Analysis of the Peeling Algorithm

Brian D. Jeffs[1] **and Sebastiaan van der Tol**[2]

1: **Brigham Young University, Electrical and Computer Engineering**
bjeffs@ee.byu.edu

2: **TU Delft, Circuits and Systems, svdt@cas.et.tudelft.nl**

BYU
BRIGHAM YOUNG UNIVERSITY

TU Delft
Delft University of Technology

The Bad News:

LOFAR is uncalibratable ...

The Good News:

with conventional algorithms.

Next Generation Widefield Instrument Calibration Challenges

- Larger apertures.
- Many more array elements.
- Wider range of frequencies.

© ASTRON

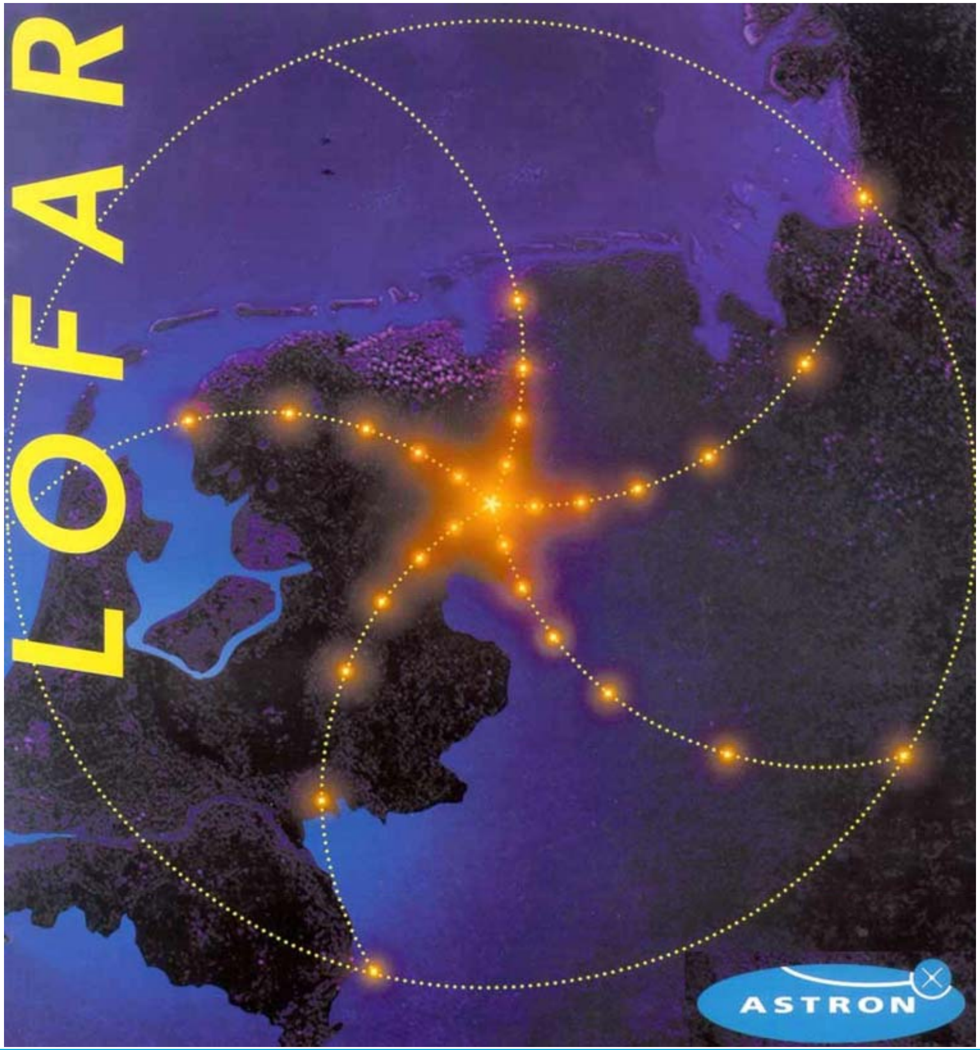
- Ionospheric interaction.
- Calibration may be source direction dependent.
- Calibrated UV data may not be possible.

LOFAR is a Widefield Instrument

- Each station antenna sees the entire sky.
- 7200 dual-pol antennas.
- Multiple simultaneous beams are formed in different directions.
- ~6° beam mainlobe

LOFAR Geometry

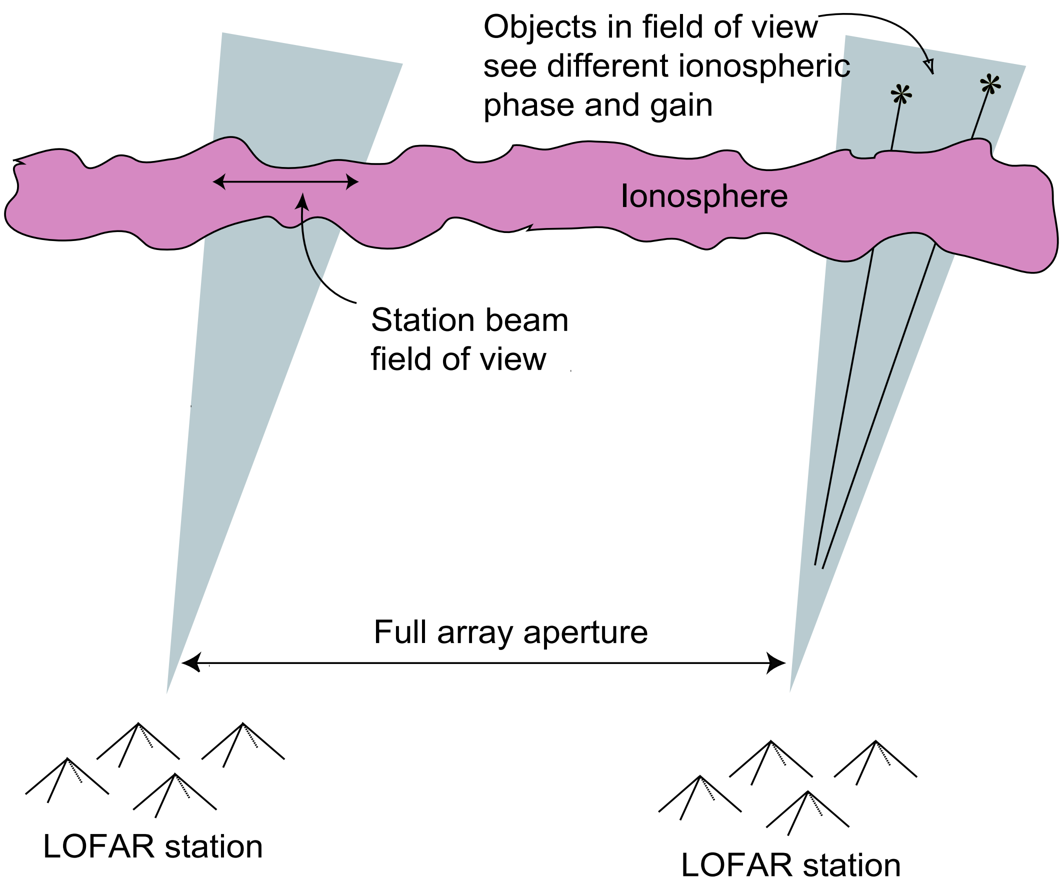
- 72 stations.
- 100 km aperture.
- Significant ionospheric variation across the array complicates calibration.
- Nonisoplanatic iono-sphere across calibration sources and stations.
- Very low frequencies: 30 - 240 MHz.



