

## SANEPIC: A MAPMAKING METHOD FOR TIME STREAM DATA FROM LARGE ARRAYS

G. PATANCHON,<sup>1,2</sup> P. A. R. ADE,<sup>3</sup> J. J. BOCK,<sup>4,5</sup> E. L. CHAPIN,<sup>1</sup> M. J. DEVLIN,<sup>6</sup> S. DICKER,<sup>6</sup> M. GRIFFIN,<sup>3</sup> J. O. GUNDERSEN,<sup>7</sup>  
M. HALPERN,<sup>1</sup> P. C. HARGRAVE,<sup>3</sup> D. H. HUGHES,<sup>8</sup> J. KLEIN,<sup>6</sup> G. MARSDEN,<sup>1</sup> P. G. MARTIN,<sup>9,10</sup> P. MAUSKOPF,<sup>3</sup>  
C. B. NETTERFIELD,<sup>10,11</sup> L. OLMI,<sup>12,13</sup> E. PASCALE,<sup>11</sup> M. REX,<sup>6</sup> D. SCOTT,<sup>1</sup> C. SEMISCH,<sup>6</sup> M. D. P. TRUCH,<sup>14</sup>  
C. TUCKER,<sup>3</sup> G. S. TUCKER,<sup>14</sup> M. P. VIERO,<sup>10</sup> AND D. V. WIEBE<sup>11</sup>

Received 2007 May 29; accepted 2007 November 27

### ABSTRACT

We describe a mapmaking method that we have developed for the Balloon-borne Large Aperture Submillimeter Telescope (BLAST) experiment, but which should have general application to data from other submillimeter arrays. Our method uses a maximum likelihood-based approach, with several approximations, which allows images to be constructed using large amounts of data with fairly modest computer memory and processing requirements. This new approach, Signal and Noise Estimation Procedure Including Correlations (SANEPIC), builds on several previous methods but focuses specifically on the regime where there are a large number of detectors sampling the same map of the sky, and explicitly allowing for the possibility of strong correlations between the detector time streams. We provide real and simulated examples of how well this method performs compared with more simplistic mapmakers based on filtering. We discuss two separate implementations of SANEPIC: a brute-force approach, in which the inverse pixel-pixel covariance matrix is computed, and an iterative approach, which is much more efficient for large maps. SANEPIC has been successfully used to produce maps using data from the 2005 BLAST flight.

*Subject headings:* balloons — methods: data analysis — submillimeter — techniques: image processing

### 1. INTRODUCTION

The problem of optimal “mapmaking” or “image reconstruction” is complex and multifaceted, with the basic procedures and even the terminology differing dramatically between different subfields of astronomy. The method adopted depends on the form in which the data are gathered and on the dominant source of systematic effects. From the optical to the near-IR one talks about combining “frames,” along with measurements of “darks” and “flat fields.” For the reduction of cosmic microwave background (CMB) data the now conventional method is to start from the principle that there is a linear algebra approach to solving the maximum likelihood problem. However, this has only been feasible up until now because of the limited number of detectors in the typical CMB experiment, and the fact that correlated signals among the detectors can be effectively ignored. Because of the rapid development of large bolometer arrays, the question that

arises is, how does one adapt the CMB approach to dealing with substantial numbers of detectors and where there are significant cross-correlations of noise between the detector time streams?

Data from the Balloon-borne Large Aperture Submillimeter Telescope (BLAST; Devlin et al. 2004) represent a new challenge for bolometric time stream-to-map algorithms. Recent CMB experiments that use detectors similar to those used on BLAST, such as BOOMERANG (Crill et al. 2003) and Archeops (Benoit et al. 2002), only use a handful of separate bolometers. Furthermore, these experiments’ off-axis designs lead to small correlations between detectors. Consequently, the correlations could be ignored at the mapmaking stage, and each detector time stream could be treated as an independent subset of the data. This has changed for BLAST, which has up to 139 detectors per band, with significant correlations induced by the on-axis design, as well as the higher frequencies of the observations. Just by itself, the large number of channels increases the impact of even small time stream correlations, the contribution from which does not integrate down with increasing number of detectors, unlike the uncorrelated noise. The high level of correlation (largely induced by temperature drifts in the obscuring secondary supports in BLAST) makes it important that the correlations be handled carefully in the mapmaking process.

This paper describes a Signal and Noise Estimation Procedure Including Correlations (SANEPIC) that has been developed for the analysis of BLAST. This algorithm will also have application to many next-generation experiments that will involve both noise correlations between channels (including correlations from the atmosphere) and very large numbers of detectors. This includes the next generation of larger format arrays for use in ground- and balloon-based instruments at microwave and millimeter wavelengths, which will have typically thousands of detectors.

This paper is arranged as follows. We next describe the pertinent aspects of BLAST. In § 3, the longest section, we set out our basic mapmaking method. Simulations we use for testing the method are described in § 4, with results presented in § 5,

<sup>1</sup> Department of Physics and Astronomy, University of British Columbia, Vancouver, BC V6T 1Z1, Canada.

<sup>2</sup> Laboratoire APC, 75205 Paris, France; patanchon@apc.univ-paris7.fr.

<sup>3</sup> Department of Physics and Astronomy, Cardiff University, Cardiff CF24 3AA, UK.

<sup>4</sup> Jet Propulsion Laboratory, Pasadena, CA 91109-8099.

<sup>5</sup> Observational Cosmology, MS 59-33, California Institute of Technology, Pasadena, CA 91125.

<sup>6</sup> Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104.

<sup>7</sup> Department of Physics, University of Miami, Coral Gables, FL 33146.

<sup>8</sup> Instituto Nacional de Astrofísica Óptica y Electrónica (INAOE), 72000 Puebla, Mexico.

<sup>9</sup> Canadian Institute for Theoretical Astrophysics, University of Toronto, Toronto, ON M5S 3H8, Canada.

<sup>10</sup> Department of Astronomy and Astrophysics, University of Toronto, Toronto, ON M5S 3H4, Canada.

<sup>11</sup> Department of Physics, University of Toronto, Toronto, ON M5S 1A7, Canada.

<sup>12</sup> Istituto di Radioastronomia, I-50125 Florence, Italy.

<sup>13</sup> Physics Department, University of Puerto Rico, Rio Piedras Campus, San Juan, PR 00931.

<sup>14</sup> Department of Physics, Brown University, Providence, RI 02912.

demonstrating the benefit of accounting for the correlated noise in BLAST-like observations. Finally, some of the maps obtained from the 2005 June BLAST flight are presented in § 6.

## 2. BLAST OBSERVATIONS OF THE SUBMILLIMETER SKY

The mapmaking procedure presented in this paper has been used to analyze the data from the BLAST 2005 flight. We use BLAST as a specific example for the application of SANEPIC throughout the paper. We describe BLAST in § 2.1 and then summarize the preprocessing of the data prior to mapmaking in § 2.2. In § 2.3 we derive a model of the data that we use for the mapmaking.

### 2.1. BLAST Instrument and Observations

BLAST incorporates a 2 m primary mirror and large-format bolometer arrays operating at 250, 350, and 500  $\mu\text{m}$ , designed to have 144, 96, and 48 bolometers, respectively (of which 139, 88, and 43, respectively, were used). The instrument is described in detail in Pascale et al. (2008). The low atmospheric opacity at the operating altitude of  $\sim 38$  km allows BLAST to map the sky very quickly compared to ground-based experiments and to conduct large-area shallow surveys as well as very deep surveys of the sky (Devlin et al. 2004). The BLAST wavelengths are near the peak of the spectral energy distribution of cold galactic dust, which gives BLAST the ability to conduct unique extragalactic and Galactic submillimeter surveys with high spatial resolution and sensitivity. BLAST thus enables studies of the distribution of very high redshift galaxies and of star-forming regions in our Galaxy.

The typical observing strategy consists of scanning the telescope back and forth in azimuth, covering the entire field by slowly varying the elevation. Cross linking of the data is assured by scanning the same field at another time of the day. Typical scanning strategies are given in Pascale et al. (2008).

The first scientific flight of BLAST took place in 2005 June from the Esrange Arctic base in Sweden to the Canadian Arctic. A total of  $\sim 100$  hr of data were taken in a variety of Galactic fields. They include a star-forming region (Vulpecula) over  $4 \text{ deg}^2$ , described in Chapin et al. (2008), three other fields of similar size in the Galactic plane (which will be the focus of future papers), an integration toward the ELAIS-N1 field (see Oliver et al. 2000), the Cas A supernova remnant over about  $0.5 \text{ deg}^2$  (P. C. Hargrave et al. 2008, in preparation), and several compact Galactic and extragalactic sources (Truch et al. 2008). Hereafter we refer to these as the BLAST05 data, to distinguish them from the data taken during the 2006 December Antarctic flight.

### 2.2. Time-ordered Data Preprocessing

The processing of BLAST data from detector time streams to the final map product involves several steps prior to mapmaking. Each of these steps is designed to remove a particular (or several) artifact(s) from the data and sometimes requires iterating, since some effects need to be removed simultaneously. In the following, we summarize the main processing stages leading to the time-ordered segments that are used as inputs for the mapmaking process.

We start by identifying events in the data that are sharply localized in time, such as spikes from cosmic-ray hits and other spurious sources. We use a method that allows us to discriminate between the different events depending on their signature in the data. Spikes that involve only a single sample are flagged, and the corrupted samples are replaced by the average value of the

samples in the vicinity. The data are deconvolved from the low-pass filter applied by the readout electronics. This filter has a frequency cutoff of approximately 35 Hz and is designed to avoid high-frequency noise aliasing. The deconvolution is performed in Fourier space. In addition, we have applied a low-pass cosine filter that limits the noise power from increasing at very high frequency (above 38 Hz) due to deconvolution. We have checked that the noise power spectrum is relatively flat after these deconvolution operations. Finally, cosmic-ray hits and other localized artifacts in the data time streams are detected and cut out. In order to avoid biasing our data products by having systematically more false event detections located where the sky is bright (e.g., when scanning a point source), we iterate this process: we make maps starting from data that have been cleaned using the process described above, subtract the maps from each original data set, and reprocess the data. The maps calculated at this intermediate stage are obtained by simply rebinning the data into pixels after strong high-pass filtering. The filter applied is a Butterworth filter with a frequency cutoff of 0.5 Hz, which is of the order of the knee frequency of the noise. Even if this operation suppresses most of the intermediate to large scales of the sky signal in the maps, it does not very much alter the signal from localized sources, at least to the level of accuracy needed at this stage, and it has the advantage of removing most of the stripes in the maps due to  $1/f$  noise. We have verified that the resulting bias for bright calibrators is less than 1%.

About 2% of the data from the BLAST05 flight were removed due to cosmic-ray events. Most of the events affected a single detector time stream, although some events affected a whole array at the same time. Detected spikes (from cosmic rays but also other spurious effects) are flagged over small time intervals of typically 1 s in the data. One second is too large an interval to simply ignore, so the corrupted data need to be replaced with random noise generated in a way that as much as possible preserves the statistical properties of the data. In the later mapmaking stage, which is partly performed in Fourier space, we assume continuity of the data. We could perform this gap filling with a constrained noise realization (see, e.g., Stompor et al. 2002). However, since the gap intervals are significantly smaller than the inverse of the knee frequency of the noise power spectrum (0.3 Hz), the noise can be well approximated by the sum of a white component plus a straight line of some slope across the gap. Specifically, we generate white noise in each gap with a standard deviation measured from the data in the vicinity of the gap and add a baseline with the parameters fitted using 20 samples on each side. The white-noise generation is for restoring as best as possible the stationarity of the data (generated samples are not reprojected to the map at the end).

After having filled the gaps in the time streams, we filter out very low frequency drifts that are poorly accounted for in the mapmaking procedure. A fifth-order polynomial is fitted to the data and removed from each data segment in order to reduce fluctuations on timescales larger than the length of the considered segment, which, depending on the specific case, varies from 30 minutes to a few hours. These fluctuations are poorly described because of the limited number of Fourier modes and would cause leakage at all timescales (for instance, a gradient in the time stream is described by a wide range of Fourier modes), degrading the efficiency of mapmaking. Note that we experimented with various polynomials and other effective high-pass filters; we found that the results were not very sensitive to precisely how this is done, but some such filtering is certainly required. The degree of the polynomial was chosen empirically as a compromise between suppressing the artifacts and keeping the large-scale signals in the final

maps. Using simulations, we have checked that the effect on the transfer function of the signal in the final map is weak. The resulting data segments are then corrected for the time-varying calibration (see Truch et al. 2008) using measurements of an internal calibration lamp (Pascale et al. 2008). Finally, the data segments are apodized at the edges over  $\sim 2000$  samples and are high-pass filtered at  $5 \times 10^{-3}$  Hz with a Butterworth filter. This filtering has very little effect on the final maps, since modes in the data at lower frequencies that are cut in this way are not expected to contribute significantly to the signals. We discuss the choices of filters further in § 3.2.

Accurate pointing reconstruction is a complicated procedure for balloon-borne telescopes, and this affects the mapmaking task through the pointing matrix (defined in § 2.3). The pointing reconstruction procedure is described in detail in Pascale et al. (2008). The next important step in reducing the BLAST data, which is calibration of the detectors, is detailed in Truch et al. (2008).

### 2.3. Model of the Data

Having performed the cleaning procedure described in the previous subsection, the resulting data time streams can be modeled very accurately as the sum of pure signal and pure noise contributions. The data for detector  $i$  observing at a given wavelength and at time sample  $t$  can be written as

$$d_{it} = [A_i]_{tp} s_p + n_{it}, \quad (1)$$

where  $p$  labels the pixels in the final map,  $A_i$  is the pointing matrix for bolometer  $i$  (whose elements, indexed with time  $t$  and pixel  $p$ , give the weight of the contribution of pixel  $p$  to the sample at time  $t$  for bolometer  $i$ ),  $s_p$  is the signal amplitude at pixel  $p$ , and the noise amplitude at time  $t$  for bolometer  $i$  is  $n_{it}$ . Summation over repeated indices is assumed here.

