Mon. Not. R. Astron. Soc. **357,** 420–428 (2005) doi:10.1111/j.1365-2966.2005.08675.x

The *INTEGRAL* spectrometer SPI: performance of point-source data analysis

P. Dubath,[1,2⋆] J. Knödlseder,[3] G. K. Skinner,[3] P. Connell,[4] I. Kreykenbohm,[1,5] A. Strong,[6] P. Sizun,[7] D. Attié,[7] S. Schanne,[7] B. Cordier,[7] L. Bouchet[3] and A. von Kienlin[6]

[1]*INTEGRAL Science Data Centre, ch. d'Ecogia 16, 1290 Versoix, Switzerland*
[2]*Observatoire de Genève, ch. des Maillettes 51, 1290 Sauverny, Switzerland*
[3]*CESR, 9, avenue du Colonel Roche, B. P. 4346, F-31028 Toulouse Cedex 4, France*
[4]*GACE – ICUMV, University of Valencia, Apdo 22085, E-46071 Valencia, Spain*
[5]*IAAT – Astronomie, Sand 1, 72076 Tübingen, Germany*
[6]*MPE, Giessenbachstr. 1, 85748 Garching, Germany*
[7]*Service d'Astrophysique, DSM/DAPNIA/SAp, CEA/Saclay, F-91 191 Gif-sur-Yvette, Cedex, France*

Accepted 2004 November 24. Received 2004 November 24; in original form 2004 October 19

ABSTRACT

The performance of the SPI point-source data analysis system is assessed using a combination of simulations and of observations gathered during the first year of *INTEGRAL* operations. External error estimates are derived by comparing source positions and fluxes obtained from independent analyses. When the source detection significance provided by the SPIROS imaging reconstruction program increases from ∼10 to ∼100, the errors decrease as the inverse of the detection significance, with values from ∼10 to ∼1 arcmin in positions, and from ∼10 to ∼1 per cent in relative flux. These errors are dominated by Poisson counting noise. Our error estimates are consistent with those provided by the SPIROS program. With higher detection significance, the accuracy is ultimately limited to ∼0.5 arcmin in position and ∼1 per cent in relative flux by other types of errors. In these cases, SPIROS underestimates the true errors as it only takes into account the Poisson counting noise. At the low signal-to-noise ratio end, SPIROS is also too optimistic, and the number of spurious detections is significantly higher than would be expected from statistical noise fluctuations only. The analysis results do not depend significantly on the target off-axis angle, or on the number of pointings considered, provided that this number is larger than ∼15. Realistic source confusion tests are carried out by adding simulated data to the observation of the Crab nebula and pulsar. Reliable flux values can be obtained for sources separated by angles as small as 0.5 deg if their positions are known and kept fixed in the data deconvolution process. However, when SPIROS is searching for best source positions, if two sources are separated by less than ∼2 deg, it may only find a single spurious source located between the real ones (with a flux equal to the sum of the fluxes). Finally, users of the SPIROS program can find a number of important recommendations in the conclusion.

Key words: instrumentation: miscellaneous – methods: data analysis – gamma rays: observations.

1 INTRODUCTION

Measuring photons with energies in the gamma-ray range from a few tens of keV to several MeV is difficult. Reliable energy determination demands large detectors with enough mass to stop the photons, while precise localization can only be obtained with physically small detectors. These contradictory requirements are at the origin of the development of the two different main instruments for the *INTEGRAL* (Winkler et al. 2003) gamma-ray satellite launched in 2002 October. The IBIS imager (Ubertini et al. 2003) with two layers of large numbers of small detectors is optimized for high angular resolution sky imaging, and the SPI spectrometer (Vedrenne et al. 2003) with only 19 large hexagonal germanium detectors is optimized for accurate energy measurements. The two instruments are otherwise analogous, with active detector shielding and coded-mask optics.

The IBIS instrument, with a better angular resolution and sensitivity, is often regarded as the prime *INTEGRAL* instrument for

⋆E-mail: Pierre.Dubath@obs.unige.ch

measuring point sources. After all, celestial gamma-ray point sources such as binaries with neutron stars and black hole candidates are not in general expected to exhibit particularly narrow spectral features. However, SPI also has excellent capabilities for point-source studies: it may be more sensitive at energies above a few hundred keV; its calibration and response is for the time being better understood and characterized; it may also be best at detecting any spectral features of point sources, such as possible cyclotron lines in pulsars; and it provides in any case additional independent information to be compared with, and to complement, IBIS data.

The *INTEGRAL* Science Data Centre (ISDC) (Courvoisier et al. 2003) distributes both data and related scientific analysis software to users. For extracting point-source information from the SPI data, the key element of the ISDC software pipeline is the SPIROS imaging program (Skinner & Connell 2003). Solving for the minimum number of required parameters, SPIROS determines through a so-called Iterative Removal Of Source (IROS) method a 'sky model' which best explains the observed detector counts. The sky model consists of a list of source positions and intensities, and of background values. SPIROS analyses simultaneously as large a number of telescope pointings as possible thus taking full advantage of the 'dithered' *INTEGRAL* observation strategy (see, e.g. section 2 of Courvoisier et al. 2003). However, even with this increased number of measurements the analysis remains very challenging. In particular, a complete background solution cannot be derived from the data. Additional information on the background behaviour has to be introduced. SPIROS generally uses a model computed with the SPI_OBS_BACK program.

Given the difficulties in deriving point-source positions and intensities from a signal amounting often to a few per cent of the background, and with such a limited number of detectors, questions inevitably arise regarding the reliability of the results. To what extent do they depend on the number and positions of the telescope pointings used? How accurate are the error estimates when the background strongly dominates the signal?

On the other hand, although the total number of detected gamma-ray sources is not very large (of the order of 200), they accumulate near the Galactic centre. In many cases, IBIS found sources with angular separation smaller than the theoretical SPI resolution of about 2.5 deg. This adds yet another difficulty to the data deconvolution process, and it is certainly useful to characterize the resulting degradations in crowded fields.

These important questions are addressed in this paper. Errors are assessed by comparing the results of independent analyses, and the point-source analysis performance of the 'SPI telescope plus related software' ensemble is derived in a robust way. The intent is to provide information useful to users from the science community, both for a better interpretation of the results and for proposing best-tailored new observations. It may especially help users unfamiliar with the SPI instrument, and thus contributes to making *INTEGRAL* a better, more widely used, observatory.

