Astronomy & Astrophysics manuscript no. paper
(DOI: will be inserted by hand later)
March 6, 2014

Yet another StefCal algorithm for direction-dependent radio interferometry calibration

N. Bourbaki

Address(es) of author(s) should be given

Received ¡date¿ / Accepted ¡date¿

Abstract.

1. Complex derivative

In this section we try to describe the shape and structure of the Jacobian over the complex scalar field, instead of what is usually done, by considering the real and imaginary separatly. Discarding the polarisation effects, the scalar Radio Interferometry Measurement Equation (RIME), we can write the visibility measured on baseline (pq), at time t and frequency ν as:

$$\mathrm{v}_{(pq)t\nu} = \sum_d g^d_{pt\nu}.\left(g^d_{qt\nu}\right)^*.k^d_{(pq)t\nu}.\mathrm{s}_d \tag{1}$$

$$\text{with } k^d_{(pq)t\nu} = \exp\left(-2i\pi\left(ul+vm+w\left(\mathrm{n}-1\right)\right)\right) \tag{2}$$

$$\text{and } n = \sqrt{1-l^2-m^2} \tag{3}$$

$$\tag{4}$$

with $[u,v,w]^T$ is the baseline vector between antennas p and q in wavelength units, and $\boldsymbol{s}_d = [l,m,n=\sqrt{1-l^2-m^2}]^T$ is a sky direction later labeled as d. In order to compute a Jacobian, we need to choose a derivative definition for complex numbers. If we write a complex number as $z = x+iy$, then the Wirtinger complex derivative operator writes as:

$$\frac{\partial}{\partial z} = \frac{\partial}{\partial x} - i\frac{\partial}{\partial y} \tag{5}$$

$$\tag{6}$$

where x and y are the real and imaginary parts respectively. The Wirtinger has a trivial but remarquable property that a scalar and its complex conjugate can be viewed as independent variables, and in particular:

$$\frac{\partial z^*}{\partial z} = 0 \tag{7}$$

Considering the sky, gain, and geometry relation given in Eq. 1, according to the property of Wirtinger derivative of complex conjugate (Eq. 7), we can see that:

$$\frac{\partial \mathrm{v}_{(pq)t\nu}}{\partial g^d_{pt\nu}} = \left(g^d_{qt\nu}\right)^*.k^d_{(pq)t\nu}.\mathrm{s}_d \tag{8}$$

$$\text{and } \frac{\partial \mathrm{v}_{(pq)t\nu}}{\partial g^d_{qt\nu}} = 0 \tag{9}$$

$$\tag{10}$$

We now consider the visibility vector $\mathbf{v}_{(pq)}$ for all time frequency block within a given interval. According to the definition of Wirtinger derivative, we can write the complex Jacobian $\boldsymbol{\mathcal{J}}_{\mathbf{v}_{pq}}$ of $\mathbf{V}_{(pq)}$ of size $[(n_t n_\nu)\times(n_a n_d)]$, where n_t, n_ν, n_a, and n_d are the number of time, frequency, antenna and directions. Each cell of $\boldsymbol{\mathcal{J}}_{\mathbf{v}_{pq}}$ can be written by taking a line corresponding to the measurement at $(t\nu)$, and a column $i = d + a.n_d$ (for given antenna a and direction d) corresponds to the derivative against g^d_a. We have:

$$\left[\boldsymbol{\mathcal{J}}_{\mathbf{v}_{pq}}\right]_{t\nu,i} = \begin{cases} \left(g^d_{qt\nu}\right)^*.k^d_{(pq)t\nu}.\mathrm{s}_d \text{ for } a = p \\ 0 \text{ otherwise} \end{cases} \tag{11}$$

We can see that non-zero columns are the ones corresponding to all direction for antenna p. The Jacobian for all baselines is written in a similar way, by superposing the $\boldsymbol{\mathcal{J}}_{\mathbf{v}_{pq}}$ for all (pq) pairs as follows:

$$\boldsymbol{\mathcal{J}}_{\mathbf{v}} = \begin{bmatrix} \vdots \\ \boldsymbol{\mathcal{J}}_{\mathbf{v}_{pq}} \\ \vdots \end{bmatrix} \tag{12}$$

which have size $[(n_{bl}n_t n_\nu)\times(n_a n_d)]$, where n_{bl} is the number of baselines and is typically $n_{bl} = n_a(n_a-1)/2$. Although it has large dimensions, $\boldsymbol{\mathcal{J}}_{\mathbf{v}}$ is sparse.

