Chapitre 3

Détermination des distributions d'échantillons cibles (article 2)

Ce chapitre donne l'intégralité de l'article Boone (2002b) dans lequel une analyse est développée pour la détermination des distributions cibles d'échantillons de Fourier (voir introduction du chapitre précédent).

3.1 Introduction

The aim of this paper is to determine the distributions of uv-plane samples that should be taken as targets when optimizing an array. The choice of a target distribution depends upon the scientific goals purchased. In the present analysis we concentrate on the case where the goal is to make a reliable clean image of the source. The same problem was addressed in Holdaway (1996) and it was found that in many cases a centrally condensed distribution of uv-plane samples is better than a uniform coverage. Here, we would like to derive an exact and general description of the ideal distribution. More precisely, the question that is addressed is the following : how should be distributed the uv-plane samples of a given interferometer array in order to allow reliable imaging with the maximum of sensitivity at a given resolution and with a given level of side-lobes ? Note that the notion of reliability is subjective. For the following we will state that reliable imaging is performed when the visibility function is estimated everywhere in the uv-domain of interest with an error not much greater than the noise level in the measurements.

This paper is complementary to (Boone 2001, hereafter Paper I) which covers the optimization issue. Together, these two papers cover the problem of interferometric array design for imaging : the model distributions derived from the analysis developed here can be used as targets in the optimization method introduced in Paper I. This complete approach of the array design is applied to the ALMA project in Boone (2002a) which can be referred to for illustration.

To be complete, the analysis should take into account the possibility of interpolating and extrapolating in the uv-plane. Indeed, a great effort has been made since the beginning of interferometry to compensate for incomplete Fourier coverage and widely used algorithms such as CLEAN or MEM have proven their efficiency. The specification on the distribution of samples should be eased accordingly. However, evaluating the performance of each existing algorithm is a complex task (it also depends upon the kind of source observed as described in Cornwell et al. (1999)) which is beyond the scope of this paper. Instead, two parameters are defined which represent the level of interpolation and extrapolation that can be achieved. All the results are therefore very general : they can be adapted to any real situation by adjusting the parameters appropriately.

Note that the purpose of this analysis is different from the one of Woody (2001). In the latter reference a precise and thorough statistical analysis of the relationship between the sidelobes and the distribution of *uv*-plane samples of real arrays (pseudo-random arrays) was developed. Here, idealized distributions only are considered i.e. distributions that are not associated with any real array. They are only ideal solutions to which the real distributions should tend. Moreover, the aim is to define how they should be for some specifications on the clean map (i.e. after deconvolution) and the aspect of the dirty map is not considered. The main argument for this omission is, that the new generation arrays will provide high densities of samples (higher than Nyquist in most of the *uv*-plane). It will therefore be possible to remove the sidelobes due to small scale bumps in the distribution by adjusting appropriately the weights of the data and at a low expense in sensitivity. Also, it is probable that linear algorithms (e.g. Non Negative Least Squares), which are well adapted to complete coverage, will be more widely used (this was also suggested by Briggs (1995)), and such algorithms are less sensitive to the problem of sidelobes in the dirty map.

The structure of the paper is the following : in Sect.3.2 the parameters characterizing the process of image reconstruction (i.e. the level of interpolation and of extrapolation that can be achieved and the loss of sensitivity) are introduced. In Sect.3.3, the number of *uv*-plane samples required to allow a distribution of samples equal to the Fourier transform of the wanted clean beam is computed. It is shown in Sect.3.4 that when there are less samples than this limit the sampling constraint imposes a distribution of samples implying a loss of sensitivity. The relationship between the shape of the distribution minimizing the loss of sensitivity and the number of samples is derived. The feasibility of such distributions is also studied by means of simulations. In Sect.3.5 it is shown that, for given imaging specifications, the choice of the distribution is actually driven by the size of the instrument (i.e. the length of the largest baseline). Two characteristic array sizes are introduced and computed for various existing and future arrays. The conclusions are drawn in Sect.3.6.

3.2 Parameterization of the imaging process

As seen from Fourier plane, the image reconstruction process consists in the estimation of the visibility function from the measurements and in the weighting of this function according to the Fourier transform of the clean beam. In this section some definitions are given and some assumptions are made to allow a simplified description of the various processes involved. Two parameters are introduced to characterize the levels of interpolation and extrapolation that can be achieved in the uv-plane. The loss of sensitivity implied by the imperfection of the distribution of samples is also derived. All the notations defined along the paper are gathered in Appendix 3.7.

Let us first define the **density of samples** at a point (u, v) in the Fourier plane and considered at the scale δu as the number of samples situated in the disc $\delta s(u, v, \delta u)$ of center (u, v) and of radius $\delta u/2$, divided by the area of the disc :

$$\mathcal{D}(u, v, \delta u) = \frac{4}{\pi \delta u^2} \int_{\delta s(u,v,\delta u)} f(u, v)\, ds, \tag{3.1}$$

where f is the sampling function defined by :

$$f(u, v) = \sum_{i=1}^{N} \delta(u - u_i, v - v_i), \tag{3.2}$$

with N the number of samples and (u_i, v_i) the coordinates of the samples. The function $\mathcal{D}$ will often be referred to as the **distribution of samples**.

A region $\mathcal{S}$ is said **well sampled at δu scale** when there exists a regular grid of step δu such that, for each knot of the grid in $\mathcal{S}$, there is a sample at a distance less or equal $\delta u/4$. In terms of density it is equivalent to :

$$\begin{aligned} &\forall (u, v) \in \mathcal{S},\ \forall\, i < \sqrt{A}/\sqrt{\pi}\delta u, \\ &\mathcal{D}\left(u, v, (i + 1/2) \times \delta u\right) \geq \tfrac{4}{\pi \delta u^2} \left(\tfrac{i}{i+1/2}\right)^2 \end{aligned} \tag{3.3}$$

where A is the area of the uv-domain $\mathcal{S}$. Then, the **sampling accuracy**, δu_{s}, of a given sampling function within a given region $\mathcal{S}$ is defined as the smallest scale at which $\mathcal{S}$ is well sampled.

We shall assume in the following that interpolation, i.e. the use of the neighboring data for estimation of the visibility function at some coordinates in a uv-region $\mathcal{S}$, is possible with a reasonable precision (i.e. with an error close to the noise level) if and only if $\mathcal{S}$ is well sampled at δu_{M} scale i.e. if and only if $\delta u_{\mathrm{s}} \leq \delta u_{\mathrm{M}}$. In the absence of prior knowledge δu_{M} is the Shannon interpolation interval (aka Nyquist sampling interval). More generally, the value of δu_{M} depends on the prior knowledge of the source observed, on the performance of the algorithm used for interpolation (which is generally based itself on *a priori* properties of the astronomical sources) and whether a mosaic is made or not. δu_{M} will be called the **sampling accuracy required** for interpolation. For Shannon interpolation, which can be used when the size of the source ξ is the only information available, we have $\delta u_{\mathrm{M}} = 1/\xi$. The justification of this requirement is emphasized by some simulations in Appendix 3.8.

This assumption is a simplification that allows to consider image reconstruction independently of the method used. It is mostly inspired by the error behavior in Shannon interpolation but it is also based on the postulate that, to allow estimation of the brightness distribution at a given scale and along a given direction, a minimum of information is required on the structure of the source at nearby scales and along nearby directions. It is therefore believed that for any method used there is a condition that can be written like Eq. (3.3). For some interpolation techniques δu_{M} might be variable in the uv-plane but this would not change the results of the analysis developed here : the distributions considered in the following are all decreasing functions of the radius in the uv-plane and the limit δu_{M} actually concerns the density of samples at the edges only. (This can however alter the value of the sensitivity loss calculated below).

Estimation of the data outside the largest region $\mathcal{S}$ that is well sampled at δu_{M} scale is defined as extrapolation. Then, from the definition of δu_{M} even if some data are present outside $\mathcal{S}$ they cannot be used to estimate the visibility function, they are "lost". In the following we will assume that we always want a circularly symmetric beam and we will therefore only consider circular regions : Δu_{s} will refer to the radius of the largest well sampled region $\mathcal{S}$ and Δu_{e} will refer to the radius up to which the visibility function can be extrapolated. Note that extrapolation is considered as forbidden by some theoretical approaches. For example in the deterministic approach of Lannes et al. (1994), a condition for the stability of the solution is that the data outside the sampled uv-region be set to zero.