BYU
BRIGHAM YOUNG UNIVERSITY

TU Delft

The LOFAR Calibration Problem



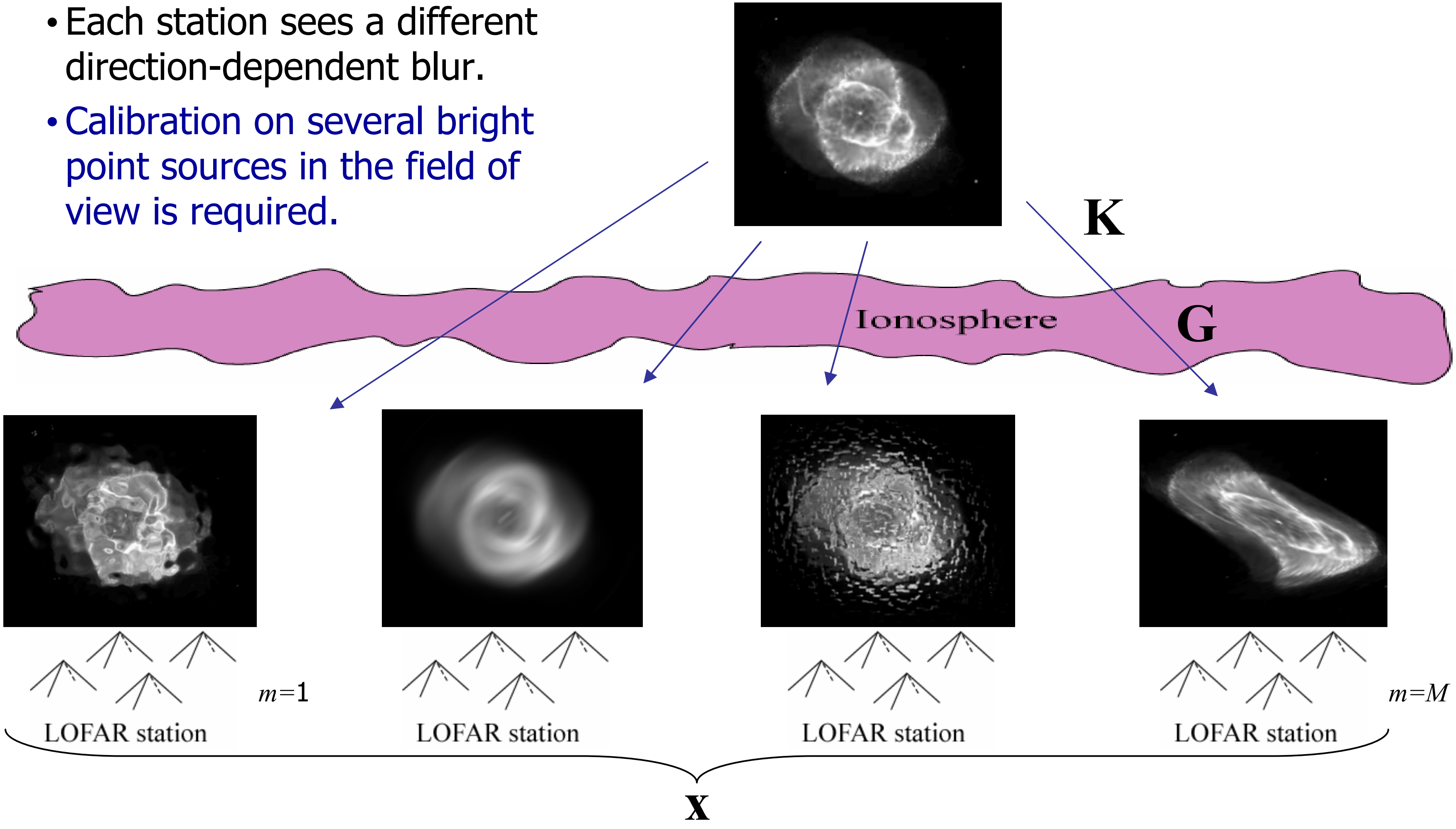
- At low frequencies the ionosphere perturbs phase and gain.
- Calibration terms must be estimated for each bright source & station.
- Calibration for other objects is interpolated.
- Physical constraints must be applied.

BYU BRIGHAM YOUNG UNIVERSITY

TU Delft

Calibration is Direction Dependent

- Each station sees a different direction-dependent blur.
- Calibration on several bright point sources in the field of view is required.



Matrix Form Data Model

V: visibility matrix, computed over a series of time-frequency intervals. Observed.

G: calibration complex gain matrix. One column per calibrator source. Unknown.

K: Fourier kernel, geometric array response. $\mathbf{s}_q$ is source direction vector. $\mathbf{r}_m$ is station location. Known.

B: Calibrator source intensity. Known.

D: Noise covariance. Unknown.

$$\mathbf{V} = E\{\mathbf{x}[n]\mathbf{x}^{\mathrm{H}}[n]\}$$

$$= (\mathbf{G} \odot \mathbf{K})\mathbf{B}(\mathbf{G} \odot \mathbf{K})^{\mathrm{H}} + \mathbf{D}$$

$$\mathbf{G} = \begin{bmatrix} g_{1,1} & \cdots & g_{1,Q} \\ \vdots & & \vdots \\ g_{M,1} & \cdots & g_{M,Q} \end{bmatrix}$$

$$\mathbf{K} = \begin{bmatrix} k_{1,1} & \cdots & k_{1,Q} \\ \vdots & & \vdots \\ k_{M,1} & \cdots & k_{M,Q} \end{bmatrix}, \quad k_{m,q} = \exp\{i\tfrac{2\pi f}{c}\mathbf{s}_q \cdot \mathbf{r}_m\}$$

$$\mathbf{B} = \begin{bmatrix} b_1 & & \\ & \ddots & \\ & & b_Q \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_M \end{bmatrix}$$

The Single Snapshot Calibration Ambiguity

- For conventional arrays without direction dependent ionospheric phase perturbation calibration is possible with a single $\mathbf{V}_{k,n}$ observation.
- Not so for LOFAR, there is an essential ambiguity.

$$\mathbf{V} = (\mathbf{G} \odot \mathbf{K})\mathbf{B}^{\frac{1}{2}}\mathbf{U}\,\mathbf{U}^{\mathrm{H}}\mathbf{B}^{\frac{1}{2}}(\mathbf{G} \odot \mathbf{K})^{\mathrm{H}} + \mathbf{D}$$

Rotation by $\mathbf{U}$ is invisible in $\mathbf{V}$

$$\mathbf{U}\,\mathbf{U}^{\mathrm{H}} = \mathbf{I}, \qquad \text{a unitary matrix}$$

$$\widetilde{\mathbf{G}}(\mathbf{U}) = ((\mathbf{G} \odot \mathbf{K})\mathbf{B}^{\frac{1}{2}}\mathbf{U}^{H}\,\mathbf{B}^{-\frac{1}{2}}) \odot \mathbf{K}^{\bullet -1}$$

Each $\mathbf{U}$ produces a different calibration

- Calibration is Impossible with a single visibility snapshot!

Solutions to Calibration Ambiguity

- Time-frequency diversity
  - Fringe rotation over time and frequency changes visibilities while calibration gains are nearly constant.
  - Low order polynomial fitting over time-frequency.
  - Peeling.
- Single snapshot calibration
  - Compact core.
  - Deterministic Frequency dependence.
  - Known gain magnitudes, $|\mathbf{G}|$.

The Direct Least Squares Solution

$$\Theta = \left[\theta_{1,1}^T, \cdots, \theta_{K,N}^T\right]^T$$

$$\hat{\Theta} = \underset{\theta_{1,1},\cdots,\theta_{K,N}}{\arg\min} \sum_k \sum_n \left\| \mathbf{V}_{k,n} - (\mathbf{G}\{\theta_{k,n}\} \odot \mathbf{K}_{k,n})\mathbf{B}(\mathbf{G}\{\theta_{k,n}\} \odot \mathbf{K}_{k,n})^H \right\|^2$$

- Problems:
  - Direct optimization is computationally intractable.
  - Too many parameters.
  - Requires good initialization. Where do you get it?
  - Does not exploit known smoothness structure over k,n.
  - Due to ambiguity, solution is not unique if $\theta_{k,n}$ has same degrees of freedom as $\mathbf{G}_{k,n}$.

The Peeling Approach

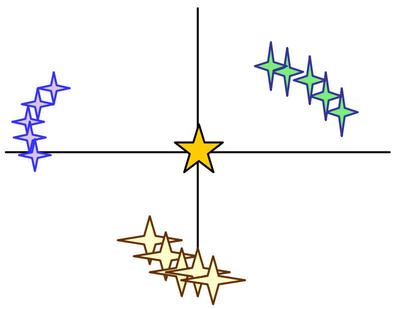
- Current proposed LOFAR calibration method.
- Replace joint estimation of $\mathbf{G}$ for Q sources with a series of single source calibration problems.
- Exploit relative fringe rotation rates among calibrator sources.
- Assume calibration gains are constant over a t-f cell.
- Computationally efficient.
- References:

1. J.E. Noordam, "LOFAR calibration challenges," *Proceedings of the SPIE, vol. 5489*, Oct. 2004.
2. J.E. Noordam, "Peeling the Visibility Onion, the optimum way of self-calibration," ASTRON tech. report MEM-078 June 2003.

Basic Peeling Algorithm Steps

- Over a time-frequency cell of nearly constant gains, rotate all visibilities to phase center the brightest remaining source.

