Robust Calibration of Radio Interferometers in Non-Gaussian Environment

V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal, M. Pesavento

SATIE - L2S
Université Paris-Saclay

LEME
Université Paris-Ouest

TU-D
Technische Universität Darmstadt

CEA Saclay

MAGELLAN

Outline

- **Introduction to radio astronomy and motivation of this work**
- **Data model**
  - Non-structured Jones matrices
  - Structured Jones matrices (3DC calibration regime)
- **Estimation procedure for non-structured Jones matrices**
  - Principle of the proposed calibration technique
  - Use of the EM algorithm
  - Use of the BCD algorithm
- **Estimation procedure for structured Jones matrices (3DC calibration regime)**
- **Numerical simulations**
  - Under SIRP noise assumption
  - Under realistic model
  - Under 3DC calibration regime
- **Conclusion**

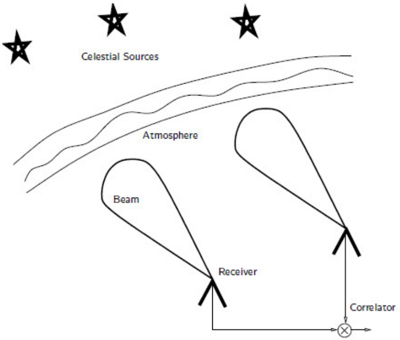
Van der Veen et al. (2013)

- **Goal**
  - Measure **electromagnetic waves** impinging on the Earth thanks to astronomical instruments
  - Deduce **spectral flux density** from maps which measure strength of radiation
  - Study **physical phenomena** to handle cosmological issues
- **Astronomical instruments**
  - Large telescope dishes: **expensive**, lack of flexibility
  - **Interferometric array**: higher angular resolution
  - **Phased-array system**: cheaper dipole antennas, no moving parts (software telescope), huge collecting area

- **Radio astronomy**
  - Study **radio emissions**
  - Case of **LOFAR** (LOw Frequency ARray):
    LBA (Low Band Antenna) 30-80 MHz
    HBA (High Band Antenna) 120-240 MHz

- **Calibration**
  - Estimation of all **perturbations** introduced along the signal path
  - Essential to perform **image reconstruction** with no distorsions

Hamaker et al. (1996)
Kazemi and Yatawatta (2013)



- **Non-Gaussianity assumption**
  - Presence of **outliers** in the data (weak unknown sources, RFI, ...)
  - Robust calibration in the literature: only **Student's t**

Data model

Non-structured Jones matrices

Mathematical model for non-structured Jones matrices

Hamaker et al. (1996), Smirnov (2011)

▷ D sources, M antennas, 2 orthogonal polarization directions (x, y)

- $\mathbf{s}_i = [s_{i_x}, s_{i_y}]^T$ i-th **incoming radiation**
- $\bar{\mathbf{v}}_{i_p}(\boldsymbol{\theta}) = [v_{i_{px}}(\boldsymbol{\theta}), v_{i_{py}}(\boldsymbol{\theta})]^T$ **generated voltage** at p-th antenna
- $\mathbf{J}_{i_p}(\boldsymbol{\theta})$ **Jones** matrix, **parametrized** by unknown vector $\boldsymbol{\theta}$

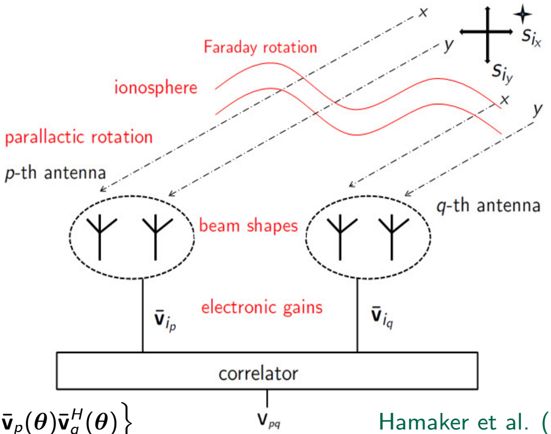
$$\Downarrow$$

$$\boxed{\bar{\mathbf{v}}_{i_p}(\boldsymbol{\theta}) = \mathbf{J}_{i_p}(\boldsymbol{\theta})\mathbf{s}_i}$$

$\Longrightarrow$ Unknown elements = **entries of all Jones matrices**

Noise-free case

$$\mathbf{V}_{pq}(\boldsymbol{\theta}) = \sum_{i=1}^{D} \mathbf{J}_{i_p}(\boldsymbol{\theta})\mathbf{C}_i\mathbf{J}_{i_q}^H(\boldsymbol{\theta}) \quad \text{for} \quad p < q, \quad p, q \in \{1, \ldots, M\}$$



