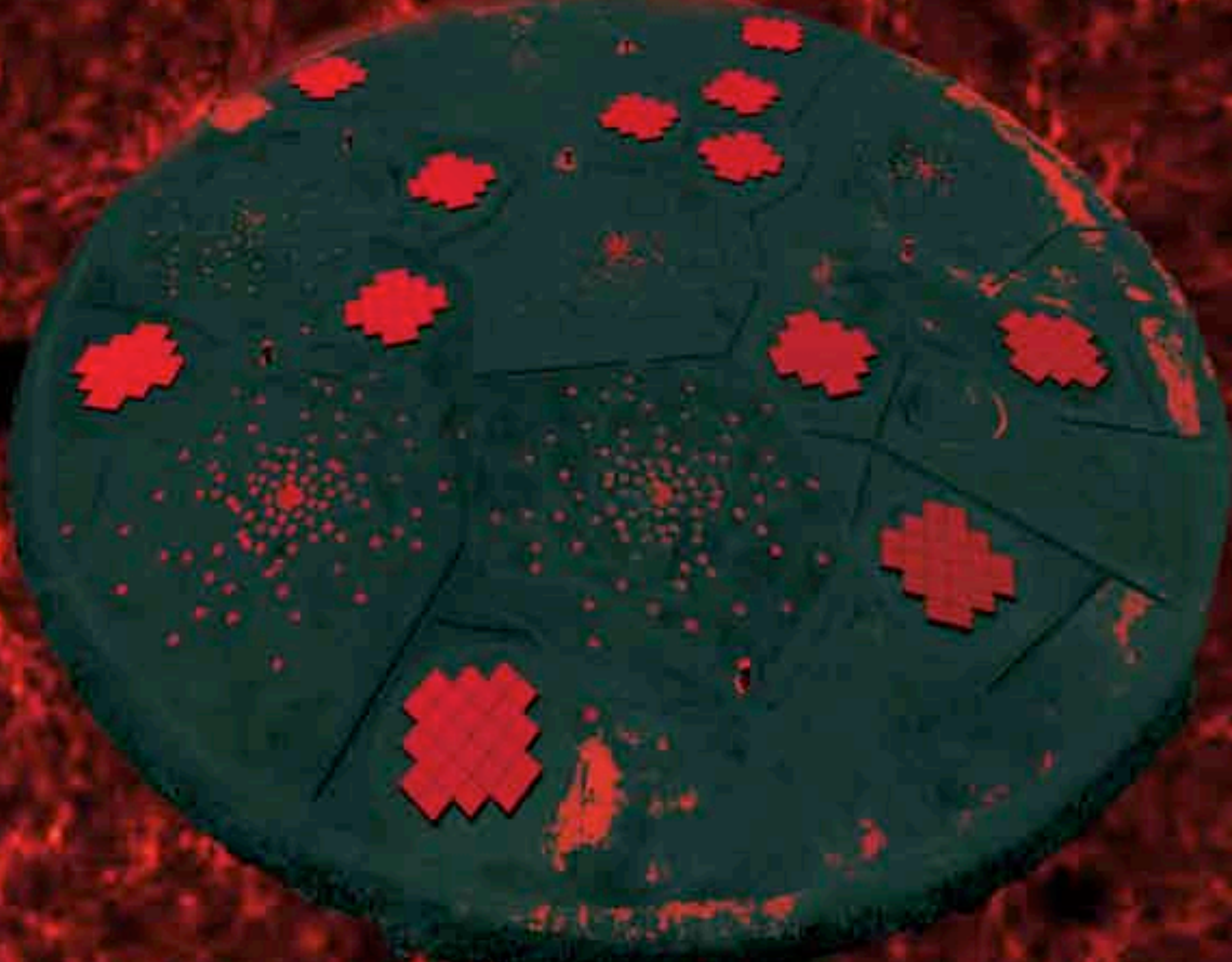
university of groningen

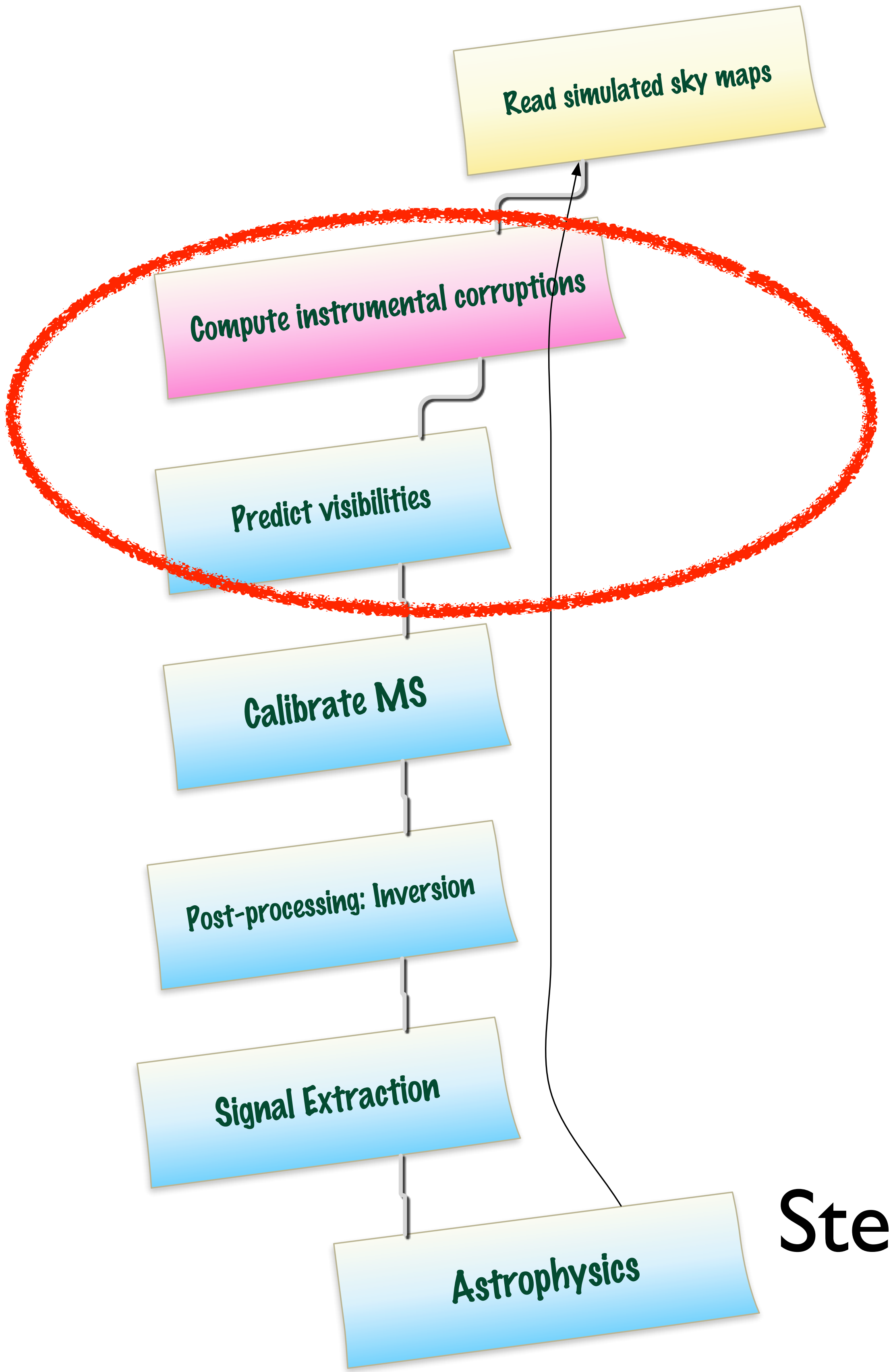
ASTRON

Simulations, Calibration and Inversion of LOFAR EoR data



