StEFCal – An Alternating Direction Implicit Method for Fast Full Polarization Array Calibration

Stefano Salvini[1] and Stefan J. Wijnholds*[2]

[1]Oxford e-Research Centre (OeRC)
7 Keble Road, OX1 3QG Oxford, United Kingdom. E-mail: stef.salvini@oerc.ox.ac.uk

[2]Netherlands Institute for Radio Astronomy (ASTRON)
Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands. E-mail: wijnholds@astron.nl

Abstract

Alternating Direction Implicit (ADI) methods provide a computationally efficient way to solve for antenna based gains in full polarization. In this paper, we analyze the convergence of such methods in simulations. We show that convergence of a basic implementation can be quite slow and we propose two forms of relaxation to improve convergence behavior. The algorithm can be shown to be statistically efficient in low-SNR scenarios, which makes this approach particularly suitable for calibration of large radio astronomical arrays like the Square Kilometre Array (SKA). This also led to the name Statistically Efficient and Fast Calibration (StEFCal). The algorithm was implemented in several stages of the Low Frequency Array (LOFAR) calibration pipeline. We report on calibration performance improvement achieved with StEFCal in this pipeline.

1. Introduction

Alternating Direction Implicit (ADI) methods provide a computationally efficient way to solve for antenna based gains. This was recognized by Johan Hamaker, who suggested to use it for full polarization calibration of radio astronomical arrays [1]. The Murchison Wide-field Array (MWA) uses an ADI approach for full-polarization calibration of the array in their real-time calibration pipeline, in which only one iteration is used using the result from the previous iteration as starting point and averaging the two solutions [2]. Salvini et al. showed that averaging after every second iteration is actually essential to ensure robust and fast convergence of the ADI iterations [3]. Since this approach turned out to be not only computationally efficient, but also statistically efficient [4], the algorithm now lives by the name StEFCal: Statistically Efficient and Fast Calibration.

In this paper, we propose two modifications to the basic full polarization algorithm with averaging after every second iteration. The first modification exploits 6-point averaging instead of 2-point averaging. The second modification is to use an update gain solution as soon as it becomes availabe. The second modification precludes parallellization of the algorithm over antenna index. However, in practice, radio astronomical data consists of many time and frequency samples and their processing is usually parallellized along those dimensions. The basic algorithm and the proposed modifications are described in more detail in the next section. In section 3, we assess convergence of the basic algorithm and its modified variants showing that the proposed modifications significantly improve speed of convergence. The algorithm was implemented in several calibration pipelines for the Low Frequency Array (LOFAR). We report on the achieved calibration performance improvement achieved with StEFCal in section 4. We conclude that StEFCal can calibrate radio astronomical arrays in full polarization with $O(N^2)$ numerical complexity, where N is the number of antennas to be calibrated. This scaling makes the algorithm particularly attractive for large arrays like LOFAR [6] and the Square Kilometre Array (SKA) [7].

2. Description of algorithm

The mathematical problem consists of finding the minimum with respect to the gain matrix G of

$$\min_{G}\left\|D-G^{H}MG\right\|_{F}$$

where D represents the observed visibilities (the data), M the model visibilities and G^H denotes the Hermitian transpose of the gain matrix G. G is diagonal for scenarios without polarization, or the scalar case, with G_{jj} corresponding to the complex gain for the jth antenna. In cases with full polarization, G is a block-diagonal matrix, whose jth 2-by-2 block corresponding to the two feeds of the jth antenna can be described as

$$G_j = \begin{bmatrix} G_{2j-1,2j-1} & G_{2j-1,2j} \\ G_{2j,2j-1} & G_{2j,2j} \end{bmatrix}$$

where the off-diagonal terms correspond to the crosstalk between the dipoles.

