unique detections in the maps from 30 to 217 GHz, whereas the MHW2 detects more objects at the highest frequencies (see Table 3). A deeper comparison is made when distinguishing between unique detections in highly contaminated regions (e.g. in the vicinity of the Galactic plane) and cleaner areas in the sky. Obviously, as mentioned above, the former will correspond (at high probability) with spurious detections associated with extended objects, whereas the latter will correspond most probably with true point sources. In order to identify the contaminated regions, we generate two types of masks. The first one includes the 15 per cent of the brightest pixels in each map. The second one, more conservative than the other, includes the 25 per cent of brightest pixels in each map. We use the less conservative 15 per cent masks between 30 and 217 GHz and the most conservative one between 353 and 857 GHz, where the Galactic dust emission extends to higher latitudes. In practice, these masks identify very well the Galactic emission and a few complex regions across the sky (e.g. Magellanic clouds, Orion, Ophiucus). In Table 3 one can also see that BAF detects more unique objects both inside and outside the mask between 30 and 217 GHz, whereas in the most contaminated channels at 353 GHz and above, it also detects more unique objects outside the mask while the MHW2 detects up to three times more unique objects inside the mask, many of which are most likely galactic extended emission rather than extragalactic point sources.

Finally, in order to give an idea of the spatial distribution in the sky of the values of the filter *index g*, the optimal *scale R* and the noise estimation σ_ω of the filtered maps that we have obtained, we have constructed a set