Experimental Evaluation Using VLA Datasets of RFI Mitigation Performance over Long ngVLA Baselines

Michael Lambert(1), Urvashi Rao Venkata(2), Mitchell C. Burnett(1), and Brian D. Jeffs(1)

(1) Brigham Young University, Provo, UT USA, bjeffs@ee.byu.edu

(2) NRAO VLA Operations Center, Socorro, NM, USArurvashi@nrao.edu

USNC-URSI

Boulder, CO

Jan. 9-12, 2019

NRAO

BYU Electrical & Computer Engineering

IRA A. FULTON COLLEGE OF ENGINEERING

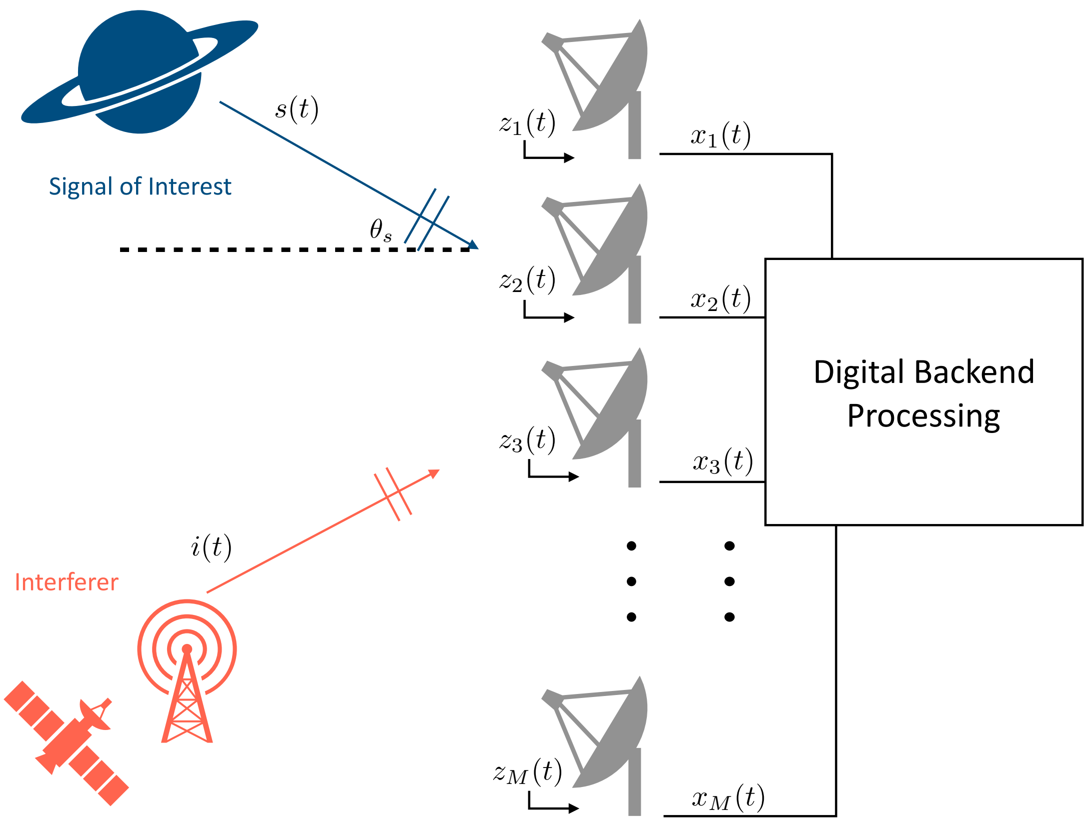
Background

- Present work motivated by 2017 BYU ngVLA community Study
- Sub-Array Processing (SAP): RFI mitigation algorithm suitable for ngVLA

NRAO

BYU Electrical & Computer Engineering
IRA A. FULTON COLLEGE OF ENGINEERING

Narrowband Sensor Array Signal Model

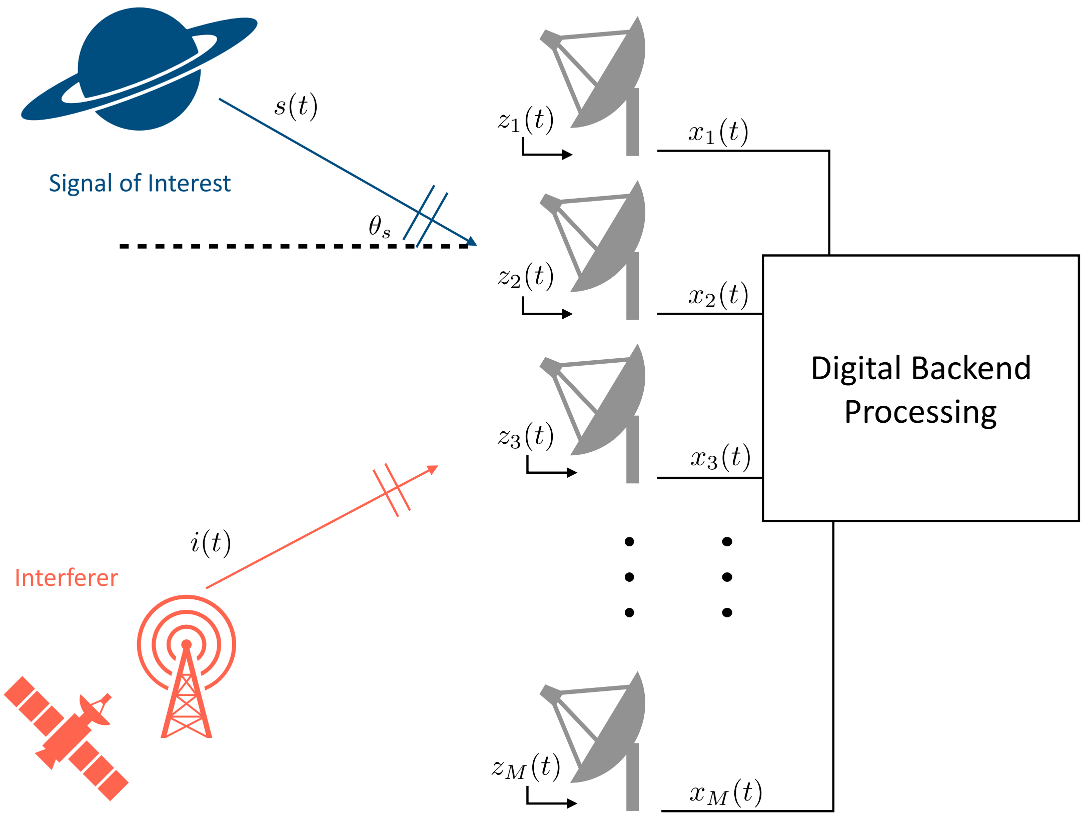


$$\mathbf{x}(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_M(t) \end{bmatrix} = s(t)\mathbf{a}(\theta_s) + i(t)\mathbf{a}(\theta_i) + \mathbf{z}(t)$$

$$\left.\begin{matrix} s(t) \\ i(t) \\ \mathbf{z}(t) \end{matrix}\right\} \text{Gaussian Proper Complex Random Processes}$$

$$\left.\begin{matrix} \mathbf{a}(\theta_s) \\ \mathbf{a}(\theta_i) \end{matrix}\right\} \text{Steering Vectors}$$