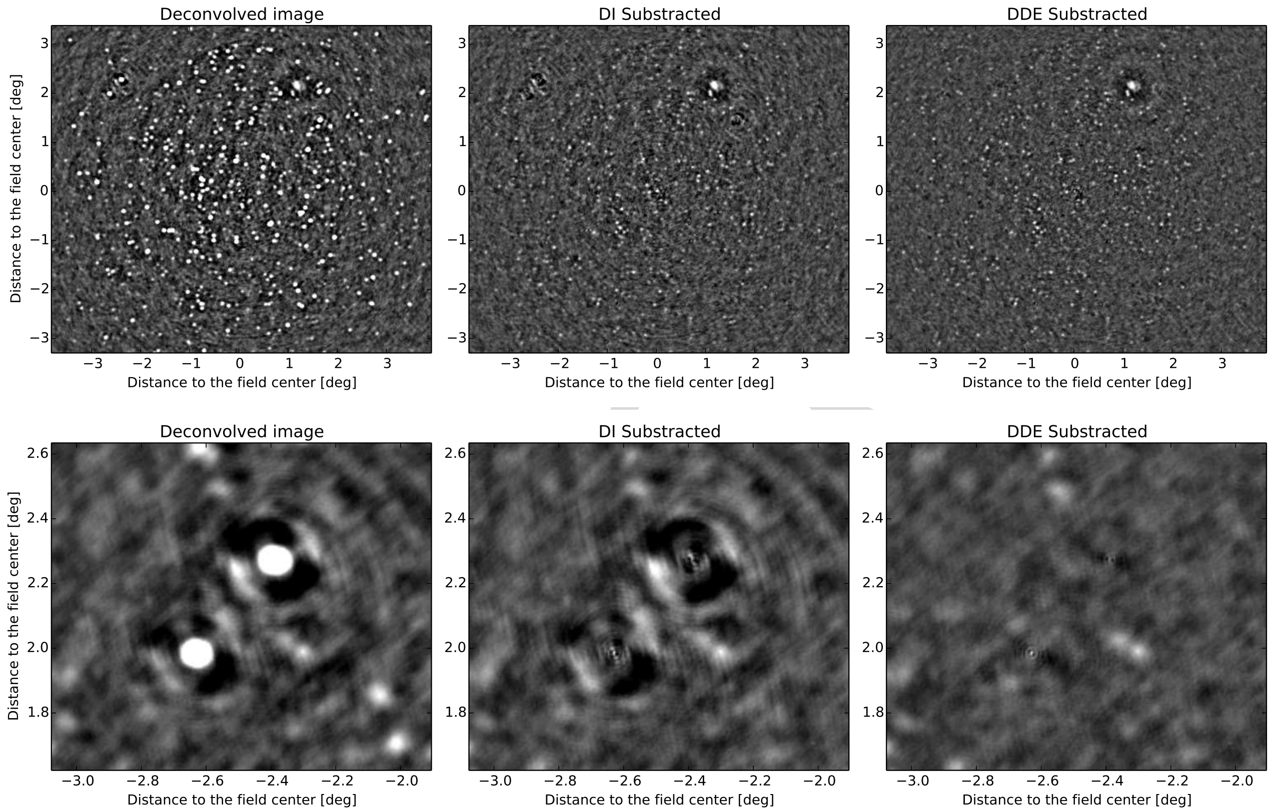

Fig. 1. This figure shows compares the image (left), the residuals data after simple skymodel substraction (center), and the residuals data after substracting the sky model corrupted by the direction-dependent solution (right).