- $\mathbf{V}_{pq}(\boldsymbol{\theta}) = \mathrm{E}\left\{\bar{\mathbf{v}}_p(\boldsymbol{\theta})\bar{\mathbf{v}}_q^H(\boldsymbol{\theta})\right\}$ Hamaker et al. (1996)
- $\mathbf{C}_i = \mathrm{E}\{\mathbf{s}_i\mathbf{s}_i^H\} = \begin{bmatrix} I_i + Q_i & U_i + jV_i \\ U_i - jV_i & I_i - Q_i \end{bmatrix}$ **intrinsic source coherency** matrix

Full visibility vector

$$\mathbf{V}_{pq}(\boldsymbol{\theta}) = \sum_{i=1}^{D} \mathbf{J}_{i_p}(\boldsymbol{\theta})\mathbf{C}_i \mathbf{J}_{i_q}^H(\boldsymbol{\theta})$$

$$\Downarrow$$

$$\tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta}) = \mathrm{vec}\Big(\mathbf{V}_{pq}(\boldsymbol{\theta})\Big) = \sum_{i=1}^{D} \mathbf{u}_{i_{pq}}(\boldsymbol{\theta})$$

- $\mathbf{u}_{i_{pq}}(\boldsymbol{\theta}) = \Big(\mathbf{J}_{i_q}^{\star}(\boldsymbol{\theta}) \otimes \mathbf{J}_{i_p}(\boldsymbol{\theta})\Big)\,\mathrm{vec}(\mathbf{C}_i)$

Noisy correlation measurements

$$\mathbf{v}_{pq} = \tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta}) + \mathbf{n}_{pq}$$

$$\Longrightarrow \quad \boxed{\mathbf{x} = [\mathbf{v}_{12}^T, \mathbf{v}_{13}^T, \ldots, \mathbf{v}_{(M-1)M}^T]^T = \sum_{i=1}^{D} \mathbf{u}_i(\boldsymbol{\theta}) + \mathbf{n}}$$

- $\mathbf{u}_i(\boldsymbol{\theta}) = \Big[\mathbf{u}_{i_{12}}^T(\boldsymbol{\theta}), \mathbf{u}_{i_{13}}^T(\boldsymbol{\theta}), \ldots, \mathbf{u}_{i_{(M-1)M}}^T(\boldsymbol{\theta})\Big]^T$
- $\mathbf{n} = \Big[\mathbf{n}_{12}^T, \mathbf{n}_{13}^T, \ldots, \mathbf{n}_{(M-1)M}^T\Big]^T \Longrightarrow$ **Gaussian** noise & **outliers**

Noise modeling

Spherically invariant random process (SIRP)

$$\mathbf{n}_{pq} = \sqrt{\tau_{pq}}\, \mathbf{g}_{pq}$$

- τ_{pq} positive real random variable (**texture**)
- $\mathbf{g}_{pq}$ complex zero-mean Gaussian process (**speckle**)

$$\mathbf{g}_{pq} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Omega}) \text{ s.t. } \operatorname{tr}\{\mathbf{\Omega}\} = 1$$

$\Longrightarrow$ Remove **scaling ambiguities**

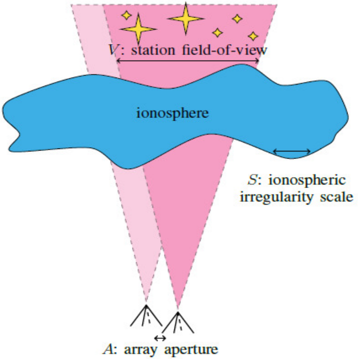
Data model

Structured Jones matrices

3DC calibration regime

Direction dependent distorsions with compact array

Lonsdale (2004)



- **Closely** packed group of **similar** antennas
- **Wide field of view** of individual elements

Noordam (1996), Smirnov (2011), Yatawatta (2012)

$$\mathbf{J}_{i_p}(\boldsymbol{\theta}_{i_p}^{\mathrm{3DC}}) = \mathbf{G}_p(\mathbf{g}_p)\mathbf{H}_{i_p}\mathbf{Z}_{i_p}(\boldsymbol{\alpha}_i)\mathbf{F}_i(\vartheta_i)$$

with $\boldsymbol{\theta}_{i_p}^{\mathrm{3DC}} = [\vartheta_i, \mathbf{g}_p^T, \boldsymbol{\alpha}_i^T]^T$

- **Ionospheric delay** matrix

$$\mathbf{Z}_{i_p}(\boldsymbol{\alpha}_i) = \exp\left\{j\varphi_{i_p}\right\}\mathbf{I}_2$$

where $\varphi_{i_p} = \eta_i u_p + \zeta_i v_p$,

- $\boldsymbol{\alpha}_i = [\eta_i, \zeta_i]^T$ source offset
- $\mathbf{r}_p = [u_p, v_p]^T$ known antenna position in units of wavelength

