J Math Chem (2013) 51:1153–1163
DOI 10.1007/s10910-012-0078-7

ORIGINAL PAPER

Non-linear parallel solver for detecting point sources in CMB maps using Bayesian techniques

P. Alonso · F. Argüeso · R. Cortina · J. Ranilla · A. M. Vidal

Received: 1 August 2012 / Accepted: 2 September 2012 / Published online: 16 October 2012

Abstract In this work we present a suitable computational tool to deal with large matrices and solve systems of non-linear equations. This technique is applied to a very interesting problem: the detection and flux estimation of point sources in Cosmic Microwave Background (CMB) maps, which allows a good determination of CMB primordial fluctuations and leads to a better knowledge of the chemistry at the early stages of the Universe. The method uses previous information about the statistical properties of the sources, so that this knowledge is incorporated in a Bayesian scheme. Simulations show that our approach allows the detection of more sources than previous non-Bayesian techniques, with a small computation time.

Keywords Cosmic microwave background · Bayesian · Non-linear systems · Efficiency

P. Alonso · F. Argüeso
Department of Mathematics, University of Oviedo, 33203 Gijón, Spain
e-mail: palonso@uniovi.es

F. Argüeso
e-mail: argueso@uniovi.es

R. Cortina · J. Ranilla (✉)
Department of Computer Science, University of Oviedo, 33203 Gijón, Spain
e-mail: ranilla@uniovi.es

R. Cortina
e-mail: raquel@uniovi.es

A. M. Vidal
Department of Computer Systems and Computation, Universitat Politècnica de València, 46022 Valencia, Spain
e-mail: avidal@dsic.upv.es

1 Introduction

The Cosmic Microwave Background (CMB) is a diffuse radiation which started to propagate freely in the Early Universe, about 380,000 years after the Big Bang. The CMB is thus a fossil radiation which carries very important information about the fundamental properties of the Universe and, in particular, about the chemistry of the Early Universe. Since its discovery in 1964 [13], the CMB has been detected and surveyed by instruments aboard balloons and satellites, such as the NASA satellites COBE in 1992 [15] and WMAP in 2003 [16].

The ESA satellite Planck was launched in 2009 and has been measuring the CMB fluctuations with unprecedented accuracy, resolution and frequency coverage. The first data obtained by Planck were published in 2011 [5].

There has been a great interest (see [7,11,14]) in studying the effect of primordial chemistry on the CMB fluctuations. Line absorption, photoionozation and photodissociation leave their imprint on the CMB power spectrum. Determining with great accuracy the CMB fluctuations is very important to estimate the cosmological parameters and, from the point of view of chemistry, to put constraints on the proportions of different molecules in the Early Universe. Thereby, the chemistry at the early stages of the Universe can be studied through accurate knowledge of CMB fluctuations.

In order to extract all the relevant information from the CMB data, one must remove the contamination produced by extragalactic sources, e.g. galaxies that also emit microwave radiation which has a non-cosmological origin. Several techniques: matched filters, wavelets, etc. have been proposed in the literature (see [9] and [10] for detailed reviews) to deal with the detection problem.

Recently, Bayesian methods have been used to increase the number of detected sources with a low percentage of spurious detections [3,4]. The advantage of these Bayesian techniques is that they incorporate statistical information about the flux distribution of the sources and their number and positions. We will put forward the method proposed in [3] in Sect. 2, this method requires the handling of large matrices, at least when large patches of the sky are studied, as is the case for the data collected by the Planck instruments, and the solution of non-linear systems. In Sect. 3, the computational approach is presented based on the methodology exhibited in [2], which allows us to perform the calculations in a very efficient way and with a small computation time. In Sect. 4, the approach is tested with simulations resembling the observational data collected by the Low Frequency Instrument (LFI) of the Planck satellite. We show that the Bayesian technique, with a suitable mathematical approach and high performance implementations, compares favorably with the classical model. Finally, the main conclusions are drawn in Sect. 5.

2 Description of the problem

In a region of the sky, we assume to have an unknown number of point-like sources, whose emission is mixed with that of the CMB, $f(x, y)$. A model for the total emission as function of the position (x, y) is given by

$$\tilde{d}(x, y) = f(x, y) + \sum_{\alpha=1}^{n} a_\alpha \delta(x - x_\alpha, y - y_\alpha) \tag{1}$$

where $\delta(x, y)$ is the 2D Dirac delta function, the pairs are the locations of the point sources in our region of the celestial sphere, and a_α are their fluxes. We observe this radiation through an instrument, with beam pattern $b(x, y)$, and a sensor that adds a random noise $n(x, y)$ to the signal measured.