Let us then define two parameters α and β representing the level of interpolation and respectively extrapolation that can be achieved :

$$\alpha = \frac{\delta u_{\mathrm{M}}}{\delta u_{\mathrm{o}}}, \quad \text{and} \quad \beta = \frac{\Delta u_{\mathrm{e}}}{\Delta u_{\mathrm{s}}}, \tag{3.4}$$

where δu_{o} is the sampling accuracy required if no prior information can be used for interpolation : $\delta u_{\mathrm{o}} = 1/\Delta\theta_{\mathrm{PB}}$, with $\Delta\theta_{\mathrm{PB}}$ the angular size of the field of view (it actually uses the information specifying that the source as seen through the primary beam cannot be larger than $\Delta\theta_{\mathrm{PB}}$).

The constraint on the sampling accuracy is $\delta u_{\mathrm{s}} \leq \alpha\, \delta u_{\mathrm{o}}$. Setting α to a value greater than one implies that at least one of the following conditions is fulfilled : an effective algorithm can be used, some prior information on the source are available (e.g. the source size), a mosaic of the source is made. Mosaicing technique (see e.g., Cornwell 1988; Cornwell et al. 1993) can indeed allow to recover information at distances up to δu_{o} around the samples. This implies that the samples can be separated by $2\delta u_{\mathrm{o}}$ and therefore $\alpha = 2$. Note that when the source size is known to be half of the field of view we also have $\alpha = 2$.

The next important assumption made is that the accuracy on the estimate of the visibility function at some point in uv-plane depends on the number of samples $n(u,v)$ situated at a distance less than $\delta u_{\mathrm{M}}/2$ from this point. This assumption is actually related to the first one that concerned the sampling accuracy required : the maximum gap allowed is close to δu_{M} and it implies that the samples separated by such a distance are more or less independent. We will also assume for simplicity that all measurements are made with the same level of noise σ. More precisely it will be assumed that the error, $\Delta(u,v)$, on the estimate of the visibility function at (u,v) follows :

$$\Delta(u,v) = \frac{\sigma}{\sqrt{n(u,v)}} = \sigma \left(\frac{\pi \delta u_{\mathrm{M}}^2}{4} \mathcal{D}(u,v,\delta u_{\mathrm{M}}) \right)^{-1/2} \tag{3.5}$$

The value of the visibility function at a point (u,v) is involved in the reconstructed map with a weight $w(u,v)$. The weighting function w is merely the Fourier transform of the clean beam, it will be called the **clean weighting function**. (Note that it is different from the weight that can be given to the measurements : it is the weight of the estimated Fourier components after deconvolution.) Then, the error on the estimate of the flux of an unresolved source at the center of the map is (see e.g., Thompson et al. 2000; Wrobel & Walker 1999) :

$$\sigma_{\rm M}^2 = \frac{\int_{\mathcal{S}} w^2(u,v)\Delta^2(u,v)\,ds}{\left(\int_{\mathcal{S}} w(u,v)\,ds\right)^2}. \tag{3.6}$$

Hence, from Eq. (3.5), if the clean weighting function is normalized such that $\int_{\mathcal{S}} w\,ds = 1$:

$$\sigma_{\rm M}^2 = \sigma^2 \int_{\mathcal{S}} w^2(u,v)/n(u,v)\,ds. \tag{3.7}$$

The highest precision achievable with N samples is $\sigma_{\rm M} = \sigma/\sqrt{N}$. This is the case when $n(u,v) = Nw(u,v)$, i.e. when the samples are distributed according to the Fourier transform of the clean beam. Otherwise some sensitivity is lost : the instrument favored some Fourier components which did not require so much sensitivity relatively to the others. Note that we are considering the image reconstruction process as a whole not only the weighting process as it is usually understood (a process preceding the deconvolution). If the weighting of the data is indeed the only processing applied (no interpolation and no extrapolation) then the latter result corresponds to the well known fact that optimum sensitivity is achieved by "natural" weighting of the data. The relative loss of sensitivity is given by :

$$W = 1 - \frac{1}{\sqrt{N\int_{\mathcal{S}} w^2(u,v)/n(u,v)\,ds}} \tag{3.8}$$

In the following the shape of the distribution of samples to be used as target in the optimization of an array will be parametrized by α and β representing the interpolation and extrapolation levels that can be demanded to the imaging process. W will give an estimation of the loss of sensitivity due to the imperfection of the distribution. A perfect distribution obtained by setting α and β to one and for which $W = 0$ would allow instantaneous imaging with the full sensitivity of the instrument.

3.3 Full sensitivity imaging

In the following it will be assumed that the N uv samples can be distributed arbitrarily, whereas in fact there are only $n_{\rm a} - 1$ degrees of freedom. The characteristics of an ideal distribution of samples allowing to reconstruct images without loss of sensitivity are dictated by three independent considerations :

1. the image specifications in terms of beam : profile (Gaussian, spheroidal, ...) and level of side-lobes.

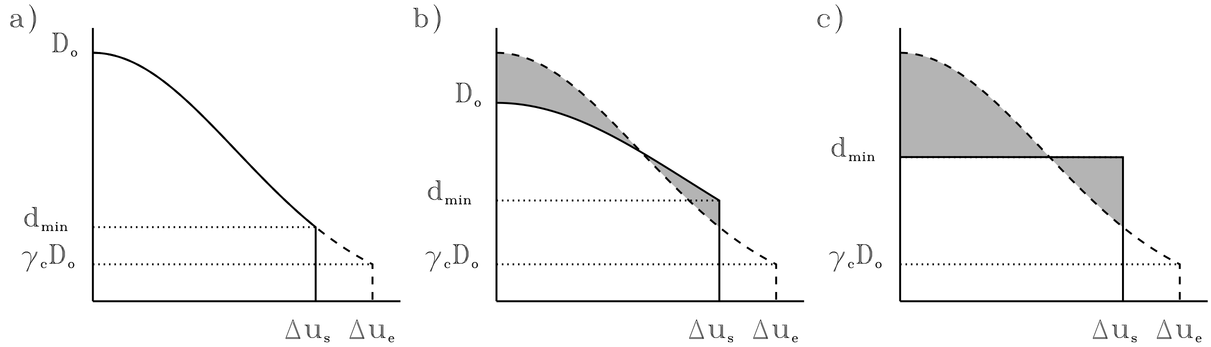

FIG. 3.1 – The solid lines show the radial profile of the model distribution of samples, $\mathcal{D}$. D_o stands for $\mathcal{D}(0)$. The dashed line shows the clean weighting function multiplied by the number of samples for a given β and a given γ_c. The difference between the model and clean weighting function which causes a loss of sensitivity is highlighted by a dark area. **a)** $N = N_o$, the distribution can be taken equal to the clean weighting function times the number of samples. **b)** $N_{\min} < N < N_o$, the Gaussian distribution which equals $d_{\min}$ at Δu_s minimizes the loss of sensitivity. **c)** $N = N_{\min}$, the distribution has to be uniform to ensure the sampling requirement to be fulfilled everywhere.

2. the levels of interpolation and extrapolation that can be achieved in uv-plane : α and β.
3. the SNR of the measurements.

The first point is clearly a constraint on the shape of the distribution of samples which should be the Fourier transform of the specified beam. Thus, if the beam is wanted to be Gaussian the distribution of samples should also be Gaussian. The acceptable level of side-lobes imposes the level of truncation, γ_c, of the clean weighting function. At the radius Δu_e up to which data is extrapolated the clean weighting function should obey $w(\Delta u_e)/w(0) = \gamma_c$. For simplicity we assume that the clean beam is uniform across the field of view unlike what can happen with deconvolution algorithms such as MEM.

The parameter, α, imposes the uv-disc to be well sampled at $\alpha\,\delta u_o$ scale. It constrains the distribution according to Eq. (3.3) with $\delta u = \alpha\,\delta u_o$. But we are now looking for idealized distributions of samples to be used as model for the optimization. We can therefore assume a perfectly smoothed distribution of samples and this condition reduces to : $\mathcal{D}(u, v, \alpha\,\delta u_o) \geq d_{min_1} = 4/\pi(\alpha\,\delta u_o)^2$. It follows that the model distribution should not decrease smoothly to zero at the edges but should rather be truncated when the density d_{min_1} is reached : the information gathered by samples distributed at a lower density would be useless (for the reliability of the map) as mentioned in the previous section. The factor, β, determines the radius up to which data can be extrapolated ($\Delta u_e = \beta \Delta u_s$) and therefore the shape of the distribution allowing to achieve the beam specifications within this radius.

The SNR of the data gives another constraint on the minimum density of samples in the case where an individual sample is too noisy to be usable. In this case more than one sample per $\pi(\alpha\,\delta u_o)^2/4$ area is necessary and the minimum density is $d_{min_2} > d_{min_1}$. In general this is not the case and in the following $d_{\min}$ will refer to d_{min_1}.