of figures assigning at each HEALPix pixel (NSIDE = 8) the average value of g, R and σ_ω of the detected sources that fall into it (see Figs 13–15). It is interesting to note that, as expected, the values of the optimal *index g* decrease a bit in the galactic plane at 30 GHz and significantly increase in and around the galactic plane at 217 GHz and above, while they remain fairly homogeneous across the sky between 70 and 143 GHz. A similar
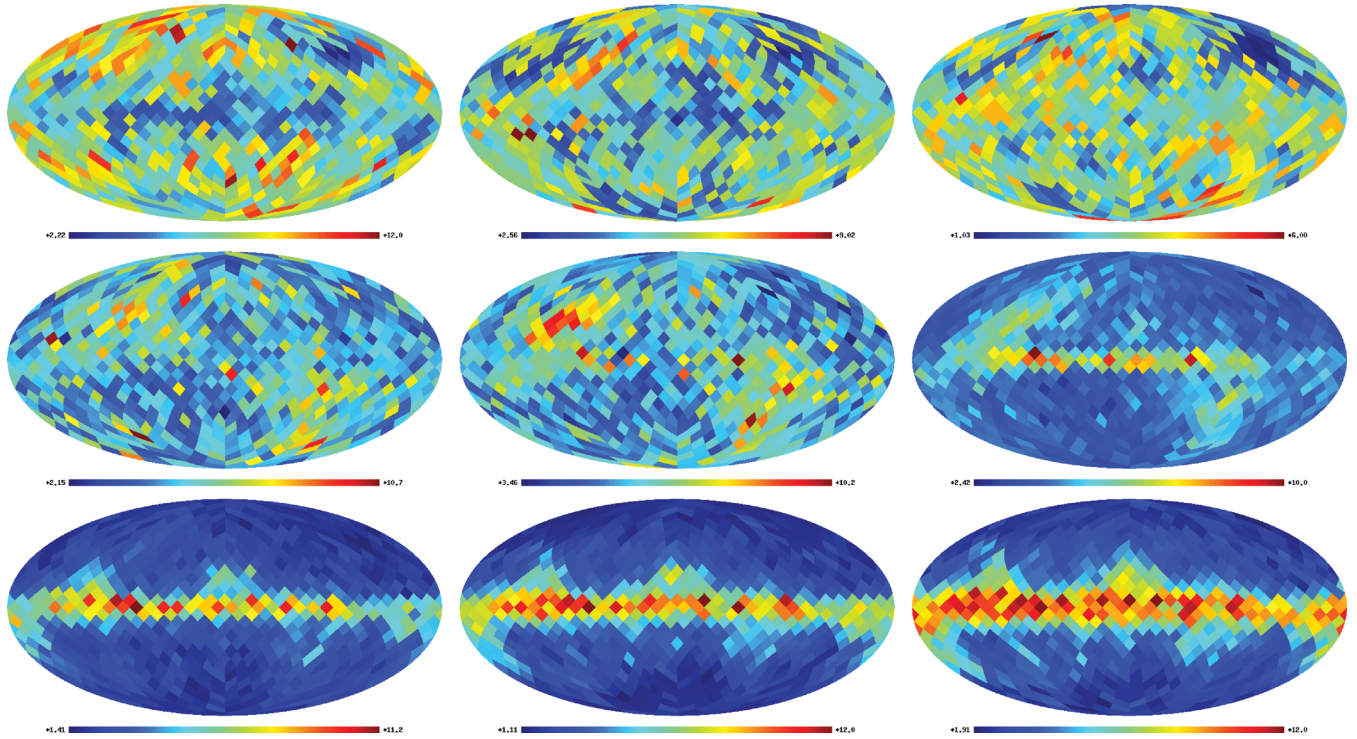
Figure 13. Spatial distribution across the sky of the values of the filter *index g* found during the analysis of nine simulations between 30 and 857 GHz. The upper panels show the 30, 44 and 70 GHz cases, the middle panels show the 100, 143 and 217 GHz cases and the lower panels show the 353, 545 ad 857 GHz cases.
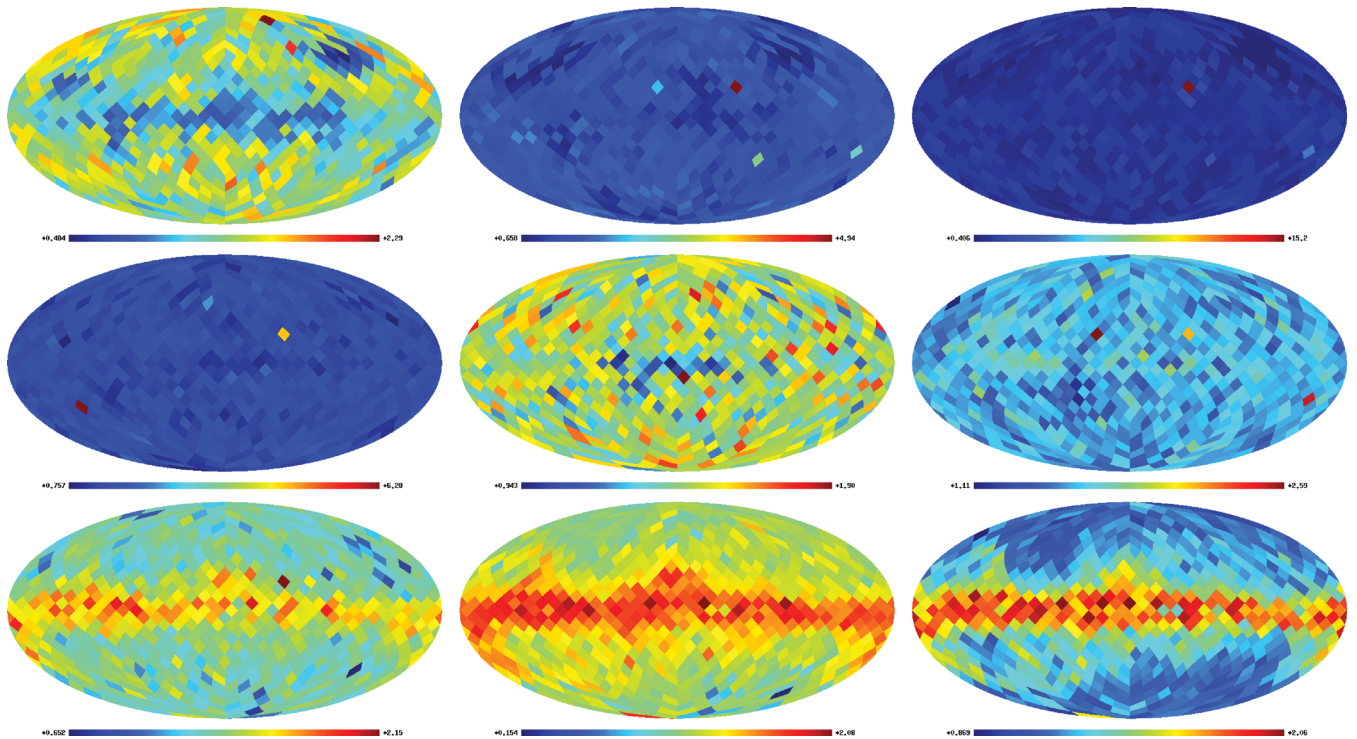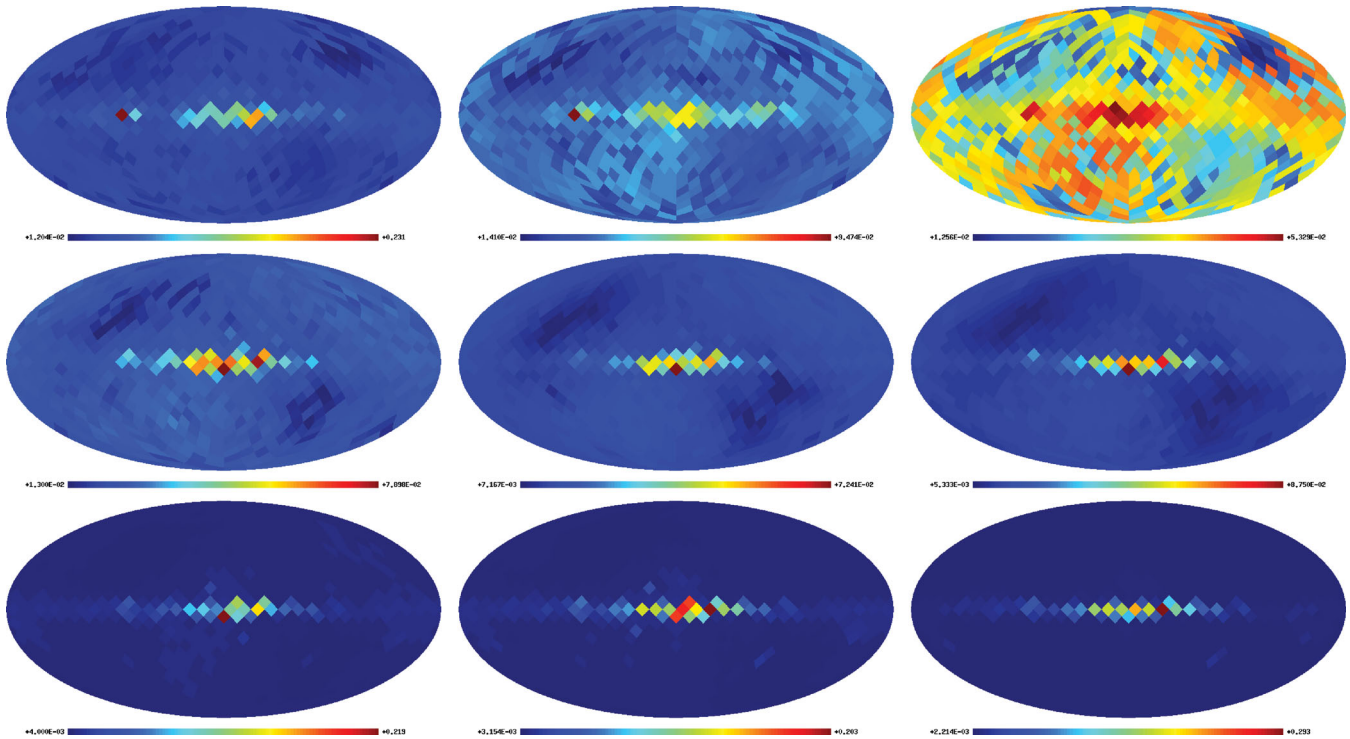