2 DATA AND METHOD

Most of the data used in this study come from the *INTEGRAL* observations of the Crab nebula and pulsar collected in 2003 February 19–27 (revolutions 43, 44 and 45) and August 14–17 (revolution 102). During these two calibration periods, SPI was operated in the standard scientific mode with, in particular, telemetry of all measured characteristics of individual events. These observations include a total of 266 pointings (i.e. of periods with stable spacecraft attitude), with a typical duration of 35–40 min – with a few longer ones – amounting to a total exposure time of the order of 7.5 d.

Data from 152 pointings taken during revolutions 71 and 72 around the high-galactic latitude position (J2000.0) of $12^{h}\,59^{m}\,49^{s}$, $+\,27^{\circ}\,58'\,50''$ are also used, as an 'empty field' observation. In a moderate observation time no significant sources are detected in these data. They are used in some of our simulations as background data. They were taken with the same SPI configuration, and the pointing durations are also of the order of 35–40 min.

As the strongest stable source in the gamma-ray sky, the Crab provides enough signal to apply a straightforward method for error assessment. In all tests, the data set was divided into subsets in different ways. The analysis of each subset leads to an independent set of results. The resulting distributions provide 'external' statistical error measurements, which can be compared with the 'internal' errors provided by the SPIROS program. This method allows the reliability questions raised in the introduction to be addressed.

The systematic errors that may affect the analysis of each subset in the same way can be quantified by comparing our mean results with reference values. For the flux, getting a good reference value for the Crab is difficult. It requires a comparison with Crab spectra obtained from other missions and a critical assessment of the SPI response, and this is the subject of other publications. In this paper, we discuss both statistical and systematic errors on source position determinations, and relative statistical errors on flux measurements.

All analyses were made with the ISDC science analysis pipeline (see the user manuals and cook-books from ISDC Web pages for detailed information). The pipeline calls successively (1) SPI_OBS_POINT for gathering pointing related information, (2) SPI-BOUNDS for definition of the energy bin, (3) SPI_OBS_HIST for event binning, (4) SPI_OBS_BACK for defining a background model, and (5) SPIROS for image reconstruction.

The so-called single detector events (which hit only one detector) were primarily used. Including double-detector and triple-detector events did not lead to significant improvements, since at the relatively low energies selected in our investigations (20–300 keV) they contribute only a minor fraction of the total signal (their contribution can however dominate at MeV energies).

The model of the sky determined by SPIROS consists of a list of source positions and fluxes, and of background parameters, one value for each detector/pointing/energy band combination in the most general case. The solution is derived by minimizing – assuming a chi-squared statistic in our case – a measure of the deviations between the observed counts and those computed from the model using the instrumental response. SPIROS starts by locating the brightest source through a simple cross-correlation process. It refines its position and flux with the chi-squared optimization and subtracts the corresponding counts from the original observed ones. The new count data set is used in a second iteration, again searching for the (second) brightest source and deriving its precise location and flux. The iterations continue until one of two criteria given as input parameters is met: either SPIROS has found the specified maximum number of sources, or no source above the given minimum significance is detected anymore.

Unfortunately, the most general case with one background value per detector/pointing/energy band cannot be solved as there would then be more unknowns than measurements. Additional constraints have to be introduced, and in all our analyses, we assume that the background is proportional to the saturated germanium events (SPI_OBS_BACK option GEDSAT). The background time variability is thus fixed, and only one scaling coefficient per detector/energy band

Table 1. The different groupings of the Crab spacecraft pointings.

Number of pointings per group	Number of groups
1	266
5	52
10	26
15	17
20	13

needs to be derived by SPIROS (SPIROS background option 2).[1] For a given energy band, the final solution includes 19 background coefficients, in addition to the list of source positions and fluxes.

SPIROS assumes that input counts are affected by Poisson noise only. Counting errors are propagated through the whole process to compute the errors on the final results. A measure of the goodness of the SPIROS fit is output to the log file as chi-squared residuals, given as a single total value, one value for each detector, and one value for each pointing.

For several of the 266 pointings, the residuals are much larger than typical values obtained from the other pointings. This indicates that the difference between the measured detector counts and those computed from the final 'source + background model' solution are much larger than the expected errors from Poisson noise. Other types of errors must be important; most probably the background model used is inadequate for these few pointings. These 'bad' pointings were identified from a preliminary analysis of all 266 pointings, and excluded from the final solutions presented in this paper.

3 POINTING GROUPS OF DIFFERENT SIZES

When analysing one or a few pointings together, the number of measurements considered is small because of the limited number of SPI detectors, and as a consequence, results can be affected by edge effects and imperfect signal coding. These effects average out with more measurements. The results presented in this section characterize the accuracy improvements obtained when analysing larger and larger numbers of pointings together.

In order to separate the possible sources of errors, photon counting noise was first minimized by selecting cases with high signal-to-background (S/B) ratio. This was done by considering an energy band of 28–48 keV. In this band, the SPIROS Crab detection significance is always very large (larger than 70 standard deviations even for single pointings).

The full data set from the Crab observations of revolutions 43, 44, 45 and 102 was divided into independent groups of N pointings. The five cases shown in Table 1 were considered. For each of the five cases, the different groups were analysed separately. For example, the 17 groups of $N = 15$ pointings lead to 17 independent results. Fig. 1 shows the Crab positions obtained from the groups of single, 5, 10 and 20 pointings in the four different panels.

The distributions of the results obtained with groups of $N = 10$ are shown in Fig. 2 as an example. Means and standard deviations are also displayed in each panel. These distributions provide a measure of both systematic and statistical errors (except for the systematic on the Crab flux). The largest residual systematic error, observed in RA at the level of 0.4 arcmin, is not significant, and anyway represents only 0.3 per cent of the full-width at half-maximum of

[1] Except in single pointing analysis where only one scaling coefficient is derived for all detectors and pointings (SPIROS background option 3).