Therefore, the output of our instrument is:

$$d(x, y) = \sum_{\alpha=1}^{n} a_\alpha b(x - x_\alpha, y - y_\alpha) + (f * b)(x, y) + n(x, y) \tag{2}$$

where the point sources and the CMB have been convolved with the beam. In our application, we are interested in extracting the locations and the fluxes of the point sources and consider the rest of the signal as just a disturbance superimposed to the useful signal. If $c(x, y)$ is the signal which does not come from the point sources, model (2) becomes

$$d(x, y) = \sum_{\alpha=1}^{n} a_\alpha b(x - x_\alpha, y - y_\alpha) + c(x, y). \tag{3}$$

If our data set is a discrete map of N pixels, the above equation can easily be rewritten in vector form, by letting d be the lexicographically ordered version of the discrete map $d(x, y)$, a be the n-vector containing the positive source intensities a_α, c the lexicographically ordered version of the discrete map, $c(x, y)$, and ϕ be an $N \times n$ matrix whose columns are the lexicographically ordered versions of n replicas of the map $b(x, y)$, each shifted on one of the source locations. Equation (3) thus becomes

$$d = \phi a + c. \tag{4}$$

Looking at Eqs. (3) and (4), we see that our unknowns are the number n, the list of locations (x_α, y_α), with $\alpha = 1, \ldots, n$ and the vector a. It is apparent that, once n and (x_α, y_α) are known, matrix ϕ is perfectly determined. Let us then denote the list of source locations by the $n \times 2$ matrix R, containing all their coordinates.

If we want to adopt a Bayesian approach, we must write the posterior probability density of our unknowns. According to Bayes theorem, this posterior probability can be written as

$$p(n, R, a|d) \propto p(d|n, R, a) p(n, R, a) \tag{5}$$

where $p(d|n, R, a)$ is the likelihood function, derived from our data model.

For the CMB plus noise we can assume that c is a Gaussian random field with zero mean and known covariance ξ. Thus, the likelihood function is

$$p(d|n, R, a) \propto \exp\Big(-(d - \phi a)^t \xi^{-1} (d - \phi a)/2\Big). \tag{6}$$

To find the prior density $p(n, R, a)$ we need to make a number of assumptions:

(1) Both R and a depend on n through the number of their elements. On the other hand, fluxes and positions are, in principle, independent. Thus, we can write

$$p(n, R, a) = p(R, a|n)p(n) = p(R|n)p(a|n)p(n). \tag{7}$$

(2) A priori, it is reasonable to assume that all the possible combinations of locations occur with the same probability. Therefore

$$p(R|n) = \frac{n!(N-n)!}{N!}, \tag{8}$$

since $N!/(n!(N-n)!)$ is the number of possible distinct lists of n locations in a discrete N-pixel map.

(3) It has been checked (see [3]) that the flux distribution can be modeled by a Generalized Cauchy Distribution

$$p(a|n) \propto \prod_{\alpha=1}^{n} \left[1 + \left(\frac{a_\alpha}{a_0}\right)^p\right]^{-\frac{\gamma}{p}}, \tag{9}$$

with γ and p positive numbers. This distribution obviously assumes that the fluxes of the different sources are mutually independent. In order to work with non-dimensional magnitudes, we define $x_\alpha = a_\alpha/a_0$; we also assume that we will detect point sources above a minimal flux a_i, that leads to the following normalized distribution

$$p(x|n) = \frac{p}{B\left(\frac{1}{1+x_i^p}; \frac{\gamma-1}{p}, \frac{1}{p}\right)} \prod_{\alpha=1}^{n} \left(1 + x_\alpha^p\right)^{-\frac{\gamma}{p}}, \tag{10}$$

where B is the incomplete beta function. The values of a_0, p and γ, can be determined by fitting this formula to the point source distribution given by the De Zotti counts model (see [6]).

(4) We assume that the number of sources in a given sky patch follows a Poisson distribution, with a known average number of sources λ

$$p(n) = \frac{\lambda^n e^{-\lambda}}{n!}. \tag{11}$$

A detailed account of all these assumptions can be seen in [3].

