Mon. Not. R. Astron. Soc. **000**, 1–15 (2013) Printed 8 April 2014 (MN LaTeX style file v2.2)

Signals Correlation Algorithm For Cheaper Surveys: Using Windowing Functions.

M. T. Atemkeng[1], O. M. Smirnov[12⋆], C. Tasse[123], G. Foster[124], J. Jonas[12]
[1]*Department of Physics and Electronics, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa*
[2]*SKA South Africa, 3rd Floor, The Park, Park Road, Pinelands, 7405, South Africa*
[3]*GEPI, Observatoire de Paris, CNRS, Universite Paris Diderot, 5 place Jules Janssen, 92190 Meudon, France*
[4]*xxxxx*

in original form 1988 October 11

ABSTRACT
This paper investigates the use of baseline dependent windowing functions in interferometry data to minimize the loss of signal amplitude (smearing) when the correlated data is averaged over wide bandwidth and long time. In radio interferometry smearing is reduced when a cross-correlator averages the correlated data over narrower bandwidth and shorter integration times. Unfortunately, this leads to a huge amount of data to manage and it is becoming a bottleneck for further data processing such as calibration and imaging. With future generation surveys, it is important to investigate the reduction of the output data rate. Therefore, the focus of this paper is on the use of baselines dependent windowing functions to keep smearing down at an acceptable extent and at the same time significantly suppress signals from out field of view sources, while the nominal sensitivity is conserved.

Key words: Instrumentation: interferometers, Methods: data analysis, Methods: numerical, Techniques: interferometric

1 INTRODUCTION

The recent radio astronomy techniques is to build a single, gigantic instrument called *interferometer*, from the combination of several small parabolic antennas separated over kilometres (**?**). The signal from each antenna is combined at the level of a cross-correlator to form the interferometer data output. The cross-correlator carries out data reduction and filters out an amount of noise by averaging the signal of each baseline over discrete time and/or frequency bins. It is well known in interferometry that averaging can lead to the loss of signal amplitude when the cross correlator integrate over a longer period of time and a wider bandwidth. This effect is known as time-average smearing and bandwidth smearing (Thompson et *al.* **?**). The above effects cause the distortion of sources within the field of interest by decreasing their intensity.

To keep smearing down at acceptable levels, a correlator must cross-correlate the signal over a shorter period of time and a narrower bandwidth, hence producing a large amount of data for subsequence stage such as imaging (Martí-Vidal & Marcaide 2008; Linfield 1986), calibration (**?**), etc. This huge amount of data is becoming an increasingly serious problem and becoming more challenging as the computational demands of the next generation radio telescopes will rise significantly (see the SKA phase 1 specification **?**). Similarly, the next generation of radio telescopes will require an unprecedented level of SNR while mapping large regions of the sky. Thus, a substantial increase in *SNR* can only be achieve by observing for longer time at wider bandwidth without loss of signal: this is not realisable with averaging. Therefore, it becomes urgent to develop new decorrelation algorithm techniques that will allow the required SNR of the future radio telescopes.

In this paper, we investigate the efficiency of correlator windowing functions for the reduction of interferometric data and the recovery of interferometers arrays desire FoV, with the ultimate goal of reaching higher SNR. The main idea is to achieve a high SNR by conserving the astrophysical signal and by limiting the noise. Thermal noise can be driven arbitrary low by increasing the observing time[1], but in radio astronomy confusion noise is a major problem and can even cause calibration to fail. Therefore, we seek to use a windowing function that will conserve the useful signal while

⋆ E-mail: o.smirnov@ru.ac.za (OMS); m.atemkeng@gmail.com (MTA); cyril.tasse@obspm.fr (CT); griffin.foster@gmail.com (GF); j.jonas@ru.ac.za (JJ)

[1] ...

also limiting sidelobes confusion from out of FoV sources.

$$SNR = \frac{S_{use}}{N_{ter} + N_{con}} \tag{1}$$

- "*Useful signal*", S_{use} the signal from source in the field of interest. These sources should be accurately recovered over the instrument entire FoV : correlator windowing functions maximized this signal by allowing the interferometer array to map a large region of the sky.
- "*Sidelobe confusion*", N_{con} signal from out field of view sources received from their sidelobes. These sources are not of interest and should be removed : correlator windowing functions acts as a remover of these signals when the array is mapping a large region of the sky.
- "*Thermal noise*", N_{ter} the thermal noise from the instrument, ionosphere, etc. Averaging presents theoretically a maximum sensitivity, but the use of extended correlator windowing functions can reduce or eliminate the loss of the nominal sensitivity.

The proposed techniques are applied to MeerKAT (Karoo Array Telescope) **?** and the Very Large Array (VLA)**?** and could also be used for future radio telescopes such as the SKA.

2 OVERVIEWS AND DEFINITIONS

2.1 Visibility and relation with the sky

An interferometer array measured a quantity $V = V(u, v, w)$ known classically as the visibility function (see **?**). The variables u, v and w are in unit of wavelength and they are the coordinates of the vector of which the norm is the distance between two antennas, known in interferometry as a baseline. A source in the sky will see u and v oriented towards the direction East-West and South-North respectively and w is directed towards the phase centre of the source plane or image plane. The projection of u and v in the image plane are l and m respectively. They are the observed source coordinates, measured in radian. The ideal measurement of interferometric wide-field imaging also known as the van Cittert-Zernike theorem (Thompson et *al*. 2001, Eq.6) is given by

$$V_{pq} = \int\int \frac{I(l, m)}{\sqrt{1 - l^2 - m^2}} e^{-2\pi i\phi(u,v,w)} dl dm, \tag{2}$$

where $I(l, m)$ is the sky brightness and $\phi(u, v, w) = u.l + v.m + w.(\sqrt{1 - l^2 - m^2} - 1)$ is a term from the cross-correlator that models the direction in the sky and the separation of the two antennas. The term $\sqrt{1 - l^2 - m^2}$ is the result of the projection of the celestial sphere on the image plane.

2.2 Averaging and convolution

The Earth rotation causes the phase, $\phi(u, v, w)$ to variate in time. The baseline coordinates are defined in units of wavelength, and making $\phi(u, v, w)$ to variate in frequency. To take this effect into account, Eq. 2 is rewritten as an integration over time and frequency interval. If we consider that $[t_s, t_e]$ is the time integration interval and $[\nu_s, \nu_e]$ the frequency integration interval, then Eq.2 can be rewritten as:

$$V_{pq}^{avg} = \frac{1}{\Delta t \Delta \nu} \int_{t_s}^{t_e} \int_{\nu_s}^{\nu_e} V_{pq,(t,\nu)} d\nu dt \tag{3}$$

V_{pq} is a continuous function, in reality we know only the sampled visibility, $V_{pq,(t,\nu)}^{samp} = S_{pq,(t,\nu)} V_{pq,(t,\nu)}$ at a specific time and frequency. $S_{pq,(t,\nu)}$ is a sampling function that indicates where the (u, v) data for the baseline (p, q) is measured during the time and frequency integration. Therefore, Eq.3 holds for many sources, when the signal at the centre frequency, ν_c and at the centre time, t_c is restricted to a narrow frequency interval and to a short time interval $[t_s, t_e]$ respectively, this is the current efficient observing mode. However, this is expressed mathematically as

$$V_{pq}^{avg} = \frac{1}{n_t n_\nu} \sum_{i=1}^{n_t} \sum_{j=1}^{n_\nu} S_{pq,(t_i,\nu_j)} V_{pq,(t_i,\nu_j)}. \tag{4}$$

Here, n_t and n_ν are the number of discrete times within the time interval and the number of discrete frequency within the frequency interval respectively. For convenience, lets introduce a normalized *Boxcar* windowing function, $\Pi_{pq,(t_c-t,\nu_c-\nu)}$ that will attribute an equal weight (in this case $1/n_t n_\nu$) to all sampling visibilities points. We can therefore rewrite Eq. 4 as:

$$V_{pq}^{agv} = \sum_{i=1}^{n_t} \sum_{j=1}^{n_\nu} \Pi_{pq,(t_c-t_i,\nu_c-\nu_j)} S_{pq,(t_i,\nu_j)} V_{pq,(t_i,\nu_j)} \tag{5}$$

$$= \sum_{i=1}^{n_t} \sum_{j=1}^{n_\nu} \Pi_{pq,(t_c-t_i,\nu_c-\nu_j)} V_{pq,(t_i,\nu_j)}^{samp}. \tag{6}$$

It is worth noting that Eq.6 is a naturally weighted visibility and is a two dimensional convolution between the *Boxcar* windowing function and the sampled visibility V_{pq}^{samp}. Thus, averaging is equivalent to convolving the sampled visibility with a *Boxcar* windowing function. Mathematically, this is described as follows:

$$V_{pq}^{avg} = c_{pq,(t,\nu)} \cdot \left(\left(\Pi_{pq} \circ V_{pq}^{samp} \right)_{(t,\nu)} \right). \tag{7}$$

Here, $c_{pq,(t,\nu)}$ is a function that samples the result of $\left(\Pi_{pq} \circ V_{pq}^{samp} \right)_{(t,\nu)}$ at the centre time interval and centre frequency interval.

2.3 Effect of time and bandwidth averaging

During imaging, Eq.7 is inverse Fourier transform, and after applying the convolution theorem , the *sinc* function ($\mathcal{F}^{-1}\Pi_{pq,(t,\nu)} = sinc(\pi \Delta u_{pq} ||\mathbf{l}||)$) multiply the sky. Thus, the sky map is tapered by the *sinc* function in the l and m direction, the response is maximal for sources at the phase centre ($l = 0, m = 0$) while for off-phase centre sources, the response is smeared (decreased) for larger Δu_{pq} (Δu_{pq} is a function of Δt and $\Delta \nu$). The main lobe of the *sinc* extends from $-1/2\Delta u_{pq}$ to $1/2\Delta u_{pq}$ while the amplitude gradually dies out, and the larger the Δu_{pq}, the narrower the central peak and the oscillations. Therefore, the degree of smearing increases with the position of a source and the baseline length (see Thompson et *al*.)[2].

(i) Fig.1 shows the attenuation of a source at various coordinates in the sky for various integration time interval. We measure more than 90% of the source brightness for integrations less than or equal to $25s$ when the source is within the Field of view and more than 90% when the source is out of the Field of View. As mention above, for small integration like $25s$ we produce large data and maintain strong sidelobes contamination from out field of view sources.
(ii) Fig.3

Fortunately, since the response is maximal only for sources at the phase centre, an interesting approach is achieved by convolving the observed visibility with a windowing function that depends on (u, v) coordinates spacing (baseline dependent windowing function). However, a windowing function with a wide dynamic range spectrum is preferable in this work.