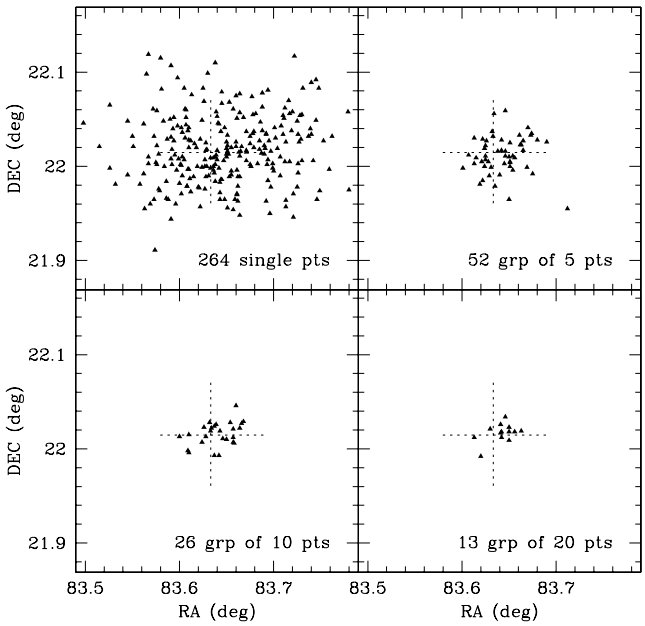

Figure 1. Each triangle in the four panels shows the Crab location resulting from one analysis of a group of either $N = 1, 5, 10$ or 20 pointings. The large dotted-line cross marks the J2000 catalogue position (83.°6332–22.°0145) of the Crab pulsar.

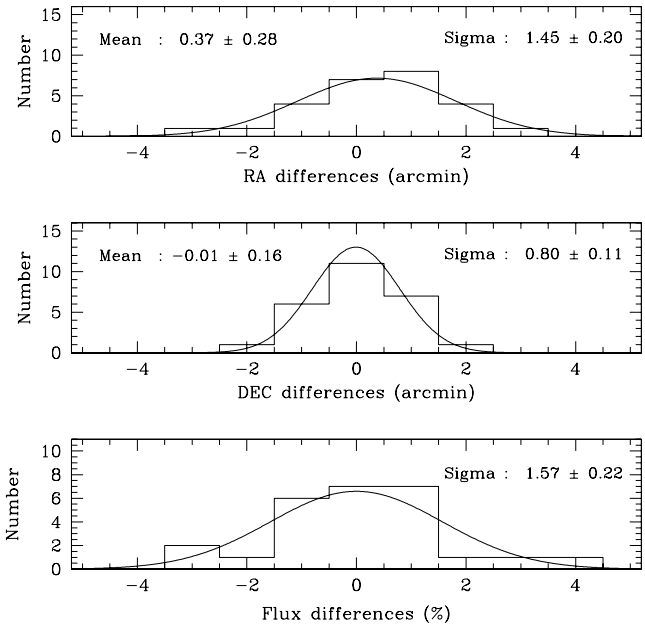

Figure 2. Distributions of the positions and fluxes obtained from analyses of 26 independent groups of 10 pointings. Positions are relative to the Crab catalogue position, while fluxes are relative to the mean flux. Means and standard deviations are displayed in each panel (except for the mean flux which is null by definition).

the response. This very good agreement is however expected given the use of an instrument misalignment matrix calibrated from the Crab and Cygnus X-1 data (Walter et al. 2003).

Fig. 3 displays the standard deviations (triangles) of all distributions as a function of N, the number of pointings in the groups. The error bars represent uncertainties due to small number statistics (i.e. $\pm$ standard deviation$/\sqrt{2 \times \text{number of groups}}$). The means of the errors provided directly by SPIROS, represented as large dots in this

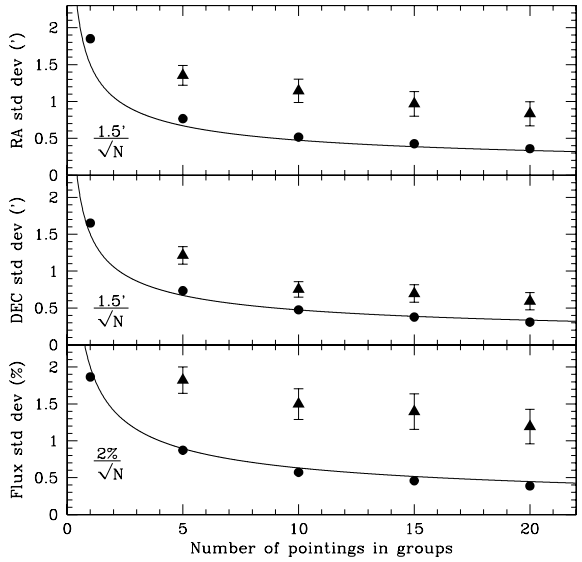

Figure 3. The standard deviations of all position and flux distributions – triangles with statistical error bars – are shown as a function of the group pointing number. Large dots represent the means of the errors provided directly by SPIROS. They are significantly smaller than the true errors because SPIROS takes only counting noise into account. Clearly, other types of errors start to be evident in these cases with very high signal-to-background (S/B) ratios. Continuous lines following 1.5-arcmin/$\sqrt{N}$ and 2 per cent/$\sqrt{N}$ laws are for illustration purposes.

figure, give a measure of the errors due to the counting noise. They certainly underestimate the true errors because of the importance of other types of errors in this case of high signal-to-background ratio. This figure shows that these 'additional errors' are of the order of 0.5 arcmin on positions and of 1 per cent on fluxes.

These values are remarkably small. They characterize differences between results derived from revolutions 43–45 and 102, and this shows that the SPI instrument is very stable. The flux error implies that the measured Crab flux is constant to better than 1 per cent over the six-month period between revolutions 43 and 102. Because the *Z*-axis of the spacecraft must always point close to the Sun (Jensen et al. 2003), there is a 180-deg spacecraft orientation difference between the two observation periods. This difference does not affect the results significantly either.

