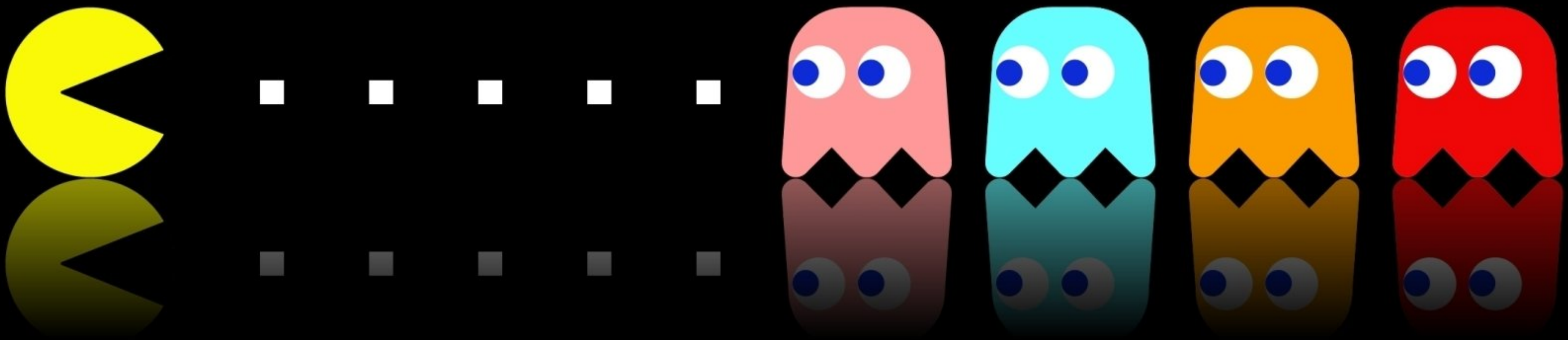
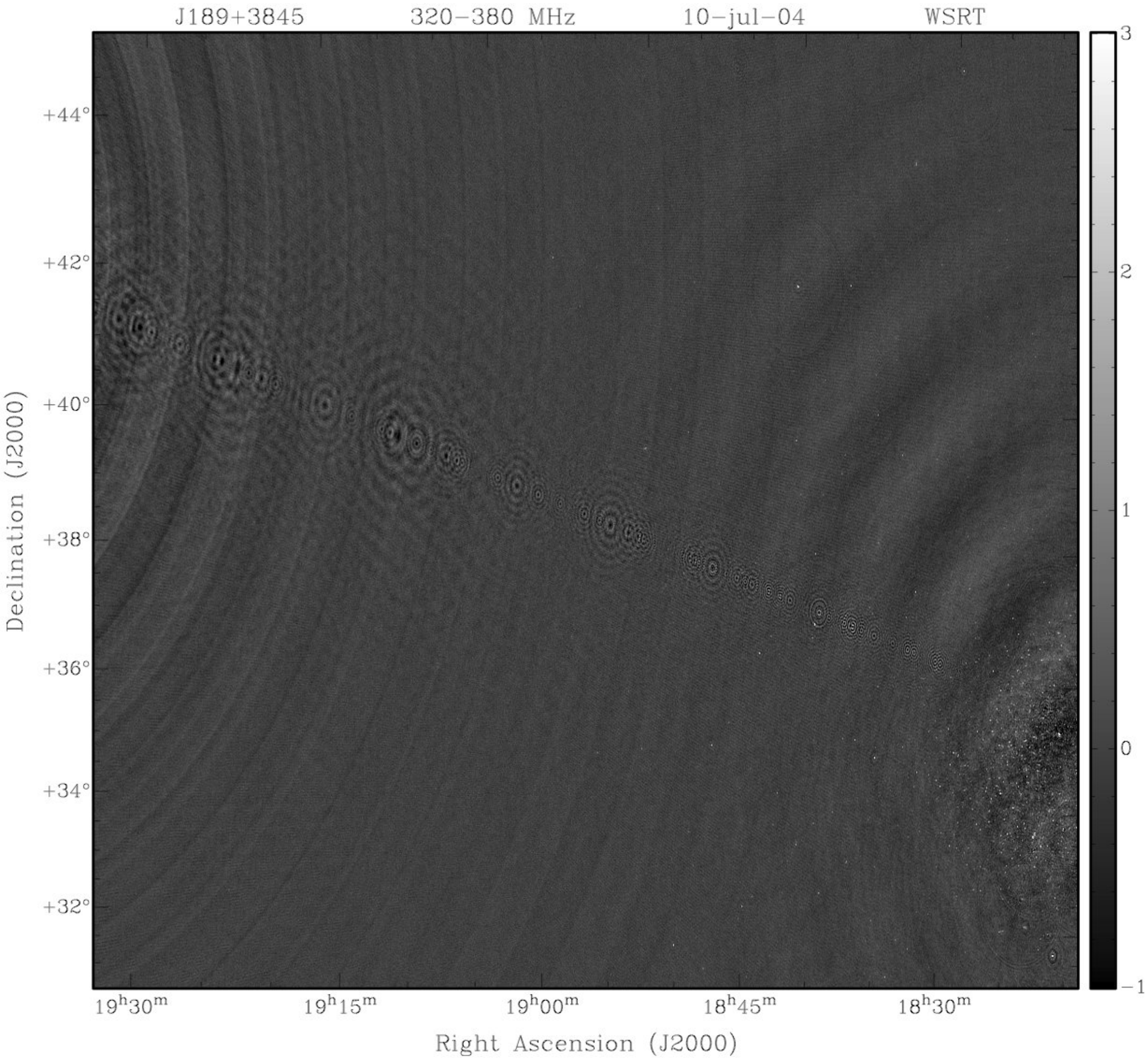
Ghostbusted