Correlator Array Covariance Matrix Visibilities



$$\begin{aligned}\mathbf{R}_x &= E\{\mathbf{x}(t)\mathbf{x}^H(t)\} \\ &= \mathbf{R}_s + \mathbf{R}_i + \mathbf{R}_z\end{aligned}$$

$$\hat{\mathbf{R}}_x = \frac{1}{N}\sum_{n=0}^{N-1}\mathbf{x}[n]\mathbf{x}^*[n]$$

- Spatial filtering techniques use the covariance matrix to estimate parameters of the RFI
- Computed at the central correlator
- Per narrowband frequency channel

Subspace Projection (SP)

- Zero forcing or null forming algorithm to cancel RFI
- Assumes interferer is the dominant signal component
- Eigenvector analysis of sample spatial covariance matrix

$$\hat{\mathbf{R}}_x = \frac{1}{N}\sum_{n=0}^{N-1} \mathbf{x}[n]\mathbf{x}^H[n]$$

$$= [\mathbf{U}_{\text{int}} | \mathbf{U}_{\text{sig+noise}}]\,\boldsymbol{\Lambda} \begin{bmatrix} \mathbf{U}_{\text{int}}^H \\ \mathbf{U}_{\text{sig+noise}}^H \end{bmatrix}$$

$$\hat{\mathbf{a}}(\theta_i) = \alpha \mathbf{U}_{\text{int}}$$

$$\mathbf{P} = \mathbf{I} - \mathbf{U}_{\text{int}}(\mathbf{U}_{\text{int}}^H \mathbf{U}_{\text{int}})^{-1}\mathbf{U}_{\text{int}}^H$$

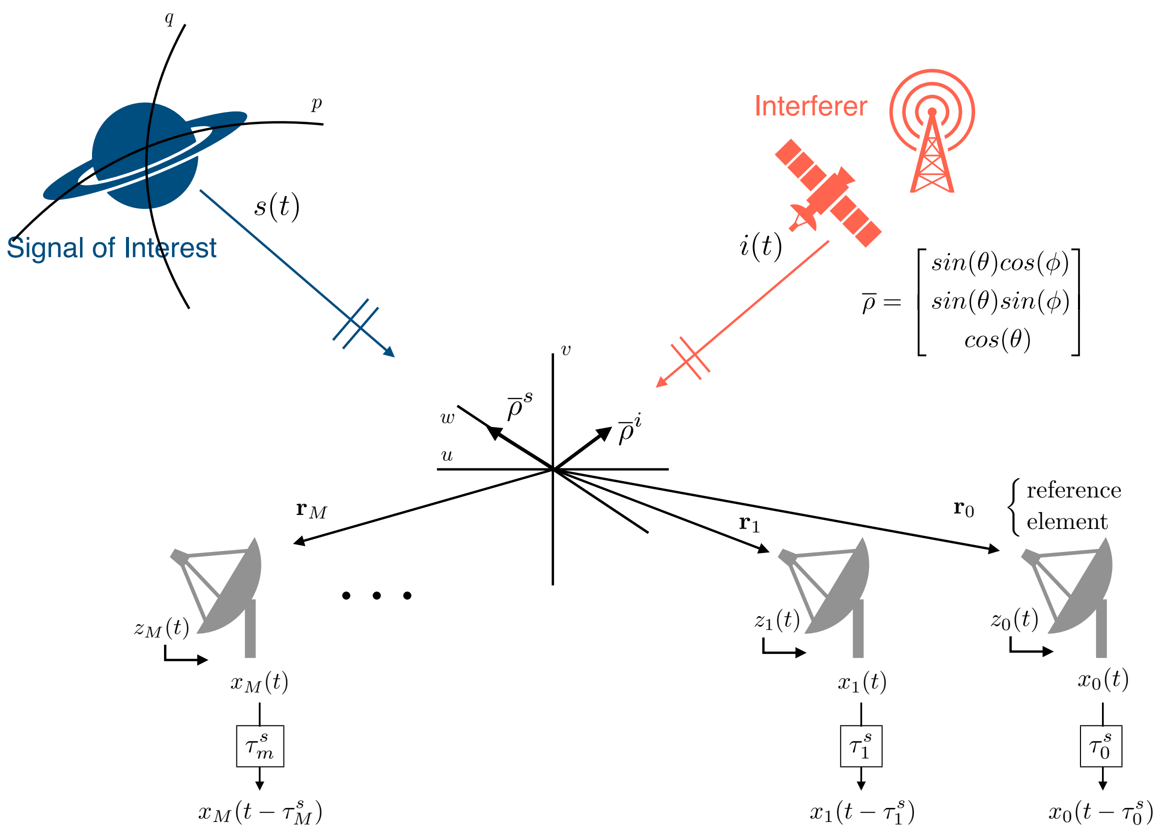
- Project RFI out of visibilities

$$\tilde{\mathbf{R}}_x = \mathbf{P}\hat{\mathbf{R}}_x\mathbf{P}^H$$

Challenges with SP on Large Interferometers

- Projection works well for small arrays (e.g. PAFs, aperture arrays)
- Spatial RFI Mitigation has been proposed for interferometer arrays (e.g. VLA) since 2000 [Leshem, van der Veen] again in 2004 [Jeffs, Li, Warnick], etc.
- **But** no significant science has been reported using SP on large arrays.
- ***Why***?
  - Rapid correlator dump-times are needed to track moving RFI (Rapid dump-times are now available at the VLA!)
  - Canceling null depth is limited due to poor RFI subspace estimates over long baselines (decorrelation of the RFI)
- We are proposing a solution!