2.4 Imaging

From the full sky Radio Interferometry Measurement Equation (RIME) formalism (see Hamaka et *al*, O.M. Smirnov (2010a)), in this case, the sampled visibilities can be presented mathematically as a $4 \times n_t \times n_\nu$ matrix of four polarizations time and frequency dependent matrices each of size $n_t \times n_\nu$.

$$\begin{aligned} \mathbf{V}^{samp}_{pq,(t,\nu)} &= \mathbf{S}_{pq,(t,\nu)} \cdot \mathbf{V}_{pq,(t,\nu)} \\ &= \left(\mathbf{V}^0_{pq,(t,\nu)}, \mathbf{V}^1_{pq,(t,\nu)}, \mathbf{V}^2_{pq,(t,\nu)}, \mathbf{V}^3_{pq,(t,\nu)} \right)^T . \end{aligned}$$

Now, consider that $\mathbf{W}_{pq,(t,\nu)}$ is a $n_t \times n_\nu$ matrix that contained the weights of the baseline (p, q) visibilities points. The weights of this matrix are given by a baseline dependent windowing function that we described in section 3 and section 4. The convolution operator is linear, therefore we can rewrite Eq.7 in terms of a series of linear transformations as follow:

$$V^{avg}_{pq} = \mathbf{W}^{block}_{pq,(t,\nu)} \cdot \mathbf{S}_{pq,(t,\nu)} \cdot \mathbf{V}_{pq,(t,\nu)}. \tag{8}$$

Here, $\mathbf{W}^{block}_{pq,(t,\nu)}$ is a block diagonal matrix of size $(4n_t n_\nu) \times (4n_t n_\nu)$ and the block elements are $\mathbf{C}_{pq,(t,\nu)} \cdot \mathbf{W}_{pq,(t,\nu)}$ of size $n_t \times n_\nu$, where $\mathbf{C}_{pq,(t,\nu)}$ is the centre time interval and centre frequency interval sampling matrix of size $n_t \times n_\nu$. This is the result of the time and frequency integration for the baseline (p, q).

For a synthesis, the baseline (p, q) made a full coverage in the (u, v) plane. Therefore, we can package into a single matrix, $\mathbf{V}^{avg}_{pq,}$ of size $(4N_t N_\nu) \times (4N_t N_\nu)$ the weighted average visibilities of the baseline (p, q) during the synthesis as follows:

$$\mathbf{V}^{avg}_{pq} = \mathbf{W}^{block,n}_{pq,(t,\nu)} \cdot \mathbf{S}^n_{pq,(t,\nu)} \cdot \mathbf{V}_{pq,(t,\nu)}, \tag{9}$$

where N_t and N_ν are the number of time sample and frequency channels entering the Fourier domain. If the synthesis time is T and the frequency range is F, then $T = N_t \times \Delta t$ and $F = N_\nu \times \Delta \nu$. The the size of $\mathbf{V}^{avg}_{pq}$ can also be written as $(4N^{pq}_v) \times (4N^{pq}_v)$, where N^{pq}_v is the number of time and frequency visibilities for the baseline (p, q)). The matrix $\mathbf{W}^{block,n}_{pq,(t,\nu)}$ is a diagonal block matrix of size $(4N^{pq}_v n_t n_\nu) \times (4N^{pq}_v n_t n_\nu)$ where each diagonal block is the block diagonal matrix $\mathbf{W}^{block}_{pq,(t,\nu)}$ defined above and n is the number of $\mathbf{W}^{block}_{pq,(t,\nu)}$. The sampled visibilities $\mathbf{S}^n_{pq,(t,\nu)} \cdot \mathbf{V}_{pq,(t,\nu)} = \mathbf{V}^{samp,n}_{pq,(t,\nu)}$ is a one row matrix of size $(N^{pq}_v 4n_t n_\nu) \times (4n_t n_\nu)$ made of $\mathbf{V}^{samp}_{pq,(t,\nu)}$ on top of each other.

$$\mathbf{V}^{avg}_{pq} = \mathbf{W}^{block,n}_{pq,(t,\nu)} \cdot \mathbf{S}^n_{pq,(t,\nu)} \mathbf{F} \cdot \mathcal{I}^{sky}_{l,m}, \tag{10}$$

if the number of pixel in the sky model is N_{pix}, then the true sky image vector $\mathcal{I}^{sky}_{l,m}$ has a size of $4N_{pix}$ and $\mathbf{F}$ is the Fourier transform operator of size $(4N_{pix}) \times (4N_{pix})$.

We are generally interested in using the total set of visibilities over baselines, time and frequencies, having $4 \times N_v$ visibilities measured over all baselines and $N_v = n_{bl} \times N^{pq}_v$ in this case. Here, n_{bl} is the number of baseline. We can write

$$\mathbf{V}^{avg}_{all} = \mathbf{A} \cdot \mathcal{I}^{sky}_{l,m}. \tag{11}$$

Here, $\mathbf{A}$ is a matrix of size $(4N_v) \times (4N_{pix})$ made of $\mathbf{W}^{block,n}_{pq,(t,\nu)} \cdot \mathcal{S}^n_{pq,(t,\nu)} \cdot \mathbf{F}$ on top of each other. The dirty image, $\mathcal{I}^D_{l,m}$ of size $4N_{pix}$ can then be derived

$$\mathcal{I}^D_{l,m} = \mathbf{F}^H \cdot \mathbf{A} \cdot \mathcal{I}^{sky}_{l,m}. \tag{12}$$

Here, H represents the the conjugate transpose operation also known as a Hermitian transpose and $\mathbf{F}^H$ is the inverse Fourier transform operator of size $(4N_{pix}) \times (4N_{pix})$.

3 DATA COMPRESSION ALGORITHM

3.1 Description

Missing spaces between sampled (u, v) coordinates has huge dependences on the baseline length. However, the spacing between longer baselines (u, v) coordinates are wider then the one on shorter baselines: this explained while sources are more distorted on longer baselines. We aims in this section, to describe the algorithm we use with a baseline-dependent windowing function to assign a proper weight to a data reference by a (u, v) coordinate considering the *spacing*[3] between the baseline (u, v) coordinates.

Fig.5 shows a snapshot coverage of an integration interval. For shorter baselines, the tracks are closer to the centre of rotation and for longer baselines the tracks are farther away from this centre. The *dot marks* are the data for a sampled (u, v) data, and the arrows indicates the separation between (u, v) coordinates and the centre (u, v) coordinate. It is trivial to see on this figure that these separations are wider

[2] The Fourier phase components $2\pi\phi(u, v, w)$ depends on the direction in the sky, the wavelength, the separation of antennas as well as the integration time and frequency. A maximal phase will occurs on longer baselines and small phase on shorter baselines.

[3] The *distance* has also huge dependences on the baseline length and allow us to formally define the data weight of a *uv* point over the entire *uv* plane.

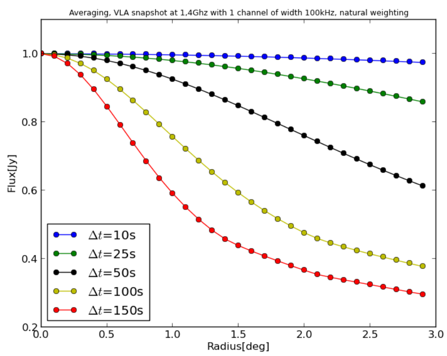

Figure 1. The fall of the intensity of a 1Jy source move from the phase centre for Δt integration synthesis at 100KHz bandwidth.

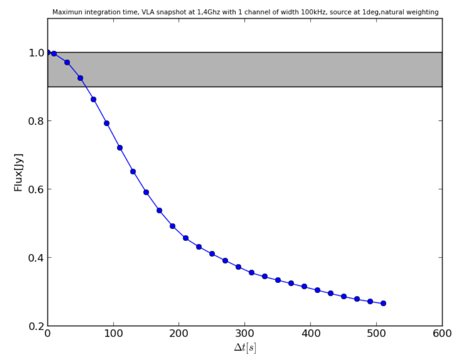

Figure 2. (b)Response to a 1Jy source at 1deg, as a function of Δt with 100KHz bandwidth

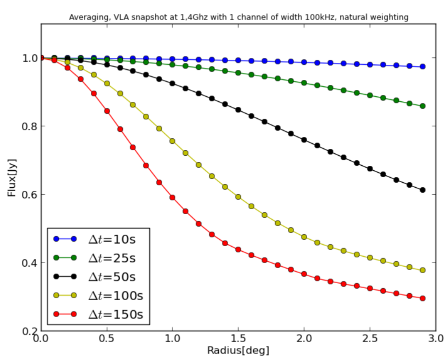

Figure 3. The fall of the intensity of a 1Jy source move from the phase centre for Δt integration synthesis at 100KHz bandwidth.

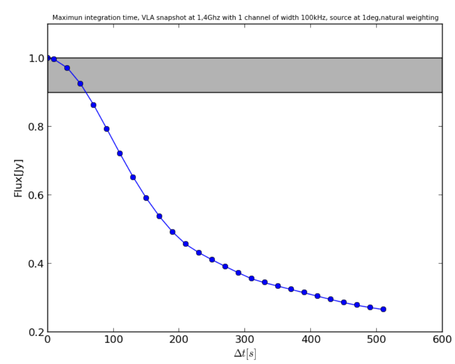

Figure 4. The fall of the intensity of a 1Jy source move from the phase centre for Δt integration synthesis at 100KHz bandwidth.

on longer baselines. The results of averaging is assigned to the centre (u, v) coordinate coloured in red.

3.2 Methods

In this section, we present the data compression algorithm we use to describe the weight of each sampled visibility during the integration time interval and and frequency interval. Now, let package the (u, v) coordinates changes of a baseline pq into a single matrix of size $n_t \times 2$.

$$\mathbf{U}_{pq,t} = \left(\mathbf{u}_{pq,t_s}, \ldots, \mathbf{u}_{pq,t_c}, \ldots, \mathbf{u}_{pq,t_e} \right)^T, \tag{13}$$

where the indexes s, c and e references the integration interval starting, centre, and ending time respectively. The elements of $\mathbf{U}_{pq,t}$ are functions of time and frequency representing a (u, v) coordinate. The frequency changes of the baseline coordinates can be package into a single vector of dimension n_ν

$$\boldsymbol{\nu} = \left(\nu_s, \ldots, \nu_c, \ldots, \nu_e \right)^T$$

We define a norm, $||\cdot||_m$ on a $n_t \times 2$ matrix as follow:

$$||\mathbf{U}_{pq,t}||_m = \left(||\mathbf{u}_{pq,t_s}||, \ldots, ||\mathbf{u}_{pq,t_c}||, \ldots, ||\mathbf{u}_{pq,t_e}|| \right), \tag{14}$$

where $||.||$ is the Euclidean norm.

