Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **416,** 2932–2943 (2011) doi:10.1111/j.1365-2966.2011.19244.x

Data reduction pipeline for the Hi-GAL survey

A. Traficante,[1,9⋆] L. Calzoletti,[2,8] M. Veneziani,[3,9] B. Ali,[5] G. de Gasperis,[1] A. M. Di Giorgio,[6] F. Faustini,[2] D. Ikhenaode,[4] S. Molinari,[6] P. Natoli,[1,2,7] M. Pestalozzi,[6] S. Pezzuto,[6] F. Piacentini,[3] L. Piazzo,[4] G. Polenta[2,8] and E. Schisano[6]

[1]*Dipartimento di Fisica, Università di Roma 'Tor Vergata', Italy*
[2]*ASI Science Data Center, I-00044 Frascati (Rome), Italy*
[3]*Dipartimento di Fisica, Università di Roma 'La Sapienza', Italy*
[4]*DIET – Dipertimento di Ingegneria dell' Informazione, Elettronica e Telecomunicazioni, Università di Roma 'La Sapienza', Italy*
[5]*NASA Herschel Science Center, Caltech, Pasadena, CA, USA*
[6]*INAF – Istituto Fisica Spazio Interplanetario, I-00133 Rome, Italy*
[7]*INFN Sezione di Tor Vergata, Rome, Italy*
[8]*INAF, Osservatorio Astronomico di Roma, Via Frascati 33, I-00040 Monte Porzio Catone, Italy*
[9]*Spitzer Science Center, Caltech, Pasadena, CA, USA*

Accepted 2011 June 13. Received 2011 June 8; in original form 2010 November 23

ABSTRACT

We present the data reduction pipeline for the *Herschel* Infrared Galactic Plane survey (Hi-GAL). Hi-GAL is a key project of the *Herschel* satellite, which is mapping the inner part of the Galactic plane ($|l| \leq 70^\circ$ and $|b| \leq 1^\circ$), using two Photodetector Array Camera and Spectrometer (PACS) and three Spectral and Photometric Imaging Receiver (SPIRE) frequency bands, from 70 to 500 μm. Our pipeline relies only partially on the *Herschel* Interactive Processing Environment (HIPE). It features several newly developed routines to perform data reduction, including accurate data culling, noise estimation and minimum variance map-making, the latter performed with the ROMAGAL algorithm, a deep modification of the ROMA code already tested on cosmological surveys. We discuss in depth the properties of the Hi-GAL science demonstration phase data.

Key words: instrumentation: photometers – methods: data analysis – techniques: image processing.

1 INTRODUCTION

The *Herschel* Space Observatory was launched from Kourou in 2009 May aboard an Ariane 5 rocket. Two of the three scientific instruments on the focal plane, the Photodetector Array Camera and Spectrometer (PACS) and the Spectral and Photometric Imaging Receiver (SPIRE), are capable of observing the infrared sky with unprecedented angular resolution and sensitivity. They provide photometric observations in six different bands: 70, 100, 160, 250, 350 and 500 μm (Pilbratt et al. 2010, and references therein).

The PACS photometer is composed of two bolometer arrays: a 64 × 32 pixel matrix arranged from eight monolithic subarrays of 16 × 16 pixels, each centred on the 70 and 100 μm wavelengths (blue and green bands), and a 32 × 16 pixel matrix organized in two subarrays for the band centred on 160 μm (red band) (see Poglitsch et al. 2010).

SPIRE comprises a three-band photometer, operating in spectral bands centred on 250, 350 and 500 μm. Every band uses a matrix of germanium bolometers (139, 88 and 43, respectively) coupled to hexagonally packed conical feed horns (Griffin et al. 2010).

To handle the science data provided by the *Herschel* instruments, including the data retrieval from the *Herschel* Science Archive, the data reduction through the standard pipelines and the scientific analysis, we use an official software environment called the *Herschel* Interactive Processing Environment (HIPE; Ott et al. 2010), which is available from the European Space Agency (ESA).

The raw data provided by the satellite are reduced in HIPE to generate the scientific data (so-called level 2) and intermediate products of the data-reduction process (level 0 and level 1 data).

In this paper, we describe the dedicated pipeline created to obtain maps for the *Herschel* Infrared Galactic Plane survey (Hi-GAL; Molinari et al. 2010a). Hi-GAL aims to homogeneously cover, with observations in five contiguous IR bands between 70 and 500 μm, a two-degree-wide stripe of Galactic plane between $l = -70^\circ$ and $l = 70^\circ$.

The Galactic plane shows emission varying from point-like sources to large-scale structures, with intensity varying over a wide dynamic range. In this paper, we show that the existing standard reduction strategy (which is based on HIPE version 4.4.0, released on 2010 November 11) is not optimized to reduce Hi-GAL data. We

⋆E-mail: alessio.traficante@roma2.infn.it

show that a dedicated pipeline can enhance the quality of the level 2 products.

After *Herschel* successfully passed the performance verification (PV) phase, two fields of the Hi-GAL survey were acquired during the science demonstration phase (SDP): a 2 deg^2 area of the Galactic plane centred on 30° of longitude (hereafter $l = 30°$) and on 59° ($l = 59°$).

We describe the data-reduction tools used to obtain high-quality maps from the SDP data, with the aim of providing a reliable environment for the routine phase (RP) data. The maps provided by our pipeline are successfully used for several works (e.g. Martin et al. 2010; Molinari et al. 2010b; Peretto et al. 2010).

The paper is organized as follows. In Sections 2 we describe the acquisition strategy for the Hi-GAL data. In Sections 3, we describe the pre-processing steps of the data-reduction pipeline, which are necessary to prepare the data for the map-making, and we describe the tools that we have developed for this purpose. In Sections 4 we describe the ROMAGAL map-making algorithm that is used to estimate the maps. ROMAGAL is used in place of the MADMAP code, which is the map-making algorithm offered in HIPE. The quality of the ROMAGAL maps for both the PACS and SPIRE instruments, related to the SDP observations, are analysed in Sections 5. We draw our conclusions in Sections 6.

2 HI-GAL ACQUISITION STRATEGY

The Hi-GAL data are acquired in PACS/SPIRE parallel mode.[1] In this mode, the same sky region is observed by moving the satellite at a constant speed of 60 arcsec s^{-1} and by acquiring images simultaneously in five photometric bands: 70 and 160 μm for PACS and 250, 350 and 500 μm for SPIRE.

The whole data acquisition is subdivided into $2° \times 2°$ fields of sky centred on the Galactic plane. Every Hi-GAL field is composed of the superposition of two orthogonal astronomical observation requests (AORs). Each of these is based on a series of consecutive, parallel and partly overlapped scan-legs covering $2° \times 2°$ deg^2. The scanning strategy adopted for Hi-GAL is fully described in Molinari et al. (2010a). The superposition is performed in order to obtain a suitable data redundancy and in order to better sample the instrumental effects, such as the high-frequency detector response (Molinari et al. 2010a).

The acquisition rate for the parallel mode is 40 Hz for PACS and 10 Hz for SPIRE, although the PACS data are averaged on-board for the effective rates of 5 and 10 Hz for the 70 and 160 μm arrays, respectively. The implications of the PACS data compression are detailed in Section 5.3.

An example of the scanning strategy of Hi-GAL is shown in Fig. 1. The coverage map of the PACS blue array is shown in the left panel. In the right panel, we highlight the superposition of one scan-leg to the following ones, by enlarging the bottom-right corner of the left image. Two calibration blocks for each AOR, during which the instrument observes two internal calibration sources located on the focal plane, were provided during the SDP observations. These appear as higher than mean coverage areas and they are marked in Fig. 1 by black and green circles for the two AORs. Higher coverage zones are also clearly visible in the slewing region at the end of each scan-leg, where the telescope decelerates and then accelerates before initiating the next scan-leg.

[1] http://herschel.esac.esa.int/Docs/PMODE/html/parallel_om.html

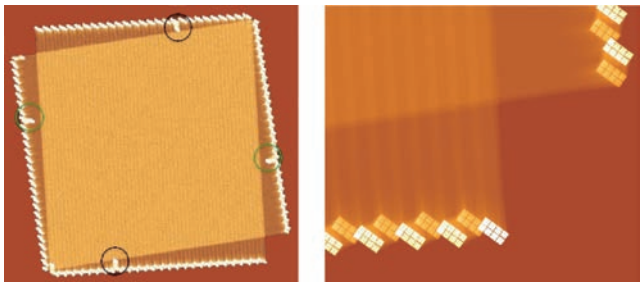

Figure 1. Left: the coverage map of the PACS blue array. Right: a zoom of the bottom-right corner where the effect of the superposition from one scan-leg to the next is clear. Black and green circles highlight the calibration blocks that, during the PV and SDP phases, were observed twice, during which the internal calibration sources are observed.