O. Smirnov, T. Grobler (Rhodes U. & SKA SA)
C.D. Nunhokee (Rhodes U.)

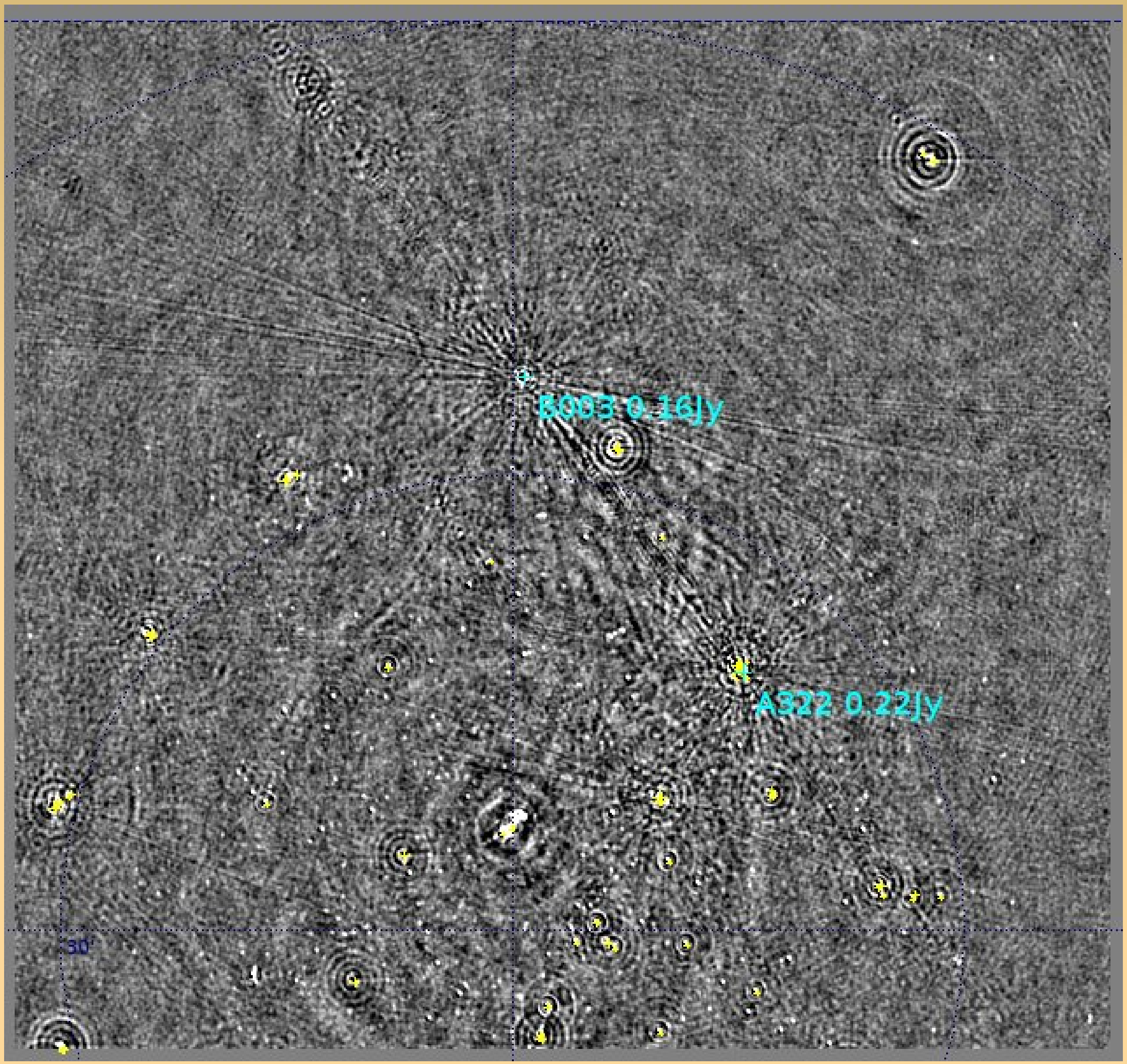
Ghosts Of Selfcal Past



The early mystery:
WSRT 92cm observation of J1819+3845 by Ger

- String of ghosts connecting brightest source to Cyg A (20° away!)
- “Skimming pebbles in a pond”
- Positions correspond to rational fractions (1/2, 1/3, 2/3, 2/5, etc...)
- Wasn't clear if they were a one-off correlator error, a calibration artefact, etc.
  - (...and if you did low-frequency in 2004, you had it coming to you anyway.)

