

RADIO INTERFEROMETRIC CALIBRATION USING A RIEMANNIAN MANIFOLD

Sarod Yatawatta

ASTRON, The Netherlands Institute for Radio Astronomy,
The Netherlands.

Email: yatawatta@astron.nl

ABSTRACT

In order to cope with the increased data volumes generated by modern radio interferometers such as LOFAR (Low Frequency Array) or SKA (Square Kilometre Array), fast and efficient calibration algorithms are essential. Traditional radio interferometric calibration is performed using nonlinear optimization techniques such as the Levenberg-Marquardt algorithm in Euclidean space. In this paper, we reformulate radio interferometric calibration as a nonlinear optimization problem on a Riemannian manifold. The reformulated calibration problem is solved using the Riemannian trust-region method. We show that calibration on a Riemannian manifold has faster convergence with reduced computational cost compared to conventional calibration in Euclidean space.

Index Terms— Calibration, Interferometry: Radio interferometry

1. INTRODUCTION

Radio interferometric calibration is the estimation of errors introduced by the propagation medium (such as the ionosphere) and by the receivers (such as the beam shape). In order to produce high fidelity and high dynamic range images, calibration is essential. While contemporary radio interferometric arrays at most have a few tens of receivers (or stations), there is a trend towards building large radio interferometers with hundreds of receivers, an example being the Square Kilometre Array (SKA). This naturally leads to data volumes that are by far greater than what is produced by contemporary radio telescopes.

The maximum likelihood estimation of calibration parameters is in fact a nonlinear optimization problem. Currently, nonlinear optimization algorithms such as the Levenberg-Marquardt (LM) method [1, 2] are used in radio interferometric calibration [3]. The cost function that is minimized during calibration is invariant to multiplication of the parameters by a 2 by 2 unitary matrix. Therefore, the solutions acquired by calibration will have a unitary matrix ambiguity [4].

In this paper, we present the 'quotient manifold' geometry [5, 6] of the calibration parameters, which is a better representation of their invariance to multiplication by 2 by 2 unitary matrices. We further develop the geometric structure of calibration parameters, first presented in [7]. Rather than minimizing the cost function in Euclidean space, as is currently done, we minimize the cost function on the developed quotient manifold. We use the Riemannian Trust-Region (RTR) method [8] for minimizing the cost function.

Optimization on matrix manifolds has developed significantly during the past decade and a complete overview can be found in [5]. In particular, when there is an underlying symmetry in the parameter space (such as the invariance to multiplication by a unitary

matrix), exploiting the geometric structure yields better performing algorithms [9, 10, 11].

Moreover, algorithms such as the LM operate in real parameter space and the cost of calibration of an interferometric array with hundreds of elements is significant, mainly due to the increased size of the Jacobian [12]. In this paper, we treat calibration parameters as complex numbers and because we employ the RTR method [8], the computational and memory costs are reduced. The novelty of the work presented in this paper (relation to prior work) is as follows: (i) We present the quotient manifold geometry of radio interferometric calibration, improving on [7]. (ii) We reformulate radio interferometric calibration as an optimization problem on a Riemannian manifold, where we derive expressions for the Riemannian gradient and the Hessian, following [9]. (iii) We apply the RTR method [8] for calibration instead of the traditional Euclidean space calibration algorithms.

The rest of the paper is organized as follows: In section 2 we give an overview of radio interferometric calibration. Next, in section 3, we present the geometric structure of calibration parameters. We present the Riemannian gradient and Hessian operators in section 4 for the calibration cost function. Simulation results are presented in section 5 where we apply the RTR method for calibration and finally, we draw our conclusions in section 6.

Notation: Matrices and vectors are denoted by bold upper and lower case letters as $\mathbf{J}$ and $\mathbf{v}$, respectively. The transpose and the Hermitian transpose are given by $(\cdot)^T$ and $(\cdot)^H$, respectively. The matrix Frobenius norm is given by $\|\cdot\|$. The set of real and complex numbers are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. The identity matrix is given by $\mathbf{I}$. The matrix trace operator is given by $\text{trace}(\cdot)$.