The 'additional errors' also explain why the differences between the measured detector counts and those computed from the final model – with a background proportional to the saturated germanium events and a source at the Crab location – are much larger than the expected errors from Poisson noise only. The chi-squared residuals are greater than the nominal value (equal to the number of degrees of freedom, $N_{\rm DOF}$), typically by four times the expected standard deviation ($\sqrt{N_{\rm DOF}}$) and sometimes by as much as 20 times. The large residuals do not come from a few pointings with particularly high background values, as those have been excluded from the analysis as explained in last section. As expected when the signal-to-noise (S/N) ratio is so large – SPIROS detection significance ranges from 110 to 250 with groups of five to 20 pointings, respectively – systematic errors become evident. They may come from instrumental problems, or from imperfections in the instrumental response description or in the background modelling. In any case, they are surprisingly small, and we show below that they are negligible in many practical cases.

Those analyses which have smaller chi-squared excesses should in principle produce better results, but we failed to find any trends or correlations between the chi-squared residuals and the quality of the results, expressed as differences from the catalogue value in the case of positions or from the mean value in the case of fluxes

Fig. 3 shows that the SPIROS error estimates follow a $1/\sqrt{N}$-type of law, as expected for noise-related errors, while systematic errors are consistent with such a law but also with a linear relation.

4 POINTING GROUPS WITH POSITION OFFSETS

In the general case, the positions of the pointings used in the analysis may not be symmetrically arranged around the target source in the sky. Because of the dithering strategy, the SPI large field of view, and the uneven coverage of the sky, the network of available pointings around any sky position is quite likely to contain biases and irregularities. It is therefore important to investigate whether the results depend significantly on the relative positions of the target and the analysed pointings.

The locations in the plane of the sky of all 266 pointings of our complete data set are displayed as triangles in Fig. 4 (with the open circle showing the Crab position). In order to test for systematic effects with large offset angles, the pointings were divided into groups following the dashed dividing lines. These 20 groups contain from six to 23 pointings, with an average of 13 pointings per group.

Fig. 5 displays the distributions of the results derived from these 20 groups. The means and standard deviations are consistent with those of Fig. 2 showing that there is no obvious degradation with increasing offset angles. The systematic error in RA is more significant, but it remains relatively small (<1 arcmin). Fig. 4 also shows the position (solid-line arrows) and flux (dashed-line arrow) error vectors, arbitrarily scaled and shifted to appear on top of their respective group locations. These error vectors are computed as the

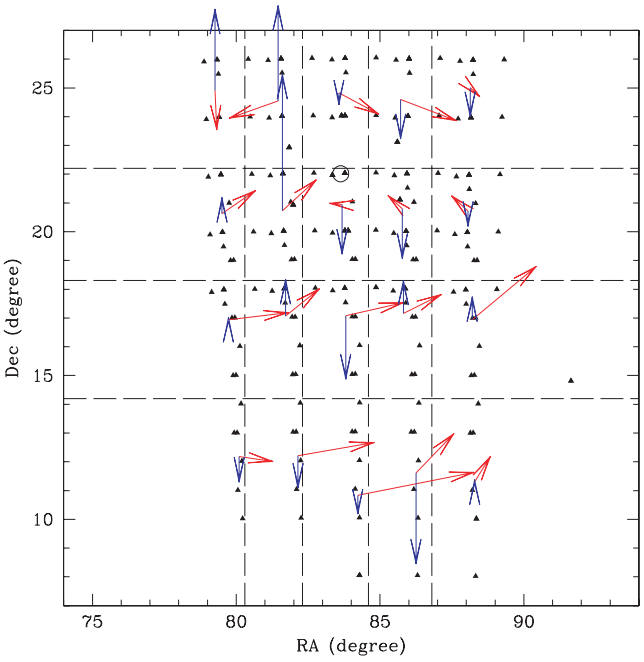

Figure 4. Triangles represent the locations of the 266 pointings of the Crab observations used in this study. These pointings are split into groups according to the boxes defined by the dashed lines. Arrows provide an illustration of the position (red) and flux (blue) error vectors resulting from the analysis of the corresponding pointing group (see text).

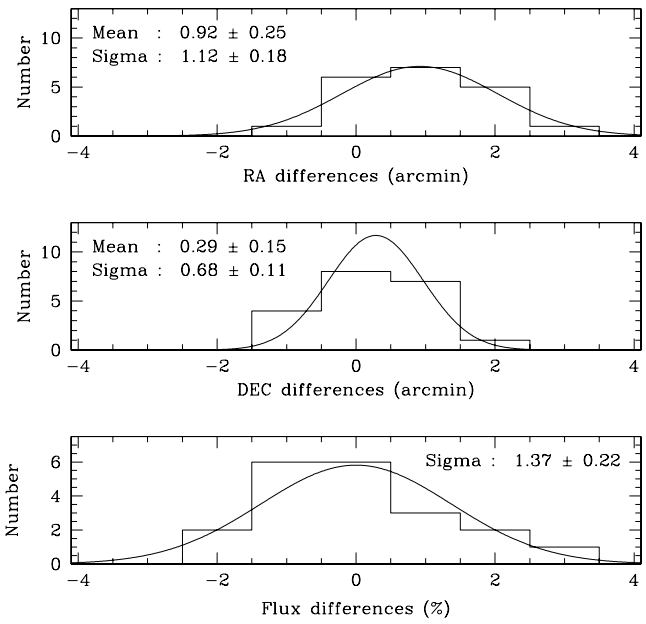

Figure 5. Distributions of the Crab position and flux results derived from the analysis of the 20 pointing groups defined by the dashed, dividing lines shown in Fig. 4. Positions are relative to the Crab catalogue position, while fluxes are relative to the mean flux. Means and standard deviations are displayed in each panel (except for the mean flux which is null by definition).

