Calibration and processing input to AWG

Athol Kemball

Department of Astronomy, and
Institute for Advanced Computing and Applications Technologies

University of Illinois at Urbana-Champaign
Calibration & processing input to AWG

- SKA needs a **composite** system design that:
  - Meets reference science mission goals
  - Falls within the overall cost envelope
  - Is technically feasible, with contingencies for identified risk

- Composite system design will be delivered *iteratively* and *incrementally* and *verified*:
  - Most SKA sub-systems contain substantial research uncertainties in areas of cost and feasibility.
  - Not a linear build-to-spec process, rather a collective, iterative narrowing of design parameter space.
  - Strong coupling of many sub-system designs
How can CPG assist AWG activities?

- **Identify** antenna design areas of special calibration and processing risk
- Help to **quantify** antenna design specifications where possible.
SKA: imaging dynamic range challenges

Reference specifications (Schillizzi et al 2007; plus subsequent revisions)

- Targeted $\lambda_{20}$cm continuum field: $10^{7-8}:1$.
- Routine $\lambda_{20}$cm continuum: $10^{6-7}:1$.
- Driven by need to achieve thermal noise limit (nJy) over plausible field integrations.
- Spectral dynamic range: $10^5:1$.

(SKA schematic)  
(de Bruyn & Brentjens (2005))
Dynamic-range and sensitivity: current state of practice

**Dynamic range**

<table>
<thead>
<tr>
<th>Year</th>
<th>Source</th>
<th>Instrument</th>
<th>Frequency</th>
<th>Dynamic Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>1982</td>
<td>Noordarm <em>et al</em></td>
<td>3C84</td>
<td>WSRT 1.4 GHz</td>
<td>10,000:1</td>
</tr>
<tr>
<td>2000</td>
<td>Geller <em>et al</em></td>
<td>1935-692</td>
<td>ATCA 1.4 GHz</td>
<td>77,000:1</td>
</tr>
<tr>
<td>2005</td>
<td>de Bruyn &amp; Brentjens</td>
<td>Perseus</td>
<td>WSRT 92 cm</td>
<td>400,000:1</td>
</tr>
<tr>
<td>2007</td>
<td>de Bruyn <em>et al</em></td>
<td>3C147</td>
<td>WSRT 1.4 GHz</td>
<td>1,000,000:1</td>
</tr>
</tbody>
</table>

**High-sensitivity deep fields**

<table>
<thead>
<tr>
<th>Year</th>
<th>Source</th>
<th>Instrument</th>
<th>Frequency</th>
<th>Sensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>2000</td>
<td>Richards</td>
<td>HDF</td>
<td>VLA 1.4 GHz</td>
<td>7.5 μJy</td>
</tr>
<tr>
<td>2005</td>
<td>Norris <em>et al</em></td>
<td>HDF-S</td>
<td>ATCA 1.4 GHz</td>
<td>10 μJy</td>
</tr>
<tr>
<td>2008</td>
<td>Middelberg <em>et al</em></td>
<td>ELAIS I</td>
<td>ATCA 1.4 GHz</td>
<td>&lt; 30 μJy</td>
</tr>
<tr>
<td>2008</td>
<td>Miller <em>et al</em></td>
<td>E-CDF-S</td>
<td>{E}VLA 1.4 GHz</td>
<td>6.4 μJy</td>
</tr>
<tr>
<td>2009</td>
<td>De Bruyn <em>et al</em></td>
<td>Targeted fields</td>
<td>WSRT 1.4 GHz</td>
<td>4-5 μJy</td>
</tr>
</tbody>
</table>

- At **central pixel** for bright sources.
- Approximately an order-of-magnitude improvement every 10-15 years.

AWG 2009
Dominant calibration and processing risks

Basic imaging and calibration equation for radio interferometry (e.g. Hamaker, Bregman, & Sault (1996ff)), here with implicit (t,ω, polzn) dependence.

\[ V_{mn} = \prod_k \left[ G_m^k \otimes G_m^{k*} \right] \int \prod_k \left[ T_m^k (\tilde{\rho}) \otimes T_m^{k*} (\tilde{\rho}) \right] e^{-2\pi j \bar{b}_{mn} \cdot (\vec{\rho} - \bar{\rho}_s)} K S(\tilde{\rho}) d\Omega \]
Dominant calibration and processing risks

Basic imaging and calibration equation for radio interferometry (e.g. Hamaker, Bregman, & Sault (1996ff)), here with implicit \((t, \omega, \text{polzn})\) dependence.