Ideally, the element  $[A_i]_{tp}$  of the pointing matrix is equal to the value of the beam response  $b[R(\mathbf{r} - \mathbf{r}_0)]$ , where  $\mathbf{r}$  points to the pixel  $p$  location,  $\mathbf{r}_0$  is the location of the beam center at time  $t$ , and  $R$  is a rotation matrix that depends on the rotation angle at time  $t$  between the telescope and sky coordinate systems. In principle, one could then recover a map of the sky *deconvolved* with the instrumental point-spread function (PSF). However, in practice this would be unacceptably noisy, as well as computationally intractable, because of the prohibitive volume of data. Although  $A_i$  might have mostly zero elements, it is nevertheless a huge matrix. It *may* be feasible to deconvolve a nontrivial beam response (e.g., like the BLAST05 beams, as shown in Truch et al. 2008) at the same time performing the mapmaking step, through an approximate treatment of the sparse pointing matrix. But we do not pursue that approach here.

In the simple case where the beam is symmetric, the mapmaking problem becomes tractable, provided that one restricts oneself to reconstructing a map of the instrument-convolved sky. We can then consider  $s_p$  in equation (1) as the map of the sky convolved with the beam, and consequently  $A_i$  indicates simply where the detector points in the sky at a given time. In this case,  $A_i$  is an ultimately sparse matrix with, in the BLAST case, simply a 1 in a single entry of each row. This approach, which has been conventional for CMB mapmaking (although with some adaptation for chopped data), is what we use in the following analysis. It gives no loss of information provided that the map pixels are sufficiently small, and one can simply assign all the flux from a bolometer's to the map pixel to which it points at each time interval. The requirement for accuracy is that the pixel size is smaller by a factor of 3 or more than the full width at half-maximum

(FWHM) of the instrumental PSF; in that case the additional convolution with the pixel shape gives negligible loss of angular resolution.

The noise term  $n_{it}$  in the model represents the sum of all contributions to the time streams that do not reproject on the sky. This will in general include instrumental noise, fluctuations in atmospheric emission and other loading, and unrecognized cosmic-ray hits. In general, some of those noise contributions will induce strong correlations between detector time streams. In this paper we adopt a very general model of the noise where the noise covariance matrix,

$$N_{ii'tt'} = \langle n_{it} n_{i't'}^t \rangle \quad (2)$$

(for bolometer indices  $i, i'$  and time indices  $t, t'$ ), has possibly nonzero elements even for  $i \neq i'$ . A key assumption, as we see later, is that the noise is Gaussian (so that  $N_{ii'tt'}$  is sufficient to describe all the statistical properties of the noise) and stationary (constraining  $N_{ii'tt'}$ ; see §§ 3.2.1 and 3.4).

In the specific case of BLAST05 observations, a very significant correlation of the noise is found in the time streams, and we have shown that a more constraining model provides a very good description of the data. An independent component analysis (Delabrouille et al. 2003) of the data enabled us to find that the noise and its correlations can be described to a high degree of accuracy by a noise component that is *not* correlated between detectors, together with a single *common-mode* component seen by all the detectors at a given wavelength (some correlations are also seen between detectors from different wavelengths, but we have chosen to treat each wavelength independently). The common part of the noise is instantaneous, meaning that the same common-mode noise is seen at the same time by all the detectors. In our model, the noise term in equation (1) is then decomposed as

$$n_{it} = \tilde{n}_{it} + \alpha_i c_t, \quad (3)$$

where the first term is the noise that is uncorrelated between detectors and the second term represents the common-mode component of the noise, rescaled by an amplitude parameter  $\alpha_i$ , which depends on the detector but not on time. This model can be generalized easily to deal with multiple noise components in time streams.

In the following section we present a method to reconstruct  $s_p$  given the data, in the framework of the linear model (eq. [1]) and in the presence of correlated noise between detector time streams.

## 3. MAPMAKING METHOD

### 3.1. Maximum Likelihood Mapmaking

The use of maximum likelihood mapmaking techniques has been developed by many authors for application to CMB data sets (Wright et al. 1996; Tegmark 1997; Borrill 1999; Prunet et al. 2000,<sup>15</sup> 2001; Tegmark et al. 2000; Ferreira & Jaffe 2000; Doré et al. 2001; Natoli et al. 2001; Dupac & Giard 2002; Stompor et al. 2002; Hinshaw et al. 2003; Yvon & Mayet 2005). Other approaches are more specific to destriping for *Planck*-like scanning strategies (Delabrouille 1998; Maino et al. 2002; Keihänen et al. 2005; de Gasperis et al. 2005; Macías-Pérez et al. 2007; Poutanen et al. 2006; Ashdown et al. 2007b). Since there are already a large number of publications on the topic, here we present only a very brief overview of the approach of maximum likelihood mapmaking techniques.

<sup>15</sup> Available at <http://moriond.in2p3.fr/J00/ProcMJ2000/>.

Assuming the simple linear model given by equation (1), the log-likelihood of the data can be calculated under the assumption that the noise is Gaussian and stationary. The solution is

$$\log L(d|s) = -\frac{1}{2}(d - As)^t N^{-1}(d - As), \quad (4)$$

where  $N \equiv \langle nm^t \rangle$  is the noise covariance matrix in the time domain (here and after we do not make explicit reference to time indices  $t$  and  $t'$ ;  $N$  means noise covariance matrix in the time domain) and superscript  $t$  denotes transpose. Maximizing the above equation with respect to the map parameters  $s$  (suppressing the pixel indices here for convenience) leads to the following well-known estimator:

$$\hat{s} = (A^t N^{-1} A)^{-1} A^t N^{-1} d. \quad (5)$$

The inverse pixel-pixel covariance matrix of the noise in the map is the term in brackets in this equation, i.e.,

$$N_{pp'}^{-1} = A^t N^{-1} A. \quad (6)$$

Computation of the solution to equation (5) is far from trivial for most astronomical applications, due to the large amount of data, and hence this poses a difficult numerical challenge. The noise covariance matrix in the time domain  $N$  is a very large matrix of size the number of samples squared, which could easily be millions, while  $N_{pp'}$  may be more reasonable in size but has no obvious symmetries and so is still difficult to invert. Nevertheless, we have implemented a method aimed at finding the maximum likelihood solution given by equation (5) when there are a large number of detectors and in the presence of possible correlations in the noise between different detector time streams.

### 3.2. Implementation

In the simple case of dealing only with independent noise between detectors, our matrix inversion method is very similar to the MADCAP method, described in Stompor et al. (2002). However, we have developed our new approach to deal efficiently with multidetector data in the presence of correlated noise between detectors (described in detail in § 3.4). In this section we summarize the basic ideas for the simpler one-detector, one-scan case.

In order to find the maximum likelihood solution of the map (eq. [5]), we have developed two different algorithms. They both allow us to solve the linear system  $N_{pp'} \hat{s} = x$ , with  $x \equiv A^t N^{-1} d$ . The first approach explicitly computes the inverse pixel-pixel covariance matrix  $N_{pp'}^{-1}$ , and we refer to this as the “brute-force algorithm.” The second approach uses iterations that converge to the maximum likelihood map without the need for computing  $N_{pp'}^{-1}$ , and we refer to this as the “iterative algorithm.” Both approaches require as a first step the computation of the inverse of the time-time noise covariance matrix  $N$ .

#### 3.2.1. Inverse Noise Covariance Matrix $N_{u'}^{-1}$

In practice, even when we have knowledge of the statistical properties of the data as described by the power spectrum  $P(\omega)$ , the brute-force inversion of  $N$  is not tractable because of its enormous size: for a single BLAST detector observing for 10 hr at a data rate of 100 Hz, the matrix has approximately  $10^{13}$  elements. However, if we make the approximation that each data segment is “circulant,” meaning that the beginning and the end of a segment are connected without discontinuity and that there are no gaps in the data, then the matrix  $N$  is also circulant (see

§ 3.6 for a description of how we treat gaps in the data). Circulant matrices are much easier to store and to invert. With this approximation the elements of the matrix can be written as

$$N_{u'} = C(|t - t'|), \quad (7)$$

where the correlation function  $C(|t - t'|)$  between samples  $t$  and  $t'$  depends only on the separation between the two samples. A circulant matrix has the property of having a diagonal matrix counterpart in Fourier space.

Let  $F$  be the discrete Fourier operator; we have

$$N = F^\dagger \Lambda F, \quad (8)$$

where superscript  $\dagger$  denotes transpose conjugation and  $\Lambda$  is a diagonal matrix whose diagonal is described by the power spectrum of the data segment,

$$\Lambda_{\omega\omega} = P(\omega). \quad (9)$$

The inverse of the noise covariance matrix is

$$N^{-1} = F^\dagger \Lambda^{-1} F, \quad (10)$$

and because  $\Lambda^{-1}$  is a diagonal matrix,  $N^{-1}$  is also circulant, so that knowledge of only one row is enough to describe the entire matrix. Then the inverse covariance matrix can be written as

$$[N^{-1}]_{u'} = \bar{C}(|t - t'|), \quad (11)$$

with

$$\bar{C}(\Delta t) = \mathcal{F}^{-1} \left[ \frac{1}{P(\omega)} \right] (\Delta t), \quad (12)$$

where  $\mathcal{F}^{-1}$  represents the inverse Fourier transform. The inverse of the covariance matrix can then be computed directly using the power spectrum of the data. This is a very fast operation [ $O(n_s \log n_s)$ , with  $n_s$  the number of samples] and does not require large memory since only one row of the matrix is computed.

The approximation that the data segments are ring shaped or circulant might seem unreasonable, but in the end this only has an effect on a small fraction of the matrix. For data with large-scale correlations (data described by a  $1/f$  power spectrum, for instance), the approximation implies an assumption that the two edges of the segment are very correlated. In reality, there is obviously little correlation at long timescales compared to short timescales, so extra striping could be introduced in the maps if one steps across the two edges of the data segment. We have addressed this problem in two ways in order to avoid introducing artifacts in the final map. In the case where we explicitly compute the inverse pixel-pixel covariance matrix  $N_{pp'}^{-1}$  (as in the brute-force inversion algorithm), for the estimation of  $N$  entering in the computation of  $N_{pp'}^{-1}$ , we constrain  $\bar{C}(\Delta t) = 0$  for  $\Delta t > n_s/2$ , with  $n_s$  being the number of samples in the segment; this is sometimes known as “the MADCAP approximation” in the literature. It is important to note that this approximation cannot be used for the computation of  $A^t N^{-1} d$  because it is performed partly in Fourier space (see § 3.2.3). Instead, we have apodized the data  $d$  at the edges, and we have removed a low-order polynomial (fifth order in practice) to reduce fluctuations having typical timescales of order (or larger than) the data segment (see § 2.2). This is reasonable, since those scales are not well described by a limited

number of Fourier modes and hence will always be hard to reconstruct.

### 3.2.2. Inverse Pixel-Pixel Covariance Matrix $N_{pp'}^{-1}$

The computation of the inverse pixel-pixel covariance matrix, which is described in this section, is required only in the brute-force inversion algorithm, or for an accurate error estimation in the pixel domain.

Since the pointing matrix has only one nonzero element per row in our simple model (one data sample is associated with a single map pixel), the matrix multiplication  $A^t N^{-1} A$  requires a single loop going across all the nonzero elements of  $N^{-1}$ . For most cases, a dominant fraction of the mapmaking computing time will be devoted to this operation. If the data are only correlated within a typical length  $\lambda_c$ , we have the property  $\bar{C}(\Delta t) \simeq 0$  for  $\Delta t > \lambda_c$ , and  $N^{-1}$  is a band-diagonal matrix (elements separated from the diagonal by more than  $\lambda_c$  are negligible). The number of elements to go through in the loop is of the order of, but smaller than,  $2n_s \lambda_c$ , which is hopefully much smaller than the size of the matrix itself.

Unfortunately, if the noise is described by a power spectrum of the form  $(1/f)^\beta$ , the correlation length of the noise is basically of the order of the length of the whole data set. However, the amplitude of the correlation is decreasing for very long timescales and becomes negligible beyond a certain scale. The function  $\bar{C}(\Delta t)$  can then be artificially set to zero for  $\Delta t > \lambda'_c$ , with  $\lambda'_c$  chosen empirically so that the correlation of the noise is low enough for scales longer than  $\lambda'_c$ , and also that there is very little constraint on the signal at those scales. For the specific case of BLAST observations, we find that  $\lambda'_c \simeq 200$  s is a good compromise, as illustrated in Figure 1.

The noise strongly dominates the signal in BLAST observations for scales longer than  $\lambda'_c$  (corresponding to frequencies smaller than  $5 \times 10^{-3}$  Hz), since (1) this frequency is well below the knee frequency of the noise power spectrum and (2) there is very little signal at frequencies smaller than  $5 \times 10^{-3}$  Hz, which is more than 10 times smaller than the scanning frequency. Therefore, we have used  $\lambda'_c = 200$  s for the computation of the noise covariance matrix. The impact of fixing  $\bar{C}(\Delta t) = 0$  for  $\Delta t > \lambda'_c$  in the initial power spectrum is shown in Figure 1. We can see a tight relation between the power spectrum and the inverse covariance matrix, since getting  $\bar{C}(\Delta t) = 0$  for  $\Delta t > \lambda'_c$  can be obtained only by modifying the power spectrum at frequencies smaller than  $1/\lambda'_c$ .

### 3.2.3. Computation of $A^t N^{-1} d$

In the computation of  $x = A^t N^{-1} d$ , which is necessary for both of our algorithms, the multiplication  $N^{-1} d$  is performed in Fourier space where the noise covariance matrix is diagonal. We obtain this vector by dividing the Fourier transform of the data by the power spectrum of the noise. Another way to represent this is by considering that since  $N^{-1}$  is circulant,  $N^{-1} d$  is a convolution operation. Assuming that this model holds, the resulting data vector  $\tilde{d} = N^{-1} d$  contains whitened noise. The remaining operation  $A^t \tilde{d}$  just performs the addition of the filtered data sample onto the pixels of  $s$  (i.e., the map) and hence is very fast.