3 MAP-MAKING PRE-PROCESSING

The aim of the pre-processing is to prepare the Hi-GAL data for map-making.

While map-making is performed using a FORTRAN parallel code borrowed from cosmological observations (see Sections 4), the pre-processing is carried out through a series of IDL and JYTHON tools, which are run on the data, one after the other. After attempting to map the Hi-GAL data using the standard tools provided within HIPE, we decided to develop our own routines, which we have tailored specifically for the reduction of data affected by a bright and irregular background, as in the Galactic plane. In fact, the high-pass filtering used in HIPE to cure the long-time drift also removes a large part of the diffuse Galactic emission. Furthermore, the standard deglitching embedded in HIPE, the multiresolution median transform (MMT; Starck, Murtagh & Louys 1995), generates false detections with respect to the very bright sources when we apply this task to the PACS data. However, the deglitching procedure, based on wavelet analysis used by SPIRE, does not affect the data, when we also compare the lower spatial resolution with that of the PACS. We therefore use the HIPE task for the SPIRE data only.

The *Herschel* data are stored in subsequent snapshots of the sky, acquired by the entire detector array, called frames. In a first processing step, the standard modules within HIPE are used to generate the level 1 data for both PACS (except for the deglitching task) and SPIRE. Thus, the data for each AOR are rearranged into one time series per detector pixel, called time-ordered data (TOD), in which the calibrated flux (Jy beam^{-1} for SPIRE and Jy sr^{-1} for PACS) and the celestial coordinates are included. At the end of this step, the TOD are exported outside HIPE in fits format. In the subsequent processing steps, the TOD are managed by a series of IDL tools, in order to produce the final TOD free of systematic effects (as a result of electronics) and of glitches and corrupted chunks of data (as a result of cosmic rays). A flag file is attached to each TOD to keep track of any flagging carried out during the data-reduction steps.

Pre-processing includes the identification of corrupted TOD (or some of the data), drift removal and deglitching. The following steps are the noise-constrained realization (NCR) and the ROMAGAL map-making, which are both described in detail in Section 4. A summary of the entire pipeline is shown in Fig. 2.

3.1 Corrupted time-ordered data

The TOD can be partially corrupted by the random hitting of charged particles (cosmic rays), which produce strong and spiky signal variations called glitches. Two different effects on the TOD can be

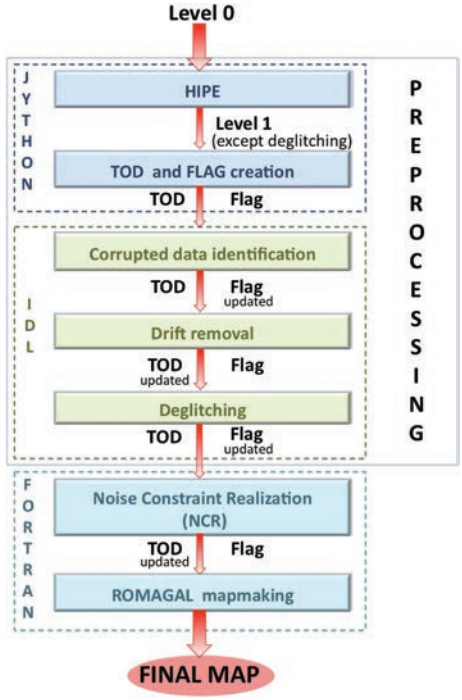

Figure 2. Schematic representation of the Hi-GAL data reduction pipeline.

identified: the glitch corrupts a single or a few consecutive samples, generating spiky changes along the TOD. This is the most common effect. In Section 3.3, we describe how to detect and mask the data for PACS, as well as to mask any possible residual glitches for SPIRE. Very powerful glitches, however, make the detector signal unstable for a considerable interval of time (see, for eaxmple, Billot et al. 2010). This effect depends on the glitch amplitude and on the time response of the bolometers. These events affect a much larger part of the TOD, which cannot be used for map-making.

The impact of these events results in a bias on the bolometer time line with a low-frequency drift, which can involve a considerable number of samples (as shown in Fig. 3).

In Fig. 3, the blue crosses represent the observed time line of one bolometer of the blue array.

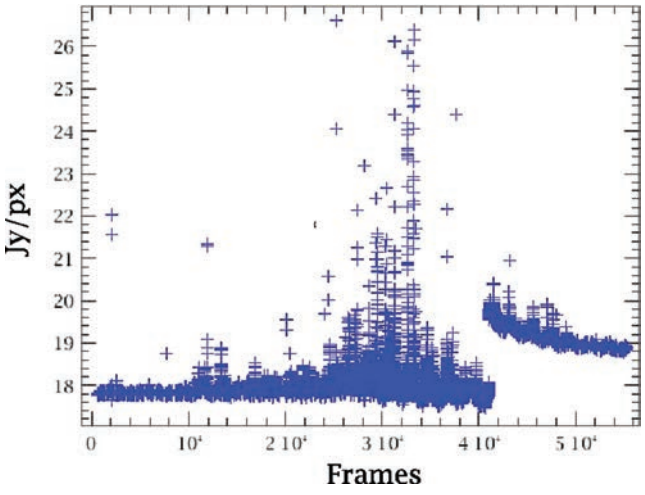

Figure 3. Time line of a PACS blue bolometer. The exponential decay illustrates the change in responsivity after frame 40 000 because of the impact of a powerful glitch.

Automatic identification of the (partially) corrupted TOD exploits the first derivative of the TOD to detect extraordinary 'jumps' in the signal. In order to determine what portion of the TOD is to be flagged, the flagging tool exploits the fact that the response of the detector pixels that have been hit by a cosmic ray is mostly exponential. Data samples, ranging from the jump to the time sample at which an exponential fit reaches 98 per cent of the signal level before the event, are identified as bad data and stored as such in the corresponding flag file. In case the exponential fit does not reach 98 per cent of the starting value before the end of the TOD, then all data starting from the hit will be flagged, as is the case in Fig. 3. This procedure is applied in the cases of both a change in responsivity or an alteration in the detector offset. For alteration in the detector offset, we estimate the fit with an exponent equal to 0. The described procedure was adopted to process both PACS and SPIRE data.

3.2 Drift removal

Having identified the corrupted data, we proceed to the elimination of changes in responsivity over time. The procedures are, in principle, identical for PACS and SPIRE data; the only differences are in the different compositions of the detector arrays and the data acquisition of the two instruments.

The signal in the PACS TOD that exit HIPE does not represent the true sky, but it is dominated by the telescope background and the (unknown) zero level of the electronics. The electronics further introduce significant pixel-to-pixel offsets. For each TOD, we mitigate the effect of pixel-to-pixel offset by calculating and then subtracting the median level for each pixel from each readout. This ensures that all pixels have a median level equal to 0. The median value is preferred over the mean because the median is a much better representation of the sky + telescope background flux, and it is much less sensitive to signals from astrophysical sources.

Also, the subtraction of this offset from all TOD does not alter the signal in the final map, but it introduces only a global offset, constant over the entire area covered in the observation. However, it should be kept in mind that bolometers are inherently differential detectors, which bear no knowledge of an absolute signal value. Besides, any optimized map-making methods, such as the one we employ (see Sections 4), produce maps with an unknown offset value, which needs to be independently calibrated. So, it is important to reduce all the bolometers to the same median value, regardless of its amount. The pixel-to-pixel median subtraction has the effect seen in Fig. 4. Diffuse emission and compact sources are clearly visible in the frame.

Still, when plotting a detector pixel time line we see that the signal decreases systematically from the start to the end of the observation.

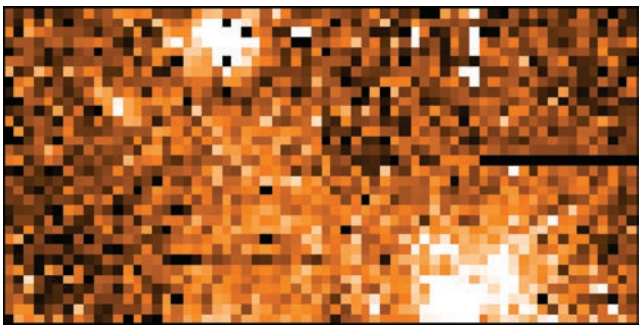

Figure 4. Blue PACS frame after the median subtraction on each pixel. The diffuse emission and compact source are visible in the frame.