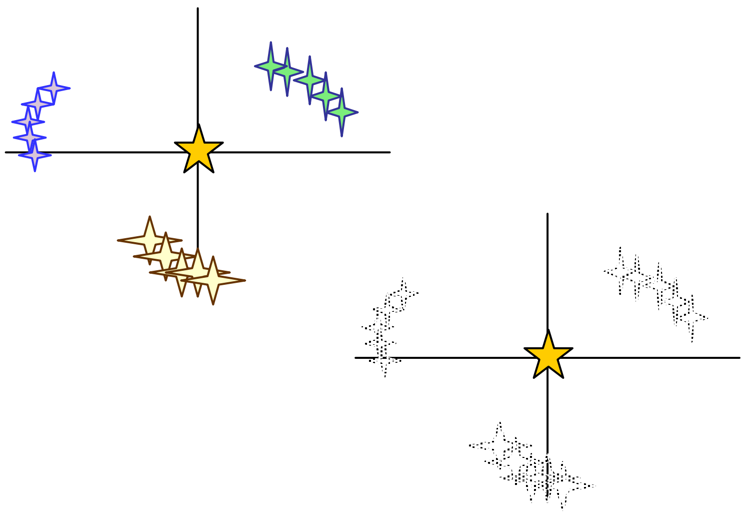
$$\widetilde{\mathbf{V}}_{k,n} = \mathrm{diag}\{\overline{\mathbf{k}}_{q,k,n}\}\mathbf{V}_{k,n}\mathrm{diag}\{\mathbf{k}_{q,k,n}\}$$



Basic Peeling Algorithm Steps

- Over a time-frequency cell of nearly constant gains, rotate all visibilities to phase center the brightest remaining source.
- Average centered visibilities to suppress non-centered sources.

$$\hat{\mathbf{V}}_q = \frac{1}{KN}\sum_k \sum_n \tilde{\mathbf{V}}_{k,n}$$

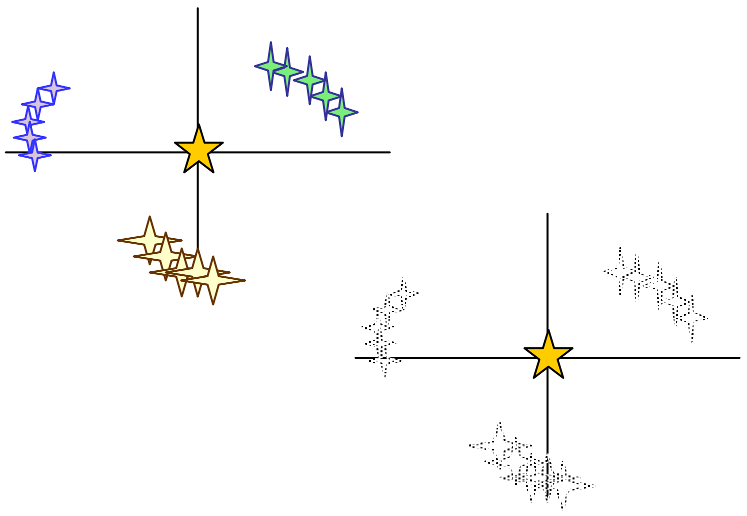


Basic Peeling Algorithm Steps

- Over a time-frequency cell of nearly constant gains, rotate all visibilities to phase center the brightest remaining source.
- Average centered visibilities to suppress non-centered sources.
- Solve as a conventional single source calibration.

$$\hat{\mathbf{g}}_q = \min_{\mathbf{g}} \left\| \hat{\mathbf{V}}_q - b_q \mathbf{g}\,\mathbf{g}^H \right\|$$

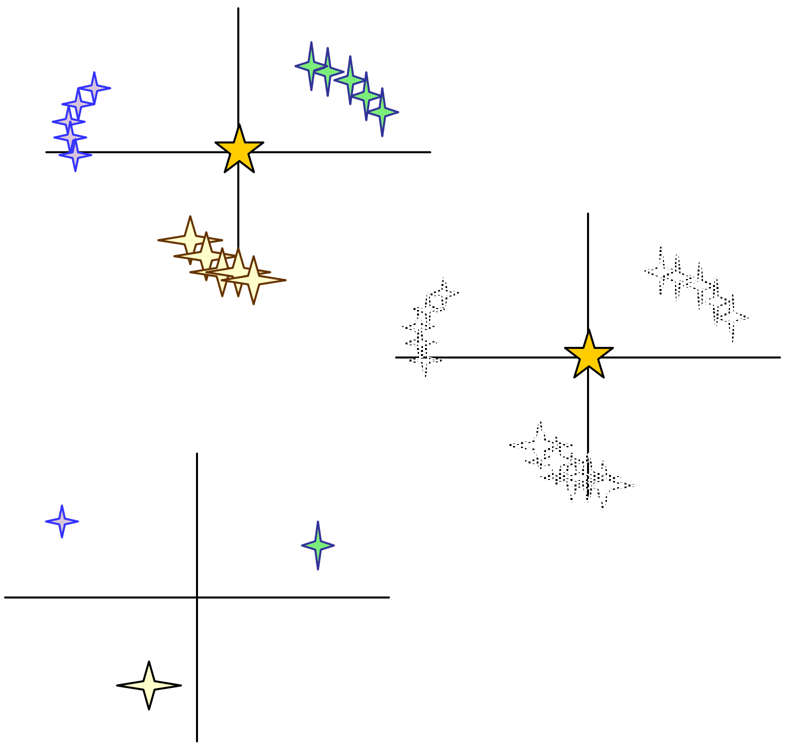
<u>"Image plane" equivalent</u>



Basic Peeling Algorithm Steps

- Over a time-frequency cell of nearly constant gains, rotate all visibilities to phase center the brightest remaining source.
- Average centered visibilities to suppress non-centered sources.
- Solve as a conventional single source calibration.
- Subtract the calibrated source from visibilities.

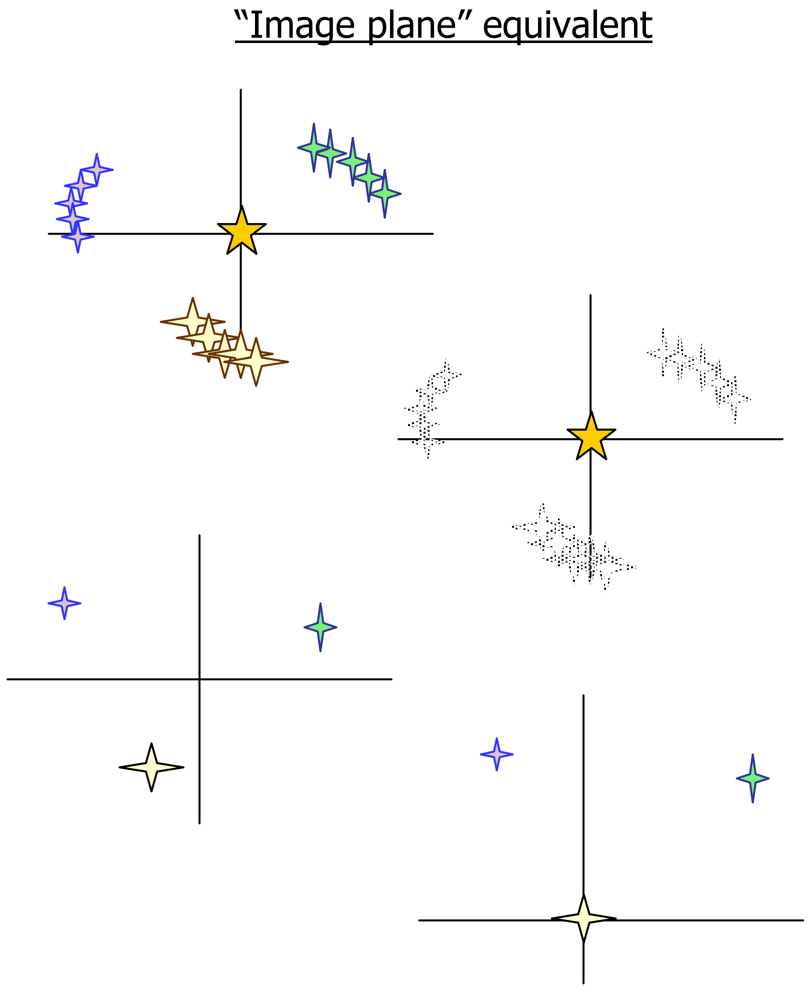
"Image plane" equivalent



$$\mathbf{V}_{k,n} = \mathbf{V}_{k,n} - b_q \operatorname{diag}\{\mathbf{k}_{q,k,n}\}\hat{\mathbf{g}}_q \hat{\mathbf{g}}_q^H \operatorname{diag}\{\overline{\mathbf{k}}_{q,k,n}\}$$

Basic Peeling Algorithm Steps

- Over a time-frequency cell of nearly constant gains, rotate all visibilities to phase center the brightest remaining source.
- Average centered visibilities to suppress non-centered sources.
- Solve as a conventional single source calibration.
- Subtract the calibrated source from visibilities.
- Repeat for next brightest source.



