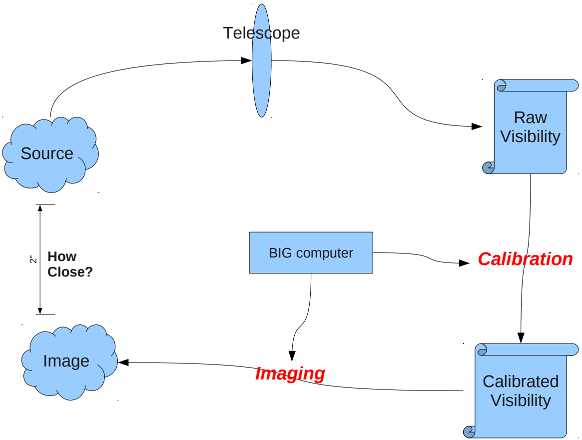
Radio interferometer calibratability and its limits

Tobia Carozzi

Onsala Space Observatory
Chalmers University, Sweden

3GC-II workshop, Albufeira Portugal, 23 Sept 2011

Calibratability



Measurement Equation

Radio interferometric measurement equation MEq is a linear relationship between Visibility and Brightness via Gains

$$\begin{bmatrix} \vdots \\ V_{pq} \\ \vdots \end{bmatrix} \longleftrightarrow \begin{bmatrix} . & . & . \\ . & G_{pqs} & . \\ . & . & . \end{bmatrix}, \begin{bmatrix} \vdots \\ B_s \\ \vdots \end{bmatrix}$$

Calibration is the process of determining **G** and applying it in the MEq above enabling *Imaging*, which is the inversion problem

$$\begin{bmatrix} \vdots \\ B_s \\ \vdots \end{bmatrix} \longleftrightarrow \mathrm{Inv} \begin{bmatrix} . & . & . \\ . & G_{pqs} & . \\ . & . & . \end{bmatrix}, \begin{bmatrix} \vdots \\ V_{pq} \\ \vdots \end{bmatrix}$$

Chimera of calibration

Conjecture

If I know my gains perfectly, then I can image perfectly :-)

Corollary

Performance of hardware is not important, so long as I know its gains (calibrate away deficiencies in software)

Counter Example

Along beam-null, NO amount of calibration will produce sensible image :-(

Chimera of calibration

Conjecture

If I know my gains perfectly, then I can image perfectly :-)

Corollary

Performance of hardware is not important, so long as I know its gains (calibrate away deficiencies in software)

Counter Example

Along beam-null, NO amount of calibration will produce sensible image :-(

Chimera of calibration

Conjecture

If I know my gains perfectly, then I can image perfectly :-)

Corollary

Performance of hardware is not important, so long as I know its gains (calibrate away deficiencies in software)

Counter Example

Along beam-null, NO amount of calibration will produce sensible image :-(

Fundamental theorem of Calibration

Definition

Calibratability (or Imagability) is the degree to which the gains in a MEq are invertible

Conjecture

In general, the conditioning of MEq sets the limits of calibratability

Fundamental theorem of Calibration

Definition

Calibratability (or Imagability) is the degree to which the gains in a MEq are invertible

Conjecture

In general, the conditioning of MEq sets the limits of calibratability

Why Calibratability is Important

- As an interferometer design tool: construction and observation scheduling
- It's the calibratability, stupid!
  - Computational muscle is not the end all of CalIm: it's applying it where/when it makes a difference
- Working out whether your existing image (using your favorite algorithm) can be improved upon
  - Sets ultimate limits of imaging
- Performance metric for your measurements

Why Calibratability is Important

- As an interferometer design tool: construction and observation scheduling
- It's the calibratability, stupid!
  - Computational muscle is not the end all of CalIm: it's applying it where/when it makes a difference
- Working out whether your existing image (using your favorite algorithm) can be improved upon
  - Sets ultimate limits of imaging
- Performance metric for your measurements

