A limit in Field-resolution ratio for interferometric arrays

Laurent Koechlin, José-Philippe Pérez

Observatoire Midi Pyrénées, 14 avenue Edouard Belin, 31400 Toulouse, France

ABSTRACT

This paper presents a theoretical limit to the field/resolution ratio for the imaging mode of aperture synthesis interferometers. This limit is a function of both the number of apertures in the interferometric array and the dynamic range of the visibility and phase data. It does not depend on the optical setup of the instrument. This work is based on the theory of information.

Keywords: Aperture synthesis, Image reconstruction, Information

1. INTRODUCTION

Radio interferometric arrays operating in the visible and IR domains have provided many stellar diameters, orbits and stellar environment parameters. These results have helped create and improve stellar physics and models, but are still limited in terms of images: so far, only simple objects ([1]) have been imaged by visible domain interferometers.

How does their limited number of apertures affect their performance for imaging? In the following section we evaluate the quantity of information available from the raw interferometric data, then estimate the quantity of information necessary to describe a reconstructed image. Finally, we explore what can be deduced by comparing these two quantities.

This work concerns diluted arrays where the individual apertures are considered punctual. A generalization to arbitrarily filled arrays should be published soon.

2. THE RAW DATA

2.1. Information given by a single measure (intuitive approach)

According to Shannon ([2]), the quantity of information provided by one measure is:

$$H = -\Sigma P_s \,\mathrm{lb} P_s \tag{1}$$

H is the Shannon entropy,
P_s is the probability of a given state (or result of a measurement),
lb stands for: log base 2,
Σ stands for summation over all the Ω possible states (values) a measure can take.

If the different values that the result of a given measure can take all have the same probability (the P_s are all equal), the information provided by this measure is simply:

$$H = \mathrm{lb}\, \Omega_{data}. \tag{2}$$

In this case, Ω is the dynamic range of the datum:

$$\Omega_{data} = \frac{maximum\, signal}{noise}. \tag{3}$$

H is expressed in bits and corresponds to the minimum number of binary bits required to store the data. The "maximum signal" is set by the range of the device used for the measurement.

Further author information: (Send correspondence to L.K.)
L.K.: E-mail: firstname.lastname@ast.obs-mip.fr, Telephone: (33) 561 33 28 87
Address: Observatoire Midi Pyrénées, 14 avenue Edouard Belin, 31400 Toulouse, France

2.2. Information given by a single measure (Noise entropy approach)

One can derive an expression similar to equation (3) in a more theoretical way. The noise in a complex visibility measurement can be expressed by: $\underline{b} = b_r + ib_i$. According to Frieden ([3]), if the noise is gaussian with a zero mean value, the noise entropy $H(N)$ is given by the following continuous expression:

$$H(N) = -\int p_b(\underline{b})\,\mathrm{lb}(p_b)\,\mathrm{d}\underline{b} \quad \text{with} \quad p_b = \frac{1}{(2\pi\sigma_{b,r})^{1/2}}\exp\left(-\frac{b_r^2}{2\sigma_{b,r}^2}\right)\frac{1}{(2\pi\sigma_{b,i})^{1/2}}\exp\left(-\frac{b_i^2}{2\sigma_{b,i}^2}\right). \quad (4)$$

As in the visibility signals, the real and imaginary parts are equally distributed; the quantities $\sigma_{b,r}$ and $\sigma_{b,i}$, the variances of the two Gauss distributions, are the same. So:

$$\sigma_b^2 = \sigma_{b,r}^2 + \sigma_{b,i}^2 = 2\sigma_{b,r}^2 = 2\sigma_{b,i}^2 \quad \text{and} \quad p_b = \frac{1}{\pi\sigma_b}\exp\left(-\frac{|\underline{b}|^2}{2\sigma_b^2}\right).$$