\[
V_{mn} = \prod_k \left[ G_m^k \otimes G_m^{k*} \right] \int \prod_k \left[ T_m^k(\tilde{\rho}) \otimes P_m^k(\tilde{\rho}) \right] e^{-2\pi i u_m (\rho - \rho_s)} K I(\tilde{\rho}) d\Omega
\]

\[
V_{mn}(u, v) = B_m^k \prod_k \left[ G_m^k \otimes G_m^{k*} \right] \int \Omega S(\eta, \xi) e^{-2\pi j (u\eta + v\xi)} d\Omega
\]

Visibility on baseline \(m-n\)

Image-plane calibration effect

Baseline m-n

Source brightness \((I,Q,U,V)\)

Direction on sky: \(\rho\)

Visibility-plane calibration effect

2-D Fourier kernel

Axisymmetric, constant over time

Discrete, unparameterized

Separable calibration and imaging

The good times…
Calibration and imaging are solved iteratively as variational regularization problem.

**CALIBRATION:** Iteratively fix $S(\tilde{\rho})$ and compute $\chi^2$ at $N$ positions (in integral equation) of unknown calibration Jones matrices. Solve using non-linear LSQ or ML methods: $\frac{\partial \chi^2}{\partial G_m} = \frac{\partial \chi^2}{\partial T_m} = 0$

**IMAGING:** Image formation: can define as update direction for unknown $S(\tilde{\rho})$: (Cornwell 1995, 2008):

$$\left[ \frac{\partial^2 \chi^2}{\partial S(\tilde{\rho}) \partial S^T(\tilde{\rho})} \right]^{-1} \frac{\partial \chi^2}{\partial S(\tilde{\rho})} \bigg|_{S(\tilde{\rho}) = S(\tilde{\rho})} + \text{regularization / deconvolution}$$
Calibration and processing feasibility risks

- Acute imaging **dynamic range** and **sensitivity** requirements (relative to current state of practice)
- **Direction-dependent** calibration terms predominate
- The SKA sky is **complex**; main-lobe confusion and sidelobe contamination are ubiquitous
- Calibration instrumental models (e.g. primary beam response) and signal propagation models (e.g. ionosphere) require **high-order, accurate parameterization**
- **Operational** models (e.g. commensal, high-speed surveys) and **data reduction** methodologies (e.g. real-time) vastly different from current arrays
- Extreme-scale **computational** needs
Direction-dependent calibration: $T_m^k$

- For SKA, direction-dependent terms cannot be ignored
- Two issues:
  - Removal of *known* prior direction-dependent corrections within imaging equation formalism:
    - Presupposes good *a priori* direction-dependent instrumental models (e.g. antenna response pattern).
      - Two approaches: unitary approximation in forward calculation; incorporate direction-dependent calibration in visibility gridding / convolution (e.g. Bhatnager et al. 2008); iterative non-linear deconvolution.
      - Software holography (e.g. Morales et al. 2009); CMB OMM model; image-plane deconvolution with data compression.
  - Solution for *unknown* direction-dependent terms
    - **Peeling**, sequential direction-dependent self-calibration over set of brightest known sources in field (or beam) in decreasing order of flux density (van der Tol *et al.* 2007 (LOFAR), Mitchell *et al.* 2008 (MWA))
    - Multivariate, generalized self-cal with constrained model (work in process).

---

Much easier; increases computational cost not feasibility.

Significant risk to feasibility; increases computational cost also.

*(Bhatnagar *et al.* 2004,2008; antenna pointing self-cal: 12µJy => 1µJy rms)*

AWG 2009
Constraints on antenna design parameters: (Uncorrected) RMS sidelobe level

- Sidelobe contamination (Perley & Clark 2003):

\[ \sigma_c = \sqrt{N_s R S_m} \]

where,

- \( N_s \) = Number of confusing point sources in beam
- \( R \) = Uncorrected RMS sidelobe level
- \( S_m \) = Mean confusing source flux density

- Goal is to achieve thermal-noise limited DR

\[ \sigma_{th} = \beta \frac{SEFD}{\sqrt{N_{ant} (N_{ant} - 1) \Delta t \Delta v}} \]

\[ DR \sim \frac{S_m}{\sigma_{th}} \]
Deriving constraints on antenna design parameters

\[ \chi^2 = \sum_{mn} \left\| V_{mn} - \prod_k \left[ G_m^k \otimes G_m^{k^*} \right] \int_{\Omega} \prod_k \left[ T_m^k(\tilde{\rho}) \otimes T_m^{k^*}(\tilde{\rho}) \right] e^{-2\pi j b_{mn} \cdot (\tilde{\rho} - \bar{\rho}_m)} K S(\tilde{\rho}) d\Omega \right\| \]