Why Calibratability is Important

- As an interferometer design tool: construction and observation scheduling
- It's the calibratability, stupid!
  - Computational muscle is not the end all of CalIm: it's applying it where/when it makes a difference
- Working out whether your existing image (using your favorite algorithm) can be improved upon
  - Sets ultimate limits of imaging
- Performance metric for your measurements

Why Calibratability is Important

- As an interferometer design tool: construction and observation scheduling
- It's the calibratability, stupid!
    - Computational muscle is not the end all of CalIm: it's applying it where/when it makes a difference
- Working out whether your existing image (using your favorite algorithm) can be improved upon
    - Sets ultimate limits of imaging
- Performance metric for your measurements

Why Calibratability is Important

- As an interferometer design tool: construction and observation scheduling
- It's the calibratability, stupid!
  - Computational muscle is not the end all of CalIm: it's applying it where/when it makes a difference
- Working out whether your existing image (using your favorite algorithm) can be improved upon
  - Sets ultimate limits of imaging
- Performance metric for your measurements

Why Calibratability is Important

- As an interferometer design tool: construction and observation scheduling
- It's the calibratability, stupid!
    - Computational muscle is not the end all of CalIm: it's applying it where/when it makes a difference
- Working out whether your existing image (using your favorite algorithm) can be improved upon
    - Sets ultimate limits of imaging
- Performance metric for your measurements

Calibratability Microcosm: polarimetry!

Basic (Jones) Measurement Equation for interferometer element is 2x2 problem

$$\mathbf{V} = \mathbf{J}\mathbf{e}$$

where $\mathbf{V}$ is measured voltages, $\mathbf{e}$ is Jones vector and $\mathbf{J}$ is "Jones" matrix. *Full polarimetric calibration* is the inversion

$$\hat{\mathbf{e}} = \mathbf{J}^{-1}\mathbf{V}$$

This seems to give perfect solutions. . .

But there's always noise and errors & the inversion is prone to errors...

Mathematically the condition number (of the Jones matrix) determines the inversions sensitivity to error propagation, i.e. calibratability.

But instead of matrix condition for calibratability (obscure to many radio astronomers due to lack of physical meaning) I suggest a related parameter to do with feed "leakiness"

Calibratability Microcosm: polarimetry!

Basic (Jones) Measurement Equation for interferometer element is 2x2 problem

$$\mathbf{V} = \mathbf{J}\mathbf{e}$$

where $\mathbf{V}$ is measured voltages, $\mathbf{e}$ is Jones vector and $\mathbf{J}$ is "Jones" matrix. *Full polarimetric calibration* is the inversion

$$\hat{\mathbf{e}} = \mathbf{J}^{-1}\mathbf{V}$$

This seems to give perfect solutions. . .

But there's always noise and errors & the inversion is prone to errors...

Mathematically the condition number (of the Jones matrix) determines the inversions sensitivity to error propagation, i.e. calibratability.

But instead of matrix condition for calibratability (obscure to many radio astronomers due to lack of physical meaning) I suggest a related parameter to do with feed "leakiness"

Calibratability Microcosm: polarimetry!

Basic (Jones) Measurement Equation for interferometer element is 2x2 problem

$$\mathbf{V} = \mathbf{J}\mathbf{e}$$

where $\mathbf{V}$ is measured voltages, $\mathbf{e}$ is Jones vector and $\mathbf{J}$ is "Jones" matrix. *Full polarimetric calibration* is the inversion

$$\hat{\mathbf{e}} = \mathbf{J}^{-1}\mathbf{V}$$

This seems to give perfect solutions. . .
But there's always noise and errors & the inversion is prone to errors...
Mathematically the condition number (of the Jones matrix) determines the inversions sensitivity to error propagation, i.e. calibratability.
But instead of matrix condition for calibratability (obscure to many radio astronomers due to lack of physical meaning) I suggest a related parameter to do with feed "leakiness"

Calibratability Microcosm: polarimetry!