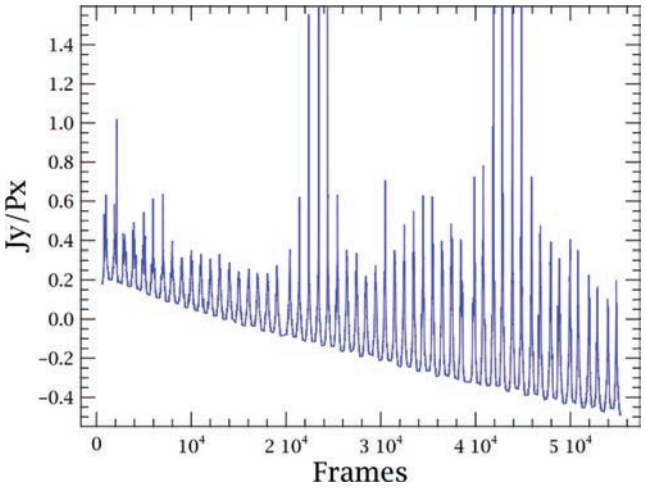

Figure 5. Median behaviour computed on the whole array for each frame. The slow-drift behaviour is a result of the electronics and the thermal bath.

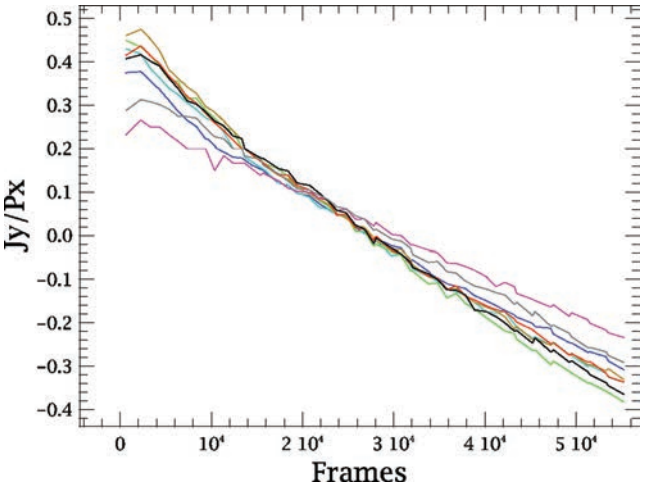

Figure 6. Interpolation of the minima of the median evaluated on every scan-leg. Each curve refers to a PACS subarray. The curves mimic the same behaviour as slopes, because of the different subarray electronics.

This trend is likely to be the result of a combination of small changes in the thermal bath of the array and of small drifts in the electronics. The former affect the entire array, while the latter affect the subunits of the detector array (in fact, PACS blue is divided into eight units and PACS red into two, electronically independent units; Poglitsch et al. 2008). The drift is then a combination of these two effects: the drifts of the entire array and the drift of a single subunit. These effects are dominant with respect to the $1/f$ noise pattern, which is described in Section 4.1.

In principle, it is not a trivial task to decide which drift has to be subtracted first, that is, the drift from the thermal bath (affecting the entire detector array) or the drift from the readout electronics (affecting subarrays differently). Ideally, both should be subtracted, if only it were possible to separate each component accurately, as the net effect in the data is the sum of both.

Our methodology for removing the correlated signal drifts (on both the bolometer module/unit level and the array level) is based on tracing the low signal envelope of the unit or array median levels. In Fig. 5, this envelope is the curve defined by the lowest signal values, which is estimated as follows.

(i) We compute the median value of the entire bolometer array/unit for each bolometer readout. Fig. 5 shows one example for the entire array.

(ii) The median values thus obtained are segmented and grouped by scan-legs. Each scan-leg is composed of ∼1000 frames and we have observed 54 scan-legs for each $2^\circ \times 2^\circ$ Hi-GAL field.

(iii) For each scan-leg, we compute the minimum value of the array/unit medians.

(iv) The resulting set of minimum median values for all scan-legs is fit with a polynomial.

The median value for each array/unit readout is chosen because it is closest to the actual sky + telescope background. However, as is clearly seen in Fig. 5, in the presence of strong astrophysical sources, the median value is incorrect for our purposes. The strong sources appear as signal spikes in Fig. 5. Hence, we take the next step of finding the minimum value from the set of medians belonging to a single scan-leg. The idea is that at some point during the scan-leg the median was indeed a true representation of the local sky + telescope, and relatively free of source emission. This step allows us to reject the sources at the expense of degrading our time-resolution to scan-leg duration (∼240 s). The polynomial fit allows us to estimate the drift behaviour at the same time resolution as the PACS signal (5 and 10 Hz for the 70- and 160-μm bands, respectively). We further note that the correlated signal drift is relatively flat over a single scan-leg. Hence, the minimum value is not significantly affected by the presence of the monotonic signal drift itself in the first place.

The minimum median method discussed above removes background levels from spatial emission structures that are of the order of, or larger than, the scan-legs, yet it preserves the spatial structures smaller than the scan-legs. In essence, information about the absolute calibration zero-point (Miville-Deschênes & Lagache 2005) is lost, but all spatial structures within our map boundaries are preserved.

In Fig. 6, we report the minimum median values of each subarray. The common downward trend is a result of the (common) thermal bath drift, while the different readout electronics are responsible for the differences in the subarray slopes.

Therefore, we have decided to subtract the subarray drifts in order to consider both the thermal bath and the readout electronics behaviours, but separately for each subunit.

Once the individual subarray drift is removed, the remaining dispersion on the whole array is only a residual scatter because of the intrinsic efficiency of the removal tool, as shown in Fig. 7.

The SPIRE array detectors are not divided into subarrays. So, every procedure that has to be run eight or two times for PACS data is only performed once per SPIRE band. SPIRE uses blind bolometers (five in total) as thermistors to evaluate the most relevant correlated noise component: the bath temperature fluctuations. A standard pipeline module uses this information to perform an effective removal of the common drift present along the scan observation. HIPE also corrects for the delay between the detector data and the telescope position along the scan, using an electrical low-pass filter response correction. However, despite these initial (and very effective) corrections, we apply the drift removal tool to SPIRE data in the same way as for PACS data. We fit a polynomial to the minimum of the median of each scan-leg (calculated over the entire detector array), which we then subtract to all the TOD. Experience of the data shows that a residual long time drift is often present in the SPIRE data.

Finally, when removing drifts it is important to know how the observational scans are oriented. In fact, as the Galactic plane is very bright, scans across the plane will give rise to an increase of the

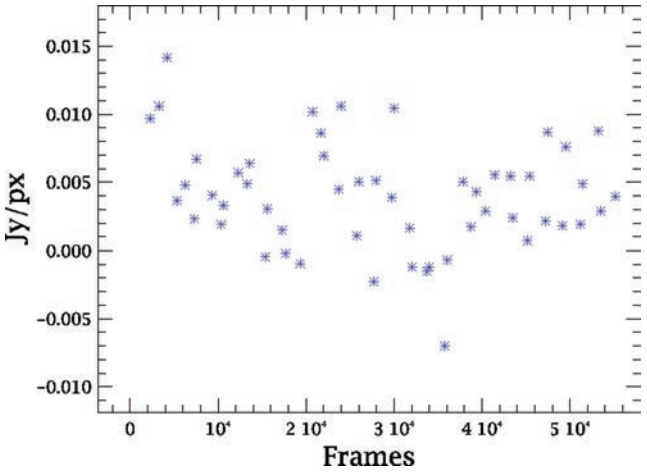

Figure 7. Minima of the median evaluated on the whole PACS blue array after the subarray drift subtraction. The dispersion is a result of the intrinsic efficiency of the drift-subtraction tool. There is no residual behaviour and the scatter is one order of magnitude below the value of the initial drift.

signal, on top of the general drift. However, when the scan is almost parallel to the plane of the Galaxy, the signal can be dominated by its bright emission, also on the evaluation of the minima of the median. In this case, only the points that are not affected by the signal are considered for the fit.

Because the procedure is not automatic, care has to be used when choosing what polynomial to fit and to subtract from the data, in order not to eliminate the genuine signal from the sky. In our experience, the best choice is a first- or a second-degree polynomial, depending on the signal behaviour observed.

A higher polynomial degree can be necessary when part of the drift has to be extrapolated in order to avoid any signal contamination. An example for one subarray is shown in Fig. 8.

Each scan-leg is bigger than the overlapping region of the two scan directions for the scanning strategy adopted by Hi-GAL (see Fig. 1). Because the Hi-GAL fields are squared regions, the slowly traversed direction of the AOR within the overlapping region has a length comparable with the scan-leg. Thus, we assume that even if

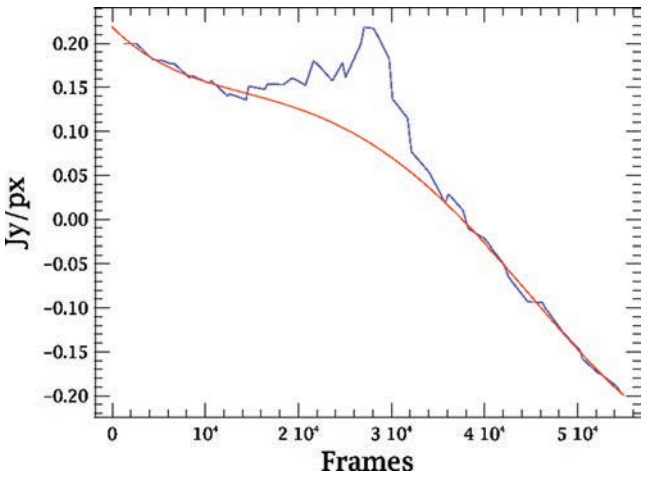

Figure 8. Blue curves: interpolation of the minima of the median for one subarray of the PACS blue channel when the scan direction is almost parallel to the Galactic plane. Red line: the fit in which we choose to evaluate the drift without considering the minima affected by the Galactic emission (central bump).

there is a signal gradient along the slowly traversed direction of the AOR, it is not filtered out by the subtraction of the array medians.

