Mon. Not. R. Astron. Soc. **000**, 1–9 (2016) Printed 5 March 2016 (MN LATEX style file v2.2)

Using Baseline Dependent Averaging and Baseline Depedent Window Functions Applied Across Equal uv-Distance for Data Compression and field-of-view shaping in radio interferometry

M.T. Atemkeng[1⋆], O.M. Smirnov[12], C. Tasse[31], G. Foster[12], 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*

in original form 2016 March 3

ABSTRACT

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

1 INTRODUCTION

In the previous chapter, we looked at the various ways in which averaging larger visibilities bins intervals can result in acceptable level of smearing i.e 5% decrease in source amplitude or less within the observation field of interest. We showed, theorically, that smearing increases on longer baselines compared to the shorter baselines, and that smearing can even be avoided, if the correlator undergo the averaging procedure withing shorter visibilities bins intervals, which result in high data rate. We made certain predictions pertaining to an equal uv-distance averaging across all baselines; in particular, averaging within a sufficiently large bins intervals for shorter baselines, while on the other hand, the longer baselines are averaged within shorter bins intervals i.e for all baselines, the visibilities are averaged across an equal uv-distance. The first question therefore arises as to whether such averaging technique (averaging over equal uv-distance on all baselines) will not only decrease smearing within the observation field of interest, but importantly, also result in reducing the data size. Indeed, one expects that averaging for a sufficiently large bins intervals on the shorter baselines, would be favorable for data compression, while averaging tiny intervals in uv-distance on the longer baselines, will result in decreasing smearing. The second question lies on the calibration issues for BDA given that calibration is a complex visibilities correction process and therefore not only for setting the scale of the overall absolute flux density. The key point for the calibration properties of BDA will be to monitor the correlator differently at each averaging intervals on different baselines. What implies that the calibration parameters will changed differently with baselines and each of the frequency and/or time intervals.

In this chapter, we show that at least BDA results in decreasing smearing over a selected field of interest, and applying BDA to BDWFs, appears to be satisfactory in a decrease in smearing ($\leqslant 1\%$ across the field of interest) and source suppression compared to the results presented in chapter **??** (section **??**). To arrive on our results, we first describe the mathematics behind BDA and the compression factor on each baseline. We identify the implementation criterias and describe three different algorithms for the implementation a BDA correlator. We implement BDA via simulations using the JVLA telescope, and discuss the results.

2 DATA STRUCTURE AND DIMENSIONS

I present a schematic description on how an interferometer array overall data are mapped into a correlation matrix. Note that this mapping is an abstraction on how the correlation is saved into computer memory. I shall come back to this in chapter **??**. In level $L1$ of Figure 1, the vertical axis represents the data for baselines, and the horizontal axis are sub-bands (spectral windows). Each sub-band on its own is a sub-dataset projected in level $L2$, where the vertical axis is the number of integration time (timeslots) and the horizontal axis is the number of channels. Each cross section (level $L3$) of times and channels is the visibility presented in Eq. **??**. In general, Eq. **??** represents a complex scalar, or 2×2 complex visibility. In terms of data storage, we restrict ourselves to 32 bits (complex data) per polarization, if N_{pol} is the number polarizations, then we have $N_{pol} \times 32$ bits per visibility. However, an interferometer of $N \times (N-1)/2$ baselines, N_{sub} sub-bands per baseline observing during a total

⋆ E-mail: m.atemkeng@gmail.com

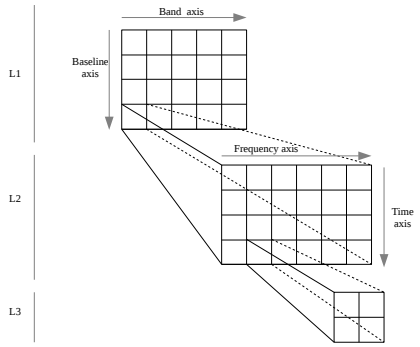

Figure 1. Interferometry data structure. A baseline is a set of sub-bands, each sub-band on it own is a set of data for timeslots and channels. The data shape for a baseline result to (N_{sub}, N_{tslot}, N_{ch}, N_{pol}).

N	N_{sub}	T_{obs}	Δt	BWD	$\Delta\nu$	$\sim$ **Size**
7	5	16	20	10^2	10^6	**0.4**
27	3	4	5	10^2	125000	**60**
64	5	4	10	10^2	10^6	**22**
1000	3	4	0.5	10^2	11000	**5846240**

Table 1. Interferometers observing parameters vs. data size

time of T_{obs}, with integration time Δt, and total bandwidth of B_w, having $\Delta\nu$ of channel width will record in memory

$$data\ size = \frac{N\times(N-1)}{2}\times N_{sub}\times\frac{T_{obs}}{\Delta t}\times\frac{B_w}{\Delta\nu}\times N_{pol}\times 32\ bits, \quad (1)$$

where, T_{obs} and Δt are in seconds, B_w and $\Delta\nu$ are measured in Hz. One can rewrite Eq. 1 as

$$data\ size = \frac{N\times(N-1)}{2}\times N_{sub}\times N_{tslot}\times N_{ch}\times N_{pol}\times 32\ bits, \quad (2)$$

where $N_{tslot} = T_{obs}/\Delta t$ and $N_{ch} = B_w/\Delta\nu$ are the number of timeslots and channels respectively. Observing for long time (above 4 hours) and large bandwidth (above 100 MHz) leads to storage issues for big interferometer array (more than 100 antennas), as well as computation load for subprocessing, this is because $\{T_{obs}, B_w\} \propto \{N_{tslot}, N_{ch}\}$ if Δt and $\Delta\nu$ remain sufficiently small (less than 10 ms and 10 kHz respectively). Table 1 is obtained from Eq. (2) and showing clearly that any increase in number of antennas, sub-bands, timeslots and channels, result in linear increased in data size and pose significant computational challenges for processing. The case of 1000 antennas is close to what is expected from the SKA (**???**), with these observing parameters, approximately 5709 Tb data has to be recorded, which is already above any modern hard disk (**?**). Since the number of baselines and sub-bands will remain constant for individual interferometers, the only way possible to reduce the data rate is using longer integration time and wider channels widths.