Basic (Jones) Measurement Equation for interferometer element is 2x2 problem

$$\mathbf{V} = \mathbf{J}\mathbf{e}$$

where $\mathbf{V}$ is measured voltages, $\mathbf{e}$ is Jones vector and $\mathbf{J}$ is "Jones" matrix.
Full polarimetric calibration is the inversion

$$\hat{\mathbf{e}} = \mathbf{J}^{-1}\mathbf{V}$$

This seems to give perfect solutions. . .
But there's always noise and errors & the inversion is prone to errors...
Mathematically the condition number (of the Jones matrix) determines the inversions sensitivity to error propagation, i.e. calibratability.
But instead of matrix condition for calibratability (obscure to many radio astronomers due to lack of physical meaning) I suggest a related parameter to do with feed "leakiness"

What's leaky and what's "bad" calibration

There's leakiness and then there's "proper" leakiness:

Figure: Is this a leaky crossed dipole feed? (ans: Yes, leaky)

Figure: Is this also a leaky feed? (ans: No, it's calibratable via coord sys transformation)

What's leaky and what's "bad" calibration

There's leakiness and then there's "proper" leakiness:

Figure: Is this a leaky crossed dipole feed? (ans: Yes, leaky)

Figure: Is this also a leaky feed? (ans: No, it's calibratable via coord sys transformation)

What's leaky and what's "bad" calibration

There's leakiness and then there's "proper" leakiness:

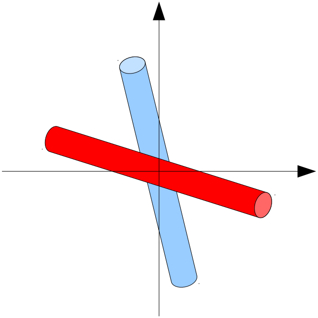


Figure: Is this a leaky crossed dipole feed? (ans: Yes, leaky)

Figure: Is this also a leaky feed? (ans: No, it's calibratable via coord sys transformation)

What's leaky and what's "bad" calibration

There's leakiness and then there's "proper" leakiness:

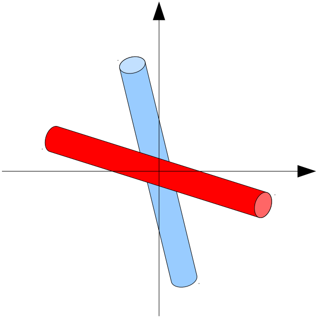


Figure: Is this a leaky crossed dipole feed? (ans: Yes, leaky)

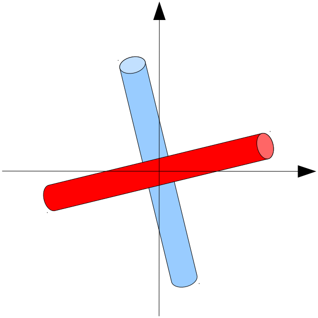


Figure: Is this also a leaky feed? (ans: No, it's calibratable via coord sys transformation)

What's leaky and what's "bad" calibration

There's leakiness and then there's "proper" leakiness:

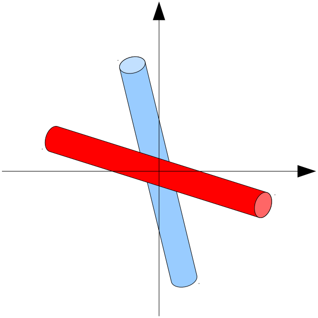


Figure: Is this a leaky crossed dipole feed? (ans: Yes, leaky)

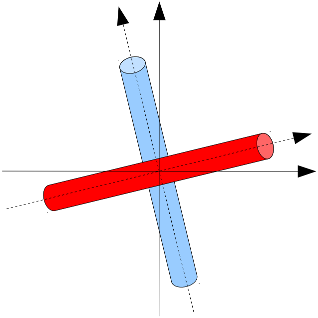


Figure: Is this also a leaky feed? (ans: No, it's calibratable via coord sys transformation)

Cross polarization ratio (XPR)...

