Wideband Super-resolution Imaging in Radio Interferometry (HyperSARA)

Abdullah Abdulaziz, Arwa Dabbech and Yves Wiaux

CIRM pre-school

January 7th, 2019

BASP Biomedical and Astronomical Signal Processing

HERIOT WATT UNIVERSITY

Outline

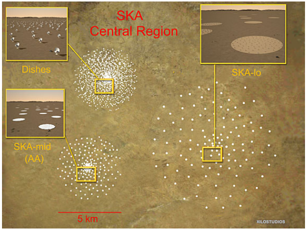
Outline

Wideband radio-interferometry (RI)

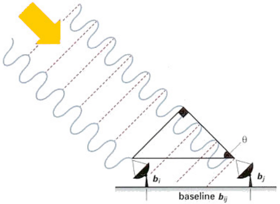
Radio-interferometer:

- A radio-interferometer is a collection of spatially separated antennas.
- Each pair of antennas identifies a baseline $\boldsymbol{b}_{ij} = (u, v, w) \in \mathbb{R}^3$.
- Each baseline gives access to a radio measurement, so-called visibility.

$$y_\nu(\boldsymbol{b}_{ij}) = \langle E_\nu(\boldsymbol{b}_i, t) E_\nu^*(\boldsymbol{b}_j, t) \rangle_{\Delta t}$$



The Square Kilometer Array
Figure courtesy: SkyWatch SA

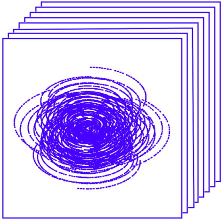


A baseline
Figure courtesy: http://hardhack.org.au

Wideband RI measurement model

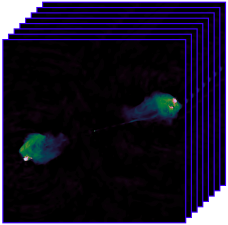
Considering L frequency channels, the wideband measurement model:

$$\mathbf{Y} = \mathbf{\Phi}(\mathbf{X}) + \mathbf{N}$$



Wideband data cube $\mathbf{Y}$

→



Wideband model cube $\mathbf{X}$

Outline

1. Wideband radio-interferometry
2. HyperSARA: optimisation problem
3. HyperSARA: algorithmic structure
4. Application to real data

HyperSARA

HyperSARA minimisation problem:

$$\min_{\mathbf{X}\in\mathbb{R}^{N\times L}} \|\mathbf{X}\|_{\omega,*} + \mu\|\mathbf{\Psi}^{\dagger}\mathbf{X}\|_{\bar{\omega},2,1} \quad \text{subject to} \quad \begin{cases} \|\boldsymbol{y}_l^b - \bar{\Phi}_l^b(\mathbf{X})\|_2 \leq \epsilon_l^b, & \forall (l,b) \\ \mathbf{X} \in \mathbb{R}_+^{N\times L} \end{cases}$$

HyperSARA

HyperSARA minimisation problem:

$$\min_{\mathbf{X}\in\mathbb{R}^{N\times L}} \|\mathbf{X}\|_{\omega,*} + \mu\|\boldsymbol{\Psi}^\dagger\mathbf{X}\|_{\bar{\omega},2,1} \quad \text{subject to} \quad \begin{cases} \|\boldsymbol{y}_l^b - \bar{\Phi}_l^b(\mathbf{X})\|_2 \leq \epsilon_l^b, & \forall(l,b) \\ \mathbf{X}\in\mathbb{R}_+^{N\times L} \end{cases}$$

Enforce:

- low rankness: $\|\mathbf{X}\|_{\omega,*} = \sum_{j=1}^{J} \omega_j \, \sigma_j(\mathbf{X}), \quad \omega_i \geq 0$

HyperSARA

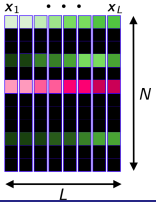
HyperSARA minimisation problem:

$$\min_{\mathbf{X}\in\mathbb{R}^{N\times L}} \|\mathbf{X}\|_{\omega,*} + \mu \|\mathbf{\Psi}^\dagger \mathbf{X}\|_{\bar{\omega},2,1} \quad \text{subject to} \quad \begin{cases} \|\boldsymbol{y}_l^b - \bar{\Phi}_l^b(\mathbf{X})\|_2 \leq \epsilon_l^b, & \forall (l,b) \\ \mathbf{X} \in \mathbb{R}_+^{N\times L} \end{cases}$$

