Revisiting the theory of interferometric wide-field synthesis

J. Pety1,2 and N. Rodríguez-Fernández1

1 IRAM, 300 rue de la Piscine, 38406 Grenoble Cedex, France
e-mail: [pety,rodriguez@iram.fr]
2 LERMA, UMR 8112, CNRS and Observatoire de Paris, 61 avenue de l’Observatoire, 75014 Paris, France

Received 13 July 2009 / Accepted 16 March 2010

ABSTRACT

Context. After several generations of interferometers in radioastronomy, wide-field imaging at high angular resolution is today a major goal for trying to match optical wide-field performances.

Aims. All the radio-interferometric, wide-field imaging methods currently belong to the mosaicking family. Based on a 30 years old, original idea from Ekers & Rots, we aim at proposing an alternate formalism.

Methods. Starting from their ideal case, we successively evaluate the impact of the standard ingredients of interferometric imaging, i.e. the sampling function, the visibility gridding, the data weighting, and the processing of the short spacings either from single-dish antennas or from heterogeneous arrays. After a comparison with standard nonlinear mosaicking, we assess the compatibility of the proposed processing with 1) a method of dealing with the effect of celestial projection and 2) the elongation of the primary beam along the scanning direction when using the on-the-fly observing mode.

Results. The dirty image resulting from the proposed scheme can be expressed as a convolution of the sky brightness distribution with a set of wide-field dirty beams varying with the sky coordinates. The wide-field dirty beams are locally shift-invariant as they do not depend strongly on position on the sky: their shapes vary on angular scales typically larger or equal to the primary beamwidth. A comparison with standard nonlinear mosaicking shows that both processing schemes are not mathematically equivalent, though they both recover the sky brightness. In particular, the weighting scheme is very different in both methods. Moreover, the proposed scheme naturally processes the short spacings from both single-dish antennas and heterogeneous arrays. Finally, the sky gridding of the measured visibilities, required by the proposed scheme, may potentially save large amounts of hard-disk space and cpu processing power over mosaicking when handling data sets acquired with the on-the-fly observing mode.

Conclusions. We propose to call this promising family of imaging methods wide-field synthesis because it explicitly synthesizes visibilities at a much finer spatial frequency resolution than the one set by the diameter of the interferometer antennas.

Key words. methods: analytical – techniques: interferometric – methods: data analysis – techniques: image processing

1. Introduction

The instantaneous field of view of an interferometer is naturally limited by the primary beam size of the individual antennas. For the ALMA 12 m-antennas, this field of view is ~9′ at 690 GHz and ~27′′ at 230 GHz. The astrophysical sources in the (sub-)millimeter domain are often much larger than this, but still structured on much smaller angular scales. Interferometric wide-field techniques enable us to fully image these sources at high angular resolution. These techniques first require an observing mode that in one way or another scans the sky on spatial scales larger than the primary beam. The most common observing mode in use today, known as stop-and-go mosaicking, consists in repeatedly observing sky positions typically separated by half the primary beam size. The improvement of the tracking behavior of modern antennas now leads astronomers to consider on-the-fly observations, with the antennas slewing continuously across the sky. The improvements in correlator and receiver technologies are also leading to techniques that could potentially sample the antenna focal planes with multi-beam receivers instead of the single-pixel receivers installed on current interferometers.

The ideal measurement equation of interferometric wide-field imaging is

\[
V(u_p, \alpha_s) = \int_{\alpha_p} B(\alpha_p - \alpha_s) \ I(\alpha_p) \ e^{-2\pi i u_p \alpha_p} \ d\alpha_p,
\]

(1)

where \(V\) is the visibility function of 1) \(u_p\) (the spatial frequency with respect to the fixed phase center) and 2) \(\alpha_s\) (the scanned sky angle), \(I\) is the sky brightness, and \(B\) the antenna power pattern of \(\alpha_p\). We propose to call this promising family of imaging methods wide-field synthesis because it explicitly synthesizes visibilities at a much finer spatial frequency resolution than the one set by the diameter of the interferometer antennas.

1 IRAM, 300 rue de la Piscine, 38406 Grenoble Cedex, France
e-mail: [pety,rodriguez@iram.fr]
2 LERMA, UMR 8112, CNRS and Observatoire de Paris, 61 avenue de l’Observatoire, 75014 Paris, France

Accepted 16 March 2010

Article published by EDP Sciences
For each $u_p$ kept constant, $V(u_p, \alpha_s)$ is the convolution of $B$ and $I$. Indeed, $I(\alpha_p, u_p = 0) = I(\alpha_p)$, so we derive
\[
V(u_p = 0, \alpha_s) = \int_{\alpha_p} B(\alpha_p - \alpha_s) I(\alpha_p) \mathrm{d} \alpha_p,
\]
\[(5)\]
i.e., the convolution equation for single-dish observations.

Ekers & Rots (1979) were the first to show that the measurement equation (Eq. (1)) enables us to recover spatial frequencies of the sky brightness at a much finer $w$ resolution than the $uw$ resolution set by the diameter of the interferometer antennas. Interestingly enough, the goal of Ekers & Rots (1979) was “just” to find a way to produce the missing short spacings of a multiplying interferometer. However, Cornwell (1988) realized that Ekers & Rots’ scheme has a much stronger impact, because it explains why an interferometer is able to do wide-field imaging. Cornwell (1988) also demonstrated that on-the-fly scanning is not absolutely necessary to interferometric wide-field imaging. Indeed, the large-scale information can be retrieved in mosaics of single-field observations, provided that the sampling of the single fields follows the sky-plane Nyquist sampling theorem.

As a result, all the information about the sky brightness is coded in the visibility function. From a data-processing viewpoint, all the current radio-interferometric wide-field imaging methods (see, e.g., Gueth et al. 1995; Sault et al. 1996b; Cornwell et al. 1993; Bhatnagar & Cornwell 2004; Bhatnagar et al. 2008; Cotton & Uson 2008) belong to the mosaicking family\(^1\) pioneered by Cornwell (1988). In this family, the processing starts with Fourier transforming $V(u_p, \alpha_s)$ along the $u_p$ dimension (i.e. at constant $\alpha_s$) to produce a set of single-field dirty images before linearly combining them and forming a wide-field dirty image. In this paper, we propose an alternate processing, which starts with a Fourier transform of $V(u_p, \alpha_s)$ along the $\alpha_s$ dimension (i.e. at constant $u_p$). We show how this explicitly synthesizes the spatial frequencies needed to do wide-field imaging, which are linearly combined to form a “wide-field $w$ plane”, i.e., one $uw$-plane containing all the spatial frequency information measured during the wide-field observation. Inverse Fourier transform will produce a dirty image, which can then be deconvolved using standard methods. The existence of two different ways to extract the wide-field information from the visibility function raises several questions: are they equivalent? What are their relative merits?

We thus aim at revisiting the mathematical foundations of wide-field imaging and deconvolution. Sections 2 to 7 propose the new algorithm, which we call wide-field synthesis: Sect. 2 first defines the notations and it then lays out the basic concepts used throughout the paper. Section 3 states the steps needed to go beyond the Ekers & Rots scheme and explores the consequences of incomplete sampling of both the $w$ and sky planes. Section 4 discusses the effects of gridding by convolution and regular resampling. Section 5 describes how to influence the dirty beam shapes and thus the deconvolution. Section 6 states how to introduce short spacings measured either from a single-dish antenna or from heterogeneous interferometers. Section 7 compares the proposed wide-field synthesis algorithm with standard nonlinear mosaicking. Some detailed demonstrations are factored out in Appendix A to enable an easier reading of the main paper, while ensuring that interested readers can follow the demonstrations. Appendices B and C then explain how the wide-field synthesis algorithm can cope with non-ideal effects: Appendix B discusses how at least one standard way to cope with sky projection problems is compatible with the wide-field synthesis algorithm. Appendix C explores the consequences of using the on-the-fly observing mode. Finally, we assume good familiarity with single-field imaging in various places. We refer the reader to well-known references: e.g. Chap. 6 of Thompson et al. (1986) and Sramek & Schwab (1989).

2. Notations and basic concepts

2.1. Notations

In this paper, we use the Bracewell (2000)’s notation to display the relationship between a function $I(\alpha)$ and its direct Fourier transform $\mathcal{T}(u)$, i.e.,
\[
I(\alpha) \ni \mathcal{T}(u),
\]
\[(6)\]
where $(\alpha, u)$ is the couple of Fourier conjugate variables. We also use the following sign conventions for the direct and inverse Fourier transforms
\[
\mathcal{T}(u) \equiv \int_{\alpha} I(\alpha) e^{-i2\pi su} d\alpha
\]
\[(7)\]
and
\[
I(\alpha) \ni \int_{u} \mathcal{T}(u) e^{i2\pi su} du.
\]
\[(8)\]
As $V$ is a function of two independent quantities ($u_p$ and $\alpha_s$), the Fourier transform may be applied independently on each dimension, while the other dimension stays constant. Several additional conventions are used to express this. First, we introduce a specific notation to state that either the first or the second dimension stays constant:
\[
V_{u_p}(\alpha_s) \equiv V(u_p = \text{const.}, \alpha_s),
\]
\[(9)\]
\[
V_{\alpha_s}(u_p) \equiv V(u_p, \alpha_s = \text{const.}).
\]
\[(10)\]
Second, we use a bottom/top line to derive the notation of the Fourier transform along the first/second dimension from the notation of the original function. Third, on the Fourier transform sign (i.e. $\ni$), we explicitly state the dimension along which the Fourier transform is computed. For instance, if $D$ is a function of $(\alpha_p, \alpha_s)$, then the Fourier transform of $D$ along the first dimension is expressed as
\[
D(\alpha_p, \alpha_s) \ni_{\alpha_p} \mathcal{D}(u_p, \alpha_s),
\]
\[(11)\]
while the Fourier transform of $D$ along the second dimension is expressed as
\[
D(\alpha_p, \alpha_s) \ni_{\alpha_s} \mathcal{D}(\alpha_p, u_s).
\]
\[(12)\]
We also use a more compact notation when doing the Fourier transform on both dimensions simultaneously, i.e.,
\[
D(\alpha_p, \alpha_s) \ni_{(\alpha_p, \alpha_s)} \mathcal{D}(u_p, u_s).
\]
\[(13)\]
Finally, the convolution of two functions $G$ and $V$ is noted and defined as
\[
\{G \star V\}(u) \equiv \int_{v} G(u - v) V(v) \mathrm{d}v.
\]
\[(14)\]

---

\(^1\) In the rest of this paper, stop-and-go mosaicking refer to the observing mode, while mosaicking alone refer to the imaging family.
2.2. Basic concepts

Figure 2 illustrates the principles underlying 1) the setup to get interferometric wide-field observations and 2) our proposition to process them. For simplicity, we display the minimum possible complexity without loss of generality. The top row displays the sky plane. The middle row represents the 4-dimensional measurement space at different stages of the processing. As it is difficult to display a 4-dimensional space on a sheet of paper, the bottom row shows 2-dimensional cuts of the measurement space at the same processing stages.

2.2.2. Processing by explicit synthesis of the wide-field spatial frequencies

All the information about the sky brightness, \( I(\alpha) \), is somehow coded in the visibility function, \( V(u_p, \alpha) \). The high spatial
Fig. 2. Illustration of the principles of wide-field synthesis, which enables us to image wide-field interferometric observations. The top row displays the sky plane. The middle row displays the 4-dimensional visibility space and the bottom row shows 2-dimensional cuts of this space at different stages of the processing. In panels b) to d), the scanned dimensions ($\alpha_s$ and $u_s$) are displayed in blue while the phased spatial scale dimensions ($u_p$) are displayed in red and the spatial scale dimensions ($u$) of the final wide-field $uv$ plane are displayed in black. The grey zones of panels b.2) and c.2) show the regions of the visibility space without measurements (missing short-spacings). In detail, panel a) shows a possible scanning strategy of the sky to measure the unknown brightness distribution at high angular resolution: for simplicity it is here just a 7-field mosaic. Panels b.1) and b.2) sketch the space of measured visibilities: the $uv$ plane at each of the 7 measured sky positions is displayed as a blue square box in panel b.1) and a blue vertical line in panel b.2). For simplicity, only 6 visibilities are plotted in panel b.1). Panels c.1) and c.2) sketch the space of synthesized visibilities after Fourier transform of the measured visibilities along the scanned coordinate ($\alpha_s$): at each measured spatial frequency $u_p$ (displayed on the blue axes) is associated one space of synthesized wide-field spatial frequencies displayed as one of the red squares in panel c.1) and the red vertical lines in panel c.2). The wide-field spatial scales are synthesized 1) on a grid whose cell size is related to the total field of view of the observation and 2) only inside circles whose radius is the primary diameter of the interferometer antennas. Panels d.1) and d.2) display the final, wide-field $uv$ plane. This plane is built by application of the shift-and-average operator along the black lines on panel c.2), lines that display the region of constant $u$ spatial frequency in the ($u_p$, $u_s$) space. Standard inverse Fourier transform and deconvolution methods then produce a wide-field distribution of sky brightnesses as shown in panel e).
frequencies (from \(d_{\text{min}}\) to \(d_{\text{max}}\)) are clearly coded along the \(u_p\) dimension. The uncertainty relation between Fourier conjugate quantities also implies that the typical spatial frequency resolution along the \(u_p\) dimension is only \(d_{\text{prim}}\) because the field of view of a single pointing has a typical size of \(\theta_{\text{prim}}\). However, wide-field imaging implies measuring all the spatial frequencies with a finer resolution, \(d_{\text{field}} = 1/\theta_{\text{field}}\). The missing information must then be hidden in the \(\alpha_s\) dimension.