So in the latter case, Jones matrix is factorizable as follows

$$\mathbf{J} = g \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} = g\cos\alpha \begin{pmatrix} 1 & \tan\alpha \\ -\tan\alpha & 1 \end{pmatrix} = g\cos\alpha \begin{pmatrix} 1 & d \\ -d & 1 \end{pmatrix}$$

where $d \neq 0$ is the "raw" leakage term (a.k.a d-term). (See Hamaker, Sault, Bregman)
But a change of coordinates to rotated frame (i.e. calibration of alignment) gives

$$\mathbf{J}' = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \mathbf{J} = g \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

which has $d = 0$! Thus, "raw" leakage may be possible to calibrate away

...and Intrinsic cross polarization ratio (IXR)

On the other hand, the SVD factorization is invariant to coordinate transformation: Jones matrix can always be written

$$\mathbf{J} = g\mathbf{U}\begin{pmatrix} 1 & d_{\text{intrinsic}} \\ d_{\text{intrinsic}} & 1 \end{pmatrix}\mathbf{V}^{\dagger}, \quad \mathbf{U},\mathbf{V} \text{ unitary}$$

so there is a choice of sky and feeds coord-sys for which the Jones matrix is

$$\mathbf{J}' = g\begin{pmatrix} 1 & d_{\text{intrinsic}} \\ d_{\text{intrinsic}} & 1 \end{pmatrix}\mathbf{V}^{\dagger}$$

where $d_{\text{intrinsic}}$ is related to the maximum and minimum amplitude gains g_{max}, g_{min} of the polarimeter.

Thus "proper, uncalibratable" leakage is given by the Intrinsic cross polarization ratio

$$\text{IXR} = \frac{1}{|d_{\text{intrinsic}}|^2} = \frac{g_{\text{max}} + g_{\text{min}}}{g_{\text{max}} - g_{\text{min}}} = \frac{g_{\text{max}}/g_{\text{min}} + 1}{g_{\text{max}}/g_{\text{min}} - 1} = \frac{\text{cond}(\mathbf{J}) + 1}{\text{cond}(\mathbf{J}) - 1}$$

where $\text{cond}(\mathbf{J})$ is the Jones condition number

Limits of Calibratability

Ultimately the relationship between calibratability and IXR come from the provable relationship

$$\mathrm{rel.RMS}(\hat{\mathbf{e}}) \equiv \frac{\|\Delta \mathbf{e}\|}{\|\mathbf{e}\|} \lessapprox \left(1 + \frac{2}{\sqrt{\mathrm{IXR}}} + \dots\right)\left(\frac{\|\Delta \mathbf{J}\|}{\|\mathbf{J}\|} + \frac{\|\Delta \mathbf{V}\|}{\|\mathbf{V}\|}\right),$$

where $\Delta \mathbf{V}$ is thermal noise in data and $\Delta \mathbf{J}$ is the imprecision in the Jones matrix

(These results are given in *Carozzi, Woan* IEEE TAP special issue “Future radio telescopes” June 2011)

Calibratability and Antenna Sensitivity

Calibratability is link to antenna sensitvity. Sensitivity can be extended polarimetrically

$$\frac{A_{\mathrm{eff}}}{T} \Longrightarrow \frac{\|\mathbf{M}\|}{\|\mathbf{T}\|}$$

where $\mathbf{M}$ is the effective Mueller matrix, $\mathbf{T}$ is the Stokes antenna temperature and $\|\cdot\|$ is a matrix/vector norm.
A related parameter is SNR of the Stokes estimate from the telescope

$$\frac{\|\mathbf{S}\|}{\|\Delta\mathbf{S}\|} \gtrsim \left(1 - \frac{2}{\mathrm{IXR}}\right)\left(\frac{\|\mathbf{M}\|}{\|\mathbf{T}\|}\|\mathbf{S}\| - \frac{\|\Delta\mathbf{M}\|}{\|\mathbf{M}\|}\right)$$

Mueller IXR

Equivalently in the Mueller formalism, the calibratability of

$$\mathbf{S}' = \mathbf{M}\mathbf{S}$$

where **S**, **S**′is the true and measured Stokes parameters and **M** is the telescopes Mueller matrix, is ultimately determined by

$$\mathrm{IXR}_M = \frac{G_{\mathrm{max}} + G_{\mathrm{min}}}{G_{\mathrm{max}} - G_{\mathrm{min}}}$$

the intrinsic Mueller cross-polarization ratio. Revealingly,IXR_M is identical to what is known as *instrumental polarization*