Finally, by putting together all these prior distributions and the likelihood, we can write the negative log-posterior

$$
\begin{aligned}
L(n, R, x) = \frac{1}{2}\left(x^t M x - 2e^t x\right) - \log(N-n)! - n\log(\lambda) \\
-n\log(p) + n\log B\left(\frac{1}{1+x_i^p}; \frac{\gamma-1}{p}, \frac{1}{p}\right) \\
+\frac{\gamma}{p}\sum_{\alpha=1}^{n}\log\left(1+x_\alpha^p\right), \qquad (12)
\end{aligned}
$$

with $M = a_0^2\phi^t\xi^{-1}\phi$ and $e = a_0\phi^t\xi^{-1}d$. We assume that we know a_0, p, x_i, γ and λ, so that the unknowns are: the normalized fluxes x, the number of point sources n and the positions of the point sources through the matrix $\phi(R)$.

3 Computational approach

In [2] a suitable algorithm for a non-Bayesian method with a high degree of parallelism that improves, from the computational point of view, the classical approaches for detecting point sources in CMB maps was presented.

In [2] matrix M and vector e are obtained using Cholesky decomposition, without calculating the inverse of ξ, given that $\xi \in \mathbb{R}^{N\times N}$ is a symmetric positive definite matrix. Besides, [2] shows that ϕ and ξ are symmetric Toeplitz-block Toeplitz matrices with symmetric blocks (see [18]) and suggests, but does not exploit this property, the possibility of using *fast-algorithms* (those which use the special structure of the matrix to dramatically reduce the execution time).

Now, we put forward a generalization of the algorithm presented in [2] that can be applied to the Bayesian method explained above. Also, this generalization explicitly handles the structure of the matrices using the novel fast-algorithms described in [1] and included in *StructPack* [17]. Thereby, we can deal with large matrices and solve large systems in an efficient way.

The goal of this work is to find the number of sources and their fluxes and positions by maximizing the posterior probability distribution, or equivalently and more simply, minimizing the negative log-posterior. In conclusion, we search the number of sources, their fluxes and positions, rendered more probable by the data, taking also into account the prior distributions.

Therefore, regarding the flux we minimize the negative log-posterior with respect to x, by taking the derivative and equating to zero, we obtain

$$
\sum_{\beta=1}^{n} M_{\alpha\beta}x_\beta - e_\alpha + \frac{\gamma x_\alpha^{p-1}}{1+x_\alpha^p} = 0, \quad \alpha = 1, 2, \ldots, n. \qquad (13)
$$

In order to fix the positions, we assume that the point sources are in the local maxima of e (see [3]). To determine the number of sources, we sort these local peaks from top to bottom and solve (13) successively adding a new source. At the same time, we calculate (12) and select the number of sources which produces its minimum value.

For the sake of simplicity, we redefine $\phi \equiv a_0 \phi$ and consider $p = 1$, this last assumption is justified by the flux distribution. Besides, taking into account that the thresholding process has been described in [2], in the following we will describe the algorithm without explaining that process again.

The system of non-linear equations we want to solve is

$$\sum_{\beta=1}^{n} M_{\alpha\beta} x_\beta - e_\alpha + \frac{\gamma}{1 + x_\alpha} = 0, \quad \alpha = 1, 2, \ldots, n, \tag{14}$$

with $M = \phi^t \xi^{-1} \phi$ and $e = \phi^t \xi^{-1} d$. Therefore, the non-linear system can be expressed in matrix form as

$$Mx = e - \gamma v, \tag{15}$$

with $v = (1/(1 + x_1), 1/(1 + x_2), \ldots, 1/(1 + x_n))^t$.

We consider that the matrices M, ϕ and ξ are of order N, while the vectors e and d have N rows. As $\xi \in \mathbb{R}^{N\times N}$ is a symmetric positive definite matrix, Cholesky decomposition can be used to obtain a lower triangular matrix such that: $\xi = LL^t$ [8]. Hence, vector e can be expressed as:

$$e = \phi^t \xi^{-1} d = \phi^t L^{-t} L^{-1} d = \phi^t L^{-t} c_1 = \phi^t c_2 \tag{16}$$

with $c_1 = L^{-1} d$ and $c_2 = L^{-t} c_1$.