2. RADIO INTERFEROMETRIC CALIBRATION

In this section, we present radio interferometric calibration as an optimization problem. Consider a radio interferometric array with N receivers. The observed data at a baseline formed by two receivers, p and q is given by [13]

$$\mathbf{V}_{pq} = \mathbf{J}_p \mathbf{C}_{pq} \mathbf{J}_q^H + \mathbf{N}_{pq} \quad (1)$$

where $\mathbf{V}_{pq} (\in \mathbb{C}^{2 \times 2})$ is the observed *visibility* matrix (or the cross correlations). The errors that need to be calibrated for station p and q are given by the Jones matrices $\mathbf{J}_p, \mathbf{J}_q (\in \mathbb{C}^{2 \times 2})$, respectively. The sky signal (or *coherency*) is given by $\mathbf{C}_{pq} (\in \mathbb{C}^{2 \times 2})$ and is known a priori. The noise matrix $\mathbf{N}_{pq} (\in \mathbb{C}^{2 \times 2})$ is assumed to have complex, zero mean, circular Gaussian elements.

For an array with N receivers, we can form at most $N(N-1)/2$ baselines that collect visibilities as in (1). We rewrite (1) as

$$\mathbf{V}_{pq} = \mathbf{A}_p \mathbf{J} \mathbf{C}_{pq} \mathbf{J}^H \mathbf{A}_q^T + \mathbf{N}_{pq} \quad (2)$$

where $\mathbf{J} \in \mathbb{C}^{2N \times 2}$ is the augmented matrix of Jones matrices of all stations,

$$\mathbf{J} \triangleq [\mathbf{J}_1^T, \mathbf{J}_2^T, \dots, \mathbf{J}_N^T]^T \quad (3)$$

and $\mathbf{A}_p \in \mathbb{R}^{2 \times 2N}$ (and $\mathbf{A}_q$ likewise) is the canonical selection matrix

$$\mathbf{A}_p \triangleq [0, \mathbf{0}, \dots, \mathbf{I}, \dots, 0]. \quad (4)$$

In (4), all elements of $\mathbf{A}_p$ are zero except the p -th block which is an identity matrix.

Calibration is the estimation of $\mathbf{J}$ given the visibilities as in (1). Under a Gaussian noise model, the Maximum Likelihood estimate is

$$\hat{\mathbf{J}} = \arg \min_{\mathbf{J}} f(\mathbf{J}) \quad (5)$$

where the nonlinear cost function $f(\mathbf{J})$ is

$$f(\mathbf{J}) \triangleq \sum_{p,q} \|\mathbf{V}_{pq} - \mathbf{A}_p \mathbf{J} \mathbf{C}_{pq} \mathbf{J}^H \mathbf{A}_q^T\|^2. \quad (6)$$

The sky signal almost always has very little polarization and therefore, the coherencies $\mathbf{C}_{pq}$ in (1) are diagonal matrices. Therefore, for any unitary $\mathbf{U} \in \mathbb{C}^{2 \times 2}$, we see that $f(\mathbf{J}) = f(\mathbf{J}\mathbf{U})$. In other words, for any solution $\mathbf{J}$, a feasible solution for (5) would also be $\mathbf{J}\mathbf{U}$ where $\mathbf{U}$ is unitary. Currently, a solution for (5) is obtained by well known nonlinear optimization methods such as the Levenberg-Marquardt [1, 2] method and an in-depth overview of current calibration approaches can be found in e.g., [3, 14, 15].

3. GEOMETRIC STRUCTURE OF CALIBRATION

In this section, we present the manifold geometric structure of the parameters $\mathbf{J}$ used in radio interferometric calibration. A manifold can be described as a set of entities, together with a set of mappings (or charts) that can locally describe the manifold in Euclidean space. For a more formal introduction to matrix manifolds, the reader is referred to [5]. A problem very similar to what we consider in this section (involving real symmetric positive semi-definite matrices) can be found in [9] and we follow the same approach.