Signal Model for Synthesis Imaging



Signal model and co-phased output

$$x_m(t) = s(t+\tau_m^s)a_m(\theta_s) + i(t+\tau_m^i)a_m(\theta_i) + z_m(t)$$

$$x_m(t-\tau_m^s) = s(t)a_m(\theta_s) + i(t+\tau_m^i-\tau_m^s)a_m(\theta_i) + z'_m(t)$$

Effective Time-delay for RFI relative to the SOI

$$\begin{aligned} \tau_m^i - \tau_m^s &= \frac{(\mathbf{r}_m - \mathbf{r}_0)^T\overline{\rho}^i}{c} - \frac{(\mathbf{r}_m - \mathbf{r}_0)^T\overline{\rho}^s}{c} \\ &= \frac{(\mathbf{r}_m - \mathbf{r}_0)^T(\overline{\rho}^i - \overline{\rho}^s)}{c} \end{aligned}$$

Resulting phase propagation

$$\psi_m = 2\pi(\tau_m^i - \tau_m^s)B$$

RFI Decorrelation

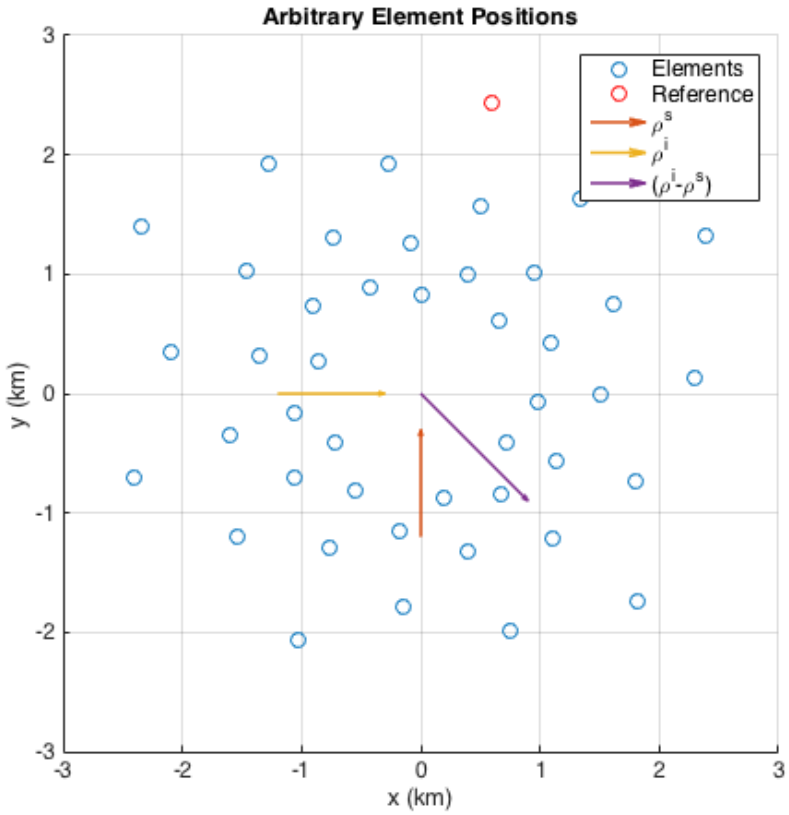
- The dot product between the baseline and propagation vectors yields an effective distance across the array relative to the reference element

$$\tau_m^i - \tau_m^s = \frac{(\mathbf{r}_m - \mathbf{r}_0)^T(\rho^i - \rho^s)}{c}$$

$$= \frac{d'_m}{c}$$

$$\psi_m = 2\pi(\tau_m^i - \tau_m^s)B$$

$$= 2\pi\frac{d'_m B}{c}$$



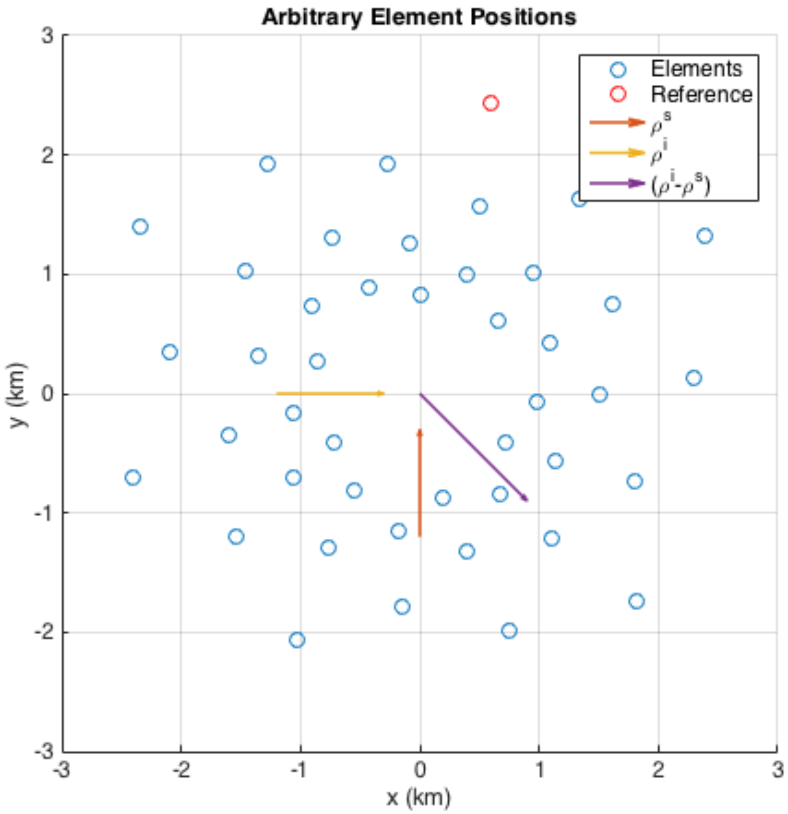
RFI Decorrelation

- The dot product between the baseline and propagation vectors yields an effective distance across the array relative to the reference element

$$\tau_m^i - \tau_m^s = \frac{(\mathbf{r}_m - \mathbf{r}_0)^T(\rho^i - \rho^s)}{c} = \frac{d'_m}{c}$$