Panagiotis Lampropoulos

EoR Simulation and testing pipeline



Step 1: Sky realizations

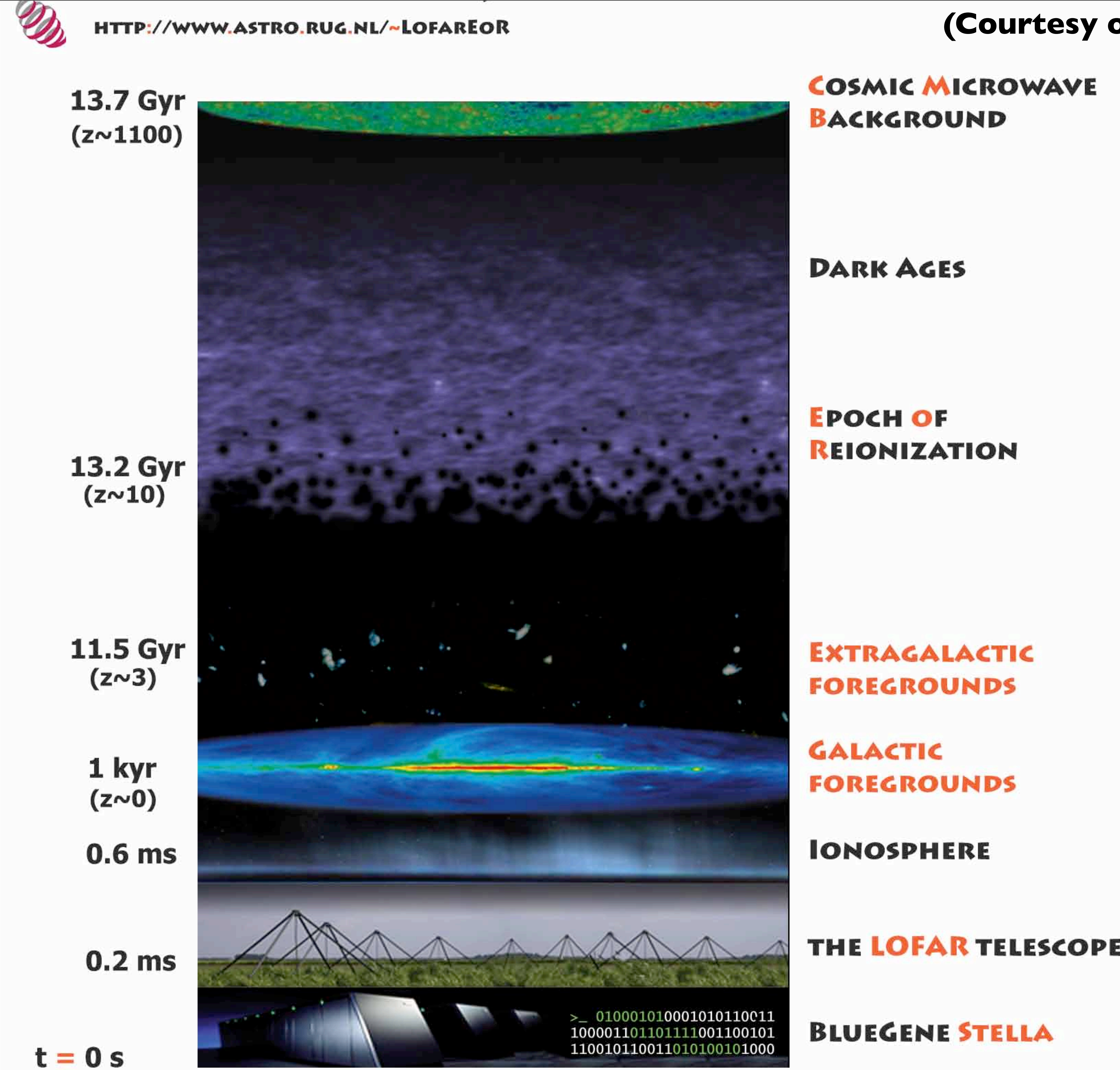
Step 2: Predict Visibilities (Part I)

