32nd URSI GASS, Montreal, 19–26 August 2017

High-compression Baseline Dependent Averaging

Stefano Salvini[(1)], and Stefan J. Wijnholds*[(2)]
(1) Oxford e-Research Centre, Oxford, UK
(2) Netherlands Institute for Radio Astronomy (ASTRON), Dwingeloo, The Netherlands

Abstract

Baseline dependent averaging (BDA) can be used to reduce the volume of visibility data significantly. Most current BDA schemes perform (weighted) averaging over a certain time interval. This quickly causes decorrelation due to time averaging. We propose to reduce this decorrelation by representing the visibilities by polynomial coefficients. The high compression made feasible by this approach may cause fast-changing calibration parameters to become undersampled. We propose the Compress-Expand-Compress (CEC) method to mitigate this. All compression and expansion methods proposed herein are very simple and cause negligible computation overhead. We demonstrate the effectiveness of our scheme in a simulation emulating a high-dynamic range imaging problem.

1 Introduction

The Square Kilometre Array (SKA) [1] will produce an unprecedented amount of visibility data. This poses significant challenges for data transport between the correlator and the science data processor, data processing and storage of the raw visibility data for selected observations. For these reasons, (effectively lossless) compression of the raw visibility data is attractive. Traditionally, correlators for radio interferometers are designed to have the same integration time on all baselines. As a result, the integration time is limited by the longest baseline of the interferometer. On shorter baselines, longer integration times are possible. This led to the idea of baseline dependent averaging (BDA), which has been successfully used in MWA [2] and can also be used to shape the imaging field of view [3].

In these BDA schemes, (weighted) averaging of the visibilities is done over a certain time interval. This results in decorrelation of the received signals due to time smearing rather quickly. To allow longer integration times (and thus higher compression rates) with the same level of decorrelation, we propose to represent each visibility by low-order polynomial coefficients. If the compression rates are very high, the integration time on the shortest baselines may become longer than the time scale over which environmental and instrumental factors that require calibration, can be assumed constant. To deal with such situations, we introduce the Compress-Expand-Compress (CEC) method. In this scheme, the data are compressed at the output of the correlator, expanded to the desired time resolution for calibration and compressed again after calibration. We successfully demonstrate both techniques in simulation.

2 Decorrelation effects

For convenience of notation, we present our analysis of the amount of decorrelation for different polynomial orders for a one-dimensional array. This array is assumed to be aligned with the u-axis of the visibility plane. Let the fringe stopping center lie in the (u,v)-plane at an azimuthal angle ϕ, which is defined such that $\phi = 0$ at meridian transit. Finally, let θ be the azimuthal angle for which we want to analyse the decorrelation.

Given a source of unit brightness it can be easily shown that the observed visibility for a baseline of length u integrated over an time interval of length T is given by:

$$\langle V \rangle = \int_{-T/2}^{T/2} \exp\{iku\,[\sin(\theta-\omega t)-\sin(\phi-\omega t)]\}\,dt \quad (1)$$

If $\omega t \ll 1$ (corresponding to integration times), then

$$\langle V \rangle \approx A \int_{-T/2}^{T/2} \exp\{iku\omega t\,[\cos\theta-\cos\phi]\}\,dt \quad (2)$$

where

$$A = \exp\{-iku\,[\sin\theta-\sin\phi\,]\}\,. \quad (3)$$

We can easily obtain:

$$\langle V \rangle \approx \frac{AT}{2}\,\mathrm{sinc}\left(ku\omega\frac{T}{2}\,(\cos\theta - cos\phi)\right) \quad (4)$$

where $\mathrm{sinc}(x)$ is the *unnormalised* sinc function defined by $\mathrm{sinc}(x) = \sin(x)/x$.

If with $\langle V \rangle_0$ we denote the integral approximated by midpoint rule for a very short integration (no decorrelation), we have

$$\langle V \rangle_0 = \frac{A \cdot T}{2} \quad (5)$$

Using the Taylor series expansion

$$\mathrm{sinc}(x) = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \ldots$$

we can write the relative error:

$$\varepsilon_r = \frac{\|\langle V\rangle - \langle V\rangle_0\|}{\|\langle V\rangle\|} \approx \frac{\|\langle V\rangle - \langle V\rangle_0\|}{\|\langle V\rangle_0\|} \quad (6)$$

$$= \frac{\left[ku\omega\frac{T}{2}(\cos\theta - \cos\phi)\right]^2}{3!} - \dots \quad (7)$$

as, in general, the higher order terms are negligible. Based on this error, we can obtain an approximate formula for the maximum integration time compatible with a given pointing direction, a given source direction, baseline and allowed maximum relative error:

$$T \leq \frac{2\sqrt{6\varepsilon_r}}{ku\omega(\cos\theta - \cos\phi)} \quad (8)$$

Similar formulas, albeit more complicated, can be derived for image plane errors. It should be noticed that the error bound obtained assumes the worst case scenario, as it calculates the error for a baseline at maximum fringe rate. This bound only applies if $\omega t \ll 1$ and if any receivers' corrupting complex gains vary very slowly. Otherwise, adequately compressed visibilities (and good images) can be obtained by retaining higher order polynomials as we will explain in the next section.