2-D Polynomial Model over time-frequency for Ionospheric Variation

- Variations in $\mathbf{G}$ are smooth over time and frequency

$$\boxed{\begin{aligned}\mathbf{G}\{\theta_{k,n}\} = {} & (\Gamma_{00} + \Gamma_{10} f_k + \Gamma_{20} f_k^2 + \Gamma_{01} t_n + \Gamma_{02} t_n^2 + \Gamma_{11} f_k t_n) \\ & \odot \exp\left\{ i(\Phi_{00} + \Phi_{10} f_k + \Phi_{20} f_k^2 + \Phi_{01} t_n + \Phi_{02} t_n^2 + \Phi_{11} f_k t_n) \right\}\end{aligned}}$$

$$\mathbf{p} = \left[\mathrm{vec}\{\Gamma_{00}\}^{\mathrm{T}}, \cdots, \mathrm{vec}\{\Gamma_{11}\}^{\mathrm{T}}, \mathrm{vec}\{\Phi_{00}\}^{\mathrm{T}}, \cdots, \mathrm{vec}\{\Phi_{11}\}^{\mathrm{T}}, \mathrm{diag}\{\mathbf{D}\} \right]^{\mathrm{T}}$$

- Estimating the smaller parameter set, $\mathbf{p}$, improves performance.
- Over a large window $\mathbf{p}$ does <u>not</u> depend on k,n.

Challenges

- Phase centered averaging does not completely remove non-centered sources.
- "Contamination" from non-centered sources biases estimate of $\hat{\mathbf{g}}_q$.
- Solution: Multiple passes of peeling.
- Computational burden eliminates some candidate approaches.
- Many local minima in polynomial optimization makes a good initial solution critical.

Why Study the Cramer-Rao Bound?

- Array calibration is a statistical parameter estimation problem.
- The CRB reveals the theoretical limit on estimation error variance.
- Absolute frame of reference: <u>no algorithm can beat</u> the CRB.
- Answers the questions:
  - Is the existing algorithm adequate?
  - Is there hope for finding a better solution, "How close are we?"
  - Permits trading off performance and computational burden.
  - Can be computed even if no algorithm exists yet.
  - The BIG question: Can LOFAR be reliably calibrated?

CRB Definition

- Notation

  $\mathbf{x}$: a vector of random samples with joint probability density $p(\mathbf{x}\,|\,\theta)$.

  $\hat{\theta}$: any unbiased estimator for θ.

  $\mathbf{C}_{\hat{\theta}}$: covariance matrix for $\hat{\theta}$.

  $\mathbf{M}$: Fisher information matrix.

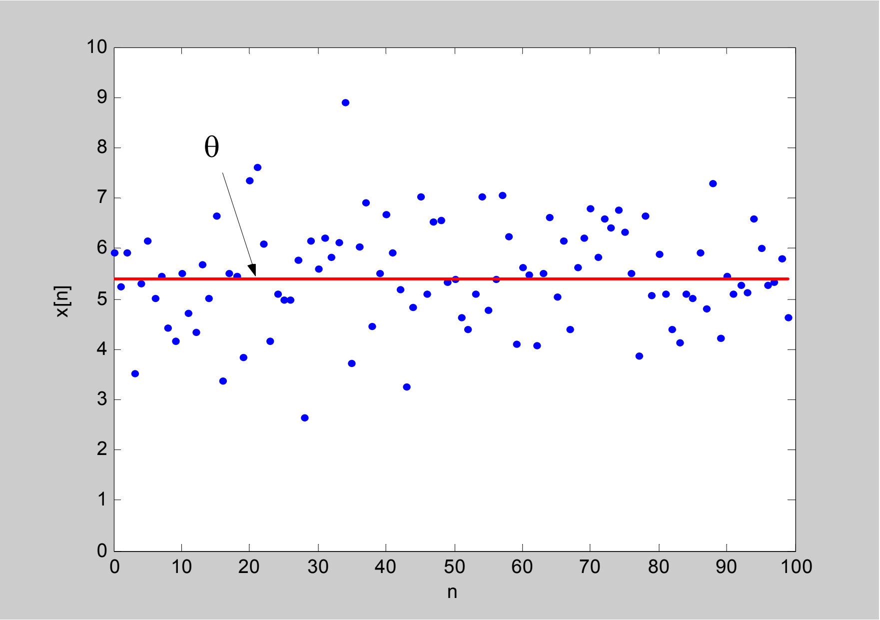
- The Cramer-Rao theorem:

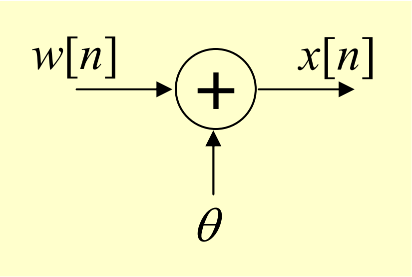
$$\mathbf{C}_{\hat{\theta}} \geq \mathbf{M}^{-1} = -\left(E\left\{ \frac{\partial^2 \ln p(\mathbf{x}\,|\,\theta)}{\partial\theta\, \partial\theta^{\mathrm{T}}} \right\} \right)^{-1}$$

- Error variance is lower bounded by $\operatorname{diag}\{\mathbf{C}_{\hat{\theta}}\}$!

A Simple Example

- Estimate a constant in additive white Gaussian noise:





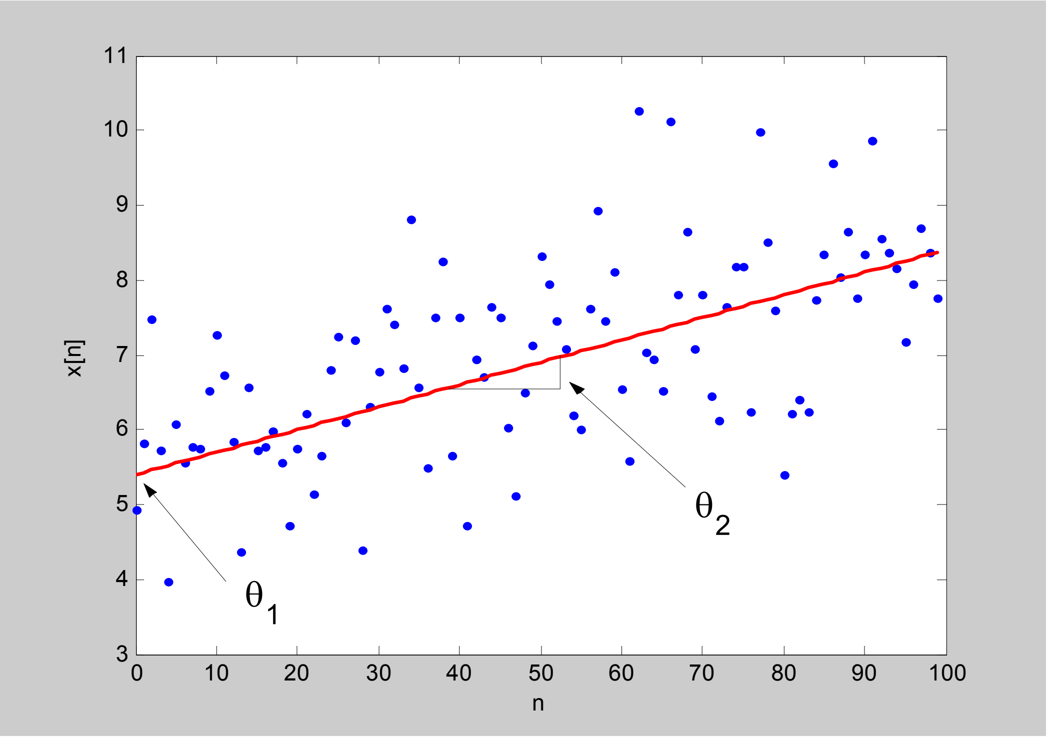
$$\mathbf{x} = [x[0], \cdots, x[N-1]]^{\mathrm{T}}$$

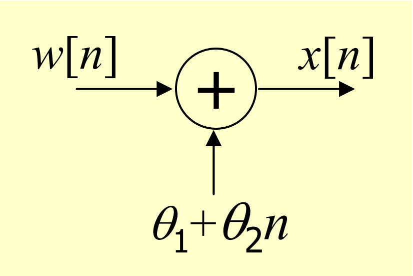
CRB result: $\mathrm{var}(\hat{\theta}) \geq \dfrac{\sigma^2}{N}.$

This is the well known variance for the sample mean!