3 PROBLEM STATEMENT AND BACKGROUND

The effect of time and bandwidth averaging is severe on longer baselines compared to the shorter baselines. This is easily understand as follows: the visibility from a baseline pq of a point source with brightness S is given by

$$V_{pq} = S\exp\{i\phi\},\ \phi = 2\pi\boldsymbol{u}_{pq}.\boldsymbol{l}, \quad (3)$$

where ϕ is the phase of the source, $\boldsymbol{u}_{pq} = (u, v, w)$ the baseline vector in unit of wavelength and $\boldsymbol{l} = (l, m, n-1)$ the source location in the sky. This implies that the larger the norm $||\boldsymbol{u}_{pq}||$ i.e the baseline length, the larger the phase and thus a much more severe attenuation of the source brightness S. Sources far away from the pointing direction have larger radius $||\boldsymbol{l}||$, thus larger phase as well i.e the fringes of sources far from the pointing direction rotates rapidly in the uv domain.

Figure 2 is a simulated observation at 1.4 Ghz of the JVLA in C configuration showing the amplitude of a sources as seen by three baselines (long baseline, medium length baseline and short baseline). In this figure, the amplitude of the source is plotted against the source distance from the phase centre. The left panel shows the amplitude of the source at low temporal resolution of 100s integration and at high spectral resolution of 125 kHz. The right panel shows the amplitude of the same source at high temporal resolution of 1 s intergration and at low spectral resolution of 10 MHz.

Figure 3 shows the amplitude loss as a function of east-west baseline length. We simulate theBDA source at 30 arcmin centered at the phase centre and observed the amplitude loss using the JVLA in C configuration. In this figure, the blue vertical line shows the level of the shortest baseline. In the top panel, the source is simulated using 100 s integration and 125 kHz channels width, while in the bottom panel the source is simulated using 1 s integration and 10 MHz channels width.

We observed from Figure 2 and Figure 3 that, smearing is severe on longer baselines compared to shorter baselines. In addition, the plot shows that smearing is also a function of source position in the sky.

Most of the data measured by any compact interferometer array are from the short baselines. A good illustration is presented in Figure 4, which is a 1 hour synthesis uv-coverage of the MeerKAT observing at 1.4 GHz. The coverage shows that the data points are more populated in the inner core compared to the outer core. The data in the inner core are from the short baselines, while the data in the outer core are from the long baselines. What if we averaged more samples in the inner core and very little at the outer core? By doing this, we avoid smearing on the longer baselines and compress more data on the shorter baselines. This method often referred as *Baseline Dependent Averaging* was first proposed by **??** as an approach for tackling wide field imaging with little to no bandwidth and time averaging effects. The next section describes the mathematical details of BDA.

3.1 Baseline Dependent Averaging

Recall from chapter **??** that an interferometer measures the average visibility over a rectangular time-frequency bin given by the time and frequency sampling intervals Δt and $\Delta\nu$ respectively, what we called the resampling bin (see chapter **??**, Eq. **??**)

$$\mathsf{B}_{kl}^{[\Delta t\Delta\nu]} = \left[t_k - \frac{\Delta t}{2}, t_k + \frac{\Delta t}{2}\right]\times\left[\nu_l - \frac{\Delta\nu}{2}, \nu_l + \frac{\Delta\nu}{2}\right], \quad (4)$$

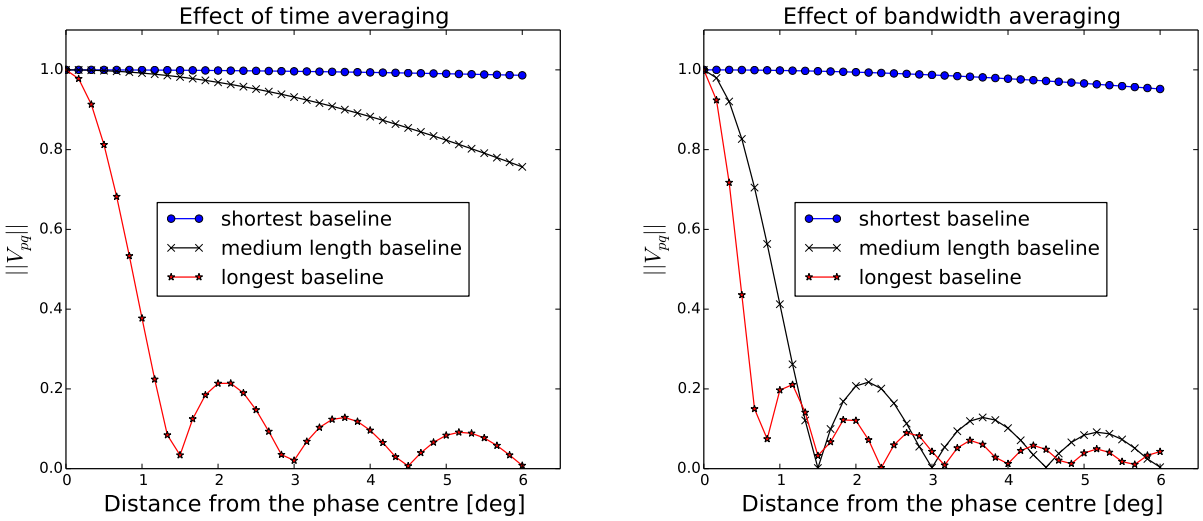

Figure 2. Amplitude loss: the apparent intensity of a 1 Jy source, as seen by JVLA-C at 1.4 GHz, as a function of distance from phase centre. (Left) time and frequency integrations fixed at 100s and 125 kHz respectively; (Right) time and frequency integrations fixed at 1 s and 10 MHz respectively.

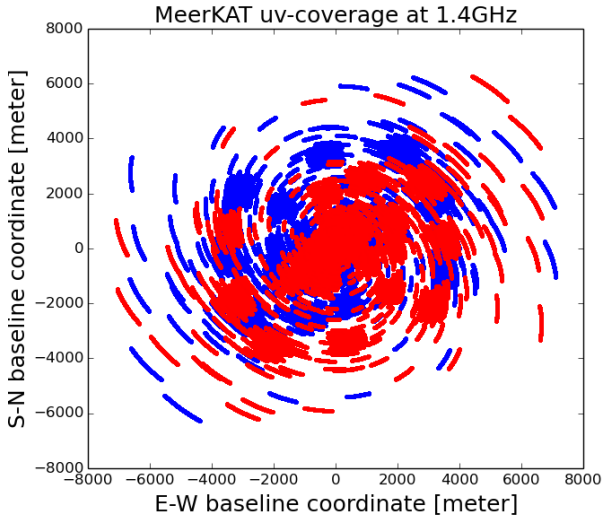

Figure 4. MeerKAT uv-coverage at 1.4 GHz, 1 hour observation showing clearly that the data is most condensed at the centre core. The data at the inner core are from the shorter baselines, while the data at the outer core are from the longer baselines.

where t_k and ν_l are the centre of Δt and $\Delta\nu$ respectively.

