A design based on an inverse approach

Proposal for the ALMA configurations, 9 January 2002

Frédéric Boone

1 Introduction

A proposal is made for the design of ALMA configurations. The locations of the pads are show in the Figures 2, 3, 4, 5, 6, 7. This report describes the way these locations were obtained and the properties of the configurations.

The design results from the application of an inverse method in which the positioning of the pads is driven by the scientific goals. The main goal of ALMA that is considered here, is to allow reliable and clean imaging with the maximum of sensitivity and with the minimum of prior knowledge of the sources. In other words, the aim is to find the design that makes the best use of the potentiality offered by the instrument.

The optimization tool used and the analysis that allows to translate the imaging specifications in terms of target distributions of samples are briefly described in Section 2. The achievement of the complete design is addressed in Section 3. Section 4 gives a description of the properties of the configurations obtained. A conclusion is drawn in Section 5.

This document is kept very synthetic, most of the points addressed are developed in other publications which can be referred to for more details (Boone 2001a,b, 2002).

2 Optimization: tool and targets

2.1 The optimization software

In Boone (2001b) a new method was introduced for the optimization of the antenna locations. It is based on the computation of pressure forces emanating from the discrepancies between the actual distribution of Fourier samples and a target one. The array elements submitted to these forces move toward optimal positions. The efficiency, rapidity and flexibility of the method are suitable to the design of large arrays.

Its implementation in a C++ library able to handle most of the observational situations like synthesis, mosaicing, multi-configuration as well as topographic constraints was described in the same reference. This software, named APO, is a powerful tool that can be used to design interferometers following an inverse approach. In such an approach the pad locations are determined by the choice of the target distributions of samples. It is therefore possible to control the minimum density of samples and the overall shape of the distribution of samples.

2.2 The target distributions

The choice of the target distributions should be driven by the scientific goals. An analysis was developed in Boone (2002) to derive the characteristics of the targets from the imaging specifications. The levels of interpolation and extrapolation that can be achieved in uv-plane were included in the analysis as free parameters. They can be adjusted to the performance of the available deconvolution algorithms, the prior knowledge of the source and the use of mosaicing.

The FWHM of the target distribution as well as the required duration of observation are determined by the length of the largest projected baseline (i.e. the size of the instrument). Below a size B_o the instrument should be optimized for a distribution of samples equal to the Fourier transform of the wanted clean beam. For sizes greater than B_o and lower than another characteristic size B_{max} the sampling constraint (coming from the limited level of interpolation that can be achieved) requires to take a target distribution different from the Fourier transform of the wanted clean beam. This implies that some sensitivity will be lost when reconstructing the wanted clean map. When a Gaussian clean beam is wanted this loss is minimized by taking a target of Gaussian shape (but wider than the Fourier transform of the wanted clean beam). For sizes greater than B_{max} reliable image reconstruction is in principle not possible and multiconfiguration observations are required. The expressions for B_o, B_{max}, the FWHM, the duration of observation required, and the loss of sensitivity are given in the reference. Some numerical applications for ALMA are given below.

2.3 Application to ALMA configurations

For simplification we will assume here that no extrapolation is allowed and that we want to produce clean images in which the clean beam is the Fourier transform of a Gaussian distribution truncated at 15 dB. (This is an arbitrary choice, but the aim is only to give an idea of the values of the parameters.) Then for 60 antennas of 12 meters in diameter, the values in meters of B_o and B_{max} are given by:

$$B_o = 100\,\alpha\,\Delta t_{max} \quad \text{and} \quad B_{max} = 615\,\alpha\,\Delta t_{max}, \tag{1}$$

where Δt_{max} is the maximum duration of observation in hours and α is the achievable level of interpolation. More precisely, α corresponds to the maximum distance between the samples in Nyquist interval units that allows a reasonable estimation of the visibility function. (It is actually a question of density of samples rather than a question of distance between the samples and a more rigorous definition is given in Boone (2002)). Hence, $\alpha = 1$ means that Nyquist sampling is required, $\alpha = 2$ means that the samples can be separated by twice the Nyquist interval, and so on... When no prior information on the source are available $\alpha = 1$, when the size of the source is known to be a fraction x of the field of view $\alpha = x$, when a mosaic is made $\alpha = 2$ (it is possible to estimate the visibility function up to Nyquist interval around the samples (Cornwell 1988)).