A Second Simple Example

- Line fitting in additive white Gaussian noise:





$$x[n] = \theta_1 + \theta_2 n + w[n]$$

We now have no intuition on estimation error for θ_1 and θ_2!

A Second Simple Example: Insights

$$\mathbf{M} = \frac{1}{\sigma^2}\begin{bmatrix} N & \frac{N(N-1)}{2} \\ \frac{N(N-1)}{2} & \frac{N(N-1)(2N-1)}{6} \end{bmatrix},$$

$$\mathrm{var}(\theta_1) = \left[\mathbf{M}^{-1}\right]_{1,1} = \frac{2(2N-1)\sigma^2}{N(N+1)}, \quad \mathrm{var}(\theta_2) = \left[\mathbf{M}^{-1}\right]_{2,2} = \frac{12\sigma^2}{N(N^2-1)}$$

$$\boxed{\lim_{N\to\infty} \mathrm{var}(\theta_1) \geq \frac{4\sigma^2}{N} \qquad \lim_{N\to\infty} \mathrm{var}(\theta_2) \geq \frac{12\sigma^2}{N^3}}$$

- Variance on constant term θ_1 is now higher.
  $\longrightarrow$ estimating more parameters increases error.
- Variance of slope term, θ_2, drops more rapidly with N.
  $\longrightarrow$ θ_2 is easier to estimate.
  $\longrightarrow$ $x[n]$ is more sensitive to θ_2 due to multiplication by n.

Now it Gets a Little Messy

$$\mathbf{M}_{k,n} = \begin{bmatrix} \mathbf{M}_{\gamma_1\gamma_1}\cdots\mathbf{M}_{\gamma_1\gamma_Q} & \mathbf{M}_{\gamma_1\varphi_1}\cdots\mathbf{M}_{\gamma_1\varphi_Q} & \mathbf{M}_{\gamma_1 d} \\ \ddots & \ddots & \vdots \\ \mathbf{M}_{\gamma_Q\gamma_1}\cdots\mathbf{M}_{\gamma_Q\gamma_Q} & \mathbf{M}_{\gamma_Q\varphi_1}\cdots\mathbf{M}_{\gamma_Q\varphi_Q} & \mathbf{M}_{\gamma_Q d} \\ \mathbf{M}_{\varphi_1\gamma_1}\cdots\mathbf{M}_{\varphi_1\gamma_Q} & \mathbf{M}_{\varphi_1\varphi_1}\cdots\mathbf{M}_{\varphi_1\varphi_Q} & \mathbf{M}_{\varphi_1 d} \\ \ddots & \ddots & \vdots \\ \mathbf{M}_{\varphi_Q\gamma_1}\cdots\mathbf{M}_{\varphi_Q\gamma_Q} & \mathbf{M}_{\varphi_Q\varphi_1}\cdots\mathbf{M}_{\varphi_Q\varphi_Q} & \mathbf{M}_{\varphi_Q d} \\ \mathbf{M}_{d\gamma_1}\cdots\mathbf{M}_{d\gamma_Q} & \mathbf{M}_{d\varphi_1}\cdots\mathbf{M}_{d\varphi_Q} & \mathbf{M}_{dd} \end{bmatrix}$$

Block Fisher information

$$\begin{aligned}
\mathbf{M}_{\gamma_p\gamma_q} &= 2\sigma_p^2\sigma_q^2 \mathrm{Re}\Big(\big(\tilde{\boldsymbol{\Phi}}_p^{\mathrm{T}}\overline{\mathbf{R}}^{-1}\bar{\tilde{\boldsymbol{\Phi}}}_q\big)\big(\mathbf{a}_p^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{a}_q\big) \\
&\quad + \big(\tilde{\boldsymbol{\Phi}}_p^{\mathrm{T}}\mathbf{R}^{-1}\bar{\mathbf{a}}_q\big)\big(\mathbf{a}_p^{\mathrm{H}}\mathbf{R}^{-1}\tilde{\boldsymbol{\Phi}}_q\big)\Big) \\
\mathbf{M}_{\varphi_p\varphi_q} &= 2\sigma_p^2\sigma_q^2 \mathrm{Re}\Big(\big(\tilde{\boldsymbol{\Gamma}}_p\overline{\mathbf{R}}^{-1}\bar{\tilde{\boldsymbol{\Gamma}}}_q\big)\big(\mathbf{a}_p^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{a}_q\big) \\
&\quad - \big(\tilde{\boldsymbol{\Gamma}}_p\mathbf{R}^{-1}\bar{\mathbf{a}}_q\big)\big(\mathbf{a}_p^{\mathrm{H}}\overline{\mathbf{R}}^{-1}\tilde{\boldsymbol{\Gamma}}_q\big)\Big) \\
\mathbf{M}_{\mathbf{dd}} &= \overline{\mathbf{R}}^{-1} \odot \mathbf{R}^{-1} \\
\mathbf{M}_{\gamma_p\varphi_q} &= 2\sigma_p^2\sigma_q^2 \mathrm{Im}\Big(\big(\tilde{\boldsymbol{\Phi}}_p^{\mathrm{T}}\overline{\mathbf{R}}^{-1}\bar{\tilde{\boldsymbol{\Gamma}}}_q\big)\big(\mathbf{a}_p^{\mathrm{H}}\mathbf{R}^{-1}\mathbf{a}_q\big) \\
&\quad + \big(\tilde{\boldsymbol{\Phi}}_p^{\mathrm{T}}\mathbf{R}^{-1}\bar{\mathbf{a}}_q\big)\big(\mathbf{a}_p^{\mathrm{H}}\mathbf{R}^{-1}\tilde{\boldsymbol{\Gamma}}_q\big)\Big) \\
\mathbf{M}_{\gamma_p\mathbf{d}} &= 2\sigma_p^2 \mathrm{Re}\big(\tilde{\boldsymbol{\Phi}}_p\overline{\mathbf{R}}^{-1} \circ \mathbf{a}_p^{\mathrm{H}}\mathbf{R}^{-1}\big) \\
\mathbf{M}_{\varphi_p\mathbf{d}} &= -2\sigma_p^2 \mathrm{Im}\big(\tilde{\boldsymbol{\Gamma}}_p\overline{\mathbf{R}}^{-1} \circ \mathbf{a}_p^{\mathrm{H}}\mathbf{R}^{-1}\big)
\end{aligned}$$

Block closed forms

Constraint Jacobian for packed central core

$$\mathbf{J}_{k,n} = \begin{bmatrix} \mathbf{1}_Q \otimes \begin{bmatrix} \mathbf{I}_{M_c} \\ \mathbf{0}_{M_r,M_c} \end{bmatrix} & \mathbf{I}_Q \otimes \begin{bmatrix} \mathbf{0}_{M_c,M_r} \\ \mathbf{I}_{M_r} \end{bmatrix} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1}_Q \otimes \begin{bmatrix} \mathbf{I}_{M_c} \\ \mathbf{0}_{M_r,M_c} \end{bmatrix} & \mathbf{I}_Q \otimes \begin{bmatrix} \mathbf{0}_{M_c,M_r-1} \\ \mathbf{I}^s_{M_r} \end{bmatrix} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I}_M \end{bmatrix}$$