\[ \frac{\partial \chi^2}{\partial S(\tilde{\rho})}_{S(\tilde{\rho}) = S_j(\tilde{\rho})} = -2 \text{Re} \sum_{mn} X_{ij}^{*T}(\tilde{\rho}) W_{mn} V_{mn}^{\text{obs}} \]

\[ X_{ij}^{*T}(\tilde{\rho}) = \prod_k \left[ G_m^k \otimes G_m^{k^*} \right] \int_{\Omega} \prod_k \left[ T_m^k(\tilde{\rho}) \otimes T_m^{k^*}(\tilde{\rho}) \right] e^{-2\pi j b_{mn} \cdot (\tilde{\rho} - \bar{\rho}_i)} K S(\tilde{\rho}) d\Omega \]

Specialization to derive zero-order constraints: a) no regularization or deconvolution \( N_a \) large; b) scalar, single-polarization calibration; c) sky model: \( S(\tilde{\rho}) = \delta(\tilde{\rho} - \bar{\rho}_k) \); d) weight: \( W_{ij} = 1 \); e) normalize approximately by \( N_{\text{vis}}^{-1} \).

Estimate variance in sky-model pixel \( S(\tilde{\rho}_k) \) as approximate estimate of inverse DR.
Constraints on antenna design parameters: (Uncorrected) pointing offsets

\[ T^k_m(\tilde{\rho}) = E^k_m(\tilde{\rho}) \]

Retain only Gaussian main-lobe response, spherical symmetry: \( E(\rho) = e^{-4\ln 2 \frac{\rho^2}{\sigma^2}} \),

[Direct PB correction in]

\[
\frac{\partial \chi^2}{\partial S(\tilde{\rho})} \bigg|_{S(\tilde{\rho})=S_i(\tilde{\rho})} = -2 \Re \sum_{mn} X^*_{ij}(\tilde{\rho}) W_{mn} V^{obs}_{mn}
\]

\[ DR \sim \frac{1}{\sigma_{S(\tilde{\rho})}} \sim \frac{1}{5} \left( \frac{\theta}{\rho} \right) \left( \frac{\theta}{\sigma_{\theta}} \right) N_a \]
Constraints on antenna design parameters: (Uncorrected) beam shape errors

Same beam correction methodology:

$$DR \sim \frac{1}{\sigma_{s(\bar{\rho}_k)}} \sim \frac{1}{2} \left( \frac{\theta}{\rho} \right)^2 \left( \frac{\theta}{\sigma_\theta} \right) N^a$$
Constraints on antenna design parameters: (Uncorrected) amplitude gain errors

\[
\chi^2 = \left\| \sum_{mn} \left[ V_{mn} - \prod_k \left[ G_m^k \otimes G_m^{k*} \right] \int_{\Omega} \prod_k \left[ T_m^k(\bar{\rho}) \otimes T_m^{k*}(\bar{\rho}) \right] e^{-2\pi \bar{\rho}_\text{m} \cdot (*\bar{\rho} - \bar{\rho}_s)} K S(\bar{\rho}) d\Omega \right\|
\]

\[
\|G_m^k\| = (1 + \varepsilon)
\]

As before, \( \frac{\partial \chi^2}{\partial S(\bar{\rho})} \bigg|_{S(\bar{\rho}) = S_0(\bar{\rho})} = -2 \text{Re} \sum_{mn} X_{ij}^{*T}(\bar{\rho}) W_{mn} V_{mn}^{\text{obs}} \)

\[
DR \sim \frac{1}{\sigma_{S(\bar{\rho}_s)}} \sim \frac{N_a}{2\sigma_\varepsilon}
\]

\( DR \) improves in this case also as \( \sqrt{M} \), where \( M \) is the number of independent calibration sets (Perley 1988). See also Condon (2009).
Order and accuracy of calibration parameterization

- Direction-dependent, image-plane effects require **much higher-order** parametrization than visibility-plane calibration terms
  - Strong implications for solver stability and performance
- Precision of *a priori* model is coupled to achieved DR
  - Accuracy and time stability of these models decay with radial distance from optical axis and frequency

\[
\text{No. free parameters} = N_a \cdot N_{\text{rms}}
\]
\[
\text{No. of data constraints} = N_a^2
\]

(VLA beam response: Bhatnagar et al 2008)
What about deconvolution?

- Non-linear; requires much more sophisticated approaches:
  - Not a trivial problem, but general heuristic is that calibration errors are more pronounced at source positions

- Research strategies:
  - **Real data** from SKA pathfinders.
  - **Simulations** for ranking of dynamic range issues
  - **Statistical methods:**
    - **Analytic** methods possible for some cases (e.g. snapshot imaging); see Wijnholds et al. (2008)
    - **Computational methods** possible for full non-linear problem using statistical resampling (Kemball and collaborators, 2005, 2008), or full Bayesian methods (Sutton & Wandelt 2005ff).