Definition 1.1. (Time direction spacing) The matrix that model the spacing between the (u, v) coordinates and the centre (u, v) coordinate of a baseline (p, q) across the time direction is defined as

$$\mathbf{U}^s_{pq,t} = \frac{\nu_c}{c} \cdot \left\{ \mathbf{U}_{pq,t} - \mathbf{H}_{pq,t} \right\},$$

where c is the speed of the light and $\mathbf{H}_{pq}$ is a matrix of size $n_t \times 2$ that model the centre uv-coordinate, $\mathbf{H}_{pq,t} = \left(\mathbf{u}_{pq,t_c}, \ldots, \mathbf{u}_{pq,t_c}, \ldots, \mathbf{u}_{pq,t_c} \right)^T$.

Definition 1.2. (Frequency direction spacing) The vector of size n_ν that model the spacing between the (u, v) coordinates and the centre (u, v) coordinate of a baseline (p, q) across the frequency direction is defined as

$$\mathbf{d}^T_\nu = \frac{||\mathbf{u}_{pq,t_c}||}{c} \cdot \left\{ \boldsymbol{\nu} - \nu_c \cdot \mathbf{g}^T_\nu \right\},$$

where $\mathbf{g}$ is a $1 \times n_\nu$ unity matrix.

Definition 1.3. (Baseline dependent windowing function) If f_{pq} is a *baseline dependent windowing function*, then:

$$\begin{aligned} f_{pq}:\{\mathcal{R},\mathcal{R}\} &\rightarrow \mathcal{R} \\ d_{t_i}, d_{\nu_j} &\mapsto \frac{w_{t_i,\nu_j}}{\sum_{i=1}^{n_t}\sum_{j=1}^{n_\nu} w_{t_i,\nu_j}}. \end{aligned}$$

where d_{t_i} is an element of the vector $\mathbf{d}_t = ||\mathbf{U}^s_{pq,t}||_m$ and d_{ν_j} is an element of the vector $\mathbf{d}^T_\nu$.

3.3 Noise analysis

For convenience, in this section, we analyse the centre pixel noise of a sky map after we had apply a baseline dependent windowing function during the time and frequency integration. The brightness of the naturally weighted central pixel is estimate by (see the inverse Fourier transform of equation (19-2) in **?**).

$$\begin{aligned} \widetilde{\mathcal{I}}(0,0) &= \frac{1}{N_v}\sum_{k=1}^{N_v} V^{avg}_{all,k} \\ &= \frac{1}{N_v}\sum_{pq,k=1}^{N_v^{pq}} V^{avg}_{pq,k} \\ &= \frac{1}{N_v}\sum_{pq,k=1}^{N_v^{pq}}\Big(\sum_{i=1}^{n_t}\sum_{j=1}^{n_\nu} f_{pq,(dt_i,d\nu_j)}V^{samp}_{pq,(t_i,\nu_j)}\Big). \end{aligned}$$

Here, the variable k run through the baseline (p,q) timeslots. The overall rms noise is given by

$$\sigma^2 = \frac{\sigma^2_{samp}}{N_v^2}\Big(\sum_{pq,k=1}^{N_v^{pq}}\sum_{i=1}^{n_t}\sum_{j=1}^{n_\nu} f^2_{pq,(dt_i,d\nu_j)}\Big),$$

where σ_{samp} is the per sampled visibility noise

4 OUT FOV SOURCES SUPPRESSION ALGORITHM

4.1 Description

In theory, windowing functions and signals generally extend to infinity. Unfortunately, in practice, filtering a signal with a low pass filter, one need to define a cut-off interval. Therefore, if one wants to achieve sufficiently an accurate estimate of the windowing function ideal spectrum, one need a wide cut-off interval as far as the spectrum approaches the ideal when the windowing function order increases. An overlap baseline dependent windowing function aims to extend the order of the baseline dependent windowing function in such a way that, we approaches the ideal spectrum. The only drawback of this technique is the increased of the time needed for processing the output sample of the signals being integrate.

4.2 Methods

The weight of a visibility is not defined by a unique baseline dependent windowing, but by the strength of the correlation between the overall overlapping baseline dependent windowing functions on the visibility. Now, consider that f^a_{pq} is an overlap-*BDWF* of width Δt and $\Delta\nu$ across the time and frequency direction respectively.

Definition 1.4. (Baseline dependent windowing function) if $\Delta_l t$ and $\Delta_l\nu$ are the overlap time interval and frequency interval of the baseline dependent windowing function $f^{a_0}_{pq}$ respectively and $\left\{f^{a_1}_{pq}, f^{a_2}_{pq}, f^{a_3}_{pq},\dots\right\}$ the set of *BDWF* overlapping on the *left hand side* of $f^{a_0}_{pq}$ then the resulting *BDWF* within $\Delta_l t$ and $\Delta_l\nu$ is defined as

$$\begin{aligned} g^{lhs}_{pq}:\{\mathcal{R},\mathcal{R}\} &\rightarrow \mathcal{R} \\ d_{t_i}, d_{\nu_j} &\mapsto \frac{1}{N_{lhs}}\left(\sum_k f^{a_k}_{pq,(d_{t_i},d_{\nu_j})} + f^{a_0}_{pq,(d_{t_i},d_{\nu_j})}\right). \end{aligned}$$

Here, N_{lhs} is the normalization term defined as

$$N_{lhs} = \sum_{i=1}^{n_{lt}}\sum_{j=1}^{n_{l\nu}}\left(\sum_k f^{a_k}_{pq,(d_{t_i},d_{\nu_j})} + f^{a_0}_{pq,(d_{t_i},d_{\nu_j})}\right),$$

where n_{lt} and $n_{l\nu}$ are the number of (u,v) coordinates changes and frequency changes within $\Delta_l t$ and $\Delta_l\nu$ respectively.

Definition 1.5. (Baseline dependent windowing function) if $\Delta_r t$ and $\Delta_r\nu$ are the overlap time and frequency interval of a *BDWF* $f^{a_0}_{pq}$ respectively and $\left\{f^{a_1}_{pq}, f^{a_2}_{pq}, f^{a_3}_{pq},\dots\right\}$ the set of *BDWF* overlapping on the *right hand side* of $f^{a_0}_{pq}$, then the resulting *BDWF* within $\Delta_r t$ and $\Delta_r\nu$ is defined as

$$\begin{aligned} g^{rhs}_{pq}:\{\mathcal{R},\mathcal{R}\} &\rightarrow \mathcal{R} \\ d_{t_i}, d_{\nu_j} &\mapsto \frac{1}{N_{rhs}}\left(f^{a_0}_{pq,(d_{t_i},d_{\nu_j})} + \sum_k f^{a_k}_{pq,(d_{t_i},d_{\nu_j})}\right). \end{aligned}$$

Here, N_{rhs} is the normalization term defined as

$$N_{rhs} = \sum_{i=1}^{n_{rt}}\sum_{j=1}^{n_{r\nu}}\left(f^{a_0}_{pq,(d_{t_i},d_{\nu_j})} + \sum_k f^{a_k}_{pq,(d_{t_i},d_{\nu_j})}\right),$$

where n_{rt} and $n_{r\nu}$ are the number of (u,v) coordinates changes and frequency changes within $\Delta_r t$ and $\Delta_r\nu$ respectively.

From the above definitions, the following derivation is trivial

$$\left\{\Delta t,\ \Delta\nu\right\} = \left\{\Delta_l t\cup\Delta_u t\cup\Delta_r t,\ \ \Delta_l\nu\cup\Delta_u\nu\cup\Delta_r\nu\right\},$$

where $\Delta_u t$ and $\Delta_u\nu$ are $f^{a_0}_{pq}$ uncorrelated time and frequency interval respectively. They follows the below rules

$$\Delta_u t = \begin{cases} \cup\{t_i\}_{i=s',\ s'\geqslant s+1}^{c',\ c'\leqslant c-1} & \text{if } n_{lt}+n_{rt} < n_t \\ \{t_c\} & \text{if } n_{lt}+n_{rt} = n_t \\ \emptyset & \text{otherwise} \end{cases}$$

$$\Delta_u \nu = \begin{cases} \cup\{\nu_i\}_{i=s',\ s'\geqslant s+1}^{c',\ c'\leqslant c-1} & \text{if } n_{l\nu}+n_{r\nu} < n_\nu \\ \{\nu_c\} & \text{if } n_{l\nu}+n_{r\nu} = n_\nu \\ \emptyset & \text{otherwise} \end{cases}$$

The resulting *BDWF* becomes g_{pq} described as

$$g_{pq} = \begin{cases} g^{lhs}_{pq} & \text{if } (t,\nu)\in(\Delta_l t,\Delta_l\nu) \\ f^{a_0}_{pq} & \text{if } (t,\nu)\in(\Delta_u t,\Delta_u\nu) \\ g^{rhs}_{pq} & \text{if } (t,\nu)\in(\Delta_r t,\Delta_r\nu) \end{cases}$$

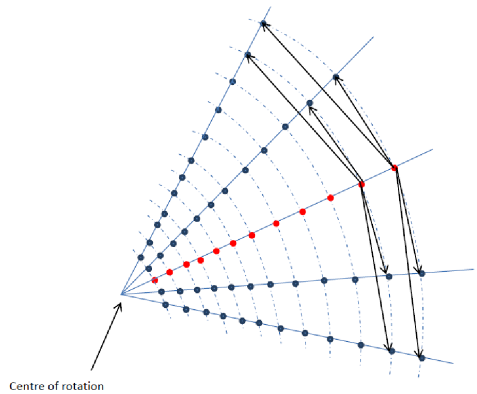

Figure 5. Snapshot coverage

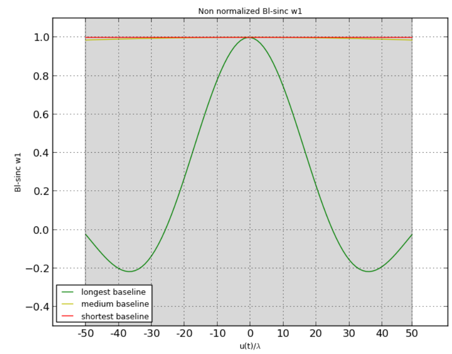

Figure 6. *Bl-sincw*1 used to convolve the visibility data obtained from the long, medium and short baseline

4.3 Noise analysis

The overall rms noise is given by

$$\sigma^2 = \frac{\sigma^2_{samp}}{N^2_v}\Big(\sum_{pq,k=1}^{N_v^{pq}}\sum_{i=1}^{n_t}\sum_{j=1}^{n_v} g^2_{pq,(dt_i,dv_j)}\Big),$$

4.4 Windowing functions

In signal processing a windowing function is a mathematical function that is zero-values outside of some chosen interval, and when another function or a signal is multiplied by the windowing function, the product is also zero-values outside the interval. In this section, we evaluation the Peak Sidelobe Level (PSL), the Main Lobe width (MLW) and the Sidelobes Roll-off (SLR) of some windowing functions spectrum. This study will allow us to make a suitable choice of the window that by tapering with the sky, we conserve the brightness of sources in the field of interest and attenuate sidelobes confusion from strong sources out of the field of interest.