Then, if the size of the instrument B is less than B_o the FWHM of the target distribution (written δ_s) and the duration of observation required (in hours) are given by:

$$\delta_s = 0.896\,\frac{B}{\lambda} \quad \text{and} \quad \Delta t = 0.01\,\frac{B}{\alpha}. \tag{2}$$

If B is greater than B_o then, the duration of observation required is Δt_{max} and the FWHM of the distribution (δ_s) is solution of:

$$\frac{2a^2\lambda^2\delta_s^2}{B^2}\left[\exp\left(\frac{B^2}{2a^2\lambda^2\delta_s^2}\right) - 1\right] - \frac{B}{1.44\,B_{max}} = 0 \tag{3}$$

where $a = 1/2\sqrt{2\ln 2}$.

From Eq.1, if the maximum duration of observation equals 8 hours and $\alpha = 2$ (e.g. if mosaicing is considered as the default mode of observation) the most extended configuration should not exceed 10 km in diameter.

2.4 Conclusion

Finally, provided that the wanted clean beam (the truncation level of its Fourier transform) is specified and that the achievable levels of interpolation and extrapolation are estimated, the target distribution is exactly determined for any configuration. Running the optimization software for the targets thus obtained ensures the design to be optimal in regard to the imaging specifications.

To get the design of the whole array, the number of configurations, the size of the configurations and their positions on the topographic mask of the site also have to be specified. Some recommendations were made at the PDR (Guilloteau & the Configuration PDR Review Committee 2001) but some of these parameters still had to be determined.

3 Designing the configurations

The derivation of the whole design was detailed in Boone (2001a). This section summarizes the steps that were followed and the main choices that were made in the determination of the parameters.

3.1 The compact configuration

The compact configuration rises the particular issues of shadowing and access to the antennas.

The APO software can handle shadowing. Two different modes of constraint are possible. Either a single source elevation above which no shadowing is wanted can be specified. Or several source elevations with a weight for each one related to the tolerance of shadowing can be given. The first mode of constraint was chosen and the no-shadowing limit was set to 50 deg of elevation.

In order to allow imaging without losing too much of sensitivity, it was decided to optimize the compact configuration for a Gaussian distribution of samples. To allow access to the antennas it was found optimal to take a truncation level of 12 dB for the distribution and a maximum baseline of 200 m. According to the results of Boone (2002) Nyquist sampling can be achieved with such a configuration within 1.2 hours of observation.

The configuration was optimized for the mean elevation (i.e. 60 deg) then roads of 20 m in width were inserted by hand, then it was re-optimized with the roads as ground constraints.

3.2 The intermediate configurations

From the recommendations of the PDR, the intermediate configurations should expand up to 3 km in diameter. To allow a maximum of flexibility in resolution, it was decided that the sizes of the configurations should correspond to regular increments from the compact configuration up to 3 km. As the total number of pads is of 250, the compact and intermediate configurations should not use more than 200 pads.

The number of intermediate configurations and the target distributions that should be used to optimize them are constrained by the total number of pads. The more centrally condensed the configurations are, the greater is the number of pads that can be shared between them. But more centrally condensed configurations imply target distributions truncated at a lower level. This in turn implies a lower density of samples at the edges of the distributions (because the time of integration is limited). Therefore, more intermediate configurations implies poorer sampling (especially for the most extended ones). A compromise has to be found.

It was found that four intermediate configurations optimized for the target distributions as specified in Table 1 (see FWHM) would fit within the number of pads limit. This choice implies that 3 configurations yield samples separated by 4 times the Nyquist interval at the edges of the distribution (column α in Table 1). But as the distributions are Gaussians truncated at a low level (15 dB) the peripheral regions in which the distance between the samples is greater than twice Nyquist interval (acceptable when mosaicing) is small. The required durations of observation are also given in the table.

The configurations were optimized for the mean declination (i.e. $-10\,\mathrm{deg}$) in crescent size order trying to use as much as possible the pads of the more compact configurations. In each optimization the previous configuration was taken as the input configuration. The compact and intermediate configurations use 196 pads.

3.3 The extended configuration

In order to achieve the best resolution within the ground constraint represented by the existing roads the PDR recommended an extended configuration of 14 km. As an alternative we propose a configuration of 10 km. The argumentation is that this represents the size limit for sampling at twice the Nyquist interval (see Eq.1 for $\Delta t_{max} = 8$ h) which is justified when mosaicing or when the source is twice smaller than the field of view. Moreover, it seems also to be the size limit for the extended configuration to fit within the 5-degrees mask. The construction of the pads would therefore imply a lower cost than for a configuration optimized on the 10-degrees mask.