Enforce:

- low rankness: $\|\mathbf{X}\|_{\omega,*} = \sum_{j=1}^{J} \omega_j \; \sigma_j(\mathbf{X}), \quad \omega_i \geq 0$
- joint sparsity: $\|\mathbf{X}\|_{\bar{\omega},2,1} = \sum_{n=1}^{N} \bar{\omega}_n \; \|\boldsymbol{x}_n\|_2, \quad \bar{\omega}_n \geq 0$

  $\mathbf{\Psi}$ is a concatenation of averaged 9 orthogonal bases; the Dirac basis and the 8 first Daubechies wavelet bases.



HyperSARA

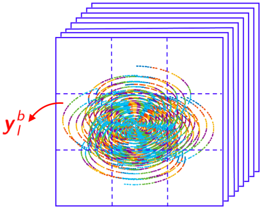
HyperSARA minimisation problem:

$$\min_{\mathbf{X}\in\mathbb{R}^{N\times L}} \|\mathbf{X}\|_{\omega,*} + \mu \|\boldsymbol{\Psi}^\dagger \mathbf{X}\|_{\bar{\omega},2,1} \quad \text{subject to} \quad \begin{cases} \|\boldsymbol{y}_l^b - \bar{\Phi}_l^b(\mathbf{X})\|_2 \leq \epsilon_l^b, & \forall (l,b) \\ \mathbf{X} \in \mathbb{R}_+^{N\times L} \end{cases}$$

Enforce:

- low rankness: $\|\mathbf{X}\|_{\omega,*} = \sum_{j=1}^{J} \omega_j \, \sigma_j(\mathbf{X}), \quad \omega_i \geq 0$
- joint sparsity: $\|\mathbf{X}\|_{\bar{\omega},2,1} = \sum_{n=1}^{N} \bar{\omega}_n \, \|\boldsymbol{x}_n\|_2, \quad \bar{\omega}_n \geq 0$

  $\boldsymbol{\Psi}$ is a concatenation of averaged 9 orthogonal bases; the Dirac basis and the 8 first Daubechies wavelet bases.
- data fidelity: inter-channel & intra-channel.



HyperSARA

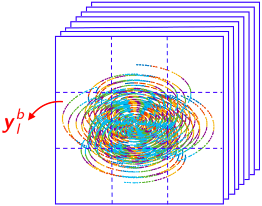
HyperSARA minimisation problem:

$$\min_{\mathbf{X}\in\mathbb{R}^{N\times L}} \|\mathbf{X}\|_{\omega,*} + \mu \|\boldsymbol{\Psi}^\dagger\mathbf{X}\|_{\bar{\omega},2,1} \quad \text{subject to} \quad \begin{cases} \|\boldsymbol{y}_l^b - \bar{\Phi}_l^b(\mathbf{X})\|_2 \leq \epsilon_l^b, & \forall (l,b) \\ \mathbf{X} \in \mathbb{R}_+^{N\times L} \end{cases}$$

Enforce:

- low rankness: $\|\mathbf{X}\|_{\omega,*} = \sum_{j=1}^{J} \omega_j \, \sigma_j(\mathbf{X}), \quad \omega_i \geq 0$
- joint sparsity: $\|\mathbf{X}\|_{\bar{\omega},2,1} = \sum_{n=1}^{N} \bar{\omega}_n \, \|\boldsymbol{x}_n\|_2, \quad \bar{\omega}_n \geq 0$

  $\boldsymbol{\Psi}$ is a concatenation of averaged 9 orthogonal bases; the Dirac basis and the 8 first Daubechies wavelet bases.
- data fidelity: inter-channel & intra-channel.
- reality & positivity.



Outline

Primal-dual algorithm

Advantages:

- Full splitting of the terms in the minimisation problem.
- Parallel updates.
- Randomised updates.
- Multiple functions unlike ADMM.
- No inversion of the linear operators unlike SDMM.