Now, in order to construct the matrix M, we calculate

$$M = \phi^t \xi^{-1} \phi = \phi^t (LL^t)^{-1} \phi = (L^{-1}\phi)^t (L^{-1}\phi) = Z^t Z, \tag{17}$$

with $Z = L^{-1}\phi$.

Next, we compute the QR decomposition of $Z = QR$, with $Q \in \mathbb{R}^{N\times N}$, orthogonal, and $R \in \mathbb{R}^{N\times n}$, upper triangular, and the matrix M can be expressed as

$$M = Z^t Z = (QR)^t (QR) = R^t Q^t QR = R^t R. \tag{18}$$

In order to solve (15) we will use the classical Newton-Raphson method and Armijo's rule, obtaining a sequence

$$x^{(k+1)} = x^{(k)} - DF\left(x^{(k)}\right)^{-1} F\left(x^{(k)}\right), \tag{19}$$

where

$$F(x) = Mx - e + \gamma v, \;\; DF(x) = M - \gamma w, \tag{20}$$

with $v = (1/(1+x_1), 1/(1+x_2), \ldots, 1/(1+x_n))^t$ and $w = (1/(1+x_1)^2, 1/(1+x_2)^2, \ldots, 1/(1+x_n)^2)^t$.

We will use the solution of the linear system $Mx = e$ as our initial condition ($x^{(0)}$). Thus, vector $x^{(0)}$ can be computed by solving the triangular linear systems $R^t y = \phi c_2$ and $Rx = y$. These ideas can be summarized in Algorithm 1.

Algorithm 1 Solution based on Cholesky decomposition

Require: ϕ, ξ, d, with $\phi, \xi \in \mathbb{R}^{N\times N}$, $d \in \mathbb{R}^{N\times 1}$ **return** x

1: Compute ξ: $\xi = L * L^t$
2: Obtain e: $L * c_1 = d$, $L^t * c_2 = c_1$, $e = \phi^t * c_2$
3: Obtain M: $Z = L^{-1} * \phi$, $Z = Q * R$, $M = R^t * R$
4: Calculate $F(x) = (R^t * R) * x - \phi^t * c_2 + \gamma * v$; $v = \left(\frac{1}{1+x_1}, \frac{1}{1+x_2}, \ldots, \frac{1}{1+x_n}\right)^t$
5: Calculate $DF(x) = (R^t * R) - \gamma * w$; $w = \left(\frac{1}{(1+x_1)^2}, \frac{1}{(1+x_2)^2}, \ldots, \frac{1}{(1+x_n)^2}\right)^t$
6: Solve non-linear system:
7: Calculate $x^{(0)}$: $R^t * y = \phi * c_2$, $R * x = y$
8: Apply Newton-Raphson: $x^{(k+1)} = x^{(k)} - DF(x^{(k)})^{-1} * F(x^{(k)})$
9: Find the number of point sources that minimizes the negative log-posterior

As mentioned, ϕ and ξ are symmetric Toeplitz-block Toeplitz matrices with symmetric blocks and, therefore, Cholesky decomposition can be replaced by the approach proposed in [1]. For example, obtaining c_2 in step 2 can be implemented as *solve* $\xi c_2 = d$, by using *Subroutine dtspg_sv* from StructPack [17]. This affects to steps 1, 2 and 7, but also to the manner in which step 3 is computed because *QR* factorization is not necessary. Therefore, we have one algorithm with two approaches: the former based on Cholesky decomposition and the latter using StructPack.

The results shown in Sect. 4 demonstrate that the Bayesian method fixes the number of detected sources in a non-arbitrary way and that the new method detects more sources than [2]. Furthermore, results show that considering the structure of the matrix improves the execution time and reduces the memory consumption.

4 Simulations and results

The experimental results shown in this section were obtained on a board with two Intel® Xeon® E5420 processors with 4 cores and 6 Mb. of cache memory each, thus resulting in a total of 8 cores for the tests. We used Intel® Composer XE 12.1, Intel® MKL 10.3 and StructPack 1.2. When possible, the operations are carried out by all cores in parallel, by calling a threaded implementation of Intel® MKL 10.3 and StructPack 1.2; or by using an OpenMP parallel directives.