Let us now define the resampling bin for the case of BDA, where the sampling intervals Δt, $\Delta\nu$ varies across baselines. Suppose that the sampling intervals for a baseline pq are $\Delta_{pq}t$, $\Delta_{pq}\nu$. We now defined the sampling bin in time and frequency for a baseline pq as follows

$$\mathsf{B}_{kl}^{[\Delta_{pq}t,\Delta_{pq}\nu]} = \left[t_k - \frac{\Delta_{pq}t}{2}, t_k + \frac{\Delta_{pq}t}{2}\right] \times \left[\nu_l - \frac{\Delta_{pq}\nu}{2}, \nu_l + \frac{\Delta_{pq}\nu}{2}\right] \tag{5}$$

The resampling bins, $\mathsf{B}_{kl}^{[\Delta_{pq}t,\Delta_{pq}\nu]}$ is a set of time and frequency bins that are part of the uv-track draw by the baseline pq during the integrations $\Delta_{pq}t$ and $\Delta_{pq}\nu$. Figure 5 shows a typical resampling bin for BDA and the resampling bin for simple averaging.

The set of samples B_{kl}^{pq}, indexes ij corresponding to the baseline pq *resampling bin*, is given as

$$\mathsf{B}_{kl}^{pq} = \{ij : \ t_i\nu_j \in \mathsf{B}_{kl}^{[\Delta_{pq}t,\Delta_{pq}\nu]}\}, \tag{6}$$

Note from this that, if $dim\{\mathsf{B}_{kl}^{pq}\}$, is the number of samples in the resampling bin then the baseline sweeps the distance $\mathsf{D}^{[\Delta_{pq}t,\Delta_{pq}\nu]}$ within this sampling bin

$$\mathsf{D}^{[\Delta_{pq}t,\Delta_{pq}\nu]} = \sum_{ij\in\mathsf{B}_{kl}^{pq}} ||\boldsymbol{u}_{pq}(t_i - t_k, \nu_j - \nu_l)||. \tag{7}$$

For simplicity, we will now refer a baseline from its east-west component. Indeed, all baselines have an east-west and south-north components. During a time/frequency integration, an east-west components has high rotation speed compared to the south-north component. The different speeds rate for baselines components indicates that, for the same time/frequency integrations, the uv-distance sweeps by two equal lengths baselines oriented differently, may not be equal. This is illustrated in Figure 6.

Eq. (8) and (15) must be satisfied for BDA: for all East-West baselines $\alpha\beta \neq pq$ with $||\boldsymbol{u}_{\alpha\beta}|| \neq ||\boldsymbol{u}_{pq}||$:

$$\mathsf{D}^{[\Delta_{\alpha\beta}t,\Delta_{\alpha\beta}\nu]} = \mathsf{D}^{[\Delta_{pq}t,\Delta_{pq}\nu]} \tag{8}$$

$$dim\{\mathsf{B}_{kl}^{[\Delta_{\alpha\beta}t,\Delta_{\alpha\beta}\nu]}\} \neq dim\{\mathsf{B}_{kl}^{[\Delta_{pq}t,\Delta_{pq}\nu]}\}.labeleq : bda2 \tag{9}$$

The above conditions show that, more samples are averaged on the shorter baselines compared to the longer baselines. For example, suppose $p = 1$, $q = 2$ are the indexes for the longest baseline and $p = 2$, $q = 3$ the indexes for the medium length baseline and $p = 3$, $q = 4$ the indexes for the shortest baseline. If these three baselines are East-West baselines or East-West components, we have

$$\mathsf{D}^{[\Delta_{12}t,\Delta_{12}\nu]} = \mathsf{D}^{[\Delta_{23}t,\Delta_{23}\nu]} = \mathsf{D}^{[\Delta_{34}t,\Delta_{34}\nu]} dim\{\mathsf{B}_{kl}^{[\Delta_{12}t,\Delta_{12}\nu]}\} < dim\{\mathsf{B}_{kl}^{[\Delta_2} \tag{10}$$

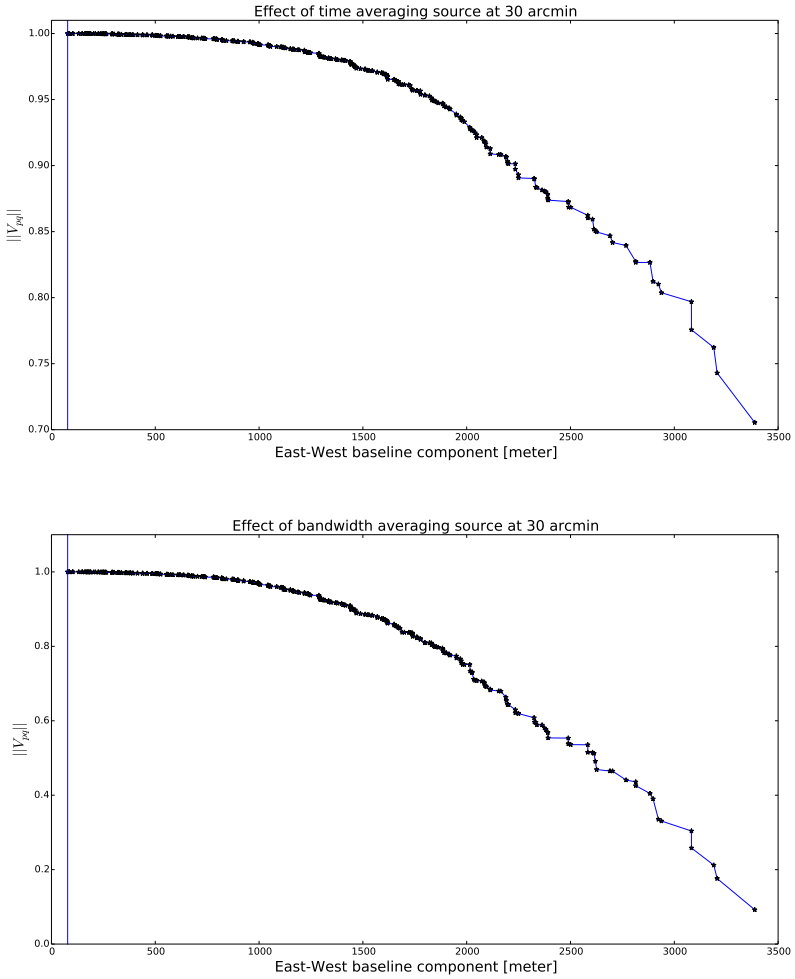

Figure 3. Amplitude loss: the apparent intensity of a 1 Jy source, as seen by JVLA-C at 1.4 GHz, as a function of Est-West baseline components. (Top) time and frequency integrations fixed at 100s and 125 kHz respectively; (Right) time and frequency integrations fixed at 1 s and 10 MHz respectively.