3 Compress-Expand-Compress Method and Polynomial Approach

A potential issue with BDA is that different baselines are compressed by different ratios. As a result, the sampling instances may vary per visibility. Calibration algorithms thus need to carry out appropriate interpolation to obtain visibility data at the same time instances. Also, the integration time on the shortest baselines may be longer than the update rate of the fastest-changing calibration parameters.

Fortunately, snapshot calibration algorithms, like StEFCal [4], have a typical complexity of $\mathcal{O}\left(N_e^2\right)$, where N_e is the number of elements of the interferometer. The initial calibration stage is therefore expected to have much lower computational cost than other components in a standard pipeline such as sky model subtraction and gridding. This makes it viable to perform calibration on data sampled at the full rate (before BDA) without dramatically increasing the overall computational costs.

The Compress-Expand-Compress (CEC) method relies on this assumption and consists of three steps:

1. *Compression:* After any appropriate flagging on the raw visibility data, BDA is applied. This reduces the amount of data to be transferred from the cross-correlator to the imaging engine.
2. *Expansion:* The averaged visibilities are then used to reconstruct approximately the visibilities at the original dump time. First-stage calibration is then carried out on those reconstructed visibilities.
3. *Compression:* After calibration, the corrected visibilities are compressed again by applying a second stage BDA. This compression reduces the amount of data that needs to be gridded. As fast-varying instrumental and environmental effects are removed by calibration, compression could be higher than in the first compression stage.

For compression, most BDA schemes perform (weighted) averaging over a certain time interval. This results in decorrelation of the received signals due to time smearing rather quickly. We therefore propose to represent each visibility by a set of low-order polynomial coefficients for each short term integration (STI) interval. Note that the length of this STI interval is typically baseline dependent. It is therefore convenient to collect the visibility data for all baselines with the same compression C for each correlator dump (indexed by k) into a vector $\mathbf{r}_k^{(C)}$. The data for C correlator dumps can be collected in a matrix $\mathbf{R}^{(C)} = \left[\mathbf{r}_1^{(C)}, \cdots, \mathbf{r}_C^{(C)}\right]$. For each of these baseline sets, we can construct an orthonormal matrix $\mathbf{Z}^{(C)}$ with C rows and $p+1$ columns, where p is the order of the interpolating polynomial. This construction can, for example, be done by the modified Gramm-Schmidt process.

The averaged visibilities can now be represented by their coefficients

$$\mathbf{C}_{\mathbf{R}^{(C)}}^{(C)} = \mathbf{R}^{(C)}\mathbf{Z}^{(C)}, \quad (9)$$

from which an estimate of the full time resolution visibilities can be obtained by

$$\widehat{\mathbf{R}}^{(C)} = \mathbf{C}_{\mathbf{R}^{(C)}}^{(C)}\left(\mathbf{Z}^{(C)}\right)^H \quad (10)$$

during the expansion step. This procedure is carried out for the respective sets of baselines for each compression ratio.

During the second compression stage, the calibrated visibilities can be compressed by weighted averaging before passing them to the gridding (imaging) stage, where there is no need (or desire) to reconstruct visibilities at the original dump rate.

In the next section, we will show that using up to second order polynomial representations in the first compression stage (immediately after the correlator) can lead to a significant increase in the compression rate of the visibility data to be gridded.

4 Simulation results

For our simulations, we used a 1-D array of 194 15-m dishes spaced along an East-West line with a longest baseline of 100 km. The "central core" consisted of a number of regularly spaced antennas; the outer antennas were spaced logarithmically. This configuration allows us to assess the maximum error due to time smearing by studying a source during meridian transit, i.e., at maximum fringe rate. We

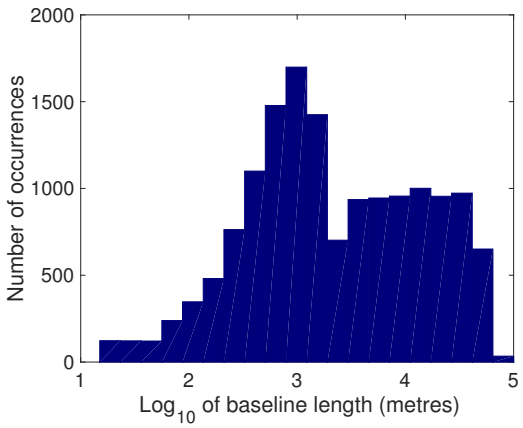

Figure 1. Baseline distribution.