- **Diagonal electronic gain** matrix

$$\mathbf{G}_p(\mathbf{g}_p) = \mathrm{diag}\{\mathbf{g}_p\}$$

- **Known** matrix $\mathbf{H}_{i_p}$
  - Electromagnetic simulations
  - A priori knowledge (calibrator sources & antenna positions)

- **Ionospheric Faraday rotation** matrix

$$\mathbf{F}_i(\vartheta_i) = \begin{bmatrix} \cos(\vartheta_i) & -\sin(\vartheta_i) \\ \sin(\vartheta_i) & \cos(\vartheta_i) \end{bmatrix}$$

  - Faraday rotation angle ϑ_i

Estimation procedure

Non-structured Jones matrices

Maximum likelihood method

$$f(\mathbf{x}|\boldsymbol{\theta},\boldsymbol{\tau},\boldsymbol{\Omega}) = \prod_{pq} \frac{1}{|\pi \tau_{pq} \boldsymbol{\Omega}|} \exp\left\{-\frac{1}{\tau_{pq}} \mathbf{a}_{pq}^H(\boldsymbol{\theta}) \boldsymbol{\Omega}^{-1} \mathbf{a}_{pq}(\boldsymbol{\theta})\right\}$$

- $\boldsymbol{\tau} = \left[\tau_{12}, \tau_{13}, \ldots, \tau_{(M-1)M}\right]^T$
- $\mathbf{a}_{pq}(\boldsymbol{\theta}) = \mathbf{v}_{pq} - \tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta})$

$$\Downarrow$$

$$\log f(\mathbf{x}|\boldsymbol{\theta},\boldsymbol{\tau},\boldsymbol{\Omega}) =$$

$$-4B\log\pi - 4\sum_{pq}\log\tau_{pq} - B\log|\boldsymbol{\Omega}| - \sum_{pq}\frac{1}{\tau_{pq}}\mathbf{a}_{pq}^H(\boldsymbol{\theta})\boldsymbol{\Omega}^{-1}\mathbf{a}_{pq}(\boldsymbol{\theta})$$

Iterative ML algorithm

▷ Optimization w.r.t. each unknown parameter: **concentrated ML** estimator

▷ Probability density function of all τ_{pq} not specified: **relaxed ML** estimator

Proposed algorithm

1. **Initialize** $\hat{\mathbf{\Omega}} = \mathbf{\Omega}_{\text{init}}, \hat{\boldsymbol{\tau}} = \boldsymbol{\tau}_{\text{init}}$
2. **Estimation of θ**

$$\hat{\theta} = \underset{\theta}{\text{argmin}} \left\{ \sum_{pq} \frac{1}{\hat{\tau}_{pq}} \mathbf{a}_{pq}^H(\theta) \hat{\mathbf{\Omega}}^{-1} \mathbf{a}_{pq}(\theta) \right\}$$

3. **Estimation of Ω**

$$\hat{\mathbf{\Omega}} = \frac{4}{B} \sum_{pq} \frac{\mathbf{a}_{pq}(\hat{\theta}) \mathbf{a}_{pq}^H(\hat{\theta})}{\mathbf{a}_{pq}^H(\hat{\theta}) (\hat{\mathbf{\Omega}})^{-1} \mathbf{a}_{pq}(\hat{\theta})}$$

$$\hat{\mathbf{\Omega}} = \frac{\hat{\mathbf{\Omega}}}{\operatorname{tr}\{\hat{\mathbf{\Omega}}\}}$$

'

4. **Estimation of τ**

$$\hat{\tau}_{pq} = \frac{1}{4} \mathbf{a}_{pq}^H(\hat{\theta}) \hat{\mathbf{\Omega}}^{-1} \mathbf{a}_{pq}(\hat{\theta})$$

5. **Repeat steps 2 to 4 until stop criterion reached**

Non-structured Jones matrices

- **Partition per source**

$$\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, \dots, \boldsymbol{\theta}_D^T]^T = [\boldsymbol{\theta}_{1_1}^T, \dots, \boldsymbol{\theta}_{1_M}^T, \dots, \boldsymbol{\theta}_{D_1}^T, \dots, \boldsymbol{\theta}_{D_M}^T]^T$$

- $\mathbf{J}_{i_p}(\boldsymbol{\theta})$ parametrized by **path from i-th calibrator source to p-th sensor**

$$\mathbf{J}_{i_p}(\boldsymbol{\theta}) = \mathbf{J}_{i_p}(\boldsymbol{\theta}_{i_p})$$

with $\boldsymbol{\theta}_{i_p} \in \mathbb{R}^{8\times 1}$

EM algorithm

Expectation-Maximization

Yatawatta et al. (2009), Kazemi et al. (2011)