In Sect. 3, we show that Fourier transforming the measured visibilities along the \(\alpha_s\) dimension (i.e. at constant \(u_p\)) can synthesize the missing spatial frequencies, because the \(\alpha_s\) dimension is sampled from \(-\theta_{\text{field}}/2\) to \(+\theta_{\text{field}}/2\), implying a typical spatial-frequency resolution of the \(u_s\) dimension equal to \(d_{\text{field}}\). Conversely, the \(\alpha_s\) dimension is probed by the primary beams with a typical angular resolution of \(\theta_{\text{prim}}\), implying that the \(u_s\) spatial frequencies will only be synthesized inside the \([-d_{\text{prim}}/2, +d_{\text{prim}}/2]\) range. Panels c.1) and c.2) illustrate the effects of the Fourier transform of \(V(u_p, u_s)\) along the \(\alpha_s\) dimension, in 4 and 2 dimensions, respectively. The red subpanels or vertical lines display the \(u_s\) spatial frequencies around each constant \(u_p\) spatial frequency.

In panels d.1) and d.2) (i.e. after the Fourier transform along the \(\alpha_s\) dimension), \(\overline{V}(u_p, u_s)\) contains all the measured information about the sky brightness in a spatial frequency space. However, the information is ordered in a strange and redundant way. Indeed, we show that \(\overline{V}(u_p, u_s)\) is linearly related to \(\overline{T}(u_p + u_s)\). To first order, the information about a given spatial frequency \(u\) is stored in all the values of \(\overline{V}(u_p, u_s)\) which verifies \(u = u_p + u_s\) (black lines on panel c.2).

A shift operation will reorder the spatial scale information and averaging will compress the redundancy (illustrated by the halving of the number of the space dimensions). The use of a shift-and-average operator thus produces a final \(w\) plane containing all the spatial scale information to image a wide field in an intuitive form. We thus call this space the wide-field \(w\) plane. Panels d.1) and d.2) display this space, where the minimum relevant spatial frequency is related to the total field of view, while the maximum one is related to the interferometer resolution.

Sections 3 and 4 show that applying the shift-and-average operator to \(\overline{V}\) produces the Fourier transform of a dirty image, which is a local convolution of the sky brightness by a slowly varying dirty beam. As a result, inverse Fourier transform of \(\overline{V}\) and deconvolution methods will produce a wide-field distribution of sky brightness as shown in panel e) at the top right of Fig. 2.

3. Beyond the Ekers & Rots scheme

In the real world, the visibility function is not only sampled, but this sampling is incomplete for two main reasons. 1) The instrument has a finite spatial resolution, and the scanning of the sky is limited, implying that the sampling in both planes has a finite support. 2) The \(w\) coverage and the sky-scanning coverage can have holes caused either by intrinsic limitations (e.g. lack of short spacings or small number of baselines) or by acquisition problems (implying data flagging). The incomplete sampling makes the mathematics on the general case complex. We thus start with the ideal case where we assume that the visibility function is continuously sampled along the \(u_p\) and \(\alpha_s\) dimension. We then look at the general case.

3.1. Ideal case: infinite, continuous sampling

Starting from the measurement Eq. (1), Ekers & Rots (1979) first demonstrated (see Sect. A.1) that

\[
\forall (u_p, u_s), \quad \overline{V}_{u_p}(u_s) = \overline{B}(-u_s) \overline{T}(u_p + u_s). \tag{15}
\]

For each constant \(u_p\) spatial frequency, the Fourier transform thus synthesizes a function, \(\overline{V}_{u_p}(u_s)\), which is simply related to \(\overline{T}(u_p + u_s)\), the Fourier components of the sky brightness around \(u_p\). \(\overline{V}_{u_p}(u_s)\) is only defined in the \([-d_{\text{prim}}/2, +d_{\text{prim}}/2]\) interval along the \(u_s\) dimension because \(\overline{B}(-u_s)\) is itself only defined inside this interval, since \(\overline{B}(-u_s)\) is the autocorrelation of the antenna illumination.

We search to derive a single estimate of the Fourier components \(\overline{T}(u)\) of the sky brightness. Equation (15) indicates that the fraction \(\overline{V}(u_p, u_s)/\overline{B}(-u_s)\) gives us an estimate of \(\overline{T}(u)\) for each couple \((u_p, u_s)\) that satisfies \(u = u_p + u_s\). However, the information about \(\overline{T}\) is strangely ordered. There are two possible ways to look at this ordering. 1) Starting from the measurement space, the Ekers & Rots scheme synthesizes frequencies around each \(u_p\) measure inside the interval \([u_p - d_{\text{prim}}, u_p + d_{\text{prim}}]\) at the \(d_{\text{field}}\) spatial frequency resolution. 2) Starting from our goal, we want to estimate \(\overline{T}\) at a given spatial frequency \(u\) with a \(d_{\text{field}}\) spatial frequency resolution. We thus search for all the couples \((u_p, u_s)\) satisfying \(u = u_p + u_s\), which are displayed in panel c.2) of Fig. 2 as the diagonal black lines. It immediately results that 1) there are several estimates of \(\overline{T}\) for each spatial frequency \(u\) and 2) the number of estimates varies with \(u\). We can average them to get a better estimate of \(\overline{T}(u)\).

This last viewpoint thus suggests averaging in the \((u_p, u_s)\) space along linepaths defined by \(u = u_p + u_s\). Such an operator can mathematically be defined as

\[
\langle F(u) \rangle \equiv \int_{u_p + u_s} \delta[u - (u_p + u_s)] W(u_p, u_s) F(u_p, u_s) \, du_p du_s, \tag{16}
\]

where \(F\) is the function to be averaged and \(W\) is a normalized weighting function. Using the properties of the Dirac function, we can reduce the double integral to

\[
\langle F(u) \rangle = \int_{u_p} W(u_p, u - u_p) F(u_p, u - u_p) \, du_p. \tag{17}
\]

In this equation, we easily recognize a shift-and-average operator. The normalized weighting function plays a critical role in the following formalism, and we propose clever ways to define \(W\) in Sect. 5. In the ideal case studied here, \(W\) can be defined as

\[W(u_p, u_s) \equiv 1/(2 \sqrt{d_{\text{prim}}}^2)\quad \text{for } u_s \in [-d_{\text{prim}}, +d_{\text{prim}}],\]

\[W(u_p, u_s) \equiv 0 \quad \text{for other values of } u_s.\]

In other words, we have just normalized the integral by the constant length \((2 \sqrt{d_{\text{prim}}}^2)\) of the averaging linepath.
3.1.1. Wide-field dirty image, dirty beam and image-plane measurement equation

Section 3.2 shows that the incomplete sky and $uv$ sampling forbids us to apply the shift-and-average operator to the $\overline{\nabla}(u_p, v_p) \overline{B}(-u_p)$ function. To guide us in this general case, we thus explore the consequences of applying this operator to $\overline{\nabla}$ in the ideal case. It is easy to demonstrate that the result is the Fourier transform of a dirty image, i.e.,

$$\overline{I_{\text{dirty}}}(u) = \langle \overline{\nabla}(u) \rangle .$$

(18)

Indeed, substituting $\overline{\nabla}(u)$ with the help of Eqs. (17) and (15) and taking the inverse Fourier transform, we get

$$\overline{I_{\text{dirty}}}(u) = D(u) I(u),$$

(19)

with

$$D(u) \equiv \int_{u_{p}} W(u_p, u - u_p) \overline{B}(u_p - u) \, du_p .$$

(20)

Here, $I_{\text{dirty}}$ conforms to the usual idea of dirty image, i.e., the convolution of a dirty beam by the sky brightness:

$$I_{\text{dirty}}(\alpha) = (D \ast I)(\alpha) .$$

(21)

In contrast to the usual situation for single-field observations, the mix between a Fourier transform and a convolution of Eq. (1), associated with the specific processing$^3$ changes the image-plane measurement equation from a convolution of a dirty beam with the product $BI$ to a convolution of a dirty beam with $I$. The dependency on the primary beam is still there. It is just transferred from a product of the sky brightness distribution into the definition of the dirty beam.

3.1.2. Summary and interpretation

In summary, a theoretical implementation of wide-field synthesis implies

1. the possibility of Fourier transforming the visibility function along the $\alpha_s$ dimension (i.e. at constant $u_p$), which gives us a set of synthesized $uv$ planes;
2. the possibility of shifting-and-averaging these synthesized $uv$ planes to build the final, wide-field $uv$ plane containing all the available information.

Using those tools, we are able to write the wide-field image-plane measurement equation as a convolution of a wide-field dirty beam ($D$) by the sky brightness ($I$), i.e.,

$$I_{\text{dirty}}(\alpha) = \int_{\alpha'} D(\alpha - \alpha') I(\alpha') \, d\alpha' .$$

(22)

We can write a convolution equation in this ideal case because the wide-field response of the instrument is shift-invariant; i.e., $D$ only depends on differences of the sky coordinates.

It is well-known that for a single-field observation, the dirty beam is the inverse Fourier transform of the sampling function. The shape of this sampling function is due to the combination of aperture synthesis (the interferometer antennas give a limited number of independent baselines) and Earth-rotation synthesis (the rotation of the Earth changes the projection of the physical baselines onto the plane perpendicular to the instantaneous line of sight). By analyzing via a Fourier transform, the evolution of the visibility function with the sky position, the Ekers & Rots scheme synthesizes visibilities at spatial frequencies needed to image a larger field of view than the interferometer primary beam. We thus propose to call this specific processing: wide-field synthesis.

3.2. General case: incomplete sampling

Reality imposes limitations on the synthesis of spatial frequencies. Indeed, we have already stated that the visibility function is incompletely sampled both in the $uv$ and sky planes. To take the sampling effects into account, we introduce the sampling function $S(u_p, \alpha_s)$, which is a sum of Dirac functions at measured positions$^4$. The sampling function cannot be factored into the product of two functions, each only acting on one plane. Indeed, the Earth rotation happening during the source scanning implies a coupling of both dimensions of the sampling function. In other words, the $uv$ coverage will vary with the scanned sky coordinate. This leads us to a shift-dependent situation, precluding us from writing the wide-field image-plane measurement equation as a true convolution. We nevertheless search for a wide-field image-plane measurement equation as close as possible to a convolution because all the inversion methods devised in the past three decades in radioastronomy are tuned to deconvolve images. The simplest mathematical way to generalize Eq. (22) to a shift-dependent situation is to write it as

$$I_{\text{dirty}}(\alpha) = \int_{\alpha'} D(\alpha - \alpha', \alpha) I(\alpha') \, d\alpha' .$$

(23)

In this section, we show how the linear character of the imaging process allows us to do this. Section 3.2.1 derives the impact of incomplete sampling on the Ekers & Rots equation, and Sect. 3.2.2 derives the wide-field measurement equation in the $uv$ plane. Section 3.2.3 interprets these results.

3.2.1. Effect on the Ekers & Rots equation

The sampled visibility function, $SV$, is defined as the product of $S$ and $V$ and $\overline{SV}$ its Fourier transform along $\alpha_s$, i.e.,

$$SV(u_p, \alpha_s) \equiv S(u_p, \alpha_s) V(u_p, \alpha_s) ,$$

(24)

and

$$\overline{SV}(u_p, u_s) \equiv \overline{SV}(u_p, u_s) .$$

(25)

Because $SV_{u_p}$ is the product of two functions of $\alpha_s$, we can use the convolution theorem to show that $\overline{SV}_{u_p}$ is the convolution of $\overline{SV}_{u_p}$ by $\overline{V}_{u_p}$, i.e.,

$$\overline{S}_{u_p}(u_s) = \int_{u_s'} \overline{S}_{u_p}(u_s - u_s') \overline{V}_{u_p}(u_s') \, du_s' .$$

(26)

By replacing $\overline{V}_{u_p}$ with the help of the Ekers & Rots relation (Eq. (15)), we derive

$$\overline{S}_{u_p}(u_s) = \int_{u_s'} \overline{S}_{u_p}(u_s - u_s') \overline{B}(-u_s') I(u_p + u_s') \, du_s' .$$

(27)

$^3$ i.e. direct Fourier transform along the $\alpha_s$ dimension, shift-and-average to define a final wide-field $uv$ plane, and inverse Fourier transform.

$^4$ Loosely speaking, the sampling function can be thought as a function whose value is 1 where there is a measure and 0 elsewhere.
As \( \overline{B} \) is bounded inside the \([-d_{\text{prim}},+d_{\text{prim}}] \) interval, \( \Sigma \overline{V}_{u_p}(u_s) \) is a local average, weighted by \( \overline{S}_{u_p}(u_s) \overline{B}(-u') \), of \( \overline{T}(u_p,u') \) around the \( u_p \) spatial frequency.

As expected, we recover Eq. (15) for the ideal case (i.e., infinite, continuous visibility function) because then \( \overline{S}_{u_p}(u_s-u'_\alpha) = \delta(u_s-u'_\alpha) \). A more interesting case arises when the visibility function is continuously sampled over a limited sky field of view, i.e.,

\[
\forall u_p, \quad S(u_p,\alpha) = 1 \quad \text{if} \quad |\alpha_s| \leq \theta_{\text{field}}/2, \\
\forall u_p, \quad S(u_p,\alpha) = 0 \quad \text{if} \quad |\alpha_s| > \theta_{\text{field}}/2. \tag{28}
\]

After Fourier transform this gives

\[
\forall u_p, \quad \overline{S}(u_p,u_s) = \frac{1}{d_{\text{field}}} \sin\left(\frac{u_s}{d_{\text{field}}}\right). \tag{30}
\]

In this case, the local average of the sky brightness Fourier components happens on a typical \( u \) scale equal to \( d_{\text{side}} \). However, the sinc function is known to decay only slowly. Some observing strategy (e.g. quickly observing outside the edges of the targeted field of view to provide a bandguard) could be considered to apodize the sky-plane dependence of the sampling function, resulting in faster decaying \( \overline{S} \) functions, hence in less mixing of the wide-field spatial frequencies.