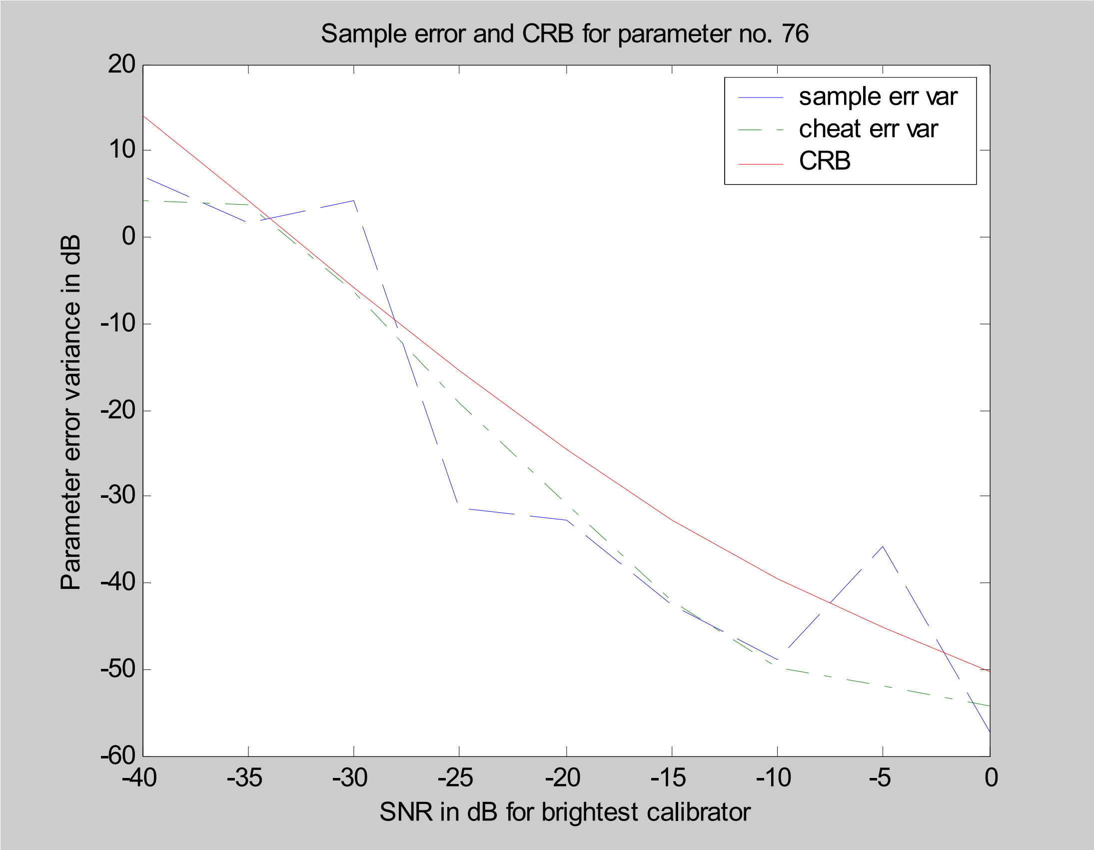
Now it Gets a Little Messy

The important points:

- Closed form CRB expressions have been derived for most important LOFAR calibration models.
- Though expressions are complex, computer codes have been developed to evaluate them.
- These solutions exist now and could be made available for astronomers to predict calibration performance for a given observations.

Peeling Simulation Results

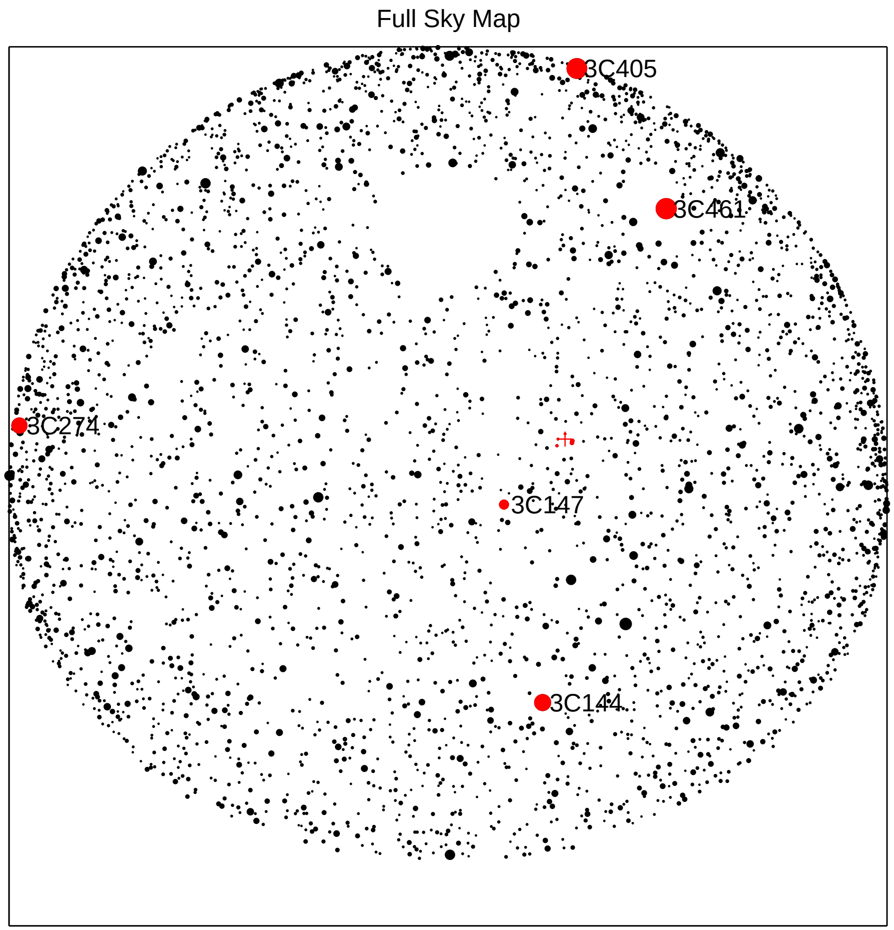
- Polynomial fit over a 10 s by 100 kHz "snippet" window.
- Two sources.
- Plot is for zero order coefficient, 16th station, 1st source.
- 30 station array with 100 km aperture.
- Performance is typical of all other parameters.
- Three peeling passes.



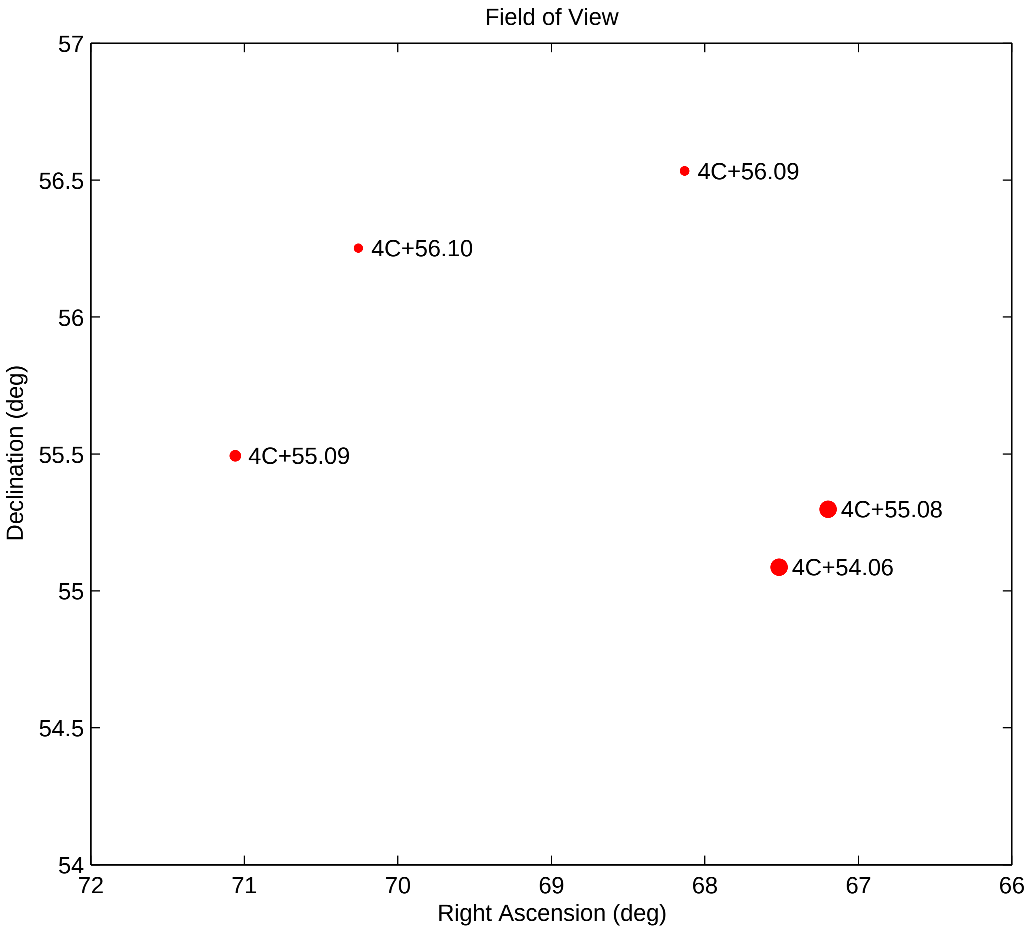
CRB Analysis for a realistic Scenario

- Point LOFAR beam in arbitrary direction.
- Model station beam pattern and noise level accurately.
- Calibrate on 5 brightest sources in beam mainlobe and 5 brightest in sidelobes.
- Use 2-D 1st order polynomial fitting over a "snippet" of 10 seconds and 500 kHz.
- Calculate CRB for polynomial coefficients.
- Project coefficient CRB to corresponding complex gain error variance.