Such a configuration was optimized. It forms a ring around Cerro El Chascon and uses 7 pads of the most extended intermediate configuration.

3.4 Conclusion

A design including 249 pad locations was build conforming to the PDR recommendations and following the analysis presented in the previous section. A compromise was found between the number of intermediate configurations and the quality of the sampling at the edges of the distributions.

The quality of the configurations was also checked in Boone (2001a). A summarize is given in the next section and the problem of reconfiguration is addressed.

4 Properties of the design

4.1 Controlled minimum density of samples

In principle, with the method of optimization used, any hole in the distribution that can be filled with samples is filled. Thus, the density of samples is in principle greater or equal the minimum of the target distribution everywhere in the uv-plane. This point is important as below a critical density which depends on the possible level of interpolation, the information contained in the samples can not be used for the reconstruction of reliable maps (this is one of the key points of the analysis developed in Boone (2002)).

In the design proposed the compact configuration was optimized for a sampling everywhere better than Nyquist ($\alpha = 1$). With such a configuration no, or very little interpolation (to correct for unavoidable small defects in the distribution) is required, i.e. imaging is possible through a simple Fourier transform (with a possible weighting of the data to correct for residual small scale bumps in the distribution).

The first intermediate and extended configurations were optimized for the samples to be everywhere separated by less than twice the Nyquist interval ($\alpha = 2$). It means that if the extension of the source is larger than half of the field of view mosaicing is required to allow reliable imaging without prior knowledge of the source. But, if algorithms able to take into account known properties of the sources are used, this condition is not required. Moreover, sources twice smaller than the field of view may represent a non negligible fraction of the projects.

The three other configurations (2nd, 3rd and 4th intermediate) are optimized for the samples to be everywhere separated by less than 4 times the Nyquist interval ($\alpha = 4$). Unless the source is known to be 4 times smaller than the field of view an effective algorithm is required to benefit from the information contained in the peripheral samples. Note that anyway such distributions are much better than those obtained with the existing interferometers: due to the Gaussian shape of the distribution only a small fraction of the samples, at the edges of the distribution, are separated by more than twice Nyquist interval. Most of the uv-plane is still sampled at a rate higher than Nyquist.

4.2 Beams

The optimization method also ensures the distributions of samples to be as close as possible to the targets. The choice of Gaussian target distributions truncated at a low level implies that clean maps can be obtained with very little loss of sensitivity. This requires estimation of the visibility function (i.e. interpolation and/or extrapolation by means of a deconvolution algorithm if necessary) and weighting of the estimated visibility function. However, due to the high density of samples (described above) it is possible to produce close to clean maps without interpolation, simply by weighting properly the data. Indeed, the sidelobes due to some unavoidable bumps in the distributions can be removed by giving weights inversely proportional to the local density excess (e.g. Briggs 1995).

Thus, it was demonstrated in Boone (2001a) that beams with sidelobes lower than 2% could be obtained for all the configurations at the expense of 5% of the sensitivity (except for the inner sidelobes of the extended configuration: it was optimized for a distribution truncated at a high level).

4.3 Reconfiguration

The design proposed is made of 6 configurations of 60 antennas for a total of 250 pads. This implies that in average 22 pads are shared between consecutive configurations and that 38 antennas have to be moved to go from one configuration to the next one. It means that for half a cycle (i.e. going through all the configurations from the compact to the extended or inversely) 190 antenna moves have to be accomplished. If it is assumed that 3 antennas can be moved within a day then, 64 days are required.

A possible strategy would then be to spend the 118 remaining days of the half-year observing in the optimized configurations proposed, that is, 19 days per configuration. Such a strategy would however not be very flexible: if, while reconfiguring the instrument the atmospheric conditions are exceptionally good or for some reasons the transport is impossible (wind or mechanical problems) it might be desirable to stop the reconfiguration and observe with all the antennas. Thus, observing with the hybrid configurations should be integrated in the reconfiguration strategy and the quality of these configurations should be studied.

In the design proposed here, the pads are distributed almost continuously from the compact configuration to the 4th intermediate configuration. It is therefore possible to get hybrid configurations of quality comparable to the optimized ones. One can afford to spend a significant observing time in the hybrid configurations. Then, there is a great flexibility in the reconfiguration strategy.