It follows:

$$H(N) = -\iint p_b(\underline{b})\,\mathrm{lb}(p_b)\,\mathrm{d}b_r\,\mathrm{d}b_i = -\int_0^\infty \frac{1}{\pi\sigma_b \mathrm{lb}(2)}\exp\left(-\frac{|\underline{b}|^2}{2\sigma_b^2}\right)\left[\left(-\frac{|\underline{b}|^2}{2\sigma_b^2}\right) - \ln(\pi\sigma_b)\right] b\,\mathrm{d}b \int_0^{2\pi}\mathrm{d}\theta, \quad (5)$$

or:

$$\frac{2}{\sigma_b\,\mathrm{lb}(2)}\int_0^\infty \exp\left(-\frac{|\underline{b}|^2}{2\sigma_b^2}\right)\left[\left(\frac{|\underline{b}|^2}{2\sigma_b^2}\right) + \ln(\pi\sigma_b)\right]\underline{b}\,\mathrm{d}\underline{b}\ ,$$

Finally, after integration:

$$H(N) = \frac{1}{\ln 2}[1 + \ln(\pi\sigma_n)]\,. \quad (6)$$

It is interesting to note that this noise entropy is entirely determined by the noise variance σ_n^2.

2.2.1. Conditional entropy

With an additive noise N, the conditional entropy $H(D|S)$ is equal to noise entropy $H(N)$, because the probability density $p(D|S)$, when the signal is fixed, is the same as the noise probability density $p_N(N|S)$. So

$$p(D|S) = p_N(N|S) = p_N(N)$$

Thus:

$$H(N|S) = \frac{1}{\ln 2}[1 + \ln(\pi\sigma_n)]. \quad (7)$$

2.2.2. Maximum Data entropy and capacity of the interferometer

The capacity of the channel is given by the maximum of the mutual information $I(S, D)$:

$$I(S, D) = H(D) - H(D|S) = H(D) - H(N) \quad \text{and} \quad C = Max\{H(D) - H(N)\}.$$

The maximal value of $H(D)$ is obtained for a Gaussian distribution. Consequently, the corresponding probability of D can be derived the same way as the noise is derived in equation (4):

$$p_D(d) = \frac{1}{\pi\sigma_d}\exp\left[-\frac{(|d| - d_m)^2}{2\sigma_d^2}\right] \quad \text{and} \quad Max\{H(D)\} = \frac{1}{\ln 2}[1 + \ln(\pi\sigma_d)]. \quad (8)$$

where σ_d^2 is the variance of the data d, and d_m the mean value. We deduce the capacity expression:

$$C = \frac{1}{\ln 2}[1 + \ln(\pi\sigma_d)] - \frac{1}{\ln 2}[1 + \ln(\pi\sigma_n)], \quad (9)$$

which simplifies to:

$$C = \mathrm{lb}\left(\frac{\sigma_d}{\sigma_n}\right). \quad (10)$$

If one compares σ_d and σ_n to the amplitudes of the signal and the noise, respectively, the ratio (σ_d/σ_n) corresponds to the dynamic range of the data (Ω_{data}), and equation (10) corresponds to equation (3).

3. APPLICATION TO INTERFEROMETRY

With interferometric arrays operating from the ground, the beams are usually not cophased. The raw data are the fringe visibility moduli and phase closures. An interferometric array operating in this "coherenced" mode with N_{tel} telescope apertures can collect $N_{tel}(N_{tel}-1)/2$ visibility moduli, plus $(N_{tel}-1)(N_{tel}-2)/2$ phase closures, that is:

$$(N_{tel}-1)^2 \quad \text{independent measurements.}$$

If one can stabilize the path lengths to better than a fraction of wavelength, for example in space, the raw data are the modulus and phase for each covered frequency. This is the "cophased" observing mode: $N_{tel}(N_{tel}-1)/2$ visibilities and the same quantity of phases, that is:

$$N_{tel}(N_{tel}-1) \quad \text{independent measurements.}$$

This difference between cophased and coherenced observing modes is significant for small arrays.

Beyond the technological challenge of controlling wave delays at wavelengths a thousand times shorter than in the radio domain, there are other differences that make imaging at visible wavelengths more difficult.