The averaged intervals are function of baselines length and direction, with longer baselines having shorter integration time and narrower channels width, while on the other hand, shorter baselines are averaged over longer integration time and wider channels widths.

Recall that in chapter **??** (section **??**), we defined the averaged or *resampled* visibilities for a baseline pq as the discrete sum

$$V^{(\mathrm{m})}_{pqkl} = \frac{1}{n_{pq}} \sum_{ij \in \mathsf{B}^{pq}_{kl}} V^{\mathrm{S}}_{pqij}, \tag{11}$$

where $dim\{\mathsf{B}^{pq}_{kl}\} = n_{pq}$ is the number of samples that has been averaged. It was then shown that averaging is a convolution at the centre of the resampling bin interval

$$V^{(\mathrm{m})}_{pqkl} = [\mathcal{V}_{pq} \circ \Pi^{[uv]}_{pq}](\boldsymbol{u}_{pq}(t_k, \nu_l)), \tag{12}$$

and this results to Eq. (13), when imaging the per-baseline visibilities:

$$\mathcal{I}^{\mathrm{D}} = \sum_{pqkl} W_{pqkl} \mathcal{P}_{pqkl} \circ (\mathcal{I} \cdot \mathcal{T}_{pq}), \quad \mathcal{T}_{pq} = \mathcal{F}^{-1}\{\Pi^{[uv]}_{pq}\}. \tag{13}$$

Now let us see what Eq. (13) becomes with BDA. For BDA, the uv-space boxcar windows $\Pi^{[uv]}_{pq}$, are approximately equal across all baselines and therefore, not any more a function of baseline length and/or direction. In other words, for any baseline $\alpha\beta \neq pq$, $\Pi^{[uv]}_{\alpha\beta} \approx \Pi^{[uv]}_{pq}$. This derivation, $\Pi^{[uv]}_{\alpha\beta} \approx \Pi^{[uv]}_{pq}$ is valid and correct, but one thing to keep in mind. Although the size of all boxcar windows are the same in uv-space for BDA, but they are sampled differently. The boxcar window are well sampled on the shorter baselines compared to the longer baselines, and which may result in different image plane taper.

However, if the pre-averaged visibilities are sampled at significantly high time and spectral resolution, then we can assume that all these boxcar at different baseline are sampled equaly. Taking this assumption into account, we can write

$$\mathcal{T}_{pq} \approx \mathcal{T}_{\alpha\beta}, \tag{14}$$

which shows that, the time and bandwidth smearing represent the effect of a single taper in the image. With this

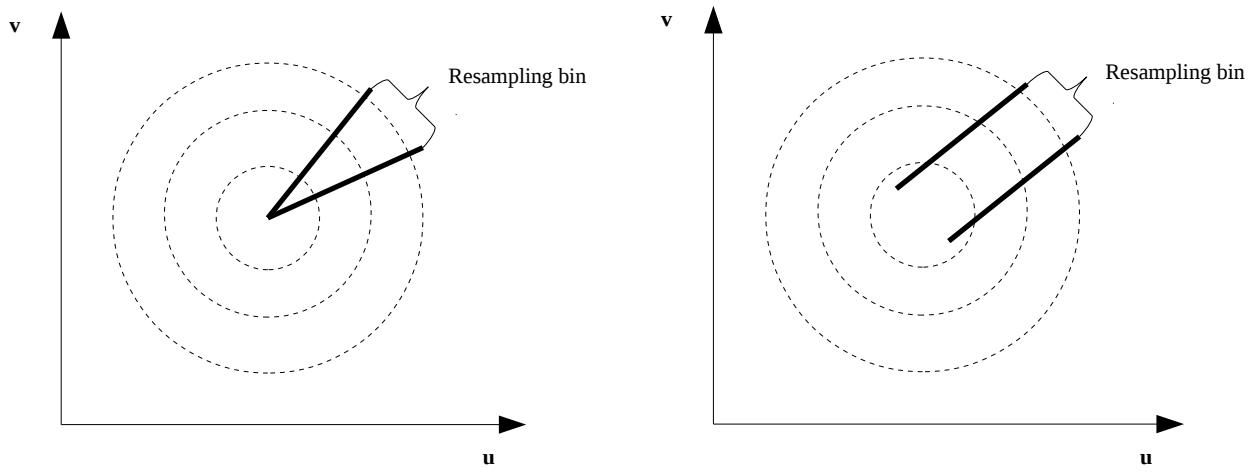

Figure 5. Applying baseline dependent averaging or normal averaging is equivalent to convolving the resampling bin by a boxcar windowing function. An Est-West interferometer array, the baseline dependent averaging (left panel) corresponds to an equal convolution kernel and resampling bin across all baselines.