4.4.1 *Boxcar window: natural weighting*

This windowing function take a hunk of the data without modification, and this leads to discontinuities at the edges[4]. For a cut-off frequency interval $[-\nu_a, \nu_a]$ the boxcar windowing function is defined as:

$$\Pi_u = \begin{cases} 1 & -\nu_a \leqslant \nu \leqslant \nu_a \\ 0 & \text{otherwise} \end{cases} \tag{15}$$

Fig.7 and Fig.8 gives the graph of Π_u and its spectrum respectively. The blue and the red curve of Fig.8 are the spectrum of Π_u for a frequency cut-off interval, $[-\nu_a, \nu_a]$ and $[-\nu_a/2, \nu_a/2]$ respectively. We noticed that when the cut-off interval is large, the MLW of the spectrum is narrower, the PSL is lower and the SLR drop faster.

[4] unless it happens that the signal fit exactly with the windowing function width. Nevertheless, it is rare to find such a situation.

4.4.2 *Gaussian window*

A Gaussian windowing function centred at mean zero with standard deviation, σ can be described as

$$G_u = e^{-b\nu^2}. \tag{16}$$

Here, $b = (2\sigma^2)^{-1}$ and $\mathcal{F}^{-1}\{G_u\} = \sqrt{\frac{b}{\pi}}e^{-cl^2}$, where $c = \pi^2/b$. This shows us that the inverse Fourier transform of a Gaussian with standard deviation σ is a Gaussian with a standard deviation $\sigma' = (2\pi\sigma)^{-1}$.

Fig.13 and Fig.14 gives the graph of G_u and its spectrum respectively, where G_u is truncate within the cut-off frequency interval $[-\nu_a, \nu_a]$, with $b = 3$ for the blue curve and $b = 5$ for the red one. We noticed that when the standard deviation is large, the MLW of the spectrum is narrower, the PSL is higher and the SLR drop slowly compare to a smaller standard deviation.

4.4.3 *Butterworth*

The frequency response of the Butterworth filter is flat in the pass band, and rolls off towards zero in the stop band, and it is characterized by two independent parameters, the cut-off frequency ν_a and the order p. These two parameters controls the bandwidth and side lobes attenuation. The frequency response of the Butterworth is given by

$$B_u = \Big(1 + (\nu/\nu_a)^{2p}\Big)^{-1}. \tag{17}$$

For the same frequency interval $[-\nu_a, \nu_a]$, we plotted three curves $\{p = 1, p = 3, p = 5\}$ of B_u in Fig.11 and their corresponding spectrum in Fig.12. We noticed that when the other p is getting bigger, the MLW of the spectrum is conserved, while the PSL is getting higher and the SLR is dropping faster.

4.4.4 *Bessel Function of the First Kind*

4.4.5 *First order prolate spheroidal wave function*

4.4.6 *Sinc window*

In theory, the sinc window is an infinitely large convolution filter as it is non zero everywhere, and its inverse Fourier

Transform produces the ideal filter kernel (the boxcar windowing function). The sinc is defined as follow:

$$S_u = sinc\big(\pi bu(t/\lambda_a)\big). \tag{18}$$

However, in practice some one need to defined a cut-off interval where the window is considered to be non zero, and zero out of this cut-off interval. Fig.9 and Fig.10 gives the graph of S_u and its spectrum respectively, where S_u is truncate within the cut-off frequency interval $[-\nu_a, \nu_a]$. We noticed that when the cut-off frequency interval is large (Fig.10, blue curve), the spectrum becomes perfectly flat at the passband while the MLW becomes narrower, the PSL becomes lower and the SLR drop faster compare to a cut-off frequency of $[-\nu_a, \nu_a]$ (Fig.10, red curve).

Windows	**MLL (-3db)**	**PSL (db)**	**SLR (db/octave)**
Π_u	$\approx 0,073$	$-6,68$	$-6,78$
$\mathcal{F}^{-1}\{\Pi_u\}$	$\approx 0,306$	$-11,22$	$-12,42$
G_u	$\approx 0,0736$	$-30,28$	$-14,5$
B_u	$\approx 0,079$	$-10,08$	$-15,39$
P_u	$\approx -$	$-$	$-$

A windowing function with a narrower main lobe width for a better spectral resolution, lower PSL to have less masking of nearby sources, and faster SLR to have less masking for far away sources is preferably in this work. Nevertheless, The energy is more concentrated in the frequency domain main lobe when the lobe width is narrower. Therefore, the frequency domain of the boxcar and the Butterworth window look similar, but the other of the Butterworth frequency domain can be control with the goal to concentrate more energy in the frequency main lobe. However, the sinc window is preferable as we expect in this work, signals in a wide dynamic range.

4.4.7 *Noise and comparison*

The methods described in section (3) and section (4) are use in this subsection to evaluate the theoretical noise, f_{pq} and g_{pq} are replaced in Eq.15 and Eq.15 respectively by the windowing functions described above. Table (4.4.7) and (4.4.7) summarize the theoretical noise where these functions are used as a baseline dependent windowing and as an overlap baseline dependent windowing functions respectively.

f_{pq}	**Theoretical noise**
Π_u	1,066
$\mathcal{F}^{-1}\{\Pi_u\}$	1,066
G_u	1,334
B_u	1,066
$\mathcal{F}^{-1}\{B_u\}$	1,066
P_u	1,334

g_{pq}	**Theoretical noise**
Π_u	1,066
$\mathcal{F}^{-1}\{\Pi_u\}$	1,066
G_u	1,334
B_u	1,066
$\mathcal{F}^{-1}\{B_u\}$	1,066
P_u	1,334

1.) In Figure 1.) In Figures (**??**) and (**??**) we represented the ratio $\frac{\sigma_{meas,pq}}{\sigma_{av,pq}}$ as a function of $\sum_i^n u_i((t_c - t_i)/\lambda))$ ($\sigma_{meas,pq}, \sigma_{av,pq}$ are the per baselines CWF and averaging rms noise respectively). We could also represented it as a function of baseline length. But, a uv-track corresponding to baseline aligned along the Est-West direction has longer tracts compare to one aligned along the South-Nordth for the same integration and frequency band. Therefore, $\frac{\sigma_{meas,pq}}{\sigma_{av,pq}}$ v/s baselines length is ambiguous. We consider five baselines dependent windowing sinc function, *Bl-sinc wk* (with an extended width of $(k-1)n$ time intervals and/or frequency channels and $k \geqslant 1$). The experiment is done for two cases, figure (**??**) is the one for *Bl-sinc wk* over both time-frequency and figure (**??**) over time. These figures shows that, the noise increases with baselines length.

2.) Figure (**??**) shows that, with *Bl-sinc wk*, the noise is lower on shorter baselines compared to the longer baselines. This is because on shorter baselines $\frac{\sigma_{meas,pq}}{\sigma_{av,pq}} \approx \frac{n}{n+k}$, it is the same case with Figure (**??**) where $\frac{\sigma_{meas,pq}}{\sigma_{av,pq}} \approx \sqrt{\frac{n}{n+k}}$ (see the proof). With *Bl-Sinc wk* ($k > 1$), the noise drops with the extended number of time intervals and/or frequency channels of the window. The rate of this drop is non-linear with baselines length and also with the overlap time interval and/or frequency channels (see figure (**??**b) and (**??**b), the variation of the noise rate between baselines).

3.) In spite of the overlapping, with the theoretical results, the noise of longer baselines do not drop when the overlap samples is increased (figure **??**). The reason is, few windows are overlapping in a visibility point when we extended in a unique direction (the time interval in this case) compare to two directions (time interval and frequency channels). The theoretical derivation for the overall noise of *Bl-sinc wk* and the simulated one are quantified similarly but the pattern of the per baseline simulated rms noise do not look the same with the theoretical one. The number of k determines the amount of noise of CWF.

5 SIMULATION AND RESULTS

In order to test the algorithms described in section and section, we performed multiple tests on JVLA simulated measurement sets (MS). In this section we summarize and discussed those results. Several windowing functions were described in the previous section, the sinc window and the Bessel of the first kind approaches well our specifications, and they are taken under consideration. Two MS is used in our simulation, a high resolution MS (HR-MS) that contained the observed JVLA data of short integration time and frequency, a low resolution MS (LR-MS), where the simple averaged or the weight averaged data are saved. To apply the second algorithm described in section 4, the following conditions are satisfied:

(i) If t^{hrms}_{start} and t^{lrms}_{start} are the starting time of the HR-MS and LR-MS respectively. $n_{t,ovlp}$ the number of timeslots extended across the time direction, and Δt^{hrms} the HR-MS integration time, then $t^{lrms}_{start} \geqslant t^{hrms}_{start} + n_{t,ovlp} \times \Delta t^{hrms}$.

(ii) If t^{hrms}_{end} and t^{lrms}_{end} are the ending time of the HR-MS and LR-MS respectively, then $t^{lrms}_{end} \leqslant t^{hrms}_{end} + n_{t,ovlp} \times \Delta t^{hrms}$.

(iii) If ν^{hrms}_{start} and ν^{lrms}_{start} are the starting frequency of the HR-MS and LR-MS respectively. $n_{\nu,ovlp}$ the number of chan-

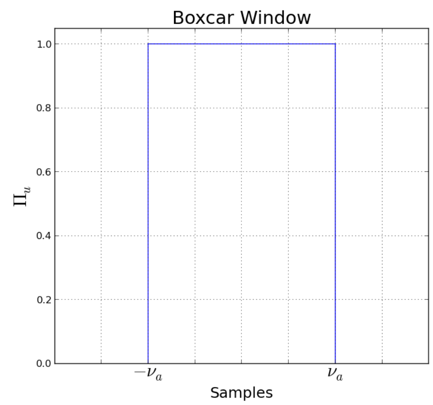

Figure 7. Boxcar windowing function.