### 3.2.2. uv-plane wide-field measurement equation

Because we aim at estimating the Fourier component of \( \overline{T} \), we introduce the following change of variables \( u' \equiv u_p + u'_\alpha \) and \( du' = du'_\alpha \), to derive

\[
\overline{S} \overline{V}_{u_p}(u_s) = \int u_p \overline{S}(u_p+u_s-u') \overline{B}(u_p-u') \overline{T}(u') du'. \tag{31}
\]

We then shift-and-average \( \overline{S} \overline{V}(u_p,u_s) \) to build the Fourier transform of a wide-field dirty image

\[
\overline{T}_{\text{dirty}}(u) \equiv \left\langle \overline{S} \overline{V}(u) \right\rangle, \quad \text{with} \quad u = u_p + u_s. \tag{32}
\]

Substituting the shift-and-average operator by its definition and using Eq. (31) to replace \( \overline{S} \overline{V}_{u_p}(u_s) \), we derive

\[
\overline{T}_{\text{dirty}}(u) = \int_{u_p} \int_{u_s} W(u_p,u-u_p) \overline{S}(u_p+u-u') \overline{B}(u_p-u') \overline{T}(u') du_p du'. \tag{33}
\]

This uv-plane wide-field measurement equation can be written as

\[
\overline{T}_{\text{dirty}}(u) = \int_{u_p} \overline{D} (u',u-u') \overline{T}(u') du', \tag{34}
\]

if we enforce the following equality

\[
\overline{D}(u',u-u') \equiv \int_{u_p} W(u_p,u-u_p) \overline{S}(u_p,u-u') \overline{B}(u_p-u') du_p. \tag{35}
\]

This is one way to define \( \overline{D} \), which is convenient though unusual. It is implicit in this definition that we need to make a change of variable \( u'' = u-u' \) to derive

\[
\overline{D} (u',u'') \equiv \int_{u_p} W(u_p,u'+u''-u_p) \overline{S}(u_p,u'') \overline{B}(u_p-u') du_p. \tag{36}
\]

In the following, we use either one or the other definition of \( \overline{D} \), depending on convenience.

### 3.2.3. Interpretation

 Appendix A.2 demonstrates that the image and uv-plane wide-field measurement equations (Eqs. (23) and (34)) are equivalent if

\[
D (u_p,\alpha_s) \overline{D} (u_p,\alpha) \overline{T}(u_p,u'). \tag{37}
\]

The image-plane wide-field measurement equation (Eq. (23)) can be written as

\[
I_{\text{dirty}}(\alpha) = [D^\alpha \ast I](\alpha). \tag{38}
\]

Its interpretation is straightforward: the sky brightness distribution is convolved with a dirty beam, \( D(\alpha',\alpha'') \), which varies with the sky coordinate \( \alpha'' \). This raises the question of the rate of change of the dirty beam with the sky coordinate. This question is addressed in Sects. 4.2 and 5.

### 4. Gridding by convolution and regular resampling

We want to Fourier transform the raw visibilities along the sky dimension \( \alpha_s \) at some constant value in the \( u_p \) dimension. The raw data, however, is sampled on an irregular grid in both the \( u \) and sky planes. We need to grid the measured visibilities in both the \( u \) and the sky planes before Fourier transformation for different reasons. First, the gridding in the \( u \) plane will handle the variation in the spatial frequency as the sky is scanned, i.e., the difficulty and perhaps the impossibility of Fourier-transforming at a completely constant \( u_p \) value. Second, the gridding along the sky dimension allows the use of Fast Fourier Transforms. As usual, we grid through convolution and regular resampling.

#### 4.1. Convolution

##### 4.1.1. Definitions

We first define a gridding kernel that depends on both dimensions, \( G(u_p,\alpha_s) \). This gridding kernel can be chosen as the product of two functions, simplifying the following demonstrations:

\[
G(u_p,\alpha_s) \equiv g(u_p) \gamma(\alpha_s). \tag{39}
\]

We then define the sampled visibility function gridded in both the \( u \) and sky planes as

\[
S V^G(u_p,\alpha_s) \equiv \left( G \ast S V \right)(u_p,\alpha_s) \equiv \int_{u_p}^{u_p} g(u_p-u_p') \gamma(\alpha_s-u_p') S V(u_p',\alpha_p') du_p' \tag{40}
\]

Finally, when assessing the impact of the gridding on the measurement Eq. (34), a new function,

\[
\Sigma(u_p,\alpha_s,\alpha_s') \equiv S(u_p,\alpha_s) B(\alpha_s'-\alpha_s), \tag{42}
\]

and its Fourier transforms naturally appear in the equations. Defining the following Fourier transform relationships

\[
\Sigma(u_p,\alpha_s,\alpha_s') \overline{D} (u_p,\alpha_s) \overline{T}(u_p,u'). \tag{36}
\]

In the following, we use either one or the other definition of \( \overline{D} \), depending on convenience.
we easily derive
\[ \Sigma(u_p, \alpha_s, u'_s) = S(u_p, \alpha_s) \overline{B}(u'_s) e^{-i2\pi u'_s \alpha_s}, \] (45)
and
\[ \Sigma(u_p, u_s, u'_s) = \overline{S}(u_p, \alpha_s + u'_s) \overline{B}(u'_s). \] (46)
Using these notations, we have before gridding,
\[ S V(u_p, \alpha_s) = \int_{u_p} \Sigma(u_p, \alpha_s, \alpha_p) I(\alpha_p) e^{-i2\pi u_p u_s} \, d\alpha_p, \] (47)
and
\[ \overline{G}(u', u-u') = \int_{u_p} W(u_p, u-u_p) \sum_{u_p, u_u} (u_p, u-u_p, u'-u') \, d\alpha_p. \] (48)

4.1.2. Conservation of the wide-field measurement equation

Appendix A.3 demonstrates that the wide-field dirty image is here again the convolution of the sky brightness \( I \) by a wide-field dirty beam \( D \) or, in the Fourier plane,
\[ T_{\text{dirty}}(u) = \int \overline{G}(u', u-u') \overline{I}(u') \, du' \] (49)
with
\[ \overline{G}(u', u-u') = \int W(u_p, u-u_p) \sum_{u_p, u_u}(u_p, u-u_p, u'-u') \, d\alpha_p, \] (50)
where
\[ \sum_{u_p, u_u}(u_p, u-u_p, u'-u') = \int_{u_p} g(u_p-u_p) \gamma(\alpha_s-\alpha_s') \sum_{u_p, u_u}(u_p, u-u_p, u'-u') \, du'_p d\alpha'_s. \] (51)
We thus have equations that resemble those containing the sampling function alone, except for 1) the replacement of the generalized sampling function \( \overline{G} \) by its gridded version \( \sum_{u_p, u_u} \) and 2) the way the variables are linked together both in the gridding of \( \overline{G} \) (i.e., Eq. (51)) and in the averaging of \( \sum_{u_p, u_u} \) (i.e., Eq. (50)).

4.2. Regular resampling

It is well known that too low a resampling rate in one space implies power aliasing in the conjugate space (see e.g. Bracewell 2000; Press et al. 1992). Aliasing must be avoided as much as possible because it folds power outside the imaged region back into it. Table 3 defines the intervals of definition of the different functions we are dealing with (i.e., visibilities, primary beam, dirty image, and dirty beam), as well as the associated sampling rates needed to enforce Nyquist sampling. The boundary values of the definition intervals \([u_{\text{max}}(\alpha_s)\) and \([\alpha_{\text{max}}(\alpha_s)\) are related to the sampling rates \((\Delta \alpha \) and \(\Delta u \), respectively) through
\[ |u_{\text{max}}| \cdot \Delta \alpha = |\alpha_{\text{max}}| \cdot \Delta u = \frac{1}{n_{\text{samp}}}, \] (52)
where \(n_{\text{samp}}\) is an integer characterizing the sampling rate. Nyquist sampling implies \(n_{\text{samp}} = 2\). However, slight oversampling (e.g. \(n_{\text{samp}} = 3\) is often recommended because the measures suffer from errors and the deconvolution is a nonlinear process. In this section, we examine the properties of the different functions to define their associated sampling rates.

<table>
<thead>
<tr>
<th>Functions</th>
<th>Intervals</th>
<th>Samplings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Visibilities</td>
<td>[</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td>[</td>
<td>\alpha</td>
</tr>
<tr>
<td>Primary beam</td>
<td>[</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td>[</td>
<td>\alpha</td>
</tr>
<tr>
<td>Dirty image</td>
<td>[</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td>[</td>
<td>\alpha</td>
</tr>
<tr>
<td>Dirty beam</td>
<td>[</td>
<td>u</td>
</tr>
<tr>
<td></td>
<td>[</td>
<td>\alpha</td>
</tr>
</tbody>
</table>

4.2.1. The \(\alpha_s\) sampling rate of the visibility function

When Fourier transforming the measurement Eq. (1) along the \(\alpha_s\) axis, we derive the Ekers & Rot's Eq. (15). This equation implies that \(V(u_p, \alpha_s)\) is bounded inside the \([-d_{\text{prim}}+d_{\text{prim}}] \) spatial frequency interval along the \(u_s\) axis. As a result, the visibility function needs to be regularly resampled at a rate of only 0.5\(d_{\text{prim}}\) to satisfy the Nyquist theorem. This was first pointed out by Cornwell (1988). This sampling rate is equal to \(\theta_{\text{prim}}/2\) or \(\sim \theta_{\text{whm}}/2.4\). The “usual”, “wrong” habit of sampling at \(\theta_{\text{whm}}/2\) is indeed undersampling with aliasing as a consequence. Mangum et al. (2007) discuss the consequences of undersampling in-depth in the framework of single-dish imaging.

4.2.2. The \(U_p\) sampling rate of the visibility function

Now, the Fourier transform of the measurement Eq. (1) along the \(u_p\) axis gives
\[ V(\alpha_p, \alpha_s) = B(\alpha_p-\alpha_s) I(\alpha_p), \] (53)
where
\[ V(\alpha_p, \alpha_s) \overset{\alpha_p}{\supset} V(u_p, \alpha_s). \] (54)
We use the tilde sign under \(V\) to denote the inverse Fourier transform of \(V\) along its first dimension. A well-known Fourier transform property implies that \(B\) has infinite support because \(\overline{B}\) is bounded. The resampling rate along the \(u_p\) axis therefore depends on the properties of the product of \(B(\alpha_p-\alpha_s)\), which is interesting as a function of \(\alpha_p\). While no unique answer exists, three facts help us to find the right sampling rate: 1) \(B\) falls off relatively quickly; 2) the result depends on the spatial distribution of the sky brightness and in particular on the dynamic range in brightness needed to accurately image it; 3) the measure of \(V(\alpha_p, \alpha_s)\) has a limited accuracy owing to thermal noise, phase noise, and other possible systematics (e.g. pointing errors). For simplicity, we quantify the measurement accuracy by a single number, namely the maximum instrumental fidelity measured in the image plane as defined in Pety et al. (2001). There are two cases:

1. the maximum instrumental fidelity limits the dynamic range in brightness. For instance, Pety et al. (2001) showed that the
The conjugate uv distance and associated uv sampling then are

\[ d_{\text{image}} = \frac{d_{\text{field}}}{1 + \frac{d_{\text{alias}}}{d_{\text{field}}}} \quad \text{and} \quad \partial u = \frac{d_{\text{image}}}{n_{\text{samp}}} \]

4.2.4. The \( u' \) and \( u'' \) sampling rates of \( \mathcal{T}(u', u'') \)

The \( u'' \) axis must thus be sampled at the same rate as the second dimension of the definition space of \( \mathcal{T} \), i.e., as \( u_s \). Moreover, \( u' \) has in this equation a behavior \( (u' = u_p + u_s) \) similar to \( u \) (\( = u_p + u_s \)). It must thus have the same sampling behavior as \( u \). This sampling rate (\( \partial u' = d_{\text{image}}/n_{\text{samp}} \)) is quite high. Some de-convolution methods (see below) allow us to relax this sampling rate.

4.3. Absence of gridding “correction”

Imaging of single-field observations goes through the following steps: 1) convolution by a gridding kernel; 2) regular resampling; 3) fast Fourier transform; and 4) gridding “correction”. The so-called gridding “correction” is a division of the dirty beam and dirty image by the Fourier transform of the gridding kernel used in the initial convolution. This step is mandatory when imaging single-field observations to keep the image-plane measurement equation as a simple convolution equation (see e.g. Sramek & Schwab 1989). When imaging wide-field observations, as proposed here, the Fourier transform along the \( \alpha \) dimension, followed by the shift-and-average operation, freeze the convolution kernel into the dirty beam of the wide-field measurement equation. This is why the gridding “correction” step is irrelevant here.

---

### Table 4. Minimum sizes of the dirty beam images to get an image fidelity or a dynamic range greater than a given value.