$\boxed{\psi_m \text{ determines RFI decorrelation!}}$

$$\psi_m = 2\pi(\tau_m^i - \tau_m^s)B = 2\pi \frac{d'_m B}{c}$$



NRAO

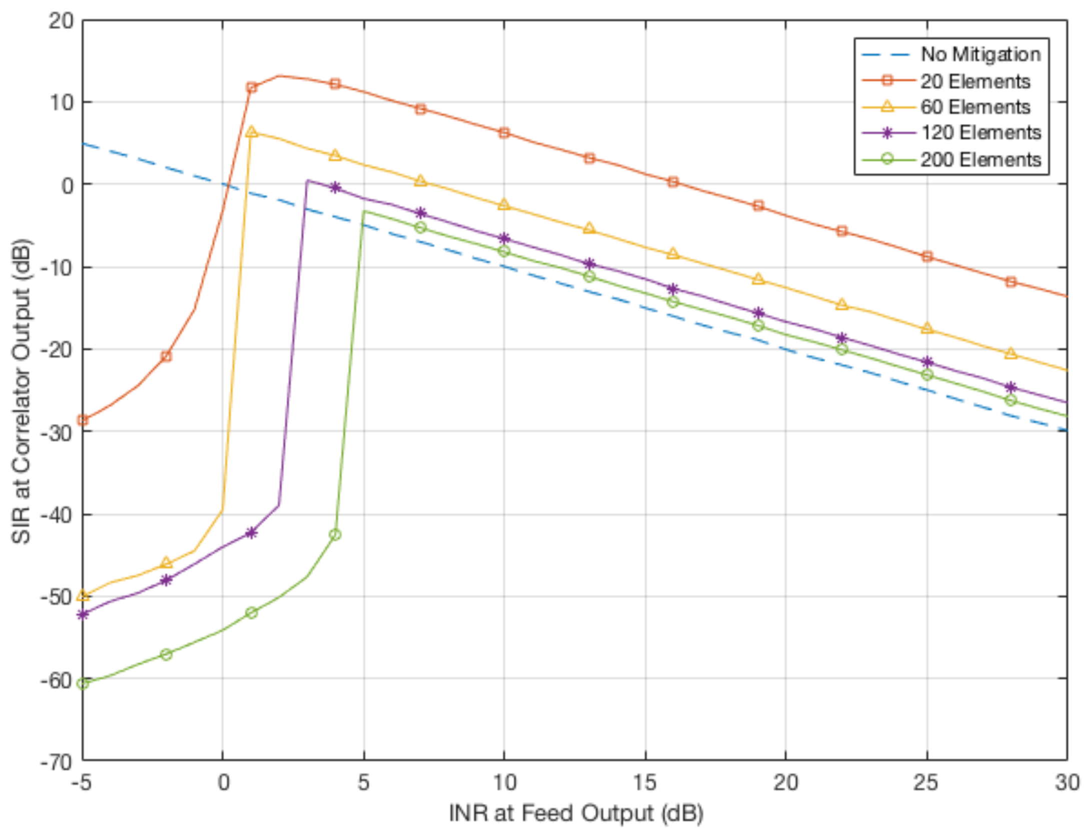
BYU Electrical & Computer Engineering

Subspace Projection Across Long Baselines

- Subspace projection becomes less effective as the baseline lengths increase
- Sample estimation error from baselines where the RFI has become decorrelated corrupts estimate of the RFI subspace
- RFI mitigation is ineffective over the full array.

<u>Processing Bandwidth:</u>
15 kHz
<u>Interferer Frequency:</u>
2.6 GHz

<u>Max Baseline Length:</u>
20 El. – 5.7 km
60 El. – 17.7 km
120 El. – 35.7 km
200 El. – 59.7 km



The Solution: Subarray Processing (SAP)

- Using the total phase propagation as a metric partition the full array into sections of K-subarrays

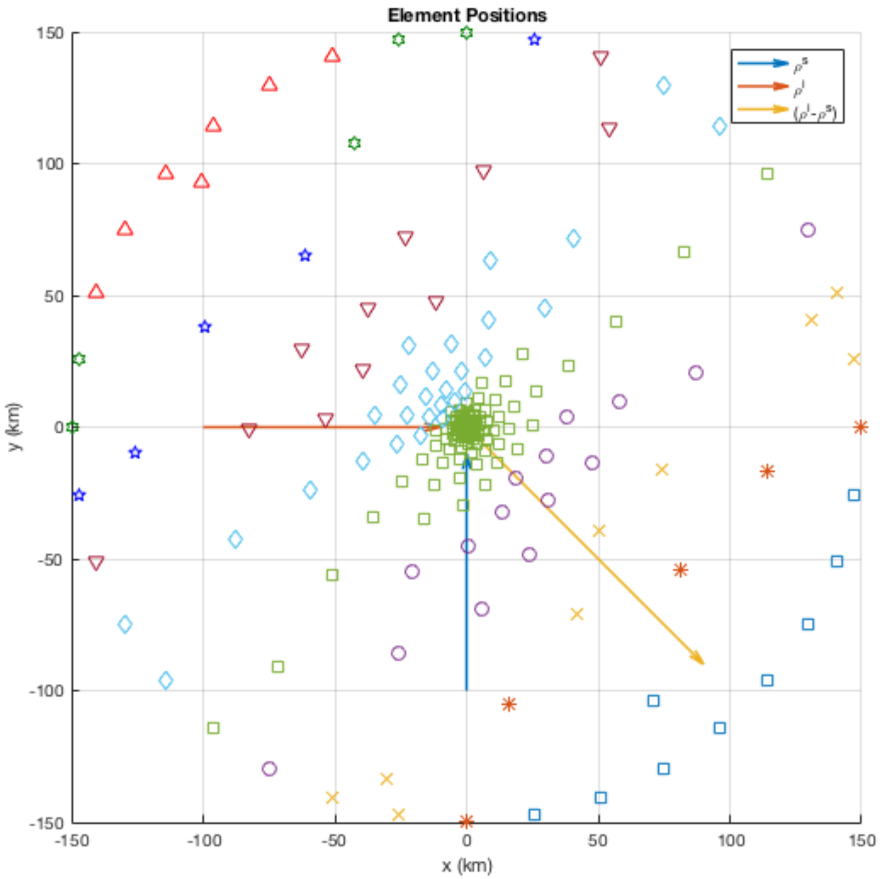
$$L_k = \{\psi_m \ : \gamma_{k-1} \leq \psi_m < \gamma_k\} \quad \forall \quad \psi_m$$

$$\psi_m = 2\pi(\tau_m^i - \tau_m^s)B$$

$$\gamma_k = \gamma_{k-1} + \psi_{\text{thresh}}$$

$$\gamma_0 = 0$$

- Grouping elements in this way creates arrays of elements perpendicular to the effective propagation vector
- These groups are such that there is high mutual RFI correlation to better estimate the interference subspac