Given the solution to (5), i.e. $\mathbf{J}$, we know that $\mathbf{J}\mathbf{U}$ is also a feasible solution. We say $\mathbf{J}$ and $\mathbf{J}\mathbf{U}$ are *similar*, i.e.,

$$\mathbf{J} \sim \mathbf{J}\mathbf{U} \quad (7)$$

when $\mathbf{U}$ is any unitary matrix. Therefore, the whole set of feasible solutions $\mathbf{J}\mathbf{U}$ where $\mathbf{U}$ is any unitary matrix can be represented by one of its elements, $\mathbf{J}$. We consider $\overline{\mathcal{M}}$ to be the manifold of all $2N \times 2$ complex matrices ($\mathbb{C}^{2N \times 2}$). While the whole set of feasible solutions lie on $\overline{\mathcal{M}}$, using the *quotient* manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$ we can represent the whole set by a single point as shown in Fig. 1.

The mapping π (canonical projection) is defined such that any matrix $\mathbf{J}\mathbf{U}$ on $\overline{\mathcal{M}}$ is mapped onto a single point, $\pi(\mathbf{J})$ on $\mathcal{M}$. With this mapping, we define the equivalence class

$$\pi^{-1}(\pi(\mathbf{J})) \triangleq \{\mathbf{J}\mathbf{U} : \mathbf{U}\mathbf{U}^H = \mathbf{U}^H\mathbf{U} = \mathbf{I}, \mathbf{U} \in \mathbb{C}^{2 \times 2}\} \quad (8)$$

of solutions represented by a single point on $\mathcal{M}$. In order to make $\mathcal{M}$ a Riemannian manifold, we introduce the (smooth) inner product $g_{\mathcal{J}}(\cdot, \cdot)$ to its tangent space $\mathcal{T}_{\mathcal{J}}\mathcal{M}$ as

$$g_{\mathcal{J}}(\xi_{\mathcal{J}}, \eta_{\mathcal{J}}) \triangleq \text{trace}(\xi_{\mathcal{J}}^H \eta_{\mathcal{J}} + \eta_{\mathcal{J}}^H \xi_{\mathcal{J}}), \xi_{\mathcal{J}}, \eta_{\mathcal{J}} \in \mathcal{T}_{\mathcal{J}}\mathcal{M}. \quad (9)$$

With (9), we can decompose $\mathcal{T}_{\mathcal{J}}\mathcal{M}$ into two complementary vector spaces as

$$\mathcal{T}_{\mathcal{J}}\mathcal{M} = \mathcal{V}_{\mathcal{J}}\mathcal{M} \oplus \mathcal{H}_{\mathcal{J}}\mathcal{M} \quad (10)$$

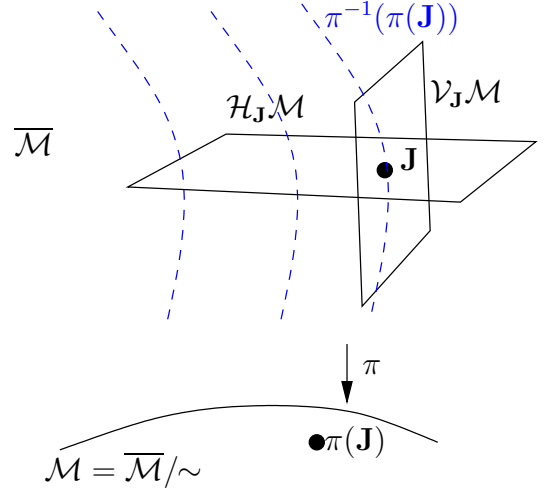


Fig. 1. The quotient manifold geometry of the calibration parameters. The dashed (blue) line (on $\overline{\mathcal{M}}$) represents the equivalence class of all solutions that are related to $\mathbf{J}$ by a unitary ambiguity. This equivalence class is represented by a single point on the quotient manifold $\mathcal{M} = \overline{\mathcal{M}}/\sim$. The vertical space $\mathcal{V}_{\mathcal{J}}\mathcal{M}$ is the vector space tangential to the equivalence class and the horizontal space $\mathcal{H}_{\mathcal{J}}\mathcal{M}$ is the orthogonal complement.