From now on it is supposed that the wanted beam is Gaussian but all the calculations can be made in a similar way for any other function. Then, the clean weighting function is also a Gaussian of FWHM δ_c such that $w(\Delta u_\mathrm{e})/w(0) = \gamma_\mathrm{c}$ that is :

$$\delta_\mathrm{c} = \beta \Delta u_\mathrm{s} / a \sqrt{-2 \ln \gamma_\mathrm{c}}, \tag{3.9}$$

with $a = 1/2\sqrt{2 \ln 2}$. Thus, specifying β and γ_c is equivalent to specifying a FWHM for the clean weighting function, i.e. a resolution for the clean map. That is the way it is usually done when restoring a map after CLEAN : a Gaussian is fitted to the dirty beam and its FWHM fixes the resolution of the clean beam.

The distribution of samples is ideal (the loss of sensitivity is minimum) when $\mathcal{D}(u,v) = N\,w(u,v)$, but the density at the edges (at radius Δu_s) should be greater than $d_{\min} = 4/\pi(\alpha\,\delta u_\mathrm{o})^2$. This condition can be satisfied when the number of samples is greater than N_o given by :

$$N_\mathrm{o} = \frac{-\pi}{\ln \gamma_\mathrm{c}} \left(\frac{\beta}{\alpha}\right)^2 \left[\exp\left(-\frac{\ln \gamma_\mathrm{c}}{\beta^2}\right) - 1\right] \left(\frac{\Delta u_\mathrm{s}}{\delta u_\mathrm{o}}\right)^2 . \tag{3.10}$$

If B is the length of the largest baseline of the array projected on the sky, then the radius of the sampled uv-region assumed to be a disc is $\Delta u_\mathrm{s} = B/\lambda$, where λ is the observing wavelength. At this wavelength the full angular size of the primary beam is related to the antenna diameter, D, by : $\Delta\theta_\mathrm{PB} = 2\lambda/D$. Hence,

$$\frac{\Delta u_\mathrm{s}}{\delta u_\mathrm{o}} = \Delta u_\mathrm{s} \Delta\theta_\mathrm{PB} = \frac{2B}{D} . \tag{3.11}$$

Then, Eq. (3.10) becomes :

$$N_\mathrm{o} = 4\pi \, \frac{\beta^2}{b\Gamma_\mathrm{c}} \left[\exp\left(\frac{b\Gamma_\mathrm{c}}{\beta^2}\right) - 1\right] \left(\frac{B}{\alpha D}\right)^2 \tag{3.12}$$

where $b = \ln 10 \,/\, 10$ and Γ_c is the truncation level, γ_c, expressed in dB. For illustration, the condition for an interferometer array to allow imaging without relying on any deconvolution method or *a priori* knowledge on the source ($\alpha = \beta = 1$) by means of a Gaussian distribution of samples truncated at 10 dB is :

$$N > N_\mathrm{o} = 49.1 \left(\frac{B}{D}\right)^2 . \tag{3.13}$$

3.4 Constrained target distributions

If the number of samples is less than N_o all the specifications cannot be satisfied and some of them have to be relaxed. However, one should stick as long as possible to the sampling condition for imaging and use a model distribution with a density higher

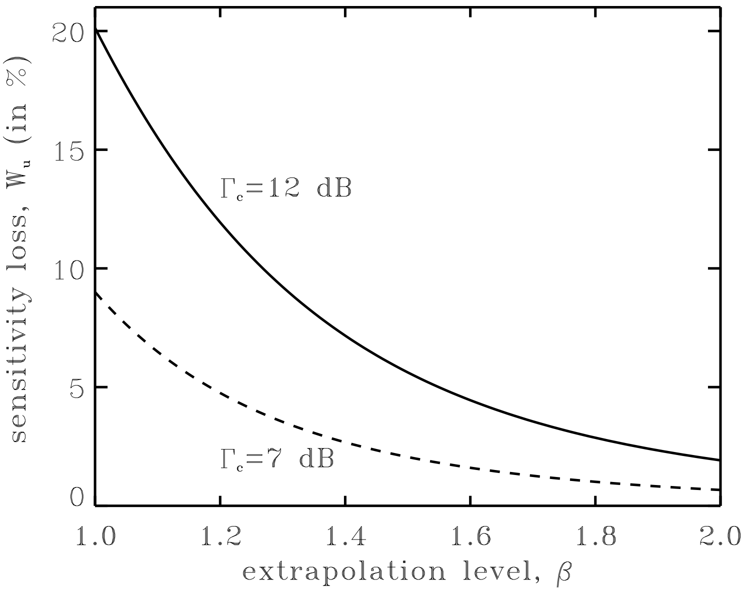

Fig. 3.2 – Loss of sensitivity due to the weighting of the estimated visibility function in the clean map when the initial distribution of samples is uniform ($N = N_{\mathrm{min}}$) and for a cut-off of the clean Gaussian distribution at 12 dB and 7 dB.

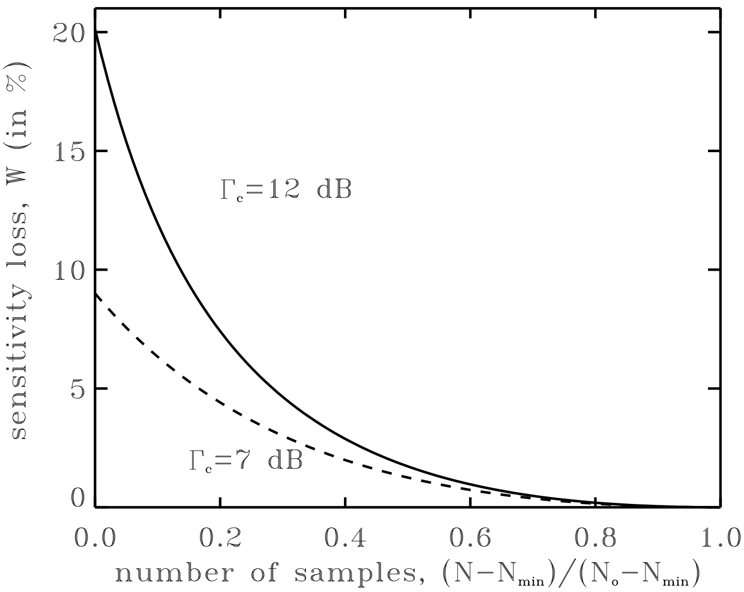

Fig. 3.3 – Sensitivity loss as a function of the number of samples in the case where $\beta = 1$ and for a cut off at 12 dB and 7 dB.