We have tested our algorithm by applying it to simulated CMB maps with the characteristics of the 30 GHz channel of the Planck satellite. Our maps include several components: point sources, simulated according to the De Zotti model [6], a CMB map generated by using the power spectrum which produces the best fit to the WMAP 5-year maps [12], and the instrumental noise. The CMB map and the point source map

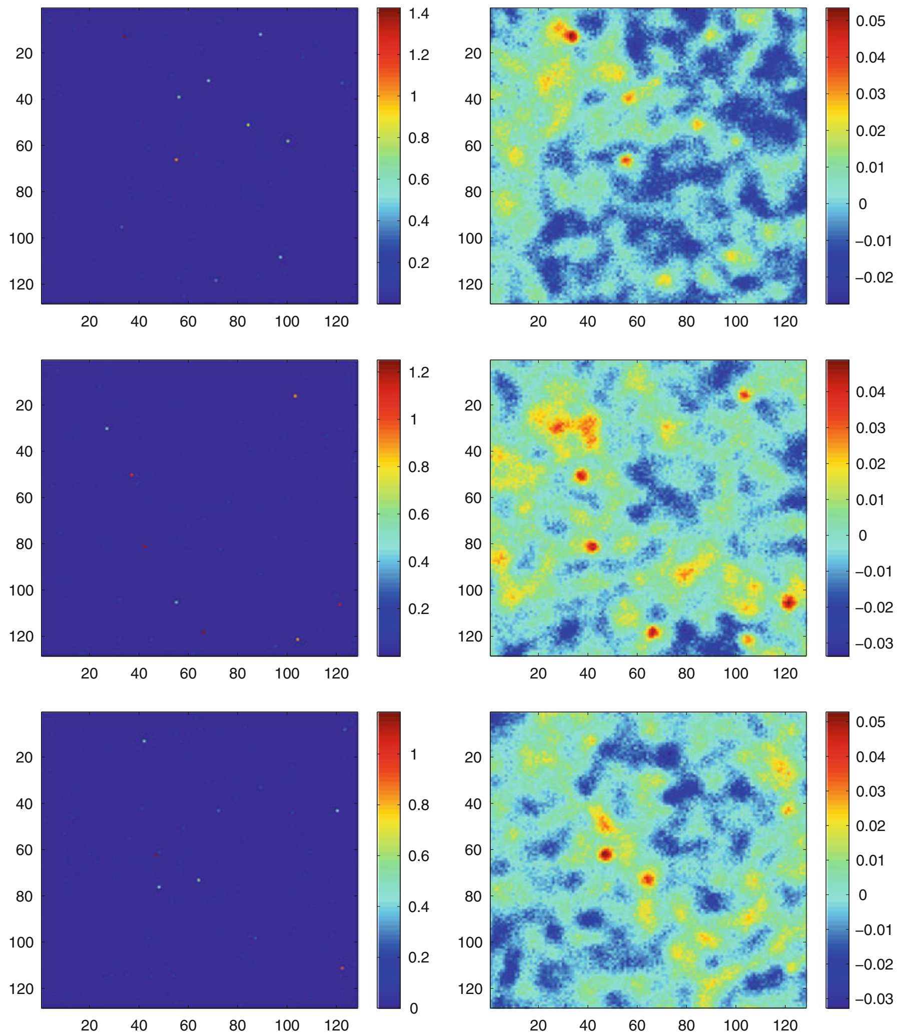

Fig. 1 Three simulations at 30 GHz. *Left panel* simulations of point sources. *Right panel* the same simulations including the CMB plus noise

are added and convolved with the instrument beam. Then, the instrumental noise is added to complete the final map.

Our simulations are flat patches of $N = 128 \times 128$ pixels, so that the size of the map is 14.66×14.66 square degrees. We have used our algorithm to determine the number of sources, n, their positions and to estimate their fluxes. We compare the results of the Bayesian method with the approach taken in [2] where a non-Bayesian method was applied and the peaks above a given threshold were selected as sources.