The ADI formalism keeps G^H fixed at each iteration while minimizing G. It can be shown, that the ADI approach splits the problem into N separate 2-by-2 Linear Least Squares problems solvable by, for example, the normal equations method:

$$\min_{G_j} \left\| Z_j^H D_j - \left(Z_j^H Z_j \right) G_j^{[k]} \right\|_F, \forall j, j = 1, 2, ..., N$$

where $Z_j = Z_{:,2j-1:2j}$ are the two columns of the matrix $Z = G^{[k-1]}M$ for the dipoles of the jth antenna, D_j and M_j are the corresponding columns of the matrices D and M respectively, and k is the iteration number.

In general, this simple algorithm shows poor convergence. However, this can be improved very considerably in a number of ways:

1-basic	Average the odd and even iterations – highly parallel
1-relax	As 1-basic, but also build a linear combination with previous iterations $G_j^{[k-2]}$ and $G_j^{[k-4]}$ – highly parallel
1-monitor	As 1-relax with heuristics to regularize convergence – highly parallel
2-basic	Use the $G_j^{[k]}$ as soon as they are computed within each iteration – parallel dependencies
2-relax	As 2-basic, but using a linear combination with the previous iterate $G_j^{[k-2]}$ – parallel dependencies

Here, highly parallel means that within each iterations the G_j can be computed in any order irrespective of the others (hence suitable, for example, for GPUs); when there are parallel dependencies, the computation of the G_j within each iterations requires the G_i, for all $i < j$ (unsuitable for many-core platforms, suitable for single threaded operation). All these algorithms

- have a small memory footprint, only requiring the visibility matrices (in full square form as explained in the next point) and two to four comples vectors of length N;
- have access to data (hence memory) exclusively by unit strides (consecutive memory storage) thus allowing very high levels of performance;
- require only 44 $(2\ N)^2$ *real* operations for each iteration.

3. Simulation results

We studied the convergence of the five algorithms for a test case with 100,000 sources following an exponential power distribution and including a few polarized sources. We included the brightest sources (about 30) in the sky model to build the model visibilities. The array configuration consisted of 256 dual-polarized antenas with a random distribution. The convergence results are reported in Figure 1. These results highlight the poor rate of convergence for the 1- and 2-basic algorithms and show that the other algorithms have comparable rates (i.e. slopes) of convergence after the initial steps. An important improvement to the 1- algorithms is to introduce some *bootstrapping*, i.e. to solve a simpler but related problem to achieve a much larger convergence initially, and using the results to start off the computation of the main iteration, for example by using a narrow band of the visibility matrices. This has proved very effective in practice.

Because of the characteristics of the StEFCal algorithms both high levels of performance and excellent scaling with problem size can be achieved, as shown in Figure 2. All results shown have been obtained by optimised code written by the authors, in double precision, running on a dual CPU node with Intel Xeon 2650 8-cores 2.0GHz CPUs. All performance data are for running *on a single core*. The figure reports scalability with size normalised to the number of iterations required and clearly show the very good scaling achieved as well as the fact that the computational costs are $O(N^2)$. In all cases, a convergence tolerance of 10^{-8} was set, most likely exceeding the demands of any practical application. Performance of the order of 50% of peak was achieved, which confirms the efficiency of StEFCal.

We have also tested a realistic scenario of full polarization calibration of a proposed SKA Low Frequency Aperture Array (LFAA) station, comprising of 256 antennas (512 dipoles), for 1024 frequencies [8]. The code was