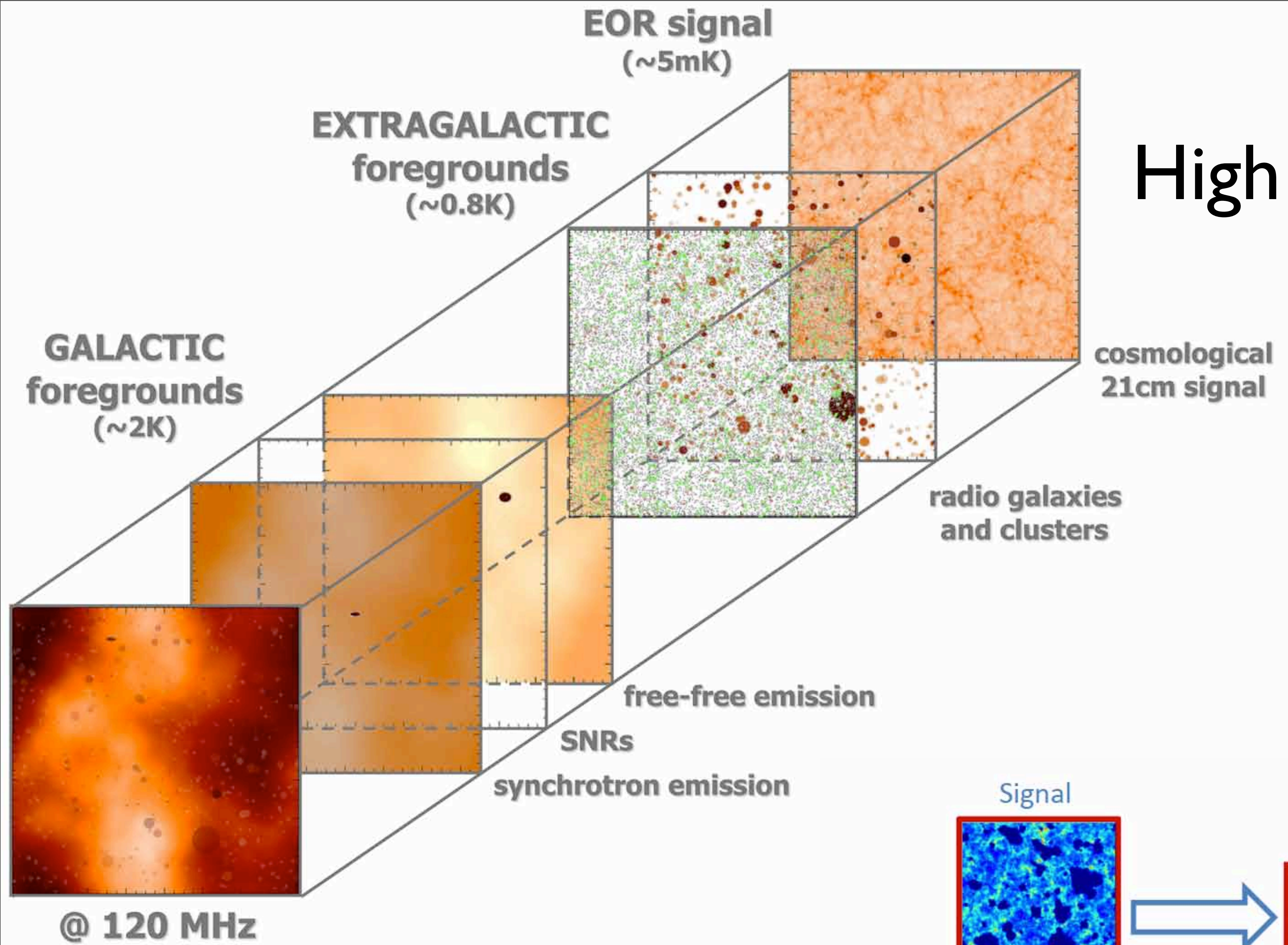
Step 3: Self-cal

Step 4: Inversion (Part II)