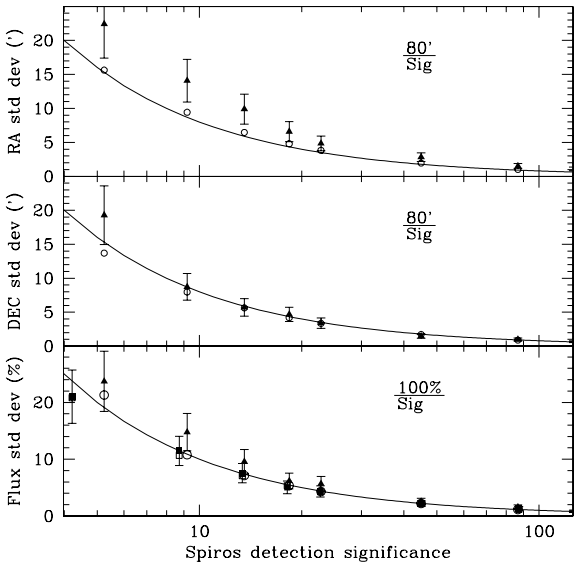

Figure 6. Results of the 'empty' field + simulated source analyses. The standard deviations of all position and flux distributions – triangles with statistical error bars – are shown as a function of the SPIROS detection significance. The means of the errors provided directly by SPIROS are represented as open circles. In the flux panel, the values obtained when fixing the source position in the analysis are also displayed as filled and open squares.

difference between the positions resulting from the analyses and the catalogue position or, for the flux, as the difference from the mean value. Again, no obvious systematic error is observed in this figure.

5 ANALYSES WITH INCREASING NOISE LEVELS

The SPI detectors measure events coming from sky sources and from the background. The main difficulty of the data analysis is to separate the two, which are both affected by counting noise statistics. The background does not depend on the pointing direction and remains always at approximately the same level.[2] The Crab is the strongest steady gamma-ray source in the sky. Consequently, the signal-to-background ratio of a typical SPI observation is smaller than those considered in the last two sections, with lower S/N ratios for the source contributions and similar S/N in the background.

Two different types of low S/N data analyses have been carried out. First, cases of an increasingly faint source on a constant background were studied using an empty field observation and a source count simulator. Second, analyses of the Crab data were carried out at higher energy, or in narrower energy bands, in order to explore cases with smaller S/N ratios but approximately constant S/B, as both the signals from the background and from the source are decreasing in roughly the same way.

5.1 Faint sources with high background level

The data from the 150 pointings taken during revolutions 71 and 72 around the high-galactic latitude position (J2000.0) of $12^{\rm h}\,59^{\rm m}\,49^{\rm s}$, $+27^{\circ}\,58'\,50''$ were used as a source of background. Several analyses were carried out to confirm the absence of bright gamma-ray sources in this data set. The contribution of a simulated source located at the centre of the field of view was then computed with the ISDC program SPI_ADD_SIM, using the SPI instrumental response and the pointing information. The computed counts were added to a copy of the 150 pointing binned-event data set.

Ten different cases of simulated source strength, ranging from 20 to 1000 mCrab, were considered, and in each of the 10 cases, 10 analyses of 10 independent groups of 15 pointings were carried out. All analyses were made with parameter values identical to those used for the above Crab analyses; in particular the energy band selected was 28–48 keV. Ten distributions of results were derived and Fig. 6 shows the standard deviations of these distributions as a function of the mean SPIROS simulated-source detection significance. Also displayed on this figure are the mean errors provided by SPIROS.

5.2 Faint sources with low background level

Using the 17 groups of 15 consecutive pointings from the Crab observation data set, analyses in narrower and higher energy bands were carried out to explore cases of decreasing signal-to-noise ratios. All results were obtained at relatively low energy (<300 keV) where the instrument response does not change much with energy. In each case, standard deviations were derived from the resulting RA, Dec., and flux distributions. These standard deviations and their statistical error bars are shown as a function of the mean detection significance given by SPIROS in Fig. 7.

5.3 Counting noise errors and SPIROS chi-squared residuals

Figs 6 and 7 show very consistent results. For a SPIROS detection significance of $\sim 10\,\sigma$, errors are of the order of 10 arcmin in position and 10 per cent in flux. These errors are much larger than those in the high S/N cases displayed in Fig. 3. They are consequently dominated by counting noise errors, and as expected for this type of

[2] The background, however, exhibits small variations which must be properly modelled in any data analysis.

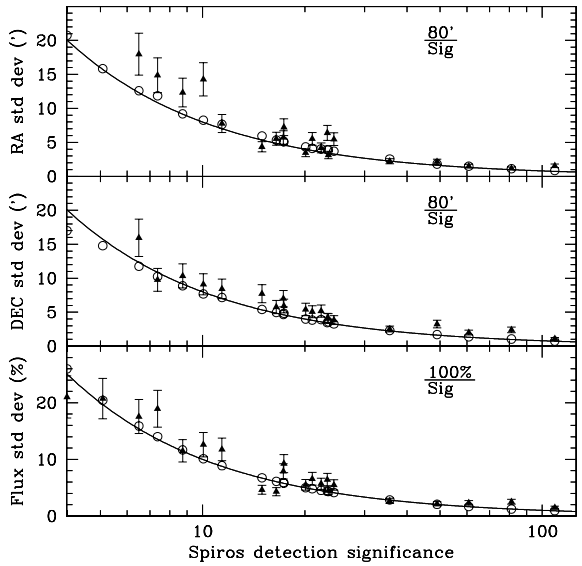

Figure 7. Results of the Crab reduced S/N data analyses. The standard deviations of all position and flux distributions – triangles with statistical error bars – are shown as a function of the SPIROS detection significance. The means of the error estimates provided directly by SPIROS are represented as open circles.

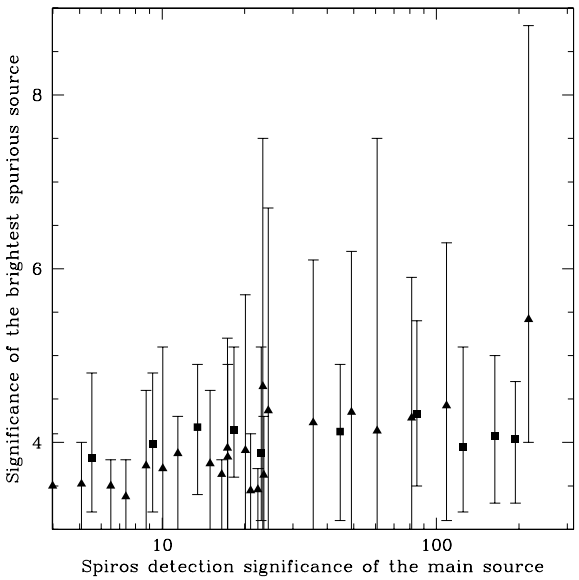

Figure 8. Significance of the spurious (second) source detected by SPIROS as a function of the detection significance of the main source. Squares show results from the 'empty' field + simulated source analysis, while triangles indicate results from the Crab analyses. Error bars do not represent the usual standard deviation (displayed in all other plots), but the total range from the smallest to the largest obtained values.

error, they decrease as the inverse of the SPIROS detection significance until they reach about 1 arcmin in position and 1 per cent in flux. At this level, as explained in the previous section, other types of errors not related to counting noise are becoming evident, and errors no longer depend much on S/N.

