Array signal processing for radio-astronomy imaging and future radio telescopes

Amir Leshem
School of Engineering
Bar-Ilan University

Faculty of EEMCS
Delft University of Technology

Joint work with Alle-Jan van der Veen, Chen Ben-David, Ronny Levanda

Some data and simulations were provided by
A-J. Boonstra, S. van der Toll, S. Wijnholds, Huib Intema and T. Clarke
Presentation overview

- Introduction
- The measurement equation
- RFI mitigation using multi-channel techniques
- Some ideas related to parametric imaging
  - LS-MVI
  - Joint imaging and calibration
- Simulated imaging results
- An example with the LOFAR test station
- Dirty images of Abell 2256
References

- C. Ben David and A. Leshem. “Parametric high resolution techniques for radio astronomical image formation”. JSTSP.


Interferometers

Fig. 3. Photograph of the east-west arm of the $\lambda = 1.7$ m instrument built in 1957 with which nearly 5000 sources were located.

From Martin Ryle - Nobel lecture, 1974
Fig. 10. The powerful radio galaxy in the constellation of Cygnus mapped with the 5 km telescope. The compact outer components are exceedingly bright - (31 and 41 contours). The central component - which corresponds to the nucleus of the optical galaxy is very weak and is drawn with contours spaced at 1/5 the interval.
Radio interferometers

Westerbork, Netherlands
1970, 14 dishes, 3 km

Very Large Array (VLA), New Mexico
1980, 27 dishes, 36 km
Cygnus A – Radio galaxy

Location: 600 million light years away
Double radio lobes, spanning over 500,000 light years, which are fed by jets of energetic particles beamed from the compact radio core

VLA (22cm) observation by John Conway and Philip Blanco
Next generation telescopes

We need a similar improvement in algorithms!
The importance of interference mitigation

Play SETI movie
The parametric approach to radio astronomical imaging

- We proposed a parametric approach to solve the image deconvolution.
- The new formalism enables to combine spatio-temporal interference mitigation techniques with imaging.
- Better interference blanking.
- Higher dynamic range.
- High resolution.
- Combination with calibration in a computationally feasible way.
Extended sources example

(a) CLEAN  
(b) LS_MVI  
(c) Original image
Signals are stacked into vectors $\mathbf{x}(t) = [x_1(t), ..., x_p(t)]^T$

Location of antenna $i$ at time $t$ is $\mathbf{r}_i(t)$

Baseline between antennas $i, j$ is $\mathbf{r}_i(t) - \mathbf{r}_j(t)$
Imaging - the correlation process

\[ R_{k,\omega} = E \left[ x_{\omega} (t_k) x_{\omega} (t_k)^* \right] = \frac{1}{N} \sum_{n=0}^{10s/T_s} x_{\omega} (t_k + nT_s) x_{\omega} (t_k + nT_s)^* \]

\( \left( R_{k,\omega} \right)_{i,j} \) corresponds to the visibility \( V \left[ r_i (t_k), r_j (t_k) \right] \) at frequency \( \frac{\omega}{2\pi} \).
Correlations are measured only for baselines $\mathbf{r}_i(t) - \mathbf{r}_j(t)$.
Imaging equation: Planar array

\[ V_f(u, v) = \iint A^2(l, m)I_f(l, m)e^{-2\pi j(ul + vm)} \, dl \, dm \]

- We limit ourselves to single frequency.
- \( A(l, m) \) is the "known" amplitude response of the antennas.
- \( I_f(l, m) \) is the intensity at location \((l, m)\).
- \( V_f(u, v) \) is the visibility function.
Imaging equation: Non-coplanar array

\[ V_f(u, v, w) = \int \int \frac{1}{\sqrt{1-l^2-m^2}} A^2(l, m, n) I_f(l, m) e^{-2\pi j (ul+vm+wn)} \, dldm \]

- We limit ourselves to single frequency.
- \( A(l, m) \) is the "known" amplitude response of the antennas.
- \( I_f(l, m) \) is the intensity at location \((l, m)\).
- \( V_f(u, v, w) \) is the visibility function.
The measurement equation – discrete form

\[ V_{ijk} = V \left[ r_i(t_k), r_j(t_k) \right] = \sum_{l=1}^{d} I(s_l) e^{-j s_{i}^{T} (r_i-r_j)} \]

- \( I(\cdot) \) is the brightness image (‘map’) of interest
- \( s_{\ell} \) is the unit direction vector of the \( \ell \)-th source (assuming discrete source model)
- \( V_{ijk} \) is the measured correlation \((R_k)_{ij}\) between antennas \(i\) and \(j\) at time \(t_k\)

Classical Fourier-based imaging

Given many samples of \(V_{ijk}\), we can compute \(I(s_{\ell})\) via an inverse Fourier transform:

- “dirty image”:
  \[ I_D(s) := \sum_{i,j,k} V_{ijk} e^{j s^{T} (r_i-r_j)} =: \sum_{\ell} I(s_{\ell}) B(s-s_{\ell}) = I \ast B \]
- “dirty beam”:
  \[ B(s-s_{\ell}) := \sum_{i,j,k} e^{j (s-s_{i})^{T} (r_i-r_j)} \]

