Advanced Calibration Methods for the Radio Astronomy context

Martin Brossard[a,b], Mohammed Nabil El Korso[c], Marius Pesavento[b], Rémy Boyer[d], Pascal Larzabal[e] and Stefan Wijnholds[f]

[a]ENS Cachan, France; [b]Nachrichtentechnische Systeme, Germany; [c]Université Paris Ouest Nantere La défense, France; [d]L2S, Gif-Sur-Yvette, France; [e]SATIE, France; [f]ASTRON, The Netherlands

Introduction and motivations

- Radio astronomy is currently making a technological transition toward the use of larges radio interferometers [1].
- Calibration is a mandatory task, otherwise strong distortions would appear during imaging.
- The proposed calibration algorithm estimates the positions of the calibrators, the gains of the array elements and their noise powers, in a parallel manner, and benefits from coherence across wavelength to improve calibration.

Data model and problem statement

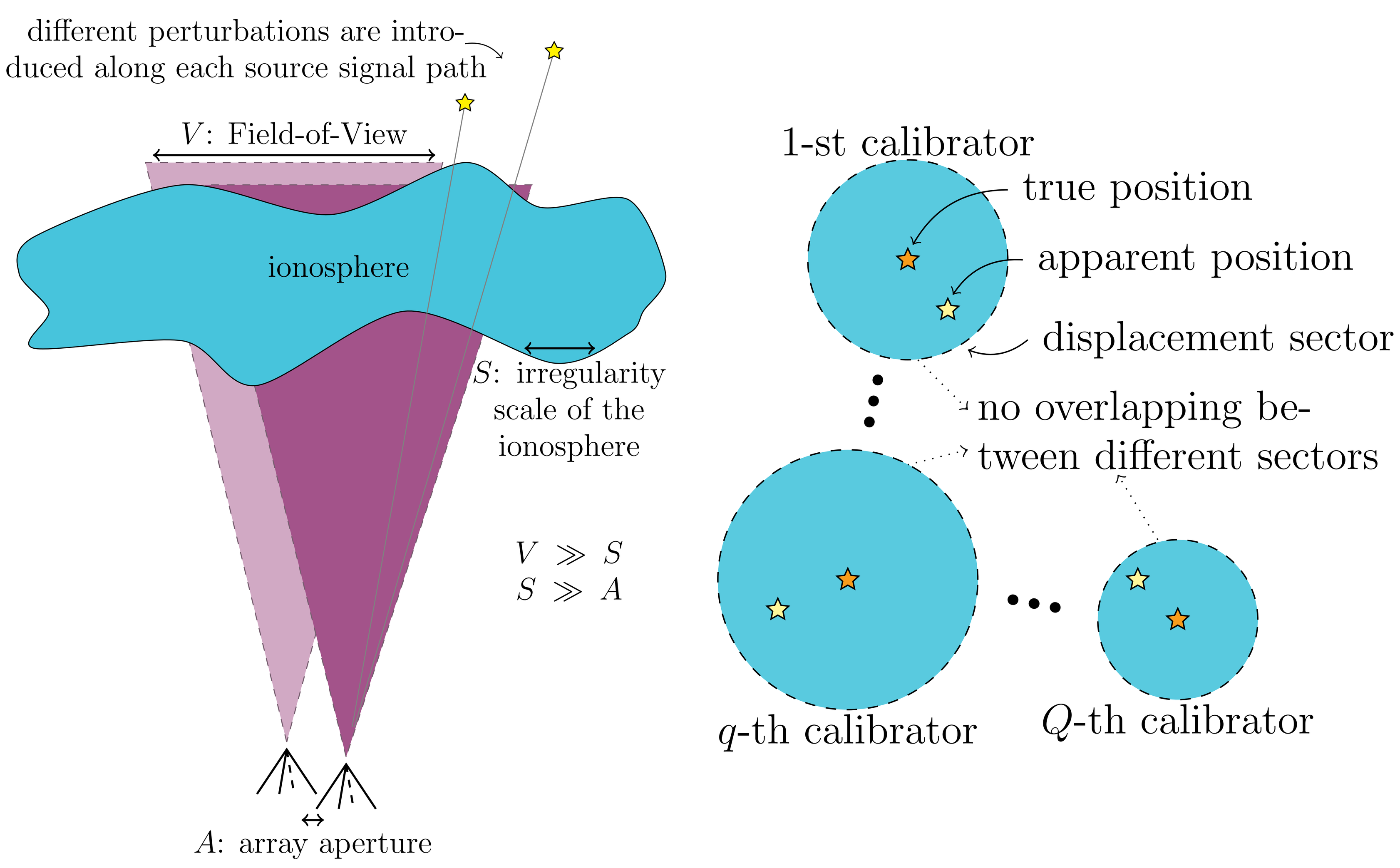


Figure 1: considered scenario (left) and calibrator assumptions (right)

1. Model of the covariance matrix

- The scenario corresponding to Figure 1 (left) is considered, i.e., the antennas have directional complex responses and the propagation mediums (mainly the ionosphere) cause different shifts for each calibrator position.
- In recent and future interferometers, the observations consist in a perturbed version of the covariance matrix $\mathbf{R}_\lambda$ of the P antenna signals, which depends on the wavelength λ, and is given by

$$\mathbf{R}_\lambda = \mathbf{G}_\lambda \mathbf{A}_{\mathbf{D}_\lambda} \boldsymbol{\Sigma}_\lambda^{\frac{1}{2}} \mathbf{M}_\lambda \left(\mathbf{G}_\lambda \mathbf{A}_{\mathbf{D}_\lambda} \boldsymbol{\Sigma}_\lambda^{\frac{1}{2}} \right)^H + \mathbf{R}_\lambda^{\mathrm{u}} + \boldsymbol{\Sigma}_\lambda^{\mathrm{n}} \in \mathbb{C}^{P\times P}, \text{ in which} \quad (1)$$