Primal-dual algorithm

Advantages:

- Broad range of convex prior functions.
  - Forward-backward iterations for non smooth functions.

$$\mathbf{Z}_{n+1} = \underbrace{\text{prox}_{\gamma_n f_1}}_{\text{backward step}} \underbrace{(\mathbf{Z}_n - \gamma_n \nabla f_2(\mathbf{Z}_n))}_{\text{forward step}}$$

where f_1 is a non-smooth function and f_2 is a smooth function.

The proximity operator of a function f, relative to a metric induced by a strongly positive, self-adjoint linear operator U, is defined as:

$$\text{prox}_f^U(\bar{\mathbf{Z}}) \triangleq \underset{\mathbf{Z} \in \mathbb{R}^{N \times L}}{\text{argmin}} \; f(\mathbf{Z}) + \frac{1}{2}(\mathbf{Z} - \bar{\mathbf{Z}})^\dagger U(\mathbf{Z} - \bar{\mathbf{Z}})$$

HyperSARA approach

Given $\mathbf{X}^{(0)}, \tilde{\mathbf{X}}^{(0)}, \mathbf{P}^{(0)}, \mathbf{A}_d{}^{(0)}, \boldsymbol{v}_l^{b(0)}, \bar{\boldsymbol{\omega}}^{(0)}, \boldsymbol{\omega}^{(0)}, \epsilon_l^{b(0)}, \vartheta_l^{b(0)}, \beta^{(0)}$

For $k = 1, \ldots$

Solve HyperSARA minimisation problem using Adaptive PPD

$$\left[\mathbf{X}^{(k)}, \tilde{\mathbf{X}}^{(k)}, \mathbf{P}^{(k)}, \mathbf{A}_d{}^{(k)}, \boldsymbol{v}_l^{b(k)}, \epsilon_l^{b(k)}, \vartheta_l^{b(k)}, \beta^{(k)}\right] = AdaptivePPD(\cdots)$$

Update the weights simultaneously

Update the re-weighted $\ell_{2,1}$ norm weights

$$\bar{\omega}_n^{(k)} = \frac{\bar{\gamma}^{(k)}}{\bar{\gamma}^{(k)} + \left\| \left(\boldsymbol{\Psi}^\dagger \mathbf{X}^{(k-1)}\right)_n \right\|_2}, \quad \forall n \in \{1, \ldots, N\}$$

Update the nuclear norm weights

$$\omega_j{}^{(k)} = \frac{\gamma^{(k)}}{\gamma^{(k)} + \sigma_j{}^{(k-1)}}, \quad \forall j \in \{1, \ldots, J\}$$

Adaptive preconditioned primal-dual algorithm

For $t = 1, \ldots$

Update dual variables simultaneously

Enforce data fidelity

$\forall (l, b) \in \{1, \cdots, L\} \times \{1, \cdots, B\}$ *(parallel)*

$$\tilde{\boldsymbol{v}}_l^{b(t)} = \boldsymbol{v}_l^{b(t-1)} + \mathbf{U}_l^b \bar{\mathbf{\Phi}}_l^b (\tilde{\mathbf{X}}^{(t-1)})$$

$$\boldsymbol{v}_l^{b(t)} = \mathbf{U}_l^{b\,1/2} \left(\boldsymbol{\mathcal{I}} - \boldsymbol{\mathcal{P}}_{\mathcal{E}\left(\boldsymbol{y}_l^b, \epsilon_l^{b(t-1)}\right)} \right) \left(\mathbf{U}_l^{b\,-1/2} \; \tilde{\boldsymbol{v}}_l^{b(t)} \right)$$

Adjust the ℓ_2 bounds

$$\rho_l^{b(t)} = \left\| \boldsymbol{y}_l^b - \bar{\mathbf{\Phi}}_l^b (\tilde{\mathbf{X}}^{(t-1)}) \right\|_2$$

if $\left(\beta^{(t-1)} < \lambda_1\right) \& \left(t - \vartheta_l^{b(t-1)} > \bar{\vartheta}\right) \& \left(\frac{\left|\rho_l^{b(t)} - \epsilon_l^{b(t-1)}\right|}{\epsilon_l^{b(t-1)}} > \lambda_2 \right)$

$$\epsilon_l^{b(t)} = \lambda_3 \; \rho_l^{b(t)} + (1 - \lambda_3) \; \epsilon_l^{b(t-1)}$$

$$\vartheta_l^{b(t)} = t$$

Adaptive preconditioned primal-dual algorithm

For $t = 1, \dots$

Update dual variables simultaneously

Promote low rankness

$$\mathbf{P}^{(t)} = \left(\mathcal{I} - \mathcal{S}^{*}_{\boldsymbol{\omega}^{(k-1)}/\kappa_1}\right)\left(\mathbf{P}^{(t-1)} + \tilde{\mathbf{X}}^{(t-1)}\right)$$