A possible scheme of reconfiguration would be to observe in the hybrid configurations during half of the time required by the transportations. For example one could move 6 antennas in 2 days then, observe during one day and alternate 2 days of transportation with 1 day of observation until the next optimized configuration is reached . This schedule could be adjusted to the atmospheric conditions, e.g. one could instead observe with all the antennas during 3 days in a row and move the antennas during the following six days. Or again, if the atmospheric conditions were bad for observations one could move the 38 antennas in a row (about 13 days required) to get directly to the next configuration.

In such a scenario 16 days would remain for observations in each optimized configuration.

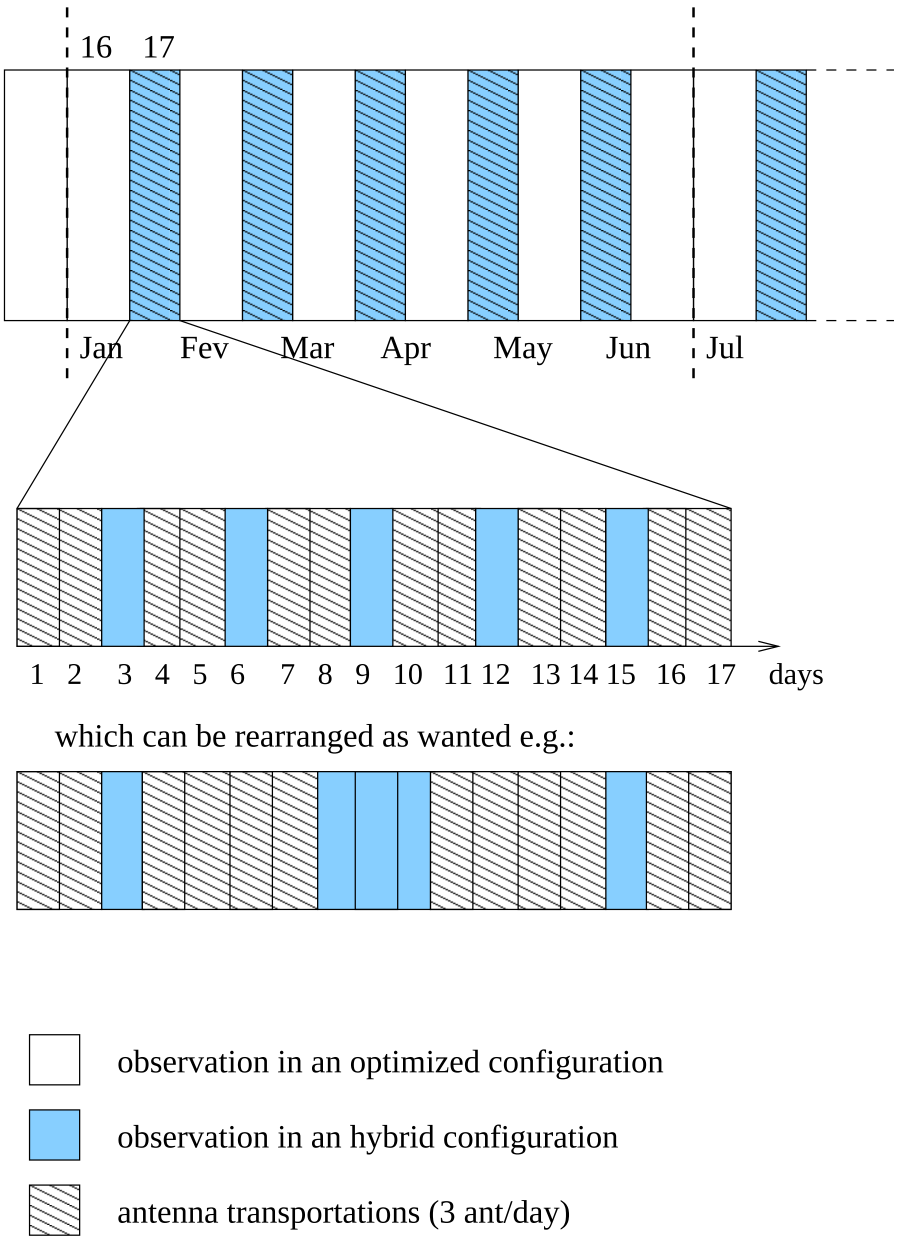

Figure 1: General scheme for the reconfiguration

More time could be spent in the hybrid configurations but it might be a good thing to have periods of about 2 weeks without antenna transportations. This would allow to take advantage of the periods of best atmospheric conditions. This might also be an advantage for the management of the staff and the maintenance of the instrument.