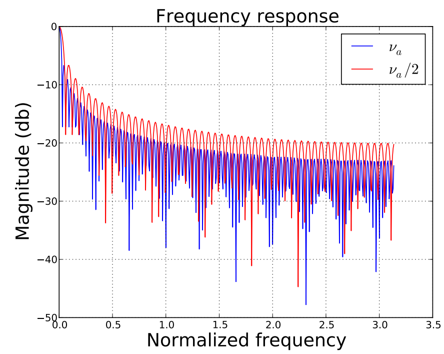

Figure 8. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

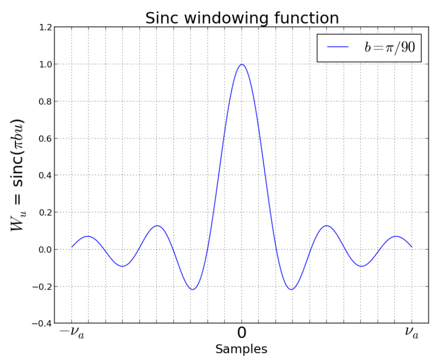

Figure 9. Overlap *BDWF's*: $\Delta_u t = \{250\}$.

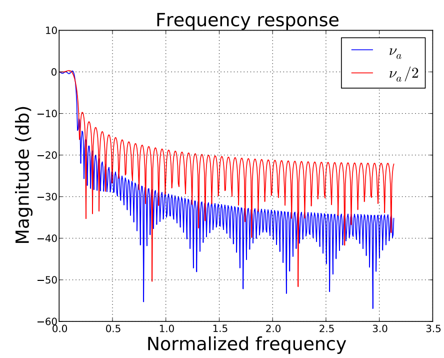

Figure 10. Overlap *BDWF's*: $\Delta_u t = \emptyset$.

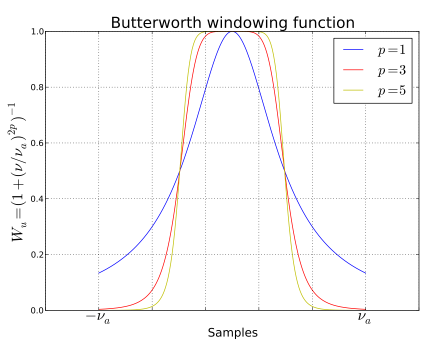

Figure 11. Overlap *BDWF's*: $\Delta_u t = \{250\}$.

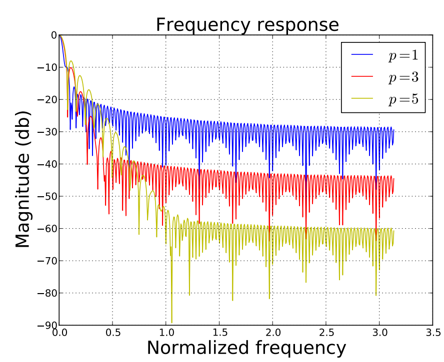

Figure 12. Overlap *BDWF's*: $\Delta_u t = \emptyset$.

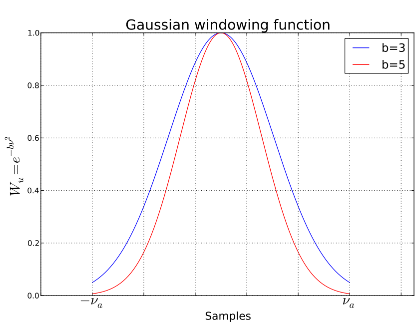

Figure 13. Overlap *BDWF's*: $\Delta_u t = \{250\}$.

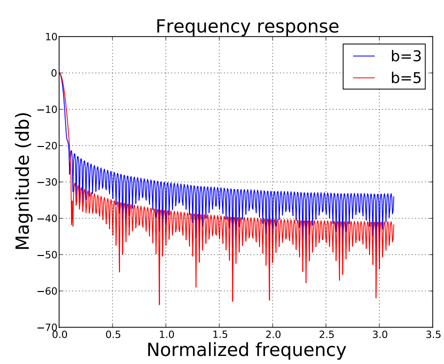

Figure 14. Overlap *BDWF's*: $\Delta_u t = \emptyset$.

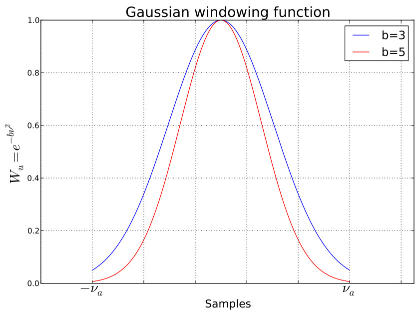

Figure 15. Overlap *BDWF's*: $\Delta_u t = \{250\}$.

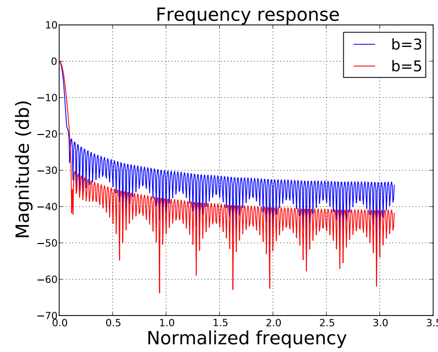

Figure 16. Overlap *BDWF's*: $\Delta_u t = \emptyset$.

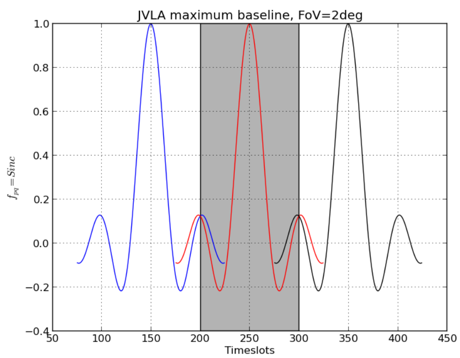

Figure 17. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

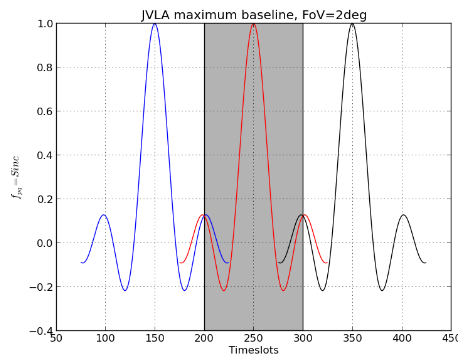

Figure 18. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

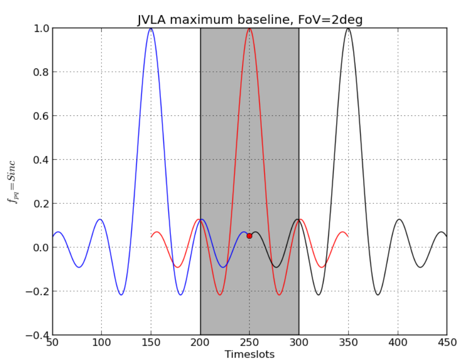

Figure 19. Overlap *BDWF's*: $\Delta_u t = \{250\}$.

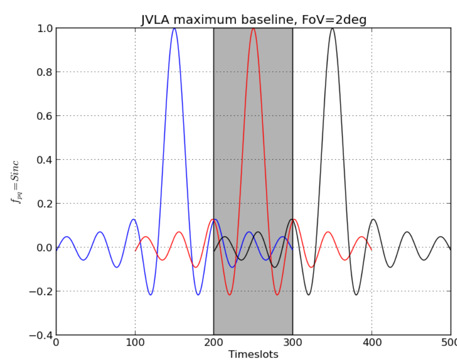

Figure 20. Overlap *BDWF's*: $\Delta_u t = \emptyset$.

nels extended across the frequency direction, and $\Delta\nu^{hrms}$ the HR-MS width, then $\nu^{lrms}_{start} \geqslant \nu^{hrms}_{start} + n_{\nu,ovlp} \times \Delta\nu^{hrms}$.

(iv) If ν^{hrms}_{end} and ν^{lrms}_{end} are the ending frequency of the HR-MS and LR-MS respectively, then $\nu^{lrms}_{end} \leqslant \nu^{hrms}_{end} + n_{\nu,ovlp} \times \Delta\nu^{hrms}$.

5.1 Smearing elimination and out FoV suppression

We varied the coordinates of a 1Jy source, and simulated a HR-MS of 7min30s snapshot synthesis, with a $\Delta t^{hrms} = 1.5s$ integration time at 1.4GHz, with 150 channels of width $\Delta\nu^{hrms}$=125kHz. The results of averaging and weighted average is save into a LR-MS of 1min30s synthesis, with a 150s integration time, with 1 channels of width 6.25MHz. We consider, $\{t^{hrms}_{start}, \nu^{hrms}_{start}\} = \{0s, 125kHz\}$, $\{t^{lrms}_{start}, \nu^{lrms}_{start}\} = \{1min30s, 6250kHz\}$, $\{t^{hrms}_{end}, \nu^{hrms}_{end}\} = \{7min30s,\text{xxxkHz}\}$, $\{t^{lrms}_{end}, \nu^{lrms}_{end}\} = \{6min, xxxkHz\}$, $n_{t,ovlp} = \{0, 100, 150\}$ and $n_{\nu,ovlp} = \{0, 25, 50\}$.

Fig.23, Fig.24, Fig.25 and Fig.26 show the results of the simulation, where we represented the brightness of the source as a function of the source coordinates in the sky. In these figures, we considered three cases of the sinc window, and three of the Bessel first kind (*Bl-sinc* $Wn_{t,ovlp} \times n_{\nu,ovlp}$, *Bl*-J$_0$ $Wn_{t,ovlp} \times n_{\nu,ovlp}$ with $n_{t,ovlp} = \{0, 100, 150\}$ and $n_{\nu,ovlp} = \{0, 25, 50\}$). We evaluated the loss in signal amplitude with longer LR-MS integration time interval

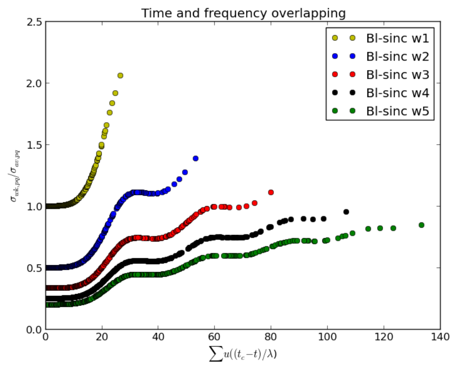

Figure 21. Noise ratio and rate of *Bl-sinc-wk*: time interval and frequency channels.