Step 5: Extraction and interpretation

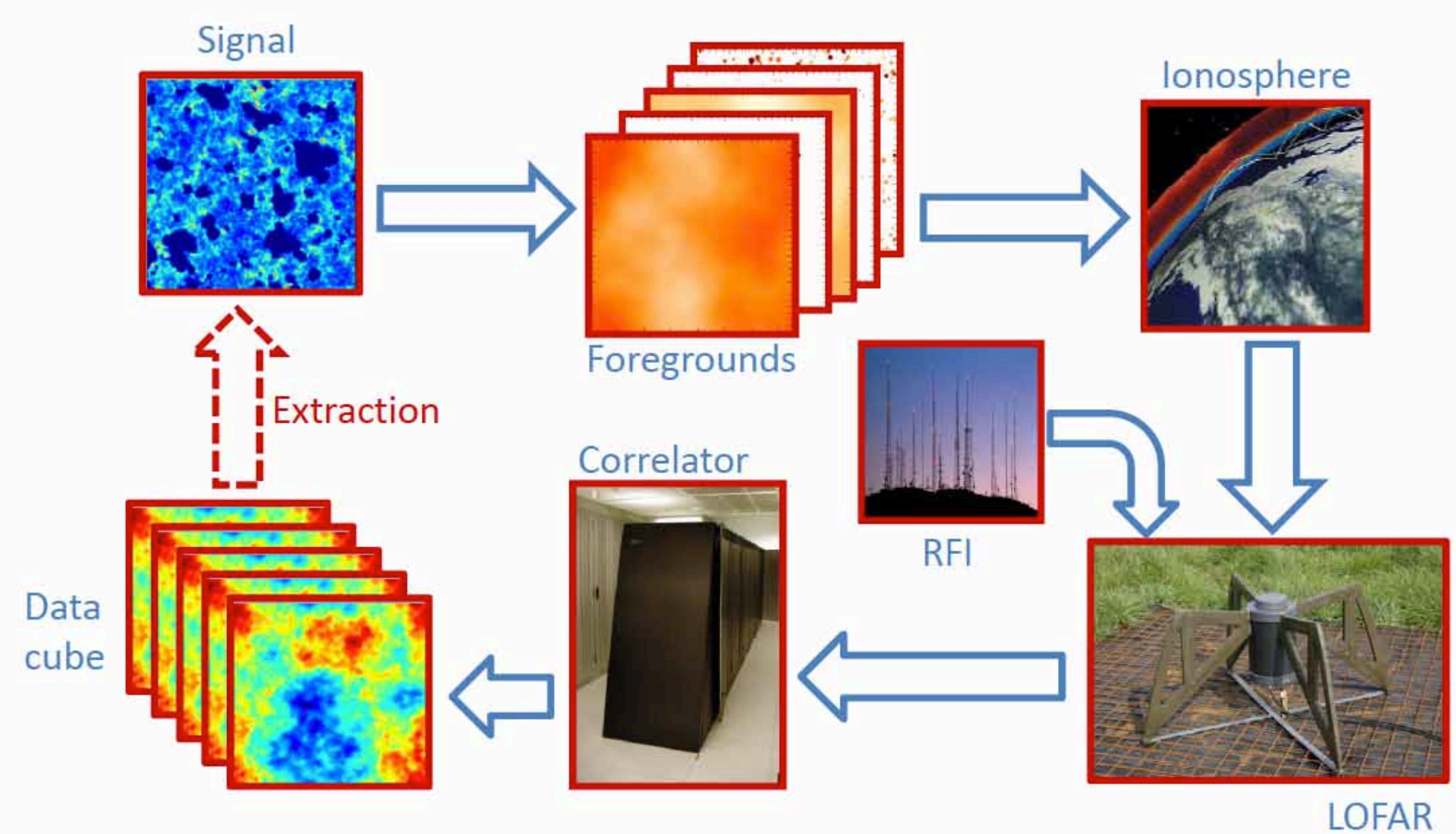
(Courtesy of V. Jelic)