▷ **Motivation**

- Decrease computational cost
  Optimization for single source problems of smaller dimensions
  $\Longrightarrow$ optimization w.r.t. $\boldsymbol{\theta}_i \in \mathbb{C}^{4M\times 1}$ instead of $\boldsymbol{\theta} \in \mathbb{C}^{4DM\times 1}$
- Ensure convergence

Complete data vector $\mathbf{w} = [\mathbf{w}_1^T, \dots, \mathbf{w}_D^T]^T$

$$\boxed{\mathbf{w}_i = \mathbf{u}_i(\boldsymbol{\theta}_i) + \mathbf{n}_i \quad \text{s.t.} \quad \mathbf{x} = \sum_{i=1}^{D} \mathbf{w}_i}$$

▷ **E-step**

Goal: conditional expectation of complete data $\hat{\mathbf{w}} = \mathrm{E}\{\mathbf{w}|\mathbf{x};\boldsymbol{\theta},\boldsymbol{\tau},\boldsymbol{\Omega}\}$

- $\mathbf{n} = \sum_{i=1}^{D} \mathbf{n}_i \sim \mathcal{CN}(\mathbf{0},\boldsymbol{\Psi})$

$$\boldsymbol{\Psi} = \begin{bmatrix} \tau_{12}\boldsymbol{\Omega} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \tau_{(M-1)M}\boldsymbol{\Omega} \end{bmatrix}$$

- $\mathbf{n}_i \sim \mathcal{CN}(\mathbf{0},\beta_i\boldsymbol{\Psi})$
- $\sum_{i=1}^{D} \beta_i = 1$

$$\boxed{\hat{\mathbf{w}}_i = \mathrm{E}\{\mathbf{w}_i|\mathbf{x};\boldsymbol{\theta},\boldsymbol{\tau},\boldsymbol{\Omega}\} = \mathbf{u}_i(\boldsymbol{\theta}_i) + \beta_i\left(\mathbf{x} - \sum_{l=1}^{D}\mathbf{u}_l(\boldsymbol{\theta}_l)\right)}$$

▷ **M-step**

Goal: estimation of $\boldsymbol{\theta}_i$

$$f(\hat{\mathbf{w}}|\boldsymbol{\theta},\boldsymbol{\tau},\boldsymbol{\Omega}) = \prod_{i=1}^{D} \frac{1}{|\pi\beta_i\boldsymbol{\Psi}|} \exp\left\{-\left(\hat{\mathbf{w}}_i - \mathbf{u}_i(\boldsymbol{\theta}_i)\right)^H (\beta_i\boldsymbol{\Psi})^{-1}\left(\hat{\mathbf{w}}_i - \mathbf{u}_i(\boldsymbol{\theta}_i)\right)\right\}$$

$$\boxed{\phi_i(\boldsymbol{\theta}_i) = \left(\hat{\mathbf{w}}_i - \mathbf{u}_i(\boldsymbol{\theta}_i)\right)^H (\beta_i\boldsymbol{\Psi})^{-1}\left(\hat{\mathbf{w}}_i - \mathbf{u}_i(\boldsymbol{\theta}_i)\right)}$$

- Numerical example: **Levenberg-Marquardt (LM)** algorithm
- Analytical method: **Block Coordinate Descent (BCD)** algorithm

BCD algorithm

Block Coordinate Descent

Friedman et al. (2007), Hong et al. (2016)

- Perform optimization of ϕ_i w.r.t. $\boldsymbol{\theta}_{i_p} \in \mathbb{C}^{4\times 1}$ with fixed $\boldsymbol{\theta}_{i_q}$, $q \neq p$

$$\phi_i(\boldsymbol{\theta}_{i_p}) = \sum_{\substack{q=1\\q>p}}^{M} \left(\mathbf{w}_{i_{pq}} - \mathbf{u}_{i_{pq}}(\boldsymbol{\theta}_{i_p})\right)^H (\beta_i \tau_{pq} \boldsymbol{\Omega})^{-1} \left(\mathbf{w}_{i_{pq}} - \mathbf{u}_{i_{pq}}(\boldsymbol{\theta}_{i_p})\right) + \sum_{\substack{q=1\\q<p}}^{M} \left(\mathbf{w}_{i_{qp}} - \mathbf{u}_{i_{qp}}(\boldsymbol{\theta}_{i_p})\right)^H (\beta_i \tau_{qp} \boldsymbol{\Omega})^{-1} \left(\mathbf{w}_{i_{qp}} - \mathbf{u}_{i_{qp}}(\boldsymbol{\theta}_{i_p})\right) + \mathrm{Cst}$$