In the case of BLAST observations, since we are not attempting to recover useful information from timescales larger than 200 s, we perform a high-pass prefiltering of the input data  $d$  at  $5 \times 10^{-3}$  Hz.

### 3.2.4. Matrix Inversion Algorithm

In the matrix inversion algorithm,  $N_{pp'}^{-1}$  is directly computed, as described in § 3.2.2. The next step is to solve the linear system

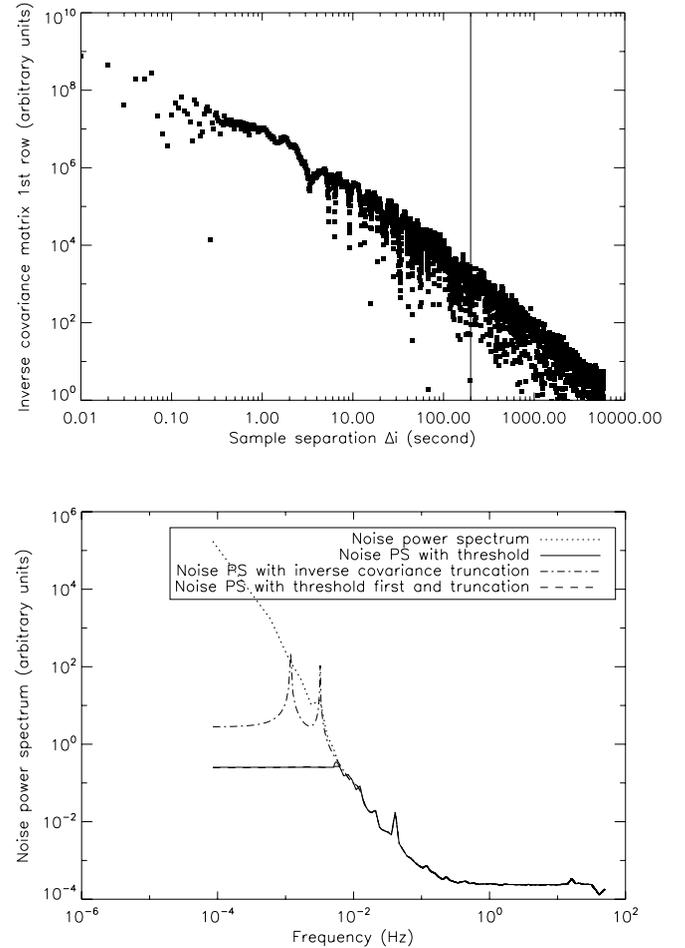


FIG. 1.—*Top*: Absolute value of the first row of the inverse covariance matrix given by  $\bar{C}(\Delta t)$  (see § 3.2.1) for a typical BLAST observation. The vertical line indicates the value of  $\lambda'_c$ , such that for  $\Delta t > \lambda'_c$ ,  $\bar{C}(\Delta t)$  is set to zero for the computation of  $N^{-1}$ . *Bottom*: Power spectrum of the noise in a typical BLAST observation (*dotted line*) and after thresholding at frequencies smaller than  $5 \times 10^{-3}$  Hz, i.e.,  $1/\lambda'_c$  (*solid line*). The dot-dashed line is obtained by inverting eq. (12) after forcing the values of  $\bar{C}(\Delta t)$  to zero for  $\Delta t > \lambda'_c$ , with  $\bar{C}$  obtained from the initial power spectrum of the noise (*dotted line*). To find the power spectrum that corresponds to the dashed line, the same procedure is applied, but starting from the thresholded power spectrum (*solid line*). All the lines are very similar for frequencies larger than  $1/\lambda'_c$ , but the dot-dashed line begins to diverge at smaller frequencies, while the dashed and solid lines lie very close to each other for all frequencies. This shows the tight relation between the power spectrum and the inverse covariance matrix. The noticeable peak in the power spectrum is located at the scanning frequency of about  $4 \times 10^{-2}$  Hz.

$\hat{s} = [N_{pp'}^{-1}]^{-1} x$ . For small maps, in which  $N_{pp'}^{-1}$  can be stored in memory, we perform a Cholesky decomposition. For larger maps, we write the matrix to disk and perform an iterative inversion of the system using a conjugate gradient method with preconditioner.

This algorithm allows one to easily perform multiple Monte Carlo simulations, since the pixel-pixel covariance matrix, being independent of the data realization, can be computed once (this is assuming that the noise power spectrum is known and does not have to be estimated from the data themselves) and used for all the realizations of simulated data. Another advantage of this approach is that it allows for the exact computation of the errors in the map, which are given by the covariance matrix.

However, the matrix inversion algorithm is generally slower than the iterative algorithm, which we discuss next, and can be used only for relatively small maps; we found that we were limited

to around 200,000 pixels if the matrix is written to disk, or less than 20,000 pixels if the matrix is stored in memory (for a processor with 4 GB RAM).

### 3.2.5. Iterative Algorithm

We now present an iterative algorithm based on conjugate gradient with preconditioner to obtain the maximum likelihood solution for the map (a similar algorithm has been used in Ashdown et al. 2007a). Let us rewrite equation (5), which relates the best estimate of the map with the data, after multiplying both sides by the pixel-pixel covariance matrix:

$$A^t N^{-1} A \hat{s} = A^t N^{-1} d. \quad (13)$$

If we define  $\hat{s}_k$  as an estimate of the map at iteration  $k$ , the conjugate gradient method allows us to solve the linear system by minimizing iteratively the following criterion:

$$\Psi = r^t N_{pp}^{-1} r, \quad (14)$$

where

$$r \equiv (A^t N^{-1} A \hat{s}_k - A^t N^{-1} d). \quad (15)$$

This criterion is indeed minimum and equal to zero if  $\hat{s}_k$  is the maximum likelihood solution.

One can interpret  $\Psi$  as the weighted variance of the difference between two map vectors. The first of these vectors,  $A^t N^{-1} A \hat{s}_k$ , is the inverse pixel-pixel covariance matrix times the current estimate of the map, while the second vector,  $A^t N^{-1} d$ , is a map constructed by simply co-adding prewhitened data. We decide that convergence is reached when the quantity  $r^t r$  (which is much easier to compute than  $\Psi$ , and also converges to zero) gets smaller than a predefined value. In practice, the number of iterations required for convergence is of the order of 100.

The conjugate gradient method is not described here, since it is a fairly standard numerical tool, and the interested reader can find many descriptions in the literature.<sup>16</sup> Instead of describing the details, we focus here on aspects that are specific to our mapmaking process. In particular, let us describe the computation of  $A^t N^{-1} A \hat{s}_k$ , which is the time-consuming part of the optimization and has to be performed at each iteration (the computation of  $A^t N^{-1} d$  is also time consuming but needs to be done only once, since none of the parameters are changing through the iterations); the other operations for updating the map at each iteration are significantly faster.

One advantage of this iterative algorithm is that the computation of the full pixel-pixel covariance matrix is not required, and the operation  $A^t N^{-1} A \hat{s}_k$  can be done step by step. Indeed, we start by computing  $\hat{d} = A \hat{s}_k$ , which is an estimate of a “signal” time stream. This operation is equivalent to scanning over the current estimate of the map using the pointing solution. The subsequent operation  $A^t N^{-1} \hat{d}$  (which should now be familiar) is carried out in Fourier space, as described in § 3.2.3 (without applying any extra filtering), and in § 3.4 for the case of correlated noise between detectors.

This iterative approach is in general much faster than the brute-force inversion approach because the most time-consuming operations are performed in Fourier space. It also requires less memory, since  $N_{pp}^{-1}$  is not explicitly computed. Of course, if there are found to be (or known to be) nontrivial correlations in  $N_{pp}^{-1}$ ,

then it may have to be calculated explicitly, hence requiring the brute-force approach. However, provided that the pixel-pixel correlations only involve relatively few pixels, it should be possible to calculate a restricted part of (or perhaps an approximation for)  $N_{pp}^{-1}$  in a modified iterative approach. A related concept is discussed in § 3.8.

### 3.3. Multiscan, Multidetector Case

In the previous subsection we presented the general method for the simple case where only a single continuous observation is considered. We now describe how we combine observations from different detectors at the same wavelength, as well as different data segments obtained over different “visits” during the flight, where by “visit” we mean a period in the data that starts after a sufficiently long gap, or after the observation of a different region of the sky.

For convenience, the data vector  $d$  in our model (eq. [1]) now contains all the individual data segments from different detectors, and also within a single channel, concatenated end to end. The noise vector  $n$  in equation (1) is defined in a similar manner. The matrix  $A$  in equation (1) is then the result of stacking individual pointing matrices. The maximum likelihood solution is also written as in equation (5), with  $N$  becoming the full covariance matrix of the noise, including all the channels and data chunks. To start with, let us assume that there is no correlation of the noise between data segments. This is a very good assumption if we consider data segments obtained over different visits, but it is certainly not a good approximation for segments obtained simultaneously with different channels, since we found a very strong common-mode noise between detectors. We consider the simple no-correlation case first, and then in the next subsection (§ 3.4) we generalize the mapmaking method to account for noise correlations from different detectors.

In the absence of correlations between data segments, the time-time noise covariance matrix  $N$  is block diagonal and each block can be inverted separately. Defining  $N_\ell$  as the subcovariance matrix for the data in segment number  $\ell$ , and  $A_\ell$  as the subpointing matrix going from the map to the data segment  $\ell$ , the inverse pixel-pixel covariance matrix can be written as

$$N_{pp}^{-1} = \sum_{\ell} A_{\ell}^t N_{\ell}^{-1} A_{\ell}. \quad (16)$$

The computation time for obtaining this matrix is proportional to the number of data segments. In this simple case the computation of  $x = A^t N^{-1} d$  can be written as

$$x = \sum_{\ell} A_{\ell}^t N_{\ell}^{-1} d_{\ell}, \quad (17)$$

where the computation of each term is fast and can be performed partly in Fourier space.

### 3.4. Detector-Detector Correlated Noise

We now allow for the presence of correlations in the noise between different detectors. In the case of multiple visits, the noise covariance matrix,  $N$ , still has null cross terms for samples from two different data visits. Therefore, if the data vector is sorted by visit, then  $N$  is block diagonal and each block contains the correlation coefficients between all the detectors for the samples within the time interval defined as a single visit. Each visit can be treated independently, since the submatrices can be inverted separately, and equation (16) is still valid, but in this case  $\ell$  is the label for the blocks in  $N$ . In the following, we therefore focus on a

<sup>16</sup> For example, the 1994 article by J. R. Shewchuk: <http://www.cs.cmu.edu/~quake-papers/painless-conjugate-gradient.pdf>.

single visit and consider observations by all the detectors; the generalization to multiple visits should be clear.

To simplify the notation, let  $N$  denote the noise covariance matrix for the visit being considered, with  $d$  and  $n$  being the data and noise vectors, respectively, containing the time stream segments for all the detectors put end to end, and  $N_{ij}$  being a block of  $N$  of size  $n_s \times n_s$ , corresponding to the noise correlations between detectors  $i$  and  $j$ . Let us define  $\bar{F}$  as the multichannel Fourier transform operator such that

$$\tilde{n} = \bar{F}n, \quad (18)$$

with  $\tilde{n}$  containing end-to-end Fourier transforms of each data segment.  $\bar{F}$  is a block-diagonal matrix, and each block is the Fourier transform operator  $F$  for one data segment.

In Fourier space, the noise covariance matrix  $R$  can be written as

$$R = \bar{F}N\bar{F}^\dagger. \quad (19)$$

If we consider a single block of the noise covariance matrix for detectors  $i$  and  $j$ , we obtain

$$R_{ij} = FN_{ij}F^\dagger. \quad (20)$$

Under the assumption that the data are stationary and continuous at the edges (see § 3.2 for a discussion),  $N_{ij}$  is a circulant matrix, since each element  $[N_{ij}]_{tt'}$  depends only on the time interval  $|t - t'|$ .  $R_{ij}$  is then a diagonal matrix with the diagonal given by the cross-power spectrum of the noise between detectors  $i$  and  $j$ :

$$[R_{ij}]_{\omega\omega'} = \begin{cases} P_{ij}(\omega), & \omega = \omega', \\ 0, & \text{otherwise.} \end{cases} \quad (21)$$

Here  $P(\omega)$  is the noise covariance matrix of size  $n_d \times n_d$  for a given mode  $\omega$ , where  $n_d$  is the total number of detectors. The computation of the inverse of  $R$  is straightforward, since each Fourier mode can be treated independently. If  $P^{-1}(\omega)$  is the inverse noise covariance matrix for mode  $\omega$ , the same relation as in equation (21) applies between  $R^{-1}$  and all  $P^{-1}$ .

From equation (19), we can calculate the inverse covariance matrix of the noise in real space:

$$N^{-1} = \bar{F}^\dagger R^{-1} \bar{F}. \quad (22)$$

Then, a block of  $N^{-1}$  between detectors  $i$  and  $j$  can be written as

$$[N^{-1}]_{ij} = F^\dagger [R^{-1}]_{ij} F. \quad (23)$$

Because  $[R^{-1}]_{ij}$  is a diagonal matrix (as discussed previously),  $[N^{-1}]_{ij}$  is circulant and is related to the inverse of the matrix containing the cross- and auto-power spectra of the noise:

$$[N^{-1}]_{ijtt'} = \mathcal{F}^{-1} \left\{ [P^{-1}]_{ij} \right\} (t' - t). \quad (24)$$

From this relation, we can see that in practice  $N^{-1}$  is relatively easy to construct, since each of its blocks (referring to each pair of detectors) is a circulant matrix, so only a row of each submatrix needs to be calculated using the fast Fourier transform. Finally, in the case in which the noise covariance matrix is used for multiplication in real space (as in the brute-force algorithm), the same approximation described in § 3.2 is performed on each block of  $N^{-1}$ , i.e.,  $[N^{-1}]_{ijtt'} = 0$  for  $|t - t'| > \lambda'_c$  (or for

$|t - t'| > n_s/2$  if  $\lambda'_c > n_s/2$ ). The global structure of the final inverse noise covariance matrix is illustrated in Figure 2.