Interferometer IXR

Continuing the preceding treatment of polarimetric calibratability to interferometry, we have

$$\mathbf{S}_{pq} = \mathbf{M}_{pq}\mathbf{S}^{\mathrm{bri}}$$

where $\mathbf{S}_{pq}$ is the Stokes visibility (complex), $\mathbf{S}^{\mathrm{bri}}$ is the Stokes brightness (real), and $\mathbf{M}_{pq}$ is the interferometer Mueller matrix (complex, not real!). Again an intrinsic value can be analogously assigned

$$\mathrm{IXR}_I = \frac{G^{pq}_{\mathrm{max}} + G^{pq}_{\mathrm{min}}}{G^{pq}_{\mathrm{max}} - G^{pq}_{\mathrm{min}}}$$

Simple vector MEq

Simple model of MEq: for $N = n(n-1)$ scalar visibilities and m point sources and equal gains G, then

$$\begin{bmatrix} V_{12} \\ \vdots \\ V_{pq} \\ \vdots \\ V_{(n-1)n} \end{bmatrix} = G \begin{bmatrix} e^{iu_{12}l_1} & \dots & e^{iu_{12}l_s} & \dots & e^{iu_{12}l_m} \\ \vdots & . & \vdots & . & \vdots \\ e^{iu_{pq}l_1} & \dots & e^{iu_{pq}l_s} & \dots & e^{iu_{pq}l_m} \\ \vdots & . & \vdots & . & \vdots \\ e^{iu_{(n-1)n}l_1} & \dots & e^{iu_{(n-1)n}l_s} & \dots & e^{iu_{(n-1)n}l_m} \end{bmatrix} \begin{bmatrix} B_1 \\ \vdots \\ B_s \\ \vdots \\ B_m \end{bmatrix}$$

$$\mathbf{V} = G\mathbf{AB}$$

Formal solution is

$$\mathbf{B} = G^{-1}\mathbf{A}^{-1}\mathbf{V}$$

If $\mathbf{A}$ is singular we can use its pseudo-inverse instead so

$$\mathbf{B} = G^{-1}\mathbf{A}^{+}\mathbf{V}$$

Simple error model

However in practice there are errors, due noise, incomplete knowledge of gains and pointing errors. A simple model for errors in previous MEq is just

$$\mathbf{V} + \Delta\mathbf{V} = (G + \Delta G)(\mathbf{A} + \Delta\mathbf{A})(\mathbf{B} + \Delta\mathbf{B})$$

The relative error can be shown to be

$$\frac{|\Delta\mathbf{B}|}{|\mathbf{B}|} \leq \underbrace{\|\mathbf{A}\|\,\|\mathbf{A}^{-1}\|}_{\mathrm{cond}(\mathbf{A})}\left(\frac{|\Delta\mathbf{V}|}{|\mathbf{V}|} + \frac{\|\Delta\mathbf{A}\|}{\|\mathbf{A}\|} + \frac{\Delta G}{G}\right)$$

The crucial parameter here is the condition of **A**, which is in turn dictated by it's singular value spectrum.

Singular value decomposition of MEq

$$\mathbf{V} = \mathbf{U}\mathbf{D}\mathbf{W}^{\dagger}\mathbf{B}$$

where $\mathbf{U}$,$\mathbf{W}$ are unitary matrices and $\mathbf{D}$ is a (positive semi-definite) diagonal matrix. Let $\mathbf{U}^{\dagger}\mathbf{V} = \mathbf{V}'$ and $\mathbf{W}^{\dagger}\mathbf{B} = \mathbf{B}'$ then

$$\mathbf{V}' = \mathbf{D}\mathbf{B}'$$

Solution is simply

$$\mathbf{B}' = \mathbf{D}^{-1}\mathbf{V}'$$

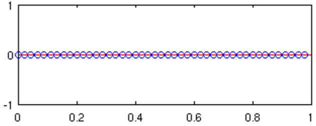
but error is this inversion is factored by

$$\|\mathbf{D}\| \left\|\mathbf{D}^{-1}\right\|$$

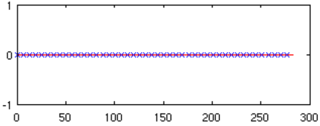
Let us see what the spectrum of singular values is in concrete cases...