BCD algorithm

- Estimate the Jones matrix associated with **path from i-th calibrator source to p-th sensor**

$$\hat{\theta}_{i_p} = \begin{cases} (\mathbf{\Sigma}_i^H \mathbf{A}_{i_p} \mathbf{\Sigma}_i + \mathbf{\Upsilon}_i^H \tilde{\mathbf{A}}_{i_p} \mathbf{\Upsilon}_i)^{-1} (\mathbf{\Sigma}_i^H \mathbf{A}_{i_p} \mathbf{w}_{i_p} + \mathbf{\Upsilon}_i^H \tilde{\mathbf{A}}_{i_p} \tilde{\mathbf{w}}_{i_p}) & \text{for } 1 < p < M \\ (\mathbf{\Sigma}_i^H \mathbf{A}_{i_p} \mathbf{\Sigma}_i)^{-1} \mathbf{\Sigma}_i^H \mathbf{A}_{i_p} \mathbf{w}_{i_p} & \text{for } p = 1 \\ (\mathbf{\Upsilon}_i^H \tilde{\mathbf{A}}_{i_p} \mathbf{\Upsilon}_i)^{-1} \mathbf{\Upsilon}_i^H \tilde{\mathbf{A}}_{i_p} \tilde{\mathbf{w}}_{i_p} & \text{for } p = M \end{cases}$$

  - $\mathbf{c}_i = \mathrm{vec}(\mathbf{C}_i) = [c_{i_1}, c_{i_2}, c_{i_3}, c_{i_4}]^T$
  - $\mathbf{J}_{i_p}(\boldsymbol{\theta}_{i_p}) = \begin{bmatrix} p_{i_1} & p_{i_2} \\ p_{i_3} & p_{i_4} \end{bmatrix}$ and $\mathbf{J}_{i_q}(\boldsymbol{\theta}_{i_q}) = \begin{bmatrix} q_{i_1} & q_{i_2} \\ q_{i_3} & q_{i_4} \end{bmatrix}$, $\boldsymbol{\theta}_{i_p} = [p_{i_1}, p_{i_2}, p_{i_3}, p_{i_4}]^T$ and $\boldsymbol{\theta}_{i_q} = [q_{i_1}, q_{i_2}, q_{i_3}, q_{i_4}]^T$
  - $\mathbf{\Sigma}_{i_q} = \begin{bmatrix} \alpha_{i_q} & \beta_{i_q} & 0 & 0 \\ 0 & 0 & \alpha_{i_q} & \beta_{i_q} \\ \gamma_{i_q} & \rho_{i_q} & 0 & 0 \\ 0 & 0 & \gamma_{i_q} & \rho_{i_q} \end{bmatrix}$ and $\mathbf{\Upsilon}_{i_q} = \begin{bmatrix} \lambda_{i_q} & \mu_{i_q} & 0 & 0 \\ \nu_{i_q} & \xi_{i_q} & 0 & 0 \\ 0 & 0 & \lambda_{i_q} & \mu_{i_q} \\ 0 & 0 & \nu_{i_q} & \xi_{i_q} \end{bmatrix}$
  - $\alpha_{i_q} = q_{i_1}^* c_{i_1} + q_{i_2}^* c_{i_3}$, $\beta_{i_q} = q_{i_1}^* c_{i_2} + q_{i_2}^* c_{i_4}$, $\gamma_{i_q} = q_{i_3}^* c_{i_1} + q_{i_4}^* c_{i_3}$ and $\rho_{i_q} = q_{i_3}^* c_{i_2} + q_{i_4}^* c_{i_4}$
  - $\lambda_{i_q} = q_{i_1} c_{i_1} + q_{i_2} c_{i_2}$, $\mu_{i_q} = q_{i_1} c_{i_3} + q_{i_2} c_{i_4}$, $\nu_{i_q} = q_{i_3} c_{i_1} + q_{i_4} c_{i_2}$ and $\xi_{i_q} = q_{i_3} c_{i_3} + q_{i_4} c_{i_4}$
  - $\mathbf{w}_{i_p} = [\mathbf{w}_{i_{p(p+1)}}^T, \dots, \mathbf{w}_{i_{pM}}^T]^T$ and $\mathbf{A}_{i_p} = \mathrm{bdiag}\{\beta_i \tau_{p(p+1)} \mathbf{\Omega}, \dots, \beta_i \tau_{pM} \mathbf{\Omega}\}^{-1}$
  - $\tilde{\mathbf{w}}_{i_p} = [\mathbf{w}_{i_{1p}}^{*\,T}, \dots, \mathbf{w}_{i_{(p-1)p}}^{*\,T}]^T$ and $\tilde{\mathbf{A}}_{i_p} = \mathrm{bdiag}\{\beta_i \tau_{1p} \mathbf{\Omega}^*, \dots, \beta_i \tau_{(p-1)p} \mathbf{\Omega}^*\}^{-1}$
  - $\mathbf{\Sigma}_i = [\mathbf{\Sigma}_{i_{p+1}}^T, \cdots, \mathbf{\Sigma}_{i_M}^T]^T$ and $\mathbf{\Upsilon}_i = [\mathbf{\Upsilon}_{i_1}^{*\,T}, \cdots, \mathbf{\Upsilon}_{i_{p-1}}^{*\,T}]^T$