In our model of the data (eq. [3]) in which all the correlations between detectors are described by a single common mode,  $P^{-1}(\omega)$  can be related to the power spectra of the common mode and of the uncorrelated part of the noise between detectors:

$$P^{-1}(\omega) = \left[ \alpha \langle c(\omega)^\dagger c(\omega) \rangle \alpha^t + \langle \tilde{n}(\omega)^\dagger \tilde{n}(\omega) \rangle \right]^{-1}. \quad (25)$$

This relation can be generalized if multiple common-mode components are present:  $\alpha$  would then be a mixing matrix and  $\langle c(\omega)^\dagger c(\omega) \rangle$  becomes the covariance matrix of these noise components.

Having computed the inverse noise covariance matrix of the time streams, we can express (using eq. [6]) the noise covariance matrix in the pixel domain:

$$N_{pp'}^{-1} = \sum_{ij} A_i^t [N^{-1}]_{ij} A_j, \quad (26)$$

where, as before,  $i$  and  $j$  label the detectors. The computation time for  $N_{pp'}^{-1}$  is now proportional to the number of detectors squared. The calculation of  $x = A^t N^{-1} d$  is straightforward:

$$x = \sum_{ij} A_i^t [N^{-1}]_{ij} d_j, \quad (27)$$

and this is fast, since (as already shown) the operation  $[N^{-1}]_{ij} d_j$  is a convolution, which can be performed in Fourier space. One can see from equation (24) that  $x$  can be expressed directly with respect to the cross- and auto-power spectra of the noise:

$$x = \sum_{ij} A_i^t \mathcal{F}^{-1} \left\{ [P^{-1}]_{ij}(\omega) \bar{d}_j(\omega) \right\}. \quad (28)$$

The formalism presented above can also be generalized easily to deal with detectors operating at different wavelengths. The map vector  $s$  could be merging different maps at different wavelengths, and the noise matrix  $N$  would account for all the correlations of the noise between detectors. The joint multiband mapmaking would be suitable in practice when some contaminations from thermal fluctuations in the instrument or atmospheric emission are present because they correlate the noise at all wavelengths.

### 3.5. Noise Power Spectrum

The maximum likelihood solution for the final map depends on the noise power spectra for each data set (through  $N$  in eq. [5]), which are assumed to be perfectly known. However, in practice the noise power spectra have to be inferred from the data themselves and some uncertainties are associated with this iterative process.

In practice, a first (approximate) estimate of the noise power spectrum can be obtained by rebinning the power spectra of each data segment, neglecting the contribution of the astrophysical signal in the time lines. Indeed, for most of the fields observed with BLAST, and in particular for blank extragalactic fields (like the ELAIS-N1 field in BLAST's flight from Sweden), the noise is highly dominant over the sky signal at all frequencies. However, this is not true for measurements of bright regions in the Galactic plane. Therefore, for the first iteration's noise estimate we focus on the data taken for about 6 hr while scanning the deepest extragalactic field (ELAIS-N1) and use this to estimate the noise power spectra, which then become the noise input for making the

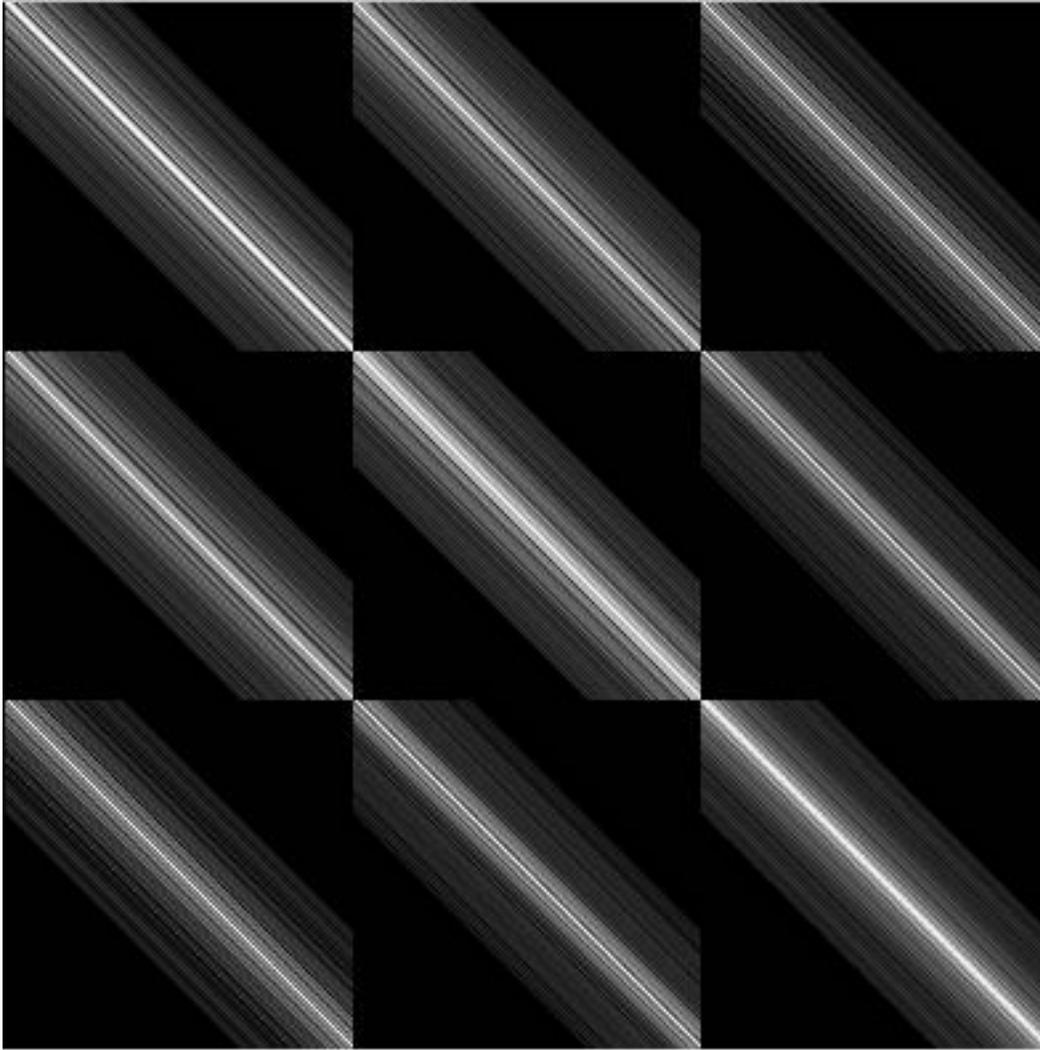


FIG. 2.—Inverse noise covariance matrix ( $N^{-1}$  in the text) for three detectors and for only 10 minutes of data (corresponding to 60,000 samples per detector). The matrix is computed following eq. (24), and the approximations described in § 3.4 are applied. The cross- and auto-power spectra of the noise, used for the calculation of  $N^{-1}$ , are computed from the data themselves (one of the auto-power spectra is shown in Fig. 3; the cross-power spectra are an unbiased measure of the common-mode signal shown in the same figure). Each submatrix is circulant and corresponds to a particular pair of detectors. One can see that the off-diagonal submatrices have amplitudes of the same order as that of the diagonal submatrices. This is due to the very high level of noise correlation between detectors.

first set of maps of *all* of our fields. This approach is not entirely satisfactory, since the noise is not stationary during the flight: the noise power spectra are seen to vary over long time-scales, although they are quite constant within a single visit of each field. This nonstationarity appears most obvious when there are variations of the scanning strategy between different visits (a change of the scanning frequency induces a variation of the location of peaks in the power spectrum) but can also occur due to variations of detector loading or the detector bias being changed during the flight.

Because of the observed nonstationarity of the noise, we would like an independent estimate of the noise power spectra for each visit for each of the observed fields. After starting from a first estimate of the noise power spectra based on ELAIS-N1 data as described above, we adopt an iterative approach between maps and noise power spectra. At each iteration, the estimated maps are subtracted from the data, prior to noise power spectrum estimation. We can summarize our procedure in the following steps:

1. Estimate  $P_0(\omega)$  from  $d$  using a field known to have little signal.

2. Compute  $\hat{s}$  from equation (5) (and also eq. [24] or eq. [12], depending on the noise correlation being considered) using  $P_0(\omega)$  as input.

3. Estimate  $P(\omega)$  from  $d - A\hat{s}$ .

4. Reestimate  $\hat{s}$  from equation (5) using  $P(\omega)$ , and iterate on these last two steps until convergence is achieved.

We stop iterating when the noise power spectra do not vary by more than 1 per 1000 from iteration to iteration (and find that in practice only three to six iterations are necessary to reach convergence).

We now focus on how we estimate the noise power spectra  $P(\omega)$  in steps 1 and 3. For the simplest case, where no correlations are assumed between detectors, we simply compute a bin-averaged power spectrum for each data segment:

$$P_\ell(q) = \frac{1}{n_q} \sum_{\omega_{\min}(q)}^{\omega_{\max}(q)} \tilde{d}_{\ell\omega}^* \tilde{d}_{\ell\omega}, \quad (29)$$

where, for bin number  $q$ ,  $n_q = \omega_{\max}(q) - \omega_{\min}(q) + 1$ ,  $\ell$  labels the data segment, and  $\tilde{d}$  is the data vector from which an estimate

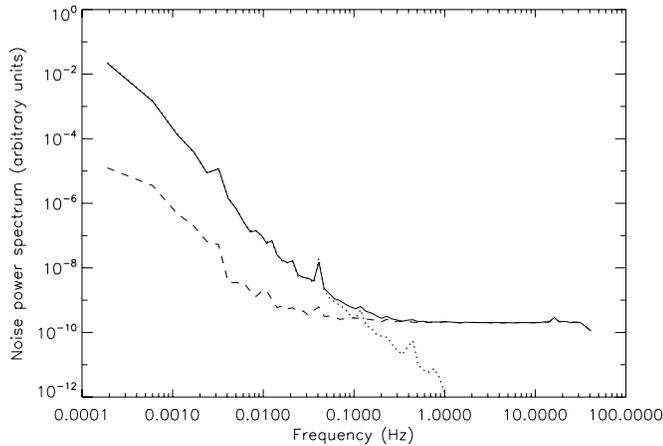


FIG. 3.—Total noise power spectrum (*solid line*) for one of the 250  $\mu\text{m}$  detectors, using about 3 hr of data, corresponding to scans of the ELAIS-N1 field. The dotted line corresponds to the estimated power spectrum of the common mode between detectors at 250  $\mu\text{m}$ , rescaled by an amplitude factor for the specific detector being considered ( $\alpha_i$  in the text, where  $i$  is the detector index), also estimated using the data themselves. The dashed line represents the estimated power spectrum of the uncorrelated part of the noise for this detector. The common mode is very strong in these data and dominates over the uncorrelated noise at frequencies lower than about 0.1 Hz. Most of the low-frequency noise excess comes from this common mode. The uncorrelated part of the noise shows a knee frequency of less than 0.02 Hz, which is 5–10 times smaller than for the total noise power spectrum. The power of the uncorrelated noise is thus reduced by a factor of about 100 at low frequencies. The excess signal at the scanning frequency (peak around 0.04 Hz in the power spectrum) is completely common between detectors.

of the map has already been subtracted. We have chosen logarithmic spacing between bins, and an estimate of  $P(\omega)$  for each  $\omega$  mode is obtained by logarithmic interpolation, which leads to a smooth power spectrum estimate.

In the more complicated case where correlations between detectors are assumed to be important and therefore are not neglected, we must estimate, for every iteration, each cross- and auto-power spectrum of the data between detectors,  $P_{ij}(\omega)$ , which enter into the computation of the inverse noise covariance matrix (eq. [24]). Each cross-power spectrum could be directly estimated as in equation (29) (using  $\tilde{d}_{i\omega}$  and  $\tilde{d}_{j\omega}$  in the formulae for detectors  $i$  and  $j$ ), but instead we choose to reduce the number of parameters to estimate at each step, by assuming that the data are described by a common mode between detectors plus *independent* noise (eq. [3]). In the framework of this model, the expected cross- and auto-power spectra depend directly on the following parameters:  $\alpha$ , the amplitude of the common mode in each channel;  $\langle c(\omega)^* c(\omega) \rangle$ , the power spectrum of the common-mode part; and  $\langle \tilde{n}(\omega)^* \tilde{n}(\omega) \rangle$ , the power spectrum of the noise component, which is independent between detectors. The relation between the model of  $P(\omega)$  and the parameters has been shown in equation (25). These parameters are typically not known a priori and must be measured using the data themselves. We use a blind “component separation” method developed for an entirely different problem in Delabrouille et al. (2003). This allows us to obtain a single estimate of all the parameters described previously, by simultaneously using all the observed time streams of a given field for all the detectors in a specified channel (i.e., at a single frequency). The method is known to be the maximum likelihood solution for a Gaussian and stationary model of both the noise and the common mode. The cross- and auto-power spectra  $P(\omega)$  are then computed following equation (25), using these same estimated parameters. Figure 3 shows the estimated noise power spectra in a sample of BLAST data for one representative detector using 3 hr of time

streams during scans of the ELAIS-N1 field (which is known to be essentially devoid of signal).

The auto-power spectrum is shown, as well as its decomposition in terms of the common-mode power spectrum and the uncorrelated noise power spectrum.

Because our estimate of the converged map of the sky  $\hat{s}$  is not perfect and contains contributions from residual noise, then in subtracting a simulated signal time line from the data to estimate the noise power spectrum we reintroduce some noise to the data, which could potentially bias our estimate of the noise power spectrum (Ferreira & Jaffe 2000; Hamilton 2003). However, this effect is greatly reduced by the large redundancy in each pixel of the final maps, as a result of the many repeated scans and the large number of detectors at each wavelength. The bias can be neglected to first order for BLAST, since the noise level per pixel in the final map is much smaller than the noise in individual detector time streams.