Example 1D MEqs: Uniform - Uniform

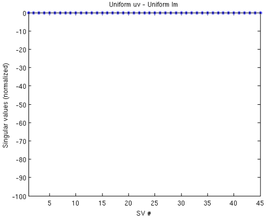
Uniform uv-sampling



Uniform lm-sampling

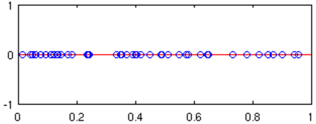


Results

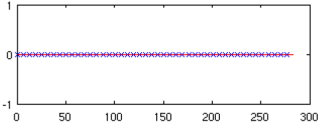


Example 1D MEqs: Poisson - Uniform

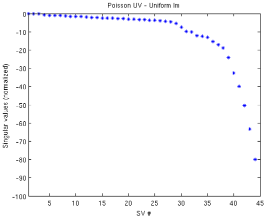
Poissonian uv-sampling



Uniform lm-sampling

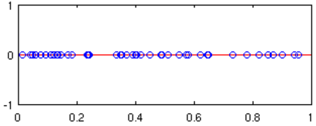


Results

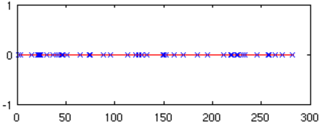


Example 1D MEqs: Poisson - Poisson

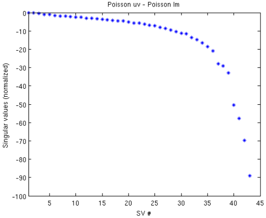
Poissonian uv-sampling



Poissonian lm-sampling

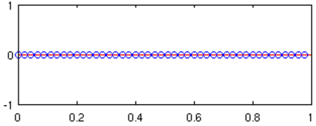


Results

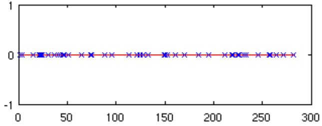


Example 1D MEqs: Uniform - Poisson

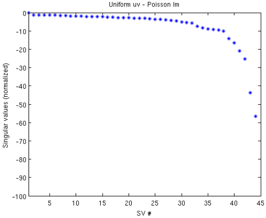
Uniform uv-sampling



Poissonian lm-sampling



Results



MEq performance metric: Synthesized Beam pattern

If we extend the number of visibilities and brightness samples by using masking matrices (essentially appropriate zero-padding)

$$\begin{bmatrix} \mathrm{diag}(\mathbf{w}) & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \vdots \end{bmatrix} = G\mathbf{U}_{DFT} \begin{bmatrix} \mathbf{B} \\ \vdots \end{bmatrix}$$

where $\mathbf{U}_{DFT}$ is a discrete Fourier transform matrix and $\mathbf{w}$ is a weights vector of length $n(n-1)$.

MEq performance metric: MEq Conditioning

Rather than FoMs based on synthesized beam shape, the conditioning of a MEq (with a given source positions and given gains) gives the rms relative error in final image estimate.

MEq full matrix condition may not a be directly sensible number in radio astronomy, so work is underway to develope a related parameter (like IXR) that makes more sense. Current idea is to use the amount of information transfered through MEq matrix.

Ultimatively, one can used the final rms relative error for the estimated image. Compared to dynamic range, this performance metric includes the "error bars" on the fluxs.

Conclusions

- Neither computational muscle nor algorthimic might is all there is to Cal & Im in future "software telescopes"
  - Bad telescope design can never be replaced by clever software
  - Some things can never be "calibrated away"
- IXR characterizes polarimetric calibratability
- Condition full RIME is better alternative to FoMs based on beam shape
  - since it gives images total rms relative error