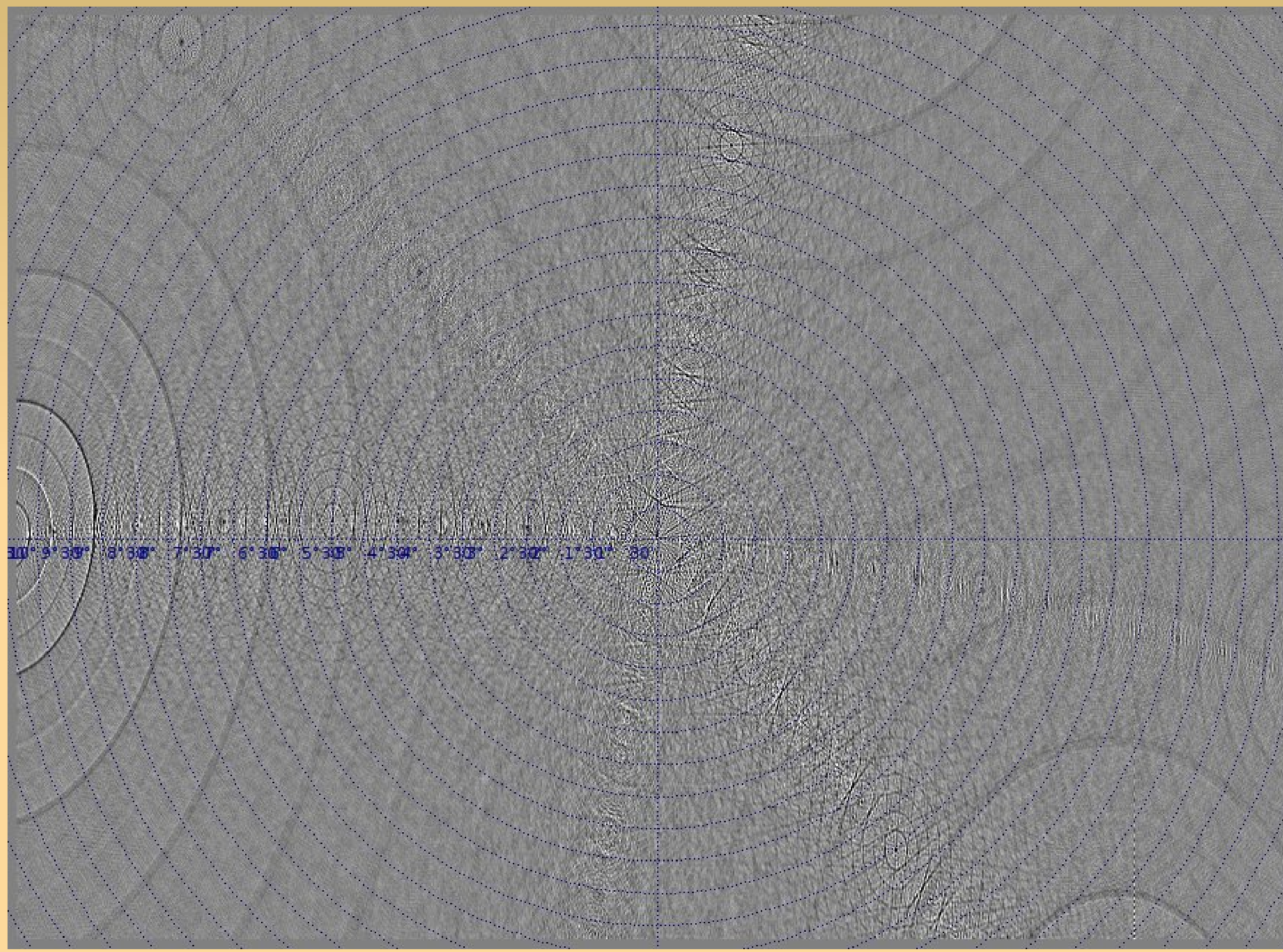
2010: Ghosts Return



WSRT 21cm observation (QMC2 field)

- ...with intentionally high pointing errors
- String of ghosts through dominant sources A (220 Jy) and B (160 mJy)
- Second, fainter, string from source A towards NNE
- Qualitatively similar to CygA ghosts
- Went away after DD calibration & repeated selfcal

Ghostbusters (CALIM 2010)