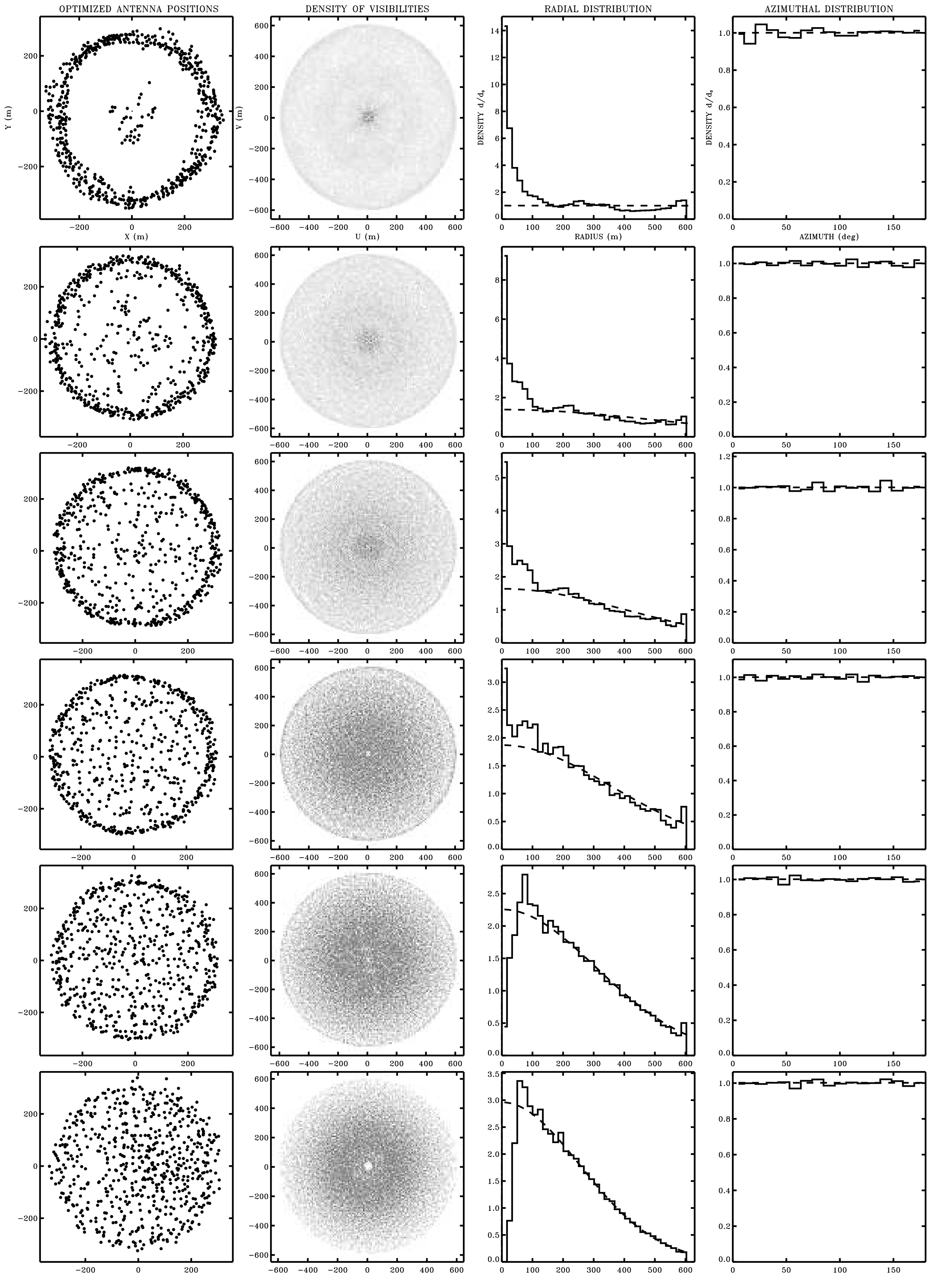

FIG. 3.4 – Optimization of 10 configurations of 64 antennas for $\delta_s/\Delta u_s = 40$, 2, 1.6, 1.4, 1.2, 1. In each row is represented a superposition of the 10 optimized configurations, the average density of samples from these 10 configurations, its radial and azimuthal profile

than d_{min}. Breaking this condition would imply the lack of some essential information and would prevent any reliable image reconstruction. The minimum number of samples required for the uv-disc to be well sampled at $\alpha\,\delta u_{\mathrm{o}}$ scale is :

$$N_{\mathrm{min}} = \pi \Delta u_{\mathrm{s}}{}^2 d_{\mathrm{min}} = 4\pi \left(\frac{B}{\alpha D}\right)^2 . \tag{3.14}$$

If $N = N_{\mathrm{min}}$ the samples have to be distributed uniformly across the uv-disc to satisfy the sampling constraint everywhere (Fig.3.1c). From Eq. (3.8) and Eq. (3.14) the relative loss of sensitivity due to the difference between the distribution and the clean weighting function is :

$$W_u = 1 - \frac{\beta^2}{b\Gamma_{\mathrm{c}}} \sqrt{2\,\frac{1-\exp\frac{-b\Gamma_{\mathrm{c}}}{\beta^2}}{1+\exp\frac{-b\Gamma_{\mathrm{c}}}{\beta^2}}} . \tag{3.15}$$

The variation of this loss as a function of the extrapolation parameter β is drawn Fig.3.2 for $\Gamma_{\mathrm{c}} = 12$ dB and 7 dB. This figure shows clearly that a uniform distribution of samples would imply a non-negligible loss of sensitivity ($\gtrsim 10\%$) unless extrapolation is possible up to more than 20% the sampled uv-disc radius.

When $N_{\mathrm{min}} < N < N_{\mathrm{o}}$ the ideal distribution is not attainable but the sampling constraint can be fulfilled and the $N - N_{\mathrm{min}}$ samples can be distributed in a way that minimizes the inevitable loss of sensitivity. It can be shown from the calculus of variations that the integral $\int_{\mathcal{S}} w^2(u,v)/n(u,v)\, ds$ in Eq. (3.8) is minimized when $w^2(u,v)/n(u,v)$ is a Gaussian function. Therefore, if $w(u,v)$ is a Gaussian, the loss of sensitivity is minimized when $n(u,v)$ is itself a Gaussian, i.e. when the distribution of samples is Gaussian. Let us write δ_{s} the FWHM of such a distribution. Its value is imposed by the requirement that the density of samples be greater or equal d_{min} at the edges and that the integral of the distribution be equal N. Hence, δ_{s} is solution of :

$$\int_0^{\Delta u_{\mathrm{s}}}\int_0^{2\pi} d_{\mathrm{min}} \exp\left(-\frac{r^2 - \Delta u_s{}^2}{2a^2 \delta_{\mathrm{s}}{}^2}\right) r\, d\theta\, dr = N, \tag{3.16}$$

leading to :

$$\frac{\beta^2}{b\Gamma_{\mathrm{c}}}\frac{\delta_{\mathrm{s}}^2}{\delta_{\mathrm{c}}^2}\left[\exp\left(\frac{b\Gamma_{\mathrm{c}}}{\beta^2}\frac{\delta_{\mathrm{c}}^2}{\delta_{\mathrm{s}}^2}\right) - 1\right] = \frac{N}{N_{\mathrm{min}}}, \tag{3.17}$$

where δ_{c} is the FWHM of the clean weighting function given by Eq. (3.9). Then, from Eq. (3.8), the loss of sensitivity is :

$$W = 1 - \frac{\delta_{\mathrm{c}}^2\left[1-\exp\frac{-b\Gamma_{\mathrm{c}}}{\beta^2}\right]}{\delta_{\mathrm{s}}\delta_{\mathrm{cs}}\sqrt{\left[1-\exp\frac{-b\Gamma_{\mathrm{c}}}{\beta^2}\frac{\delta_{\mathrm{c}}^2}{\delta_{\mathrm{s}}^2}\right]\left[1-\exp\frac{-b\Gamma_{\mathrm{c}}}{\beta^2}\frac{\delta_{\mathrm{c}}^2}{\delta_{\mathrm{cs}}^2}\right]}}, \tag{3.18}$$

with $\delta_{\mathrm{cs}}^2 = (2/\delta_{\mathrm{c}}^2 - 1/\delta_{\mathrm{s}}^2)^{-1}$. The loss is minimum when $\delta_{\mathrm{s}} = \delta_{\mathrm{c}}$, i.e. when the distribution of samples has the same shape as the clean weighting function.

Equation 3.17 has been solved for δ_s and the loss computed according to Eq. (3.18). The loss is plotted versus the number of samples in Fig.3.3 for $\beta = 1$ (no extrapolation), $\Gamma_c = 12$ dB and $\Gamma_c = 7$ dB.

For a given uv-disc radius, the less the number of samples the greater the FWHM of the Gaussian distribution. More generally, whatever the kind of model distribution chosen, as the number of samples decreases and tends to $N_{\rm min}$, it should tend to be uniform to ensure the density to be higher than $d_{\rm min}$ everywhere. However, as mentioned in Paper I it is not possible to get a uniform distribution of samples. Furthermore, for a number of samples close to the minimum required, $N_{\rm min}$, non-uniformity implies necessarily under sampling of some regions. The number of samples actually required to ensure the uv-disc to be well sampled is therefore greater than $N_{\rm min}$ as given Eq. (3.14). In order to evaluate the largest ratio $\delta_s/\Delta u_s$ (i.e. the maximum flatness of the distribution) allowing the sampling requirement to be fulfilled, and the corresponding number of samples required $N'_{\rm min}$, 10 configurations have been optimized for various ratios $\delta_s/\Delta u_s$. The optimizations have been performed using the APO software introduced in Paper I. The results for $\delta_s/\Delta u_s = 40$, 2, 1.6, 1.4, 1.2 and 1 are drawn Fig.3.4. From these simulations, the highest ratio $\delta_s/\Delta u_s$ allowing the density of samples to be greater than $d_{\rm min}$ everywhere is estimated to be around 2. From Eq. (3.17) it can be shown that the number of samples required, $N'_{\rm min}$, for such a model distribution to be feasible is :

$$N'_{\rm min} = 1.44\, N_{\rm min}\,. \tag{3.19}$$

It can be noted that for $\delta_s/\Delta u_s < 2$, if the sampling constraint is satisfied, the actual distribution of samples is still different from the model and this difference increases with $\delta_s/\Delta u_s$. It is then clear that in practice the loss of sensitivity can increase more rapidly than shown Fig.3.3 for N decreasing toward $N'_{\rm min}$. This effective loss is however difficult to estimate because the quality of the distribution, i.e. the resemblance to the model distribution, depends upon the position of the source, the latitude of the array site and the time of integration (Paper I). The theoretical value of Eq. (3.8) is a lower limit.