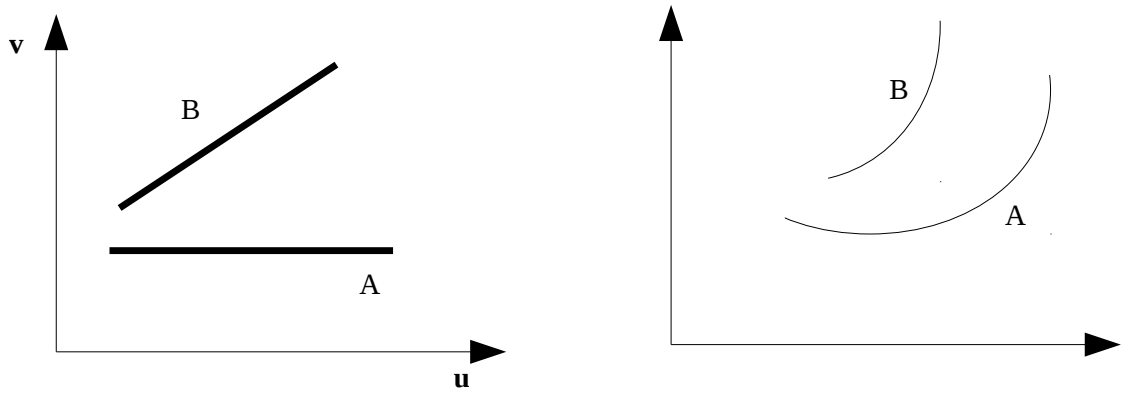

Figure 6. Two equal lengths baselines A and B with different orientation. The baselne A is a typical East-West baseline (the South-North component is zero) and B has a non zero South-North and East-West components. The distances sweep by these baselines are different during the same time integration.

condition, Eq. (13) becomes

$$\mathcal{I}^{\mathrm{D}} \sim \sum_{pqkl} W_{pqkl}\mathcal{P}_{pqkl} \circ \mathcal{I} \cdot \mathcal{T}, \tag{15}$$

where $\mathcal{T} \approx \mathcal{T}_{pq} \approx \mathcal{T}_{\alpha\beta}$ is the smearing response, constant on all baselines.

3.2 Compression factor

The compression factor is defined as the ratio between the pre-averaged data (high-res data) size and the average data (low-res data) size. If a complex value is stored with 32 bits, the data size of the high-res measurement set is defined as

$$2Data\ size\ hires = \frac{N \times (N-1)}{2} \times N_{sub} \times N_{tslt}^{hires} \times N_{ch}^{hires} \times N_{pol} \times 32\ [bits], \tag{16}$$

where N is the number of antennas of the interferometer array, N_{sub} the number of sub-bands, N_{pol} is the number of polarizations, N_{tslt}^{hires} and N_{ch}^{hires} are the number of timeslots and channels of the high-res measurement set respectively. In the case of simple averaging, the data size of the average measurement set is given by

$$Data\ size\ avg = \frac{N \times (N-1)}{2} \times N_{sub} \times \frac{N_{tslt}^{hires}}{n_t} \times \frac{N_{ch}^{hires}}{n_\nu} \times N_{pol} \times 32\ [bits], \tag{17}$$

where n_t is the number time bins and n_ν is the channels averaged on each baseline. Note that this formulation applies for an equal compression across all baselines, therefore the compression factor is defined by

$$2Compression\ factor\ avg = \frac{Data\ size\ hires}{Data\ size\ avg} = n_t \times n_\nu \tag{18}$$

The space savings which is defined as the reduction in size relative to the high-res data size therefore follows

$$2space\ savings\ avg = 1 - \frac{Data\ size\ avg}{Data\ size\ hires} = 1 - \frac{1}{n_t \times n_\nu} \tag{19}$$

Let us now define the compression factor for BDA. Recall that in this case, the compression rate varies across baselines. Following the analogy for BDA resampling bins, the number of samples in the resampling bin for a baseline pq is defined as $dim\mathsf{B}_{kl}^{[\Delta_{pq}t,\Delta_{pq}\nu]} = n_t^{pq} \times_\nu^{pq}$, where n_t^{pq} and n_ν^{pq} are the number of time and frequency samples respectively. The interferometer array data size for BDA then follows

$$2Data\ size\ bda = \sum_{pqkl} N_{sub} \times \frac{N_{tslt}^{hires} \times N_{ch}^{hires}}{dim\{\mathsf{B}_{kl}^{pq}\}} \times N_{pol} \times 32\ [bits] = \sum_{pq} N_{sub} \times \frac{N_{tslt}^{hires}}{n_t^{pq}} \times \frac{N_{ch}^{hires}}{n_\nu^{pq}} \times N_{pol} \times 32\ [bits] \tag{20}$$

The interferometer array compression factor for BDA is then given by

$$2Compression\ factor\ bda = \frac{N(N-1)}{2} \times \left(\sum_{pq} \frac{1}{n_t^{pq} \times n_\nu^{pq}} \right)^{-1}, \tag{21}$$

and the space savings is

$$2space\ savings\ bda = 1 - \left(\frac{N(N-1)}{2} \right)^{-1} \times \sum_{pq} \frac{1}{n_t^{pq} \times n_\nu^{pq}} \tag{22}$$

In the rest of this chapter, we refer to compression factor as $CF = x_t \times y_\nu$, where x_t is the compression factor in time and y_ν is the compression factor in frequency. The notation, $CF = x_t \times 1$ implies the data is compressed only in time by a factor of x_t, while $CF = 1 \times y_\nu$ implies that data is compressed only in frequency by a factor of y_ν.

4 IMPLEMENTATION DETAILS

The principal difficulty with BDA is that, the visibility plane contains either repeatedly averaged values or flagged data points. The correlation matrix is a time and frequency regular grid. Averaging entries in this correlation matrix over long times for short baselines and short times for long baselines result to an irregular grid. In practice, a better idea is to map this irregular correlation matrix onto a regular grid by either flagging out the supplementary points, or duplicating the averaged values on these supplementary points. We explain these processes in the following.

4.1 Flagging

In the BDA procedure, one has to make sure that the resampling bin interval contains an odd number of data points in time and frequency i.e,

$$dim\{\mathsf{B}_{kl}^{[\Delta_{\alpha\beta}t,\Delta_{\alpha\beta}\nu]}\} = (2k_t+1)(2k_\nu+1),\ \ k_t\ and\ k_\nu\ are\ integers \tag{23}$$

where $2k_t+1$ is the number of visibilities to average in time and $2k_\nu+1$ the number of visibilities to average in frequency. This condition must hold on all baselines otherwise the average baseline vector may not coincide with the mid time and frequency vector and will lead to a phase shift. If this condition is satisfied, the average value is assigned to the mid

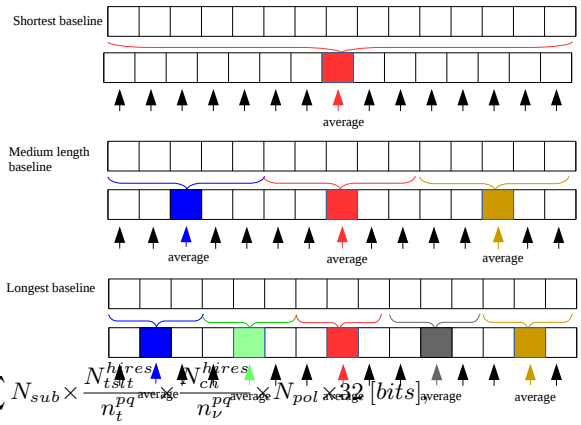

Figure 7. Baseline dependent averaging with flagging: the bins are averaged and the average value is assign at the centre of the resampling bin interval, while a flagging value is assign to others points.