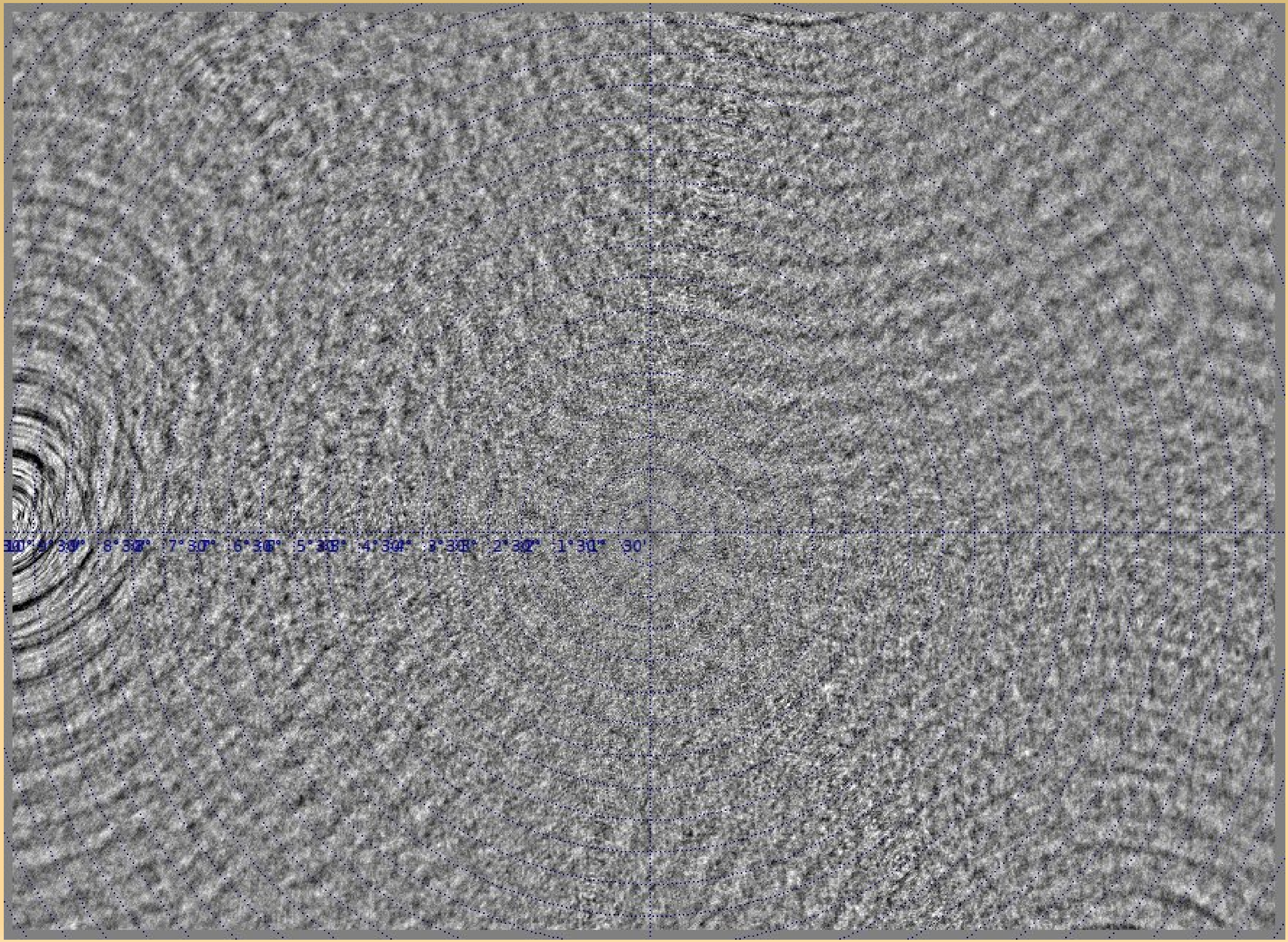


- Ghosts reproduced via simulations

Ghosts In The (Selfcal) Machine

- Ghosts arise due to *missing flux* in the calibration sky model
- Mechanism: selfcal solutions try to compensate for this by moving flux around
  - Not enough DoFs to do this perfectly
  - ...so end up dropping flux all over the map
- Regular structure suspected to be due to WSRT's redundant layout
  - JVLA, MeerKAT: “random” (but not Gaussian!)

JVLA Ghost Sim



Ghastly Mysteries

- Shown empirically (2010) but not understood:
  - Why do they form on lines passing through unmodelled sources?
  - Why do they sit on rational fractions?
  - Why do they have different PSFs?
  - Why do they seem to scale with the missing flux, but not with the model flux?

Understanding Ghosts

- Results suggest ghosts are fundamental to selfcal
  - ...and we really couldn't let Ger retire with the mystery unsolved
- Trienko Grobler and Ridhima Nunhokee worked on the problem

Fundamentals

- Calibration: $\mathcal{R} \leftarrow G\mathcal{M}G^H, \quad G = \operatorname{diag} \vec{g}$

$$\begin{bmatrix} r_{11} & \cdots & r_{1n} \\ \vdots & \text{observed} & \vdots \\ r_{n1} & \cdots & r_{nn} \end{bmatrix} \leftarrow \begin{bmatrix} g_1 & \cdots & 0 \\ \vdots & \text{gains} & \vdots \\ 0 & \cdots & g_n \end{bmatrix} \begin{bmatrix} m_{11} & \cdots & m_{1n} \\ \vdots & \text{model} & \vdots \\ m_{n1} & \cdots & m_{nn} \end{bmatrix} \begin{bmatrix} g_1^* & \cdots & 0 \\ \vdots & \text{gains} & \vdots \\ 0 & \cdots & g_n^* \end{bmatrix}$$

...or equivalently:

Element-by-element multiplication

$$\mathcal{R} \leftarrow \begin{bmatrix} g_1 g_1^* m_{11} & \cdots & g_1 g_n^* m_{1n} \\ \vdots & \ddots & \vdots \\ g_n g_1^* m_{n1} & \cdots & g_n g_n^* m_{nn} \end{bmatrix} = \mathcal{G} \odot \mathcal{M}, \quad \mathcal{G} = \vec{g}\vec{g}^H$$

Element-by-element inverse

- Correction: $\mathcal{R}_{\text{corr}} = G^{-1}\mathcal{R}G^{-H} = \mathcal{G}^{\top} \odot \mathcal{R}$

The Simplest Case

- *R*: two point sources (1 Jy at centre, <1 at l_0, m_0)
  M: 1 Jy source at centre = matrix of all ones

$$\mathcal{R} \leftarrow \mathcal{G}, \quad \mathcal{G} = \vec{g}\vec{g}^H$$

"calibrated sky"

- Conventional calibration: "←" is an LSQ fit of off-diagonal terms (=Gaussian ML)
  - (though: see robust calibration, Kazemi & Yatawatta)
  - what does it do? God only knows, very difficult to understand analytically...
  - ...some crucial insights were needed

Some Familiar Names...

GAIN DECOMPOSITION METHODS FOR RADIO TELESCOPE ARRAYS

A.J. Boonstra[1,2] *and A.J. van der Veen*[2]

[1] ASTRON, Netherlands Foundation for Research in Astronomy
Oude Hoogeveensedijk 4, 7991 PD Dwingeloo, The Netherlands
Tel: +31(0)521 595100, fax: +31(0)521 597332, email: boonstra@astron.nl

[2] Department of Electrical Engineering, Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands
Tel: +31(0)15 2781372, fax: +31(0)15 2786190, email: allejan@cas.et.tudelft.nl

(Boonstra A.J., van der Veen A.J., 2003, IEEE Trans. Sig. Proc., 51, 25)

“The phantom memo”

7. REFERENCES

[1] H. van Someren-Greve. Logarithmic least square gain decomposition algorithm for the WSRT. IWOS software documentation, ASTRON internal document, 1980.

[2] R.A. Perley F.R. Schwab and A.H.Bridle. Synthesis imaging in radio astronomy. *Astronomical Society of the Pacific Conference Series*, 6, 1994.

[3] A.B.Smolders and M.P. van Haarlem, editors. *Perspectives on Radio Astronomy: Technologies for Large Antenna Arrays*

ALS Calibration

$$\mathcal{R} \leftarrow \mathcal{G}, \quad \mathcal{G} = \vec{g}\vec{g}^H$$

- *R* is rank two (for two sources), *G* is rank one by construction
- ALS – "deranking" – builds *G* by taking just the largest eigenvalue/eigenvector of *R*
  - Not exactly the same as off-diagonal LSQ...
  - ...but we've empirically shown that this produces similar ghost patterns
  - ...and deranking can be studied analytically

Regular spacing

- WSRT is regularly-spaced: there's always a "common quotient baseline" (CQB) $\boldsymbol{b}_0$ such that for all baselines there is a whole-number scaling relationship:

whole number

$$\vec{u}_{pq} = \vec{b}_0 \phi_{pq}$$

"array geometry matrix"

$$\begin{bmatrix} \vec{u}_{11} & \cdots & \vec{u}_{1n} \\ \vec{u}_{21} & \cdots & \vec{u}_{2n} \\ \vdots & \ddots & \vdots \\ \vec{u}_{n1} & \cdots & \vec{u}_{nn} \end{bmatrix} = \vec{b}_0 \begin{bmatrix} 0 & \phi_{12} & \cdots & \phi_{1n} \\ \phi_{21} & 0 & \cdots & \phi_{2n} \\ \vdots & \vdots & 0 & \vdots \\ \phi_{n1} & \phi_{n2} & \cdots & 0 \end{bmatrix} = \vec{b}_0 \mathbf{\Phi}$$

Per-baseline "calibrated sky"

- Deranking allows us to work out ***G*** analytically...
  - (...lots of math skipped, see paper...)
- **<u>Key result</u>**: the "calibrated sky" seen by each baseline *pq* is an infinite string of delta-functions of varying intensity, placed at intervals inversely proportional to ϕ_{pq}:

$$I_{pq}(l,m) = \sum_{j=-\infty}^{\infty} c^{\mathcal{G}}_{j,pq}\,\delta(l + \frac{jl_0}{\phi_{pq}}, m + \frac{jm_0}{\phi_{pq}})$$

ghost intensity coefficients
(can be worked out numerically)

Putting It Together

- Each baseline "sees" its own ghost string with intervals of ϕ_{pq}
- The combined effect is some sort of average (depends on imaging weights, etc.)
- Because of the whole-number scaling relationship, ghosts occupy a discrete set of positions (i.e. rational fractions of l_0, m_0)
  - redundancy means that some positions are "preferred"
  - Amplitude coefficients differ per baseline, hence each ghost position exhibits its own "GSF" ≠ PSF

Distilled Ghost Pattern

- This pattern is then translated into the "corrected sky" image, given by

$$\mathcal{R}_{\mathrm{corr}} = G^{-1} \mathcal{R} G^{-H} = \mathcal{G}^{\top} \odot \mathcal{R}$$