Figs 6 and 7 also indicate that our results are consistent with the error provided by SPIROS, although there are some systematic differences (notably in RA) at the 1σ level below a detection significance of about 20. In any case, SPIROS appears to propagate counting noise errors correctly and it produces reliable error estimates.

Contrary to the high S/N cases, chi-squared residuals of all the low S/N analyses are perfectly consistent with the numbers of degrees of freedom. In other words, the detector counts predicted by SPIROS from the model representing the final solution are consistent, within the counting noise errors, with the observed counts. This means that the background modelling approach is appropriate. Assuming that the background is proportional to the germanium detector saturated counts, and deriving one coefficient for each detector through the SPIROS solution, produces an adequate model. Note however that our observations span a limited observation time. On much longer time-scales, the saturated germanium counts may not always provide a satisfactory background model.

Errors displayed in Figs 6 and 7 are all consistent with a 80-arcmin/significance law for the positions, and a 100 per cent/significance for the flux. Comparing triangles and squares in the lower panel of Fig. 6 also points to the important improvement obtained when extracting fluxes at a known position rather than at the position found by SPIROS. Whenever possible the source position should be fixed in the SPIROS analysis.

5.4 SPIROS detection of 'spurious' sources

The 150 pointings of our 'empty field' observation were analysed in 10 groups of 15 pointings. In eight of the 10 cases, a source with a significance above 3 is detected, and one of the values is as high as 4.9 standard deviations. These detections appear, however, at eight incompatible, apparently random locations in the sky. Consequently, they are probably statistical noise fluctuations. The mean flux extracted at those positions is 27 mCrab.

Even though both sets of low S/N analyses from the two previous subsections only include one source in the field of view, SPIROS can be instructed to search for other sources in the image. In more than half of the images, SPIROS finds a second source with a detection significance larger than 3σ. These sources are clearly spurious as they appear at random locations in the sky. Fig. 8 shows their significance as a function of the SPIROS detection significance of the main source, i.e. the simulated source or the Crab. The error bars in this case show the total range of the results, from the smallest to the largest values. This figure shows that spurious sources with significance up to 7–8 can be seen with SPIROS when another source is present in the field of view. The correlation with the significance of the main source is weak, although spurious sources with significance larger than 5 are not seen when the main source has a significance smaller than $\sim$20. This indicates that the spurious sources are unlikely to be simply ghosts of the main source, as a strong correlation would be expected in this case.

All these spurious sources are most probably statistical noise fluctuations, and this is an indication that SPIROS errors are somewhat optimistic for low detection significance. Assuming a Gaussian noise distribution in the image, a peak just larger than 3σ should only be seen in one out of three images on average (and peaks at 5–8σ should never be seen) as one of our images typically contains of the order of 100 independent points in the sky.

6 SOURCE CONFUSION TESTS

In order to investigate the expected result degradations when sources are increasingly close to one another, a mock source was simulated and added to the Crab observation. In this approach, the simulated part is kept to a minimum as the background and one of the two