Forming and Applying Projections

- Create a selection matrix, S, that orders the output time-series into their corresponding subarrays

$$\mathbf{x}'(t) = \mathbf{S}\mathbf{x}(t)$$

$$\mathbf{x}(t) = \mathbf{S}^T\mathbf{x}'(t)$$

- Form a block diagonal projection matrix where the diagonal elements are the projection matrix for each subarray

$$\mathbf{P} = \begin{bmatrix} P_1 & 0 & 0 & 0 \\ 0 & P_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & P_k \end{bmatrix}$$

- Now project RFI out of visibilities

$$\tilde{\mathbf{R}}_x = \mathbf{S}^T\mathbf{P}\mathbf{S}\hat{\mathbf{R}}_x\mathbf{S}^T\mathbf{P}^H\mathbf{S}$$
$$= \mathbf{P}'\hat{\mathbf{R}}_x\mathbf{P}'^H$$

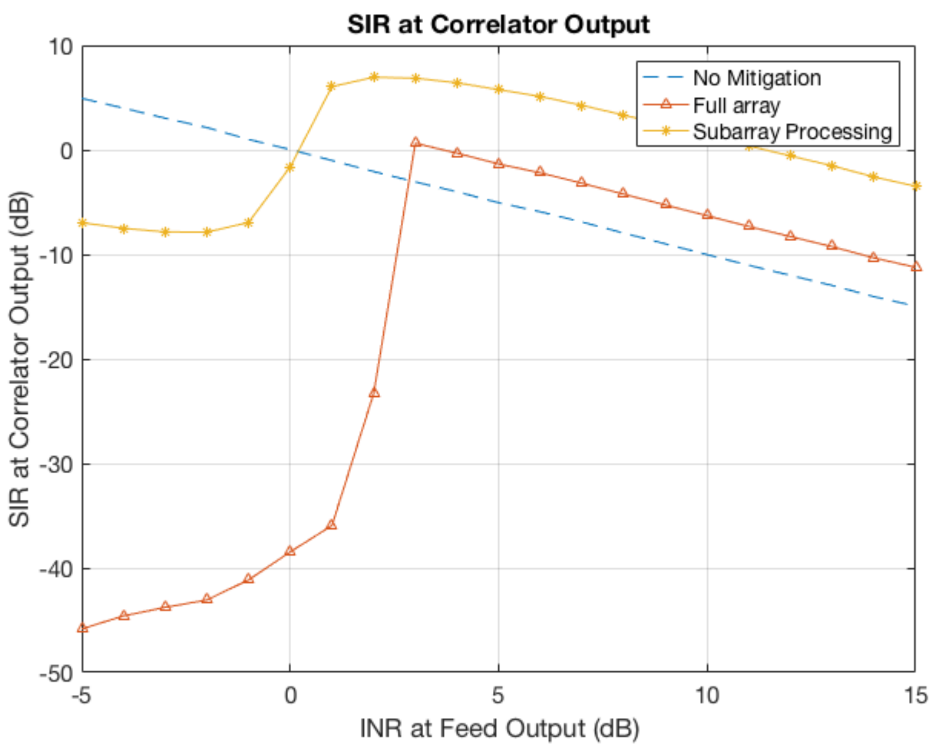
$$\mathbf{P}' = \mathbf{S}^T\mathbf{P}\mathbf{S}$$

Note:

$$\mathbf{P}\mathbf{P} = \mathbf{P}$$

P is still Idempotent. i.e. a projection matrix

Subarray Processing Results



- Simulation scenario
  - Full ngVLA configuration
  - 2.8 GHz Center frequency
  - 15 kHz Bandwidth
  - Wideband interferer
  - 25m dishes
- Subarray processing provides better RFI mitigation across the whole array as compared to subspace projection

RFI Canceling Performance Analysis Using Real VLA Data

- Evaluate SP and SAP RFI canceling performance with VLA data
- Attempt to emulate ngVLA-like long baseline RFI decorrelation

NRAO

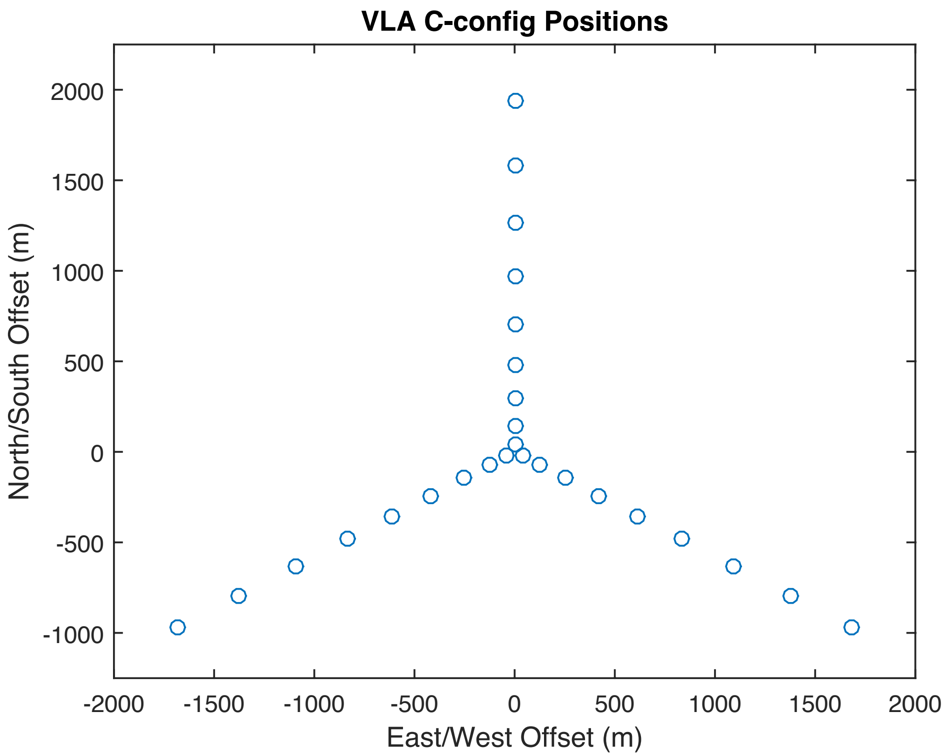
BYU Electrical & Computer Engineering
IRA A. FULTON COLLEGE OF ENGINEERING