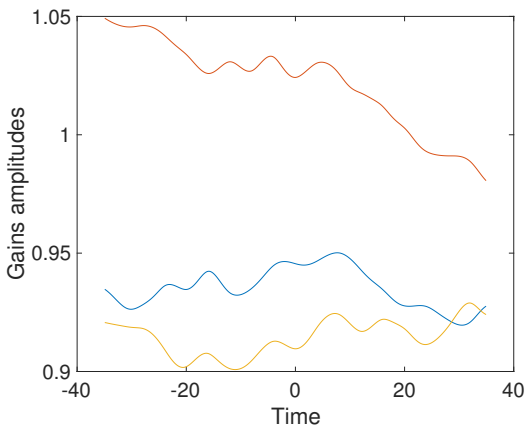

Figure 2. Gains amplituedes for receivers 1, 6 and 13.

simulated the effect of six sources: a 1-Jy source placed at the edge of the FoV, about 0.8 degrees away from the phase centre, surrounded by three 10^{-5}-Jy sources very close to it, and by two 10^{-6} sources in very close vicinity of the first four. The observing frequency was 700 MHz, whilst the interferometer was pointing 10 degrees off transit. The dump time was 0.14 seconds. The observed visibilities are obtained for each dump time integrating numerically evaluating Eq. 2 for each baseline. Figure 1 shows the distribution of the baselines.

The visibilities were corrupted by receiver-based complex-valued gains that were generated by using a fractional Brownian motion process parameterised by Allen variance and smoothed using a Hanning window. Gain amplitude and phases had Hurst parameters equal to 0.8 and 0.5 respectively. The corresponding values for the Allan variance were 0.01 and 2 degrees and a Hanning window of length 10 was used.

These gains were generated at smaller time intervals than the dump time and cubic splines were used to interpolate the gains between the points at which they were defined to avoid discontinuity of the gains and their derivatives. Figures 2 and 3 show the amplitudes and phases of the corrupting gains for receivers 1, 5, and 13.

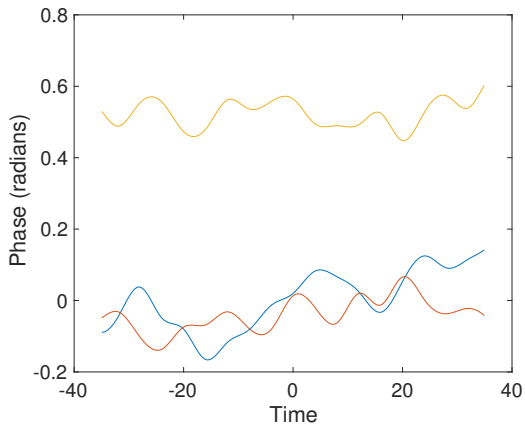

Figure 3. Phases of complex gains for receiver 1, 6 and 13 in radians.

Quite obviously, these generated complex gains vary conspicuously over the simulation interval, probably far more than in real instruments. However they are used to show the power of a polynomial-based approach.

We have used two compression regimes, the second using a higher level of compression. Table 1 shows the range of baseline compression rates (BCRs), the overall stage 1 (post-crosscorrelation) compression for order 0, 1 and 2 polynomials and the overall compression for stage 2 (pre-gridding). Note that compression rates in stage 2 are the same for all orders of polynomials. The BCR should be interpreted in the following way, as, to first approximation, the decorrelation time depends on the inverse of the baseline lengths: given the integer part (floor) of the ratio between the current and maximum baseline $r = \lfloor \frac{u}{u_{max}} \rfloor$, the compression for the current baseline is given by $c(u) = c(u_{max}) + r \cdot \Delta c$, where Δc is the increment of compression rates.

	Scheme 1 (Fig. 4)		Scheme 2 (Fig. 5)	
	Stage 1	Stage 2	Stage 1	Stage 2
BCR	[1:1:10]	[1:1:100]	[2:2:20]	[4:4:200]
Order 0	6	10	17	39
Order 1	3.5	10	8.5	39
Order 2	2.5	10	5.5	39

Table 1. Compression rates for the two schemes examined

Figures 4 and 5 show images (obviously one-dimensional) for different levels of compression for both the first (at correlator output) and the second (pre-gridding) stage of BDA, including only images for polynomials of order 1 and 2. Figure 4 also includes the BDA for polynomials of order 0 (straight averaging) and shows that the image obtained is far from satisfactory for this compression scheme.