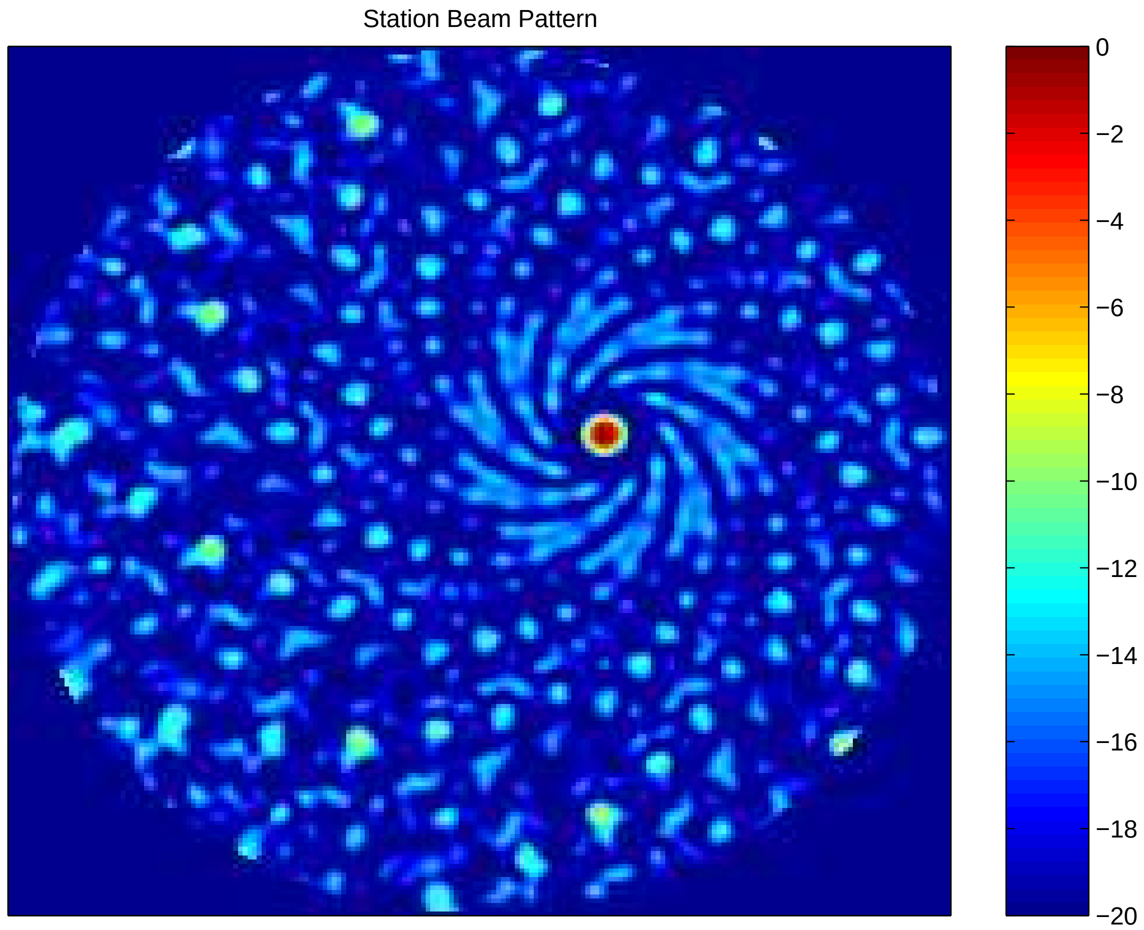
Full Sky Map



Field of View

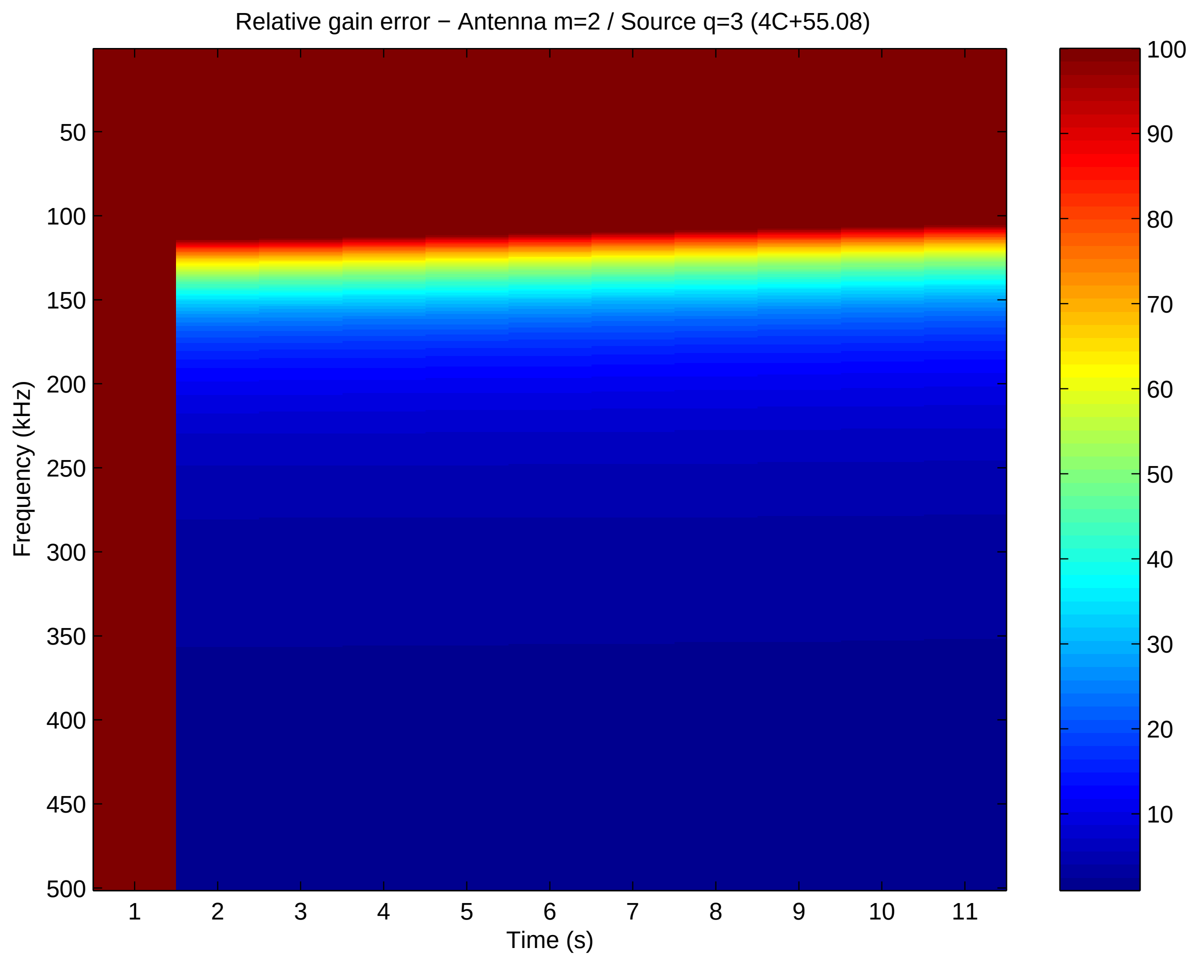


Station Beam Pattern



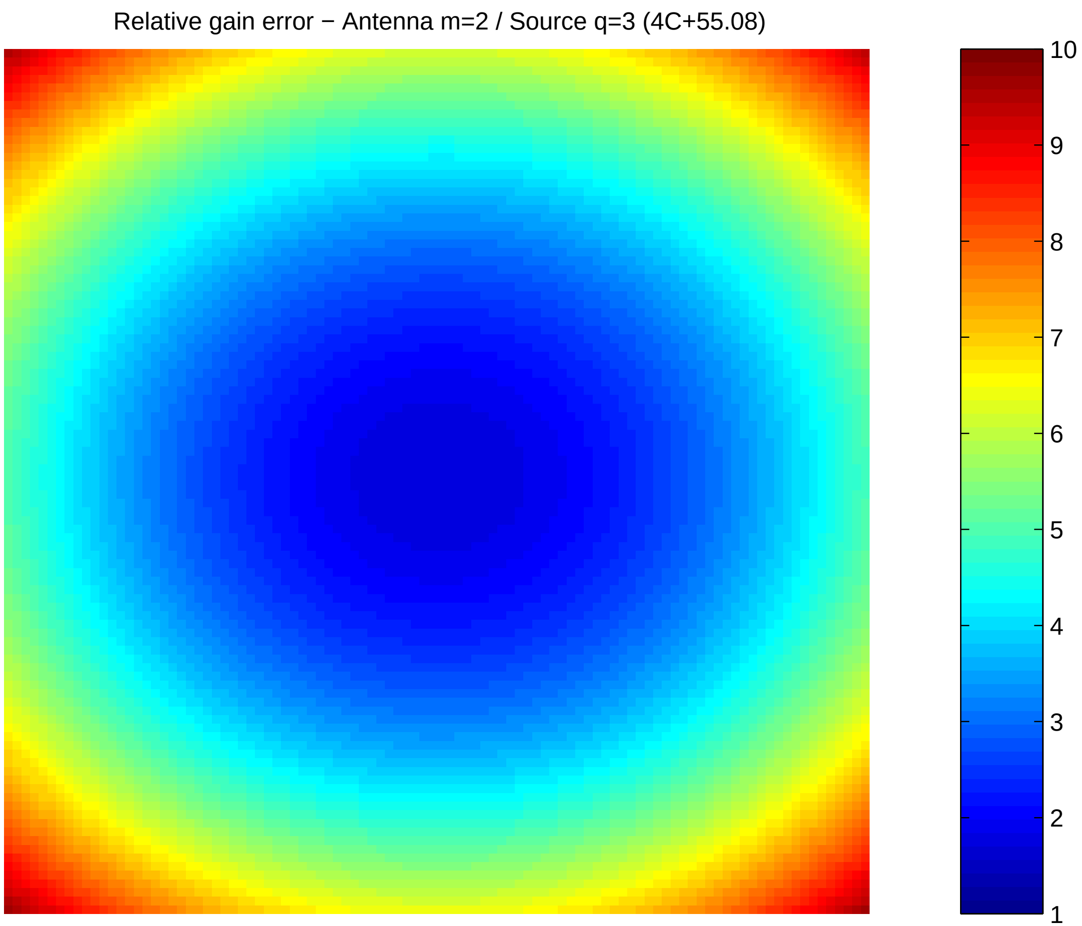
Multiple Source CRB for 2-D Polynomial Coefficients

- Coefficient error variance relative to conventional single source calibration.
- Note that calibra-tion fails without at least two seconds and 150 kHz of diversity.



Multiple Source Complex Gain Error CRB

- Use estimated polynomial coefficients to calculate complex gains.
- Calculate error variance CRB over full time frequency range.
- Largest error is near domain edges.



Conclusions

- The BIG answer: Yes, LOFAR can be calibrated.
- Given a range of time-frequency observations and compact core geometry: there are no theoretical roadblocks to achieving useful calibration estimates.