3.3 Deglitching

To remove outliers in the time series of the bolometers, we exploit the spatial redundancy provided by the telescope movement. This ensures that each sky pixel of the final map is observed with different bolometers. The detection of outliers is performed using the standard sigma-clipping algorithm. Given a sample of N values, estimates for the mean and for the standard deviations are first derived. Then, all the values that differ from the mean by more than n standard deviations are considered outliers, and these are removed.

For this algorithm, the choice of n, the parameter that defines the threshold above which a value is considered an outlier, is usually arbitrary. A certain n is chosen, very often equal to 3, without any statistical justification. Recently, Pezzuto (2010) has derived a formula that, starting from the properties of the error function for a Gaussian distribution and exploiting the discreteness of the number of available measures, relates n to the size of the sample. The formula is

$$n = -0.569 + \sqrt{-0.072 + 4.99 \log(N)}. \quad (1)$$

As a consequence, in the central region of the map, where the coverage (and so n) is high, the number of standard deviations is larger than in the outskirts of the map where the coverage is low. For instance, if a sky pixel has been observed with 40 bolometers, the above formula gives $n = 2.25$. So, once we have estimated the mean m and the standard deviations σ, all the values x_i such that ABS$(x_i - m) > 2.25\sigma$ are flagged as outliers. If a pixel has been observed with 20 bolometers, the threshold lowers to 1.96σ. The detection of outliers is performed in this way for both instruments. However, for SPIRE, as explained before, we also make use of the standard deglitching algorithm (wavelet-based) implemented in the official pipeline. We found some weak glitches left in the SPIRE TOD so we decided to run our deglitching algorithm on SPIRE data also.

The number of glitches found is, on average, about 15 per cent, a value that is likely larger than the real percentage. For PACS, we are now working on a different way to associate each bolometer to the sky pixels, taking into account the finite size of the sky pixels. For the first test cases we run, the percentage of detected glitches is now around 5–6 per cent.

4 ROMAGAL MAP-MAKING ALGORITHM

The ROMAGAL algorithm is based on a generalized least-squares (GLS) approach (Lupton 1993). Because the TOD are a linear combination of signal and noise, we can model our data set $\boldsymbol{d}_k$ for each detector k as (Wright 1996):

$$\boldsymbol{d}_k = \mathbf{P}\boldsymbol{m} + \boldsymbol{n}_k. \quad (2)$$

Here, $\mathbf{P}$ is the pointing matrix, which associates with every sample of the time line a direction in the sky, $\boldsymbol{m}$ is our map estimator of the 'true' sky and $\boldsymbol{n}_k$ is the noise vector.

The observed sky, $\mathbf{P}\boldsymbol{m}$, is the map estimator of the 'true sky' convolved with the instrumental transfer function and the optical beam. However, for a circularly symmetric beam profile, $\boldsymbol{m}$ is a beam-smeared, pixelized image of the sky.

In this case, the pointing matrix only has a non-zero entry per row corresponding to the sky pixel observed at a given time. Because the beam profiles for PACS (Poglitsch et al. 2008) and SPIRE (Griffin

et al. 2008b) are only weakly asymmetric, we can put ourselves in the symmetric beam approximation. Note that the transpose of the $\mathbf{P}$ operator performs a data binning (without averaging) into the sky pixels.

Equation (2) holds only if the noise vector of each detector $\boldsymbol{n}_k$ is composed of statistical random noise, with a Gaussian distribution and null average. All the relevant systematic effects (offset, glitches) then have to be removed with accurate data pre-processing before map production, as explained in Sections 3.

The formalism can be easily extended in the case of multidetector analysis. In this case, the vector $\boldsymbol{d}$ contains the data relative to each detector. In fact, we also have to take care to upgrade the noise vector $\boldsymbol{n}$, according to the correct noise value for each detector.

The GLS algorithm produces minimum noise variance sky maps. The noise properties for each detector have to be estimated first, and provided as input to the algorithm, as described in Section 4.1.

The GLS estimate for the sky, $\tilde{\boldsymbol{m}}$, is (Natoli et al. 2001)

$$\tilde{\boldsymbol{m}} = (\mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\boldsymbol{d}. \quad (3)$$

Here, $\mathbf{N} = \langle \boldsymbol{n}\boldsymbol{n}^{\mathrm{T}} \rangle$ is the noise covariance matrix, which takes into account the noise time correlation between different samples. Such a correlation is particularly high at low frequencies because of the $1/f$ (or long memory) noise. For uncorrelated noise (or white noise), the $\mathbf{N}$ matrix becomes diagonal, and the problem is greatly simplified. If we further assume stationary uncorrelated noise, equation (3) reduces to

$$\tilde{\boldsymbol{m}} = (\mathbf{P}^{\mathrm{T}}\mathbf{P})^{-1}\mathbf{P}^{\mathrm{T}}\boldsymbol{d}. \quad (4)$$

$\mathbf{P}^{\mathrm{T}}\mathbf{P}$ is the number of observations of a pixel of the map. So, we average the different TOD values regarding pixels, assigning the same weight to each sample. We refer to this map estimate as 'naïve' or 'binned' in the following.

When non-negligible noise correlation is present, as in the case of PACS (Poglitsch et al. 2008) and SPIRE (Schulz et al. 2008), equation (3) must be solved. This is a challenging computational task as it requires, in principle, the inversion of the large (of the order of the number of pixels in the map) matrix $\mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{P}$, which is the covariance matrix of the GLS estimator (Lupton 1993). One key simplifying assumption is to invoke the fact that the noise is stationary. In this case, the $\mathbf{N}$ matrix has a Toeplitz form, which can be approximately treated as circulant, ignoring boundary effects (Natoli et al. 2001). A circulant matrix is diagonal in Fourier space, and its inverse is also circulant. So, the product between $\mathbf{N}^{-1}$ and a vector is a convolution between the same vector and a filter provided by any of the rows of the matrix. In the following, we refer to any of these rows as a noise filter. Its Fourier transform is the inverse of the noise frequency power spectrum.

Considering the conditions listed above, the GLS map-making algorithm performs the following operations, starting with rewriting equation (3) in the form

$$(\mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\mathbf{P})\boldsymbol{m}_0 - \mathbf{P}^{\mathrm{T}}\mathbf{N}^{-1}\boldsymbol{d} = \boldsymbol{r}. \quad (5)$$

Here, $\boldsymbol{m}_0$ is the starting map used at the first iteration, generally the naïve map.

The $\mathbf{P}\boldsymbol{m}_0$ product projects the map on to a time line: (i) the application of $\mathbf{N}^{-1}$, which is a convolution that can be performed in Fourier space; (ii) the application of $\mathbf{P}^{\mathrm{T}}$, a step that projects the convolved time line back into a map.

The second step performs the convolution with the filter (applying $\mathbf{N}^{-1}$ to the data vector $\boldsymbol{d}$ in Fourier space), and then the projection of the convolved time line into a map (applying $\mathbf{P}^{\mathrm{T}}$ to the product $\mathbf{N}^{-1}\boldsymbol{d}$).

Then, we need to evaluate the residual $\boldsymbol{r}$. If the residual is higher than a fixed threshold, it is used to produce a new map, $\boldsymbol{m}_1$, as described in Hestenes & Stiefel (1952). This map is considered instead of $\boldsymbol{m}_0$ for evaluating equation (5) again, until convergence. This is achieved by running a conjugate gradient algorithm, an iterative method that is useful to obtain a numerical solution of a system (Hestenes & Stiefel 1952), until convergence is reached with a residual lower than the threshold.

The algorithm outlined is the same as that described in Natoli et al. (2001) – 'unroll, convolve and bin' – and it is implemented in the ROMAGAL code. In Section 4.1, we explain the strategy employed to estimate the noise filters used by ROMAGAL directly from in-flight data.

4.1 Noise estimation

In order to estimate the noise filters for ROMAGAL, we need to investigate the statistical properties of the noise in the time lines. Data are mostly affected by two types of statistical noise: $1/f$ noise, resulting from both the electronics and the thermal background radiation from the telescope or the instruments, and photon noise (see Poglitsch et al. 2008; Schulz et al. 2008).