Experiment Strategy

- Demonstrate SP RFI canceling with VLA small configuration C data
- Identify decorrelation effects in VLA large configuration A data
- If necessary to induce decorrelation, use relatively large channel bandwidth, B
- Attempt to demonstrate sub-array processing performance with A config. data.
- Two observation campaigns: summers of 2017 and 2018

Image courtesy of NRAO/AUI

Data set 1, Summer 2017, VLA Configuration C

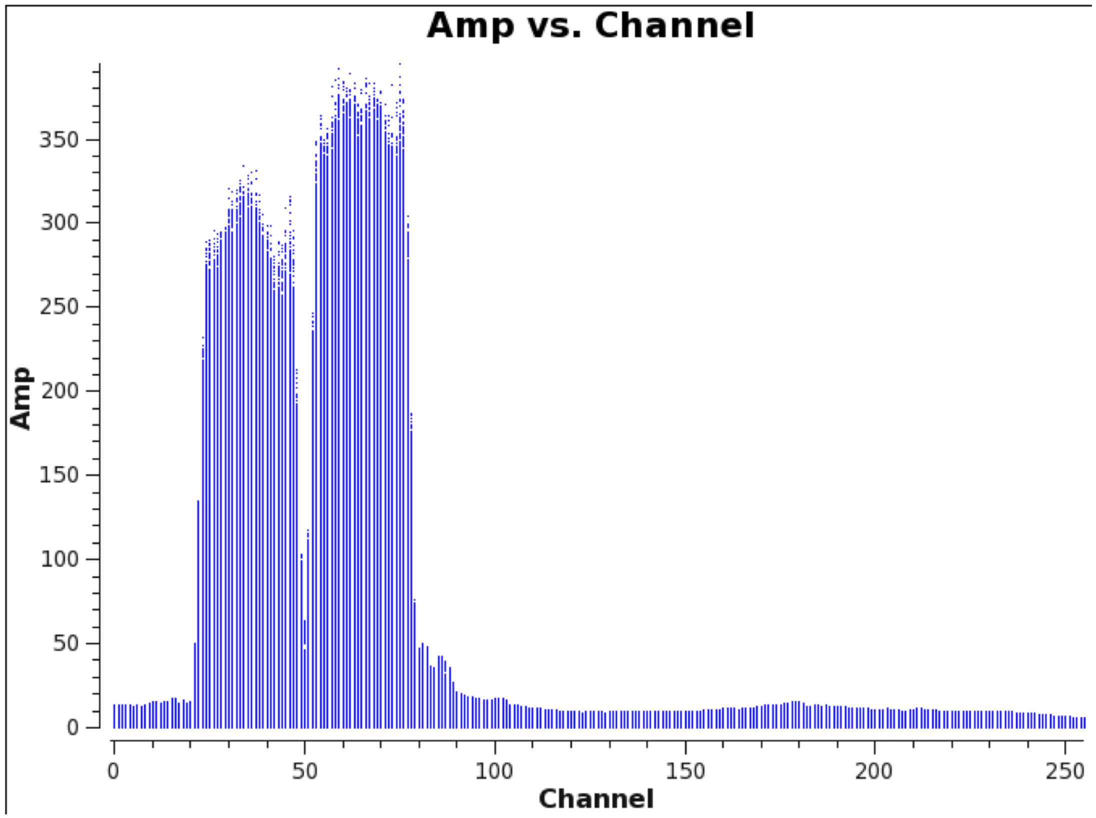


- 3.4 km longest baseline
- Source: 3C138: a point calibrator in southern sky
- RFI: Geostationary SiriusXM sat.
- Processing band center: 2.34 GHz
- Correlator channel bandwidth: 62.5 kHz, 256 chans., 16 MHz total
- Correlator dump interval: 20 ms
- 2 arcsec pixels, spans 6 integrations

BYU Electrical & Computer Engineering

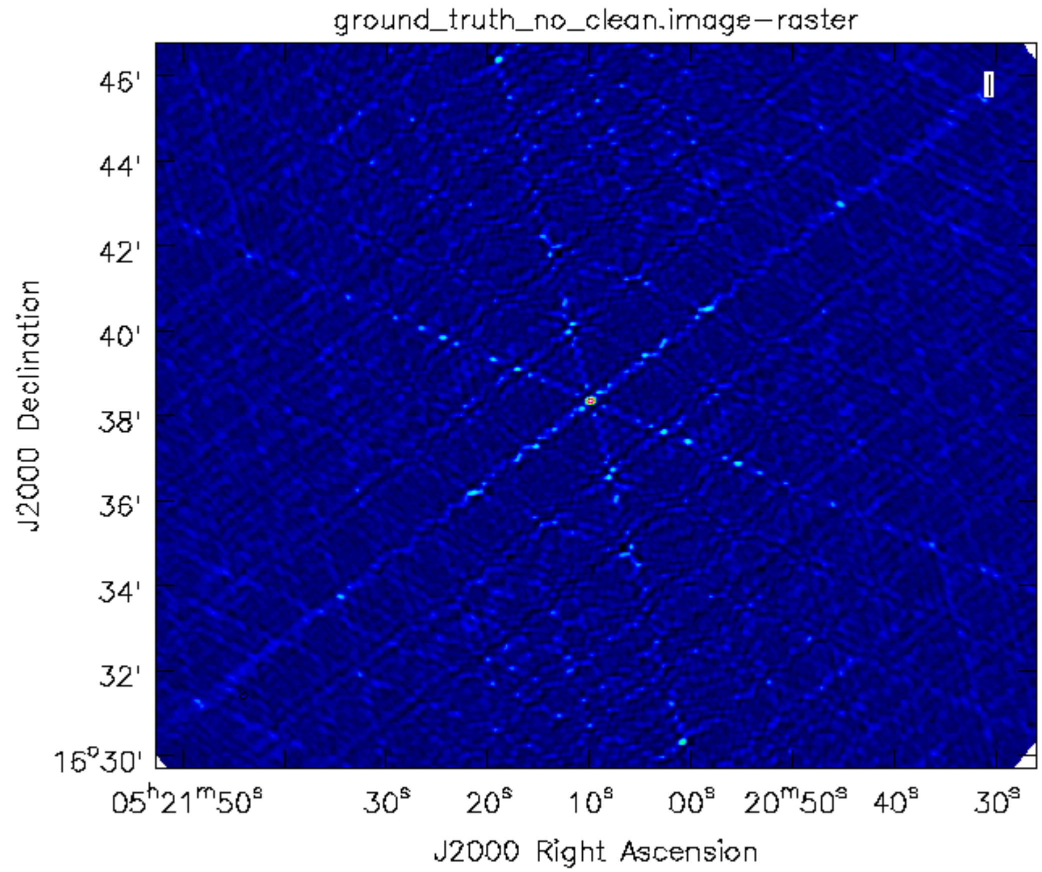
RFI Spectrum

- Scatter plot of correlator self-powers of every baseline and dump time vs channel.
- Channels 15 – 105 are severly corrupted by RFI
- Some RFI residual power is present outside allocated downlink band due to filter spectral leakage