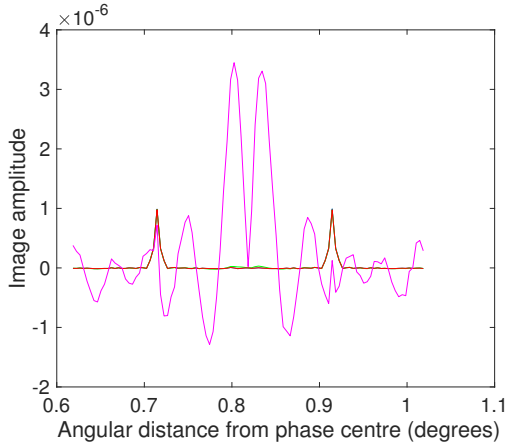

Figure 4. First compression scheme (see text)

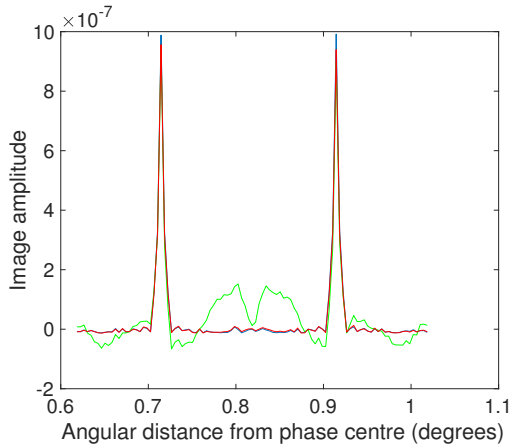

Figure 5. Second compression scheme (see text)

These figures report the difference between the calibrated sky image and the model sky image consisting of the four brightest sources to show whether the 10^{-6}-Jy sources can still be observed when BDA is applied. The blue, magenta, green and red line show the following

blue No BDA is applied. This gives a lower bound on the errors (basically 0 apart from numerical noise).

magenta BDA scheme based on simple mid-point averaging (order 0 polynomial). It is only shown in figure 4.

green and red BDA schemes based first and second order polynomials, respectively.

Figure 4 shows that, while the zero order approach is obviously inadequate, polynomials of order 1 and 2 give very good results, identifying the weak sources. A comparison between non-BDA and polynomial BDA of orders 1 and 2 bears this out.

Figure 5 shows that polynomial BDA (particularly for order 2) can employ high compression rates for both stages. While the first order polynomial BDA shows some artefacts, order 2 must be viewed as completely satisfactory. Order 0 results are not shown.

5 Conclusions

This work shows that BDA using a polynomial approximation to the visibilities can produce high-quality results while allowing a very high compression of the visibilities transmitted from the crosscorrelator to the imaging machine (stage 1) and even more so for stage 2 (pre-gridding). The results shown here are very encouraging and show that significant reductions in gridding costs can be achieved very easily. Almost a factor 40 (possibly, it could be pushed further) compression has been achieved for the pre-gridding stage.

We think that a similar approach could be useful in generating the model visibilities when a large number of sources are included. We also believe that monitoring the variations of the visibilities and their approximate derivatives could lead to a better dynamical estimation of compression rates.

6 Acknowledgements

This work is funded by the UK funding agency STFC, IBM, ASTRON, the Dutch Ministry of Economic Affairs and the Province of Drenthe.

The authors would like to thank Tony Willis, Ben Mort, Fred Dulwich, Jeroen Stil for the fruitful discussions during the development of the material presented in this paper.

References

[1] P. E. Dewdney, P. J. Hall, R. T. Schilizzi, and T. J. L. W. Lazio, "The Square Kilometre Array," *Proceedings of the IEEE*, **97**, 8, August 2009, pp. 1482–1496, doi: 10.1109/JPROC.2009.2021005.

[2] D. A. Mitchell et al., "Real-Time Calibration of the Murchison Widefield Array," *IEEE Journal of Selected Topics in Signal Processing*, **2**, 5, October 2008, pp. 707–717, doi: 10.1109/JSTSP.2008.2005327.

[3] M. T. Atemkeng, O. M. Smirnov, C. Tasse, G. Foster, and J. Jonas, "Using baseline-dependent window functions for data compression and field-of-interest shaping in radio interferometry," *Monthly Notices of the Royal Astronomical Society*, **462**, 3, July 2016, pp. 2542–2558, doi: 10.1093/mnras/stw1656.

[4] S. Salvini, and Stefan J. Wijnholds, "Fast gain calibration in radio astronomy using alternating direction implicit methods: Analysis and applications," *Astronomy & Astrophysics*, **571**, A97, November 2014, pp. 1–14, doi: 10.1051/0004-6361/201424487.