### 3.6. Dealing with Gaps in the Data

In order to derive the formalism presented so far, we have assumed that each data segment is stationary and hence consists of a continuous series of data points. However, we have seen in § 2.2 that the BLAST05 data contain multiple gaps of typical size less than 1 s. The amount of data in these gaps is a few percent of the total. In order to reasonably restore the continuity of the data, we have filled the gaps with random noise, as described in § 2.2. The data samples generated in the gaps are not reprojected into the final map but are directed to “dummy” pixels. In principle, the optimal approach would be to create one dummy pixel per flagged data sample, avoiding the possibility of several simulated samples falling on the same pixel (through the rebinning of flagged pixels we do not want to introduce any spurious constraints for the mapmaking process that could arise from adding crossings over different time intervals). However, the approach of using one dummy pixel per gap sample is impractical because the total number of dummy pixels would be excessive for typical BLAST observations, and the pixel-pixel covariance matrix (which has size the total number of pixels squared) would be prohibitive to store and compute.

We have adopted a simpler approach, which does not lead to the mathematically exact solution but comes very close (as has been shown in simulations). This consists of rejecting (for the computation of  $N_{pp}^{-1}$ ) all the elements of  $N_{ii}^{-1}$  associated with flagged samples. This is equivalent to removing from  $N_{pp}^{-1}$  the rows and columns corresponding to the dummy pixels before the inversion of the matrix, as opposed to after inversion, which would be the correct treatment discussed above. This approach is also equivalent to assuming null off-diagonal terms for those rows and columns. However, such dummy pixels are obviously correlated to some degree with the real pixels in the map, and hence this cannot be entirely correct.

Nevertheless, we have verified that this approximation has a small impact (a few percent only) on the final map for most of the angular scales that are sampled, although we found some differences at very large angular scales, sizes of the order of the map size. For now, we have not put much effort into recovering those very large scales because they are subject to other effects, as discussed in § 5.2.2. The minimal impact on the final map has been verified using pure signal simulations and by comparing the results obtained between our simple approach and the correct mapmaking solution. This approach works to a high degree of accuracy on most scales because the gaps are small and do not introduce important discontinuities in the time streams.

We have used this simple approach because it gives sufficiently accurate results over the relevant angular scales, while being simple and fast to implement. However, another iterative procedure could be adopted, which would lead to the exact solution. In this approach, we define two maps. The first map “A” is made from only the uncorrupted (i.e., real sky) samples, while the second map “B” is obtained from projecting the simulated (i.e., for gap filling) samples. The difficulty arises in deciding what to do when simulated data from different scans or detectors fall in the same pixel: one might want the “generated signal” (and not necessarily the noise) to be identical in both measurements, in order to satisfy the mapmaking hypothesis. If this condition is not satisfied, then some artifacts may be introduced into both maps. Here is a solution to this problem:

1. Generate a first set of maps A and B after filling the gaps in the time streams with white noise plus a linear baseline.
2. Fill the gaps in the data with the best estimate of the signal in map A together with white noise plus a baseline that is fitted in the gap vicinity of the data “minus” signal time stream.
3. Recompute maps A and B. Step 2 ensures that the signal is the same for each generated sample falling in the same pixel of map B.

This approach can also be coupled with the procedure for estimating the noise power spectra described in § 3.5. Preliminary results indicate that this approach works in practice. Detailed studies will be presented in a future publication.

### 3.7. Pixel Constraints

For some specific observed fields we may have strong priors about the sky emission at a given location. For instance, we know that over some regions the astronomical signal should vary very smoothly or should be very weak with respect to the noise, at least outside some localized region. This is the case in particular when we map bright extragalactic sources in order to calibrate the detectors and estimate the beams; in these cases, regions beyond some predefined distance from the beam center can be assumed to have null flux (or a constant relative flux in the map, since we do not have access to the DC level in maps). If we really have strong prior knowledge that we are dealing with a bright localized region, then we can take a further drastic step: we can constrain the map to have the same value in some domains of the sky by defining a single pixel containing all the data samples falling in that region.

In practice, we define a small box centered at the source location and constrain the part of the map outside this box to have a constant value. This is a very efficient way to remove stripes from the map, since the extremities of all the paths across the map are readjusted. We have used this technique to make maps of the isolated calibrators observed by BLAST (Truch et al. 2008).

### 3.8. Error Estimation

The variance of the noise in each pixel of the final map and its correlations are directly given by the pixel-pixel covariance matrix  $N_{pp'} = (A'N^{-1}A)^{-1}$ . This is true given the following assumptions: that our model of the data holds, in particular that the noise is a purely Gaussian random process, which may not be the case in practice at low frequencies; and that our estimate of the sample-sample noise covariance matrix  $N_{tt'}$  is accurate enough that the errors do not propagate significantly into the final map. As already mentioned, we never explicitly compute the covariance matrix, but rather its inverse. The direct inversion would take a prohibitive computation time for most applications. However, to first order, we can obtain an estimate of the errors by

inverting the diagonal part of  $N_{pp'}^{-1}$  only, neglecting the off-diagonal terms. This is equivalent to assuming that the noise in the final map is white. We have checked with the help of simulations that this very simple approximation is accurate to better than 10% for all our BLAST05 fields, even for those with very poor cross linking.

Provided that the size of the map is reasonably small, so that we are able to explicitly calculate  $N_{pp'}$ , we can obtain an accurate estimate of the errors for a limited (and small) number of pixels in the map. The variance for pixel  $p$ , and the covariance with respect to the other pixels of the map, can be computed by solving the linear system:

$$\langle n_p^\dagger n_{p'} \rangle = N_{p'p} u_p, \quad (30)$$

where  $u_p$  is a unitary vector with a single 1 for pixel  $p$ . If the Cholesky decomposition of the matrix has already been performed for the mapmaking procedure, the computation of equation (30) is relatively fast and hence can be carried out for a grid of non-adjacent pixels, for example. This can be used to check the validity of the error prediction approximation described in the previous paragraph.

### 3.9. Computational Requirements

For the brute-force inversion algorithm (in which the full inverse pixel-pixel covariance matrix  $N_{pp'}^{-1}$  is computed), 5 minutes of computation with a single 3 GHz processor are needed to process 2 hr of data from a single detector at a rate of 100 Hz. The computational time is proportional to the number of samples if this is longer than the assumed correlation length of the noise in the data (which has been evaluated to be  $\lambda_c = 200$  s in BLAST05 time streams). If noise correlations between detectors are also to be accounted for, the computational time is proportional to the square of the number of detectors.

Most of the computing time is spent on calculating  $N_{pp'}^{-1}$ . Inversion of the linear system to estimate the map  $s$  is relatively fast (a few minutes to a few hours for maps of several square degrees in size).

For the iterative algorithm, about 2 minutes of computational time is required for 2 hr of data (under the same conditions described above). This assumes that 100 iterations are necessary to reach convergence (which is a realistic number for most applications), and the algorithm scales with  $n_s \log n_s$ . The situation is much better than for the brute-force inversion algorithm if correlations between detectors are included. In that case, the algorithm scales with the square of the number of detectors if this exceeds about 40. If there are fewer than about 40 detectors, then the algorithm scales linearly with the number of detectors. As an example, if there are 100 detectors, then including noise correlations between the detectors increases the computational time by a factor of 4 with respect to the “no-correlation” case. The full processing of 10 hr of BLAST05 data at all wavelengths, including detector-detector correlations, can be done with a single processor in a few days.

In terms of memory, the brute-force inversion algorithm requires storage of the full  $N_{pp'}^{-1}$  matrix. However, for the iterative algorithm, only vectors of the size of the maps need to be kept in memory, which is much less demanding.

## 4. SIMULATIONS FOR TESTING SANEPIC

We now focus on the application of SANEPIC to data. Our aim is to develop tests to validate our method using simulated BLAST observations. We derive conclusions about how well

low-frequency noise in the maps can be reduced, depending on observational parameters such as scanning strategy, and we compare the results obtained with those from simpler methods based on filtering the data, e.g., common-mode subtraction. In this section we describe the simulations performed to test the SANEPIC method.

We have generated several different sets of simulations of BLAST time streams. Each set of simulated time streams, representing one particular observed field, is generated for all the BLAST detectors used for the analysis of real data (132 at  $250\ \mu\text{m}$ , 78 at  $350\ \mu\text{m}$ , and 39 at  $500\ \mu\text{m}$ ) and is the sum of simulated astrophysical signal, independent noise, and common-mode noise between all detectors. The noise is generated randomly with Gaussian statistics, given fixed power spectra derived from real BLAST05 data. Figure 3 shows an example of the power spectrum of the noise in the data used as input to the simulations for one of the fields. The part of the noise that is independent between detectors is generated for every detector time stream and has a power spectrum well described on average by a relatively flat plateau for frequencies larger than about 0.05 Hz, and by a part proportional to  $(1/f)^{2.5}$  for smaller frequencies (these characteristics vary slightly from detector to detector). Our knowledge of the real bolometer noise power spectrum at low frequency is limited by the very dominant common mode. In simulations, the common-mode noise is generated once for all detectors and has a power spectrum very well fitted by a power law with an index equal to approximately 2.5, together with some broad peaks, the largest being at the scanning frequency (the amplitude of the peak depends on the scanning strategy and the observed field). The common-mode power spectrum has an amplitude such that it reaches the level of the independent noise at about 0.3 Hz (see Fig. 3). The generated common-mode time stream is multiplied by an amplitude factor that varies from detector to detector by  $\sim 10\%$  and is added to the simulated detector time streams. The amplitude factors used for the simulations have been estimated from the data themselves.

In order to represent the astrophysical signal, we have simulated simple maps of diffuse emission with a power spectrum proportional to  $k^{-3}$ , as for typical Galactic cirrus emission (e.g., Miville-Deschênes et al. 2007). Maps are generated following Gaussian statistics with a resolution of  $1''$ , much higher than the typical pixel sizes in the final maps, in order to reduce artifacts due to repixelization. The amplitude of the fluctuations of the simulated map is chosen to match the expected level of signal in each observed field. The simulated maps are scanned using BLAST05 pointing, and pure signal time streams are generated for each detector. Signal and noise time streams are added at the end of the procedure (but see § 5 for an explanation of why this operation is not always carried out).

We have generated two sets of simulations that correspond to two different fields observed by BLAST. We selected two fields that were observed with very different scanning strategies, since the performance of the mapmaking procedure is very dependent on scanning strategy; this allows us to test SANEPIC in two very different configurations. In the first case the scanning was performed mainly in a single direction over a short time interval, while in the second case the field was observed several times during the flight at different scanning angles, to achieve significant cross linking in the map.

The first data set uses observations of the Cas A supernova remnant emission, which comprises about 20 minutes of data. BLAST observations of this field and derived conclusions will be described in detail in P. C. Hargrave et al. (2008, in preparation). The rectangular region mapped has a size of the order of  $0.5\ \text{deg}^2$

and was scanned two times back and forth over a short time interval. We have generated simulations corresponding to all the detectors at  $250\ \mu\text{m}$  (we used a total of 132 detectors).

The second data set reproduces the observations of the intermediate-velocity cloud (IVC) G86.5+59.6 (hereafter “G86”). Simulations include four different visits of the field performed during the flight at very different time intervals (ranging from a few hours to more than a day). Each continuous observation segment has a size that varies from 1 to 2 hr. Two scanning directions are dominant, which form an angle close to  $50^\circ$ . The region covered has a size of about  $2\ \text{deg}^2$  on the sky. Simulations for this field are performed specifically for all the  $500\ \mu\text{m}$  detectors (41 detectors used). Similar Monte Carlo simulations at  $250\ \mu\text{m}$  would have taken a factor of 10 longer, while we believe that the conclusion would remain unchanged.

A total of 20 sets of simulations of the observations for each field have been performed. For each set we vary the realization of the noise and of the signal input map. About 4 hr of computing time are needed to create one realization of a full set of simulations of G86 with a single processor, compared to a few minutes for the Cas A simulations. This is using the precomputed full pixel-pixel covariance matrix, which was also used to analyze the real data.

## 5. RESULTS FROM SIMULATIONS

We now present the results obtained with SANEPIC applied to the two sets of simulated data. In each case, we compare the final map with other maps obtained using simpler mapmaking procedures. For these tests we have assumed that the noise power spectra are perfectly known, rather than estimated separately from each data set; in practice, we fix the noise power spectra to be the ones from the simulations. We have verified that relaxing this constraint has almost negligible change on the final maps.

For these simulated data sets, we have applied some preprocessing of the time streams before applying SANEPIC, just as we do for the real data. We have systematically removed a fifth-order polynomial from each time stream segment and weakly high-pass filtered the data, as described in § 2.2. Finally, for the gap filling we flag the simulated data at the same locations as in the real data in order to check the influence of the flagging procedure in the final maps.

In the following we have applied SANEPIC independently to pure noise time streams (containing independent noise and common-mode noise, but without simulated astrophysical signal) and to pure signal time streams. This procedure allows us to easily derive conclusions about the noise properties in the final maps, as well as about the signal, without biasing the results, because SANEPIC is a linear method (as shown by eq. [5]). This is only strictly true if the noise power spectrum is fixed, as done here, and not estimated simultaneously along with the maps. Then, applying SANEPIC on pure noise time streams and pure signal time streams independently and adding the two final maps is rigorously equivalent to applying SANEPIC on signal plus noise time streams. This has been checked numerically, and we find that the difference is consistent with double floating precision error. An important consequence of this is that the properties of the noise in the final map are independent of the signal-to-noise ratio.