Algorithm 1

1. **Initialize** $\hat{\boldsymbol{\Omega}} = \boldsymbol{\Omega}_{\text{init}}, \hat{\boldsymbol{\tau}} = \boldsymbol{\tau}_{\text{init}}, \hat{\boldsymbol{\theta}}_i = \boldsymbol{\theta}_{i_{\text{init}}}, i = 1, \ldots, D$
2. **EM algorithm** (for $i = 1, \ldots, D$) to repeat until stop criterion reached

   E-step: Estimation of $\mathbf{w}_i$

$$\hat{\mathbf{w}}_i = \mathbf{u}_i(\hat{\boldsymbol{\theta}}_i) + \beta_i \left(\mathbf{x} - \sum_{l=1}^{D} \mathbf{u}_l(\hat{\boldsymbol{\theta}}_l) \right)$$

   M-step: Estimation of each $\boldsymbol{\theta}_{i_p}$ for $p = 1, \ldots, M$ (to repeat iteratively)

$$\hat{\boldsymbol{\theta}}_{i_p} = \begin{cases} (\boldsymbol{\Sigma}_i^H \mathbf{A}_{i_p} \boldsymbol{\Sigma}_i + \boldsymbol{\Upsilon}_i^H \tilde{\mathbf{A}}_{i_p} \boldsymbol{\Upsilon}_i)^{-1} (\boldsymbol{\Sigma}_i^H \mathbf{A}_{i_p} \mathbf{w}_{i_p} + \boldsymbol{\Upsilon}_i^H \tilde{\mathbf{A}}_{i_p} \tilde{\mathbf{w}}_{i_p}) & \text{for } 1 < p < M \\ (\boldsymbol{\Sigma}_i^H \mathbf{A}_{i_p} \boldsymbol{\Sigma}_i)^{-1} \boldsymbol{\Sigma}_i^H \mathbf{A}_{i_p} \mathbf{w}_{i_p} & \text{for } p = 1 \\ (\boldsymbol{\Upsilon}_i^H \tilde{\mathbf{A}}_{i_p} \boldsymbol{\Upsilon}_i)^{-1} \boldsymbol{\Upsilon}_i^H \tilde{\mathbf{A}}_{i_p} \tilde{\mathbf{w}}_{i_p} & \text{for } p = M \end{cases}$$

3. **Estimation of $\boldsymbol{\Omega}$ and $\boldsymbol{\tau}$**

$$\hat{\boldsymbol{\Omega}} = \frac{4}{B} \sum_{pq} \frac{\mathbf{a}_{pq}(\hat{\boldsymbol{\theta}}) \mathbf{a}_{pq}^H(\hat{\boldsymbol{\theta}})}{\mathbf{a}_{pq}^H(\hat{\boldsymbol{\theta}}) (\hat{\boldsymbol{\Omega}})^{-1} \mathbf{a}_{pq}(\hat{\boldsymbol{\theta}})} \qquad \hat{\boldsymbol{\Omega}} = \frac{\hat{\boldsymbol{\Omega}}}{\text{tr}\{\hat{\boldsymbol{\Omega}}\}} \qquad \hat{\tau}_{pq} = \frac{1}{4} \mathbf{a}_{pq}^H(\hat{\boldsymbol{\theta}}) \hat{\boldsymbol{\Omega}}^{-1} \mathbf{a}_{pq}(\hat{\boldsymbol{\theta}})$$

4. **Repeat steps 2 to 3 until stop criterion reached**

Estimation procedure

Structured Jones matrices

Structured Jones matrices

- **Estimation of the antenna gains**

$$\hat{\mathbf{g}}_p = \underset{\mathbf{g}_p}{\operatorname{argmin}} \sum_{i=1}^{D} ||\hat{\mathbf{J}}_{i_p} - \mathbf{G}_p(\mathbf{g}_p)\mathbf{H}_{i_p}\mathbf{Z}_{i_p}\mathbf{F}_i||_F^2$$

- **Estimation of the source shifts due to the ionosphere**

$$\hat{\varphi}_{i_p} = \underset{\varphi_{i_p}}{\operatorname{argmin}} ||\hat{\mathbf{J}}_{i_p} - \mathbf{G}_p\mathbf{H}_{i_p}\mathbf{Z}_{i_p}(\varphi_{i_p})\mathbf{F}_i||_F^2$$

with $\varphi_{i_p} = \eta_i u_p + \zeta_i v_p$ and $\boldsymbol{\alpha}_i = [\eta_i, \zeta_i]^T$