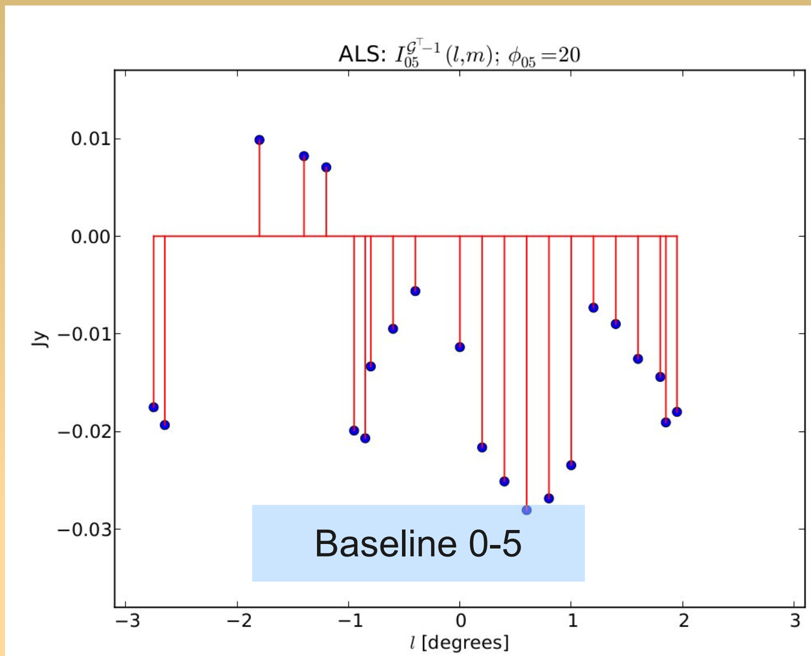
  - ...because $\mathcal{G}^{\top}$ actually has the same string-like structure (with different values for the *c* coefficients)
- The interesting thing is the "distilled" ghost pattern:

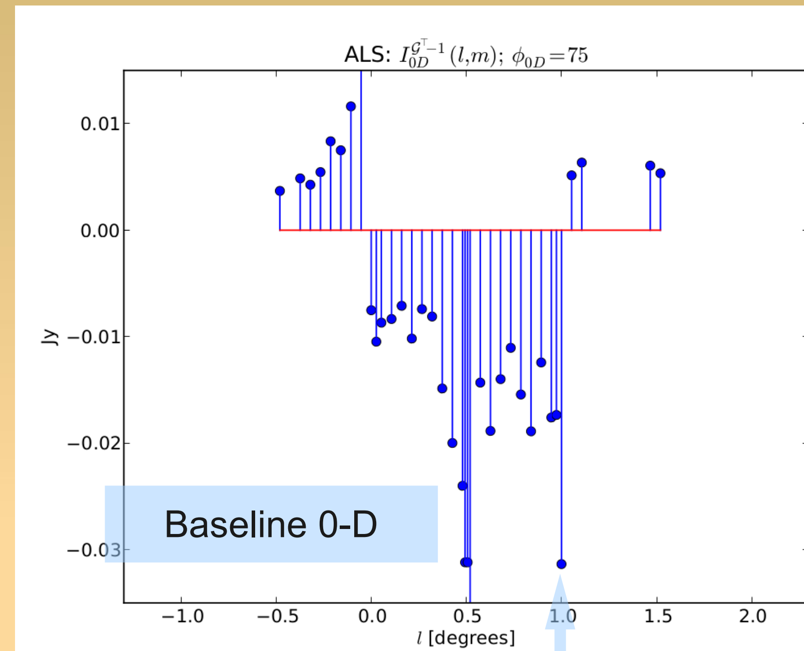
convolution

$$\mathcal{R}_{\mathrm{corr}} - \mathcal{R} = (\mathcal{G}^{\top} - \mathbf{1}) \circ \mathcal{R}$$