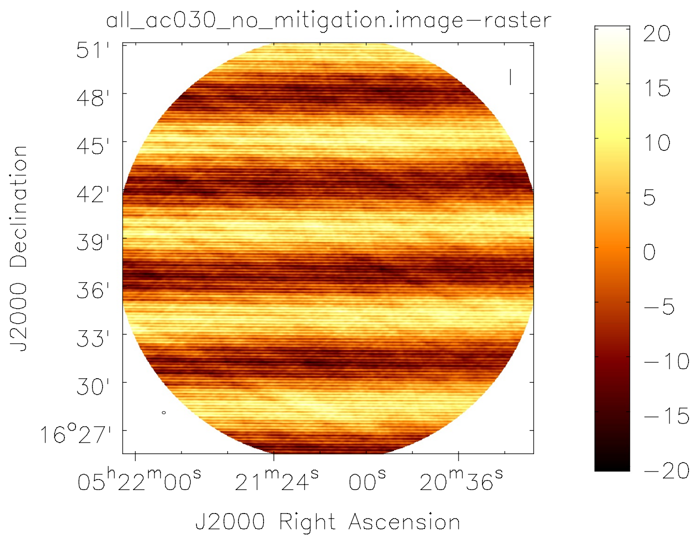


3C128 "Ground Truth" Image

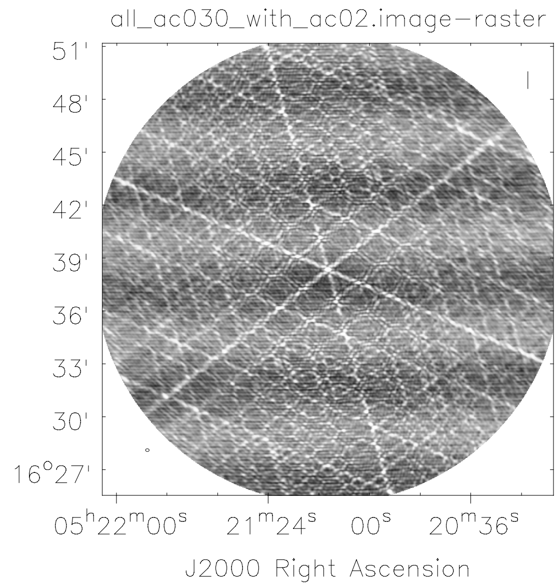
- From RFI corrupted data set
- Only relatively RFI-free channels used (106 – 256)
- This is the Dirty image; no CLEAN deconvolution
- Formed from 60s of data
- Artifacts are dirty beam sidelobe pattern



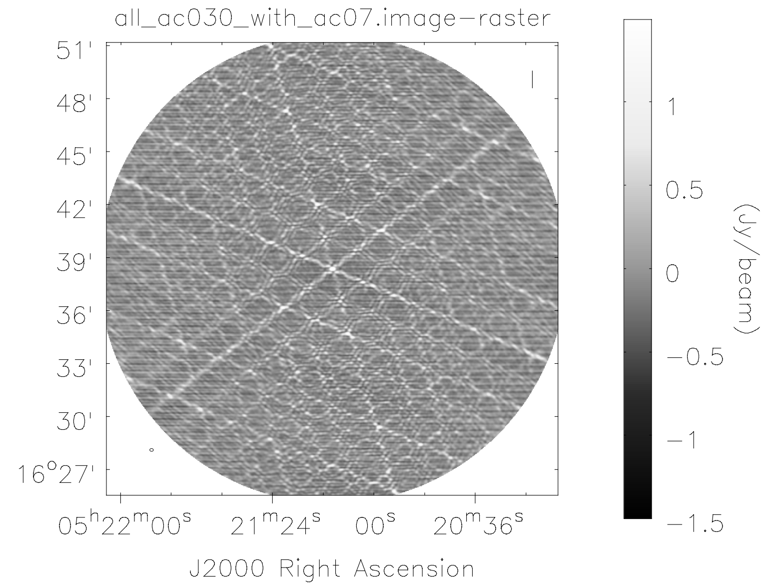
Images from RFI Corrupted Channel 30



No RFI mitigation



SP: 2 eigenvectors removed



SP: 7 eigenvectors removed

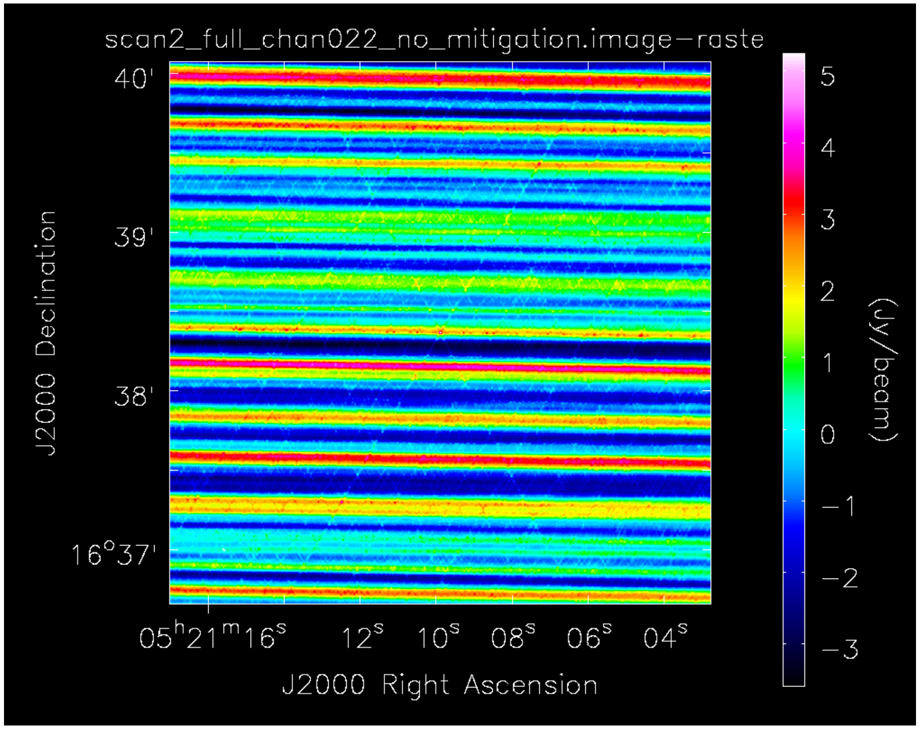
Data set 2, Summer 2018, VLA configuration A