resampling bin interval point i.e at the $k_t^{th}+1$ and $k_\nu^{th}+1$ visibility point. The others entries of the resampling bin are assigned a flag value which is generally unity. This flag value will account for missing samples during post-processing. Let us suppose that the data in Figure 7 represents 15 time-bins, for three different baselines i.e the longest, the medium length and a shortest baseline. Each bin is a 1 s integration sample and 15 s bins are averaged on the shortest baseline, 5 s on the medium length baseline and 3 s on the longest baseline. The colored pointers indicate the averaged bins and the black pointers the flagging points. Note that most of the flagging occurs on the shortest baselines because of the larger time bins averaged (15 s).

4.2 Duplication

This method consists of duplicating the average value at all entries of the resampling bins. While this process is easier to implement than the flagging method, it may not serve the purpose of data compression and/or quick computation for post-processing. Easier to implement in the sense that, one may not care or always verify that the number of visibilities points in the resampling bin is an odd number. Also note here that the data size of the resulting measurements set remains the same as the pre-averaged measurement set due to the fact that all value were duplicate along the pre-averaged measurement set. This method may be used in practice for cases where one does not want to average the uv coordinates as there are in the pre-averaged measurement set. While this method may be of little use, it is interesting due to the fact that its leads to an increased in nominal sensitivity since averaging represents maximum sensitivity. Similarly to the example in Figure 7, Figure 8 represents the duplicate method. Pointers with the same color indicate the duplicate averaged values throughout the resampling bin intervals.

4.3 Semi duplication and flagging

This method consists of combining the flagging and the duplicate methods in other to benefit of their full advantages.

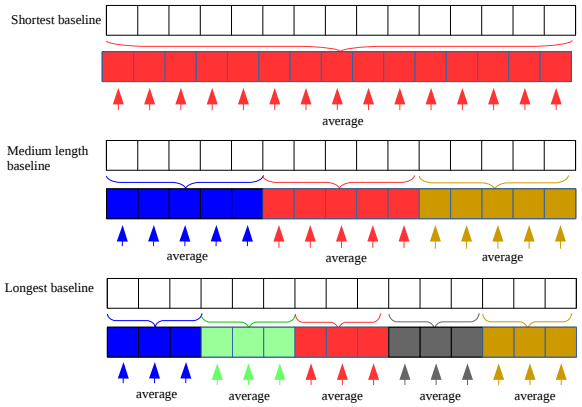

Figure 8. Baseline dependent averaging with duplication: the bins are averaged and the average value is assign at all points of the resampling bin interval.

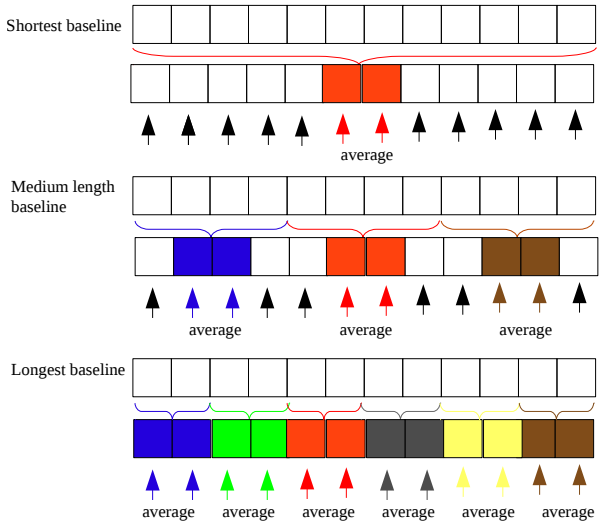

Figure 9. Baseline dependent averaging with semi duplication and flagging: the bins are averaged and the average value is duplicated only at the two central points of the resampling bin interval and the other bins are flagged.

By so doing, we seek both for data compression and quick computation while on the other hand, the implementation is easier to handle. We have learned from the flagging method that it is complex to implement but save for data compression and quick processing. We also learned from the duplication method that its does not help for compression or quick data processing but easier to implement. The semi duplication and flagging method should be able to account for both data compression aspects and easier implementation simultaneously. The idea is to duplicate the averaged bin along two central entries of the resampling bin if the total number of entries within this resampling bin is even otherwise the averaged bin is assigned only on the central bin of the resampling bin. Any other entry is then flagged as we shall learn in the example presented in Figure 9.

5 SIMULATIONS AND RESULTS

Having explored the mathematics and difficulties behind the implementation of BDA, we now turn to the simulation aspects. To probe the simulation issue, we consider our tests on a sky model of 1 Jy point source at various sky positions, with no noise or other corruptions included. Using the JVLA telescope in its C configuration, we evaluate the efficiency of a BDA correlator using two different procedure. First, we simulate the source at a fixed sky position and applied BDA, and measured the averaging effects separately on each baselines. Second, we simulate the point source at various radius from the phase centre and applied BDA, and BDWFs across equal uv-distance, hereby evaluating the interferometer array cumulative smearing effects on all baselines. Following the same procedure used in chapter **??** (section **??**), we measured the source peak amplitude in each image after averaging. Since each dirty image corresponds to a single source, the peak gives us the degree of smearing associated with a given averaging method and compression factor.

5.1 Source amplitude vs east-west baseline length

We simulate two high-res measurement sets each with a source at 30 arcmin from the phase centre of the observation. Furthermore, we generates two low-res measurement sets to receive the resampled visibilities. We present the amplitude loss as a function of Est-West baseline length. The results of BDA and simple averaging are compared in Figure 10.

- The first dataset from which the results are presented in the top panel of Figure 10 consisted of 10 frequencies channels of 125 kHz width for high spectral resolution, and 1000 timeslots of 1 s integration time. In this test, the compression factor is fixed to 48 both for BDA and simple averaging. Note that, for BDA, the shorter baselines are compressed by a lot more than 48 while the longer baselines by a lot less, this correspond to 4 s on the longest baseline and 250 s on the shortest baseline, while for simple averaging, 48 s bins are averaged in time.
- The second dataset consists of 10 timeslots of 1s integration, and 1000 frequencies channels of 125 kHz bins width. The compression factor is fixed to 32 both for BDA and simple averaging. For BDA, this result to 100kHz and 11kHz averaging bins in frequency for the shortest and the longest baseline respectively, while 30kHz frequencies bins are averaged for simple averaging.