<table>
<thead>
<tr>
<th>Minimum fidelity or dynamic range</th>
<th>( \theta_{\text{alias}}/\theta_{\text{fwhm}} ) (( f_b = 0.0625 ))</th>
<th>( \theta_{\text{alias}}/\theta_{\text{fwhm}} ) (( f_b = 0.1 ))</th>
</tr>
</thead>
<tbody>
<tr>
<td>( 10^2 )</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>( 10^3 )</td>
<td>3.5</td>
<td>3.7</td>
</tr>
<tr>
<td>( 10^4 )</td>
<td>8.4</td>
<td>13.4</td>
</tr>
<tr>
<td>( 10^5 )</td>
<td>19.8</td>
<td>&gt;20.0</td>
</tr>
</tbody>
</table>

Notes. (a) The image sizes are expressed in units of the primary beam full width at half maximum. (b) The computation is done for 3 different ratios of the secondary-to-primary diameters (i.e. \( f_b \), the antenna blockage factors). The values are derived from the modeling of the antenna power patterns shown in Fig. 3.

The criterion derived in each case gives a typical image size (\( \theta_{\text{alias}} \)), which can be converted into the desired \( u_p \) sampling rate. To be more quantitative, Fig. 3 models the normalized antenna power patterns of an antenna illuminated by a Gaussian beam of 12.5 dB edge taper and with a given blockage factor (ratio of the secondary-to-primary diameters). The top panel presents an ideal case without secondary mirror, while the middle and bottom panels present simple models of the ALMA and PdBI antennas. The red lines define the minimum angular sizes for which the antenna power pattern is less than a given fraction. The middle and bottom panels respectively model ALMA and PdBI antennas. The red lines define the minimum angular sizes for which the antenna power pattern is less than a given fraction.
5. Dirty beams, weighting, and deconvolution

In radioastronomy, the dirty beam is the response of the interferometer to a point source. In the wide-field synthesis framework, the response of the interferometer to a point source, $D$, a priori depends on the source position on the sky. $D(\alpha',\alpha'')$ can thus be interpreted as a set of dirty beams, with each dirty beam referred to by its fixed $\alpha''$ sky coordinate. These simple facts raise several questions. What are the properties of the convolution kernel? Is it possible to modify these properties? How do we deconvolve the dirty image?

5.1. A set of wide-field dirty beams

With the wide-field synthesis framework proposed here, Appendix A.6 shows that

$$D(\alpha',\alpha'') = \int_{\Omega} B(\alpha''-\alpha'-\alpha_s) \Omega(\alpha'-\alpha_p,\alpha''-\alpha_s) \Delta(\alpha_p,\alpha_s) \, d\alpha_p d\alpha_s,$$  \hspace{1cm} (57)

where

$$\Delta(\alpha_p,\alpha_s) \supset S(u_p,\alpha_s),$$  \hspace{1cm} (58)

and

$$\Omega(\alpha',\alpha'') \supset W(u',\alpha'').$$  \hspace{1cm} (59)

$\Delta(\alpha_p,\alpha_s)$ is the single-field dirty beam, associated with the $uv$ sampling at the sky coordinate $\alpha_s$. And $\Omega(\alpha',\alpha'')$ will be called the image plane weighting function, while $W(u',\alpha'')$ is the $uv$ plane weighting function. The set of wide-field dirty beams $D$ is then the double convolution of the image plane weighting function and the single-field dirty beams, apodized by the primary beam at the current sky position $\alpha_s$.

While the shape of the single-field dirty beam is directly given by the Fourier transform of the sampling function, the shape of the wide-field dirty beam depends, directly or through Fourier transforms, on the sampling function $S$, the primary beam shape $B$, and the weighting function $W$. Moreover, the wide-field dirty beam shape a priori varies slowly with the sky position, since it is basically constant over the primary beamwidth as stated in Sect. 4.2. It nevertheless varies, implying, for instance, a “slow” variation of the synthesized resolution over the whole field of view.

While the single-field and wide-field dirty beam expressions seem very different, they share the same property of expressing the way the interferometer is used to synthesize a telescope of larger diameter in the image plane. In other words, the sampling function for single-field imaging and $D$ for wide-field imaging express the sensitivity of the interferometer to a given spatial frequency. These $uv$ space functions are called the transfer functions of the interferometer (Thompson et al. 1986, Chap. 5). Modifying the transfer function has a direct impact on the measured quantity. Once the interferometer is designed and the observations are done, the only way to change this transfer function is data weighting.

An ideal set of wide-field dirty beams, $D(\alpha',\alpha'')$, would have the following properties. All the wide-field dirty beams should be identical (i.e., independent of the $\alpha''$ sky coordinate) and equal to a narrow Gaussian (its FWHM giving the image resolution). This would give the product of a wide Gaussian of $u'$ by a Dirac function of $u''$, as the ideal wide-field transfer function, $\overline{D}(u',u'')$.

5.2. Dirty beam shapes and weighting

When imaging single-field observations, giving a multiplicative weight to each visibility sample is an easy way to modify the shape of the dirty beam and thus the properties of the dirty and deconvolved images. Natural weighting (which maximizes signal-to-noise ratio), robust weighting (which maximizes resolution), and tapering (which enhances brightness sensitivity at the cost of a lower resolution) are the most popular weighting techniques (see e.g. Sramek & Schwab 1989).

In the case of wide-field synthesis, a multiplicative weight can also be attributed to each visibility sample before any processing. However, the weighting is also at the heart of the wide-field synthesis because it is an essential part of the shift-and-average operation. No constraint has been set on the weighting function up to this point, which indicates that the weighting function ($W$) gives us a degree of freedom in the imaging process. We look in turn at both kinds of weighting. In both cases, an obvious issue is the definition of the optimum weighting functions. As in the case of single-field imaging, there is no single answer to this question. It depends on the conditions of the observation and on the imaging goals.

5.2.1. Weighting the measured visibilities

Natural weighting consists of slightly changing the definition of the sampling function. It is now set to a normalized natural weight where there is a measure and 0 elsewhere. The natural weight is usually defined as the inverse of the thermal noise variance, computed from the radiometric equation, i.e., from the system temperature, the frequency resolution, and the integration time. Using this weighting scheme before computing the first Fourier transform along the $\alpha_s$ sky dimension makes sense because the observing conditions (and thus the noise) vary from visibility to visibility.

We propose to generalize this weighting scheme to other observing conditions than just the system noise. Indeed, critical limitations of interferometric wide-field imaging are pointing errors, tracking errors, atmospheric phase noise (in the (sub-)millimeter domain), etc. While techniques exist for coping with these problems (e.g., water vapor radiometer, direction-dependent gains: Bhatnagar et al. 2008), they are not perfect. The usual way to deal with the remaining problems is to flag the source data based on a priori knowledge of the problems, e.g., pointing measurement, tracking errors, rms phase noise on calibrators, etc. However, flagging involves the definition of thresholds, while reality is never black and white. It can thus be asked whether some weighting scheme could be devised to minimize the effect of pointing errors, tracking errors or phase noise on the resulting image. We propose to modulate natural weighting based on the a priori knowledge of the observing conditions.

5.2.2. Weighting the synthesized visibilities

Robust weighting or tapering the measured visibilities do not make sense in wide-field synthesis because the dirty image is made from the synthesized visibilities after the first Fourier transform along the $\alpha_s$ sky dimension. A weighting function $W$ then appears naturally as part of the shift-and-average operator. Its optimum value depends on the properties of the measured sampling function. Here are a few examples.

Infinite, continuous sampling. This is the ideal case studied in Sect. 3.1. Knowing that the Ekers & Rots Eq. (15) links the
quantity we want to estimate, i.e., \( \hat{I} \), to many noisy\(^5\) measurements, \( \vec{V}(u_p, u_o) \), via a product by \( \vec{B} \) (assumed to be perfectly defined), we can invoke a simple least-squares argument (see e.g. Bevington & Robinson 2003) to demonstrate that the optimum weighting function is

\[
W(u_p, u - u_p) = \frac{w(u_p, u - u_p) \vec{B}(u_p - u)}{\int_{u_p} w(u_p, u - u_p) \vec{B}(u_p - u) \, du_p},
\]

(60)

with \( w(u_p, u_o) \) the weight computed from the inverse of the noise variance of \( \vec{V}(u_p, u_o) \). Using Eq. (20), it is then easy to demonstrate that \( \vec{B}(u) = 1 \), and then \( I_{\text{dirty}}(\alpha) = I(\alpha) \). The dirty image is a direct estimate of the sky brightness; i.e., deconvolution is superfluous.

**Complete sampling.** The signal is Nyquist-sampled, but it has a finite support in both the \( w \) and sky planes, implying a finite synthesized resolution and a finite field of view. In contrast to the previous case, this may have practical applications, e.g., observations done with ALMA in its compact configuration. Indeed, the large number of independent baselines coupled to the design of the ALMA compact configuration ensure a complete, almost Gaussian, sampling for each snapshot. In this case, the best choice may be to choose the weighting function so that all the dirty beams are identical to the same Gaussian function. In this case, the deconvolution would also be superfluous.

**Incomplete sampling.** This is the more general case studied in Sect. 3.2. The signal not only has a finite support but it also is undersampled (at least in the \( w \) plane). The deconvolution is mandatory. The choice of the weighting function thus will depend on imaging goals.

If the user needs the best signal-to-noise ratio, some kind of natural weighting will be needed. It is tempting to use Eq. (60) as a natural weighting scheme. However, the main condition for derivation of this weighting function, i.e., the Ekers & Rots Eq. (15), is not valid anymore, as the noisy measured quantity \( J \) is now linked to the quantity we want to estimate \( \hat{I} \) by a local average (see Eq. (31)). This is why it was more appropriate to try to get a Gaussian dirty beam shape in the complete sampling case. If the signal-to-noise ratio is high enough, the user has two choices. Either he/she wants to maximize angular resolution power and needs some kind of robust weighting, or he/she wants to get the more homogeneous dirty beam shape over the whole field of view. This requirement cannot always be fully met. The Ekers & Rots scheme enables us to recover unmeasured spatial frequencies only in regions near to measured ones, because \( \vec{B} \) has a finite support.

5 The noise is assumed to have a Gaussian probability distribution function.

5.3. Deconvolution

Writing the image-plane measurement equation in a convolution-like way is very interesting because all the deconvolution methods developed in the past 30 years are optimized to treat deconvolution problems (see e.g. Högbom 1974; Clark 1980; Schwab 1984; Narayan & Nityananda 1986). For instance, it should be possible to deconvolve Eq. (23) with just slight modifications to the standard CLEAN algorithms. Indeed, Eq. (23) can be interpreted as the convolution of the sky brightness by a set of dirty beams, so that the only change, once a CLEAN component is found, would be the need to find the right dirty beam in this set in order to remove the CLEAN component from the residual image.

Following Clark (1980) and Schwab (1984), most algorithms today deconvolve in alternate minor and major cycles. During a minor cycle, a solution of the deconvolution is sought with a simplified (hence approximate) dirty beam. During a major cycle, the current solution is subtracted either from the original dirty image using the exact dirty beam or from the measured visibilities, implying a new gridding step. In both cases, the major cycles result in greater accuracy. The iteration of minor and major cycles enables one to find an accurate solution with better computing efficiency. In our case, the approximate dirty beams used in the minor cycle could be 1) dirty beams of a much smaller size than the image; or 2) a reduced set of dirty beams (i.e., guessing that the typical variation sizescale of the dirty beams with the sky coordinate is much larger than the primary beamwidth); or 3) both simultaneously. The model would be subtracted from the original visibilities before re-imaging at each major cycle. The trade-off is between the memory space needed to store a full set of accurate dirty beams and the time needed to image the data at each major step. Some quantitative analysis is needed to know how far the dirty beams can be approximated in the minor cycle.

It is worth noting that the accuracy of the deconvolved image will be affected by edge effects. Indeed, the dirty brightness at the edges of the observed field of view is attenuated by the primary beam shape. When deconvolving these edges, the deconvolved brightness will be less precise, because the primary beam has a low amplitude there. This only affects the edges, because inside the field of view, every sky position should be observed a fraction of the time with a primary beam amplitude between 0.5 and 1. This edge effect is nevertheless expected to be much less troublesome than the inhomogeneous noise level resulting from standard mosaicking imaging (see Sect. 7.1).

6. Short spacings

6.1. The missing flux problem

Radio interferometers are bandpass instruments; i.e., they filter out not only the spacings longer than the largest baseline length but also the spacings shorter than the shortest baseline length, which is typically comparable to the diameter of the interferometer antennas. In particular, radio interferometers do not measure the visibility at the center of the \( w \) plane (the so-called “zero spacing”), which is the total flux of the source in the measured field of view.

The lack of short baselines or short spacings has strong effects as soon as the size of the source is more than about 1/3 to 1/2 of the interferometer primary beam. Indeed, when the size of the source is small compared to the primary beam of the interferometer, the deconvolution algorithms use, in one way or another, the information of the flux at the lowest measured spatial frequencies for extrapolating the total flux of the source. The extreme case is a point source at the phase center for which the amplitude of all the visibilities is constant and equal to the total flux of the source: extrapolation is then exact. However, the larger the size of the source, the worse the extrapolation, which then underestimates the total source flux. This is the well-known problem of the missing flux that observers sometimes note when comparing a source flux measured by a mm interferometer with the flux observed with a single-dish antenna.

Page 11 of 21
Wide-field synthesis does not recover the full short spacings. Let us assume that the visibility function is continuously sampled from \(d_{\min}\) to \(d_{\max}\), with \(d_{\min} \approx 1.5 \, d_{\text{prim}}\). The length of the averaging linepath\(^6\), \(L(u)\), can be interpreted as the number of measures that contribute to the estimate of \(\mathcal{T}(u)\).