- 36.4 km longest baseline
- All other parameters same as Set 1:
  - Source: 3C138, point calibrator in southern sky
  - RFI: Geostationary SiriusXM sat.
  - Processing band center 2.34 GHz
  - Correlator channel bandwidth: 62.5 kHz, 256 chans., 16 MHz total
  - Correlator dump interval: 20 ms
  - 2 arcsec pixels, spans 6 integration

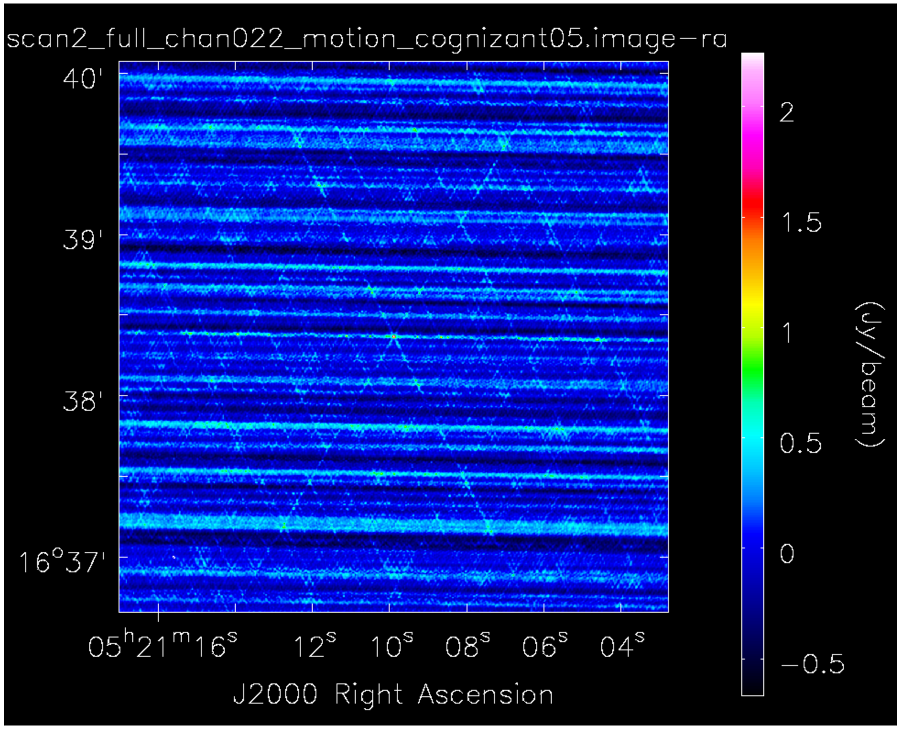
- Significant RFI decorrelation on longer baselines (ngVLA-like)
- Only one subarray with sufficient mutually RFI correlation to form projection matrices: inner core of 9 antennas. (Not ngVLA-like)
- Full array SP did reduce RFI, but performed much poorer than C config.

Full Configuration A array SP Results

Full Configuration A array SP Results

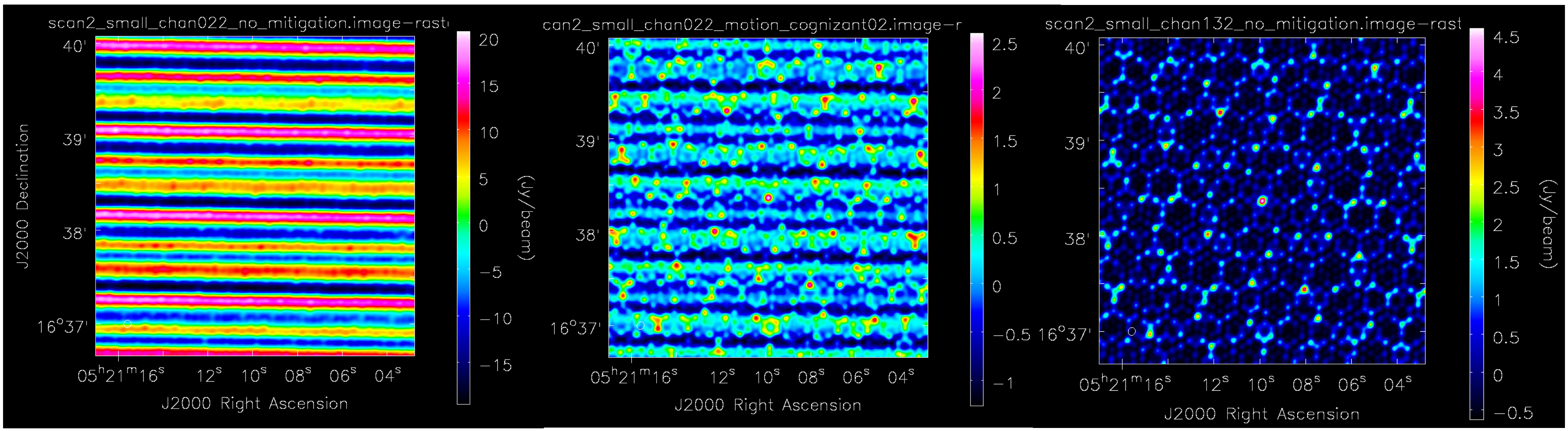


No mitigation



5 eigenvectors removed

9-element Core Subarray SP Results



No mitigation

3 eigenvectors removed

"Ground truth" dirty image from channel with no RFI

Conclusions

- For compact interferometric arrays, subspace projection (SP) is effective.
- Longer baseline arrays will perform poorly with SP due to RFI decorrelation.
- In simulation, the subarray processing (SAP) algorithm works well for a proposed geometry of the ngVLA
- Very large numbers of elements (as with ngVLA) are needed to form effective SAP subarrays for arbitrary RFI DOAs.
- RFI decorrelation for large arrays can be mitigated by using proportionally narrower correlator channel bandwidths
- When not doing SP or SAP, RFI decorrelation helps, but we have demonstrated scenarios where SAP is a much better option.