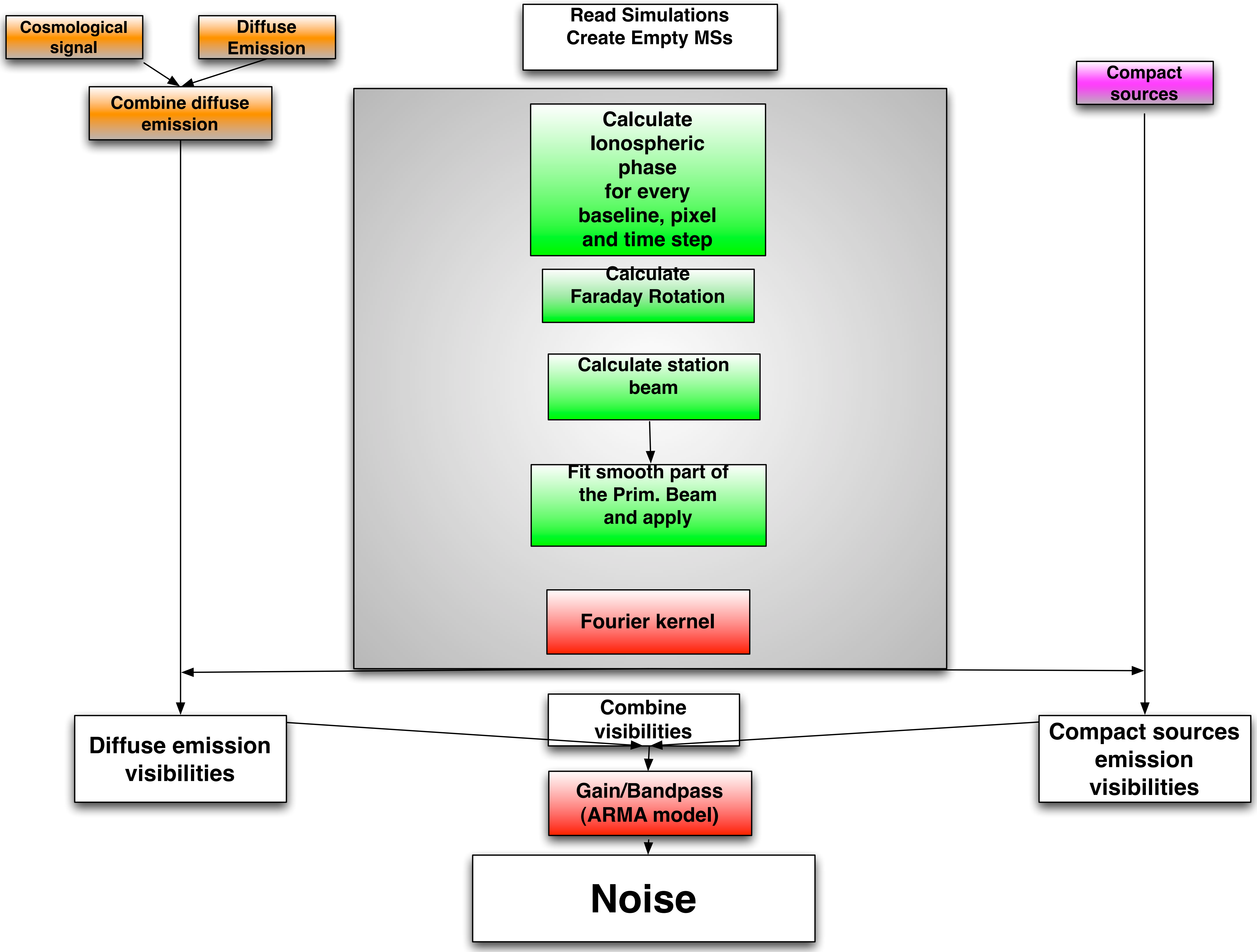
High dynamic range

(Jelic et al.)



Complex signal path

End To End Pipeline: Prediction

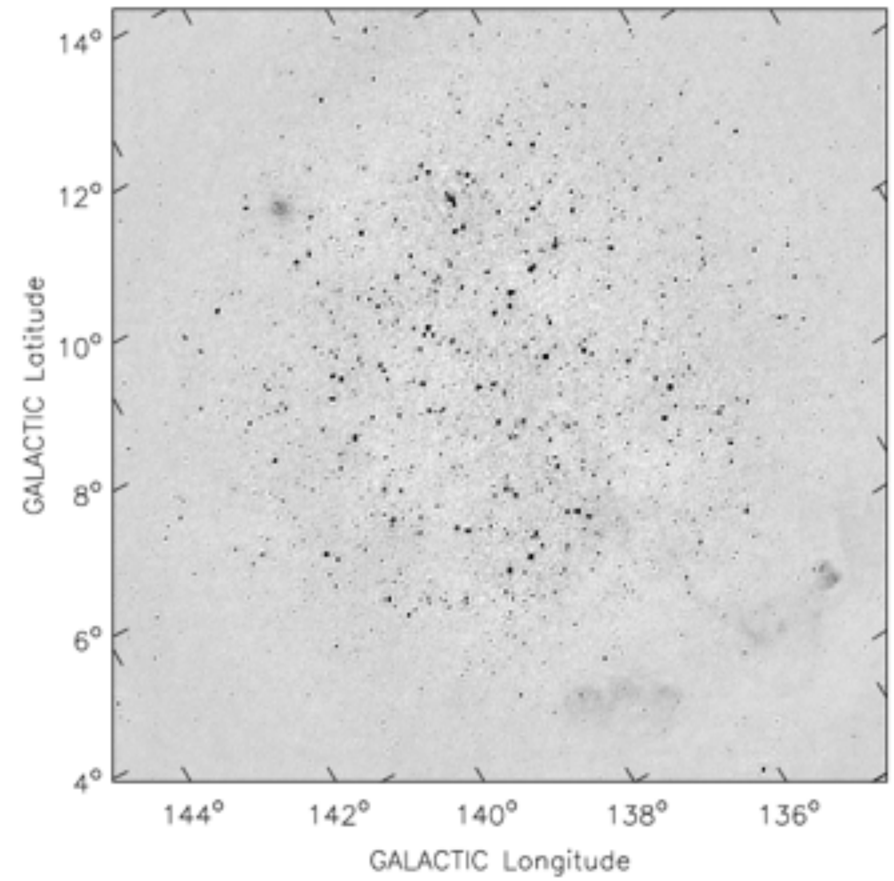


Simulation Specifications

Coordinates of the 3C196 field centre (J2000.0)	$\alpha = 8^h13^m36^s.0$, $\delta = 48°13'03''$
Galactic coordinates	$l \simeq 171°$, $b \simeq 33°$
Number of spectral bands	128
Frequency coverage (MHz)	120 - 184
Width of each band (MHz)	1
Frequency resolution (MHz)	0.5
Time resolution (sec)	30
FoV	$\sim$ 10 deg
Noise at 150 MHz (mK)	840 mK
Obs. duration (hrs)	4