3.5 Parameterization by the size of the array

In the previous section it was shown that for a given uv-disc radius the shape of the distribution of samples (i.e. the FWHM for a Gaussian distribution) is determined by the number of samples available. But the radius of the uv-disc depends on the size of the instrument (i.e. the largest projected baseline of the instrument) and the number of samples depends on the size of the instrument and on the duration of the observation. Therefore, for a given duration of observation the shape of the distribution of samples is entirely determined by the size of the instrument. If the size of the array is less or equal a value B_o, there are enough samples for the distribution to be ideal : its FWHM can be the same as that of the clean weighting function. For greater sizes the FWHM has to be adapted for the sampling constraint to be fulfilled everywhere. For a size greater than a value $B_{\rm max}$ the sampling constraint cannot be fulfilled anymore. In this section the

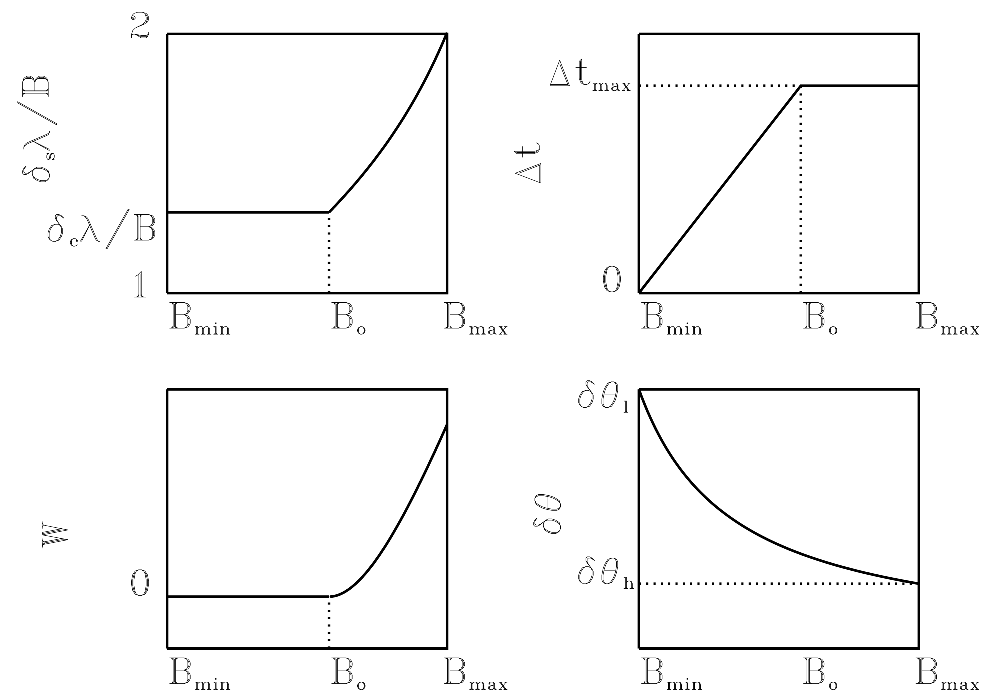

Fig. 3.5 – The FWHM of the model distribution in uv-disc radius units, $\delta\lambda/B$, the duration of observation required, Δt, the loss of sensitivity, W, and the resolution of the clean map, $\delta\theta$, as functions of the largest baseline, B.

Tab. 3.1 – Characteristic array sizes as given by Eq. (3.24), Eq. (3.25) and Eq. (3.26) for various existing or future instruments and for $\alpha = \beta = 1$, i.e. without input of prior information in the imaging process, and $\Delta t_{\rm max}$ =8 hour. B_{10dB} and B_{7dB} are the largest baselines the instrument can have to yield a Gaussian distribution of samples truncated at respectively 15 dB and 10 dB as given by Eq. (3.25). $B_{\rm max}$ is the largest baseline the instrument can have to ensure full sampling. δt is the maximum averaging time for the value of $B_{\rm max}$ to hold. It gives a constraint on the source flux through Eq. (3.22).

array	D	$n_{\rm a}$	$B_{\rm min}$	B_{15dB}	B_{10dB}	$B_{\rm max}$	δt
PdB	15m	6	37m	8m	19m	52m	3960s
BIMA	6m	9	18m	8m	18m	50m	1650s
VLA	25m	27	130m	330m	750m	2.0km	169s
ALMA	12m	64	100m	910m	2.1km	5.6km	29.5s
ATA	4m	350	75m	9.2km	21km	57km	1.0s

two characteristic baselines $B_{\rm o}$ and $B_{\rm max}$ are computed and the value of the FWHM as a function of the size of the instrument B is derived for $B_{\rm o} < B < B_{\rm max}$.

The number of samples depends on the number of antennas, $n_{\rm a}$, the duration of the observation, Δt, and the averaging time interval, δt. This last parameter has a specific meaning in the present context. Indeed, it was assumed that the density of samples could reach the minimum value $d_{\rm min}$ at the edges of the uv-disc. Then, δt should be such that neighbor samples measured with the same antenna pair be separated by $\alpha\,\delta u_{\rm o}$ (note that it is a theoretical constraint on the way to count the samples, and not on the way to effectively perform the measurements). During an observation a baseline describes an elliptical track in uv-plane (e.g., Thompson et al. 2000). When the source is at the north or south pole the curve is a circle and the rotation velocity is : $\omega_{\rm E} = \pi/12$ rd hour^{-1}. But most of the time the track is not circular and the linear velocity of the baseline is less than $\Delta u_{\rm s}\omega_{\rm E}$. This linear velocity is difficult to evaluate in a general way as it depends upon the source position, the latitude of the array site and the baseline orientation. However it is sufficient for the following analysis to use an average velocity, $\Delta u_{\rm s}\omega_{\rm E}/2$. Then δt is given in hour by :

$$\delta t = \alpha\,\delta u_{\rm o}\frac{2}{\Delta u_{\rm s}\omega_{\rm E}} = \frac{12\,\alpha D}{\pi B}\,. \tag{3.20}$$

It must be reminded here that the SNR of one sample was assumed to be sufficiently high (say > 3) for the information to be usable. This implies the averaging time interval, δt to be higher than a value $\tau_{\rm a}$ depending upon the antenna temperature, $T_{\rm A}$, the system temperature, $T_{\rm S}$, and the bandwidth of the intermediate frequency, $\Delta\nu_{I\!F}$. For a point source, the averaging time required to get SNR=3 is given by (Thompson et al. 2000) :

$$\tau_{\rm a} = \frac{9}{2\Delta\nu_{I\!F}}\left(\frac{T_{\rm S}}{T_{\rm A}}\right)^2 \tag{3.21}$$

for two polarizations. In other words Eq. (3.20) holds only for sources yielding an antenna temperature, $T_{\rm A}$, such that :

$$T_{\rm A} > \frac{3T_{\rm S}}{\sqrt{2\Delta\nu_{I\!F}\delta t}}\,. \tag{3.22}$$

In this case the number of samples is given by :

$$N \equiv n_{\rm a}(n_{\rm a}-1)\frac{\Delta t}{\delta t} = n_{\rm a}(n_{\rm a}-1)\frac{\pi\,B\,\Delta t}{12\,\alpha\,D}\,, \tag{3.23}$$

where Δt is expressed in hour. Then, the largest size of the instrument, $B_{\rm max}$, allowing to get $N'_{\rm min}$ samples is solution of $N = N'_{\rm min}$, where N is the number of samples obtained for the maximum duration of observation $\Delta t_{\rm max}$. From Eq. (3.14), Eq. (3.19) and Eq. (3.23), we have :

$$B_{\rm max} = \frac{n_{\rm a}(n_{\rm a}-1)\,\alpha\,D\,\Delta t_{\rm max}}{48\times 1.44}\,. \tag{3.24}$$

Similarly from Eq. (3.12) and Eq. (3.23) the largest baseline, $B_{\rm o}$, of an instrument allowing to get $N_{\rm o}$ samples is given by :

$$B_o = \frac{1.44}{\chi}B_{\rm max} = \frac{n_{\rm a}(n_{\rm a}-1)\,\alpha\,D\,\Delta t_{\rm max}}{48\,\chi} \tag{3.25}$$

with $\chi = \beta^2/b\Gamma_\mathrm{c}\,(\exp(b\Gamma_\mathrm{c}/\beta^2) - 1)$. B_max and B_o are realistic instrument sizes as long as they are larger than the shortest diameter an array can be packed into. This diameter is approximately given by :

$$B_\mathrm{min} = \sqrt{n_\mathrm{a}} D\,. \tag{3.26}$$

It can be noted that B_max, B_o and B_min are independent of the observing wavelength. They depend upon the instrument only through the number of antennas and their diameter. B_max and B_o are proportional to the square of the number of antennas (in the approximation of large n_a). It is consequently much more efficient (but also more costly) to increase the number of antennas than their diameter to increase the imaging resolution of an array.