For long wavelengths (radio), the complex amplitude can be directly recorded, and a given beam can be correlated numerically with any number of other beams, without any loss of information.

At short wavelengths (optical and IR), the number of photons per wave from natural sources is too small to allow recording of the amplitude: only the intensity of the wave can be measured. Thus, the beams have to be combined optically, and the dynamic range of a given fringe signal is reduced when combining more two beams. Working at short wavelengths domains provides less raw data than in the radio domain, for an equivalent number of apertures.

3.1. Information transmitted by a radio domain observation

The information from independent sources adds up ([2]). Assuming an equal dynamic range for both the modulus and the closure phase data, a snapshot observation in the long wavelength mode the following quantity of information:

$$H_{snap} = N_{tel}(N_{tel}-1)\mathrm{lb}\,\Omega_{data}. \tag{11}$$

In order to improve the frequency (UV) coverage, a series of N_{obs} snapshot observations can be made with different baseline configurations. This is achieved by moving the telescopes to different positions, or by observing at different times, as the projected baselines change due to diurnal rotation. In that case:

$$H_{set} = N_{obs}N_{tel}(N_{tel}-1)\mathrm{lb}\,\Omega_{data}. \tag{12}$$

One should note that even in the case of no baseline modification between snapshots, the quantity of information brought up by a series of observations is still given by equation (14). The quality of the UV coverage will play a role on how much information is lost at the image reconstruction stage.

3.2. Information transmitted by a short wavelength domain observation

The beams have to be combined optically, the dynamic range of a fringe measurement, and/or the number of actual baselines is reduced when N_{tel} is grater than 2, hence the inequality in the two following expressions.

Assuming an equal dynamic range for both the modulus and the closure phase data, a snapshot observation in the visible or IR provides the following quantity of information:

$$H_{snap} \leq (N_{tel}-1)^2\mathrm{lb}\,\Omega_{data}. \tag{13}$$

and if N_{obs} snapshot observations are combined,

$$H_{set} \leq N_{obs}(N_{tel}-1)^2\mathrm{lb}\,\Omega_{data}. \tag{14}$$

4. EFFECT OF THE RECONSTRUCTION PROCESS

The image reconstruction process, would it be numerical such as Clean ([4]), Maximum Entropy ([5]), and Wipe ([6]), or optical such as in a hypertelescope ([7]), ([8]), may discard, but never add information by itself. Hence, after reconstruction, the quantity of information in the image is limited by the following inequality (for visible and IR arrays):

$$H_{image} \leq N_{obs}(N_{tel}-1)^2 \mathrm{lb}\,\Omega_{data}. \tag{15}$$

If a model of the object is taken into account during the reconstruction process, this *a priori* information contributes to the total, and we have:

$$H_{image} \leq N_{obs}(N_{tel}-1)^2 \mathrm{lb}\,\Omega_{data} + H_{model}. \tag{16}$$

Such an *a priori* information can be the knowledge of the maximum support of the object: all the resolution elements (resels) set to zero outside a given region, as for the reconstruction of a multiple star image in which individual components are known to be unresolved to the chosen resolution.

5. INFORMATION CONTAINED IN AN IMAGE

Let's now look at the other side of the problem and estimate the information contained in (*i.e.* necessary to describe) an arbitrary image. First, we need to find the number of independent samples provided by this image: it corresponds to the number of resels in the field covered by this image:

$$N_{resels} = \left(\frac{field}{resolution}\right)^2 \tag{17}$$

Assuming all the resels in the image have the same maximum possible intensity I_{max} and the same resolution in intensity ΔI, the dynamic range Ω_{resel} of each resel is:

$$\Omega_{resel} = \frac{I_{max}}{\Delta I}\,, \tag{18}$$

and the corresponding information can be expressed

$$H_{resel} = \mathrm{lb}\,\Omega_{resel}. \tag{19}$$

The total information required to describe this image is then

$$H_{image} = \mathrm{lb}\,\Omega_{resel}\,(field/resolution)^2\,. \tag{20}$$

6. CONSTRAINTS ON THE FIELD-RESOLUTION RATIO

Combining the expressions from the previous sections: the information available from the data after image reconstruction in eq. (20), and the quantity of information implied by the size and dynamic range of this image in eq.(15), one can derive the following inequality:

$$\left(\frac{field}{resolution}\right)^2 \leq N_{obs}(N_{tel}-1)^2 \frac{\mathrm{lb}\Omega_{data}}{\mathrm{lb}\,\Omega_{resel}}\,. \tag{21}$$