1. $\mathbf{G}_\lambda = \operatorname{diag}(\mathbf{g}_\lambda) \in \mathbb{C}^{P\times P}$ models the undirectional antenna gains
2. The steering matrices $\mathbf{A}_{\mathbf{D}_\lambda}$ are functions of the positions $\mathbf{D}_\lambda$ of Q calibrators
3. $\boldsymbol{\Sigma}_\lambda = \operatorname{diag}(\boldsymbol{\sigma}_\lambda)$ represents the real and known real powers of the calibrators
4. The directional amplitude gains toward calibrators are stacked in $\mathbf{M}_\lambda = \operatorname{diag}(\mathbf{m}_\lambda) \in \mathbb{R}^{Q\times Q}$
5. $\mathbf{R}_\lambda^{\mathrm{u}}$ is the covariance matrix of non-calibration sources (background noise)
6. $\boldsymbol{\Sigma}_\lambda^{\mathrm{n}} = \operatorname{diag}(\boldsymbol{\sigma}_\lambda^{\mathrm{n}}) \in \mathbb{R}^{Q\times Q}$ contains the antenna noise powers

2. Model effects of the wavelength

- We assume that the antennas and source parameters of the covariance matrix are wavelength dependent. Specifically:
  - $\mathbf{g}_\lambda$ has smooth variations across wavelength, representing by $\mathbf{g}_\lambda = \mathbf{B}_\lambda \boldsymbol{\alpha}$, where $\mathbf{B}_\lambda \in \mathbb{R}^{P\times KP}$ is a basis and $\boldsymbol{\alpha} \in \mathbb{C}^{KP}$ is the smooth parameter.
  - $\mathbf{m}_\lambda$ is proportional to λ^{-2} and the position shifts are proportional to λ^2.
  - Calibrators are well separated (see Figure 1, right).

3. Problem statement

- The aim of calibration is the estimation of $\{\mathbf{g}_\lambda, \mathbf{D}_\lambda, \mathbf{m}_\lambda, \boldsymbol{\sigma}_\lambda^{\mathrm{n}}\}_{\lambda\in\Lambda}$, from the $J \geq K$ sample covariance matrices $\{\hat{\mathbf{R}}_\lambda\}_{\lambda\in\Lambda}$, where Λ represents the set of the J available wavelengths.