6 The notion of averaging linepath has been introduced in Sect. 3.1 (see in particular Eq. (16)).

![Fig. 4.](image) Length of the averaging linepaths displayed as black lines in panel c.2) of Fig. 2, as a function of the spatial scale in the final, wide-field \(uv\) plane. In the case of a continuous sampling of \(u_p\) between \(d_{\min}\) and \(d_{\max}\), these quantities can be interpreted as the number of measures that contribute to the estimate of \(\mathcal{T}(u)\).

### 6.3. Processing short spacings from single-dish measurements

The single-dish measurement equation can be written as

\[
I_{sd}(\alpha) = S_{sd}(\alpha) \int_{\alpha'} B_{sd}(\alpha' - \alpha) \, I(\alpha') \, d\alpha',
\]

where \(I_{sd}\) is the measured single-dish intensity, \(S_{sd}\) the single-dish sampling function, and \(B_{sd}\) the single-dish antenna power pattern. As already stated in the introduction, the above integral is identical to the ideal measurement equation of interferometric wide-field imaging taken in \(u_p = 0\). If we define a single-dish visibility function as

\[
V_{sd}(u_p = 0, \alpha) \equiv \int_{\alpha'} B_{sd}(\alpha' - \alpha) \, I(\alpha') \, d\alpha',
\]

we can thus write the measured single-dish intensity as

\[
I_{sd}(\alpha) = S_{sd}(\alpha) \, V_{sd}(u_p = 0, \alpha).
\]

The recognition that the single-dish measurement equation is a particular case of the interferometric wide-field measurement equation opens the way to treating both the single-dish and interferometric data sets through exactly the same processing steps. We just have to define a hybrid sampling function, \(S_{hyb}\), as

\[
S_{hyb}(u_p \neq 0, \alpha) = S(u_p, \alpha)
\]

\[
S_{hyb}(u_p = 0, \alpha) = S_{sd}(\alpha),
\]

the Fourier transform of the hybrid primary beam, \(\mathcal{B}_{hyb}\), as

\[
\mathcal{B}_{hyb}(u_p \neq 0, u') = \mathcal{B}(u_p - u')
\]

\[
\mathcal{B}_{hyb}(u_p = 0, u') = \mathcal{B}_{sd}(-u'),
\]

and a hybrid weighting function, \(W_{hyb}\), as

\[
W_{hyb}(u_p \neq 0, u' + u'' - u_p) = W_{hyb}(u_p, u' + u'' - u_p)
\]

\[
W_{hyb}(u_p = 0, u' + u'') = W_{sd}(u' + u'').
\]

All the processing steps described in the previous sections (including a potential gridding step of single-dish, on-the-fly data) can then be directly applied to the hybrid data set. Using the wide-field synthesis formalism, we can easily write

\[
\mathcal{T}_{hyb}(u) = \int_{u_p} \mathcal{D}_{hyb} (u', u - u') \, \mathcal{T}(u') \, du',
\]

with

\[
\mathcal{T}_{hyb}(u) = \mathcal{T}_{dmy}(u) + W_{sd}(u) \, \mathcal{T}_{sd}(u)
\]

and

\[
\mathcal{D}_{hyb}(u', u'') = \mathcal{D}(u', u'') + W_{sd}(u' + u'') \, S_{sd}(u'') \, \mathcal{B}_{sd}(-u').
\]

We thus see that \(\mathcal{T}_{hyb}\) is a linear combination of the information measured by the single-dish (\(\mathcal{T}_{sd}\)) and by the interferometer (\(\mathcal{T}_{dmy}\)). There, \(W_{sd}(u)\) plays a particular role for two reasons. First, its dependency on the spatial frequency is exclusive of the single-dish antenna with low accuracy. Second, it is well-known that the relative weight of the single-dish to interferometric data is a critical parameter in the processing of the short spacings from single-dish data (see e.g. Rodríguez-Fernández et al. 2008).
Table 5. Definition of the symbols used to expose the processing of the short spacings.

<table>
<thead>
<tr>
<th>Symbol &amp; Definition</th>
<th>Plane(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$I_{sd}$</td>
<td>sky</td>
</tr>
<tr>
<td>$B_{sd}$</td>
<td>sky</td>
</tr>
<tr>
<td>$S_{sd}$</td>
<td>sky</td>
</tr>
<tr>
<td>$V_{sd}$</td>
<td>sky</td>
</tr>
<tr>
<td>$W_{se}$</td>
<td>$uv$</td>
</tr>
<tr>
<td>$b_i$</td>
<td>sky</td>
</tr>
<tr>
<td>$B_{ij}$</td>
<td>sky</td>
</tr>
<tr>
<td>$V_{ij}$</td>
<td>$uv$ &amp;sky</td>
</tr>
<tr>
<td>$b_{hyb}$</td>
<td>sky</td>
</tr>
<tr>
<td>$B_{hyb}$</td>
<td>sky</td>
</tr>
<tr>
<td>$S_{hyb}$</td>
<td>$uv$ &amp;sky</td>
</tr>
<tr>
<td>$W_{hyb}$</td>
<td>$uv$ &amp; $uv$</td>
</tr>
<tr>
<td>$D_{hyb}$</td>
<td>sky &amp;sky</td>
</tr>
</tbody>
</table>

Notes. This table uses similar conventions as Table 1. The top part defines symbols related to single-dish measurements. The middle part defines symbols related to heterogeneous-array measurements. The bottom part defines hybrid symbols, which result from combinations of single-dish and heterogeneous-array measurements.

This relative weight is a free parameter within the restrictions set by the noise level (i.e., we want the single-dish data to bring information and not just noise to the interferometric data), and a criterion must therefore be defined to adjust it to an optimal value. We refer the reader to the discussion of Sect. 5, which also applies here.

6.4. Processing short spacings from heterogeneous arrays

A heterogeneous array is an interferometer composed with antennas of different diameters. ALMA and CARMA are two such examples. The measurement equation for a heterogeneous array is

$$V_{ij}(\alpha_p, \alpha_s) = \int b_i(\alpha_p - \alpha_s) b_j^*(\alpha_p - \alpha_s) I(\alpha_p) e^{-i2\pi \nu x_{pq}} d\alpha_p,$$

(73)

where $b_i$ and $b_j$ are the voltage reception patterns of the antenna pair that forms the $ij$ baseline and the asterisk denotes the complex conjugate (Thompson et al. 1986, Chap. 3). The formalism developed in the previous sections holds as long as we redefine

$$B_{ij}(\alpha) \equiv b_i(\alpha) b_j^*(\alpha).$$

(74)

A simple application of the correlation theorem implies that

$$\overline{B_{ij}}(u) = \int u' \overline{b_i(u + u')} \overline{b_j(u')} du'.$$

(75)

The use of the baseline indices $ij$ must be generalized throughout the equations because the knowledge of the antenna type must be attached to each individual data point (visibility). As a result, the wide-field synthesis formalism can be easily adapted to heterogeneous arrays at the price of additional bookkeeping.

6.5. Two textbook cases: IRAM-30 m + PdBI and ALMA + ACA

Figure 5 sketches why wide-field synthesis naturally handles the short spacings in two textbook cases. In the ideal case, the Fourier transform along the $\alpha_s$ dimension produces visibilities, which are related to the wide-field spatial frequencies of the source brightness weighted by the transfer function of the interferometer. In this sense, Fig. 5 displays the natural weighting of the synthesized wide-field visibilities at the position of each measured visibility. Handling visibilities from antennas of different sizes just implies that the natural weighting function of the synthesized visibilities will have a different shape.

The top panel of Fig. 5 displays how the IRAM-30 m single-dish is used to complement the Plateau de Bure interferometer visibilities. The bottom panel displays how ACA is used to produce the short spacing information for ALMA. The four 12 m antennas will provide the single-dish information, while the 12 additional 7 m-antennas will form with ALMA a heterogeneous array. In the first design, ACA and ALMA form two independent interferometers; i.e., they are not cross-correlated. The single-dish antennas, ACA and ALMA, thus appear as three different instruments. It is thus possible to decompose the hybrid set of wide-field dirty beams obtained by processing the 3 sets of data together in 3 different sets of dirty beams

$$D_{hyb}(u', u'') = D_{zm}(u', u'') + D_{im}(u', u'') + D_{sd}(u', u''),$$

(76)
with
\[
\mathbf{\Omega}^{\text{hyb}}(u', u'') \equiv \int_{u_p} W_{\text{hyb}}(u_p, u' + u'' - u_p) \mathbf{S}_{\text{hyb}}(u_p, u'') \mathbf{\Omega}^{\text{hyb}}(u_p, u') \, du_p.
\]

For a multiplying interferometer,
\[
\forall \left| u_p \right| < d_{\text{prim}}, \quad S(u_p, \alpha) = 0 \quad \text{and} \quad \mathbf{S}(u_p, u_c) = 0.
\]

This implies that \( \mathbf{\Omega}^{\text{d}} \) contributes at \( u_p = 0 \) in the sum over \( u_p \) in Eq. (77), \( \mathbf{\Omega}^{\text{mos}} \) contributes for \( 9 \, m < u_p \leq 40 \, m \) and \( \mathbf{\Omega}^{\text{12 m}}(u', u'') \) contributes for \( 15 \, m < u_p < 150 \, m \) in the most compact configuration of ALMA.

7. Comparison with standard nonlinear mosaicking

7.1. Mosaicking in a nutshell

Several excellent descriptions of the mosaicking imaging and deconvolution algorithms can be found (see e.g. Cornwell 1988; Cornwell et al. 1993; Sault et al. 1996b). Here, we summarize the approach implemented in the GILDAS/MAPPING software used to image and deconvolve the data from the Plateau de Bure Interferometer. This approach is based on original ideas by F. Viallefond in the early 90s (Gueth et al. 1995).

The basic ideas of nonlinear mosaicking are 1) imaging the different fields of the mosaic independently; 2) linearly adding the single-field dirty images into a dirty mosaic; and 3) jointly deconvolving the dirty mosaic.

7.1.1. Single-field imaging

For simplicity, we skip the gridding convolution in the following equations because the gridding step does not change the nature of the equations. Imaging the fields individually means that we will work at constant \( \alpha_s \). We first define the single-field dirty image of the \( \alpha_s \)-field as
\[
I^{\text{std}}(\alpha_p, \alpha_s) \equiv \mathbf{S}(u_p, \alpha_s) \cdot \mathbf{\Omega}^{\text{std}}(u_p, \alpha_s),
\]
where the Fourier transform of the single-field dirty image is the product of the sampling function \( S(u_p, \alpha_s) \) and the visibility function \( V(u_p, \alpha_s) \):
\[
\mathbf{T}^{\text{std}}(u_p, \alpha_s) \equiv SV(u_p, \alpha_s).
\]
From the previous equations, it is easily demonstrated that
\[
I^{\text{std}}(\alpha_p, \alpha_s) = \int_{\alpha_p} \Delta(\alpha_p - \alpha_p', \alpha_s) \left[ B(\alpha_p' - \alpha_s) \mathbf{I}(\alpha_p') \right] \, d\alpha_p',
\]
where the single-field dirty beam is defined as
\[
\Delta(\alpha_p, \alpha_s) \equiv \frac{1}{S(\alpha_p, \alpha_s)}.
\]
We can rewrite the previous equation as
\[
I^{\text{std}}_p(\alpha_p) = \{ \Delta^{\text{std}} \ast (B^{\text{std}}) \}(\alpha_p),
\]
meaning that the single-field dirty images can be written as a local convolution of \( B^{\text{std}} \) and \( \Delta^{\text{std}} \), the single-field dirty beam associated to the currently imaged field.

7.1.2. Mosaicking the dirty images

In GILDAS/MAPPING\(^7\), the single-field dirty images are formed on the same grid (in particular the same pixel size and the same image size covering about twice the mosaic field of view). These single-field dirty images are then linearly averaged as
\[
I^{\text{mos}}(\alpha_p) \equiv \int_{\alpha_s} \Omega^{\text{mos}}(\alpha_p, \alpha_s) I^{\text{std}}(\alpha_p, \alpha_s) \, d\alpha_s,
\]
where
\[
\Omega^{\text{mos}}(\alpha_p, \alpha_s) \equiv \frac{w(\alpha_s) B(\alpha_p - \alpha_s)}{\int_{\alpha_s} w(\alpha_s) B^2(\alpha_p - \alpha_s) \, d\alpha_s}
\]
and \( w(\alpha) \) is the sky plane weighting function, i.e.,
\[
w(\alpha_s) = \sum_i \delta(\alpha_s - \alpha_i) \frac{1}{\sigma_i^2}.
\]
In the previous equation, the \( \alpha_i \) are the positions of each skyplane measurement, and \( \sigma_i \) is the rms noise associated with \( I^{\text{std}}_i \). Cornwell et al. (1993) demonstrates that the noise in the mosaic image naturally varies across the field as
\[
N^{\text{mos}}(\alpha_p) = \frac{1}{\sqrt{\int_{\alpha_s} w(\alpha_s) B^2(\alpha_p - \alpha_s) \, d\alpha_s}}.
\]
In particular, it rises sharply at the edges of the mosaic.

7.1.3. Joint deconvolution

Standard algorithms of single-field deconvolution must be adapted to the mosaicking case because both the dirty beam and the noise vary across the mosaic field of view. We describe here the adaptations made in GILDAS/MAPPING of the simplest CLEAN deconvolution method, described in Högbom (1974). Adaptations of more evolved CLEAN deconvolution methods are also implemented following the same basic rules.