Every point source excites a beam \(B(\cdot)\) centered at its location \(s_{\ell}\)
Matrix formulation of the imaging equation

Recall

\[(R_k)_{i,j} \equiv V_{ijk} = \sum_{\ell=1}^{d} I(s_{\ell}) e^{-js_{\ell}^T(r_i-r_j)} = \sum_{\ell=1}^{d} e^{-js_{\ell}^T r_i} I(s_{\ell}) e^{-js_{\ell}^T r_j}\]

• In "our" notation this translates to

\[R_k = A_k B A_k^H\]

where

\[A_k = [a_k(s_1), \ldots, a_k(s_d)], B = \begin{bmatrix} I(s_1) & 0 \\ \vdots & \ddots \\ 0 & I(s_d) \end{bmatrix}\]

• \[a_k(s) = \begin{bmatrix} e^{-js_1^T r_1(t_k)} \\ \vdots \\ e^{-js_p^T r_p(t_k)} \end{bmatrix}\]
The non co-planar array case

\[ R_k = A_k B A_k^H, \]

\[ R_k \equiv R(t_k), \quad A_k = [a_k(l_1, m_1), \ldots, a_k(l_d, m_d)], \]

\[ a_k(l, m) = \begin{bmatrix} e^{-2\pi i (u_{10}(t_k)l + v_{10}(t_k)m + w_{10}(t_k)n)} \\ \vdots \\ e^{-2\pi i (u_{p0}(t_k)l + v_{p0}(t_k)m + w_{p0}(t_k)n)} \end{bmatrix} \]

\[ B = \text{diag} \left[ \frac{I(l_1, m_1)}{\sqrt{1 - l_1^2 - m_1^2}}, \ldots, \frac{I(l_d, m_d)}{\sqrt{1 - l_d^2 - m_d^2}} \right]. \]
Matrix formulation of the imaging equation

\[ R_k = \Gamma(t_k)A_kBA_k^H\Gamma(t_k)^H + R_{nn}(t_k) \]

- \( \Gamma(t_k) \) is a self-calibration slowly time varying pxp matrix
- \( R_{nn}(t_k) \) is the noise covariance matrix. Typically AWGN but when interference exists it dominates
Interference blanking
Typically interference is much stronger than astronomical signals

$$R = P_i a_i a_i^H + R_{\text{astronomical}} + \sigma^2 I$$

$$\Rightarrow$$ Largest eigenvalue is close to $P_i \gg \sigma^2$
More than a single interferer

\[ R = P_{i_1} a_{i_1} a_{i_1}^H + P_{i_2} a_{i_2} a_{i_2}^H + R_{\text{astronomical}} + \sigma^2 I \]

\[ \Rightarrow \text{Two largest eigenvalues are } >> \sigma^2 \]
Eigen-structure based techniques
Multichannel blanking
Multichannel blanking can improve performance on the 3C48 absorption line with interference.
Deconvolution

Classical Fourier imaging

\[ I_D(s) = \sum_{k=1}^{K} a_k^H(s)R_k a_k(s) \]

Super-resolution MVDR based imaging

Pseudo-spectrum

\[ I_D'(s) = \sum_{k=1}^{K} w_k^H(s)R_k w_k(s) \]

MVDR criterion

\[ \hat{w}_k(s) = \arg \min_w I_D'(s) \text{ such that } w_k^H(s)a_k(s) = 1 \]

Solution

\[ \hat{w}_k(s) = \beta_k R_k^{-1} a_k(s) \text{ where } \beta_k = \frac{1}{a_k^H(s)R_k^{-1}a_k(s)} \]

\[ I_D'(s) = \sum_{k=1}^{K} \frac{1}{a_k^H(s)R_k^{-1}a_k(s)} \]
MVI in pictures – ULA point sources

- Two sources 85, 100 deg.
MVI estimating the second source
MVI- Non uniform sparse array
Simulated example – High dynamic range

- 1 dimensional images – for simplicity
- Dirty images only!
- Theoretical dynamic range within the image 1000,000:1
- Weakest source is only 10x noise RMSE!
- Sources are extended off the grid

- One source is moved through FOV
Array response
Sources integrated power

Source integrated power profile

Power \( \sigma_{\text{noise}} \)

\( \theta^\circ \)
Dirty images

Classical dirty image

AAR dirty image

AAR dirty image with 3% RMS errors on each antenna element
Classical CLEAN imaging

$$\max_s \sum_{k=1}^{K} w_k^H(s) R_k w_k(s)$$

$$R_k = \sum_{n=1}^{N} b_k a_k(s_n) a_k^H(s_n) + \sigma^2 I$$

$$I_D(s) = \sum_{k=1}^{K} a_k^H(s) R_k a_k(s)$$

$$\hat{s} = \arg \max_s I_D(s)$$

$$\hat{\alpha} = \max I_D(\hat{s})$$
The adapted angular response

We would like to maintain the Gaussian receiver noise isotropic. To that end we add to the MVDR equations the requirement

$$\|w\| = 1$$

Still we want a solution of the form $w_k(s) = \alpha R_k^{-1} a_k(s)$

We obtain that

$$w = \frac{R_k^{-1} a_k(s)}{a_k^H(s) R_k^{-2} a_k(s)}$$

and the angular spectrum becomes

$$I''_D(s) = \sum_{k=1}^{K} \frac{a_k^H(s) R_k^{-1} a_k(s)}{a_k^H(s) R_k^{-2} a_k(s)}$$
Enhancing the image using independence