Thus, to summarize, in such a scenario periods of 17 days during which transportations alternate with observations in the hybrid configurations (12 days of transportation and 5 days of observations) would alternate with periods of 16 days of observation without transportations in an optimized configuration (see Fig.1). Note that the length of the periods of transportation is an average: it will be longer to reconfigure the instrument from the 4th intermediate configuration to the extended one than from an intermediate to another intermediate. It is of course always possible to observe during the days of transportation with less than 60 antennas.

The APO software is being developed to handle reconfiguration. It should allow to choose at each step of the reconfiguration which antennas should be moved and on which pads they should be fixed.

5 Conclusion

A design made of 6 optimized configurations is proposed for the ALMA project. It results from the optimization of the pad locations for some target distributions of samples. These target distributions were defined for the highest possible imaging quality and taking into account the recommendations of the PDR.

The optimized configurations show excellent imaging properties. All of them are able to sample most of the uv-plane at a rate higher than Nyquist rate. The Gaussian shapes of the distributions allow to get clean images with a minimum loss of sensitivity.

The resulting distribution of pads is almost continuous from the compact configuration up to the fourth intermediate configuration of 3 km in diameter. This allows to get hybrid configurations of good imaging quality (probably comparable to that of the optimized configurations). A great flexibility is therefore allowed in the reconfiguration strategy: a significant fraction of the observing time can be spent in the hybrid configurations.

In addition, the design proposed does not rely on any fixed geometrical concept. It is the output of a software able to take into account the topographic constraints and the target distributions analytically defined. It results from an automated process that can be rapidly executed. If the constraints or the specifications were to be changed, another result would be rapidly obtained. This point is essential as several iterations will probably be required before the construction of the pads really starts: the soil composition may appear unsuitable to the building of some of the pads, or some of the specifications may evolve. Thus, rather than a design concept it is a global approach of the design considered as an iterative process that is proposed here. The pad locations represented in the figures can be considered as the first step of this iterative process.

References

Boone, F. 2001a, in ALMA memo No. 400

—. 2001b, A&A, 377, 368

—. 2002, submitted

Briggs, D. S. 1995, in American Astronomical Society Meeting, Vol. 187, 11202+

Cornwell, T. J. 1988, A&A, 202, 316

Guilloteau, S. & the Configuration PDR Review Committee. 2001

Table 1: Characteristics of the configurations.

config.	largest baseline	FWHM of target dist.	duration of the obs.	sampling α	truncation level	resolution @100 GHz
compact	200 m	200 m	1.2 h	1	12 dB	$3.05''$
1rst	900 m	806 m	4.5 h	2	15 dB	$0.68''$
2nd	1600 m	1433 m	4.0 h	4	15 dB	$0.38''$
3rd	2300 m	2061 m	5.8 h	4	15 dB	$0.26''$
4th	3000 m	2670 m	7.5 h	4	15 dB	$0.20''$
extended	10 km	20 km	8 h	2	3 dB	$0.06''$

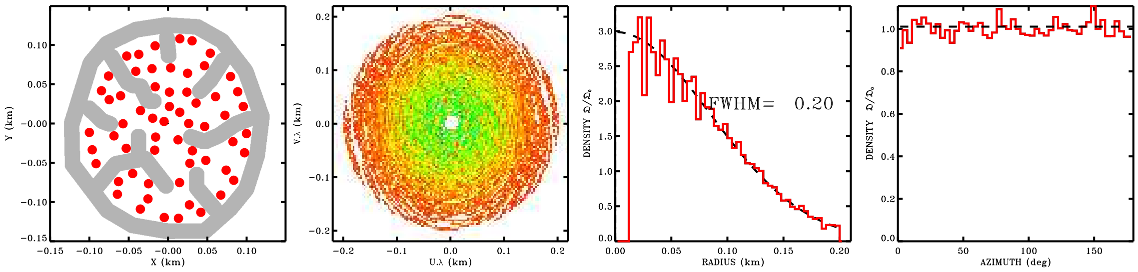

Figure 2: Compact configuration optimized for a Gaussian distribution of visibilities of FWHM= 200 m/λ (i.e. a truncation level of 12 dB) and for the mean source elevation $\epsilon_m = 60$ deg. The inner boundary condition prevents any shadowing for elevations down to 50 deg.