where $\oplus$ is the direct sum operator. We define the vertical space to be the directions tangential to the equivalence class at $\mathbf{J}$, i.e.,

$$\mathcal{V}_{\mathcal{J}}\mathcal{M} \triangleq \{\mathbf{J}\boldsymbol{\Omega} : \boldsymbol{\Omega}^H = -\boldsymbol{\Omega}, \boldsymbol{\Omega} \in \mathbb{C}^{2 \times 2}\} \quad (11)$$

and we choose the set of directions orthogonal to the equivalence class at $\mathbf{J}$ as the horizontal space $\mathcal{H}_{\mathcal{J}}\mathcal{M}$,

$$\mathcal{H}_{\mathcal{J}}\mathcal{M} \triangleq \{\xi_{\mathcal{J}} \in \mathbb{C}^{2N \times 2} : \xi_{\mathcal{J}}^H \mathbf{J} = \mathbf{J}^H \xi_{\mathcal{J}}\}. \quad (12)$$

The proof of (12) is easy to obtain: Let $\eta_{\mathcal{J}} = \mathbf{J}\boldsymbol{\Omega} \in \mathcal{V}_{\mathcal{J}}\mathcal{M}$ then by making $g_{\mathcal{J}}(\xi_{\mathcal{J}}, \eta_{\mathcal{J}}) = 0$, we get (12).

The projection of any direction $\mathbf{Z} \in \mathbb{C}^{2N \times 2}$ onto the horizontal space at $\mathbf{J}$ is given by

$$\Pi_{\mathcal{H}_{\mathcal{J}}\mathcal{M}}(\mathbf{Z}) \triangleq \mathbf{Z} - \mathbf{J}\boldsymbol{\Omega} \quad (13)$$

where $\boldsymbol{\Omega} \in \mathbb{C}^{2 \times 2}$ is skew-Hermitian and (because $\mathbf{Z} - \mathbf{J}\boldsymbol{\Omega} \in \mathcal{H}_{\mathcal{J}}\mathcal{M}$) satisfies the Sylvester equation

$$\boldsymbol{\Omega}^H \mathbf{J} + \mathbf{J}^H \boldsymbol{\Omega} = \mathbf{J}^H \mathbf{Z} - \mathbf{Z}^H \mathbf{J}. \quad (14)$$

A retraction is a mapping from $\mathcal{T}_{\mathcal{J}}\mathcal{M}$ to $\mathcal{M}$. There are many possible retractions but we choose a simple formula for the retraction as

$$R_{\mathcal{J}}(\xi_{\mathcal{J}}) \triangleq \mathbf{J} + \xi_{\mathcal{J}}. \quad (15)$$

We shall use the geometric structure of $\mathcal{M}$, in particular (13) and (15), in the following section to reformulate calibration.

4. CALIBRATION USING A RIEMANNIAN MANIFOLD

Rather than solving (5) in Euclidean space, we minimize the cost function $f(\mathbf{J})$ on $\mathcal{M}$. In order to do this, we need to compute the