For example, a reconstruction process adjusted to yield a dynamic range in the reconstructed image, same as the dynamic range in the raw visibility data: $\Omega_{resel} = \Omega_{data}$, would have the following field limit:

$$field \leq resolution \times (N_{tel}-1)\sqrt{N_{obs}}\,. \tag{22}$$

6.1. Examples of field/resolution ratios

1. At the VLTI with 4 coherent apertures (coherenced but not cophased) and in snapshot mode, the field / resolution ratio is equal to 3 (9 resels). In the case of a 2 milli arc seconds (mas) angular resolution, attainable with 200 m baseline at $\lambda = 2.2\mu$m,the field would be limited to 6 mas.

2. In a more complete VLTI configuration, with 8 cophased apertures and 25 reconfigurations (supersynthesis), the field / resolution ratio reaches 35 (1 200 resels). With a 2 mas resolution, the field would be limited to 70 mas.

Using a lower resolution and shorter baselines at the same wavelength, one could increase the field up to the order of 300 mas, but still with not more than 1200 resels total.

From eq.(22), one can derive this rule of thumb: "in snapshot mode, the limit in field/resolution ratio roughly corresponds to the number of individual apertures in the array."

7. CONCLUSION

The inequality presented above does not depend on the type of reconstruction algorithms or optics used, nor does it depend on the type of object observed. It only gives a limit, which may or may not be attained, depending on the quality of the reconstruction process and the UV coverage.
It leads to very narrow fields in the case of high-resolution arrays featuring a small number of apertures. This field limitation can be improved by:

1. Increasing the number of apertures (direct effect),

2. Increasing the number of baseline reconfigurations during a numerical reconstruction process (square root effect),

3. Accepting a dynamic range in the reconstructed image, which is lower than the dynamic range in the raw data (driven by the signal/noise ratio).
   This situation can be expressed the other way around: asking for too large a field at the image reconstruction stage will lead to an insufficient dynamic range in the reconstructed image. The resulting amplification of errors from the initial data will cause wreckage in the reconstructed image. *A contrario*, choosing a lower field/resolution at the reconstruction stage will improve the dynamic range.

4. Observing objects which can be described by a small number of parameters. The required quantity of information will then be expressed by the dynamic range and number of these independent parameters. This is equivalent to adding model information into the reconstruction process.

REFERENCES

1. J. E. Baldwin, M. G. Beckett, R. C. Boysen, D. Burns , D. F. Buscher, G. C. Cox, C. A. Haniff, C. D. Mackay, N. S. Nightingale, J. Rogers, P. A. G. Scheuer, T. R. Scott, P. G. Tuthill, P. J. Warner, D. M. A. Wilson, R. W. Wilson A&A, **306**, L13-L16, 1996.
2. C.E. Shannon, 1948, Bell System Technical Journal, **27**,379-423, 623-656
3. B.R. Frieden, *Image Enhancement and Restoration*, Berlin, Springer-Verlag, 1978.
4. W. N. Christiansen, J. A. Högbom, *Radiotelescopes, 2nd ed.* Cambridge, England, Cambridge University Press, pp. 214-216, 1985.
5. T.J. Cornwell, K.F. Evans, 1985. A&A **143**, 77-83.
6. A. Lannes, 1989, Experimental Astronomy, **1**, 47-76.
7. A. Labeyrie 1996, A&A sup., **118**, 517-524.
8. A. Boccalett, P. Riaud, C. Moutou, A. Labeyrie, Icarus, **146**, 628.