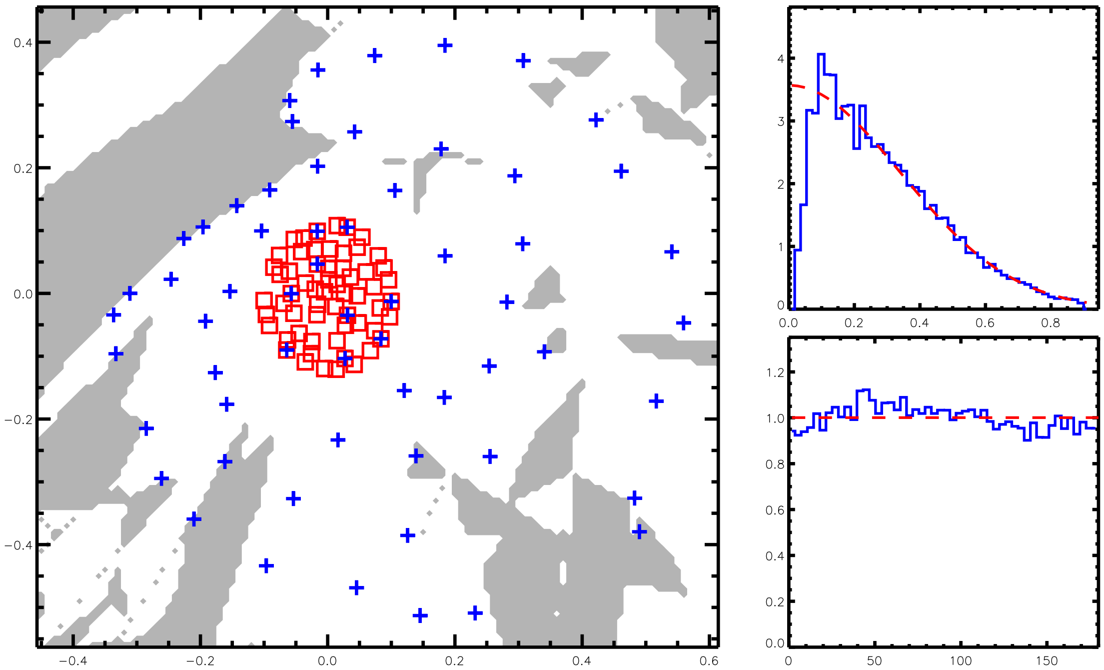

Figure 3: The blue crosses show the locations of the antennas in the first intermediate configuration. The red squares show the locations of the pads of the compact configuration. The figures on the right-hand side show the radial (top) and azimuthal (bottom) distribution of samples.