(Wijnholds et al. 2008)
Computational complexity and feasibility

- Computational cost driven primarily by data rate:

\[
\left( \frac{\dot{V}_{\text{vis}}}{TBps} \right) \sim 10^{-14} \left( \frac{NB}{D} \right)^2 \left( \frac{\Delta \nu}{\nu} \right)
\]

- where \( D \) = dish diameter, \( B \) = max. baseline, \( \Delta \nu \) = bandwidth, and \( \nu \) = frequency

- Full-field continuum imaging cost (derived from Cornwell 2004); calibration likely comparable:

\[
\frac{C}{PF} \approx 22.3 N^2 \left( \frac{B}{D} \right)^2 10^{-0.00273 \frac{B}{D^2} - 15}
\]

- Strong dependence on \( 1/D \) and \( B \).
- Data rates of Tbps and computational costs in PF are readily obtained from underlying geometric terms.
Scalability

- Moore’s Law holds, but high-performance architectures are evolving *laterally* rapidly.

- Sustained petascale calibration performance for SKA requires:
  - Demonstrated mapping of SKA calibration and imaging algorithms to modern HPC architectures, and proof of feasible scalability to petascale: \(O(10^5)\) processor cores.
  - Remains a considerable design unknown in both feasibility and cost.

- Not a fatal problem – just requires work, informed by other disciplines.
# Calibration and processing feasibility risks

<table>
<thead>
<tr>
<th>Issue</th>
<th>Solutions</th>
<th>Risks or unknowns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application of <em>known</em> direction-dependent gains</td>
<td>Full imaging equation formalism or holographic imaging.</td>
<td>Accuracy of instrumental models (e.g. primary beam response) especially at large radial distance from optical axis.</td>
</tr>
<tr>
<td>Pixelization / sky model representation</td>
<td>Centering or multi-scale basis functions</td>
<td>Optimal decompositions</td>
</tr>
<tr>
<td>Wide-band synthesis and deconvolution</td>
<td>Multi-frequency, multi-scale synthesis</td>
<td>Frequency-, time-, and polarization-dependence of net bandpass response.</td>
</tr>
<tr>
<td>Solution for <em>unknown</em> direction-dependent effects</td>
<td>Peeling or generalized direction-dependent solvers.</td>
<td>Numerical stability; convergence, and order of parametrization relative to information measure of data.</td>
</tr>
<tr>
<td>Computational cost &amp; scalability</td>
<td>Weakly-coupled decompositions (e.g. by frequency)</td>
<td>Any stronger coupling in calibration? Scalability to extreme-scales (P/EF)</td>
</tr>
<tr>
<td>Instrumental /propagation model accuracy (e.g. ionosphere)</td>
<td>Incremental refinement of auxiliary data.</td>
<td>Are models good enough?</td>
</tr>
<tr>
<td>Real-time calibration &amp; imaging (little visibility archiving)</td>
<td>Stable, routine calibration; or lossless image representations (e.g. OMM)</td>
<td>Degree of reversibility actually needed.</td>
</tr>
<tr>
<td>RFI mitigation</td>
<td>Demonstrated heuristics from existing pathfinders and telescopes</td>
<td>Non-Gaussianity of residuals in solvers.</td>
</tr>
<tr>
<td>Non-closing errors</td>
<td>Identification &amp; mitigation; solvers</td>
<td>Instrumental origin and magnitude</td>
</tr>
</tbody>
</table>

(RED: contingent on or coupled to net antenna/feed/receiver design.)
Conclusions (and questions)

- Calibration and processing analysis produces stringent requirements on antenna image-plane instrumental performance (e.g. pointing, beam shape, etc):
  - Serious for spectral-line; acute for continuum.
  - Information on antenna costs/restrictions along this design parameter axis would be very valuable.
  - We need accurate image-plane models for antenna instrumental response over angular, time, frequency, and polarization coordinates, over the full beam solid angle – how well can we predict and model these for current antenna/feed designs?

- A huge difference in calibration and processing between applying an a priori image-plane corrections, and solving for residual image-plane corrections.
  - E.g. former has just happened for VLA; latter a work in progress almost everywhere.

- Some problems are serious cost drivers but lower feasibility risk
  - Computing costs & scalability (i.e. antenna diameter).

- Mitigation is possible:
  - Targeted fields for the highest continuum DR (trade-off FOV against beam-pattern accuracy).
  - HI line survey requirements are far less acute