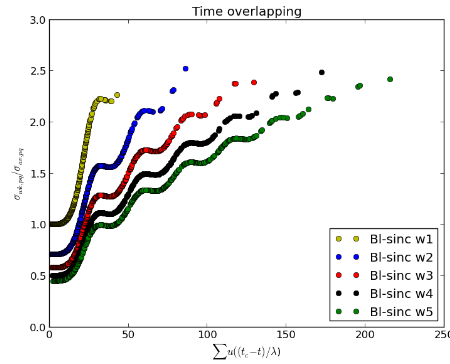

Figure 22. Noise ratio and rate of *Bl-sinc-wk*: time interval.

$\Delta t^{lrms} = 150s$ and wider LR-MS integration frequency interval $\Delta\nu^{lrms} = 6250kHz$. We furthermore evaluated the noise ratio, $\mathcal{R}_\sigma = \frac{\sigma_w}{\sigma_{Avg}}$ (with σ_w the resulting noise of the filter under consideration and σ_{Avg} the noise of simple averaging).
These results shows that $\approx$ 75% of the source brightness within the FoV is recovered with *Bl-sinc* $W0 \times 0$ (see Fig.23 and Fig.24) and with *Bl*-J$_0$ $W0 \times 0$ (Fig.25 and Fig.26). These results shows that:

- We measured more than 90% of the source brightness within $\approx$ 75% of the FoV using *Bl-sinc* $W0 \times 0$ and *Bl*-J$_0$ $W0 \times 0$. Unfortunately, the attenuation rate of the source out of the FoV is approximately the same when performing with simple averaging.
- We measured more than 90% of the source brightness within $\approx$ 95% of the FoV with the overlap filters. However, as mention in section 4.4 the reason for this is that, the overlap filter is an accurate practical evaluation of the window theoretical representation[5].
- Another obvious effect of the overlap filters is the suppression of the source when this source is out of the FoV. When looking at these figures, it appears that the curves of the overlaps filters are below the one of simple averaging. However, the overlaps filters attenuated $\approx$ 99% of the source brightness out of the FoV.
- When looking at these figures, it appears that the sinc window conserved the signal within the FoV compared to the Bessel of the first kind window, while the Bessel of the first kind suppressed the signal out of the FoV compare to the sinc. However, the reason for this is that the main lobe width of the sinc is narrower than the one of the Bessel first kind and the Bessel first kind has low sidelobes level compared to the sinc.

However, when performing the baseline dependent lters, smearing is eliminated within the FoV when we compressed the data by integrating over large time interval and wider bandwidth. When performing overlap baseline dependent windowing function, the source significantly disappears out of the FoV.

[5] windowing function theoretically extend to infinitely

5.2 Maximal integration time and frequency

We varied the coordinates of a 1Jy source, and simulated a HR-MS of 7min30s snapshot synthesis, with a $\Delta t^{hrms} = 1.5s$ integration time at 1.4GHz, with 150 channels of width $\Delta\nu^{hrms}$=125kHz. The results of averaging and weighted average is save into a LR-MS of 1min30s synthesis, with a 150s integration time, with 1 channels of width 6.25MHz. We consider, $\{t^{hrms}_{start}, \nu^{hrms}_{start}\} = \{0s, 125kHz\}$, $\{t^{lrms}_{start}, \nu^{lrms}_{start}\} = \{1min30s, 6250kHz\}$, $\{t^{hrms}_{end}, \nu^{hrms}_{end}\} = \{7min30s,$xxxkHz$\}$, $\{t^{lrms}_{end}, \nu^{lrms}_{end}\} = \{6min, xxxkHz\}$, $n_{t,ovlp} = \{0, 100, 150\}$ and $n_{\nu,ovlp} = \{0, 25, 50\}$.
Fig.23, Fig.24, Fig.25 and Fig.26 show the results of the simulation, where we represented the brightness of the source as a function of the source coordinates in the sky. In these figures, we considered three cases of the sinc window, and three of the Bessel first kind (*Bl-sinc* $Wn_{t,ovlp} \times n_{\nu,ovlp}$, *Bl*-J$_0$ $Wn_{t,ovlp} \times n_{\nu,ovlp}$ with $n_{t,ovlp} = \{0, 100, 150\}$ and $n_{\nu,ovlp} = \{0, 25, 50\}$). We evaluated the loss in signal amplitude with longer LR-MS integration time interval $\Delta t^{lrms} = 150s$ and wider LR-MS integration frequency interval $\Delta\nu^{lrms} = 6250kHz$. We furthermore evaluated the noise ratio, $\mathcal{R}_\sigma = \frac{\sigma_w}{\sigma_{Avg}}$ (with σ_w the resulting noise of the filter under consideration and σ_{Avg} the noise of simple averaging).
These results shows that $\approx$ 75% of the source brightness within the FoV is recovered with *Bl-sinc* $W0 \times 0$ (see Fig.23 and Fig.24) and with *Bl*-J$_0$ $W0 \times 0$ (Fig.25 and Fig.26). These results shows that:

- We measured more than 90% of the source brightness within $\approx$ 75% of the FoV with *Bl-sinc* $W0 \times 0$ and *Bl*-J$_0$ $W0 \times 0$.
- We measured more than 90% of the source brightness within $\approx$ 95% of the FoV with the overlap filters (*Bl-sinc* $Wn_{t,ovlp} \times n_{\nu,ovlp}$, *Bl*-J$_0$ $Wn_{t,ovlp} \times n_{\nu,ovlp}$).
- We loss $\approx$ 99% of the source brightness when the source is $\approx 1^\circ$ away from the FoV when using the overlap filters.
- The sinc window conserved the signal within the FoV compared to the Bessel of the first kind window, while the Bessel of the first kind suppressed the signal out of the FoV compare to the sinc.

This suggest that, smearing is eliminated within the FoV when we compressed the data by integrating over large time

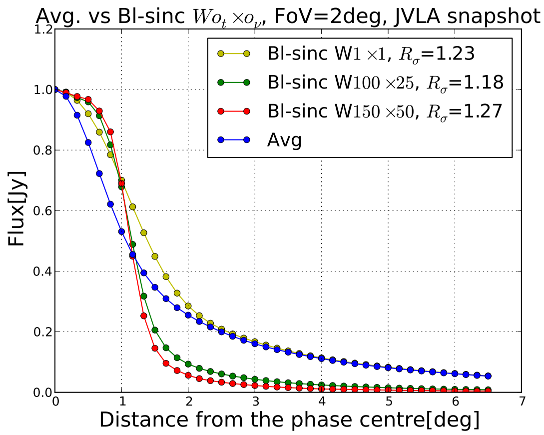

Figure 23. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

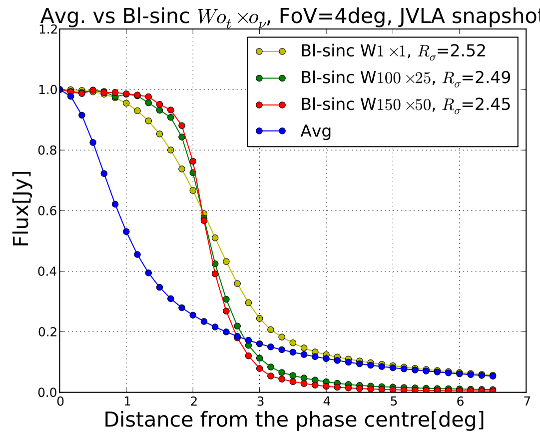

Figure 24. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

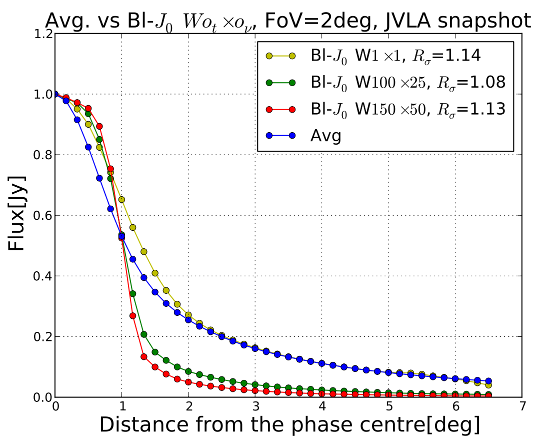

Figure 25. Overlap *BDWF's*: $\Delta_u t = \{250\}$.

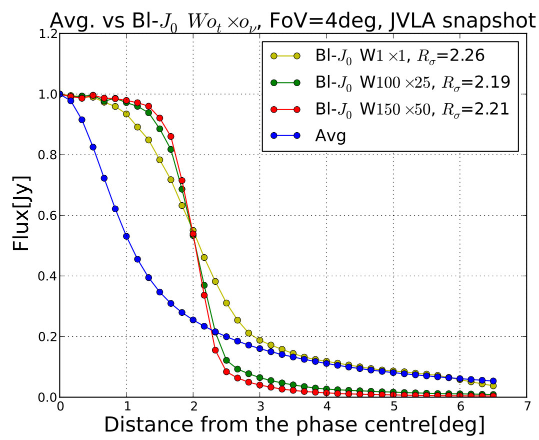

Figure 26. Overlap *BDWF's*: $\Delta_u t = \emptyset$.

interval and wider bandwidth and the FoV can therefore be extended according on the science goals.

5.3 Discussion

6 CONCLUSIONS

The goal of this paper was threefold. The first objective was to investigate **** windowing functions***
The second objective was to study ****first algorithm data compression***
The final objective was to ****second algorithm data compression and out field suppression***
Drawback and futures works*** drawback and futures works*****

ACKNOWLEDGEMENTS

REFERENCES

Linfield R., 1986, AJ, 92, 213
Martí-Vidal I., Marcaide J., 2008, A&A, 480, 289
Perlis S., 1952, Theory of matrices. Dover Publications

APPENDIX A: DEMONSTRATION

The proof of the norm used in Section 3.2 is given below.

if $||\mathbf{U}_{pq,t}||_m = \left(||\mathbf{u}_{pq,t_s}||, \dots, ||\mathbf{u}_{pq,t_c}||, \dots, ||\mathbf{u}_{pq,t_e}|| \right)$ is a $n_t \times 2$ matrix, where each element $\mathbf{u}_{pq,t_i}$ is a vector of size 2 and $||.||$ is an Euclidean norm, then $||.||_m$ is a norm.