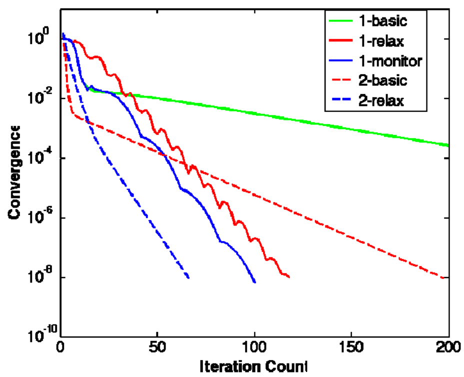

Figure 1 Convergence rate of algorithm variants

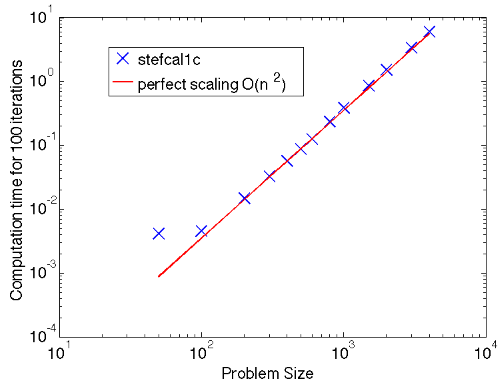

Figure 2 Algorithm scaling

N.Cores	Time (s)	GFlops	%CGEMM
1	16.63	17.9	41.8%
2	8.34	35.8	41.9%
3	5.57	53.5	42.2%
4	4.19	71.2	41.8%
8	2.90	102.9	42.0%
10	2.35	127.0	42.2%
12	1.92	155.3	41.3%
16	1.72	173.3	38.3%

Table 1 Multi-core performance results

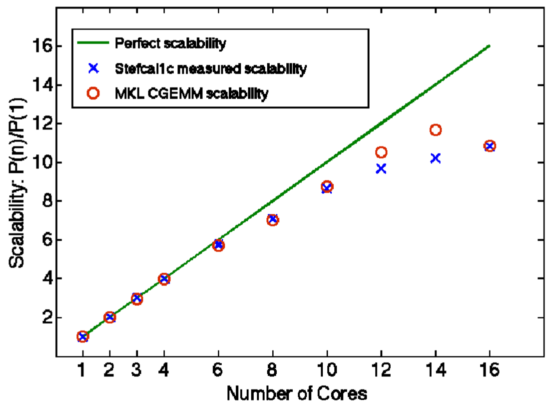

Figure 3 Multi-core performance results

parallelised over frequencies using OpenMP, whereby each core grabs the first available frequency yet to be calibrated (dynamic load balancing). All computations were carried out using single precision to a tolerance of 10^{-5}, delivering better than 1% accuracy in the complex gains, as required. Bootstrapping, as defined above, was employed.

Peak performance was defined as the performance of Intel MKL CGEMM (single precision complex matrix-matrix product), for an optimal matrix size (nominal peak is 44.8 GFlops), on the required number of cores. Notice, in particular, the decrease in scalability as the number of cores in use increases. This is due to unavoidable memory ad other contentions and affects StEFCal as much as CGEMM. In all cases, scalability and performance are very good. This clearly shows the potential of StEFCal for SKA LFAA calibration.

4. Practical results

The Amsterdam-ASTRON Radio Transient Facility and Analysis Centre (AARTFAAC) is an transient monitoring facility installed on the innermost six LOFAR stations. This facility lets these stations operate together as a single 288-antenna station with a diameter of 350 stations. AARTFAAC exploits image plane transient detection which requires near real-time calibration and imaging with a cadence of one second [5]. Calibration of the AARTFAAC system involved estimation of antenna based gains, apparent source fluxes and apparent source positions. Estimation of these parameters is done by alternatingly updating these groups of parameters. It was therefore straightforward to replace the original antenna based gain estimation step by the basic variant of StEFCal for the non-polarized case. Since the AARTFAAC system has 288 antennas, replacing the original estimator with $O(N^3)$ complexity by StEFCal with $O(N^2)$ complexity could potentially reduce the compute cost by a factor of order 100. Due to the larger number of iterations required by StEFCal, the net result was a factor 35 reduction in compute load. If the results from the previous timeslice were used as initial estimate, the number of StEFCal iterations could be reduced to one, reducing the compute load by another factor 8 for a total reduction of a factor ~200.