In the following, we comment on the results obtained for five simulations (three of them are shown in Fig. 1), that can be considered as representative of the general

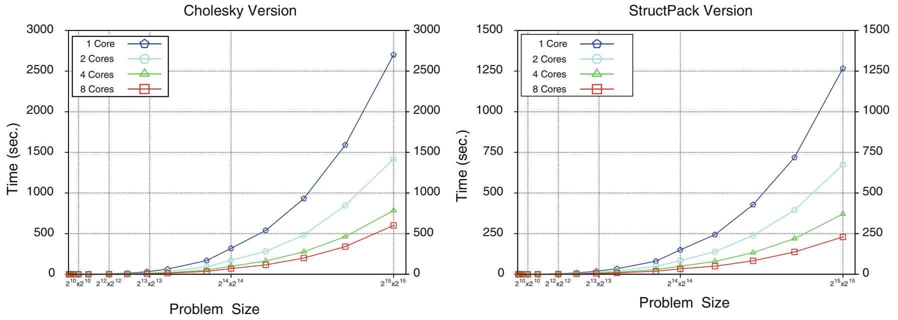

Fig. 2 Execution time with different number of cores and problem sizes. *Left* Cholesky. *Right* StructPack

performance of the method. By applying the algorithm put forward in Sect. 3, we are able to detect 37 sources. We have considered peaks above 2σ in the map e, although this particular threshold is not important, since the number of detected sources is given by the minimization of the negative log-posterior (Eq. 12). Only two of these sources are spurious, about 5 % of the total number, what agrees with the results found in [3]. A classical method, implemented by selecting the 5σ peaks of e, see [2], allows us to detect just 26 sources, all of them real. Therefore, there is about a 35 % increase in the number of real detections, at the cost of a small percentage of spurious sources. With these simulations of 14.66×14.66 square degrees, we have shown that the Bayesian method compares favorably with the classical approach, essentially a matched filter, for large patches too.

The sources are found in their real positions but for a few cases in which there is a displacement of one or two pixels and with small errors in their fluxes. The flux estimation is carried out by solving Equation 14. The average of the absolute value of the relative error in the flux estimation is 11 % in our simulations. If the fluxes are determined by solving a linear system as in [2], the errors are similar. However, note that all the sources above 0.45 Jy are detected with the Bayesian approach and by contrast, all the sources above 0.75 Jy are detected with the matched filter and a 5σ threshold. In conclusion, the source catalogue is more complete if it has been elaborated with the Bayesian technique.

From the computational point of view, the behavior of the approaches presented in Sect. 3 is shown in the next figures. As a result of time plotted in Fig. 2 it is easy to conclude that the solution based on StructPack is better than Cholesky approximation. For all sizes and number of cores StructPack's time is always less than 50 % of Cholesky's time. Nevertheless, both times are very good times when compared with those obtained with other solutions. For example, a problem of size $2^{15} \times 2^{15}$ is solved in less than a quarter of an hour using only 4 cores (less if 8 cores are used or we have an Intel® i7 or higher processor).

Additionally, memory consumptions are also lower with StructPack. Thus, if the problem size is $2^{15} \times 2^{15}$ StructPack uses about 12 GB of RAM, 4 GB less than Cholesky. In general, StructPack needs 25 % less memory than the Cholesky solution.

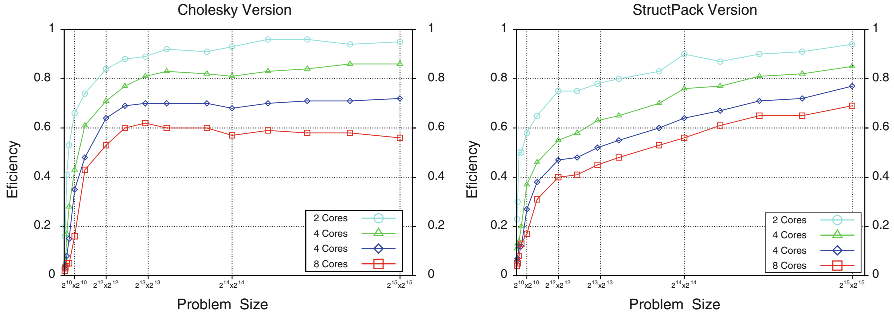

Fig. 3 Efficiency with different number of cores and problem sizes. *Left* Cholesky. *Right* StructPack

Finally, both solutions have a good efficiency (or speed-up) as seen in Fig. 3, which results in a reasonable scalability. That is, big problems (when the problem size increases) can be tackled with a reasonable increase in computational power.