Proof. (i) For all $\mathbf{u}_{pq,t_k}$, $||\mathbf{u}_{pq,t_k}|| \geqslant 0$. Therefore, all the elements of the vector $||\mathbf{U}_{pq,t}||_m$ are positives.
(ii) Let $\mathbf{U}_{pq,t}$ and $\mathbf{U}'_{pq,t}$ be two matrices with norm of $||\mathbf{U}_{pq,t}||_m = \left(||\mathbf{u}_{pq,t_s}||, \dots, ||\mathbf{u}_{pq,t_c}||, \dots, ||\mathbf{u}_{pq,t_e}|| \right)$ and norm of $||\mathbf{U}'_{pq,t}||_m = \left(||\mathbf{u}'_{pq,t_s}||, \dots, ||\mathbf{u}'_{pq,t_c}||, \dots, ||\mathbf{u}'_{pq,t_e}|| \right)$.
$||\mathbf{U}_{pq,t} + \mathbf{U}'_{pq,t}||_m = \left(||\mathbf{u}_{pq,t_s} + \mathbf{u}'_{pq,t_s}||, \dots, ||\mathbf{u}_{pq,t_c} + \mathbf{u}'_{pq,t_c}||, \dots, ||\mathbf{u}_{pq,t_e} + \mathbf{u}'_{pq,t_e}|| \right)$. For all k, $||\mathbf{u}_{pq,t_k} + \mathbf{u}'_{pq,t_k}|| \leqslant ||\mathbf{u}_{pq,t_k}|| + ||\mathbf{u}'_{pq,t_k}||$, this is trivial because $||.||$ is the Euclidean norm. Therefore, $||\mathbf{U}_{pq,t} + \mathbf{U}'_{pq,t}||_m \leqslant ||\mathbf{U}_{pq,t}||_m + ||\mathbf{U}'_{pq,t}||_m$

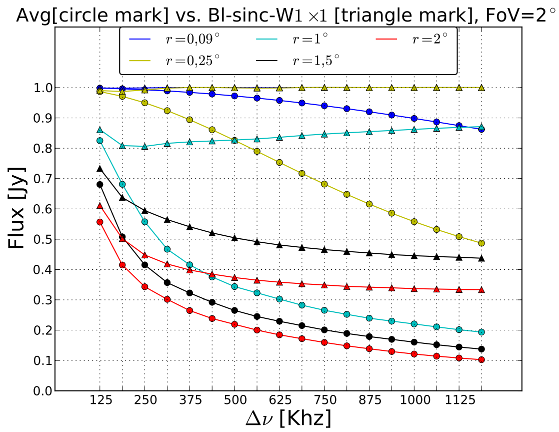

Figure 27. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

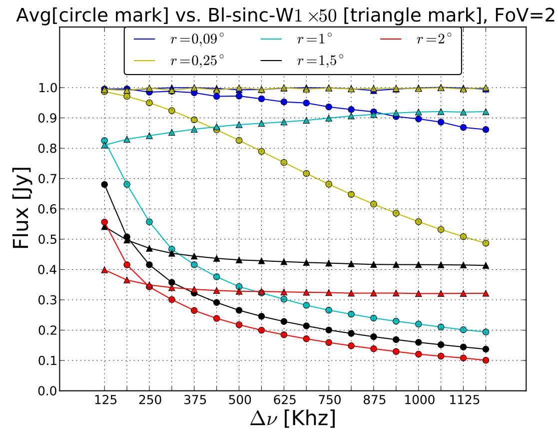

Figure 28. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

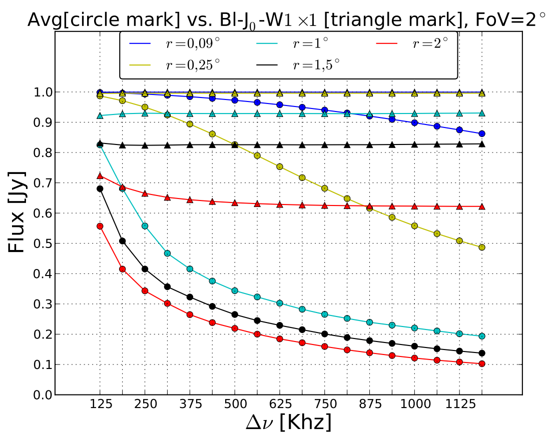

Figure 29. Overlap *BDWF's*: $\Delta_u t = \{250\}$.

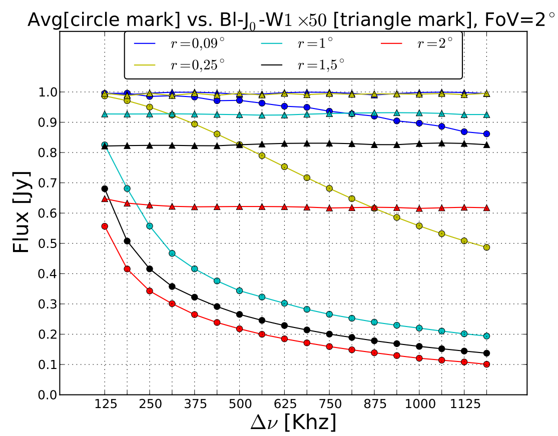

Figure 30. Overlap *BDWF's*: $\Delta_u t = \emptyset$.

□

APPENDIX B: DERIVATION OF COMPLEX MATRICES

The matrix $\mathbf{W}_{pq,(t,\nu)}$ is of size $n_t \times n_\nu$ and contained the weights of the baseline (p,q) visibilities points. The convolution operator is linear, therefore we can rewrite Eq.7 in terms of a series of linear transformations as follow:

$$V^{corr}_{pq} = \mathbf{C}^{block}_{(t,\nu)} \cdot \mathbf{W}^{block}_{pq,(t,\nu)} \cdot \mathbf{V}^{samp}_{pq,(t,\nu)}. \tag{B1}$$

Here, $\mathbf{W}^{block}_{pq,(t,\nu)}$ is a block diagonal matrix of size $(4n_t n_\nu) \times (4n_t n_\nu)$ defined as follow:

$$\mathbf{C}^{block}_{(t,\nu)} = \begin{bmatrix} \mathbf{c}_{(t,\nu)} & 0 & 0 & 0 \\ 0 & \mathbf{c}_{(t,\nu)} & 0 & 0 \\ 0 & 0 & \mathbf{c}_{(t,\nu)} & 0 \\ 0 & 0 & 0 & \mathbf{c}_{(t,\nu)} \end{bmatrix}$$

Here, $\mathbf{W}^{block}_{pq,(t,\nu)}$ is a block diagonal matrix of size $(4n_t n_\nu) \times (4n_t n_\nu)$ defined as follow:

$$\mathbf{W}^{block}_{pq,(t,\nu)} = \begin{bmatrix} \mathbf{W}_{pq,(t,\nu)} & 0 & 0 & 0 \\ 0 & \mathbf{W}_{pq,(t,\nu)} & 0 & 0 \\ 0 & 0 & \mathbf{W}_{pq,(t,\nu)} & 0 \\ 0 & 0 & 0 & \mathbf{W}_{pq,(t,\nu)} \end{bmatrix}$$

$$\mathbf{V}^{samp}_{pq,(t,\nu)} = \left[\mathcal{V}^0_{pq,(t,\nu)}, \mathcal{V}^1_{pq,(t,\nu)}, \mathcal{V}^2_{pq,(t,\nu)}, \mathcal{V}^3_{pq,(t,\nu)} \right]^T.$$

For a synthesis, the baseline (p,q) made a full coverage in the (u,v) plane. Therefore, we can package into a single matrix, $\mathbf{V}^{avg}_{pq,}$ of size $(4N_t N_\nu) \times (4N_t N_\nu)$ the weighted average visibilities of the baseline (p,q) during the synthesis as follows:

$$\mathbf{V}^{corr}_{pq} = \mathbf{C}^{block,n}_{(t,\nu)} \cdot \mathbf{W}^{block,n}_{pq,(t,\nu)} \cdot \mathbf{V}^{samp,n}_{pq,(t,\nu)}. \tag{B2}$$

where N_t and N_ν are the number of time sample and frequency channels entering the Fourier domain. If the synthesis time is T and the frequency range is F, then $T = N_t \times \Delta t$ and $F = N_\nu \times \Delta\nu$. The the size of $\mathbf{V}^{avg}_{pq}$ can also be written as $(4N^{pq}_v) \times (4N^{pq}_v)$, where N^{pq}_v is the number of time

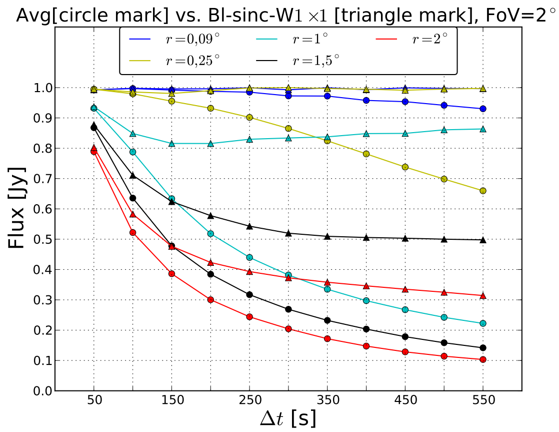

Figure 31. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

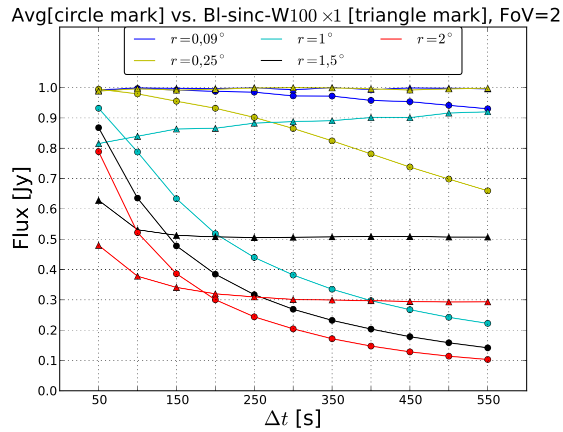

Figure 32. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

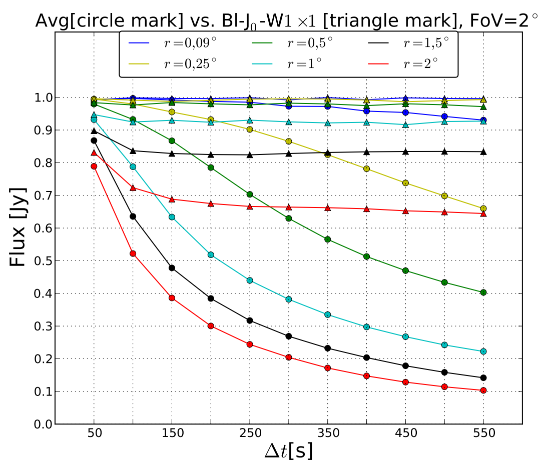

Figure 33. Overlap *BDWF's*: $\Delta_u t = \{250\}$.