- **Estimation of the Faraday rotation angle**

$$\hat{\vartheta}_i = \underset{\vartheta_i}{\operatorname{argmin}} \sum_{p=1}^{M} ||\hat{\mathbf{J}}_{i_p} - \mathbf{G}_p\mathbf{H}_{i_p}\mathbf{Z}_{i_p}\mathbf{F}_i(\vartheta_i)||_F^2$$

Algorithm 2

1. **Initialize** $\hat{\vartheta}_i = \vartheta_{i_{\text{init}}}$, $\hat{\boldsymbol{\alpha}}_i = \boldsymbol{\alpha}_{i_{\text{init}}}$, $i = 1, \ldots, D$, $\hat{\mathbf{g}}_p = \mathbf{g}_{p_{\text{init}}}$, $p = 1, \ldots, M$
2. **Estimation of** ϑ_i, $i = 1, \ldots, D$

$$\hat{\vartheta}_i = \text{argmin}_{\vartheta_i} \sum_{p=1}^{M} ||\hat{\mathbf{J}}_{i_p} - \hat{\mathbf{G}}_p \mathbf{H}_{i_p} \hat{\mathbf{Z}}_{i_p} \mathbf{F}_i(\vartheta_i)||_F^2$$

3. **Estimation of** $\mathbf{g}_p$, $p = 1, \ldots, M$

$$[\hat{\mathbf{g}}_p]_k = \left(\sum_{i=1}^{D} [\hat{\mathbf{W}}_{i_p}^*]_{k,k} \right)^{-1} \sum_{i=1}^{D} [\hat{\mathbf{X}}_{i_p}^*]_{k,k}$$

for $k \in \{1, 2\}$, with $\hat{\mathbf{X}}_{i_p} = \hat{\mathbf{R}}_{i_p} \hat{\mathbf{J}}_{i_p}^H$, $\hat{\mathbf{W}}_{i_p} = \hat{\mathbf{R}}_{i_p} \hat{\mathbf{R}}_{i_p}^H$ and $\hat{\mathbf{R}}_{i_p} = \mathbf{H}_{i_p} \hat{\mathbf{Z}}_{i_p} \hat{\mathbf{F}}_i$

4. **Estimation of** $\boldsymbol{\alpha}_i = [\eta_i, \zeta_i]^T$, $i = 1, \ldots, D$

$$\hat{\boldsymbol{\alpha}}_i^T = \frac{\hat{\boldsymbol{\varphi}}_i^T \boldsymbol{\Lambda}^H \begin{bmatrix} \sum_{p=1}^{M} v_p^2 & -\sum_{p=1}^{M} u_p v_p \\ -\sum_{p=1}^{M} v_p u_p & \sum_{p=1}^{M} u_p^2 \end{bmatrix}}{\sum_{p=1}^{M} u_p^2 \sum_{p=1}^{M} v_p^2 - (\sum_{p=1}^{M} u_p v_p)^2}$$

with $\hat{\boldsymbol{\varphi}}_i = [\hat{\varphi}_{i_1}, \ldots, \hat{\varphi}_{i_M}]^T$, $\boldsymbol{\Lambda} = \begin{bmatrix} u_1 & \ldots & u_M \\ v_1 & \ldots & v_M \end{bmatrix}$, $\exp\left\{2j\hat{\varphi}_{i_p}\right\} = \frac{\mathrm{Tr}\left\{\hat{\mathbf{M}}_{i_p}\right\}}{\mathrm{Tr}\left\{\hat{\mathbf{M}}_{i_p}^H\right\}}$ and

$\hat{\mathbf{M}}_{i_p} = \hat{\mathbf{J}}_{i_p} \hat{\mathbf{F}}_i^H \mathbf{H}_{i_p}^H \hat{\mathbf{G}}_p^H$

5. **Repeat steps 2 to 4 until stop criterion reached**

Numerical simulations

Besson et al. (2013)

- Additive noise term follows a **SIRP**
- Particular case of **Student's t** (texture $\sim$ inverse gamma)
- Compare **MSE and CRB**

$$\mathrm{MSE}([\hat{\boldsymbol{\theta}}]_k) = \mathrm{E}\left\{ \left([\hat{\boldsymbol{\theta}}]_k - [\boldsymbol{\theta}]_k \right)^2 \right\} \geqslant [\mathrm{CRB}(\boldsymbol{\theta})]_{k,k}$$

- Expression of the **FIM**

$$[\mathbf{F}]_{k,l} = 2\frac{\nu+4}{\nu+5} \sum_{pq} \Re \left\{ \frac{\partial \tilde{\mathbf{v}}_{pq}^H(\boldsymbol{\theta})}{\partial [\boldsymbol{\theta}]_k} \boldsymbol{\Omega}^{-1} \frac{\partial \tilde{\mathbf{v}}_{pq}(\boldsymbol{\theta})}{\partial [\boldsymbol{\theta}]_l} \right\}$$