1. We first initialize the residual and signal-to-noise maps from the dirty and noise maps
\[
R_0(\alpha_p) = I^{\text{mos}}(\alpha_p)
\]
and
\[
\text{SNR}_0(\alpha_p) = \frac{I^{\text{mos}}(\alpha_p)}{N^{\text{mos}}(\alpha_p)}.
\]
2. The \( k \)th CLEAN component is sought on the \( \text{SNR}_{k-1} \) map instead of the \( R_{k-1} \) map to ensure that noise peaks at the edges of the mosaic are not confused with the true signal of the same magnitude.
3. Using that the \( k \)th CLEAN component is a point source of intensity \( I_k \) at position \( \alpha_k \), the residual and signal-to-noise maps are then upgraded as follows:
\[
R_k(\alpha_p) = R_{k-1}(\alpha_p) - \gamma_k \int_{\alpha_s} \Omega^{\text{mos}}(\alpha_p, \alpha_s) \, d\alpha_s,
\]
\[
\text{SNR}_k(\alpha_p) = \frac{I^{\text{mos}}(\alpha_p)}{N^{\text{mos}}(\alpha_p)},
\]
\[
\gamma_k \equiv \frac{I_k}{\text{SNR}^2_0(\alpha_k)}.
\]

See http://www.iram.fr/IRAMFR/GILDAS for more information about the GILDAS software.
and

\[ \text{SNR}_k(\alpha_p) = \frac{R_k(\alpha_p)}{N_{\text{mos}}(\alpha_p)}. \]  

(91)

Here \( \gamma \approx 0.2 \) is the usual loop gain that ensures convergence of the CLEAN algorithms.

4. Steps 2 and 3 are iterated as long as the stopping criterion is not met.

### 7.1.4. Wide-field measurement equation

To help the comparison between mosaicking and wide-field synthesis, we now go one step further than is usually done in the description of mosaicking; i.e., we write the image-plane measurement equation as a wide-field measurement equation of the same kind (see Eqs. (23)). Substituting Eq. (81) into Eq. (84) and reordering the terms after inverting the order of the sum over \( \alpha_s \) and \( \alpha_p \), one obtains

\[ I_{\text{mos}}(\alpha_p) = \int D_{\text{mos}}(\alpha_p - \alpha_p', \alpha_p) I(\alpha_p') \, d\alpha_p', \]  

(92)

with

\[ D_{\text{mos}}(\alpha', \alpha'') = \int_{\Omega_{s+\alpha}} B(\alpha'' - \alpha' - \alpha_s) \Omega_{\text{mos}}(\alpha'', \alpha_s) \Delta(\alpha', \alpha_s) \, d\alpha_s. \]  

(93)

Taking the inverse Fourier transforms of \( D_{\text{mos}} \), we get the mosaicking transfer function

\[ \overline{D}_{\text{mos}}(u', u - u') = \int_{\Omega_{s+\alpha}} \hat{W}_{\text{mos}}(u - u_p, u_p - u') \hat{S}(u_p, u_p - u) \hat{B}(u_p - u') \, du_p \, du_s, \]  

(94)

with

\[ \Omega_{\text{mos}}(\alpha', \alpha'') \overset{(\alpha', \alpha'')}{\supset} \hat{W}_{\text{mos}}(u', u''). \]  

(95)

### 7.2. Comparison

While both mosaicking and wide-field synthesis produce image-plane measurement equations of the same kind (see Eqs. (23) and (92)), the comparison of the dirty beams (Eqs. (57) and (93)) and of the transfer functions (Eqs. (35) and (94)) immediately shows the different dependencies on the primary beams (\( B \)), the single-field dirty beams (\( \Delta \)), the image-plane weighting functions (\( \Omega \)), and their respective Fourier transforms (\( \hat{B}, \hat{S} \) and \( \hat{W} \)). This means that mosaicking is not mathematically equivalent to wide-field synthesis, though both methods recover the sky brightness. These differences come directly from the differences in the processing. If we momentarily forget the gridding steps, mosaicking starts with a Fourier transform along the \( u_p \) dimension of the visibility function, and most of the processing thus happens in the sky plane, while wide-field synthesis starts with a Fourier transform along the \( \alpha_s \) dimension, and most of the processing thus happens in the \( u_p \) plane.

Moreover, both methods are irreducible to each other. Wide-field synthesis gives a more complex dirty beam formulation in the image plane, which could give the impression that it is a generalization of mosaicking. Indeed, the wide-field image-plane weighting function can be chosen as the product of a Dirac function of \( \alpha' \) times a function \( \omega(\alpha'') \), i.e., \( \Omega(\alpha', \alpha'') = \delta(\alpha') \omega(\alpha'') \).

This implies a wide-field \( uv \)-plane weighting function independent of \( u' \); i.e., \( W(u', u'') = \delta(\alpha') \). This choice is a clear limitation because it enables us to influence the transfer function only locally (around each measured \( u_p \) spatial frequency), while weighting is generally intended to globally influence the transfer function (see Sect. 5). Eitherway, in this case, the wide-field dirty beam can easily be simplified to

\[ D(\alpha', \alpha'') = \int_{\Omega_{s+\alpha}} B(\alpha'' - \alpha' - \alpha_s) \omega(\alpha'' - \alpha_s) \Delta(\alpha', \alpha_s) \, d\alpha_s. \]  

(97)

While this simplified formulation of the wide-field dirty beam is closer to the mosaicking formulation, they still differ in a major way: \( \omega(\alpha'' - \alpha_s) \) is a shift-invariant function contrary to \( \Omega_{\text{mos}}(\alpha'', \alpha_s) \). This is the shift-dependent property of \( \Omega_{\text{mos}}(\alpha'', \alpha_s) \), which implies the additional complexity (integral over \( u_p \) in addition to the integral over \( u_s \)) of the mosaicking transfer function (Eq. (94)) over the wide-field one (Eq. (35)).

One main difference between the two processing methods is that standard mosaicking prescribes a precise weighting function, while we argue that the wide-field weighting function should be defined according to the context (see Sect. 5). Another important difference is the treatment of the short spacings, which are naturally processed in the wide-field synthesis methods, but which needs a very specific treatment in mosaicking (see Sect. 6 and references therein). Finally, while mosaicking implies a gridding only of \( u_p \) dimension of the measured visibilities, wide-field synthesis naturally requires a gridding of both the \( u_p \) and \( \alpha_s \) dimensions. As the Nyquist sampling along the \( \alpha_s \) dimension is only \( 0.5/\Delta_{\text{sample}} \), the gridding of the sky plane can result in a large reduction of the data storage space and cpu processing cost when processing on-the-fly and/or multi-beam observations.

### 8. Summary

Interferometric wide-field imaging implies scanning the sky in one way or another (e.g. stop-and-go mosaicking, on-the-fly scanning, sampling of the focal plane by multi-beams). This produces sampled visibilities \( SV \), which depends both on the \( uv \)-plane and sky coordinates (e.g., \( u_p \) and \( \alpha_s \)).

Based on a basic idea by Ekers & Rots (1979), we proposed a new way to image the interferometric wide-field sampled visibilities: \( SV(u_p, \alpha_s) \). After gridding the measured visibilities both in the \( uv \) and sky planes, the gridded visibilities \( SV^G \) are Fourier-transformed along the \( \alpha_s \) sky dimension, yielding synthesized visibilities \( SV^G \) sampled on a \( uv \) grid whose cell size is related to the total field of view; i.e., it is much finer than the diameter of the interferometer antennas. We thus proposed calling this processing scheme “wide-field synthesis”.

The Fourier transform is performed for each constant \( u_p \) value. As many independent estimates of the \( uv \) plane are produced as independent values of \( u_p \) measured. A shift-and-average operator is then used to build a final, wide-field \( uv \) plane, which translates into a wide-field dirty image after inverse Fourier transform, i.e.,

\[ \overline{S}^F_{\text{dirty}}(u) = \int_{\Omega_{s+\alpha}} W(u_p, u - u_p) \, SV(u_p, u - u_p) \, du_p, \]  

(98)
where $W$ is a normalized weighting function. Using these tools, we demonstrated that:

1. The dirty image ($i_{\text{dirty}}^D$) is a convolution of the sky brightness distribution ($I$) with a set of wide-field dirty beams ($D^F$) varying with the sky coordinate $\alpha$, i.e.,

$$i_{\text{dirty}}^D (\alpha) = \int_{\alpha'} D^F (\alpha - \alpha', \alpha) I (\alpha') \, d\alpha'. \tag{99}$$

Compared to single-field imaging, the dependency on the primary beam is transferred from a product of the sky brightness distribution into the definition of the set of wide-field dirty beams.

2. The set of gridded dirty beams ($D^G$) can be computed from the ungridded sampling function ($S$), the transfer function ($B$, the inverse Fourier transform of the primary beam), and the gridding convolution kernel (see Eqs. (42), (50) and (51)).

3. The dependence of the wide-field dirty beams on the sky position is slowly-varying, with their shape varying on an angular scale typically larger than or equal to the primary beamwidth.

Adaptations of the existing deconvolution algorithms should be straightforward.

A comparison with standard nonlinear mosaicking shows that it is not mathematically equivalent to the wide-field synthesis proposed here, though both methods do recover the sky brightness. The main advantages of wide-field synthesis over standard nonlinear mosaicking are

1. Weighting is at the heart of the wide-field synthesis because it is an essential part of the shift-and-average operation. Indeed, not only can a multiplicative weight be attributed to each visibility sample before any processing, but the on-plane weighting function ($W$, see Eq. (88)) is also a degree of freedom, which should be set according to the conditions of the observation and the imaging goals, e.g. highest signal-to-noise ratio, highest resolution, or most uniform resolution over the field of view. The $W$ weighting function thus enables us to modify the wide-field response of the instrument. On the other hand, mosaicking requires a precise weighting function in the image plane, which freezes the wide-field response of the interferometer.

2. Wide-field synthesis naturally processes the short spacings from both single-dish antennas and heterogeneous arrays along with the long spacings. Both of them can then be jointly deconvolved.

3. The gridding of the sky plane dimension of the measured visibilities, required by the wide-field synthesis, may potentially save large amounts of hard disk space and cpu processing power relative to mosaicking when handling data sets acquired with the on-the-fly observing mode. Wide-field synthesis could thus be particularly well suited to process on-the-fly observations.

The wide-field synthesis algorithm is compatible with the unufaceting technique devised by Sault et al. (1996a) to deal with the celestial projection effect, known as non-coplanar baselines (see Appendix B). Finally, on-the-fly observations imply an elongation of the primary beam along the scanning direction. These effects can be decreased by an increase in the primary beam sampling rate. However, it may limit the dynamic range of the image brightness if the primary beam sampling rate is too coarse (see Appendix C).

Acknowledgements. This work has mainly been funded by the European FP6 “ALMA enhancement” grant. This work was also funded by grant ANR-09-BLAN-0231-01 from the French Agence Nationale de la Recherche as part of the SCHISM project. The authors thank F. Gueth for the management of the on-the-fly working package of the “ALMA enhancement” project. They also thank S. Guilollet, R. Lucas and J. Usón for useful comments at various stages of the manuscript and D. Downes for editing their English. They finally thank the referee, B. Sault, for his insightful comments, which challenged us to try to write a better paper.

Appendix A: Demonstrations

A.1. Ekers & Rots scheme

Fourier-transforming the visibility function along the $\alpha_s$ dimension at constant $u_p$, we derive with simple replacements

$$\begin{align*}
\mathbf{V}_{\nu_p} (u_s) &= \int_{\nu_p} V_{\nu_p} (\nu_p) e^{-i2\pi \nu_p u_s} \, d\nu_p \quad \tag{A.1} \\
&= \int_{\nu_p, \nu_p'} B (\alpha_p - \alpha_p') I (\alpha_p') e^{-i2\pi (\nu_p + \nu_p u_s)} \, d\alpha_p d\nu_p \quad \tag{A.2}
\end{align*}$$

We then use the following change of variables $\beta = \alpha_p - \alpha_p'$. We then use the following change of variables $\beta = \alpha_p - \alpha_p'$, to get

$$\begin{align*}
\mathbf{V}_{\nu_p} (u_s) &= \int_{\nu_p} B (\beta) I (\alpha_p') e^{-i2\pi (\nu_p + \nu_p u_s - \beta \nu_p)} \, d\alpha_p d\beta \quad \tag{A.3} \\
&= \left[ \int_{\nu_p} B (\beta) e^{-i2\pi (\nu_p - \beta \nu_p)} \, d\beta \right] \left[ \int_{\alpha_p} I (\alpha_p') e^{-i2\pi \nu_p (\nu_p + \nu_p u_s)} \, d\nu_p \right] \quad \tag{A.4}
\end{align*}$$

A.2. Incomplete sampling

We here demonstrate that Eqs. (23) and (34) are equivalent. To do this, we take the direct Fourier transform of $i_{\text{dirty}}^D (\alpha)$

$$\begin{align*}
\mathbf{T}_{\text{dirty}} (u) &= \int_{\alpha} D (\alpha - \alpha', \alpha) I (\alpha') e^{-i2\pi \nu u} \, d\alpha' \quad \tag{A.6}
\end{align*}$$

and we replace $I (\alpha')$ by its formulation as a function of its Fourier transform

$$I (\alpha') = \int_{u'} \mathbf{T} (u') e^{+i2\pi \nu' u'} \, du'. \quad \tag{A.7}$$

We thus derive

$$\begin{align*}
\mathbf{T}_{\text{dirty}} (u) &= \int_{\alpha} \left[ \int_{\alpha, \alpha'} \left( D (\alpha' - \alpha, \alpha) e^{-i2\pi (\alpha - \alpha' u)} \right) \, d\alpha' \right] \mathbf{T} (u') \, du'. \quad \tag{A.8}
\end{align*}$$

Using the following change of variables $\alpha'' \equiv \alpha - \alpha'$, $\alpha'' = \alpha - \alpha''$ and $\alpha'' = -\alpha'$, the innermost integral can be written as

$$\begin{align*}
\int_{\alpha, \alpha'} \left( D (\alpha' - \alpha, \alpha) e^{-i2\pi (\alpha - \alpha' u)} \right) \, d\alpha' &= \int_{\alpha} \left[ \int_{\alpha} D (\alpha'', \alpha) e^{-i2\pi \nu (\alpha - \alpha' u)} \, d\alpha'' \right] e^{-i2\pi \nu (\alpha - \alpha' u)} \, d\alpha \quad \tag{A.9} \\
&= \int_{\alpha} D (\alpha', \alpha) e^{-i2\pi \nu (\alpha - \alpha' u)} \, d\alpha \quad \tag{A.10} \\
&= \mathbf{D} (\alpha', u - u'). \quad \tag{A.11}
\end{align*}$$

In the last two steps, we have simply recognized two different steps of Fourier transforms of $D$. Finally,

$$\begin{align*}
\mathbf{T}_{\text{dirty}} (u) &= \int_{u_p} \mathbf{D} (u', u - u') \, du'. \quad \tag{A.12}
\end{align*}$$
A.3. Gridding

The gridding kernel can be defined as the product of two functions, each one operating in its own dimension. We use this to study separately the effect of gridding in the \( uv \) and sky planes. We then use the intermediate results to get the effect of gridding simultaneously in both planes.