Proposed Parallel Calibration Algorithm

- The proposed algorithm estimates the parameters in an iterative manner:

Until *convergence* **repeat**
1. Estimate $\{\mathbf{g}_\lambda\}_{\lambda\in\Lambda}$
2. Estimate $\{\mathbf{D}_\lambda, \mathbf{m}_\lambda, \boldsymbol{\sigma}_\lambda^{\mathrm{n}}\}_{\lambda\in\Lambda}$

1. Estimation of $\{\mathbf{g}_\lambda\}_{\lambda\in\Lambda}$

- This step considers the following minimization problem

$$\hat{\boldsymbol{\alpha}}, \{\hat{\mathbf{g}}_\lambda\}_{\lambda\in\Lambda} = \underset{\boldsymbol{\alpha}, \{\mathbf{g}_\lambda\}_{\lambda\in\Lambda}}{\arg\min} \sum_{\lambda\in\Lambda} \left\| \hat{\mathbf{R}}_\lambda - \mathbf{R}_\lambda(\mathbf{g}_\lambda) \right\|_2^2 = \sum_{\lambda\in\Lambda} \kappa_\lambda(\mathbf{g}_\lambda) \quad (2)$$
$$\text{subject to } \mathbf{g}_\lambda = \mathbf{B}_\lambda \boldsymbol{\alpha}, \forall \lambda \in \Lambda$$

- (2) is reformulated as a consensus problem by use of ADMM [2]
- It minimizes in a parallel manner the augmented Lagrangian $L(\{\mathbf{g}_\lambda, \mathbf{y}_\lambda\}_{\lambda\in\Lambda}, \boldsymbol{\alpha}) = \sum_{\lambda\in\Lambda} L_\lambda(\mathbf{g}_\lambda, \mathbf{y}_\lambda, \boldsymbol{\alpha})$, by use of the following updates:

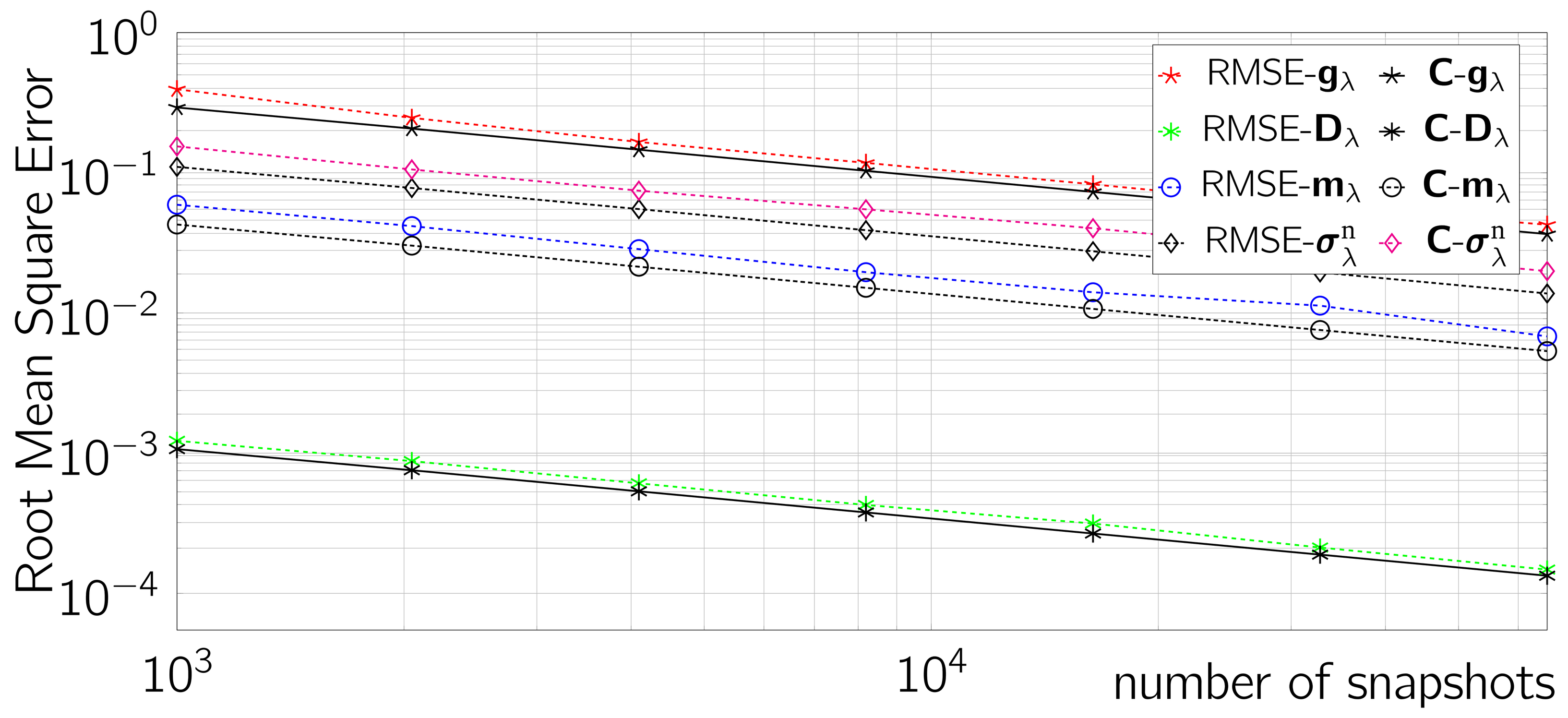
$$\mathbf{g}_\lambda^{[t+1]} = \underset{\mathbf{g}_\lambda}{\arg\min}\, L_\lambda\left(\mathbf{g}_\lambda, \mathbf{y}_\lambda^{[t]}, \boldsymbol{\alpha}^{[t]}\right), \lambda \in \Lambda, \quad (3)$$
$$\boldsymbol{\alpha}^{[t+1]} = \underset{\boldsymbol{\alpha}}{\arg\min} \sum_{\lambda\in\Lambda} L_\lambda\left(\mathbf{g}_\lambda^{[t+1]}, \mathbf{y}_\lambda^{[t]}, \boldsymbol{\alpha}\right), \quad (4)$$
$$\text{update the Lagrange multiplier } \mathbf{y}_\lambda^{[t+1]}, \lambda \in \Lambda \quad (5)$$

2. Estimation of $\{\mathbf{D}_\lambda, \mathbf{m}_\lambda, \boldsymbol{\sigma}_\lambda^{\mathrm{n}}\}_{\lambda\in\Lambda}$

- This step estimates jointly the positions $\mathbf{D}_\lambda$ and the gains $\mathbf{m}_\lambda$, by reformulating the problem as a sparse problem:
  - it substitutes $\mathbf{m}_\lambda$ by a large vector $\tilde{\mathbf{m}}_\lambda$, that represents the amplitudes for a large number of positions, and only the values of $\tilde{\mathbf{m}}_\lambda$ corresponding to the apparent positions of the calibrators are non-zero elements
  - the efficient and parallel procedure to solve the reformulated problem is based on Distributed Iterative Hard Thresholding [3]

Simulation results

- Realistic simulations are performed in a low Signal-to-Noise Radio scenario.
- The variances of the errors are closed to the Cramèr-Rao bounds, showing the proposed scheme is statistically efficient and robust to non-calibration sources.



Bibliography

[1] M. P. Van Haarlem, M. W. Wise, A. W. Gunst, et. al., "LOFAR: The LOw-Frequency ARray," *Astronomy & Astrophysics*, vol. 556, 2013.

[2] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundation and Trends in Machine Learning*, vol. 3, no. 1, pp. 1-122, 2011.

[3] S. Patterson, Y. C. Eldar, and I. Keidar, "Distributed compressed sensing for static and time-varying networks," *IEEE Transactions on Signal Processing*, vol. 62, no. 19, pp. 4931-4946, 2014.