Riemannian gradient and the Riemannian Hessian of $f(\mathbf{J})$. The Riemannian gradient $\text{grad}(f(\mathbf{J}))$ is the unique operator that satisfies

$$g_{\mathbf{J}}(\xi_{\mathbf{J}}, \text{grad}(f(\mathbf{J}))) = Df(\mathbf{J})[\xi_{\mathbf{J}}], \forall \xi_{\mathbf{J}} \in \mathcal{T}_{\mathbf{J}}\mathcal{M} \quad (16)$$

where,

$$Df(\mathbf{J})[\xi_{\mathbf{J}}] \triangleq \lim_{t \rightarrow 0} \frac{f(\mathbf{J} + t\xi_{\mathbf{J}}) - f(\mathbf{J})}{t}. \quad (17)$$

Using (6) and (9), we get

$$\begin{aligned} \text{grad}(f(\mathbf{J})) &= - \sum_{p,q} \left(\mathbf{A}_p^T (\mathbf{V}_{pq} - \mathbf{A}_p \mathbf{J} \mathbf{C}_{pq} \mathbf{J}^H \mathbf{A}_q^T) \mathbf{A}_q \mathbf{J} \mathbf{C}_{pq}^H \right. \\ &\quad \left. + \mathbf{A}_q^T (\mathbf{V}_{pq} - \mathbf{A}_p \mathbf{J} \mathbf{C}_{pq} \mathbf{J}^H \mathbf{A}_q^T)^H \mathbf{A}_p \mathbf{J} \mathbf{C}_{pq} \right) \end{aligned} \quad (18)$$

and the horizontal lift of $\text{grad}(f(\mathbf{J}))$ to $\mathcal{H}_{\mathbf{J}}\mathcal{M}$ is

$$\overline{\text{grad}(f(\mathbf{J}))} = \Pi_{\mathcal{H}_{\mathbf{J}}\mathcal{M}}(\text{grad}(f(\mathbf{J}))). \quad (19)$$

The Riemannian Hessian is defined as

$$\text{Hess}f(\mathbf{J})[\eta_{\mathbf{J}}] \triangleq \Pi_{\mathcal{H}_{\mathbf{J}}\mathcal{M}} \left(\lim_{t \rightarrow 0} \frac{1}{t} (\text{grad}f(\mathbf{J} + t\eta_{\mathbf{J}}) - \text{grad}f(\mathbf{J})) \right) \quad (20)$$

where

$$\begin{aligned} \lim_{t \rightarrow 0} \frac{1}{t} (\text{grad}f(\mathbf{J} + t\eta_{\mathbf{J}}) - \text{grad}f(\mathbf{J})) &= \\ &\sum_{p,q} \left(\mathbf{A}_p^T \left((\mathbf{V}_{pq} - \mathbf{A}_p \mathbf{J} \mathbf{C}_{pq} \mathbf{J}^H \mathbf{A}_q^T) \mathbf{A}_q \eta_{\mathbf{J}} \right. \right. \\ &\quad \left. \left. - \mathbf{A}_p (\mathbf{J} \mathbf{C}_{pq} \eta_{\mathbf{J}}^H + \eta_{\mathbf{J}} \mathbf{C}_{pq} \mathbf{J}^H) \mathbf{A}_q^T \mathbf{A}_q \mathbf{J} \right) \mathbf{C}_{pq}^H \right. \\ &\quad \left. \mathbf{A}_q^T \left((\mathbf{V}_{pq} - \mathbf{A}_p \mathbf{J} \mathbf{C}_{pq} \mathbf{J}^H \mathbf{A}_q^T)^H \mathbf{A}_p \eta_{\mathbf{J}} \right. \right. \\ &\quad \left. \left. - \mathbf{A}_q (\mathbf{J} \mathbf{C}_{pq} \eta_{\mathbf{J}}^H + \eta_{\mathbf{J}} \mathbf{C}_{pq} \mathbf{J}^H) \mathbf{A}_p^T \mathbf{A}_p \mathbf{J} \right) \mathbf{C}_{pq} \right). \end{aligned} \quad (21)$$

Note that for notational purposes we write products such as $\mathbf{A}_p \mathbf{J}$ in the above expressions but we do not actually form a matrix product because $\mathbf{A}_p$ -s are merely selection matrices.