It is pretty noticeable in Figure 10 that on shorter baselines, the smearing rate of BDA and simple averaging are equivalent. The physical interpretion of this is that, at this compression rate, visibilities averaging do not suffer from decorrelation on the shorter baselines. On the other hand, BDA suppresses smearing on the longer baselines compared to simple averaging. This is because fewer samples are averaged on the longer baselines when applying BDA, thereby causing a decrease in decorrelation.

5.2 Source amplitude vs distance from the phase centre

In this section, we simulate a high-res measurement set corresponding to a 375 s synthesis at 1 s integration, with 42.6

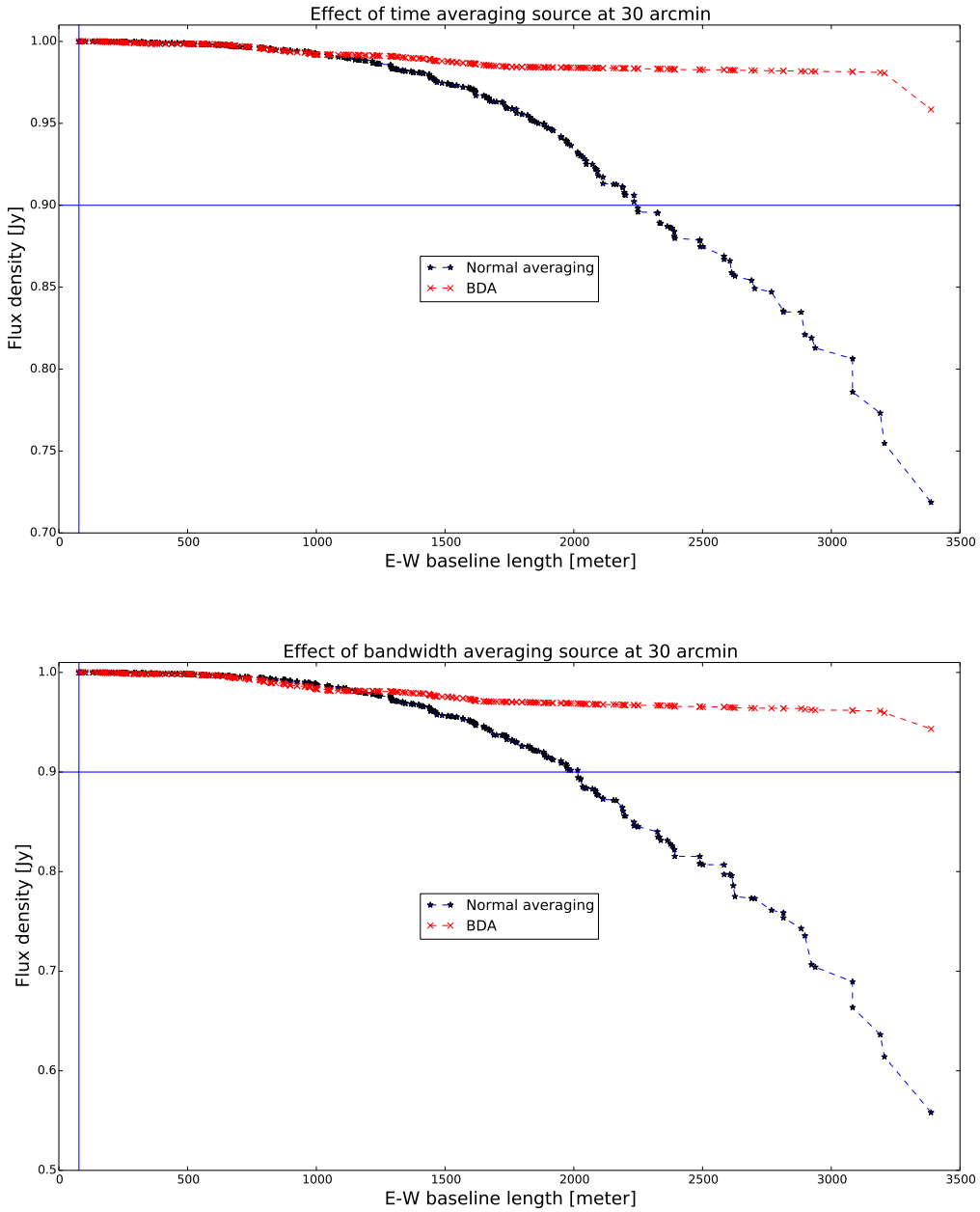

Figure 10. Amplitude loss: the apparent intensity of a 1 Jy source, as seen by JVLA-C at 1.4 GHz, as a function of Est-West baseline components. (Top) time and frequency integrations fixed at 100s and 125 kHz respectively; (Right) time and frequency integrations fixed at 1 s and 10 MHz respectively.

MHz total bandwidth centred at 1.4 GHz, divided into 511 channels of 83.4 kHz each. The sky model is a single 1 Jy point source at a given distance from the phase centre with noise free. We then generate two measurement sets to receive the resampled visibilities:

(i) A low-res measurement set to receive the resampled visibilities with simple averaging. This measurement set correspond to 25 s integration time and 2.5 MHz channels width. I remind the reader that, the synthesis time for this measurement set is 325 s splitted into 13 timeslots with 37.5 MHz total bandwidth divide into 15 channels each of width 2.5 MHz. It is pretty remarkably that, the synthesis time and the total bandwidth are less than the hight-res measurement set. We learned from chapter **??** (section **??**) that, for overlap BDWFs some number of time and frequencies bins must be allowed at the observation starting time/or frequency and ending time and/or frequency. These overlapping bin correspond here to 50 s and 2.5 MHz in time and frequency respectively.

(ii) The second measurement set is prepared to receive the sampled visibilities for BDA. This MS is a copy of the hight-res MS where one of the BDA implementation methods describe

in section 4 will be apply. It is important to note here that, for the case of overlap BDWFs, all the bins reserved for the overlap filters will be flagged at the end of the resampled procedure i.e before the resampled visibilities are imaged.