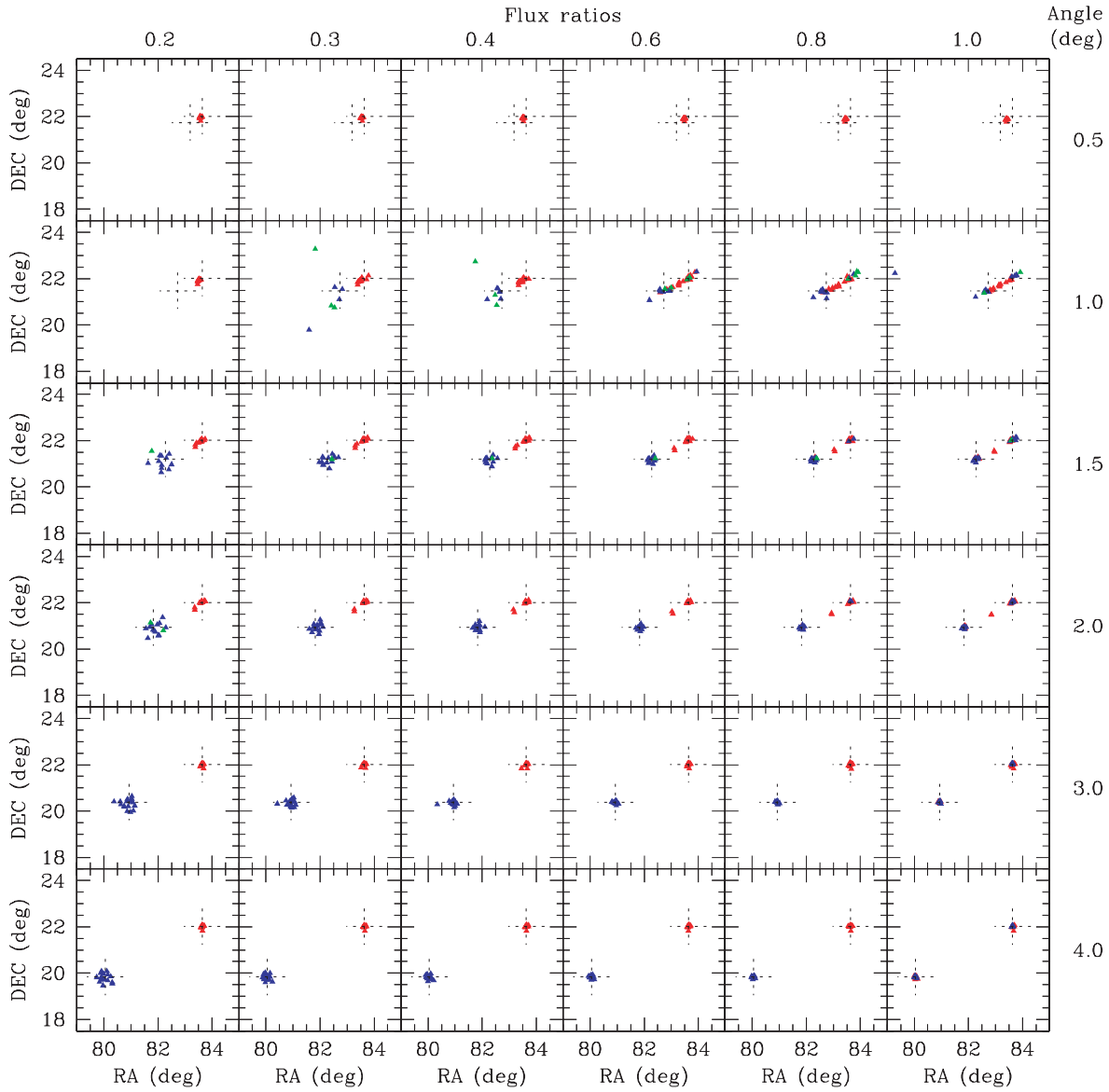

Figure 9. Locations of the (up to) three brightest sources above a threshold of 3σ found by SPIROS in the analyses of different data sets obtained by adding the contributions of a simulated source to the actual Crab *INTEGRAL* observation. The separation angles and flux ratios for the 36 studied cases are indicated around the main frame. The two large crosses show the input positions of the two real sources.

sources come from an actual observation. The detector counts expected for a source close to the Crab location were computed for all 266 pointings of our data set using the program SPI_ADD_SIM and the SPI instrumental response. 36 cases were considered, with different separation angles (0.5, 1.0, 1.5, 2.0, 3.0 and 4.0 deg) and different flux ratios (1, 0.8, 0.6, 0.4, 0.3 and 0.2 of the Crab flux). An energy band 110–116 keV was selected. In this band, the SPIROS Crab detection significance is about 46 with 15 pointings, corresponding possibly to a more typical signal-to-noise case. The computed counts were then added to the Crab data, and in each of the 36 cases, 17 analyses were carried out for independent groups of 15 pointings. Results are displayed in Figs 9 and 10.

Different zones can be identified in Fig. 9. In the lower panels, the two sources are properly separated. In some cases of 2-deg separation angle, instead of the two 'real' sources, SPIROS starts to find only a 'spurious' one at a position located between the two sources. The number of such cases increases with decreasing separation angle, and with 0.5-deg separation SPIROS detects only one source in every case. Also note that the spurious source position is half-way between the two actual sources when they have equal flux, and that this location moves towards the main real source position when the second source becomes weaker.

The left panels of Fig. 10 show that when SPIROS finds only one combined 'spurious' source, the resulting flux value is the sum of the Crab value and that of the simulated source.

A second complete set of analyses was carried out, fixing source positions in the flux extraction process. The resulting dramatic improvement is pointed out in the middle column of Fig. 10: spurious results cannot be obtained with fixed positions. However, the ranges of derived fluxes are larger when the source separation decreases. This is quantified in the right panels of Fig. 10 which shows the standard deviations of the flux values as small triangles. The mean