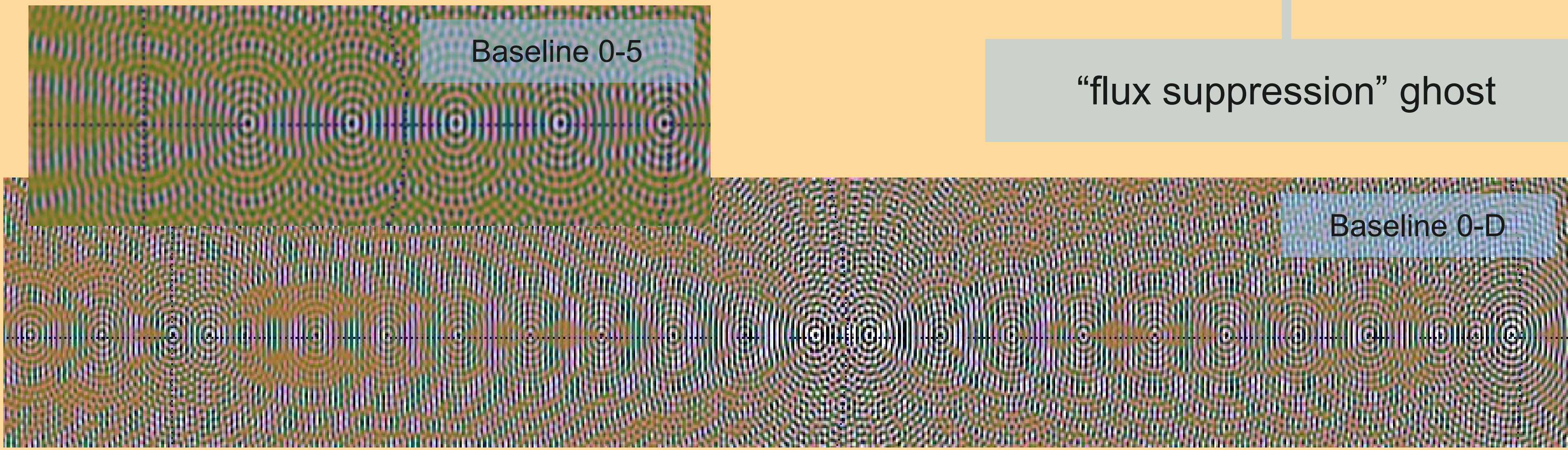
"atomic" ghost pattern

Predicted vs. Observed Patterns





“flux suppression” ghost



Ghost Intensity

- Empirical observation (2010): ghost pattern scaled with intensity of secondary source A_s, but did **not** seem to depend on primary source A_p
- This can now be explained:

Dominated by A_p

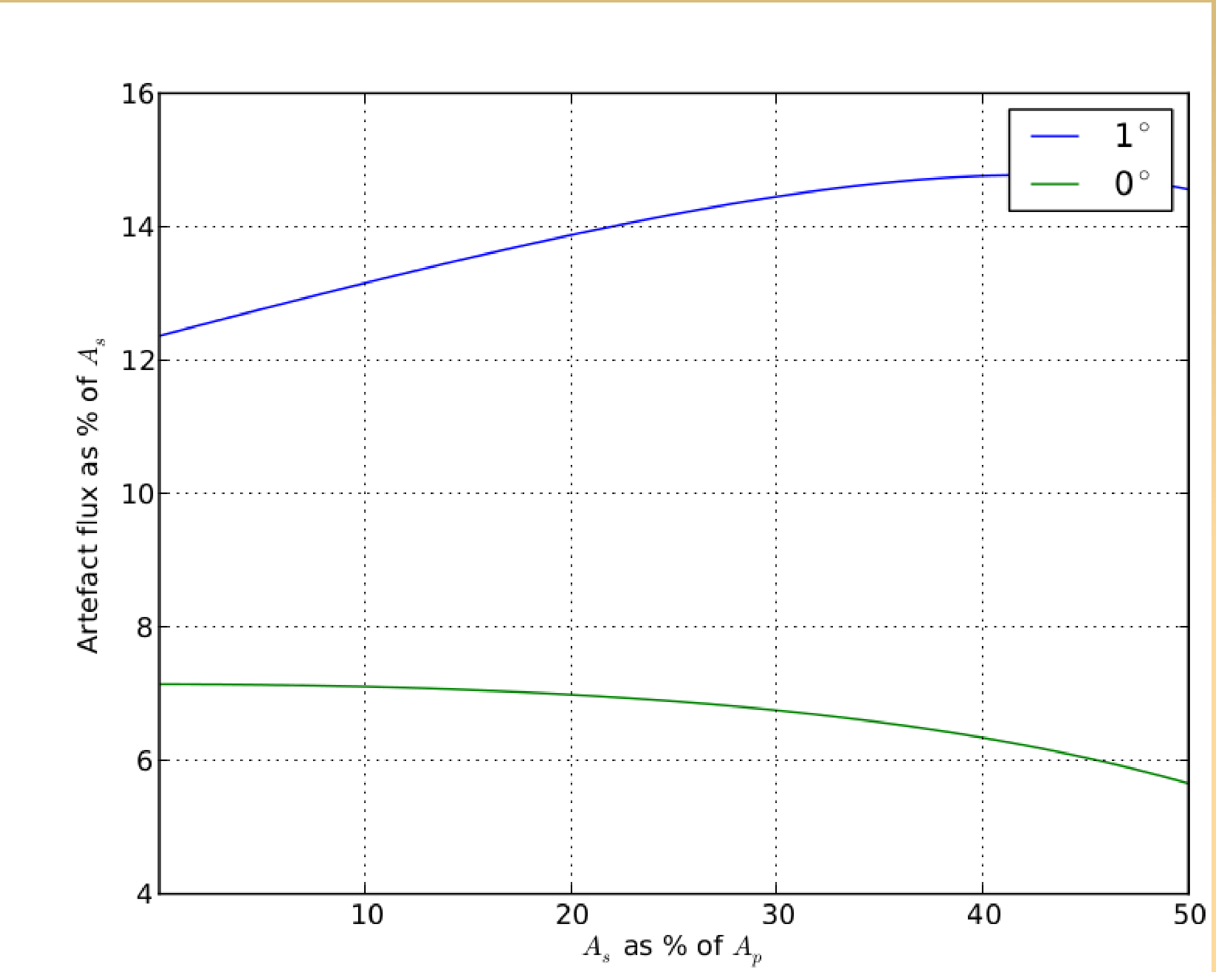
$$\mathcal{R}_{\mathrm{corr}} - \mathcal{R} = (\mathcal{G}^\top - \mathbf{1}) \circ \mathcal{R}$$

Dominated by
A_s/A_p flux ratio

- ...but the full picture is even more interesting...