Finally, the shape of the model distribution of samples for an interferometer array of a given size, B, is determined as follows :

1. if $B < B_\mathrm{o}$ as given Eq. (3.25), the array can be optimized for the Gaussian distribution of FWHM given by Eq. (3.9) :

$$\delta_\mathrm{c} = \frac{\beta}{a\sqrt{2b\Gamma_\mathrm{c}}}\frac{B}{\lambda}\,. \tag{3.27}$$

   From Eq. (3.25) the time of observation required is then given by :

$$\Delta t = \frac{48\,\chi\,B}{n_\mathrm{a}(n_\mathrm{a}-1)\,\alpha\,D}\,. \tag{3.28}$$

   If the model distribution can be closely approached by the real distribution of samples there is little loss of sensitivity in producing clean maps.

2. if $B_\mathrm{o} < B < B_\mathrm{max}$ as given Eq. (3.24), the time of observation is maximum ($\Delta t = \Delta t_\mathrm{max}$) and the FWHM, δ_s, of the model Gaussian distribution is solution of Eq. (3.17) where $N/N_\mathrm{min} = B/1.44 \times B_\mathrm{max}$ (from Eq. (3.23)) and δ_c is given by Eq. (3.27). Hence, δ_s is solution of :

$$\frac{2a^2\lambda^2\delta_\mathrm{s}^2}{B^2}\left[\exp\left(\frac{B^2}{2a^2\lambda^2\delta_\mathrm{s}^2}\right) - 1\right] - \frac{B}{1.44\,B_\mathrm{max}} = 0 \tag{3.29}$$

   As $\delta_\mathrm{s} > \delta_\mathrm{c}$, the truncation level is higher than specified for the clean weighting function. To get the clean image specified some sensitivity is necessarily lost. The theoretical loss given by Eq. (3.8) is a lower limit to the actual one.

3. if $B > B_\mathrm{max}$, the Fourier plane is under sampled and in principle a reliable image reconstruction is not possible. To overcome this limitation, more samples should be obtained by observing the source with several configurations of antennas.

In each case the resolution of the clean map is related to the largest baseline length by :

$$\delta\theta_\mathrm{c} = \frac{4\ln 2}{\pi\delta_\mathrm{c}} = \frac{2\sqrt{b\ln 2}}{\pi}\frac{\Gamma_\mathrm{c}^{\frac{1}{2}}}{\beta}\frac{\lambda}{B}\,. \tag{3.30}$$

Note that if $\beta = 1$ (no extrapolation) the resolution is higher than λ/B (i.e. $\delta\theta_c < \lambda/B$) for a truncation level of the distribution of samples, Γ_c, such that $\Gamma_c < 15.5$ dB. From Eq. (3.24) the highest clean map resolution an interferometer can provide if $\Delta t_{max} = 8$h is :

$$\delta\theta_{ch} = 2.2 \, \frac{\Gamma_c^{\frac{1}{2}}}{\alpha\,\beta} \, \frac{\lambda}{n_a(n_a - 1)\,D} \,. \tag{3.31}$$

These results are summarized Fig.3.5 where the variations of, δ_s, Δt, W and $\delta\theta_c$ versus the largest baseline length are plotted. Increasing the baseline length allows to increase the resolution but when $B > B_o$ the improvement of the resolution is paid for by a loss of sensitivity : there is a trade-off between resolution and sensitivity.

The baseline lengths B_{min}, B_o and B_{max} are given in Table 3.1 for some existing or future interferometers and when reliance upon deconvolution is minimized ($\alpha = \beta = 1$). The values show that imaging with instruments like PdB or BIMA with only one configuration, even if optimally designed, is generally paid by a loss of sensitivity ($B > B_o$) if no efficient deconvolution algorithms are used (i.e. if we are not in the case where $\alpha > 1$ and/or $\beta > 1$). The VLA could in principle produce a Gaussian beam with the Fourier plane sampled at Nyquist accuracy for baselines up to 1km. Clearly, for new generation interferometers such as ALMA and ATA, the possibility of imaging without deconvolution and loss of sensitivity will become a reality for baseline lengths up to respectively 3km and 30km.

3.6 Conclusion

For a limited duration of observation, the shape of the distribution of samples that should be taken as target is determined by the size of the instrument. For an array of n_a antennas of diameter D two characteristic sizes independent of the observing wavelength were introduced : B_{max} and B_o given by Eqs. (3.24) and (3.25) respectively. Their value depend on the effort that can be demanded to the interpolation (sampling accuracy required) and B_o also depends on the level of extrapolation allowed and on the level of side-lobes specified (or the truncation level of the clean weighting function). For an instrument smaller than B_o it is possible to achieve a distribution of samples implying no (or a little in practice) loss of sensitivity. For a size between B_o and B_{max} the sampling requirement imposes a distribution of samples which shape induces a loss of sensitivity. For a Gaussian clean beam the sensitivity loss is minimized when the distribution of samples is Gaussian. The FWHM of the distribution is solution of Eq. (3.29). The closer the size of the array to B_{max} the flatter the distribution (the higher the ratio of the FWHM by the sampled uv-disc radius). In terms of array shapes it implies that the larger the array the more it should resemble a ring (see Fig.3.4). On the contrary the smaller the array the more the configuration should be centrally condensed. For sizes greater than B_{max} imaging with a single configuration is in principle not possible. This gives a limit to the imaging resolution attainable with a single configuration, δ_{ch}, given by Eq. (3.31) when the duration of observation is limited to eight hours.

This result may appear opposite to the conclusion of Woody (2001) or Holdaway (1997), namely, that large arrays should be filled or centrally condensed and arrays made of a small number of antennas should be ring-like. Actually, both conclusions are in agreement : in our conclusion above a fixed number of antennas was considered and the array size was the free parameter. But the values of B_{max} and B_{o} are proportional to the square of the number of antennas. Therefore, increasing the number of antennas implies the possibility of getting a Gaussian distribution with good coverage up to large configurations (precisely up to B_{o}) i.e. the configurations can be filled or centrally condensed. When the number of antennas is small, B_{o} is small and for configurations larger than B_{o} the distribution should be more uniform, i.e. the configuration more ring-like.

The size, B, of the array to be designed depends on the resolution wanted $\delta\theta_{\mathrm{c}}$ through Eq. (3.30). Finally, from the expressions of B, B_{o}, B_{max} and the equation giving the FWHM (Eq. (3.29)), the shape of the distribution of samples, and therefore the shape of the array itself, is entirely determined by a set of 5 specifications :

$$\{\alpha, \beta, \gamma_{\mathrm{c}}, \delta\theta_{\mathrm{c}}, \lambda\}, \tag{3.32}$$

where α and β represent the effort that can be demanded to respectively the interpolation process and the extrapolation process as defined Sect.3.2 and γ_{c} is the truncation level of the clean weighting function (Sect.3.3) related to the level of side lobes tolerated in the clean map.

It was also shown that for new generation interferometers such as ALMA or ATA the maximum size allowed for imaging without demanding any effort to interpolation and with a single configuration was of several kilometers (5.6 km for ALMA and 56 km for ATA). Even B_{o}, the size below which it is possible to image without losing any sensitivity, is large (see Table 3.1). It implies the possibility of performing images in a way comparable to optical imaging systems (e.g. optical telescopes) : without particular algorithm of image reconstruction and with the full sensitivity of the instrument. For example, with ALMA it will be possible to get a Gaussian distribution of samples truncated at 10 dB up to 2.1 km (Table 3.1) which corresponds to a resolution of 0.23 asec at 1 mm (Eq. (3.30)).

It should however be noted that the so called *short spacings problem* was ignored throughout the analysis. It was assumed that a real distribution of samples can follow a Gaussian distribution down to the center of the uv-plane. Due to the physical size of the antennas this assumption is obviously never verified : the distribution necessarily decreases around the center and reaches zero at a radius greater or equal $(D + x)/\lambda$, where x is the minimum spacing required between the sides of neighboring antennas. But there exist some solutions such as mosaicing technique, or the use of complementary observations from a single dish telescope or another array of smaller antennas. Then, the great potentiality of the new generation arrays as depicted here requires at least one of these solutions to be applied.

3.7 Notations

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

$b =$ ln10 /10.

B **:** largest projected baseline length of the array also considered as the size of the array.

$B_{\mathbf{o}}$ **:** maximum size of the array allowing an ideal distribution of samples.

$B_{\mathbf{max}}$ **:** maximum size of the array allowing the sampling requirement to be fulfilled.

$d_{\mathbf{min}}$ **:** minimum density allowed in a model distribution of samples to ensure the sampling requirement to be fulfilled.