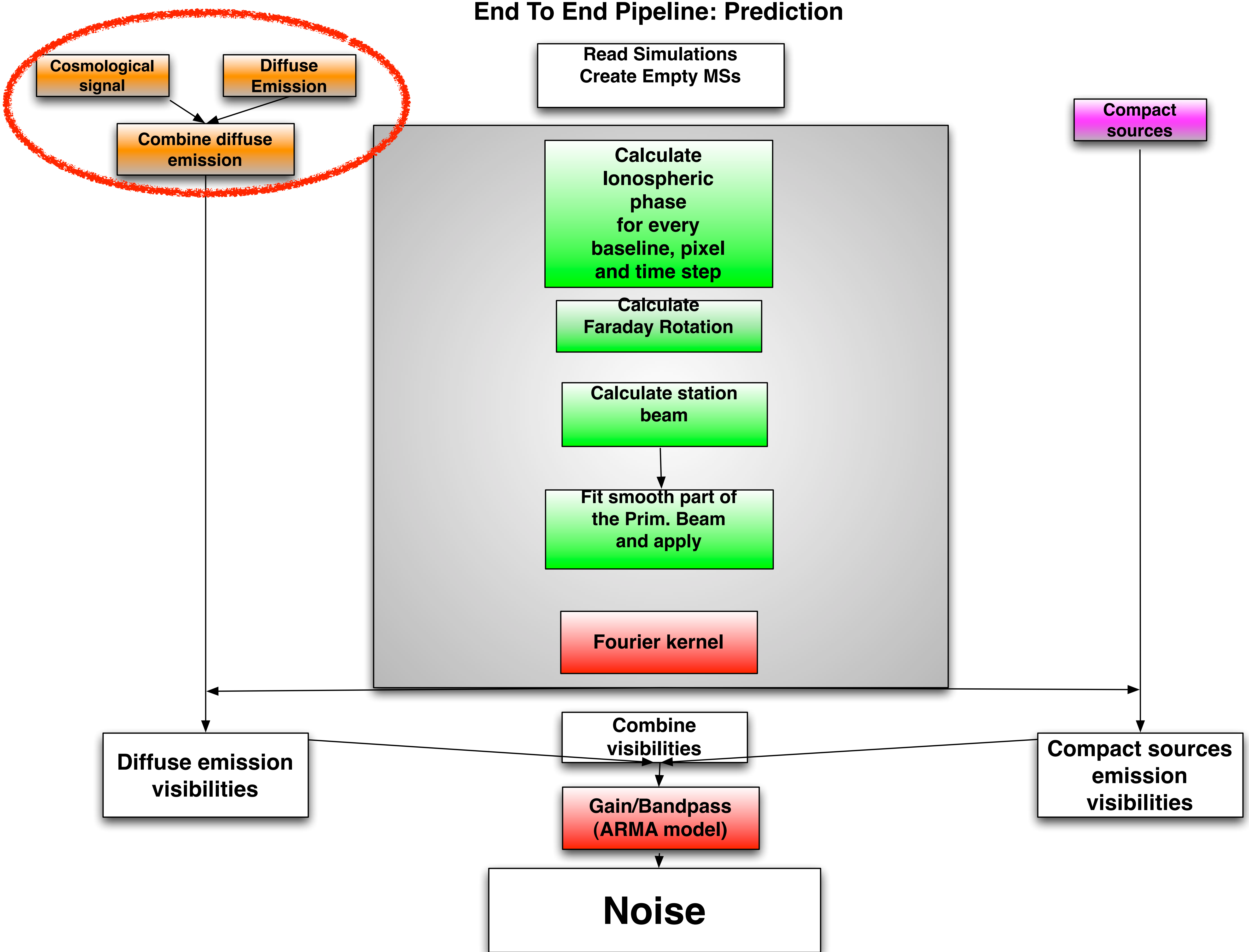
Point sources:

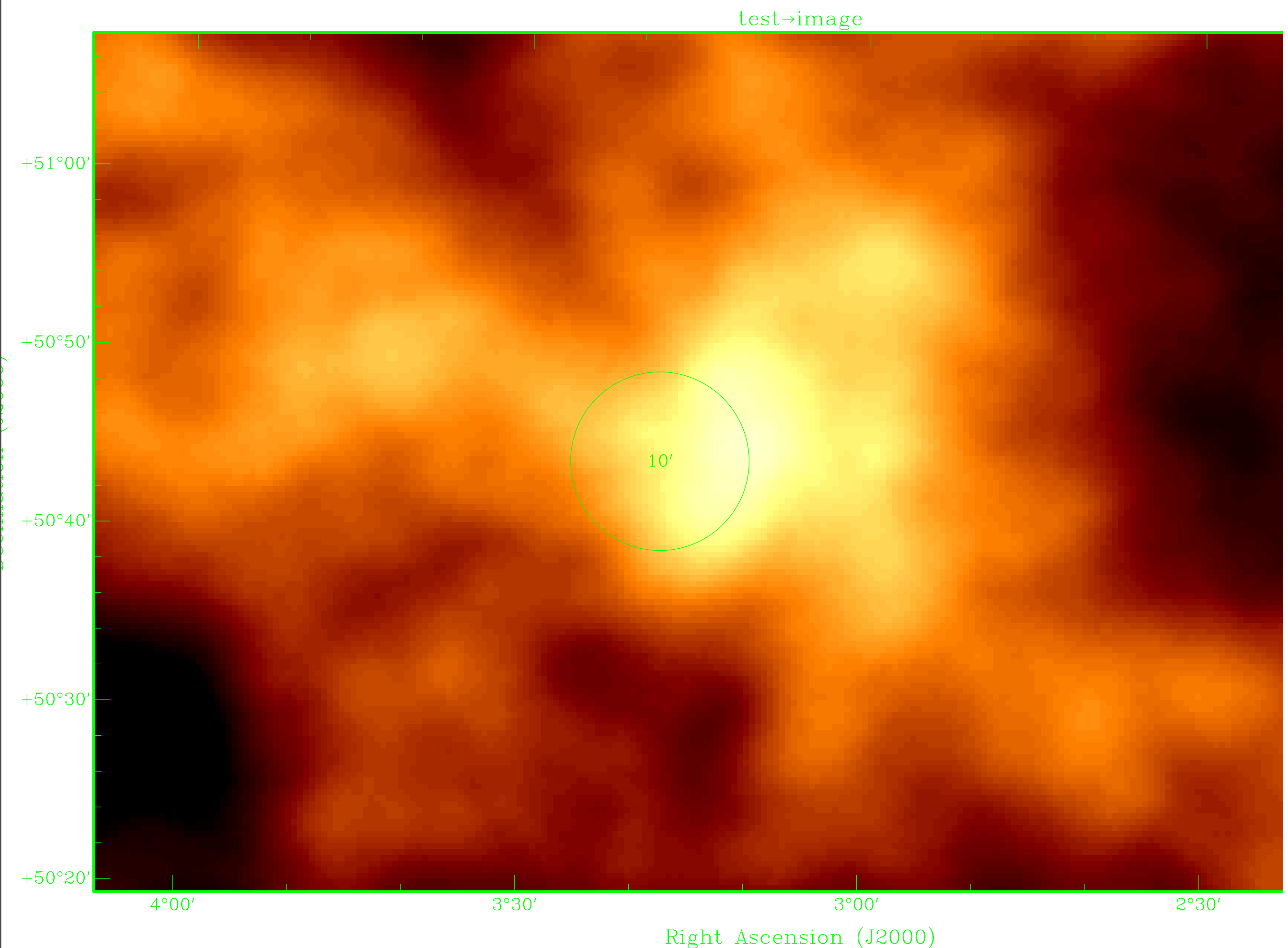
- 3C 196 (75 Jy @ 138 MHz)Field
- RA: 08 : 13 : 16.05, DEC: +48 : 13 : 02.58
- (+ A team)
- Point-sources extracted with duchamp
- Random spectral indices
- Tail of the point-source brightness distribution from Jelic et al.