Promote joint sparsity

$\forall d \in \{1, \cdots, D\}$ *(parallel)*

$$\mathbf{A}_d^{(t)} = \left(\mathcal{I} - \mathcal{S}^{\ell_{2,1}}_{\bar{\boldsymbol{\omega}}^{(k-1)}\mu/\kappa_2}\right)\left(\mathbf{A}_d^{(t-1)} + \boldsymbol{\Psi}_d^{\dagger}\tilde{\mathbf{X}}^{(t-1)}\right)$$

Update primal variable

$$\mathbf{G}^{(t)} = \kappa_1 \mathbf{P}^{(t)} + \kappa_2 \sum_{d=1}^{D} \boldsymbol{\Psi}_d \mathbf{A}_d{}^{(t)} + \kappa_3 \sum_{l=1}^{L}\sum_{b=1}^{B} \bar{\boldsymbol{\Phi}}_l^{b\dagger}\left(\boldsymbol{v}_l^{b(t)}\right.$$

$$\mathbf{X}^{(t)} = \mathcal{P}_{\mathbb{R}_+^{N\times L}}\left(\mathbf{X}^{(t-1)} - \tau \mathbf{G}^{(t)}\right)$$

$$\tilde{\mathbf{X}}^{(t)} = 2\mathbf{X}^{(t)} - \mathbf{X}^{(t-1)}$$

$$\beta^{(t)} = \frac{\|\mathbf{X}^{(t)} - \mathbf{X}^{(t-1)}\|_2}{\|\mathbf{X}^{(t)}\|_2}$$

Outline

Cyg A

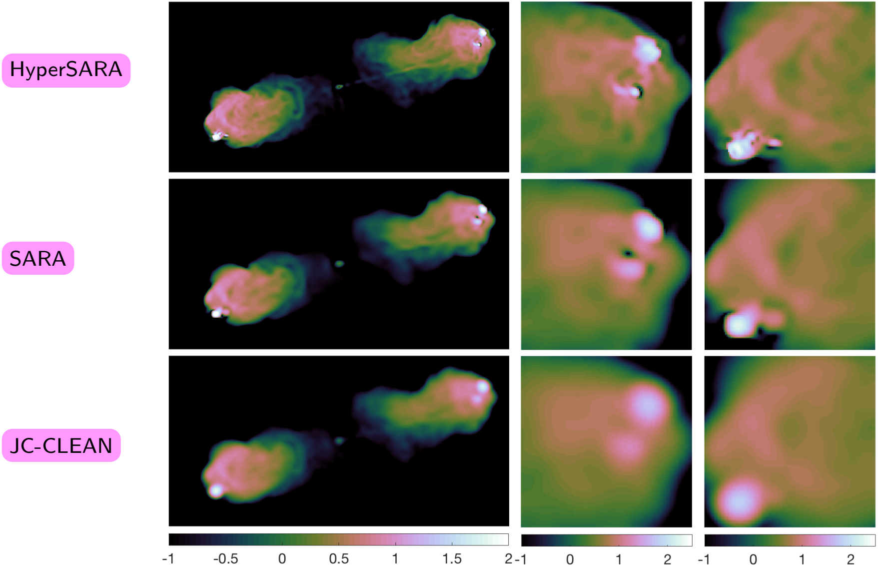
Data under scrutiny:

- Real VLA observations of the radio galaxy Cygnus A at the S band (2 - 4 GHz) and the C band (4 - 8 GHz).
- 32 channels of size 25×10^4 visibilities.
- Frequency range $2.04 - 5.96$ GHz with a frequency step 128 MHz and total bandwidth of 4 GHz.
- We consider images of size 1024×512 with a pixel size $\delta l = 0.19$ arcsec.