The detector $1/f$ noise arises in the electronic chain and it particularly affects regions with low signal-to-noise ratio (S/N), where only diffuse emission is present. In these regions, it can be of the same order of magnitude as the signal, or even higher. In these cases, a GLS treatment is particularly effective.

Photon noise is a result of statistical fluctuation in the photon counts. This process follows a Poissonian statistic, so the S/N is proportional to the square root of counts. Because a Poisson distribution tends to be Gaussian for large numbers, we can approximate photon noise as Gaussian on the map if the number of counts is large enough. As bolometers are organized into a matrix and submatrix, the signal of a bolometer can be correlated with the signal of another, generally adjacent, bolometer. These effects can be both statistical and deterministic. We have already described how to remove the deterministic common mode from the TOD (e.g. thermal bath variations; see Sections 3).

One possible source of statistical cross-correlated noise is the crosstalk between bolometers – the signal of one pixel might contaminate the signal of its neighbours through capacitive inductive coupling, which generates a common mode called 'electrical crosstalk'. In contrast, 'optical crosstalk' is a result of diffraction or aberrations in the optical system, which mean that an astronomical source can fall on to an inappropriate detector (Griffin et al. 2008a).

We then analyse the residual contribution of the statistical component of the correlated noise. We have found that the residual correlated noise level in each pixel is negligible with respect to the intrinsic detector noise level for both the PACS and SPIRE instruments, as described in the following.

In principle, the noise properties vary significantly across the array and we have had to estimate the noise power spectrum for each bolometer. To do this, we have processed the 'blank' sky mode (i.e. filled with negligible contribution from sky signal) data acquired during the PV phase.

In Fig. 9, we show a typical noise spectrum estimated for a pixel of the 160 μm PACS band (black) and the cross-spectrum between two adjacent bolometers (red). The cross-spectrum evaluates the impact of the cross-correlated noise in the frequency domain between two different bolometers. The level of the cross-correlated noise is at least four orders of magnitude below the autocorrelated noise power spectrum of each pixel. Note that this means we do not see any

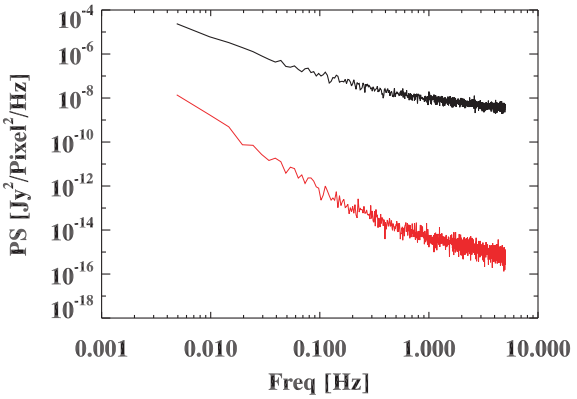

Figure 9. Black line: typical noise spectrum of a PACS 160-μm detector, estimated on blank-sky data. Red line: cross spectrum between two detectors of the same subarray. The level of the cross-correlated noise is significantly below the noise level of each single bolometer, so we can reasonably neglect it.

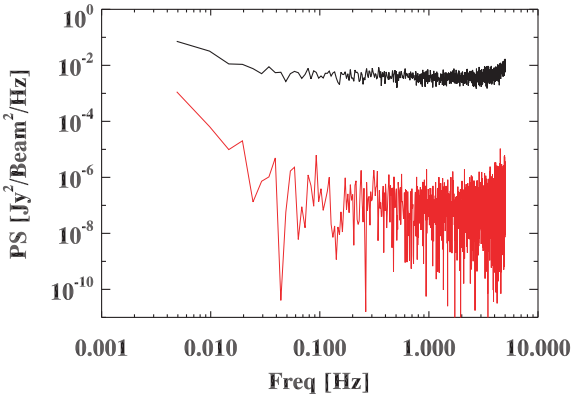

Figure 10. Same as Fig. 9 for a SPIRE 250-μm bolometer. For SPIRE also, the noise level of the cross-spectrum is reasonably negligible with respect to the autospectrum level.

relevant cross-correlated noise, despite the fact that crosstalk can be present in the time line.

In Fig. 10, we show the noise power spectra of the 250-μm SPIRE band bolometers. Also, in this case, the cross-spectrum is negligible.

The noise spectra of both PACS and SPIRE display low-frequency noise excess ($1/f$). For SPIRE spectra (Fig. 10), a high-frequency rise is also evident, which is because of the deconvolution of the bolometric response function. PACS spectra do not show this behaviour because the bolometer transfer function is not deconvoluted by the standard pipeline.

4.2 From ROMA to ROMAGAL

The ROMAGAL GLS code has been optimized to recover the features in the Hi-GAL fields with great accuracy.

Hi-GAL observes the Galactic plane where the dynamic range of sources spans over several orders of magnitudes. This puts great constraints on the map-making algorithm (i.e. both the weak diffuse emission and the bright compact sources in, for example, star-forming regions have to be recovered with great accuracy). The signal often exhibits steep gradients that are hard to follow for the GLS solver, which relies on the assumption that the sky signal does not vary significantly within a resolution element (see Section 5.3). At the same time, several systematics affect the data set. As explained previously, many of these are cured at the pre-processing level. However, their removal generates a conspicuous amount of transient flagging, which must be correctly handled by the GLS code.

Table 1. Time and minimum RAM amount required from ROMAGAL for each PACS and SPIRE band using eight BLADE processors.

Band	Total time (s)	RAM (Gb)
70 μm	∼1400	16
160 μm	∼1000	8
250 μm	∼180	4
350 μm	∼130	1
500 μm	∼100	1

The core of ROMAGAL is the same as the ROMA code (de Gasperis et al. 2005), where the input–output routines have been greatly modified to adapt to the HIPE-generated data set. ROMAGAL inputs are the TOD generated by HIPE, pointing and transient flagging. These have the same format for both PACS and SPIRE. ROMAGAL outputs are fits files containing the optimal map, arranged in standard gnomonic projection routines. The code is written in FORTRAN 95, and it relies on the MPI library for parallel calls. It runs on the Hi-GAL dedicated machine, called BLADE, a cluster of 104 processors at 2.5 GHz each and with 208 Gb RAM in total. Its nodes are interconnected with MPI-infiniBAND. The machine is located at the Institute of Physics of Interplanetary Space (IFSI), Rome.

As explained earlier, the computation of equation (3) is not feasible, because of the size of the map's covariance matrix. However, we assume the noise of each Hi-GAL field to be stationary in order to set up a fast Fourier transform (FFT) based solver, built upon a conjugate gradient iterative algorithm (see Section 4). Such a scheme can estimate the final maps with a precision of the order of $\epsilon = 10^{-8}$ in ∼150 iterations for Hi-GAL. The ROMAGAL computational time scales linearly with the size of the data set, but only weakly with the number of pixels in the output maps. The scaling with the number of processors is highly optimal in the range of cores required for the Hi-GAL analysis (<50). For the largest channels (PACS blue band), a final GLS map of a $2^\circ \times 2^\circ$ field requires about 16 Gb of RAM and 1400 s on eight cores. Because of the high number of array pixels (2048), this channel is the largest data set to analyse, as well as the most demanding in terms of computational resources. Further information on resource consumptions can be found in Table 1.

4.3 Optimal treatment of transient flagging

As mentioned above, the time lines are inspected for bad data samples, which must be excluded from map-making as part of the pre-processing pipeline. Bad data can arise for a variety of reasons. They are generally caused by transient events, either unexpected (e.g. glitches, anomalous hardware performance) or expected (e.g. detectors saturating because of a bright source, observation of a calibrator). Once identified, a flag record is generated and stored for these anomalous events, so that their contribution can be safely excluded from map-making. Flags in the TOD pose a potential problem for ROMAGAL because its solver is based on the FFT, as discussed in Section 4.2. The FFT requires the time line to be continuous and uniformly sampled. As noise in the PACS and SPIRE data is correlated, just excising the flagged samples to fictitiously create a continuous time line would interfere with noise decon-