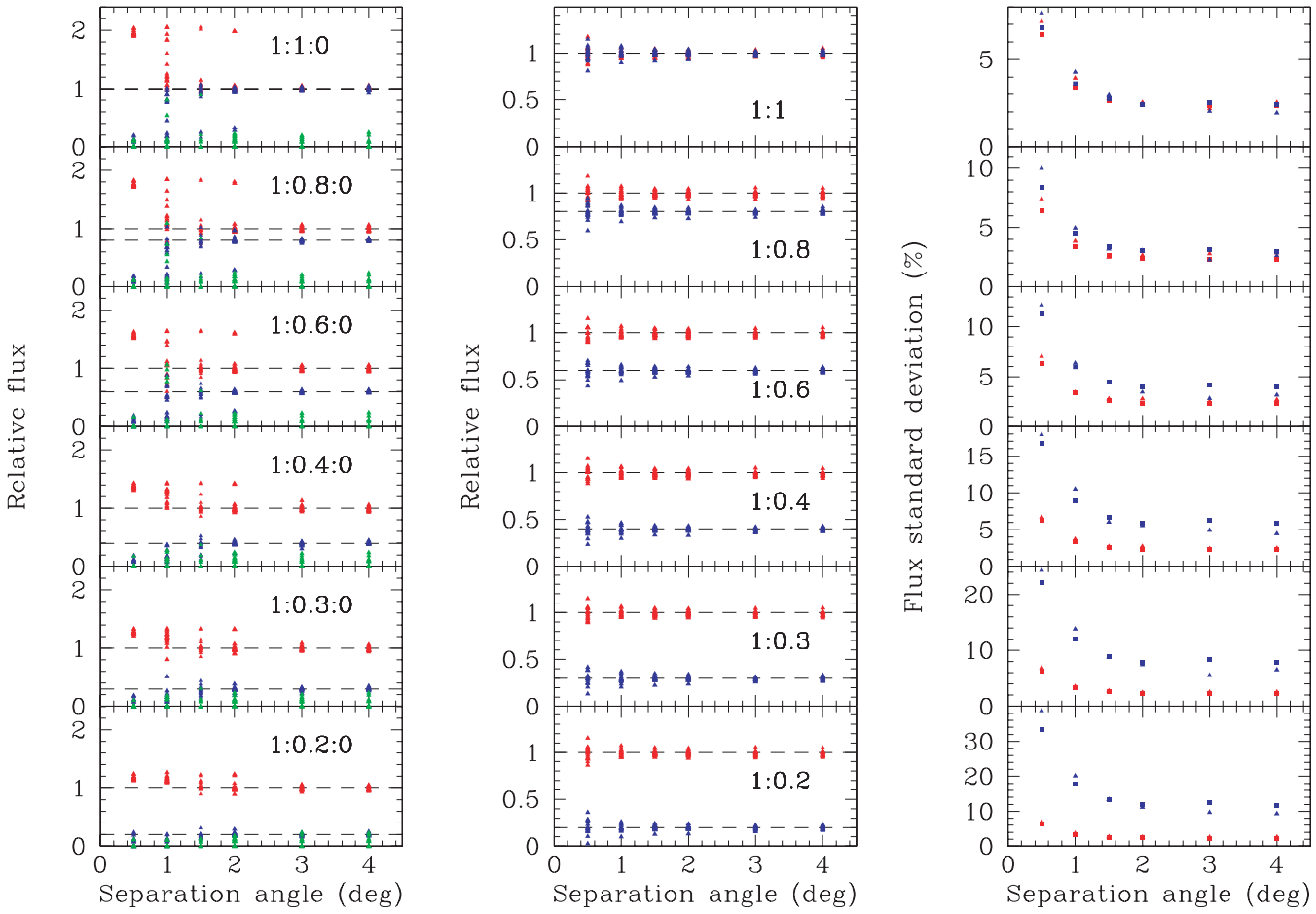

Figure 10. Left: relative fluxes extracted by SPIROS at the positions displayed in Fig. 9. Green triangles show the third (spurious) source found by SPIROS, indicating the noise level in the plots. Middle: relative fluxes extracted at the fixed known positions. Right: SPIROS errors are displayed as squares, while the triangles are the standard deviations of the flux distributions.

SPIROS errors, displayed as small squares, closely follow the triangles, indicating that SPIROS always produces reliable results, even with separation angles of 0.5 deg.

7 CONCLUSION

The performance of the SPI point-source data analysis system is remarkably good. In moderate S/N ratio cases, with SPIROS detection significance ranging from 10 to 30, errors are from 9 to 3 arcmin in position and from 10 to 3 per cent in relative flux. In this S/N range, SPIROS provides reliable solutions, with in particular, error estimates consistent with the spread observed in results of large numbers of independent analyses of the same sky field. Poisson counting noise is clearly the dominant source of errors. The observed errors decrease as the inverse of the SPIROS detection significance until they reach about 1 arcmin in position and 1 per cent in flux. At this level, other errors, not related to counting noise, become evident, and uncertainties no longer depend much on the S/N ratio. Extracted fluxes are always significantly more accurate when the source position is fixed in the deconvolution process.

The level of chi-squared residuals of the SPIROS solution is a very good reliability indicator. When the residuals are consistent with the numbers of degrees of freedom of the problem, it means that the detector counts computed by SPIROS from the final solution are consistent, within the counting noise errors, with the observed counts. Good residuals can in principle only be obtained when the sky model is appropriate and the counting noise is the dominant source of error. The sky model is the sum of the background and of the source contributions. SPIROS can be instructed to take only a limited number of sources into account, and the first check is to make sure that all significant sources are indeed included in the sky model. When this is the case, large chi-squared residuals generally reveal a non-optimum background model.

In very high S/N cases (SPIROS detection significance >110), large chi-squared residuals can also be obtained because other types of errors become larger than the counting noise. In these cases, SPIROS underestimates the true errors, as it takes only the counting noise into account. However, the effect of these additional errors is small – of the order of 0.5 arcmin in position and of 1 per cent in relative flux – and Poisson counting noise is likely to be the dominant source of error in most practical cases.

The number of pointings required to reach the above quoted level of error is not very large. Edge effects and imperfect coding are averaged out below that level as soon as the pointing number is of the order of 10. Except for long staring observation, having too few pointings is unlikely to be a real problem for most practical cases. We find that the results do not depend significantly on off-axis angle or on asymmetry in the pointing location distribution. However, it is very important to make sure all real sources are included in the analysed field of view. By design, SPIROS cannot consider a source

outside the specified image field of view. In our case, if the Crab location is outside the final image, completely spurious results are obtained, as SPIROS tries to find a source inside the image field of view to explain the counts due to the Crab.

When SPIROS detection significance is below about 10, extreme care should be taken in the analysis. In our cases, spurious sources have been detected with detection significance up to ∼5 in the empty field, and up to ∼8 in the Crab observation. SPIROS detection significance is too optimistic at the lower S/N end. The number of spurious sources which are likely to be found also depends on the size of the sky image, i.e. on the number of independent sky elements.

In crowded fields, reliable results are obtained with source angular separation down to ∼2.5 deg. With decreasing separations in our two-source experiment, it becomes more and more likely to detect only one combined spurious source located between the two real sources, with a flux equal to the sum of the two input fluxes. Whenever possible the position of known sources should be fixed in the flux/spectra extraction process. The improvement is dramatic, and reliable results are then obtained even with separation angle as small as 0.5 deg. The errors increase, but the error estimates provided by SPIROS always remain reliable, i.e. consistent with the spread of results derived from independent analyses, as can be seen in Fig. 10.

It is important to note that in the SPECTRA mode of SPIROS, spectra are built by extracting the best flux independently in one energy bin after another. Spectra are just a succession of flux values in different energy bins, and as a consequence our results can easily be extrapolated to spectra. The case of variable sources is not directly addressed in this paper. Our results apply if an observation is cut into pieces within which the source(s) can be considered as stable. SPIROS does, however, have a timing mode which was not used in the current study but which does allow one to solve for the intensities of variable sources.

As pointed out above, some of our results may be directly useful to users of *INTEGRAL*, for SPI data analysis, and the interpretation of the results. Our method can also be followed in some other cases. Separating the data set into independent parts is often possible, and it allows one to investigate whether the results behave as expected according to the errors derived by the programs. The program SPI_ADD_SIM is part of the ISDC software distribution. It can be used to add a simulated source on top of a crowded field to see whether it is found despite the confusion, and to study how the other source results may be affected.

ACKNOWLEDGMENTS

The SPI project has been completed under the responsibility and leadership of the Centre National d'Etudes Spatiales (CNES). We are grateful to ASI, CEA, CNES, DLR, ESA, INTA, NASA and OSTC for their support.

REFERENCES

Courvoisier T. J.-L. et al. 2003, A&A, 411, L53
Jensen P. L., Clausen K., Cassi C., Ravera F., Janin G., Winkler C., Much R., 2003, A&A, 411, L7
Roques J. P. et al., 2003, A&A, 411, L91
Skinner G., Connell P., 2003, A&A, 411, L123
Ubertini P. et al., 2003, A&A, 411, L131
Vedrenne G. et al., 2003, A&A, 411, L63
Walter R. et al., 2003, A&A, 411, L25
Winkler C. et al., 2003, A&A, 411, L1

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