Assessment:

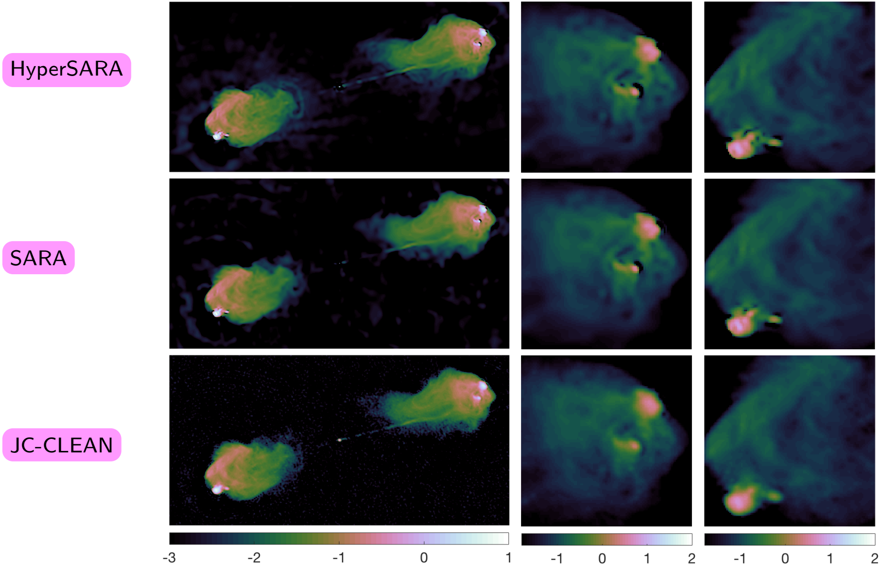
- Qualitative study through visual inspection of the reconstructed images.
- Spectral analysis of selected pixels from the different sources of the estimated wideband cubes.

Cyg A imaging results
Qualitative comparison - channel 1



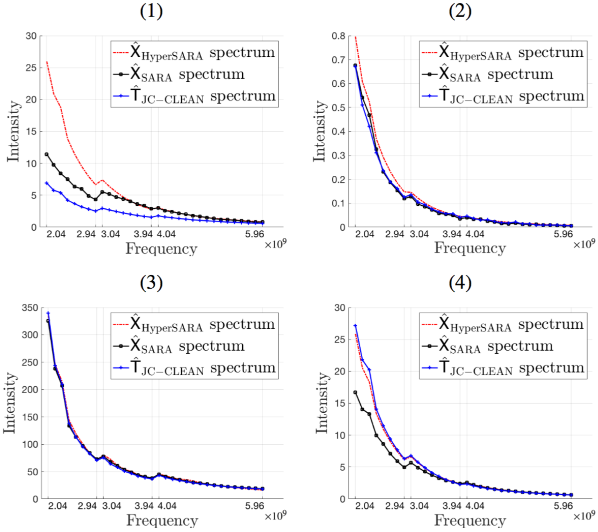
Cyg A imaging results

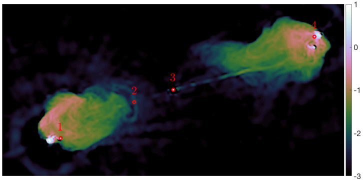
Qualitative comparison - channel 32



Cyg A Imaging results

Spectral analysis





References

- A. Abdulaziz, A. Dabbech and Y. Wiaux "Wideband Super-resolution Imaging in Radio Interferometry via Low Rankness and Joint Average Sparsity Models (HyperSARA)". *arXiv preprint*, 2018.
- A. Dabbech, A. Onose, A. Abdulaziz, R. A. Perley, O. M. Smirnov and Y. Wiaux "Cygnus A super-resolved via convex optimization from VLA data". *Monthly Notices of the Royal Astronomical Society*, 2018.
- A. Abdulaziz, A. Onose, A. Dabbech and Y. Wiaux "A distributed algorithm for wide-band radio-interferometry". *International BASP Frontiers workshop*, 2017.
- A. Abdulaziz, A. Dabbech, A. Onose and Y. Wiaux "A low-rank and joint-sparsity model for hyper-spectral radio-interferometric imaging". *EUSIPCO*, 2016.

Thank you!