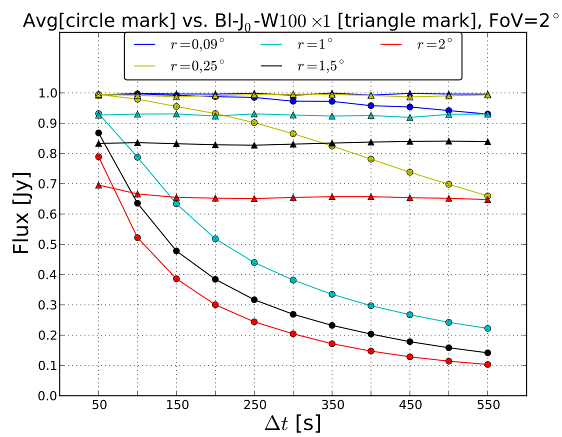

Figure 34. Overlap *BDWF's*: $\Delta_u t = \emptyset$.

and frequency visibilities for the baseline (p,q)). The matrix $\mathbf{W}^{block,n}_{pq,(t,\nu)}$ is a diagonal block matrix of size $(4N_v^{pq} n_t n_\nu) \times (4N_v^{pq} n_t n_\nu)$ defined as follow:

$$\mathbf{W}^{block,n}_{pq,(t,\nu)} = \begin{bmatrix} \mathbf{W}^{block}_{pq,(t,\nu)} & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \mathbf{W}^{block}_{pq,(t,\nu)} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & \mathbf{W}^{block}_{pq,(t,\nu)} \end{bmatrix}$$

$$\mathbf{C}^{block,n}_{(t,\nu)} = \begin{bmatrix} \mathbf{C}^{block}_{(t,\nu)} & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & \mathbf{C}^{block}_{(t,\nu)} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \dots & \mathbf{C}^{block}_{(t,\nu)} \end{bmatrix}$$

The sampled visibilities $\mathbf{V}^{samp,n}_{pq,(t,\nu)}$ is a one row matrix of size $(N_v^{pq} 4 n_t n_\nu) \times (4 n_t n_\nu)$ made of $\mathbf{V}^{samp}_{pq,(t,\nu)}$ on top of each other.

$$\mathbf{V}^{samp,n}_{pq,(t,\nu)} = \left[\mathbf{V}^{samp}_{pq,(t,\nu)}, \dots, \mathbf{V}^{samp}_{pq,(t,\nu)}, \dots, \mathbf{V}^{samp}_{pq,(t,\nu)} \right]^T.$$

$$\mathbf{V}^{avg}_{pq} = \mathbf{W}^{block,n}_{pq,(t,\nu)} \cdot \mathbf{S}^n_{pq,(t,\nu)} \mathbf{F} \cdot \mathcal{I}^{sky}_{l,m}, \tag{B3}$$

if the number of pixel in the sky model is N_{pix}, then the true sky image vector $\mathcal{I}^{sky}_{l,m}$ has a size of $4N_{pix}$ and $\mathbf{F}$ is the Fourier transform operator of size $(4N_{pix}) \times (4N_{pix})$.

We are generally interested in using the total set of visibilities over baselines, time and frequencies, having $4 \times N_v$ visibilities measured over all baselines and $N_v = n_{bl} \times N_v^{pq}$ in this case. Here, n_{bl} is the number of baseline. We can write

$$\mathbf{V}^{avg}_{all} = \mathbf{A} \cdot \mathcal{I}^{sky}_{l,m}. \tag{B4}$$

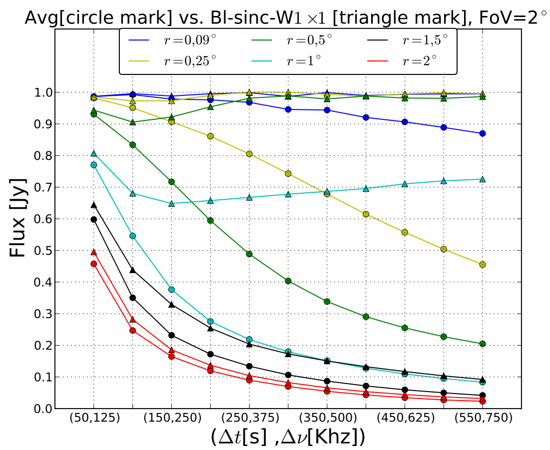

Figure 35. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

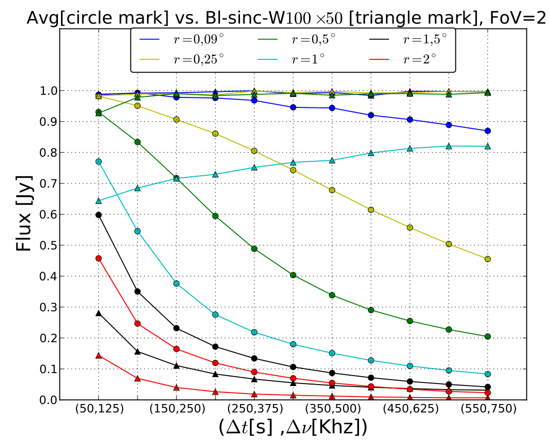

Figure 36. Overlap *BDWF's*: $\Delta_u t = [225, 250]$.

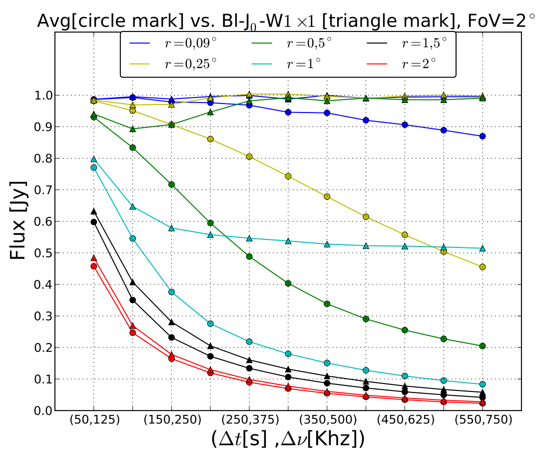

Figure 37. Overlap *BDWF's*: $\Delta_u t = \{250\}$.

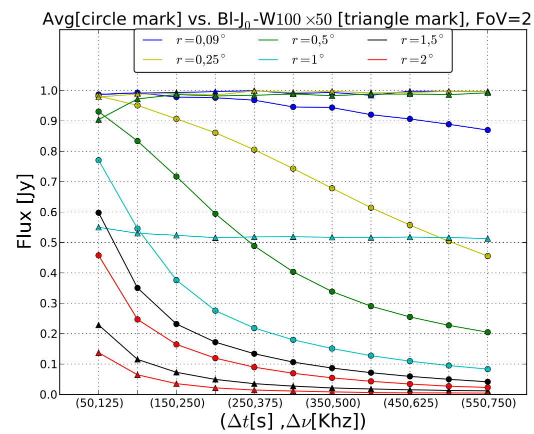

Figure 38. Overlap *BDWF's*: $\Delta_u t = \emptyset$.

Here, $\mathbf{A}$ is a matrix of size $(4N_v) \times (4N_{pix})$ defined as follow

$$\mathbf{A}_{l,m} = \begin{bmatrix} \mathbf{C}^{block,n}_{(t,\nu)} \cdot \mathbf{W}^{block,n}_{01,(t,\nu)} \cdot \mathbf{S}^{block,n}_{01,(t,\nu)} \cdot \mathbf{F} \\ \vdots \\ \mathbf{C}^{block,n}_{(t,\nu)} \cdot \mathbf{W}^{block,n}_{ik,(t,\nu)} \cdot \mathbf{S}^{block,n}_{ik,(t,\nu)} \cdot \mathbf{F} \\ \vdots \\ \mathbf{C}^{block,n}_{(t,\nu)} \cdot \mathbf{W}^{block,n}_{jl,(t,\nu)} \cdot \mathbf{S}^{block,n}_{jl,(t,\nu)} \cdot \mathbf{F} \end{bmatrix}$$

The dirty image, $\mathcal{I}^D_{l,m}$ of size $4N_{pix}$ can then be derived

$$\mathcal{I}^D_{l,m} = \mathbf{F}^H \cdot \mathbf{A} \cdot \mathcal{I}^{sky}_{l,m}. \tag{B5}$$

Here, H represents the the conjugate transpose operation also known as a Hermitian transpose and $\mathbf{F}^H$ is the inverse Fourier transform operator of size $(4N_{pix}) \times (4N_{pix})$.

APPENDIX C: SIMILAR WAY OF IMAGING

We explained how we derived the complex matrices used in section 2.4 Each baseline has his own baseline dependent windowing functions during integration. Unfortunately, we can not accurately estimate the resulting spectrum of all windowing functions that multiply the sky seen by these baselines (if we supposed that all baselines are seen the same sky). The measured sky intensity of the array is derived from the inverse Fourier transform of the sum of the sample visibilities measured at each baseline. The mathematics behind this is as follow:

$$\mathcal{I}^D_{l,m} \quad = \quad \mathcal{F}^{-1}\left\{ \sum_{pq} \mathcal{S}_{pq} \cdot \Big(c_{pq} \cdot \left(\mathcal{W}_{pq} \circ \mathcal{V} \right) \Big)_{(t,\nu)} \right\}$$

where $\mathcal{S}_{pq}$ is the sampling function of the baseline pq. This can be rewritten as

$$\mathcal{I}^D_{l,m} \quad = \quad \sum_{pq} \mathcal{B}_{pq} \circ \Big(\mathcal{F}^{-1}\{c\}_{pq} \circ \left(\mathcal{R}_{pq} \cdot \mathcal{I}^{sky} \right) \Big)_{(l,m)}$$

Here, $\mathcal{R}_{pq} = \mathcal{W}_{pq}$ is the sky response or smearing response for the baseline pq. The same sky seen by all baselines is $\mathcal{I}^{sky}$ and $\mathcal{B}_{pq}$ is the synthesized beam or point spread function of the baseline pq. $\Big(\mathcal{F}^{-1}\{c\}_{pq} \circ \left(\mathcal{R}_{pq} \cdot \mathcal{I}^{sky} \right) \Big)_{(l,m)} \quad = \quad (\mathcal{R}_{pq} \cdot$

$\mathcal{I}^{sky}\big)_{(l,m)}$. Therefore, the dirty beam can be written as:

$$\mathcal{I}^{D}_{l,m} = \left(\sum_{pq}\mathcal{B}_{pq}\circ\left(\mathcal{R}_{pq}\cdot\mathcal{I}\right)\right)_{(l,m)}$$

If the integration windowing function were the same in all baseline, then we can write:

$$\mathcal{I}^{D}_{l,m} = \left(\left(\sum_{pq}\mathcal{B}_{pq}\right)\circ\left(\mathcal{R}\cdot\mathcal{I}\right)\right)_{(l,m)}$$

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