5 Concluding remarks

We have put forward a new algorithm that can be used in a Bayesian approach to detect point sources in CMB maps. The advantages of the algorithm are the following ones:

(1) The number of detected sources is higher than in previous methods, with a low percentage of spurious detections. The minimization of the negative log-posterior allows the determination of the number of sources in a non-ambiguous way, without having to resort to fixing arbitrary thresholds.
(2) We can construct more complete source catalogs, what yields important information from the astrophysical and cosmological points of view.
(3) The algorithm can be extended to larger patches, since it handles large matrices and systems of non-linear equations in a very efficient way.
(4) We propose two software approaches. One of them, the version based on the StructPack package, is fast, scalable and needs little memory.
(5) Big problems, with large patches, can be now tackled at high resolution.

Acknowledgments This work was financially supported by the Spanish Ministerio de Ciencia e Innovación and by FEDER Projects (TEC2012–38142–C04, TIN2010-14971, TEC2009-13741 and CAPAP-H4 TIN2011–15734-E) Universitat Politècnica de València through Programa de Apoyo a la Investigación y Desarrollo (PAID-05-11) and Generalitat Valenciana through project PROMETEO/2009/013. We thank Diego Herranz for his help with the CMB simulations.

References

1. P. Alonso, D. Argüelles, J. Ranilla, A.M. Vidal, The solution of Block-Toeplitz linear systems of equations in multicore computers. J. Supercomput. (2012). doi:10.1007/s11227-012-0824-4
2. P. Alonso, F. Argüeso, R. Cortina, J. Ranilla, A.M. Vidal, Detecting point sources in CMB maps using an eficient parallel algorithm. J. Math. Chem. **50**, 410–420 (2012)
3. F. Argüeso, E. Salerno, D. Herranz, J.L. Sanz, E.E. Kuruoglu, K. Kayabol, A Bayesian technique for the detection of point sources in CMB maps. Mon. Not. Roy. Astron. Soc. **414**, 410–417 (2011)

4. P. Carvalho, G. Rocha, M. Hobson, A fast Bayesian approach to discrete object detection in astronomical images. PowellSnakes I. Mon. Not. Roy. Astron. Soc. **393**, 681 (2009)
5. Planck Collaboration, P.A.R. Ade et al. Planck early results I. The Planck mission. Astron. Astrophys. **536**, A1 (2011)
6. G. De Zotti et al., Predictions for high-frequency radio surveys of extragalactic sources. Astron. Astrophys. **431**, 893–903 (2005)
7. V.K. Dubrovich, Blurring of spatial microwave fluctuations by molecular last scattering. Astron. Lett. J. Astron. Space Astrophys. **19**, 53 (1993)
8. G.H. Golub, C.F. Van Loan, *Matrix Computations* (Johns Hopkins University Press, Baltimore, 1996)
9. D. Herranz, P. Vielva, Microwave backgound images. IEEE Signal Process. Mag. **27**, 67–75 (2010)
10. D. Herranz, F. Argüeso, P. Carvalho, Compact source detection in multichannel microwave surveys: from SZ clusters to polarized sources. Adv. Astron. (2012). doi:10.1155/2012/410965
11. R. Maoli, F. Melchiorri, D. Tosti, Molecules in the postrecombination universe and microwave background anisotropies. Astrophys. J. **425**, 372 (1994)
12. M.R. Nolta et al., Five-year wilkinson microwave anisotropy probe (WMAP) observations: angular power spectrum. Astrophys. J. Suppl. **180**, 296 (2009)
13. A.A. Penzias, R.W. Wilson, A measurement of excess antenna temperature at 4,080 Mc/s. Astrophys. J. **142**, 419–421 (1965)
14. D.R.G. Schleicher et al., Effects of primordial chemistry on the cosmic microwave background. Astron. Astrophys. **490**, 521 (2008)
15. G. Smoot et al., Structure in the COBE differential microwave radiometer first-year maps. Astrophys. J. **396**, L1–L5 (1992)
16. D.N. Spergel et al., First-year Wilkinson microwave anisotropy probe (WMAP) observations: determination of cosmological parameters. Astrophys. J. Suppl. **148**, 175–194 (2003)
17. StructPack: a high performance PACKage for STRUCTred Matrices. http://www.inco2.upv.es/structpack.html
18. M. Wax, T. Kailath, Efficient inversion of Toeplitz-block Toeplitz matrix. IEEE T. Acoust. Speech. **31**, 5 (1983)