With the Riemannian gradient and Hessian at hand, we apply the Riemannian trust-region method [8] to our problem. The trust-region method solves the problem

$$\min_{\eta_{\mathbf{J}} \in \mathcal{H}_{\mathbf{J}}\mathcal{M}} f(\mathbf{J}) + g_{\mathbf{J}}(\overline{\text{grad}(f(\mathbf{J}))}, \eta_{\mathbf{J}}) + \frac{1}{2} g_{\mathbf{J}}(\text{Hess}f(\mathbf{J})[\eta_{\mathbf{J}}], \eta_{\mathbf{J}})$$

subject to $g_{\mathbf{J}}(\eta_{\mathbf{J}}, \eta_{\mathbf{J}}) \leq \delta^2$, where δ is the trust-region radius.

The computational cost of the RTR method is significantly less compared with the LM method mainly due to the following reason. In the LM method, with N stations, the Jacobian is a matrix of size $8 \times N(N-1)/2$ by $8N$ with real entries. The multiplication of the transpose of the Jacobian with itself has cost $\mathcal{O}((8N)^2 4N(N-1))$ and the linear system solved is of size $8N$. On the other hand, in the RTR method, both the gradient and the Hessian are of size $2N \times 2$ with complex entries. Moreover, no full linear system is solved (since the truncated conjugate gradient method is used [8, 16]), except in solving (14), which is only a linear system of order 4.

5. SIMULATION RESULTS

In this section we compare the performance of the proposed calibration approach against conventional calibration. For conventional calibration, we consider two optimization algorithms: LM algorithm and Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm [16]. For the LM algorithm, we use closed form Jacobian calculation and for BFGS we use closed form gradient calculation (i.e. not using finite differences). We used the MATLAB implementation of the RTR method [17] in our simulations.

We simulate an array of N receivers where N is varied. The error matrices $\mathbf{J}_p, \mathbf{J}_q$ in (1) are generated with their elements having values drawn from a complex uniform distribution in $[0, 1]$ as $\mathcal{U}(0, 1) + j\mathcal{U}(0, 1)$. The sky signal is kept at unity, i.e. $\mathbf{C}_{pq} = \mathbf{I}$. The noise matrix $\mathbf{N}_{pq}$ is simulated to have complex circular Gaussian random variables. The variance of the noise is changed according to the signal to ratio (SNR)

$$\text{SNR} \triangleq \frac{\sum_{p,q} \|\mathbf{V}_{pq}\|^2}{\sum_{p,q} \|\mathbf{N}_{pq}\|^2}. \quad (22)$$

The initial values for the parameters are set as $\mathbf{J}_p = \mathbf{I}$ for $p \in [1, N]$. For the RTR method, the upper bound for the trust region radius $\bar{\delta}$ is chosen as

$$\bar{\delta} = \frac{1}{N} \sum_{p,q} \|\mathbf{V}_{pq}\|^2 \quad (23)$$

and the initial trust region radius is chosen as $\bar{\delta}/10$.

In Fig. 2, we show the reduction of the cost $f(\mathbf{J})$ for $N = 30$ and $\text{SNR} = 100$ for the three algorithms. The computing time was measured using a single Intel Xeon CPU core. It is evident that the RTR method takes significantly less time to reach the minimum cost. Furthermore, Fig. 2 shows that both three algorithms reach the minimum cost (i.e. they converge).

In the next simulation, we vary both N and the SNR. For each value of N , the SNR is changed to 50, 100, 150, and 200 and the computation time taken by each algorithm to reach convergence is measured. Once again, we use a single CPU core for the computations. The results are given in Fig. 3. In Fig. 3, we present the average computing time taken for all values of SNR. The superiority of the RTR method is once again highlighted in this figure.

The average residual error for all values of SNR (or the value of $f(\mathbf{J})$ at convergence) is shown in Fig. 4. It is clear that all three methods reach the same final cost at convergence.