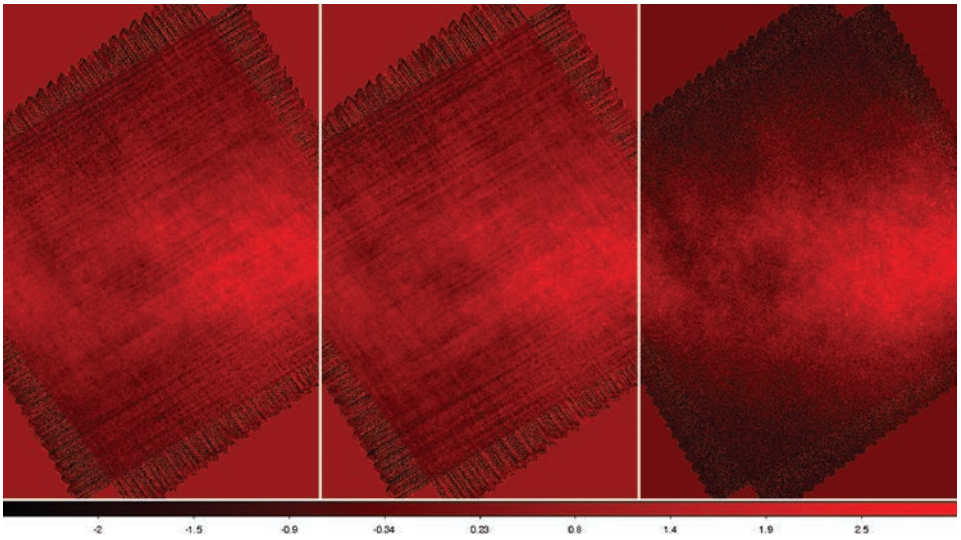

Figure 11. The results obtained in a noise-dominated regime (normal Hi-GAL noise is amplified by a factor of 100). The left panel is the map obtained by replacing flagged data samples with null values (clearly it does not work). The middle panel is the map obtained by replacing data samples with unconstrained noise realization (this does not work either). The right panel shows the map obtained using our NCR code (this does work).

volution; thus, it is not a safe option. Instead, we advocate using a suitable gap-filling technique. The rest of this section is mostly devoted to defining which technique, among the various options for gap-filling, is best suited for the Hi-GAL maps.

It is important to realize that the output map will depend on the content of the flagged sections of the TOD, even if these values are not associated with any (real) map pixel. This is because of the convolution performed within the solver, which 'drags' out data from the flagged section of the time line, even if, when the data are summed back into a map, the $\mathbf{P}$ operator is applied only to the unflagged samples. As there is no control of the content of the flagged section of the time line, a type of gap-filling must be performed. We have tested different recipes for gap-filling, making extensive use of simulations. We have treated the signal and noise components of the time lines separately, running noise-dominated and signal-dominated cases separately. This is because the behaviour towards the flags of the two components is different, as shown below.

The simplest form of gap-filling is to remove the content of the flags altogether, replacing the flagged sections with a nominal constant (usually null) value. This works very well on a signal-only simulation of the Hi-GAL field. However, it fails dramatically when noise is present, as evident from Fig. 11 (left panel), where a markedly striped pattern in the reconstructed map is seen (in this simulation, the Hi-GAL noise has been amplified by a factor of 100 to make its pattern more evident). The reason for this failure is readily understood. The GLS map-making employed requires noise stationarity (see Section 4.2), which is obviously not preserved by zeroing the gaps. A less obvious effect is that even if gaps are filled with a noise realization with the correct statistical properties, but unconstrained, the GLS map-making is also bound to fail, as shown in the middle panel of Fig. 11. A noise realization is said to be constrained when it respects certain boundary conditions (Hoffman & Ribak 1991), which in our case are represented by the noise behaviour outside the gap. Although unconstrained noise inside the gap has the correct statistical properties, it creates a border discontinuity that causes the GLS map-maker to behave suboptimally (Stompor et al. 2002). We have employed a code to create Gaussian NCRs, based on the algorithm set forth by Hoffman & Ribak (1991). The code uses as input the noise pattern and statistical properties, as measured from the time lines. The results for noisy simulations are excellent, as shown by the third (rightmost) panel in Fig. 11. Note, however, that Fig. 11 refers to a noise-dominated simulation. We now discuss the effect of a non-negligible signal contribution (a far more realistic case for Hi-GAL).

We have verified that the presence of a non-negligible signal in the time lines does not affect the results, provided that the NCR is performed using the underlying noise field as a baseline. However, measuring the latter in the presence of a signal is impractical. It would be significantly simpler if the NCR were run directly on the time lines themselves, thus constraining the (fake) noise realization within the gap to the outside noise plus signal (true) values. Unfortunately, this creates a problem for the Hi-GAL data: large signal jumps are present in the field, and the resulting gap-filling realizations are affected in a non-negligible manner by the boundary conditions, at least with the present version of ROMAGAL. This behaviour is different from what happens for experiments aimed at the cosmic microwave background (see, for example, Masi et al. 2006) where NCR codes are routinely run on the time lines as they are. In order to find a workaround that would spare us the inconvenience of estimating the underlying noise field to serve as the NCR input, we have modified the flag treatment of the ROMAGAL code, as explained in the following.

The original version of ROMAGAL makes use of a single extra pixel (dubbed the 'virtual pixel') to serve as a junk bin, where the contents of the gaps are sent when applying the $\mathbf{P}^{\mathrm{T}}$ operator within the ROMAGAL solver. This approach, as stated above, works excellently in presence of both signal and noise, irrespective of their relative amplitude, provided the NCR code assumes the underlying noise field as a baseline to perform the realization. In order to relax this assumption, we have modified ROMAGAL to take into account not a single virtual pixel but an entire virtual map. In other words, we introduce a virtual companion for each pixel of the map, and we use it as a junk bin to collect the output from the gaps they correspond to. The hope is to redistribute the content of the flagged sections more evenly, preventing artefacts. With this approach, we

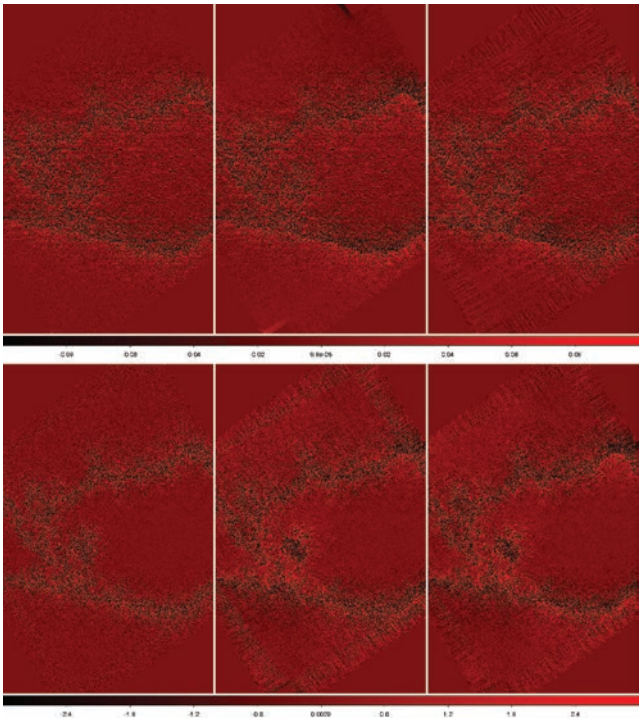

Figure 12. Top row: For a signal-dominated case, we show the relative differences between the input simulated map and the output obtained with ROMAGAL (in the 'virtual map' mode) with NCR performed on the underlying noise (left), on the original signal plus noise TOD (middle) and without NCR, after replacing the gaps with null values. The latter case is clearly striped, but no signal-related artefacts are present. Bottom row: For a noise-dominated case, we again show the relative difference versus the input map obtained by ROMAGAL with the 'virtual map', assuming NCR on underlying noise only (left), without NCR (middle) and with NCR on the signal plus noise TOD (right). As expected, the left panel shows the best residuals, but the right panel appears to be a good compromise (see also the text).

have obtained satisfactory results when the NCR code is run on the original (signal plus noise) time lines, as shown in Fig. 12.

We can summarize our findings, as follows.

(i) Using a NCR realization code is *always* necessary in order to avoid artefact striping effect into the map.

(ii) If the NCR code is run on the underlying noise-only time lines (which are cumbersome to estimate), we obtain the best-quality output map, with no signal-related artefacts and a noise standard deviation that is lower (by a factor of $\sim$2) with respect to the case in which no NCR is performed.

(iii) Running the NCR on the original time lines is possible with the 'virtual map' approach. No signal artefacts are detected in the difference maps, and the advantage in terms of noise standard deviation with respect to the case of no NCR is still present, of the order of 10–20 per cent, on average.

We have therefore chosen the NCR + 'virtual map' approach as the baseline for the Hi-GAL pipeline.

5 ROMAGAL MAPS

In this section, we analyse the final map obtained by running our dedicated pipeline. We analyse the point spread function (PSF) for the five bands in order to fix the resolution of the ROMAGAL maps. We compare the final GLS map with the naïve map, and we point out the differences and the capability to recover the diffuse emission from the GLS map. Finally, we discuss the noise residuals on the maps.

5.1 Point spread function and pixel size

The angular resolution (pixel size) of the final map is a free parameter, and choosing it is a compromise between competing requirements. A small pixel size ensures a correct sampling of the PSF.

Indeed, assuming a Gaussian profile for the PSF (which is reasonable, as discussed in the following), the Nyquist theorem imposes that to better sample a two-dimensional image we need to set a pixel size that is, at most, one-third of its FWHM value.

However, a pixel size that is too small can cause a loss of redundancy. The redundancy can be useful to reduce the white-noise level and (even) some non-observed pixels in the final map.