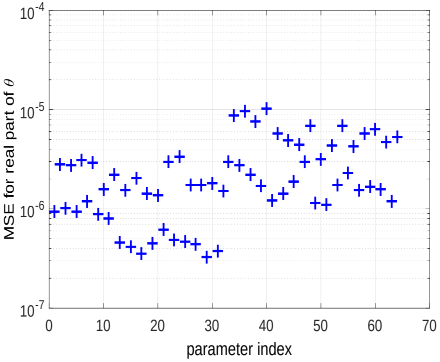


Figure 1: MSE of the real part of the 64 unknown parameters for a given SNR

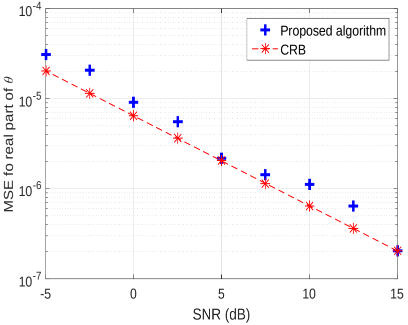


Figure 2: MSE vs. SNR for the real part of a given unknown parameter and the corresponding CRB

- $D = 2$ bright signal sources, $M = 8$ antennas
- 100 Monte-Carlo runs
  128 real unknown parameters of interest

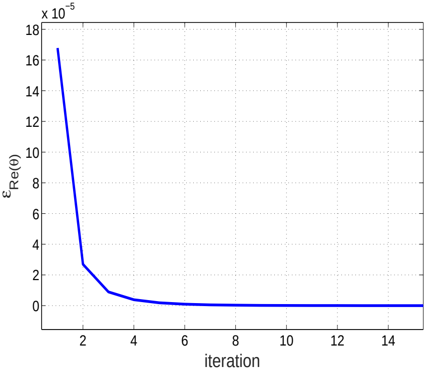


Figure 1: $\epsilon^h_{\Re\{\boldsymbol{\theta}\}}$ as function of the h-th iteration, for second loop from Algorithm 1

- Convergence properties

$$\epsilon^h_{\Re\{\boldsymbol{\theta}\}} = ||\Re\left\{\boldsymbol{\theta}^h - \boldsymbol{\theta}^{h-1}\right\}||_2^2$$

with h-th iteration

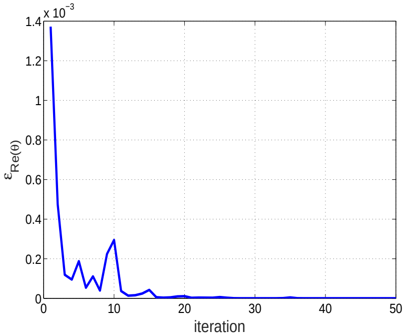


Figure 2: $\epsilon^h_{\Re\{\boldsymbol{\theta}\}}$ as function of the h-th iteration, for first loop from Algorithm 1

Realistic model

- $D = 2$ calibrator sources, $D' = 8$ weak outlier sources, background Gaussian noise
- M=8 antennas,

  128 real parameters of interest

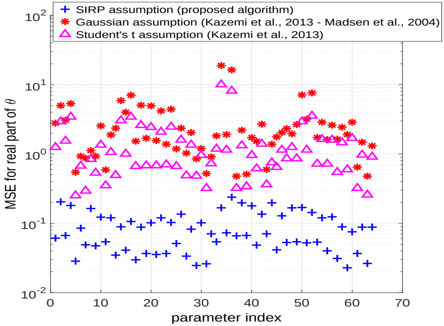


Figure: MSE of the real part of the 64 unknown parameters for a given SNR

Structured case

- $D = 2$ calibrator sources, $D' = 4$ weak outlier sources, $M = 8$ antennas
- $\mathbf{g} = [\mathbf{g}_1^T, \dots, \mathbf{g}_M^T]^T$
- Similar behavior for ϑ_1, ϑ_2, η_2, η_1, ζ_1 and ζ_2

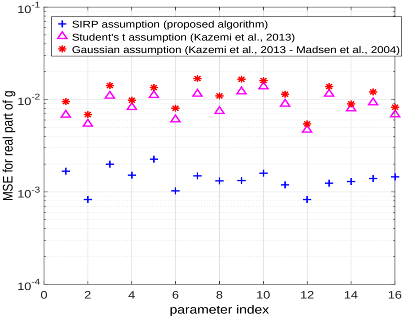


Figure: MSE of the real part of the 16 complex gains for a given SNR

- **Conclusion**
  - **Robust calibration** algorithm based on **iterative relaxed concentrated MLE** and **SIRP** modeling
  - Perturbation effects are modeled by **Jones matrices**
  - Study of non-structured and structured case (**compact arrays** like a LOFAR station)

Thank you for your attention