### 5.1. Case without Cross Linking

#### 5.1.1. Noise-only Time Streams

We first study the maps resulting from the noise-only time streams in the configuration of Cas A observations. The chosen pixel size of the map is  $25''$  and matches the pixel size of the

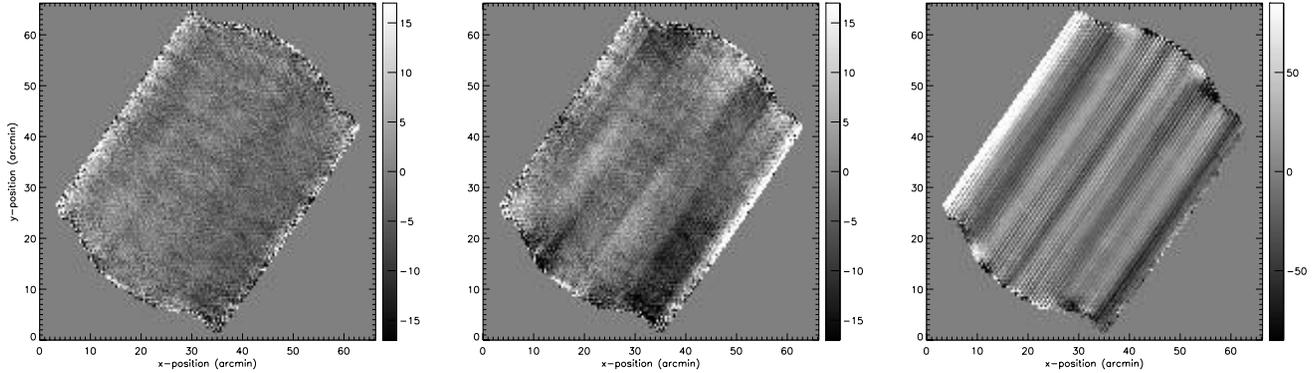


FIG. 4.—Final maps computed from simulated pure noise time streams in the configuration of the BLAST05 Cas A observations, which have a dominant scan direction. *Left to right*: Maps obtained with SANEPIC including noise correlations, SANEPIC with no noise correlations included in the model, and simple pixel binning (see text for more details). Note the extended dynamic range of the simple co-added map (*right*). The maps have a size of about  $40'$  in the cross-scan direction and about  $1^\circ$  along the scan. The pixel size is  $25''$ .

maps discussed in P. C. Hargrave et al. (2008, in preparation). We compare the noise maps obtained from three different procedures:

1. Use SANEPIC with the correct treatment of the correlated noise.
2. Use SANEPIC fixing the correlation of noise between detectors to zero and fixing the noise power spectrum for each detector to the power spectrum of the sum of uncorrelated noise and common mode. This procedure is very similar to more standard mapmakers in the literature (e.g., Stompor et al. 2002).
3. Make a simple reprojection of the data onto a pixelized map by simply averaging the data falling in each pixel, after having filtered the time stream data with the same very weak low-pass filter used for SANEPIC. This procedure is sometimes called “co-addition.”

Figure 4 shows computed noise maps for one of the realizations of the noise in each of the three cases.

As expected, the map obtained with the simple pixel binning approach contains a very large amount of low-frequency noise, with strong striping visible along the scan direction. Residual low-frequency noise can also be seen in the map obtained using SANEPIC *without* accounting for the noise correlations between detectors. We do not expect this method to be very efficient, since it is very nonoptimal in cases (such as this example) where a very large fraction of the noise is correlated between detectors. In contrast, the noise map obtained with SANEPIC is quite satisfactory, showing reduced power at low frequency as compared to the previous case. Nevertheless, some very weak excess power is seen in the cross-scan direction. This is expected, since the map is not cross linked, and very poor constraints can be put on the cross-scan directions at low spatial frequencies (two positions in the map separated by more than the size of the array in the cross-scan direction are observed far apart in time).

In order to quantify the level of low-frequency noise in the maps, we compute the one-dimensional (1D) power spectra of the maps, averaged over the 20 realizations of the simulated data. For the computation of power spectra, we take into account only the central part of each map, where the level of redundancy in the observations is high (we use only the highest signal-to-noise ratio region in the maps). To do so, we apply an apodized mask to the maps going smoothly from 0 at the edges to 1. Figure 5 shows the noise power spectra in the three cases.

The noise level in the simple reprojection map is obviously very poor at all scales. Both of the other mapmakers reach the white-noise level for scales smaller than  $3'$  and have excess power at larger angular scales. Nevertheless, the gain between

full SANEPIC and SANEPIC without correlations is very important at all scales larger than about  $2'$  and reaches a maximum value of about 10 at around  $20'$  angular scales. An interesting fact is that the knee frequency of the noise power spectrum in the optimal case here corresponds to the inverse of the physical scale of the detector array in the cross-scan direction (which is of the order of  $6'$ ). Indeed, there are no observational redundancies on scales larger than the array in the cross-scan direction in the absence of cross linking in the map. Thus, the very long timescale  $1/f$  noise present in the time streams is not efficiently removed and reprojects in the final map at large angular scales. This effect is also present along the scan direction, but with a lower amplitude as the map is scanned back and forth. The trend of the large angular scale power spectrum of the noise in the map just

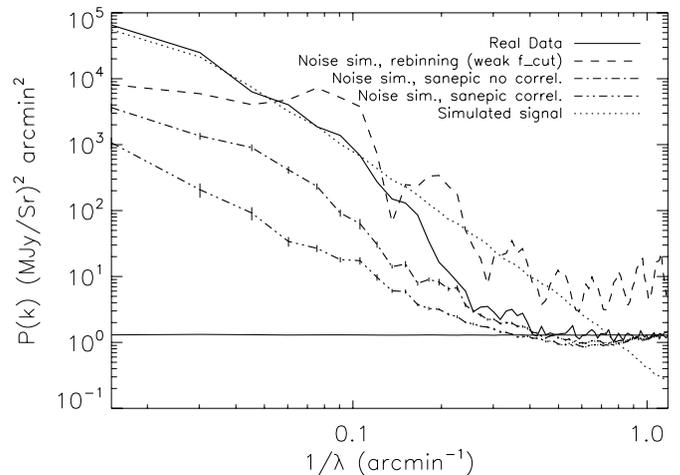


FIG. 5.—1D power spectra of the noise (rebinning in frequency) in the final noise maps after mapmaking in the BLAST05 Cas A configuration. Power spectra are averaged over the 20 realizations of the simulated data. The dashed line is for the simple reprojection map, the dot-dashed line for SANEPIC with *no* noise correlation between detectors, and the double-dot-dashed line for SANEPIC *including* a treatment of the correlations. The horizontal line indicates the level of white noise in the map predicted by the mapmaking procedure (see § 3.8). Error bars are computed from the dispersion of measurements among the realizations. For comparison, the upper dotted line (decreasing almost like a power law at all scales) represents the power spectrum of the pure simulated signal in the final map. The solid line represents the power spectrum of the final map obtained with real data using SANEPIC, with correlations included. This shows the benefit of taking into account correlations of the noise between detectors in the mapmaking procedure, reducing the noise structure far below that of the signal in the map. The real data power spectrum shows that the signal dominates at all angular scales larger than about  $3'$ , and at smaller scales we can see that white noise at the expected level dominates in the map. The drop of power at around a  $3'$  scale is due to the BLAST05 beam.

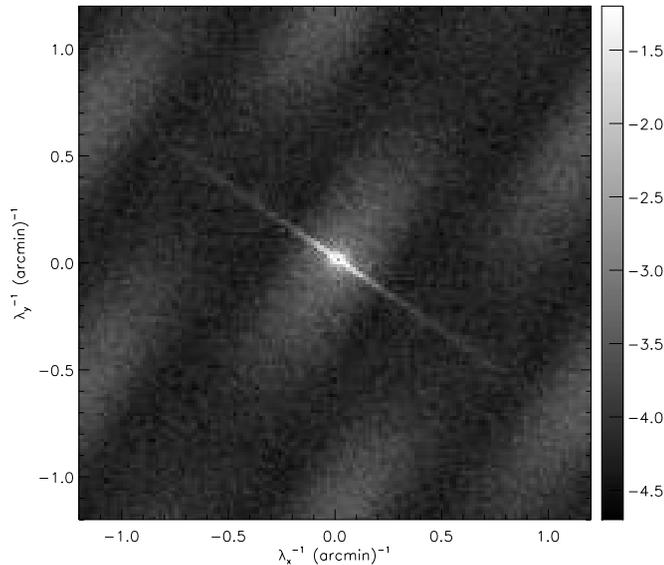


FIG. 6.—2D power spectrum of the noise maps in the BLAST05 Cas A observational configuration obtained with SANEPIC (noise correlations included) plotted on a logarithmic contrast scale.

follows the trend of the low-frequency noise power spectrum in the time streams. We see in § 5.2 that this effect is reduced when there are multiple scanning directions in the map.

In order to determine the direction in which the noise power is strongest in the map, we have also computed the two-dimensional (2D) noise power spectrum. The map of the 2D power spectrum of the noise obtained with SANEPIC (noise correlations included) is shown in Figure 6.

The large bright spot around the center corresponds to a relatively isotropic component of correlated noise (at least at large angular scales). It contains a large fraction of the noise power at large angular scales (seen in the 1D power spectrum in Fig. 5). A smaller but significant fraction of the correlated noise is concentrated in directions perpendicular to the scan direction, as can be seen in the figure. As already discussed, the reason for this excess power is that the noise in the cross-scan direction is poorly constrained. This cross-scan component of the noise is significant all the way up to the pixel scale.

### 5.1.2. Signal-only Time Streams

We now focus on the signal-only time stream simulations. In order to demonstrate the superior performance of SANEPIC relative to simpler methods based on data filtering, we compare with a mapmaking method that consists of the following: we first

remove the whole array average from each detector time stream and then make maps using SANEPIC, assuming no correlations between detectors. Removing the array average reduces the signal to almost zero for scales larger than the array and so we expect no large-scale structures to survive in the map. This SANEPIC “common-mode subtraction” method is still a better procedure than just reprojecting the data (after common-mode subtraction) with a well-chosen filtering to suppress noise drifts (at  $f_{\text{cut}} = 0.02$  Hz, for instance, since that corresponds to the knee frequency of the independent part of the noise). The latter method is commonly used in the submillimeter community and is referred to as “sky removal” in reduction of SCUBA data (Jenness et al. 1998).

Figure 7 shows the input map for one of the signal realizations (*left panel*), as well as the maps obtained with SANEPIC (correlations included; *middle panel*) and with the common-mode subtraction method (*right panel*). Results are expected to be worse in the second case because of the extra filtering and also because SANEPIC gives less weighting to modes at lower frequency that are more contaminated by independent noise.

We can see from Figure 7 that part of the very large scale fluctuations with sizes of the order of the map are removed using SANEPIC, but apart from those very large scales, the input map and the SANEPIC map look very similar. More differences can be seen in the map obtained with the common-mode subtraction method. This is quantified in Figure 8, which compares the 1D power spectra of the two output maps, averaged over 20 simulations and multiplied by  $k^3$ . Recall that the input spectrum varies as  $k^{-3}$  and so deviations from a flat line are the result of the mapmaking reconstruction. Note that the vertical scale is linear in Figure 8.

With SANEPIC, the power of the reconstructed map decreases for scales larger than about  $30'$ . There are three reasons for this: the power spectrum is computed over only a small fraction of the sky (and for an apodized map), so that structures of the order of the map size are never fully recovered; there is weak filtering of the time streams at  $f_{\text{cut}} = 5 \times 10^{-3}$  Hz and through the fifth-order polynomial subtraction; and modes in the maps that are very weakly constrained in the mapmaking procedure tend not to be reconstructed through the matrix inversion procedure, since the matrix is very ill conditioned and numerical problems occur. The last two effects are the dominant ones. As a result, modes that are preferentially filtered are those that lie perpendicular to the scan direction.

At angular scales smaller than about  $2'$ , the power slightly decreases due to the smoothing effect of the pixelization. The transfer function at those scales is well described by a sinc function.

Turning now to the common-mode subtraction method, the power in the map is significantly reduced for scales larger than

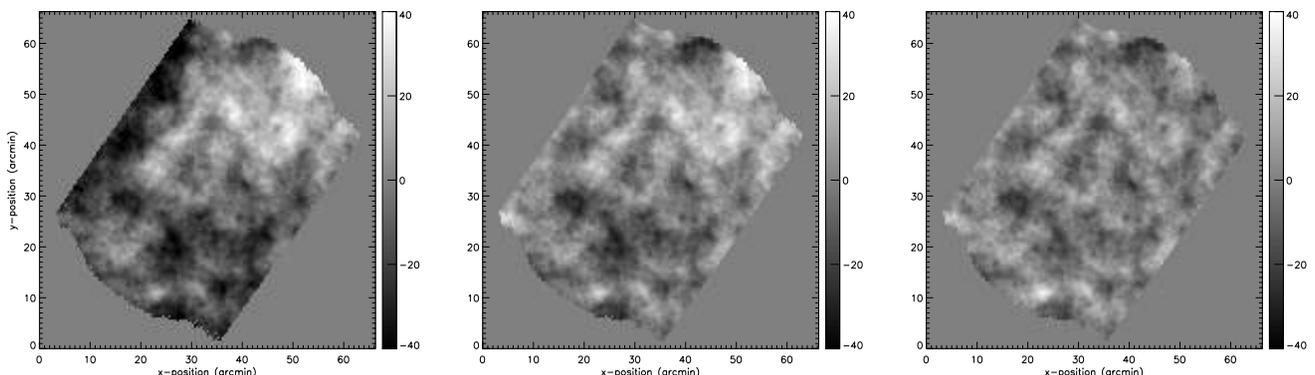


FIG. 7.—*Left to right*: Map of simulated signal used as input for one of the realizations of the simulations in the configuration of the BLAST05 Cas A field, reconstructed map using SANEPIC accounting for noise correlations, and reconstructed map using the simple common-mode subtraction method (intensity units here are arbitrary).

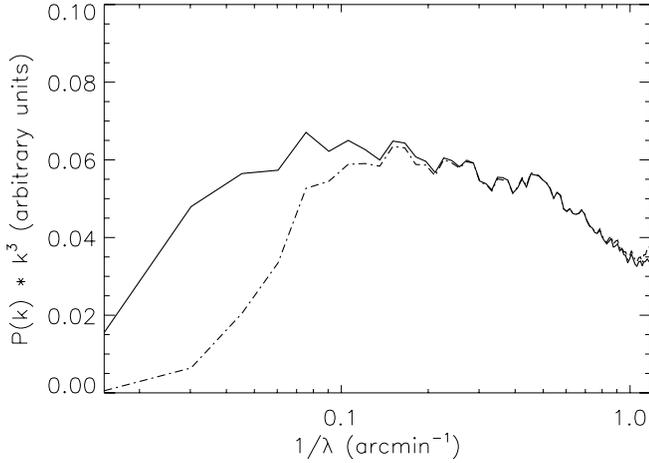


FIG. 8.— 1D power spectra of signal-only maps for the BLAST05 Cas A field reconstructed using SANEPIC (*solid line*) and using the common-mode subtraction method (*dot-dashed line*). Power spectra are multiplied by  $k^3$ , displayed on a linear scale, and averaged over 20 realizations of the simulations. The large angular scale behavior shows the effective filtering of each mapmaking procedure, while the drop-off at small scales is caused by the PSF.

about  $10'$  and drops rapidly to zero. This is because the common-mode subtraction removes power on all scales larger than the array. On smaller scales, the filtering effect is relatively weak and is reduced when the number of detectors increases.