6. CONCLUSIONS

We have presented the geometric structure in the form of a Riemannian quotient manifold that can be used in radio interferometric calibration. We have derived the Riemannian gradient and Hessian operators to minimize the cost function used in calibration. By employing the Riemannian trust-region method, we have proposed a computationally efficient calibration method. Based on simulation results, we have shown that the proposed calibration algorithm is much faster and also efficient in memory usage, compared with existing calibration algorithms that operate in Euclidean space.

7. REFERENCES

- [1] K. Levenberg, "A method for the solution of certain non linear problems using least squares," *The Quarterly Jnl. of App. Math.*, vol. 2, pp. 164–168, 1944.

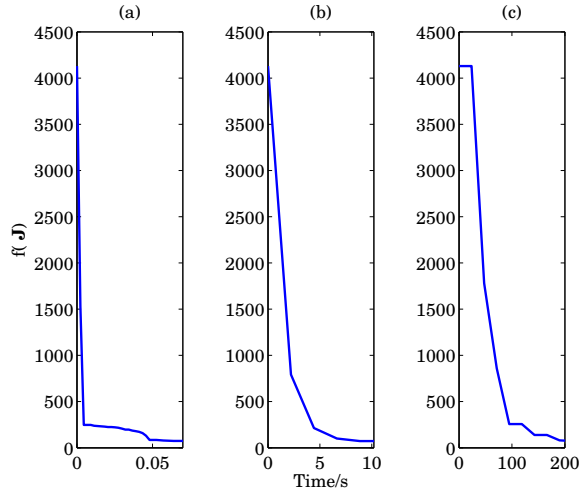


Fig. 2. Reduction of the cost function $f(\mathbf{J})$ with the time taken for the three optimization algorithms. (a) the RTR method (b) the LM method (c) the BFGS method. The number of stations $N = 30$ and the SNR = 100. It is clear that the RTR method uses much less time to reach the minimum cost while both other algorithms take significantly longer times.

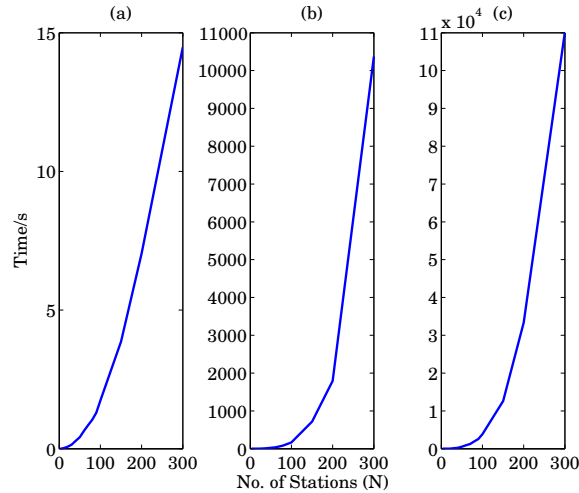


Fig. 3. Average computation time for the three algorithms for various values of N . (a) the RTR method (b) the LM method (c) the BFGS method. The noise is varied with SNR = 50, 100, 150, and 200 for each value of N . The RTR method takes significantly less time than the other two algorithms.