As was discussed in chapter **??** (section **??**), applying BDWFS to visibilities also results in different image plane tapers. For example, with a sinc-1×1, longer baselines correspond to sinc-like window function and boxcar taper in the image plane. While for shorter baselines, a sinc-1×1 correspond to boxcar and a sinc-like taper in the image plane, and therefore increased smearing. Applying BDA to BDWFS may lead to a more optimal response given that the shorter baselines and the longer baselines are adapted with the same visibility plane sinc-like convolution kernel, thus resulting to an equal boxcar taper in the image plane.

To distinguish the case of overlapping BDWFs applied across an equally uv-distance from non-overlapping ones, we will designate the overlapping window functions as *bda-WF-*$x \times y$. It implies that, for all baselines, x s and y MHz bins are overlap in time and frequency respectively. For example, *bda-sinc*-25× 2.5 is an overlap baseline dependent sinc function applies across an equal uv-distance, and where the total bins averaged are $(25 + \Delta_{pq}t)$ s and $(2.5 + \Delta_{pq}\nu)$ MHz in time and frequency respectvely.

The compression factor for the results presented in this section is fix to CF=25×30 for for all averaging methods i.e for simple averaging, 25 and 30 time and frequency bins are averaged in time and frequency respectively. For BDA, the averaged time and frequency bins are baseline dependent as presented in Appendix **??**.

In Figure 11, we compare the performance of simple averaging with a resampling bin sizes of 25s×2.5MHz (solid grey lines) to:

- *bda* (solid red lines), which is a BDA with a compression factor equal to CF=25×30
- An overlapping *bda-sinc*-25× 2.5 turned over two FoVs, 2∘ (solid blue lines) and 4∘ (dashed blue lines)
- An overlapping *bda-bessel*-25× 2.5 turned over two FoVs, 2∘ (solid green lines) and 4∘ (dashed green lines)

Based on the above discussion, we can now interpret some our main results. This results can be alternatively appreciated by regarding the performances of BDWFs apply across an equal uv-distance on all baselines.

(i) BDA provides good results in flux recovery i.e for 5% smearing, this gives us a field with radius 1.4° while with simple averaging, we can only recover a field with radius 0.9° at this compression factor. We can also notice from the result that, BDA is far to achieve source suppression.

(ii) Overlapping BDWFs apply across equal uv-distance (*bda-sinc*-25× 2.5, *bda-bessel*-25× 2.5) performance is excellent for source suppression and flux recovery within the choosing FoV. If the desired FoV size is 2° or 4°, overlapping BDWFs (*bda-sinc*-25× 2.5, *bda-bessel*-25×) provide excellent performance (1% or less) compared to averaging, the smearing performance across the FoV is better than that of simple averaging or BDA, and out-of-FoV source suppression is almost two orders of magnitude higher.

6 CONCLUSION

For a fixed time length a long baseline will cover a longer track in visibility space compared to a shorter baseline which results in the lower spatial modes being oversampled compared to higher spatial modes. This necessitates the use of baseline-dependent averaging to optimize the image plane response for each baseline. The use of baseline dependent averaging and BDWFs for an equal uv-space averaging have been investigated, this in two different aspects: mathematically and via simulations.

We have identified and established that baseline dependent averaging can only be used for purpose of data compression and FoV recovering i.e baseline dependent averaging decreases smearing over the observation FoV, while on the other hand, sources out of this field are not suppressed compared to simple averaging. We have found that combining BDA with BDWFs results to an excellent tapering behavior, which can decrease smearing to about 1% or less over a selected FoV, with out of FoV sources suppression almost two orders of magnitude higher compared to simple averaging, while the data is compress at the same rate.

We should also note that, baseline dependent averaging also distorts the PSF, and differently compared to simple averaging. An interesting avenue for future work is deriving this PSF at different sky position for the case of baseline dependent averaging in the offing to integrate into an existing imager.

We still have to keep in mind that, using baseline dependent averaging, the calibrations parameters as was described in the introduction will change during the observation, due to the varying time and channel lengths on different baselines. This is an obvious direction for future research plan with baseline dependent averaging.

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

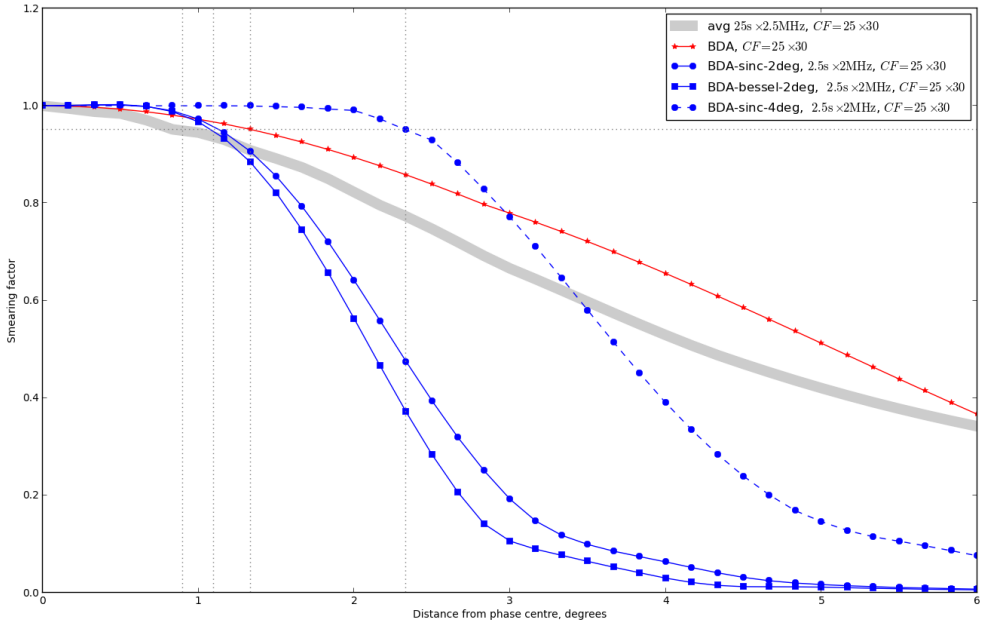

Figure 11. Amplitude loss: the apparent intensity of a 1 Jy source as seen by JVLA-C at 1.4 GHz as a function of distance from phase centre, for simple averaging with 25s×2.5MHz bins, and for BDA, and overlapping BDWFs for equal compression factor of $CF = 25 \times 30$.