- Does peeling work for direction dependent calibration?
  - Yes, so far so good.
  - Ongoing progress in reducing cross-source contamination: multiple pass peeling & demixing.
  - Low SNR performance improvements being studied.

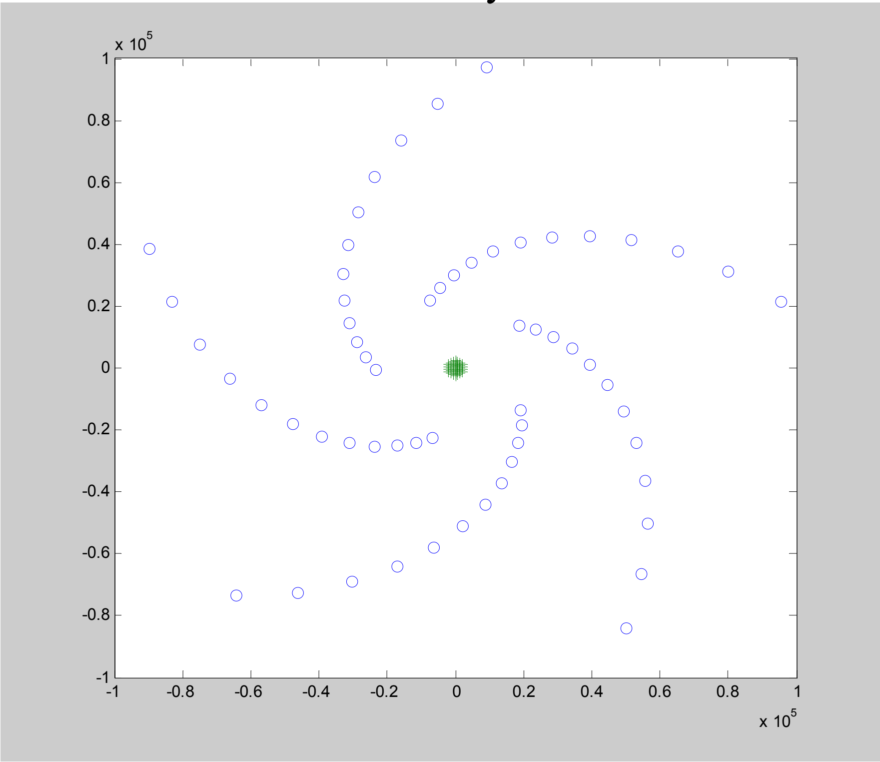
Optional Slides

- Use these only if (by some miracle) there is extra time.

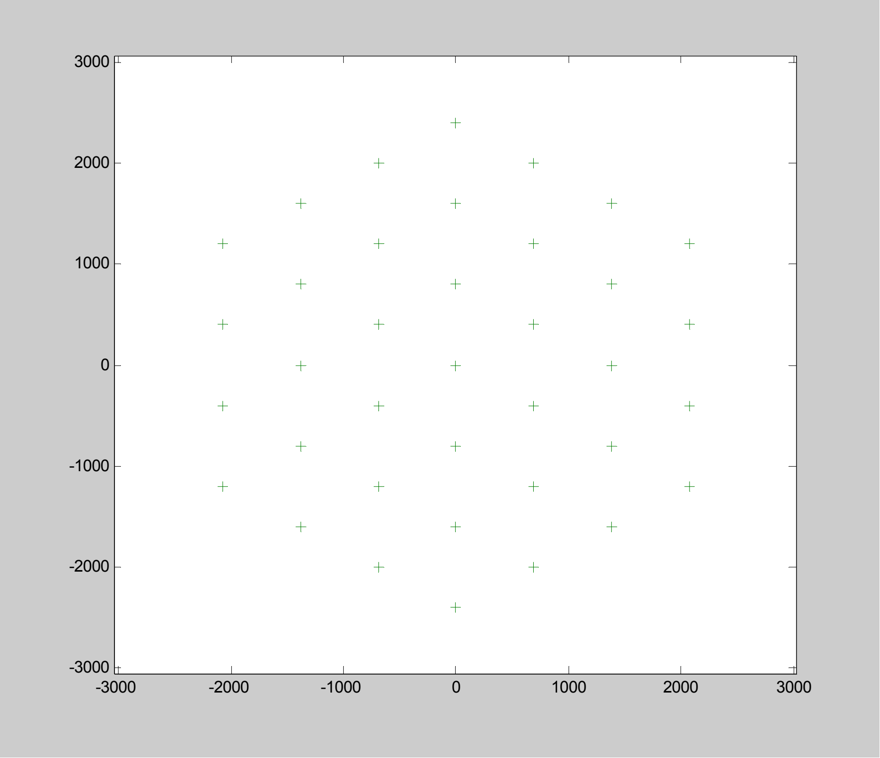
LOFAR Calibration with Compact Central Core

- The full array can be calibrated given a compact central core.
- The wisdom of the LOFAR design is confirmed by CRB analysis

Full array



Compact core



LOFAR Calibration with Compact Central Core

- Calibration succeeds for $Q \leq M_c+1$.
  Q calibrator sources and an M_c element core.