$\mathcal{D}(u, v, \delta u)$ **:** density of samples at the point (u, v) considered at the scale δu. $\mathcal{D}$ is the distribution of samples.

$f(u, v)$ **:** sampling function.

$n(u, v)$ **:** number of samples in the disc of radius $\delta_{\mathrm{M}}/2$ centered on (u, v).

$n_{\mathbf{a}}$ **:** number of antennas.

N **:** number of samples.

$N_{\mathbf{min}}$ **:** minimum number of samples required for the sampling requirement to be fulfilled everywhere in the uv-disc when the uniform distribution is assumed to be feasible.

$N'_{\mathbf{min}}$ **:** minimum number of samples required for the sampling requirement to be fulfilled everywhere if realistic distributions only are considered.

$N_{\mathbf{o}}$ **:** minimum number of samples required to allow a distribution of samples equal to the Fourier transform of the specified clean beam.

$w(u, v)$ **:** weight given to the value of the visibility function at (u, v) in the clean map, w is the clean weighting function.

W **:** loss of sensitivity.

α **:** maximum level of interpolation that can be achieved : $\alpha = \delta u_{\mathrm{s}}/\delta u_{\mathrm{o}}$

β **:** maximum level of extrapolation that can be achieved : $\beta = \Delta u_{\mathrm{M}}/\Delta u_{\mathrm{s}}$

$\chi = \beta^2/b\Gamma_{\mathrm{c}}\,(\exp(b\Gamma_{\mathrm{c}}/\beta^2) - 1)$.

$\delta_{\mathbf{c}}$ **:** FWHM of the clean weighting function.

$\delta_{\mathbf{s}}$ **:** FWHM of the distribution of samples.

δt **:** averaging time for a single measurement.

$\delta\theta_{\mathbf{c}}$ **:** angular resolution of the clean map.

$\delta\theta_{\mathbf{ch}}$ **:** the highest angular resolution of the clean map attainable by an array.

$\delta u_{\mathbf{s}}$ **:** sampling accuracy of a given distribution of samples.

$\delta u_{\mathbf{M}}$ **:** sampling accuracy required for interpolation.

$\delta u_{\mathbf{o}} = 1/\Delta\theta_{\mathrm{PB}}$.

$\Delta(u, v)$ **:** error on the estimate of the visibility function at (u, v).

Δt **:** duration of the observation.

$\Delta t_{\mathbf{max}}$ **:** maximum possible duration of the observations.

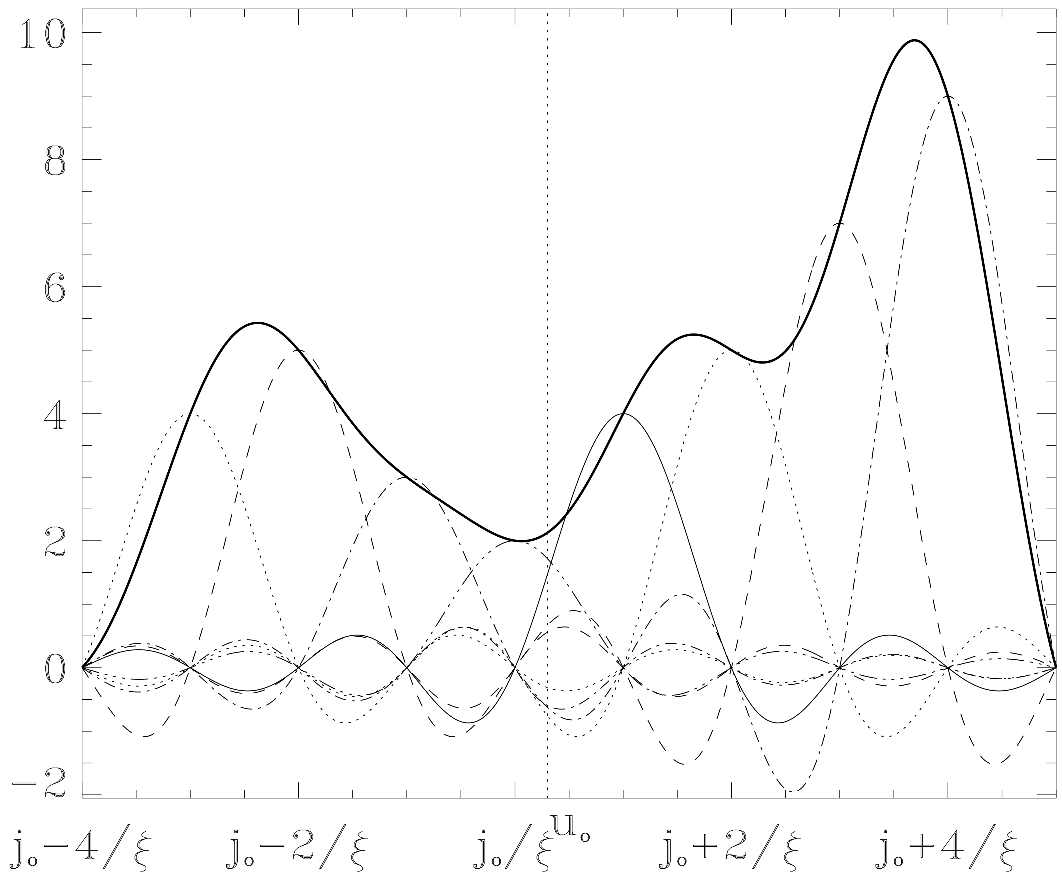

FIG. 3.6 – Illustration of Eq. (3.34) when only the 8 sinc functions the closest to a point u_o are considered. The visibility function is represented by the thick line which corresponds to the sum of the sinc functions scaled by some parameters.

$\Delta\theta_{\mathbf{PB}}$ **:** angular size of the primary beam.

$\Delta u_{\mathbf{s}}$ **:** radius of the sampled *uv*-disc.

$\Delta u_{\mathbf{e}}$ **:** radius up to which extrapolation is possible.

$\gamma_{\mathbf{c}}$ **:** level of truncation of the clean weighting function.

$\Gamma_{\mathbf{c}}$ **:** γ expressed in dB : $\Gamma_c = -10 \log \gamma_c$.

σ **:** level of noise in the visibilities.

λ **:** observing wavelength.

ξ **:** angular size of the source.

3.8 Sampling required

This appendix is aimed at illustrating the effect of the sampling on the accuracy of Shannon interpolation in Fourier plane. It shows the correctness of the assumption made