Since the covariance matrices at different epochs represent independent observations it is reasonable to assume that the overall covariance of the data is given by

\[
\bar{\mathbf{R}} = \begin{bmatrix}
\mathbf{R}_1 \\
\vdots \\
\mathbf{R}_K
\end{bmatrix}
\]

Choosing a vector \( \mathbf{w} \) for \( \bar{\mathbf{R}} \) yields \( \mathbf{w} = [\mathbf{w}_1^T, \ldots, \mathbf{w}_K^T]^T \) where

\[
\mathbf{w}_k = \mathbf{R}_k^{-1} \mathbf{a}_k(s)
\]

In this case the MVDR dirty image becomes:

\[
I_{\text{MVDR}}(s) = \frac{1}{\sum_{k=1}^{K} \mathbf{a}_k^H(s) \mathbf{R}_k^{-1} \mathbf{a}_k(s)}
\]
The AAR dirty image

\[
I_{\text{AAR}}(s) = \frac{\sum_{k=1}^{K} a_k^H(s) R_k^{-1} a_k(s)}{\sum_{k=1}^{K} a_k^H(s) R_k^{-2} a_k(s)}
\]
Dirty images

(a) Classical
(b) MVDR
(c) AAR
Extension to bandwidth synthesis is different than in classical imaging due to non-linear structure

\[ I_{\text{AAR}}(s) = \frac{\sum_{\omega} \sum_{k=1}^{K} a_{k,\omega}^H(s) R_{k,\omega}^{-1} a_{k,\omega}(s)}{\sum_{\omega} \sum_{k=1}^{K} a_{k,\omega}^H(s) R_{k,\omega}^{-2} a_{k,\omega}(s)} \]
Least Squares power estimator to remove bias

\[ \lambda^{(n)} = \arg \min_{\alpha} \sum_{k=1}^{K} \left\| R_k^{(n)} - \alpha a(s^{(n)})a^H(s^{(n)}) \right\|_F^2 \]

Independently used in the context of multi-resolution CLEAN by Bhatnagar & Cornwell 2004
The LS-MVI algorithm

Measure $R_1, \ldots, R_K$

Set $R_k^{(0)} = R_k$ $\gamma = 0.1$

Until convergence criterion

Source location estimator

$$s^{(n)} = \arg \max_s I_{AAR}^{(n)}(s)$$

Power estimator

$$\lambda^{(n)} = \arg \min \sum_{k=1}^K \| R_k^{(n)} - \alpha a(s^{(n)}) a^H(s^{(n)}) \|_F^2$$

Update $R_k^{(n)}$

$$R_k^{(n+1)} = R_k^{(n)} - \gamma \lambda^{(n)} a(s^{(n)}) a^H(s^{(n)})$$
Enforcing the non-negative definite constraint

All source powers are non-negative and residual covariance is also non-negative

\[
R_k - \sigma^2 I - \alpha a(s^{(n)})a^H(s^{(n)}) \geq 0
\]
\[
\alpha \geq 0
\]

Solution is simple:

Solve the unconstrained problem and use bisection on \( \alpha \) finding sequentially over \( k \)

\( \alpha \) that is good for all \( k \)
Extended source
Cross section through the image

Crosssection through the image

- Initial image
- CLEAN image
- LS-MVI image

Power

RA [arcsec]
LOFAR test station - dirty images

*Left: Classical beamforming; right: MVDR*

Snapshot image showing Cassiopeia A along with interference at the horizon
Example LOFAR test station

- Snapshot image
- 25 frequency bands of 156 kHz each
- At each frequency a single covariance matrix
- Initial calibration by S. Wijnhold
Original image provided by S. Wijnholds

Wijnholds 2008
Comparison of dirty images
All sky image after AAR based cleaning
Dirty images of Abell 2256 1369 MHz

Clarke & Ensslin 2006 (right)
Bridle & Fomalont 76 (left)
Abell 2256 1369 MHz

Left: Cleaned image by Huib Intema. Right: AAR dirty image. Calibrated data: T. Clarke
Extensions

- Simplified Maximum likelihood (Leshem and van der Veen)
- Maximum likelihood using EM (Lanterman)
- Robust beamforming techniques, e.g., diagonal loading
- Performance analysis
- Sensitivity to modeling assumptions
- Resolution limits
Conclusions

- Exciting science can be done with future instruments
- Science strongly depends on technology

Great challenges:
- Calibration is extremely challenging!
- Image processing can take us beyond the equipment!!
- Real time adaptive interference mitigation is crucial
- Architectures and algorithms for extremely large arrays
  - Comm – 2-4 sensors
  - Radar - thousands of sensors
  - Radio astronomy – 10,000-100,000,000 elements