A.4. Gridding in the \( uv \) plane

We define the sampled visibility function gridded in the \( uv \) plane as

\[
S_{\text{V}}^{\gamma}(u_p, \alpha_s) \equiv \vert \gamma \ast S_{\text{V}}^{\alpha_s}(u_p) \vert
= \int_{u_p'} g(u_p - u_p') S_{\text{V}}^{\alpha_s}(u_p') \, du_p'.
\]  

(A.13)

Using that the gridding is here applied on the \( u_p \) dimension, while the Fourier transform is applied on the \( \alpha_s \) dimension, it is easy to show that the gridding and Fourier-transform operations commute:

\[
\bar{S}_{\text{V}}^{\gamma}(u_p, \alpha_s) =
\int_{u_p'} g(u_p - u_p') S(u_p', \alpha_s) e^{-i2\pi u_p u_p'} \, du_p',
\]  

(A.14)

\[
\bar{S}_{\text{V}}^{\gamma}(u_p, \alpha_s) = \int_{u_p'} g(u_p - u_p') \bar{S}_{\text{V}}(u_p', \alpha_s) \, du_p'.
\]  

(A.15)

Defining the Fourier transform of the \( uv \) gridded dirty image, we derive

\[
\mathcal{T}_{\text{dirty}}^\gamma(u) \equiv \left( \bar{S}_{\text{V}}^{\gamma} \right)(u)
\]  

(A.16)

\[
= \int_{u_p'} W(u_p, u-u_p) g(u_p - u_p') \bar{S}_{\text{V}}(u_p', \alpha_s) \, du_p
du_p'.
\]  

(A.17)

Using Eq. (31) to replace \( \bar{S}_{\text{V}}(u_p', \alpha_s) \), we can write the Fourier transform of the \( uv \) gridded dirty image as

\[
\mathcal{T}_{\text{dirty}}^\gamma(u) = \int_{u_p'} \mathcal{T}^\gamma(u', u-u_p') \mathcal{T}(u') \, du',
\]  

(A.18)

with

\[
\mathcal{T}^\gamma(u', u-u') \equiv \int_{u_p} W(u_p, u-u_p) \mathcal{T}(u_p, u-u_p, u') \, du_p
\]  

(A.19)

and

\[
\mathcal{T}^\gamma(u_p, \alpha_s, u') \equiv
\int_{u_p'} g(u_p - u_p') \bar{S}(u_p', \alpha_s - u' + u_p') \bar{B}(u_p' - u') \, du_p'.
\]  

(A.20)

Using

\[
\bar{S}(u_p', \alpha_s - u' + u_p') =
\left[ \int_{u_p} S(u_p', \alpha_s) e^{-i2\pi u_p u_p'} \, du_p \right] e^{-i2\pi (u_p' - u') \alpha_s},
\]  

(A.21)

and

\[
\bar{S}(u_p, \alpha_s, u') \overset{\alpha_s}{\ast} \bar{S}(u_p, \alpha_s, u')
\]  

(A.22)

we derive

\[
\bar{S}^\gamma_{uv}(u_p, \alpha_s, u') =
\int_{u_p'} g(u_p - u_p') S(u_p', \alpha_s) \bar{B}(u_p' - u') e^{-i2\pi (u_p' - u') \alpha_s} \, du_p'.
\]  

(A.23)

or

\[
\bar{S}^\gamma_{uv}(u_p, \alpha_s, u') = \int_{u_p'} g(u_p - u_p') \bar{S}(u_p', \alpha_s, u_p' - u') \, du_p'.
\]  

(A.24)

Thus, \( \bar{S}^\gamma_{uv} \) is the \( uv \) gridded version of the generalized sampling function \( \bar{S} \).

A.4.1. Gridding in the sky plane

We define the sampled visibility function gridded in the sky plane as

\[
S_{\text{V}}^{\gamma}(u_p, \alpha_s) \equiv \{ \gamma \ast S_{\text{V}}^{\alpha_s}(u_p) \}
= \int_{u_p'} \gamma(\alpha_s - \alpha_s') S_{\text{V}}^{\alpha_s}(u_p') \, d\alpha_s'.
\]  

(A.25)

Applying the convolution theorem on the Fourier transform along the \( \alpha_s \) dimension, we derive

\[
\bar{S}_{\text{V}}^{\gamma}(u_p, \alpha_s) = \bar{\gamma}(u_p) \bar{S}_{\text{V}}(u_p, \alpha_s).
\]  

(A.26)

Defining the Fourier transform of the sky-gridded dirty image, we derive

\[
\mathcal{T}_{\text{dirty}}^\gamma(u) \equiv \left( \bar{S}_{\text{V}}^{\gamma} \right)(u)
\]  

(A.28)

\[
= \int_{u_p} W(u_p, u-u_p) \mathcal{T}(u_p) \bar{S}_{\text{V}}(u_p, \alpha_s) \, du_p.
\]  

(A.29)

Using Eq. (31) to replace \( \bar{S}_{\text{V}}(u_p, \alpha_s) \), we can write the Fourier transform of the sky-gridded dirty image as

\[
\mathcal{T}_{\text{dirty}}^\gamma(u) = \int_{u_p} \mathcal{T}^\gamma(u', u-u_p') \mathcal{T}(u', u-u_p') \, du_p
\]  

(A.30)

with

\[
\mathcal{T}^\gamma(u', u-u_p') \equiv \int_{u_p} W(u_p, u-u_p) \mathcal{T}(u_p, u-u_p, u') \, du_p
\]  

(A.31)

and

\[
\mathcal{T}^\gamma(u_p, \alpha_s, u') \equiv \bar{\gamma}(u_p) \mathcal{T}(u_p, \alpha_s + u_p') \bar{B}(u_p'),
\]  

(A.32)

or, with the definition of \( \bar{\gamma} \) (i.e., Eq. (45)),

\[
\mathcal{T}^\gamma(u_p, \alpha_s, u') \equiv \bar{\gamma}(u_p) \bar{S}(u_p, \alpha_s, u_p').
\]  

(A.33)

Using

\[
\mathcal{T}^\gamma(u_p, \alpha_s, u_p') \overset{\alpha_s}{\ast} \bar{\gamma}(u_p) \mathcal{T}(u_p, \alpha_s, u_p')
\]  

(A.34)

and the convolution theorem when taking the inverse Fourier transform of \( \bar{\gamma} \), we derive

\[
\mathcal{T}^\gamma(u_p, \alpha_s, u_p') = \int_{u_p'} \gamma(\alpha_s - \alpha_s') \bar{S}(u_p, \alpha_s, u_p') \, d\alpha_s'.
\]  

(A.35)

Thus, \( \mathcal{T}^\gamma \) is the sky gridded version of the generalized sampling function \( \bar{\gamma} \).
A.5. Gridding in both planes

Starting from the definition of $S V^G$ (Eq. (41)), we Fourier-transform it along the sky dimension at constant $u_0$. Using that the gridding along the $u_0$ dimension can be factored out of the Fourier transform, we derive

$$S V^G_{u_0}(u_0) = \int_{u_0} g(u_0 - u'_0) S V^G_{u_0}(u_0) \, du'_0.$$  
(A.36)

Using Eq. (A.27), we now replace $S V^G_{u_0}(u_0)$ in the previous equation to get

$$S V^G_{u_0}(u_0) = \mathcal{F}(u_0) \int_{u_0} g(u_0 - u'_0) S V^G_{u_0}(u_0) \, du'_0.$$  
(A.37)

or

$$S V^G_{u_0}(u_0) = \mathcal{F}(u_0) S V^G_{u_0}(u_0).$$  
(A.38)

From this relation, it is easy to deduce that

$$S V^G_{u_0}(u_0, u_0') = \mathcal{F}(u_0) \sum S(u_0, u_0', u').$$  
(A.39)

Using the convolution theorem when taking the inverse Fourier transform of $S V^G_{u_0}$ along the $u_0$ dimension and replacing $S(u_0, u_0', u')$ with Eq. (A.24), we finally derive

$$S V^G(u_0, u_0', u') \equiv \int_{u_0} g(u_0 - u'_0) \gamma(u_0, u_0', u_0 - u') \, du'_0 \, d\alpha_s.$$  
(A.40)

A.6. Wide-field vs. single-field dirty beams

The notation (59) yields $W(u', u'') = \mathcal{F}(u', u'')$. Using this in Eq. (35) gives

$$D(u', u'') = \int_{u_0} \mathcal{F}(u_0, u_0' + u' - u_0) S(u_0, u_0') B(u_0 - u') \, du_0.$$  
(A.41)

Taking the inverse Fourier transform along the $u''$ axis of Eq. (A.41) and reordering the integral to factor out the term independent of $u''$, we can write

$$D(u', u'') = \int_{u_0} \mathcal{F}(u_0, u_0' - u_0) \mathcal{F}(u_0, u_0') FT1(u_0, u_0', u'') \, du_0,$$  
(A.42)

where

$$FT1(u_0, u_0', u'') = \int \mathcal{F}(u_0, u_0' + u' - u_0) S(u_0, u_0') e^{i2\pi(u_0 - u_0') \alpha} \, du_0.$$  
(A.43)

We now introduce the following definition

$$\bar{S}(u_0, u_0') \equiv \int S(u_0, \alpha) e^{-i2\pi u_0 \alpha} \, d\alpha,$$  
(A.44)

to derive

$$FT1(u_0, u_0', u'') = \int S(u_0, \alpha) \left[ \int \bar{S}(u_0, u_0' + u' - u_0) e^{i2\pi \alpha (u_0 - u_0')} \, d\alpha \right] \, d\alpha.$$  
(A.45)

Using the following change of variables $v \equiv u'' + u' - u_0$, $u'' = v - u' + u_0$ and $dv = du''$, we get

$$FT1(u_0, u_0', u'') = \int_{u_0} S(u_0, \alpha) \left[ \int \bar{S}(u_0, u_0' + u' - u_0) e^{i2\pi \alpha \alpha} \, d\alpha \right] \, d\alpha_s.$$  
(A.46)

Substituting this result into Eq. (A.42) and taking the inverse Fourier transform along the $u'$ axis, we can write

$$D(u', u'') = \int \mathcal{F}(u_0, \alpha_0 - \alpha) S(u_0, \alpha_0) FT2(u_0, \alpha_0, u', u'') \, du_0 \, d\alpha_s.$$  
(A.47)

where

$$FT2(u_0, \alpha_0, u', u'') \equiv \int \bar{S}(u_0, \alpha_0 - \alpha) \bar{S}(u_0, \alpha_0 - \alpha) \, d\alpha.$$  
(A.48)

Substituting this result into Eq. (A.47) and re-ordering the terms, we can write

$$D(u', u'') = \int \mathcal{F}(u_0, \alpha_0 - \alpha) S(u_0, \alpha_0) e^{i2\pi \alpha \alpha} \, du_0.$$  
(A.49)

A simple application of the convolution theorem gives

$$FT3(\alpha_0, \alpha_0', \alpha_0'') \equiv \int \mathcal{F}(u_0, \alpha_0' - \alpha) \bar{S}(u_0, \alpha_0') e^{i2\pi \alpha \alpha} \, du_0.$$  
(A.50)

We now introduce the following definition

$$\Delta(\alpha_0, \alpha_0', \alpha_0'') \equiv \int \bar{S}(u_0, \alpha_0' - \alpha) \, du_0.$$  
(A.51)

We now introduce the following definition

$$\Delta(\alpha_0, \alpha_0', \alpha_0'') \equiv \int \bar{S}(u_0, \alpha_0' - \alpha) \, du_0.$$  
(A.52)

Substituting this result into Eq. (A.47), we finally derive the desired expression, i.e., Eq. (57).