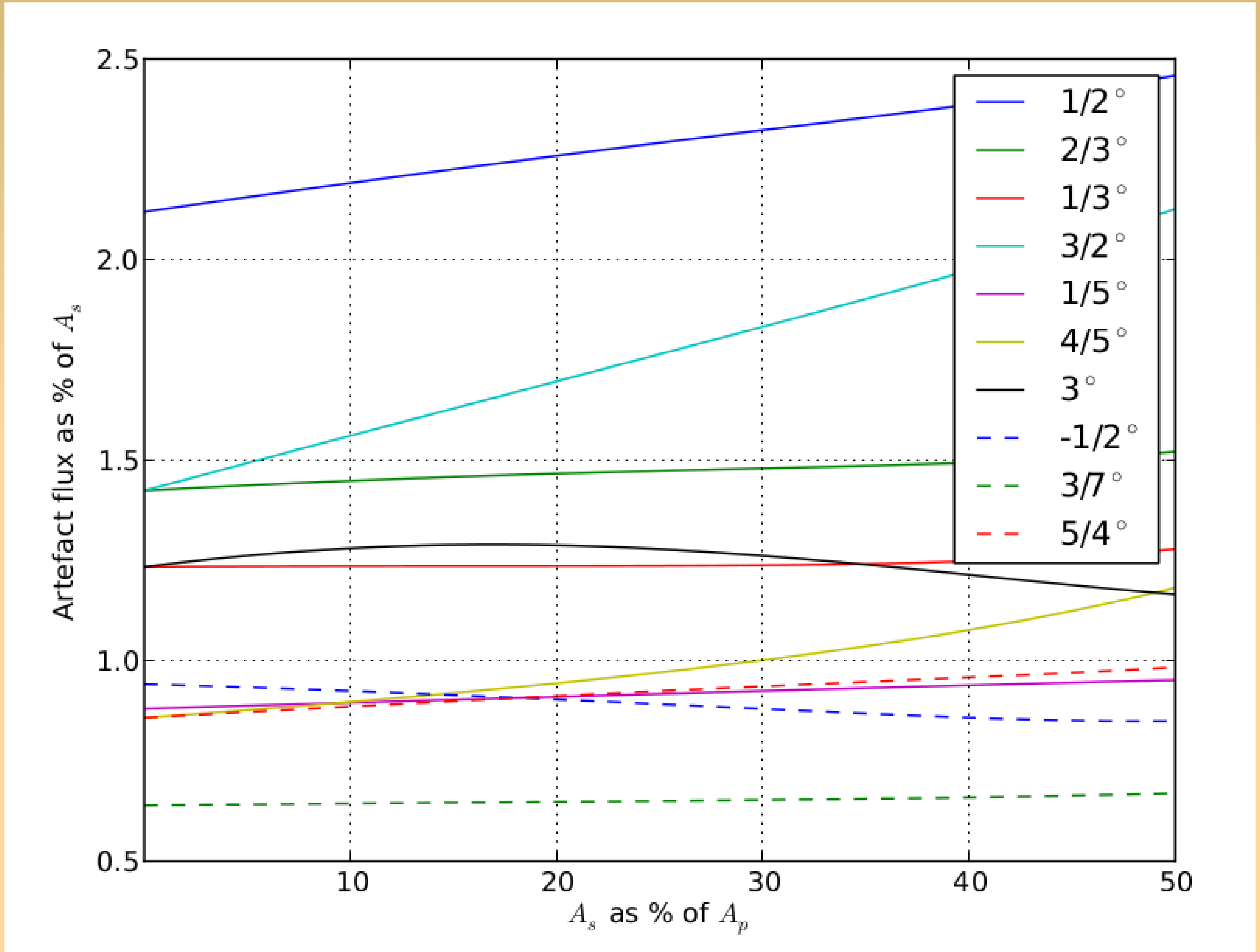
Flux Suppression vs. Flux Ratio

- Ghost at 1 deg sits on top of missing source
- Ghost at 0 deg sits on top of primary source
- Responsible for what we know as flux suppression



- Note the non-trivial dependence!

Flux Of Others



Why Is This Important?

- Affects all instruments
  - Only regularity is special to WSRT
- Ghosts will always exist (in the noise, at least), until you build up a complete sky model
  - Which is very laborious and/or compute-intensive
  - What about other calibration approaches?
- What does this do to the noise statistics?
- Shallow calibration pipelines (AARTFAAC, etc.)
  - Need to identify how deep a model is needed to keep ghosting within acceptable levels

Conclusions

- Ghosts (and WSRT regularity) explained
  - Ger can retire now
- We have a theoretical framework to predict ghost formation, which can and should be extended to other instruments