Figure 14. Spatial distribution across the sky of the values of the optimal *scale* of the filter $z = R/\sigma_{\mathrm{b}}$ found during the analysis of nine simulations between 30 and 857 GHz. The upper panels show the 30, 44 and 70 GHz cases, the middle panels show the 100, 143 and 217 GHz cases and the lower panels show the 353, 545 ad 857 GHz cases.

Figure 15. Spatial distribution across the sky of the estimated noise in the filtered maps σ_ω found during the analysis of nine simulations between 30 and 857 GHz. The upper panels show the 30, 44 and 70 GHz cases, the middle panels show the 100, 143 and 217 GHz cases and the lower panels show the 353, 545 and 857 GHz cases.

behaviour can be seen in the optimal *scale R* showing that the *scale* of the filter increases with increasing complexity of the background. In addition, one can see that the estimated noise in the filtered maps σ_ω follows the galaxy and a few other complex regions in the sky very well.

4 CONCLUSIONS

In this work we have developed and studied the performance of a new filter that maximizes the amplification of the compact sources embedded in a complex background using two free parameters, the *scale R* and the *index g* of the filter. This new filter is called BAF. To study the capabilities of this new filter we have used simulations of the microwave sky at the frequencies of *Planck* between 30 and 857 GHz.

In a first detailed analysis of three frequencies (30, 143 and 857 GHz) we have visually selected three regions of interest with increasing background complexity and obtained projected patches of $7.3 \times 7.3 \, \text{deg}^2$. We have applied the BAF to these regions and have demonstrated that we can always find a combination of R and g that maximizes the amplification of the sources in the filtered map with respect to the original map. In addition, we have found that it is important to optimize not only the *scale* of the filter but also the *index g* that adapts to scaling properties of the background, finding values of g in the range (1.8–12) for the nine discrete patches that we have studied.

In a second test, we have explored the performance of the BAF when doing a full-sky analysis. We divide the sky into 1344 over-lapping patches, $7.3 \times 7.3 \, \text{deg}^2$ each, that effectively cover the 100 per cent of the sky and apply our new filtering technique to each

one of them, detecting sources above SNR > 5. The results not only confirm what we found in the preliminary study of nine interesting regions but also show how the *index g* changes in a smooth and co-herent way when we move from low Galactic latitudes to high ones in the presence of strong Galactic emission such as the synchrotron radiation at 30 GHz or the dust emission and the far-infrared back-ground starting at 217 GHz and above. The most extreme cases are 353, 545 and 857 GHz, where the *index g* changes from very low values close to $g = 2$ for the highest Galactic latitudes up to $g = 12$ for the regions very close to the Galactic plane. Even more, in these channels that we could consider cleaner from Galactic emission, in the sense that the CMB is the dominant component of the back-ground, one can see that the behaviour of the *index* is fairly flat at all Galactic latitudes and takes values, on average, between $g = 4$ and 6. In particular, at 70 GHz, the *index* is always very close to $g = 4$, the *index* of the MHW2. In addition we have qualitatively compared the performance of the filter with that of the MHW2, in terms of the number of detections above SNR > 5 in two cases, in the whole sky and above a Galactic cut of $\pm 30°$. We find that in the first case, the BAF detects more sources than the MHW2 below 217 GHz and the MHW2 detects more source at 217 GHz and above, although from the inspection of the position of the sources in the sky one can see that most of the new detections at high frequency are very close to the Galactic plane and are likely to be spurious detections due to bright compact emission from the galaxy. In the second case, where the Galactic cut has been applied, we find the BAF detects more sources at all frequencies. Moreover, we have looked at the number of unique detections obtained by one method and by the other, inside and outside a Galactic mask, and concluded that, first, the BAF detects more unique objects at all bands inside and outside

the mask up to 217 GHz, and, secondly, that at 353 GHz and above the BAF detects less objects than the MHW2 inside the mask, where most of the detections are spurious, as mentioned above.

We have demonstrated that a tool to detect compact sources like the MHW2 is a very good compromise for the kind of backgrounds that one can find in microwave experiments, both at low and high Galactic latitudes, at all the *Planck* frequencies, but its performance can be improved if we use the BAF, a filter that explores the combination of the *scale R* and *index g* that best adapt to the profile of the sources and to the local properties of the background maximizing the amplification and the SNR of the detections.