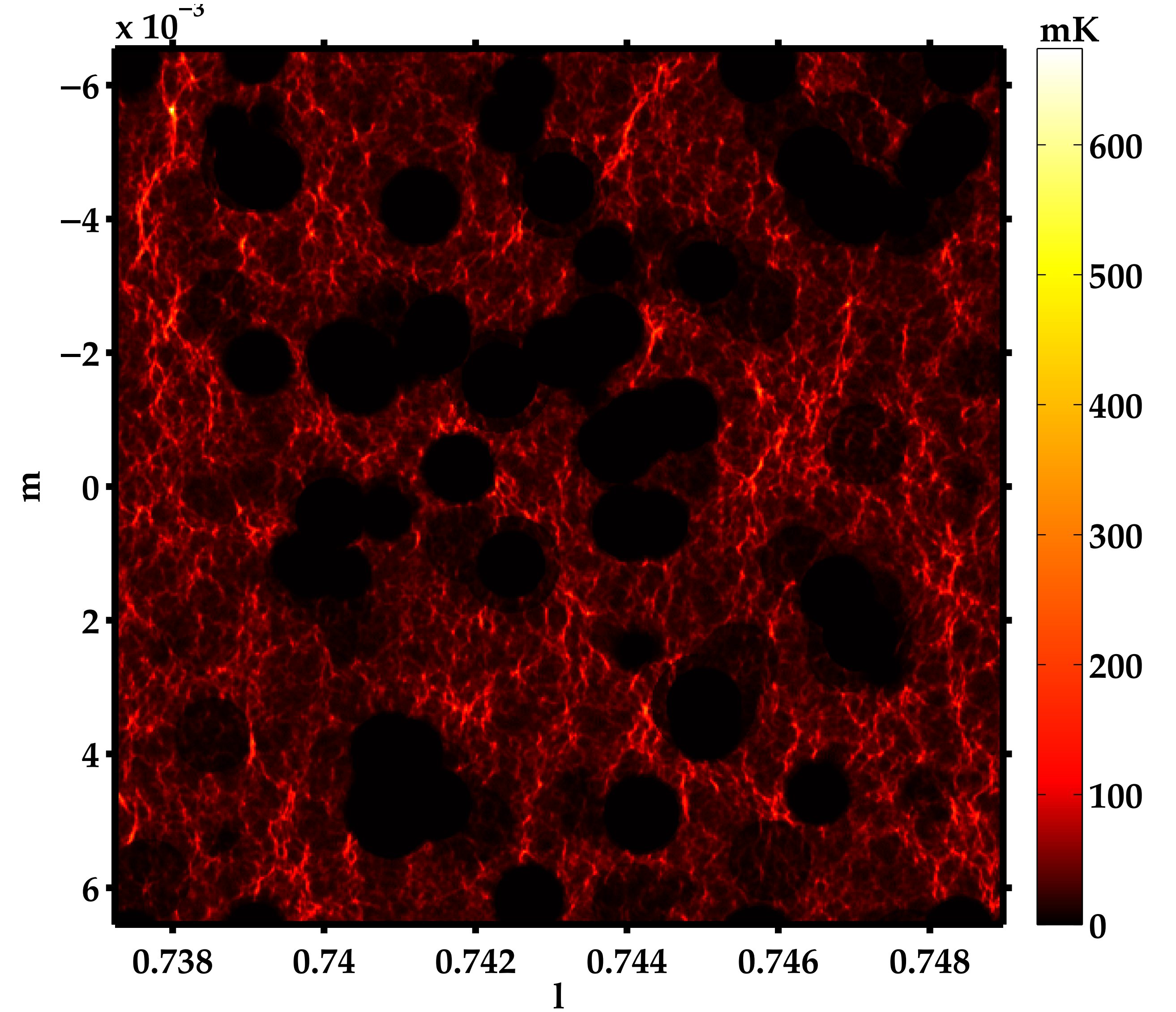
(Bernardi et al.)