For these particular simulations the common-mode subtraction method (using SANEPIC, but with no correlations) does not in fact perform *very* poorly compared to the SANEPIC optimal approach. This is because the observed field is small, with a size just a few times bigger than the array, and structures at scales smaller than the array size are not strongly affected. This particular map is also not cross linked. However, the situation is different for large cross-linked maps like the Vulpecula field, as discussed in § 6.2.

## 5.2. Case with Cross Linking

### 5.2.1. Noise-only Time Streams

We now focus on the set of simulations of the G86 field at  $500 \mu\text{m}$ . As in the previous example, we first examine the maps resulting from noise-only simulated time streams using three methods: optimal SANEPIC (with noise correlations taken into account), SANEPIC without considering noise correlations between detectors, and the simple co-add method. The chosen pixel

size for the reconstructed maps is  $1'$ , which allows for inversion of the covariance matrix with a single processor and hence rapid Monte Carlo simulations. The conclusions drawn would remain unchanged if the pixel size was reduced.

Figure 9 shows the final noise maps in the three cases for one realization of the simulations.

The simple reprojection map is obviously very stripy and would be of little use as an estimate of the signal; nevertheless, it helps to visualize the directions of scanning in the map. We can see two main directions covering the central region of the map, oriented at about  $50^\circ$  to each other. A third scanning direction is also visible but has a much smaller weight. The central part of the map is the cross-linked region, where we expect the more optimal mapmaking procedures to excel.

In the map obtained using SANEPIC *without* considering noise correlations between detectors (Fig. 9, *middle panel*), some large-scale noise is still visible in the map, even if almost no residual striping is apparent in the cross-linked region. Indeed, too much weight is given to the large timescales in the time streams (which are basically common between all the detectors) as compared to the smaller timescales for which there are more independent measurements because the degree of correlation of the noise is weaker. The residual noise at large angular scales is much weaker in the SANEPIC map in which we account for the proper correlations of the noise between detectors (Fig. 9, *left panel*). The noise in this map looks particularly white.

The noise level in each simulated map is quantified in Figure 10, showing the 1D power spectrum of residual noise averaged over 20 simulations and computed over the cross-linked region (which has a diameter of about  $100'$ ).

While several orders of magnitude are gained in the noise power at all scales using the SANEPIC “no correlation” method as compared to the simple reprojection method, accounting for the correlations with SANEPIC allows us to further reduce the noise power on scales ranging from  $20'$  to the size of the map by an additional factor of  $\sim 5$ . Toward smaller scales, the gain between the SANEPIC correlation and no-correlation test cases is still very significant down to about  $10'$ , where both methods start to approach the white-noise level. In the optimal map obtained with SANEPIC the ratio between the noise power spectrum at large scales and the white-noise level is around 20, which is relatively small. Figure 11 shows the 2D power spectra of the noise in the SANEPIC map averaged over 20 realizations. As expected, the large-scale noise is more important in directions perpendicular to the scans.

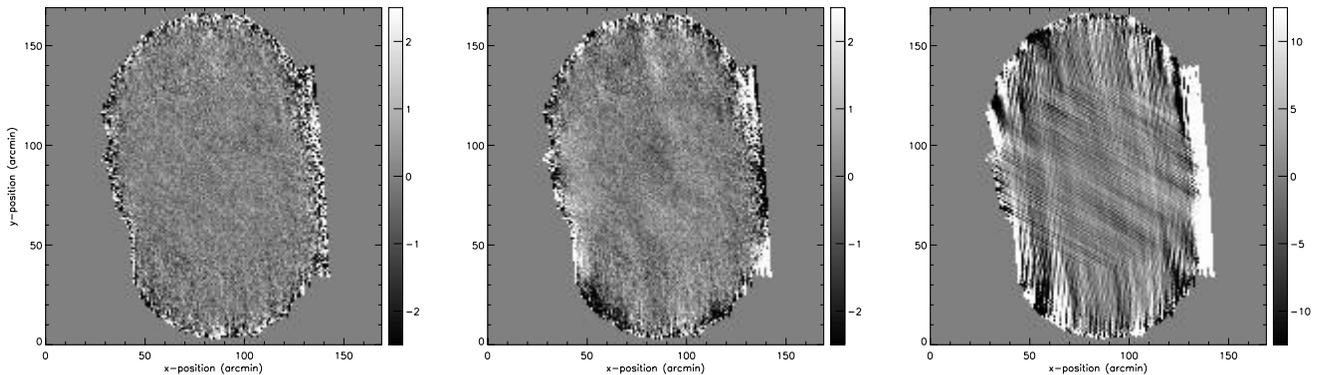


FIG. 9.— Noise-only simulations (like Fig. 4) for the BLAST05 G86 scanning configuration. The maps have a size of about  $100'$  across the short axis and  $2.5^\circ$  across the long axis. The pixel size here is  $1'$ . The three panels show the map obtained with SANEPIC including noise correlations (*left*), SANEPIC with no noise correlations in the model (*middle*), and the simple co-added map (*right*). Note the different brightness scale chosen for the last map due to its larger dynamic range.

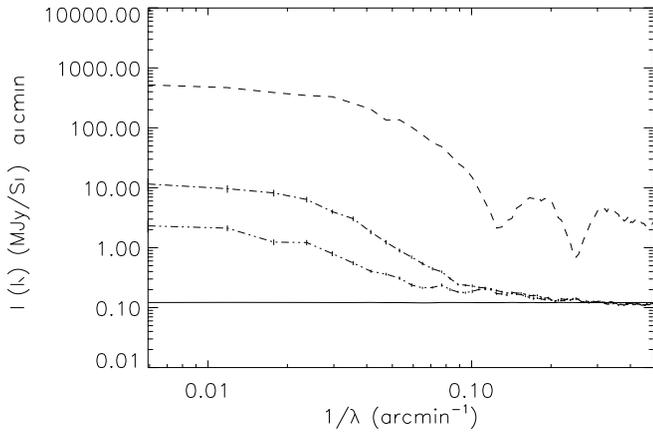


FIG. 10.—Output power spectrum comparison of noise-only simulations after mapmaking (like Fig. 5), for the BLAST05 G86 scanning configuration. These power spectra are computed only in the cross-linked region of the maps, which forms a large disk of about  $100'$  diameter and can be easily identified in Fig. 9. The dashed line is for the simple reprojection map (Fig. 9, *right panel*), the dot-dashed line is for SANEPIC *without* consideration of noise correlations between detectors (Fig. 9, *middle panel*), and the double-dot-dashed line is for SANEPIC *including* the correlation treatment (Fig. 9, *left panel*). The horizontal line indicates the level of white noise in the map predicted by the mapmaking procedure. Error bars are estimated from the dispersion among measurements for all the realizations. Residual low-frequency noise in the optimal SANEPIC map is very low, thanks to the multiple scanning directions of this field. The situation is much better than for the Cas A observational configuration, which had a single scan direction (Fig. 5).

### 5.2.2. Signal-only Time Streams

We now focus on signal-only simulations for G86. As in the case of the Cas A configuration, we compare the performance of SANEPIC with respect to the simpler common-mode subtraction method. Figure 12 shows the pure signal input map for one realization of the simulations (*left panel*) compared with two recovered maps. The first output map is obtained with SANEPIC including correlations between detectors (*middle panel*), while the second is obtained by subtracting the common mode between all detectors, followed by applying SANEPIC, but neglecting noise correlations between detectors (*right panel*). We see that the very largest scales are not recovered by SANEPIC. This is because of the weak filtering applied to the time streams and, to a greater extent, because of inversion problems with SANEPIC on scales of the order of the size of the map (these scales are very poorly constrained by the mapmaking procedure). In the common-mode subtraction method, only the very largest scales are suppressed by the mapmaking procedure itself, but the filtering effect is more dramatic and extends to much smaller scales.

The effective filtering is quantified in Figure 13, which shows the power spectra of output maps averaged over 20 simulations of pure signal time streams. Again the power spectra are multiplied by  $k^3$  for comparison purposes.

In this case SANEPIC works well on scales up to about half a degree, above which it fails to recover structure in the map; this limit corresponds to scales of about one-quarter of the map and larger. The filtering effect is much more pronounced in the map produced with the common-mode subtraction method, being strong for all scales above around  $14'$ . This is of course expected, since subtracting the average of the array strongly reduces the signal on scales larger than the array size. Therefore, in order to recover large- and intermediate-scale structure in the maps, it is beneficial to use SANEPIC instead of other methods that are based on simple filtering the data.

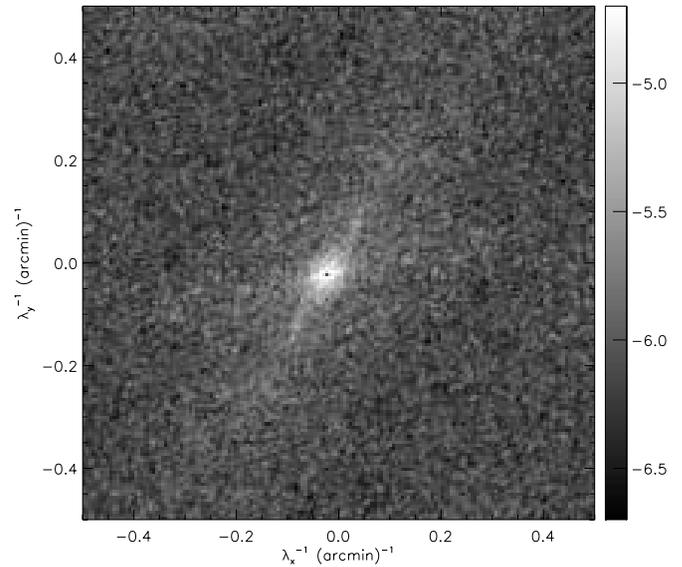


FIG. 11.—2D power spectrum of a noise-only simulation reduced using SANEPIC (like Fig. 6) for the BLAST05 G86 scanning configuration.

### 5.3. Advantages of Cross Linking

The relative level of residual noise at low spatial frequency in the maps is significantly reduced in the G86 observational configuration as compared to the Cas A configuration. The fundamental difference is that the G86 observations contain multiple (essentially two for most of the data) scanning directions, while the Cas A observations are realized with only two passes across the field in the same direction. Multiple scanning directions give a huge number of additional constraints for the mapmaking procedure. In particular, large-scale structures in the map are much better recovered in directions parallel to scans because the noise there is smaller. Thus, having multiple scanning angles allows for recovery of the sky fluctuations for all directions and ends up giving almost no weight to the individual loosely constrained cross-scan  $k$ -modes.

Differences in the results for maps from these two example scanning strategies can be quantified in two ways. First of all, for the G86 scanning strategy, the transition between white noise and “excess” large-scale noise in the map occurs at a scale around  $10'$ , while the same transition occurs at a scale of around  $3'$  for the Cas A scanning strategy. Second, the ratio between large-scale noise power and white-noise power is larger by more than 2 orders of magnitude for Cas A than for G86. On the other hand, some caution should be taken not to overinterpret this comparison because the pixel size we used is larger for the G86 map ( $1'$ ) as compared to the Cas A map ( $25''$ ), and therefore the number of crossings per pixel is greater for the G86 map. Nevertheless, this simulation exercise has demonstrated that cross linking in the map is extremely beneficial, particularly for recovering the large-scale structures in the map.

### 5.4. Mapmaking Transfer Function

When carrying out a complex data processing procedure, it is important to check whether the results are biased in any way. We have found that the transfer function of the mapmaking procedure, defined as the ratio between the amplitude of fluctuations in the output pure signal map relative to the input map, is not always exactly unity, even for intermediate and small angular scales in the map. For example, in the Cas A configuration at  $250 \mu\text{m}$  the fluctuation amplitudes in the final map are reduced

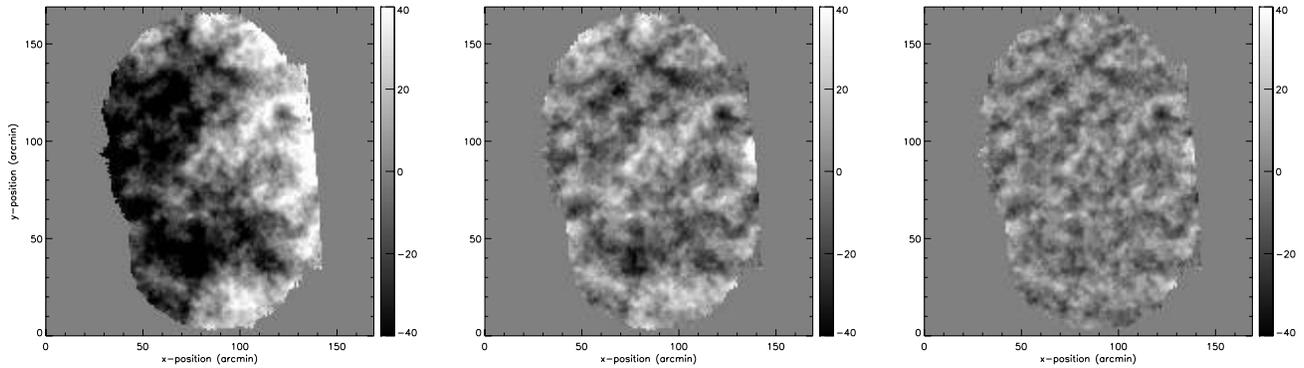


FIG. 12.—Signal-only simulations (like Fig. 7) for the BLAST G86 configuration. The three panels show the input signal map (*left*), full SANEPIC (*middle*), and simple common-mode subtraction followed by application of SANEPIC without correlations (*right*).

by 3% on average as compared to the input map, almost uniformly across spatial frequencies and directions. This global discrepancy reaches the level of 9% at  $500 \mu\text{m}$ . Moreover, it is also present for the G86 configuration. We believe that this reduction is due to the fact that the pixel-pixel covariance is ill conditioned and numerical imprecision occurs in the matrix inversion. We find that the bias tends to be smaller when the number of detectors is larger and also when the number of constraints increases, like when we have multiple scanning directions, or when we map isolated bright sources (presented in Truch et al. 2008) and constrain the data outside a defined region to have a constant flux (see § 3.7 for details of this procedure). Since this bias can be estimated using simulations, it is straightforward to correct for. We have found that it is not always important, e.g., for the large Galactic map in the Vulpecula region analyzed in Chapin et al. (2008) the bias is negligible.