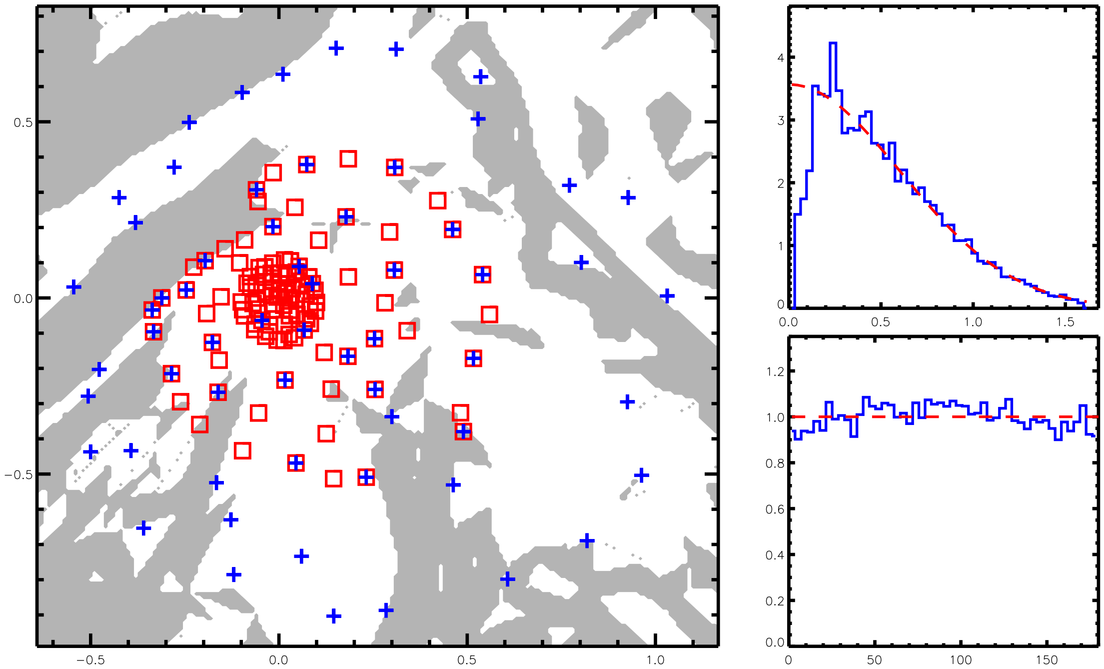

Figure 4: The blue crosses show the locations of the antennas in the second intermediate configuration. The red squares show the locations of the pads of the compact and first intermediate configurations. The figures on the right-hand side show the radial (top) and azimuthal (bottom) distribution of samples.

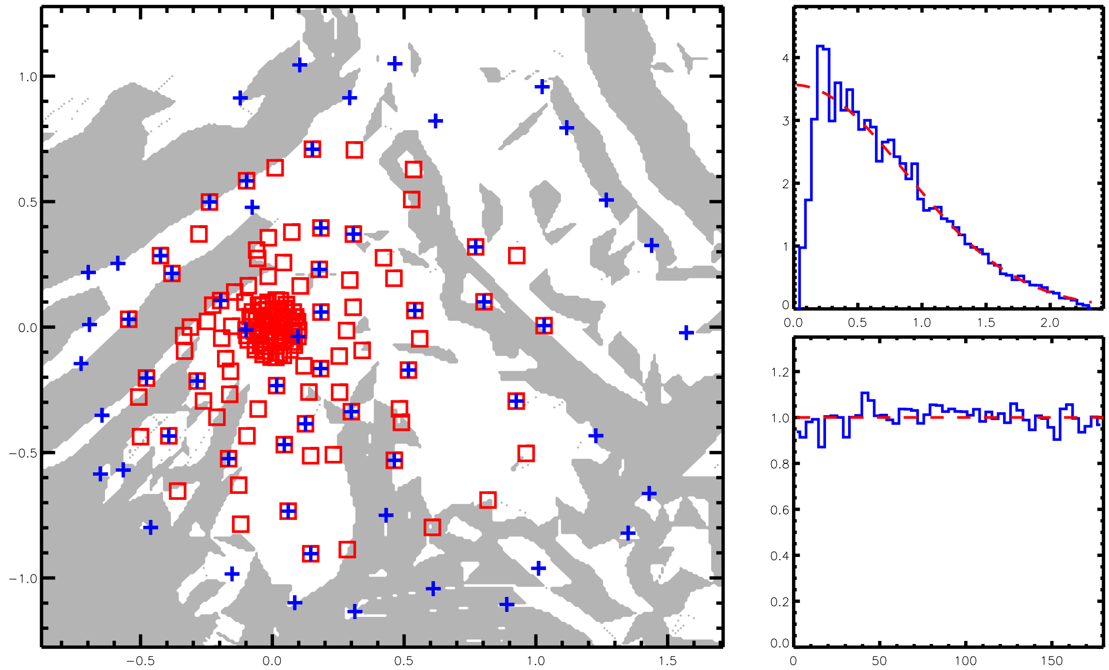

Figure 5: The blue crosses show the locations of the antennas in the third intermediate configuration. The red squares show the locations of the pads of the compact, the first and the second intermediate configurations. The figures on the right-hand side show the radial (top) and azimuthal (bottom) distribution of samples.