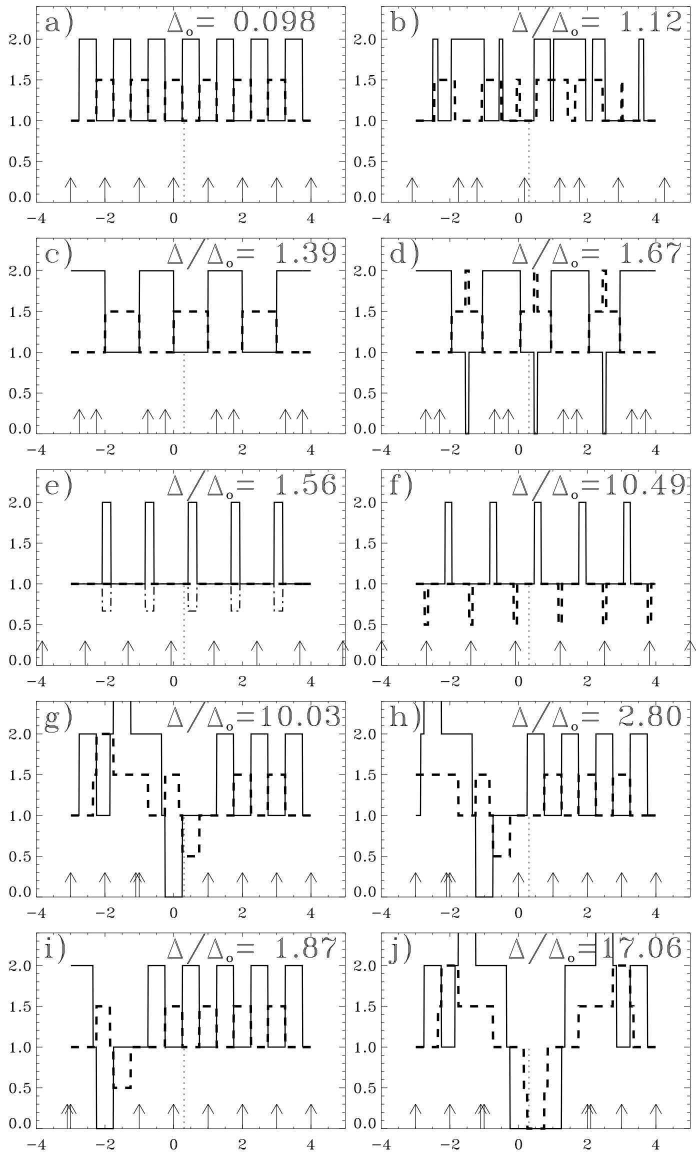

Fig. 3.7 – $d_1(u, 1/\xi)$ (plain line) and $d_2(u, 1/\xi)$ (dashed line) as defined by Eq. (3.41). In Figure e) $d_3(u, 1/\xi)$ is also plotted in dashed-doted line. Δ is the error on the estimate of $\mathcal{V}(u_\mathrm{o})$. The 8 sample positions are symbolized by the arrows. **a)** regular sampling at Nyquist rate : $\boldsymbol{u} = \boldsymbol{r}$. The corresponding error on the estimate of $\mathcal{V}(u_\mathrm{o})$ is $\Delta_\mathrm{o} \simeq \sigma = 0.1$. **b)** $\boldsymbol{u} = \boldsymbol{r} + (-0.1, 0.25, -0.21, 0.18, 0.2, -0.23, -0.1, 0.25)$. **c)** $u_i = r_i + (-1)^i \times 0.25$. **d)** $u_i = r_i + (-1)^i \times 0.3$. **e)** $u_i = (r_i - u_0) \times 1.25 + u_\mathrm{o}$. **f)** $u_i = (r_i - u_0) \times 1.3 + u_\mathrm{o}$. **g)** $\forall i \neq j_\mathrm{o},\ u_i = r_i$ and $u_{j_\mathrm{o}} = r_{j_\mathrm{o}} - 1.1/\xi$. **h)** $\forall i \neq j_\mathrm{o} - 1,\ u_i = r_i$ and $u_{j_\mathrm{o}-1} = r_{j_\mathrm{o}-1} - 1.1/\xi$. **i)** $\forall i \neq j_\mathrm{o} - 2,\ u_i = r_i$ and $u_{j_\mathrm{o}-2} = r_{j_\mathrm{o}-2} - 1.1/\xi$. **j)** $\forall i \neq j_\mathrm{o}$, and $i \neq j_\mathrm{o} + 1$ $u_i = r_i$ and $u_{j_\mathrm{o}} = r_{j_\mathrm{o}} - 1.1/\xi$ and $u_{j_\mathrm{o}+1} = r_{j_\mathrm{o}+1} + 1.1/\xi$.

in Section 3.2, namely, that for Shannon interpolation $\delta u_{\mathrm{M}} = 1/\xi$. For simplicity, one dimension only is considered. Then, if the source size is known to be ξ it can be shown that the Fourier transform of the source brightness distribution can be expanded as a sum of sinc functions :

$$\mathcal{V}(u) = \sum_{i=-\infty}^{+\infty} a_i \frac{\sin\left(\pi(u - i/\xi)\right)}{\pi(u - i/\xi)} \tag{3.33}$$

In practice a limited support only can be considered for $\mathcal{V}$. When the size of the support considered is sufficiently large as compared to $1/\xi$ and if ξ designates the size of the source convolved by the finite PSF (i.e. the leakage effect is compensated for) then, Eq. (3.33) still holds for a finite number of points. The positive part of the symmetric visibility function can be written as :

$$\mathcal{V}(u) = \sum_{i=0}^{N-1} a_i g_i(u)\,, \quad \text{with} \quad g_i(u) = \frac{\sin\left(\pi(u - i/\xi)\right)}{\pi(u - i/\xi)} \tag{3.34}$$

Owing to this equation, if the N parameters a_i are known it is possible to estimate the data at any point in the support. To estimate these parameters N measurements are enough. The aim of this appendix is to consider the impact of the sampling configuration in the vicinity of a point u_{o} on the accuracy of the Shannon interpolation at this same point u_{o}. Actually two independent variables have to be estimated : the real and imaginary part of the visibility function. In the following we will consider only one variable and $\mathcal{V}$ will stand for the real or the imaginary part only.

The sinc functions that are centered on points situated at a distance larger than a given limit from u_{o} can be considered as having a negligible influence on the value of $\mathcal{V}(u_{\mathrm{o}})$. Let n be the number of sinc functions that are close enough to affect the value of $\mathcal{V}(u_{\mathrm{o}})$ and let us assume that the measurements made outside the corresponding domain have a negligible influence on the accuracy of the estimate of $\mathcal{V}(u_{\mathrm{o}})$. We will also assume that the parameters outside the domain are known to be zero. These assumptions simplify the following error computation without changing significantly the results. Let us suppose that m measurements are made to estimate the value of the relevant a_i parameters with $m \geq n$. Each sample has a value y_i measured at a position u_i with an uncertainty σ_i. Then, the data can be modelized by the linear equations :

$$y_i = \sum_{j=j_{\mathrm{o}}-\frac{n}{2}+1}^{j_{\mathrm{o}}+\frac{n}{2}} g_{ij} a_j, \quad \text{with} \quad g_{ij} = \frac{\sin\left(\pi(u_i - j/\xi)\right)}{\pi(u_i - j/\xi)}, \tag{3.35}$$

where $j_{\mathrm{o}} = \mathrm{E}(u_{\mathrm{o}}/\xi)$. From usual linear least squares analysis if the design matrix $\boldsymbol{D}$ is defined by $D_{ij} = f_{ij}/\sigma_i$, and the vector $\boldsymbol{b}$ by $b_i = y_i/\sigma_i$, then, the vector $\boldsymbol{a}$ which components are the a_i parameters is solution of the normal equation :

$$\left(\boldsymbol{D}^T\boldsymbol{D}\right).\,\boldsymbol{a} = \boldsymbol{D}^T.\,\boldsymbol{b} \tag{3.36}$$

The variance associated with the estimate a_i is :

$$\sigma^2(a_i) = C_{ii} \tag{3.37}$$

where $\boldsymbol{C}$ is the variance-covariance matrix defined by :

$$\boldsymbol{C} = \left(\boldsymbol{D}^T\boldsymbol{D}\right)^{-1}. \tag{3.38}$$

The estimate of $\mathcal{V}(u_\text{o})$ is obtained by summing the sinc functions scaled by the a_i parameters. Then, the error on this estimate is :

$$\Delta = \sqrt{\sum_{j=j_\text{o}-\frac{n}{2}+1}^{j_\text{o}+\frac{n}{2}} g_j^2(u_\text{o})\, C_{jj}} \tag{3.39}$$

If the noise σ_i is independent of the signal y_i then Δ is independent of the visibility function and depends only upon the sampling configuration i.e. the u_i's. Let us define Δ_o the error on $\mathcal{V}(u_\text{o})$ when the sampling is regular with a step of $1/\xi$: $\boldsymbol{u} = \boldsymbol{r} = (j_\text{o} - \frac{n}{2}+1, ..., j_\text{o}+\frac{n}{2})/\xi$ (such a sampling is known as Nyquist sampling).

In one dimension the condition for the support $\mathcal{S}$ to be well sampled at δu scale is (Eq. (3.3)) :

$$\begin{array}{l} \forall\, u \in \mathcal{S},\ \forall\, i < L/\delta u, \\ \mathcal{D}\left(u, (i+1/2)\times\delta u\right) \geq \frac{1}{\delta u}\left(\frac{i}{i+1/2}\right) \end{array} \tag{3.40}$$

where L is the length of the support and $\mathcal{D}(u, \delta u)$ is the density of samples in the interval $[u - \delta u/2,\ u + \delta u/2]$. Let us define $d_i(u, \delta u)$ as :

$$d_i(u, \delta u) = \delta u \left(\frac{i+1/2}{i}\right) \mathcal{D}\left(u, (i+1/2)\times\delta u\right). \tag{3.41}$$

Then Eq. (3.40) is equivalent to :

$$\forall\, u \in \mathcal{S},\ \forall\, i < L/\delta u,\ d_i(u, \delta u) \geq 1. \tag{3.42}$$

For Shannon interpolation we assumed that the sampling accuracy required is $\delta u_\text{M} = 1/\xi$. To illustrate the correctness of this assumption d_1 and d_2 are plotted for various sampling configurations and the corresponding error Δ is given in Fig.3.7. For this example we took $n = 8$, $m = 8$, $u_\text{o} = j_\text{o}/\xi + 0.3$ and $\sigma = 0.1$. In Fig.3.7e d_3 is also plotted. In Fig.3.7a the sampling is regular, and we get $\Delta_\text{o} \simeq \sigma$. It is striking that whenever d_1 or d_2 (or d_3 in Fig.3.7e) have some values below 1 the error is increased by more than 50% (see Fig.3.7d, e, f, g, h, i, j). Such errors would affect the whole map and its reliability would not be guarantied any more.