ACKNOWLEDGMENTS

REFERENCES

Argüeso F., González-Nuevo J., Toffolatti L., 2003, ApJ, 598, 86
Argüeso F., Sanz J. L., Herranz D., López-Caniego M., González-Nuevo J., 2009, MNRAS, 395, 649
Argüeso F., Salerno E., Herranz D., Sanz J. L., Kuruoğlu E. E., Kayabol K., 2011, MNRAS, 414, 410
Beichman C. A., Neugebauer G., Habing H. J., Clegg P. E., Chester T. J., eds, 1988, IRAS Catalog and Atlases, Explanatory Supplement. GPO, Washington
Bennett C. L. et al., 2003, ApJ, 583, 1
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Carvalho P., Rocha G., Hobson M. P., 2009, MNRAS, 393, 681
Cayón L. et al., 2000, MNRAS, 315, 757
Chen X., Wright E. L., 2008, ApJ, 681, 747
Chen X., Wright E. L., 2009, ApJ, 694, 222
Condon J. J., Cotton W. D., Greisen E. W., Yin Q. F., Perley R. A., Taylor G. B., Broderick J. J., 1998, AJ, 115, 1693
Curto A., Martínez-González E., Barreiro R. B., 2009, ApJ, 703, 399
de Zotti G., Ricci R., Mesa D., Silva L., Mazzotta P., Toffolatti L., González-Nuevo J., 2005, A&A, 431, 893
Dickinson C., Davies R. D., Davis R. J., 2003, MNRAS, 341, 369
Finkbeiner D. P., Davis M., Schlegel D. J., 1999, ApJ, 524, 867
González-Nuevo J., Argüeso F., López-Caniego M., Toffolatti L., Sanz J. L., Vielva P., Herranz D., 2006, MNRAS, 369, 1603
González-Nuevo J., Massardi M., Argüeso F., Herranz D., Toffolatti L., Sanz J. L., López-Caniego M., de Zotti G., 2008, MNRAS, 384, 711
Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reimecke M., Bortclmann M., 2005, ApJ, 622, 759
Gregory P. C., Scott W. K., Douglas K., Condon J. J., 1996, ApJS, 103, 427
Griffith M. R., Wright A. E., Burke B. F., Ekers R. D., 1995, ApJS, 97, 347
Haslam C. G. T., Salter C. J., Stoffel H., Wilson W. E., 1982, A&A, 47, 1
Herranz D., Vielva P., 2010, IEEE Signal Processing Magazine, 27, 67
Hinshaw G. et al., 2007, ApJS, 170, 228
Högbom J. A., 1974, A&AS, 15, 417
Komatsu E. et al., 2003, ApJS, 148, 119
Leach S. M. et al., 2008, A&A, 491, 597
López-Caniego M., Herranz D., Barreiro R. B., Sanz J. L., 2005, MNRAS, 359, 993
López-Caniego M., Herranz D., González-Nuevo J., Sanz J. L., Barreiro R. B., Vielva P., Argüeso F., Toffolatti L., 2006, MNRAS, 370, 2047
López-Caniego M., González-Nuevo J., Herranz D., Massardi M., Sanz J. L., de Zotti G., Toffolatti L., Argüeso F., 2007, ApJS, 170, 108
López-Caniego M., Massardi M., González-Nuevo J., Lanz L., Herranz D., De Zotti G., Sanz J. L., Argüeso F., 2009, ApJ, 705, 868
Mauch T., Murphy T., Buttery H. J., Curran J., Hunstead R. W., Piestrzynski B., Robertson J. G., Sadler E. M., 2003, MNRAS, 342, 1117
Sanz J. L., Herranz D., Martínez-Gónzalez E., 2001, ApJ, 552, 484
Tauber J. A. et al., 2010, A&A, 520, A1
Tegmark M., 1997, ApJ, 501, 1
Tegmark M., de Oliveira-Costa A., 1998, ApJ, 500, L83
Toffolatti L., Negrello M., González-Nuevo J., de Zotti G., Silva L., Granato G. L., Argüeso F., 2005, A&A, 438, 475
Tucci M., Martínez-González E., Vielva P., Delabrouille J., 2005, MNRAS, 360, 935
Tucci M., Toffolatti L., de Zotti G., Martínez-González E., 2011, A&A, 533, A57
Vielva P., Martínez-González E., Cayón L., Diego J. M., Sanz J. L., Toffolatti L., 2001, MNRAS, 326, 181
Wright E. L. et al., 2009, ApJS, 180, 283

This paper has been typeset from a TEX/LATEX file prepared by the author.