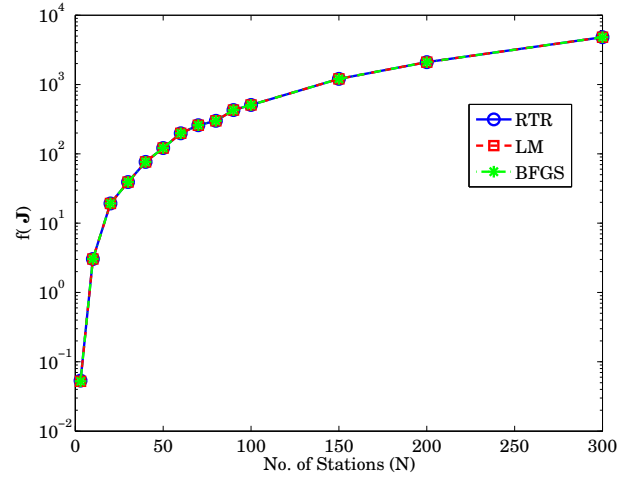


Fig. 4. Average residual error (or the value of $f(\mathbf{J})$ at convergence) for the three algorithms for various values of N . The noise is varied with SNR = 50, 100, 150, and 200 for each value of N . All three methods attain exactly the same values for $f(\mathbf{J})$ at convergence.

- [2] D. Marquardt, "An algorithm for least squares estimation of nonlinear parameters," *SIAM Jnl. of App. Math.*, vol. 11, pp. 431–441, 1963.
- [3] S. Yatawatta, S. Kazemi, and S. Zaroubi, "GPU accelerated nonlinear optimization in radio interferometric calibration," in *Proceedings of Innovative Parallel Computing (InPar 2012)*, May 2012, pp. 1–6.
- [4] J. P. Hamaker, "Understanding Radio Polarimetry IV: The full-coherency analogue of scalar selfcalibration," *Astronomy and Astrophysics Supp.*, vol. 143, no. 3, pp. 515–534, 2000.
- [5] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton Univ. Press, Princeton NJ, 2008.
- [6] S. T. Smith, "Covariance, subspace, and intrinsic Cramer Rao bounds," *IEEE Trans. on Sig. Proc.*, vol. 53, no. 5, pp. 1610–1630, May 2005.
- [7] S. Yatawatta, "On the interpolation of calibration solutions obtained in radio interferometry," *Monthly Notices of the Royal Astronomical Society*, p. 33, Oct. 2012.
- [8] P.-A. Absil, C. G. Baker, and K. A. Gallivan, "Trust-region methods on Riemannian manifolds," *Found. Comput. Math.*, vol. 7, no. 3, pp. 303–330, July 2007.
- [9] B. Mishra, G. Meyer, and R. Sepulchre, "Low-rank optimization for distance matrix completion," in *proc. 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, Orlando, FL, Dec. 2011.
- [10] P.-A. Absil, M. Ishteva, L. Lathauwer, and S. van Huffel, "A geometric Newton method for Oja's vector field," *Neural Computation*, vol. 21 no. 5, pp. 1415–1433, 2009.
- [11] B. Mishra, G. Meyer, S. Bonnabel, and R. Sepulchre, "Fixed-rank factorizations and Riemannian low-rank optimization," *ArXiv e-prints*, Sept. 2012.

- [12] S. Kazemi, S. Yatawatta, and S. Zaroubi, "Ordered-subsets acceleration of radio interferometric calibration: OS-SAGE calibration algorithm," in *proc. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2012)*, Ho Chi Minh City, Vietnam, Dec. 2012.
- [13] J. P. Hamaker, J. D. Bregman, and R. J. Sault, "Understanding radio polarimetry, paper I," *Astronomy and Astrophysics Supp.*, vol. 117, no. 137, pp. 96–109, 1996.
- [14] A.-J. Boonstra and A.-J. van der Veen, "Gain calibration methods for radio telescope arrays," *Signal Processing, IEEE Transactions on*, vol. 51, no. 1, pp. 25–38, Jan.
- [15] S. van Der Tol, B.D. Jeffs, and A.-J. van der Veen, "Self-calibration for the lofar radio astronomical array," *Signal Processing, IEEE Transactions on*, vol. 55, no. 9, pp. 4497–4510, Sept.
- [16] J. Nocedal and S. J. Wright, *Numerical Optimization*, New York USA:Springer, 1999.
- [17] P.-A. Absil, C.G. Baker, and K.A. Gallivan, "GenRTR Riemannian Optimization Package," <http://www.math.fsu.edu/~cbaker/genrtr/>, 2007.