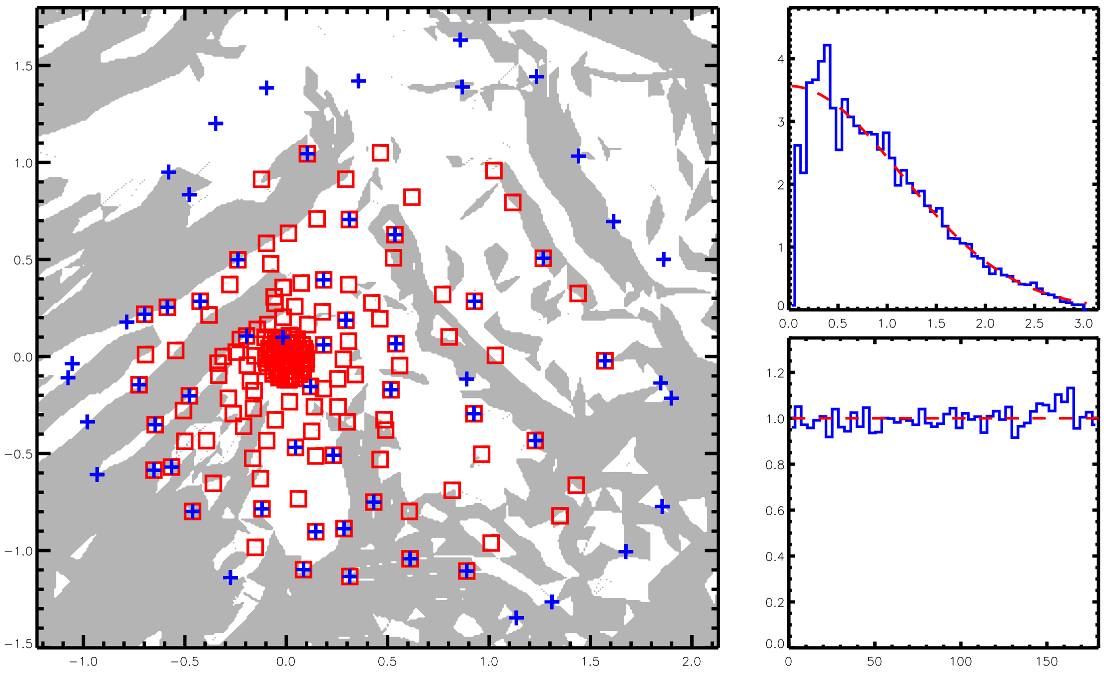

Figure 6: The blue crosses show the locations of the antennas in the fourth intermediate configuration. The red squares show the locations of the pads of the compact and the other intermediate configurations.

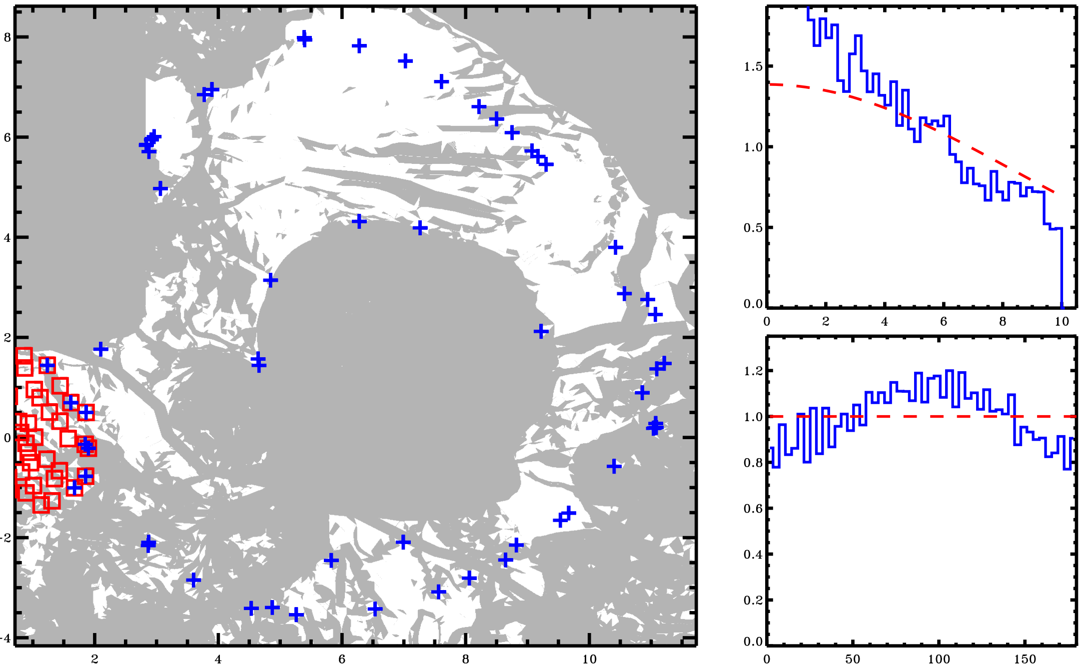

Figure 7: The blue crosses show the locations of the antennas in the extended configuration. The red squares show the locations of the pads of the compact and intermediate configurations. The figures on the right-hand side show the radial (top) and azimuthal (bottom) distribution of samples.