## 6. APPLICATION TO REAL BLAST05 DATA

Now that we have looked at signal-only and noise-only simulations, we now turn to real BLAST data. In this section we show maps of two example fields from the BLAST05 data that have been obtained using SANEPIC.

### 6.1. Cassiopeia A

Figure 14 shows the map obtained from the observations of the Cas A field at  $250 \mu\text{m}$  using SANEPIC including full con-

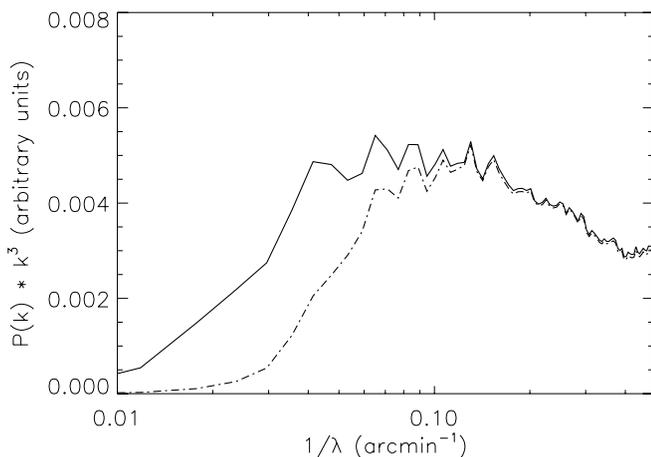


FIG. 13.—Power spectrum comparison of the SANEPIC map (*solid line*) vs. the simple common-mode removal map (*dot-dashed line*) for signal-only simulated data, as in Fig. 8, but for the BLAST05 G86 scanning configuration. Power spectra are multiplied by  $k^3$ , so that a flat line would indicate no filtering. The drop-off at small angular scales is due to the BLAST05 PSF.

sideration of the noise correlated between detectors. Detailed analysis of the maps at the three wavelengths is described in P. C. Hargrave et al. (2008, in preparation). The properties of the noise and the transfer function of the signal in the map have been studied in detail from Monte Carlo simulations in § 5.1. Results from such simulations are used to characterize the map. The power spectrum of this map has been compared to results from simulations in Figure 5.

The map structures are relatively smooth, due to the BLAST05 PSF, which has a width of the order of  $3'$ , causing the drop in the 1D power spectrum below those spatial scales. The signal clearly dominates over noise on angular scales larger than about  $3'$ , and the diffuse structure should be reliable up to a large fraction of the overall map size.

### 6.2. Vulpecula Region

Another field observed during the BLAST05 campaign is centered in the Galactic plane close to the open cluster NGC 6823 in the constellation of Vulpecula. The region mapped has a size of about  $4 \text{ deg}^2$  and was chosen for its high-mass star formation activity. Complete analysis of this observed field is presented in Chapin et al. (2008).

A few hours of these data were taken at different time intervals during the flight. By design, this field has been observed with

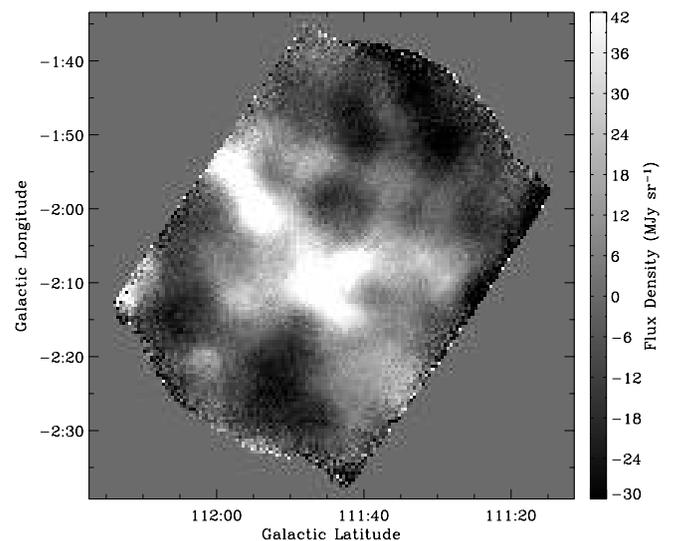


FIG. 14.—Map of the Cas A supernova remnant at  $250 \mu\text{m}$  made from BLAST05 data using SANEPIC including noise correlations between detectors. The map is represented in Galactic coordinates with  $25''$  pixels and has a size of about  $0.5 \text{ deg}^2$ .

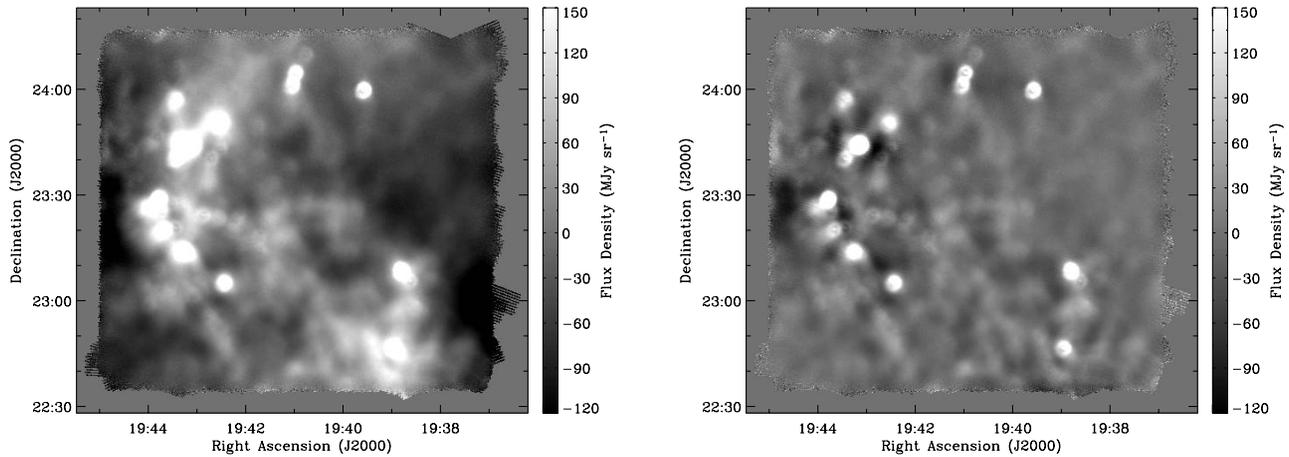


FIG. 15.—Maps of a Galactic plane region near NGC 6823 in the Vulpecula constellation derived from BLAST05 observations at  $250\ \mu\text{m}$ . The two maps are obtained using two different methods: SANEPIC (*left*) and simple reprojection after removing the array average signal at each time step from all the time streams, together with filtering (*right*). Maps are presented in equatorial coordinates with  $15''$  pixels. The region mapped has a size of about  $4\ \text{deg}^2$ . Numerous point sources have been identified in the field (see Chapin et al. 2008), their telltale shape in the map resulting from the PSF of the BLAST05 optics (see Truch et al. 2008).

very different scanning directions, and therefore it should be possible to recover diffuse large-scale structures.

The map of the observed region at  $250\ \mu\text{m}$  obtained with SANEPIC is shown in Figure 15 (*left panel*). For comparison (*right panel*), we have computed another map using a much simpler method that consists of removing the array average signal at each time step for all of the time streams and reprojecting the data onto the map after filtering. This is like the “sky removal” procedure often carried out for ground-based submillimeter data (see also § 5.1.2). One can see that it suppresses almost all the diffuse structure in the map.

No residual striping is visible in the map obtained with SANEPIC (Fig. 15, *left panel*), mainly due to the presence of multiple scanning directions for this field. Large-scale structures in the map are successfully recovered with SANEPIC, as can easily be seen by comparing with the right panel of Figure 15. This recovery applies to scales that are significantly larger than the array size. However, the resulting effective filtering after applying the “array average subtraction” method induces negative signals near bright sources in the map, while no such filtering effect is seen in the SANEPIC map (except perhaps near the edges of the map). This shows that optimal mapmaking methods (in the sense of least squares) like SANEPIC are better suited to recover point sources in the maps, as well as diffuse structures.

## 7. CONCLUSIONS

Large-format detector arrays operating at far-IR and submillimeter wavelengths are becoming the norm, rather than the exception. Ground-based instruments are plagued with common-mode emission arising from the Earth’s atmosphere. And as we have found with BLAST, the same applies to high-frequency balloon-borne instruments, where we see correlated noise from thermal and atmospheric effects. There is an expectation that even upcoming satellites might be faced with similar issues because of thermal variations in the spacecraft, for example. In general, we expect that correlated noise between detectors will be a major issue that all such experiments have to deal with, and

we expect that the SANEPIC approach, which we have described here, will be widely applicable. Indeed, there is evidence from existing arrays (e.g., SHARC-II and AzTEC) that once there are many detectors, there are *multiple* correlations between subsets of the detectors, as well as an overall common-mode term. Consequently, one sees correlations between contiguous blocks of detectors on the array, or sets of detectors that share amplifiers or are otherwise coupled through the electronics. Provided that these correlations can be investigated and their behavior modeled, it is straightforward to extend the SANEPIC approach to deal with several distinct sources of correlated noise. Hence, we expect the SANEPIC approach to be applicable to future instruments such as SCUBA-2, SPIRE, ACT, *Planck* HFI, and others. There are also many experiments being planned that use large detector arrays to perform sensitive polarization measurements, and we see no reason why the SANEPIC approach could not also be extended to polarimetry.

The BLAST collaboration acknowledges the support of NASA through grants NAG5-12785, NAG5-13301, and NNGO-6G111G, the Canadian Space Agency (CSA), Canada’s Natural Sciences and Engineering Research Council (NSERC), and the UK Particle Physics and Astronomy Research Council (PPARC). We would also like to thank the Columbia Scientific Balloon Facility (CSBF) staff for their outstanding work. We are grateful to Matthew Hasselfield and Duncan Hanson for valuable discussions that contributed to the development of SANEPIC. L. O. acknowledges partial support by the Puerto Rico Space Grant Consortium and by the Fondo Institucional para la Investigacion of the University of Puerto Rico. C. B. N. acknowledges support from the Canadian Institute for Advanced Research. This research made use of WestGrid computing resources, which are funded in part by the Canada Foundation for Innovation, Alberta Innovation and Science, BC Advanced Education, and the participating research institutions.

## REFERENCES

- Ashdown, M., et al. 2007a, *A&A*, 467, 761  
 ———. 2007b, *A&A*, 471, 361  
 Benoit, A., et al. 2002, *Astropart. Phys.*, 17, 101  
 Borrill, J. D. 1999, preprint (astro-ph/9911389)  
 Chapin, E. L., et al. 2008, *ApJ*, 681, 428  
 Crill, B. P., et al. 2003, *ApJS*, 148, 527  
 de Gasperis, G., et al. 2005, *A&A*, 436, 1159  
 Delabrouille, J. 1998, *A&AS*, 127, 555

- Delabrouille, J., Cardoso, J. F., & Patanchon, G. 2003, *MNRAS*, 346, 1089
- Devlin, M., et al. 2004, *Proc. SPIE*, 5498, 42
- Doré, O., et al. 2001, *A&A*, 374, 358
- Dupac, X., & Giard, M. 2002, *MNRAS*, 330, 497
- Ferreira, P. G., & Jaffe, A. H. 2000, *MNRAS*, 312, 89
- Hamilton, J.-Ch. 2003, preprint (astro-ph/0310787)
- Hinshaw, G., et al. 2003, *ApJS*, 148, 63
- Jenness, T., Lightfoot, J. F., & Holland, W. S. 1998, *Proc. SPIE*, 3357, 548
- Keihänen, E., Kurki-Suonio, H., & Poutanen, T. 2005, *MNRAS*, 360, 390
- Macías-Pérez, J., et al. 2007, *A&A*, 467, 1313
- Maino, D., et al. 2002, *A&A*, 387, 356
- Miville-Deschênes, M.-A., Lagache, G., Boulanger, F., & Puget, J.-L. 2007, *A&A*, 469, 595
- Natoli, P., et al. 2001, *A&A*, 372, 346
- Oliver, S., et al. 2000, *MNRAS*, 316, 749
- Pascale, E., et al. 2008, *ApJ*, 681, 400
- Poutanen, T., et al. 2006, *A&A*, 449, 1311
- Prunet, S., Netterfield, C. B., Hivon, E., & Crill, B. P. 2000, preprint (astro-ph/0006052)
- Prunet, S., et al. 2001, in *Mining the Sky*, ed. A. J. Banday, S. Zaroubi, & M. Bartelmann (Berlin: Springer), 421
- Stompor, R., et al. 2002, *Phys. Rev. D*, 65, 022003
- Tegmark, M. 1997, *ApJ*, 480, L87
- Tegmark, M., et al. 2000, *ApJ*, 541, 535
- Truch, M. D. P., et al. 2008, *ApJ*, 681, 415
- Wright, E. L., Hinshaw, G., & Bennett, C. L. 1996, *ApJ*, 458, L53
- Yvon, D., & Mayet, F. 2005, *A&A*, 436, 729