2. Iterative direction-dependent solver

2.1. Linear system for calibration

If we write the direction dependent gain vector as $\boldsymbol{g}$, which i^{th} component $i = d + a.n_d$ if the gain of antenna a in direction d, then we find the remarquable property that around the solution, Eq. 1 behaves like a linear system. Specifically, from Eq. 11 and 12, it is easy to check that:

$$\mathbf{v} = \left(\boldsymbol{\mathcal{J}}_{\mathbf{v}} \big|_{\boldsymbol{g}} \right) . \boldsymbol{g} \tag{13}$$

where $\boldsymbol{\mathcal{J}}_{\mathbf{v}} \big|_{\boldsymbol{g}}$ mean that the Jacobian is evaluated at $\boldsymbol{g}$.

Assuming a linear operator satisfying Eq. 13 is given, we can build an iterative scheme to derive an estimate $\widehat{\boldsymbol{g}}$ of $\boldsymbol{g}$. The linear operator $\boldsymbol{\mathcal{J}}_{\mathbf{v}} \big|_{\widehat{\boldsymbol{g}_i}}$ is build from the estimate $\widehat{\boldsymbol{g}_i}$ at step i. Then Eq. 13 is solved using the least-squares solution given by computing pseudo-inverse as follows:

$$\widehat{\boldsymbol{g}_{i+1}} = \left[\mathbf{A}^H \mathbf{C}^{-1} \mathbf{A} \right]^{-1} \mathbf{A}^H \mathbf{C}^{-1} \mathbf{v} \tag{14}$$

$$\text{with } \mathbf{A} = \boldsymbol{\mathcal{J}}_{\mathbf{v}} \big|_{\widehat{\boldsymbol{g}_i}} \tag{15}$$

where $\mathbf{C}$ is the covariance matrix of $\mathbf{v}$.

2.2. Convergence and averaging

As shown in Fig. ... the convergence of this algorithm is slow, and following Stef-the-great, instead of estimating $\boldsymbol{\mathcal{J}}_{\mathbf{v}}$ at $\widehat{\boldsymbol{g}_i}$, we build it at a modified location constructed from previous iterations, and Eq. 15 becomes:

$$\mathbf{A} = \boldsymbol{\mathcal{J}}_{\mathbf{v}} \big|_{\widetilde{\boldsymbol{g}_i}} \tag{16}$$

$$\text{with } \widetilde{\boldsymbol{g}_i} = (\widehat{\boldsymbol{g}_{i-1}} + \widehat{\boldsymbol{g}_i})/2 \tag{17}$$

3. Test on real data

We test the algorithm described above on a LOFAR dataset. The visibilities produced by this interferometer are predominantly affected by direction dependent effects including (i) the phased beam instability and deviation from the theoritical model, (ii) ionosphere time delays shifts, and (iii) Faraday rotation.

We first calibrate the data using BBS, and in order to build a pertinent model of the field, we substract 3C295. We extract the sources using pyBDSM. The sources are the clustered in 10 directions using Voronoi tesselation (fig. 2). In Fig 1, we compare the residuals as computed by substracting the model data in the visibility domain, and the model data affected by DDEs.

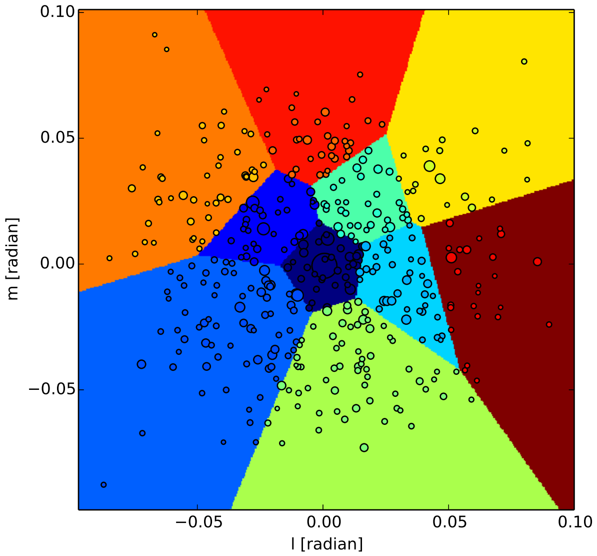

Fig. 2. In order to minimize the number of degrees of freedom, and increase the amount of signal in each direction, we cluster the sources in 10 direction using a Voronoi tesselation.

References