The diffraction limited beam of PACS at 70 μm is 5.2 arcsec. Thus, we should build the map with a pixel size of at least 1.8 arcsec. However, because of the limited bandwidth available for transmission, especially in PACS/SPIRE parallel mode, PACS frames are co-added on-board the satellite before broadcasting. For the 70-μm channel, a co-addition of eight consecutive frames is applied by on-board software. Because the acquisition rate is 40 Hz and the scanning speed for Hi-GAL is set to 60 arcsec s^{-1}, two close frames are snapshots of the sky acquired 1.5 arcsec apart. Because of co-addition, the satellite provides one averaged frame every 12 arcsec; in spatial coordinates, this is twice the beamwidth of the PACS blue channel. Thus, the measured PSF is not the diffraction-limited beam, but it results in an elongation along the scan direction because of the convolution of the co-addition with the beam. As shown in Lutz (2009), the observations of Vesta and αTau with the blue channel gave evidence of a FWHM equal to 5.86 × 12.16 $arcsec^2$ as a result of a two-dimensional Gaussian fitting, elongated in the scan direction.

The PACS 160-μm channel is also affected by the averaging on-board, but only four frames are co-added together. The nominal instrumental beam is 12.0 arcsec, while that measured is 11.64 × 15.65 $arcsec^2$ (Lutz 2009), elongated along the scan direction. However, in this case we can sample the beam without issues, and the effect of co-addition on the final map is negligible.

For the Hi-GAL fields, the scanning strategy consists of two orthogonal AORs. Therefore, the redundancy regularizes the PSF, resulting approximately in a two-dimensional symmetric Gaussian profile, as shown in Table 2.

According to the values reported in Table 2, we observe quasi-symmetric beams with an averaged ellipticity of less than 15 per cent for both blue and red channels, and the axis is oriented randomly with respect to the scan direction.

Table 2. Nominal and map-measured (two AORs) beams and ellipticity of each band.

Band	Nominal beam (arcsec)	Measured beam (arcsec)	Ellipticity	Pixel size (arcsec)
70 μm	5.2 × 5.2	∼9.7 × ∼10.7	14.6 per cent	3.2
160 μm	12.0 × 12.0	∼13.2 × ∼13.9	14.7 per cent	4.5
250 μm	18 × 18	∼22.8 × ∼23.9	8.3 per cent	6.0
350 μm	24 × 24	∼29.3 × 31.3	8.8 per cent	8.0
500 μm	34.5 × 34.5	∼41.1 × ∼43.8	9.7 per cent	11.5

We choose a pixel size of 3.2 arcsec for the PACS 70-μm band, which samples the observed beam at the Nyquist frequency. Below this threshold, the diffuse emission areas become too noisy because of the low S/N. Similarly, we can choose a pixel size of 4.5 arcsec for red band without losing redundancy.

SPIRE does not suffer from on-board co-adding, and the detectors were built to reach the band diffraction limit. In-flight data show that the SPIRE beam is well approximated by a two-dimensional asymmetric Gaussian curve, with the major axis orientation independent of the scan direction, and an ellipticity no bigger than 10 per cent (see Sibthorpe, Ferlet & Schulz 2010). We set the pixel size for each SPIRE band equal to one-third of the nominal beam. In Table 2, we also report the beam measured in the SPIRE maps.

The average ellipticity we observe agrees with that found by Sibthorpe et al. (2010). In contrast, while the FWHMs for the two axes found by Sibthorpe et al. (2010) are in agreement with the nominal values (within the error), our measured beam results in a FWHM larger than the nominal value of ∼25 per cent.

5.2 Hi-GAL SDP results

The quality of the outcome can be qualified by the comparison between the igls maps with the naïve maps. In fact, the naïve map is the simple averaging of the signal recorded by every spatial pixel, and it represents 'the least compromised view of the sky'.

Because the TOD are created at the end of the pre-processing steps, when the data are a combination of only signal and $1/f$ noise, we expect the $1/f$ residuals in the naïve map as well as a 'pure' sky map produced by ROMAGAL.

Fig. 13 shows a comparison between the naïve map and the ROMAGAL map of the $l = 30^\circ$ field at 70 μm. The GLS code is capable of removing the $1/f$ residuals without losing any signal, both on bright sources and on diffuse emission.

In particular, we choose the following three main proxies.

(i) The difference between naïve and igls should show only a pattern resulting from the $1/f$ noise residuals in the binned map. The pattern of this low-frequency noise is recognizable as stripes superimposed on the sky signal in the naïve map. The stripes are the consequence of the $1/f$ noise resulting from the adopted scanning strategy.

Fig. 14 shows the difference between the igls and naïve maps. The $1/f$ noise is removed in the igls map but not in the naïve map. The residual stripes resulting from the low-frequency noise are clearly visible.

(ii) The source fluxes should be the same in both the igls and naïve maps.

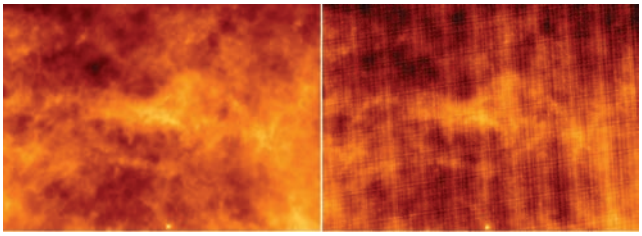

Figure 13. Left: The ROMAGAL map of the red PACS array, $l = 30^\circ$ field. Right: The naïve map, same band and field. The $1/f$ noise is evident on the naïve map, and its minimization is also evident on the GLS map without losing the signal of the diffuse component.

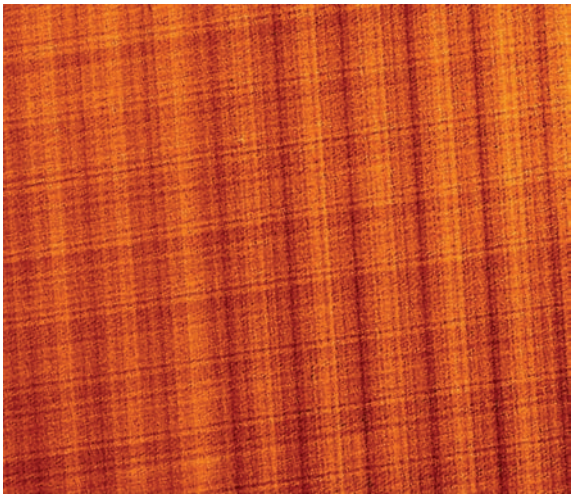

Figure 14. The map difference between the PACS 160-μm igls and naïve maps, with the same region as Fig. 13. The stripes resulting from the $1/f$ removal made in the igls are evident.

Table 3. The rms compared from the GLS and naïve map on both the SDP observations, for PACS bands, measured on a background region of 50 × 50 pixels. The last column gives the ratio between the naïve and GLS rms.

Band	rms igls (MJy pixel^{-1})	rms naïve (MJy pixel^{-1})	Ratio
PACS $l = 30^\circ$ field			
70 μm	0.0085	0.026	∼3.1
160 μm	0.047	0.102	∼2.2
PACS $l = 59^\circ$ field			
70 μm	0.004545	0.02208	∼4.9
160 μm	0.01899	0.03586	∼1.9

This item is quantified by the map difference, where the pattern is only a result of the noise without any residual signal, except across a very brilliant source. This last effect is discussed in Section 5.3.

(iii) A statistical analysis of the background noise level should show a decrease of the rms value in the igls map with respect to the naïve map.

In Tables 3 and 4, we report the rms residuals of the PACS and SPIRE maps, respectively, calculated in a diffuse emission area of each map. As the flux between the maps is conserved, a decrease of the rms noise level ensures an increase of the S/N in the ROMAGAL maps.

The ratio between the naïve and igls rms shows an improvement of a factor of ∼2–5 in the PACS ROMAGAL maps, and of a factor ∼1–2 in the SPIRE case. The difference is mostly the result of an intrinsically different $1/f$ noise level.

5.3 Consistency of data analysis assumptions

One of the assumptions of ROMAGAL, as with all Fourier-based GLS map-making codes, is that the underlying sky signal is constant within a chosen resolution element. If this is not the case, artefacts (stripes) will be generated in the final map, contributing to the

Table 4. The rms compared from the GLS and naïve map on both the SDP observations, for SPIRE bands, measured on a background region of 50 × 50 pixels. The last column gives the ratio between the naïve and GLS rms. The $1/f$ noise is less evident in the SPIRE bolometers compared to the PACS bolometer, but its effect is still remarkable.