The LOFAR central processor performs a number of steps on the raw data produced by the correlator correlating the signals from the LOFAR stations. These steps include a calibration step for direction independent (station based) gains to provide initial corrections for clock drift and propagation effects. The latter are mainly caused by the ionosphere and may require full polarization corrections on baselines to stations outside the core area. Recently, Tammo Jan Dijkema made an implementation of the basic version of full polarization StEFCal for the standard processing pipeline of LOFAR. This implementation was used to run the same pipeline on several data sets from actual LOFAR observartions twice, once with the standard Levenberg-Marquardt (LM) solver and once with the LM solver replaced by StEFCal. In all cases, the results obtained were practically identical (there were some small differences caused by the different numerics of the two solvers), but the pipeline with StEFCal was typically a factor 10 faster than the pipeline running the LM solver. Based on the material presented in this paper, we expect that we can boost improvement by at least another factor five by optimizing the implementation of StEFCal.

5. Conclusions and future work

In this paper, we showed that two forms of relaxation can be used to improve the convergence behavior of simple ADI methods for antenna based gain estimation. The first approach is to use a weighted 6-point average, the second modification is to use updated gain solutions as soon as they become available. We assessed the computational performance in simulation showing perfect quadratic scaling for array sizes ranging from 100 to 4000 antennas and demonstrating the excellent use of computing resources. The effective use of computing resources and quadratic scaling make StEFCal well suited for large radio astronomical arrays like LOFAR and SKA. This was demonstrated by implementing StEFCal in the AARTFAAC calibration pipeline (no polarization) and the standard LOFAR post-correlation pipeline (full polarization). In both cases, at least a factor 10 reduction in compute time for calibration was achieved. Despite the low number of operations per sample required by StEFCal and corresponding I/O limitations in GPU implementations, an experimental implementation on GPUs is currently being studied. This is potentially interesting for SKA LFAA station calibration, since StEFCal can potentially be integrated with the station correlator if the latter is implemented on GPUs.

6. Acknowledgments

We would like to acknowledge the many and helpful discussion with Oleg Smirnov, Ronald Nijboer, Johann Hamaker, and particularly express our thanks to Tammo Jan Dijkema for his work on implementing StEFCal in the standard LOFAR pipeline and Marzia Rivi who is developing a GPU implementation (CUDA). The research leading to this paper has received funding from the European Commission Seventh Framework Programme (FP/2007-2013) under grant agreement No. 283393 (RadioNet3).

7. References

1. J. P. Hamaker, "Understanding radio polarimetry – IV. Th full-coherency analogue of scalar self-calibration: Self-alignment, dynamic range and polarimetric fidelity," *Astronomy & Astrophysics Supplement Series*, **143**, January 2000, pp. 515-534.
2. D. A. Mitchell et al., "Real-Time Calibration of the Murchison Widefield Array," *IEEE Journal of Selected Topics in Signal Processing*, **2**, no. 5, October 2008, pp. 707-717.
3. S. Salvini, F. Dulwich, B. Mort and K. Zarb-Adami, "A New Antenna Calibration algorithm for AA single station snapshot correlation," Aperture Array Verification Programme Workshop 2011, Dwingeloo, The Netherlands, 12-16 December 2011.
4. S. Salvini and S. J. Wijnholds, *Astronomy & Astrophysics*, in prep.
5. P. Prasad, S. J. Wijnholds, F. Huizinga and R. Wijers, "Real-time calibration of the AARTFAAC array for transient detection," *Astronomy & Astrophysics*, February 2014, submitted.
6. M. P. van Haarlem et al., "LOFAR: The LOw-Frequency ARray," *Astronomy & Astrophysics*, **556**, no. A2, May 2013, pp. 1-53.
7. P. E. Dewdney, P. J. Hall, R. T. Schillizzi and T. J. L. W. Lazio, "The Square Kilometre Array," *Proceedings of the IEEE*, **97**, no. 8, August 2009, pp. 1482-1496.
8. P. E. Dewdney et al., "SKA1 System Baseline Design," SKA Project Office, Manchester, United Kingdom, Technical Report SKA-TEL-SKO-DD-001, 12 March 2013.