Appendix B: From the celestial sphere onto a single tangent plane

Equation (1) neglects projection effects, known as non-coplanar baselines. Any method which deals with interferometric wide-field imaging must take this problem into account. After a short introduction to the problem, we show how wide-field synthesis is compatible with at least one method, namely the $w$-unfacing of Sault et al. (1996b). This method tries to build a final wide-field $w$ plane from different pieces, just as our wide-field synthesis approach does. Another promising method is the projection, based on original ideas of Frater & Docherty (1980) and first successfully implemented by Cornwall et al. (2008). We did not look yet at its compatibility with wide-field synthesis.
B.1. w-axis distortion

When projection effects are taken into account, the measurement equation reads

\[ V(w, u_0, \phi) = \int B(\alpha_p - \alpha_s) \frac{I(\alpha_p)}{\sqrt{1 - \alpha_p^2}} e^{-i2\pi \omega u_0 \sqrt{1 - \alpha_p^2}} d\alpha_p. \] (B.1)

In this equation, we continue to work in 1 dimension for the sky cosine direction (\(\alpha_p\)), but we explicitly introduce the dependence along the direction perpendicular to the sky plane. This dependence appears in two ways, which is handled in very different ways. First, the factor \(\sqrt{1 - \alpha_p^2}\) can be absorbed into a generalized sky brightness function

\[ I(\alpha_p) \equiv \frac{I(\alpha_p)}{\sqrt{1 - \alpha_p^2}}. \] (B.2)

After imaging and deconvolution, \(I(\alpha_p)\) can be easily restored from the deconvolved \(I(\alpha_p)\) image. The second dependence appears as an additional phase, which is written as

\[ P(\alpha_p, u) \equiv e^{-i2\pi u \sqrt{1 - \alpha_p^2}}. \] (B.3)

Thompson et al. (1986, Chap. 4) shows that this additional phase can be neglected only if

\[ \frac{\pi \theta_{\text{field}}^2}{4 \theta_{\text{syn}}} \ll 1 \quad \text{or} \quad \frac{\pi \lambda d_{\text{max}}^2}{d_{\text{field}}^2} \ll 1. \] (B.4)

The first form of the criterion indicates that the approximation gets worse at high spatial dynamic range (i.e., \(\theta_{\text{field}}/\theta_{\text{syn}} \ll 1\)) while the second form indicates that the approximation gets worse at long wavelengths.

B.2. uvw-unfaceting

For stop-and-go mosaicking, it is usual to delay-track at the center of the primary beam for each pointing/delay of the mosaic. This phase center is also the natural center of projection of each pointing/delay. Stop-and-go mosaicking thus naturally paves the way to a uvw-unfaceting scheme. In the framework of on-the-fly observations with ALMA, D'Addario & Emerson (2000) indicate that the phase center will be modified between each on-the-fly scan while it will stay constant during each on-the-fly scan. This is a compromise between loss of coherence and technical possibilities of the phase-locked loop. Using this hypothesis, the maximum sky area covered by the on-the-fly scan must take into account the maximum tolerable w-axis distortion.

The easiest way to deal with such data is to image each pointing/delay around its phase center and then to reproject this image onto the mosaic tangent plane as displayed in Fig. 5 of Sault et al. (1996b). These authors point out that this scheme implies a typical w-axis distortion \(\epsilon\) less than

\[ \epsilon \leq (1 - \cos \theta_{\text{alias}}) \sin \theta_{\text{center}} \sim \frac{1}{2} \theta_{\text{center}} \theta_{\text{alias}}^2, \] (B.5)

where \(\theta_{\text{center}}\) is the angle from the pointing/field center and \(\theta_{\text{alias}}\) is the anti-aliasing scale defined in Sect. 4.2. In particular, \(\epsilon\) is 0 at the phase center of each pointing/field. In other words, this scheme limits the magnitude of the w-axis distortion to its magnitude on a size equal to the anti-aliasing scale (i.e., a few times the primary beamwidth) instead of a size equal to the total mosaic field of view. This scheme thus solves the projection effect as long as the w-axis distortion is negligible at sizes smaller than or equal to the anti-aliasing scale. A natural name for this processing scheme is uvw-unfaceting because it is the combination of a faceting observing mode (i.e., regular change of phase center) and a linear transform of the uv coordinates to derive a single sine projection for the whole field of view.

Sault et al. (1996b) also demonstrate that the reprojection may be done much more easily and quickly in the uvw space before imaging the visibilities because it is then just a simple transformation of the w coordinates, followed by a multiplication of the visibilities by a phase term. Finally, Sault et al. (1996b) note that it is the linear character of this wv coordinate transform which preserves the measurement Eq. (1). As the change of coordinates happens before any other processing, it also conserves all the equations derived in the previous sections to implement the wide-field synthesis.

Appendix C: On-the-fly observing mode and effective primary beam

Usual interferometric observing modes (including stop-and-go mosaicking) implies that the interferometer antennas observe a fixed point of the sky during the integration time. Conversely, the on-the-fly observing mode implies that the antennas slew on the sky during the integration time. This implies that the measurement Eq. (1) must be written as (Holdaway & Foster 1994; Rodríguez-Fernández et al. 2009):

\[ V(\hat{\alpha}, \hat{\phi}) = \frac{1}{\delta t} \int_{\delta t/2}^{\delta t/2} B(\alpha_p - \alpha_s(t)) I(\alpha_p) e^{-i2\pi u \sqrt{1 - \alpha_p^2}} d\alpha_p \, dt, \] (C.1)

where \(\delta t\) is the integration time and \(\hat{\alpha}\) and \(\hat{\phi}\) are the mean spatial frequency and direction cosine, defined as

\[ \hat{\alpha} = \frac{1}{\delta t} \int_{-\delta t/2}^{\delta t/2} \alpha_s(t) \, dt \quad \text{and} \quad \hat{\phi} = \frac{1}{\delta t} \int_{-\delta t/2}^{\delta t/2} \phi_s(t) \, dt. \] (C.2)

In this section, we analyze the consequences of the antenna slewing on the accuracy of the wide-field synthesis.

C.1. Time averaging

In all interferometric observing modes, it is usual to adjust the integration time so that \(u_T(t)\) can be approximated as \(\hat{u}\). To do this, it is enough to ensure that \(u_T(t)\) always varies less than the \(w\) distance associated with tolerable aliasing (\(d_{\text{alias}}\), see Sect. 4.2) during the integration time \(\delta t\)

\[ \delta t \ll \frac{d_{\text{alias}}}{d_{\text{max}} \omega_{\text{earth}}} \quad \text{or} \quad \frac{d\theta}{ds} \ll \frac{6900}{\theta_{\text{alias}}/\theta_{\text{syn}}}, \] (C.3)

where \(d_{\text{max}}\) is the maximum baseline length, \(\omega_{\text{earth}}\) is the angular velocity of a spatial frequency due to the Earth rotation \((7.27 \times 10^{-5} \text{ rad s}^{-1})\), \(\theta_{\text{alias}}\) and \(\theta_{\text{syn}}\) are respectively the minimum field of view giving a tolerable aliasing and the synthesized beam angular values.
Table C.1. Definition of the symbols used to explore the influence of on-the-fly scanning on the measurement equation.

<table>
<thead>
<tr>
<th>Symbol &amp; Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta t$</td>
</tr>
<tr>
<td>$\tilde{\alpha}_s$</td>
</tr>
<tr>
<td>$\tilde{\alpha}_p$</td>
</tr>
<tr>
<td>$\delta \alpha_s$</td>
</tr>
<tr>
<td>$v_{\text{slew}}$</td>
</tr>
<tr>
<td>$A$</td>
</tr>
<tr>
<td>$B_{\text{eff}}$</td>
</tr>
<tr>
<td>$\omega_{\text{earth}}$</td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

C.2. Effective primary beam

Assuming that condition (C.3) is ensured, we can write Eq. (C.1) with the same form as Eq. (1) by the introduction of an effective primary beam ($B_{\text{eff}}$); i.e.,

$$V\left(\tilde{\alpha}_p, \tilde{\alpha}_s\right) = \int_{\tilde{\alpha}_p} B_{\text{eff}}\left(\alpha_p - \tilde{\alpha}_s\right) I\left(\alpha_p\right) e^{-i2\pi \alpha_p \delta \alpha_s} d\alpha_p,$$

where

$$B_{\text{eff}}\left(\alpha_p - \tilde{\alpha}_s\right) = \frac{1}{\delta t} \int_{\delta \alpha_s \approx -\delta \alpha_s / 2}^{\delta \alpha_s / 2} B(\alpha_p - \alpha_s(t)) d\alpha_s.$$

Using the following change of variables

$$\beta \equiv \alpha_s(t) - \tilde{\alpha}_s,$$

we derive

$$B_{\text{eff}}\left(\alpha_p - \tilde{\alpha}_s\right) = \int B\left[(\alpha_p - \tilde{\alpha}_s) - \beta\right] A(\beta) d\beta,$$

with

$$A(\beta) \equiv \frac{1}{v_{\text{slew}}(\beta)} \delta t \left(\frac{\beta}{\delta \alpha_s}\right)$$

and

$$\delta \alpha_s = \int_{\delta \alpha_s / 2}^{\delta \alpha_s / 2} v_{\text{slew}}(t) dt.$$

In these equations, $v_{\text{slew}}(\beta)$ is the slew angular velocity of the telescope as a function of the sky position, $\delta \alpha_s$ is the angular distance covered during $\delta t$, $A$ is an apodizing function, and $\Pi(\beta)$ is the usual rectangle function, which reproduces the finite character of the time integration.

C.3. Interpretation

The form of the measurement equation is conserved when averaging the visibility function over a finite integration time, as long as the true primary beam is replaced by an effective primary beam, which is the convolution of the true primary beam by an apodizing function. To go further, it is important to return to the two dimensional case. Indeed, the convolution must be done along the slewing direction, resulting in an effective primary beam elongated in a particular direction.

In principle, the equations derived in Sect. 3 can be accommodated just by replacing the true primary beam by its effective associate. In practice, the probability to take into account the effective primary beam is low because its shape varies with time. Indeed, it is often assumed that the sky is slewed along a straight line at constant angular velocity. Even in this simplest case, it is advisable to slew at least two perpendicular directions to average systematic errors, implying two different effective primary beams. However, practical reasons may/may lead to complex scanning patterns: 1) the limitation of the acceleration when trying to image a square region leads to spiral or Lissajous scanning patterns; 2) the probable absence of derotators in future multi-beam receivers (B. Lazareff, private communication) implies the need to take into account the Earth rotation in the scanning patterns of the off-axis pixels.

C.4. Approximation accuracy

In the following, we thus ask what is the trade-off accuracy of using the true primary beam instead of the effective primary beam. The first point to mention is that using different scanning patterns somehow helps because the averaging process then makes the bias less systematic. Following Holdaway & Foster (1994), we quantify the accuracy lost in the Fourier plane. Indeed, the Ekers & Rots scheme tries to estimate missing sky brightness Fourier components from their measurements apodized by the Fourier transform of the primary beam. In the Fourier space, the above convolution just translates into a product. The Fourier transform of the apodizing function thus degrades the sensitivity of the measured visibility, $V(\alpha_p, \alpha_s)$, to spatial frequencies at the edges of the interval $[\alpha_p - \alpha_p, \alpha_p + \alpha_p]$. To guide us in our quantification of the accuracy lost, we now explore the simplest case of linear scanning at constant velocity, where $v_{\text{slew}}(\beta)$ is constant and $\delta \alpha_s = v_{\text{slew}}(\beta) \delta t$. The Fourier transform of the apodizing function is then a sinc function:

$$\overline{A}(u) = \text{sinc} (u \delta \alpha_s).$$

The relative error implied by the use of the true primary beam instead of the effective primary beam is then

$$\frac{B_{\text{eff}}(u) - B(u)}{B_{\text{eff}}(u)} = 1 - \frac{1}{A(u)} = 1 - \frac{1}{\text{sinc} (u \delta \alpha_s)}.$$

Figure C.1 shows this relative error as a function of the number of samples per primary beam FWHM in the image plane (i.e.,
\( \theta_{\text{FWHM}} / \delta \alpha \) for different uv distances (in units of \( d_{\text{prim}} \)). We see that we derive a 1% accuracy at all \( u \) when we sample the image plane at a rate of 5 dumps per primary beam. However getting a 0.1% accuracy needs quite high sampling rates (about 15). This must be compared with the accuracy of knowledge of \( B \).

We note that if a better accuracy is needed than the one achievable with the highest sampling rate, it is in theory possible to replace in the correlator software the rectangle apodizing function by another function which falls more smoothly. To avoid the loss of sensitivity inherent to the use of such an apodizing function (by throwing away data at the edges of the time interval of integration), would require, for instance, to half-overlap the integration intervals. This would imply more book-keeping in the correlator software and some noise correlation between the measured visibilities.

References

Bevington, P. R., & Robinson, D. K. 2003, Data Reduction and Error Analysis For the Physical Sciences, 3rd edn. (McGraw-Hill)
Cornwell, T. J., Golap, K., & Bhatnagar, S. 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 647
D’Addario, L. R., & Emerson, D. T. 2000, On-The-Fly Fringe Tracking, ALMA memo, 331
Holdaway, M. A., & Foster, S. M. 1994, On-The-Fly Mosaicing, ALMA memo, 122
Pety, J., Gueth, F., & Guilloteau, S. 2001, Impact of ACA on the Wide-Field Imaging Capabilities of ALMA, ALMA memo, 398
Rodríguez-Fernández, N. J., Pety, J., & Gueth, F. 2008, Single-dish observation and processing to produce the short-spacing information for a millimeter interferometer, IRAM memo, 2008-2