Band	rms igls (MJy beam^{-1})	rms naïve (MJy beam^{-1})	Ratio
SPIRE $l = 30^\circ$ field			
250 μm	0.1749	0.2868	∼1.6
350 μm	0.1569	0.2302	∼1.5
500 μm	0.2659	0.4065	∼1.5
SPIRE $l = 59^\circ$ field			
250 μm	0.09857	0.1123	∼1.1
350 μm	0.0734	0.08164	∼1.1
500 μm	0.1073	0.2101	∼1.9

so-called pixelization noise (Poutanen et al. 2006). In the case of Hi-GAL, the situation is complicated by several effects, as follows.

(i) The on-board co-addition of samples. Each PACS 70-μm (160-μm) frame is the result of an on-board average of eight (four) consecutive frames, reducing the effective sampling frequency of the instrument (see Section 5.1). Thus, the sky signal is low-pass filtered only by partially effective data-windowing, rather than telescope response, leaving room for signal aliasing.

The map-making code is sensitive to aliasing as it works in Fourier space. The situation is worsened by the large dynamic range of the Hi-GAL fields, especially when scanning across bright sources.

(ii) The time bolometer response induces signal distortions along the scan. While, within HIPE, the SPIRE detector response is deconvoluted from the data (Griffin et al. 2008a), the same is not true for PACS. Redundancy in each pixel is obtained by scans coming from different directions, and thus the effect contributes further signal mismatch at the pixel level.

(iii) Pointing error. As analysed in detail in Garcia Lario, Heras & Sanchez-Portal (2007), the pointing performance of *Herschel* (which means the capability to assign coordinates for each sample in a given reference frame) can be affected by several pointing error effects. The main contributor is a result of the angular separation between the desired direction and the actual instantaneous direction, based on the position-dependent bias within the star-trackers.

The goal of the *Herschel* AOCS software is for the mismatch between real coordinates and assigned coordinates along the scan-leg to be smaller than 1.2 arcsec at 1σ (Garcia Lario et al. 2007). So, a 2σ event becomes significant compared to the PSF of the PACS blue band.

All of the above effects challenge the basic assumptions (no signal aliasing and no subpixel effect) under which ROMAGAL works. Our simulations suggest that signal aliasing contributes significantly more than the other two. The net result on maps is the striping of bright sources in the GLS maps. An example is shown in Fig. 15 for the PACS blue band.

It is important to notice that the stripes are present only around point-like sources (where, of course, the signal aliasing is more evident), regardless of their magnitude. However, magnitude influences the amplitude of the stripes. For a source peak that is only 10 times higher that the rms background, the intensity of the stripes are within 1σ from the background dispersion. For the more intense sources, the magnitude of the stripes in the pixels surrounding the PSF shape can be 100σ times away from the background value.

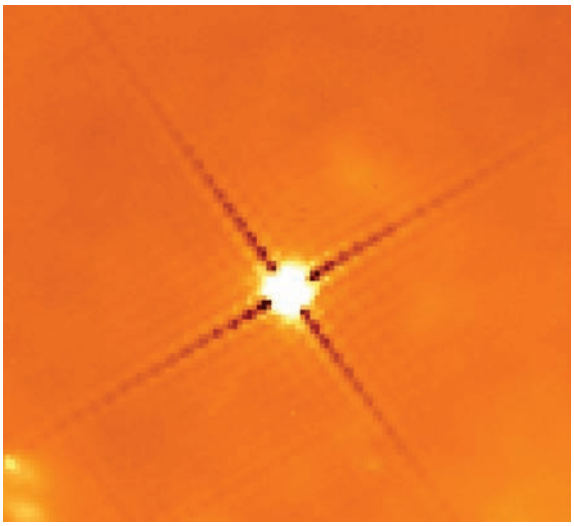

Figure 15. A zoom around a compact source for the PACS blue optimal map. The striping dragged along the scan directions is remarkable.

As these stripes are produced within the GLS solver, which performs repetitive convolutions along the scan directions, but they do not affect the naïve map, the obvious workaround is to implement dedicated map-making, which considers a different analysis around the bright sources.

A detailed accounting of the above effects and the enhanced map-making strategies to address them will be the subjects of a forthcoming paper (Piazzo et al., in preparation).

6 SUMMARY AND CONCLUSIONS

In this paper, we describe in detail all the steps in the processing of *Herschel* data, from the originally downloaded frames to high-quality maps, as used in the framework of Hi-GAL. The Hi-GAL data are taken in fast scan mode (60 arcsec s^{-1}) and simultaneously by PACS and SPIRE (parallel mode). We test our pipeline by reducing data from the SDP, and we present our results, taking as proxy for the quality of the final images their comparison with naïve maps.

We have divided the data processing into two distinct phases: pre-processing and map-making. Pre-processing aims to accurately remove systematics and random effects in the data, in order to prepare these for the ROMAGAL map-making algorithm, which implements the minimum variance GLS approach in order to minimize the noise in the data. It turns out that the NCR is a fundamental step in the pre-processing because ROMAGAL, as a FFT map-making code, needs continuous and regular data time series as input.

The noise residuals in the diffuse emission of the two test fields (SDP Hi-GAL data, two $2^\circ \times 2^\circ$ tiles centred on the Galactic plane at $l = 30^\circ$ and $l = 59^\circ$) show that we obtain optimal maps, removing glitches and systematic drifts, as well as minimizing the $1/f$ and white-noise components. The remaining effects, which do not affect the overall quality of the maps except across a bright source on the PACS 70-μm maps, are under investigation and will appear in a dedicated publication to be available shortly.

ACKNOWLEDGMENT

The authors would especially like to thank Göran Pilbratt for allowing them to use the PV phase blank data.

REFERENCES

Billot N. et al., 2010, in Holland W. S., Zmuidzinas J., eds, Proc. SPIE Vol. 7741, Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy V. SPIE, Bellingham, 774102
de Gasperis G., Balbi A., Cabella P., Natoli P., Vittorio N., 2005, A&A, 436, 1159
Garcia Lario P., Heras A. M., Sanchez-Portal M., 2007, Herschel Observer Manual
Griffin M. et al., 2008a, in Oschmann J. M., Jr, de Graauw M. W. M., MacEwen H. A., eds, Proc. SPIE Vol. 7010, Space Telescopes and Instrumentation 2008: Optical, Infrared, and Millimeter. SPIE, Bellingham., 70102Q
Griffin M. et al., 2008b, in Oschmann J. M., Jr, de Graauw M. W. M., MacEwen H. A., eds, Proc. SPIE Vol. 7010, Space Telescopes and Instrumentation 2008: Optical, Infrared, and Millimeter. SPIE, Bellingham, 701006
Griffin M. J. et al., 2010, A&A, 518, L3
Hestenes M. R., Stiefel E., 1952, J. Res. Nat. Bureau of Standards, 49, 409
Hoffman Y., Ribak E., 1991, ApJ, 380, L5
Lupton R., 1993, Statistics in Theory and Practice. Princeton Univ. Press, Princeton, NJ
Lutz D., 2009, PACS photometer point spread function, http://herschel.esac.esa.int/aotsreleasestatus.shtml
Martin P. G. et al., 2010, A&A, 518, L105
Masi S. et al., 2006, A&A, 458, 687
Miville-Deschênes M., Lagache G., 2005, ApJS, 157, 302
Molinari S. et al., 2010a, PASP, 122, 314
Molinari S. et al., 2010b, A&A, 518, L100
Natoli P., de Gasperis G., Gheller C., Vittorio N., 2001, A&A, 372, 346
Ott S., *Herschel* Science Centre, European Space Agency, 2010, arXiv e-prints
Peretto N. et al., 2010, A&A, 518, L98
Pezzuto S., 2010, poster presented at 'The Impact of *Herschel* Surveys on ALMA Early Science', ESO Garching
Pilbratt G. L. et al., 2010, A&A, 518, L1
Poglitsch A. et al., 2008, in Oschmann J. M., Jr, de Graauw M. W. M., MacEwen H. A., eds, Proc. SPIE Vol. 7010, Space Telescopes and Instrumentation 2008: Optical, Infrared, and Millimeter. SPIE, Bellingham, 701005
Poglitsch A. et al., 2010, A&A, 518, L2
Poutanen T. et al., 2006, A&A, 449, 1311
Schulz B. et al., 2008, in Duncan W. D., Holland W. S., Withington S., Zmuidzinas J., eds, Proc. SPIE Vol. 7020, Millimeter, Submillimeter Detectors and Instrumentation for Astronomy IV. SPIE, Bellingham, 702022
Sibthorpe B., Ferlet M., Schulz B., 2010, SPIRE Beam Model Release Note. ICC, London
Starck J., Murtagh F., Louys M., 1995, in Shaw R. A., Payne H. E., Hayes J. J. E., eds, ASP Conf. Ser. Vol. 77, Astronomical Data Analysis Software and Systems IV. Astron. Soc. Pac., San Francisco, p. 268
Stompor R. et al., 2002, Phys. Rev. D, 65, 022003
Wright E. L., 1996, paper presented at the IAS CMB Data Analysis Workshop (astro-ph/9612006)

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