End To End Pipeline: Prediction

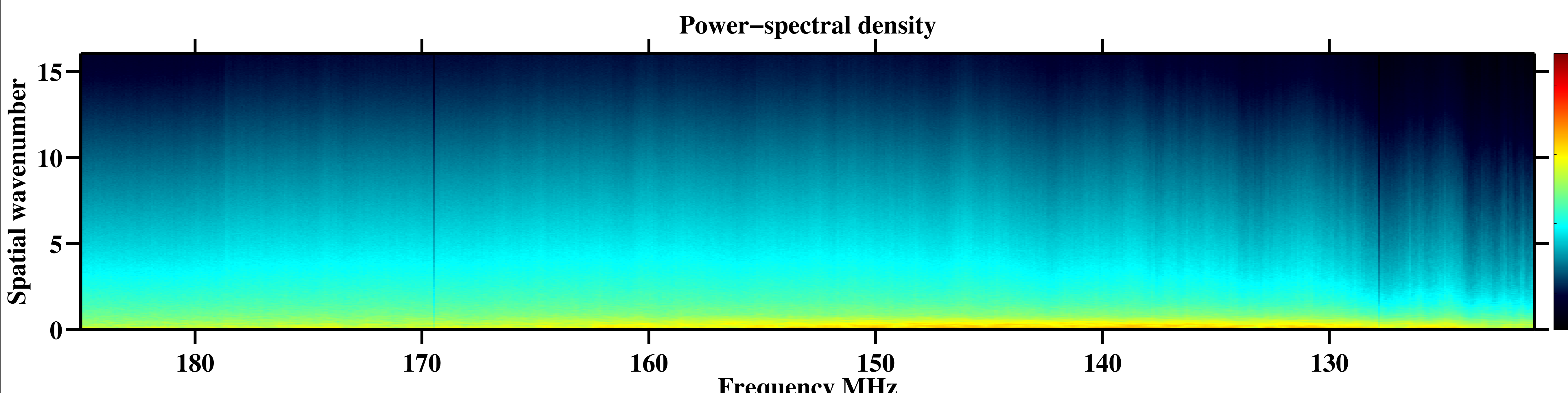




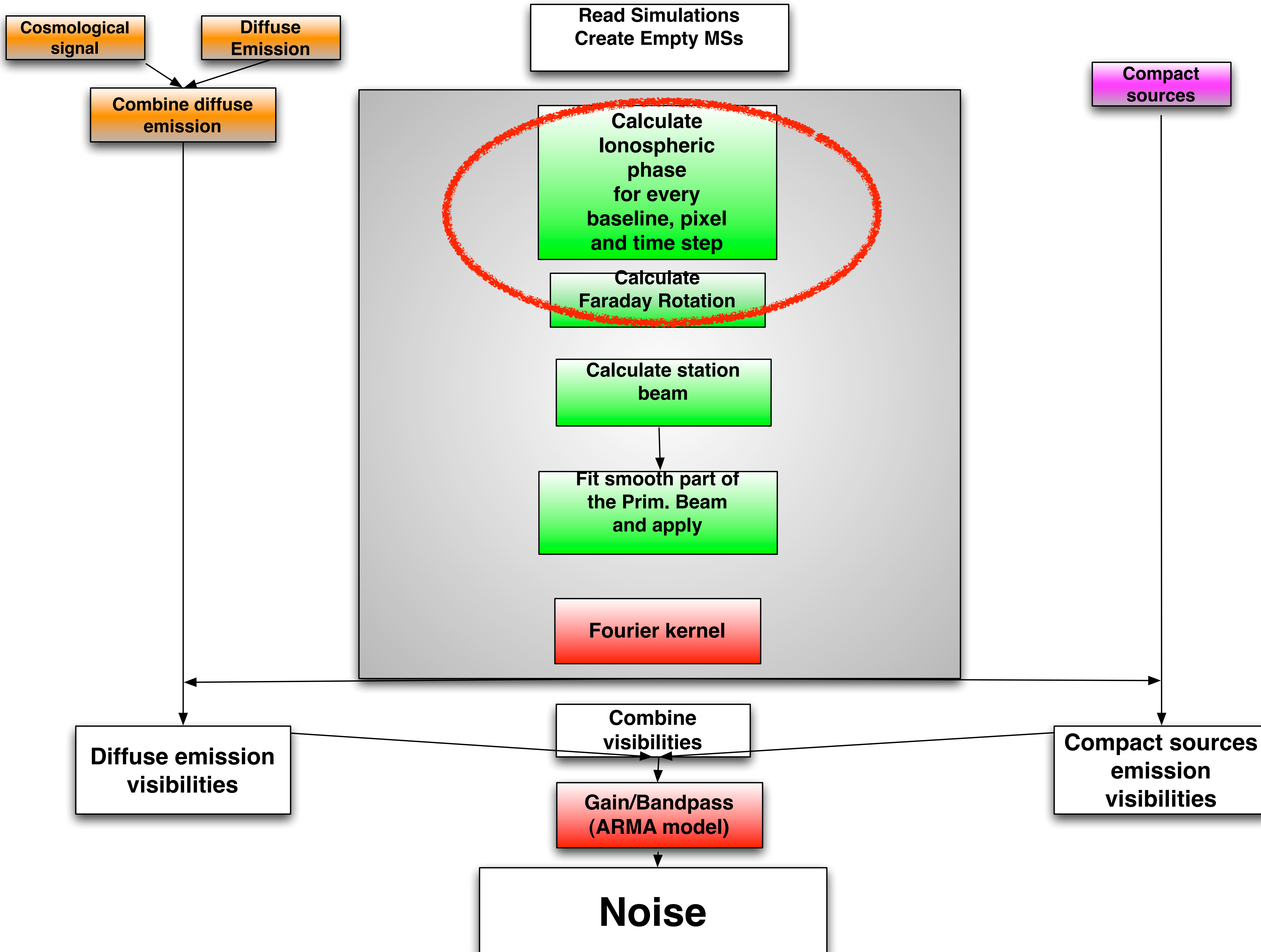
Diffuse galactic foregrounds as GRF with normalization taken from Jelic et al.



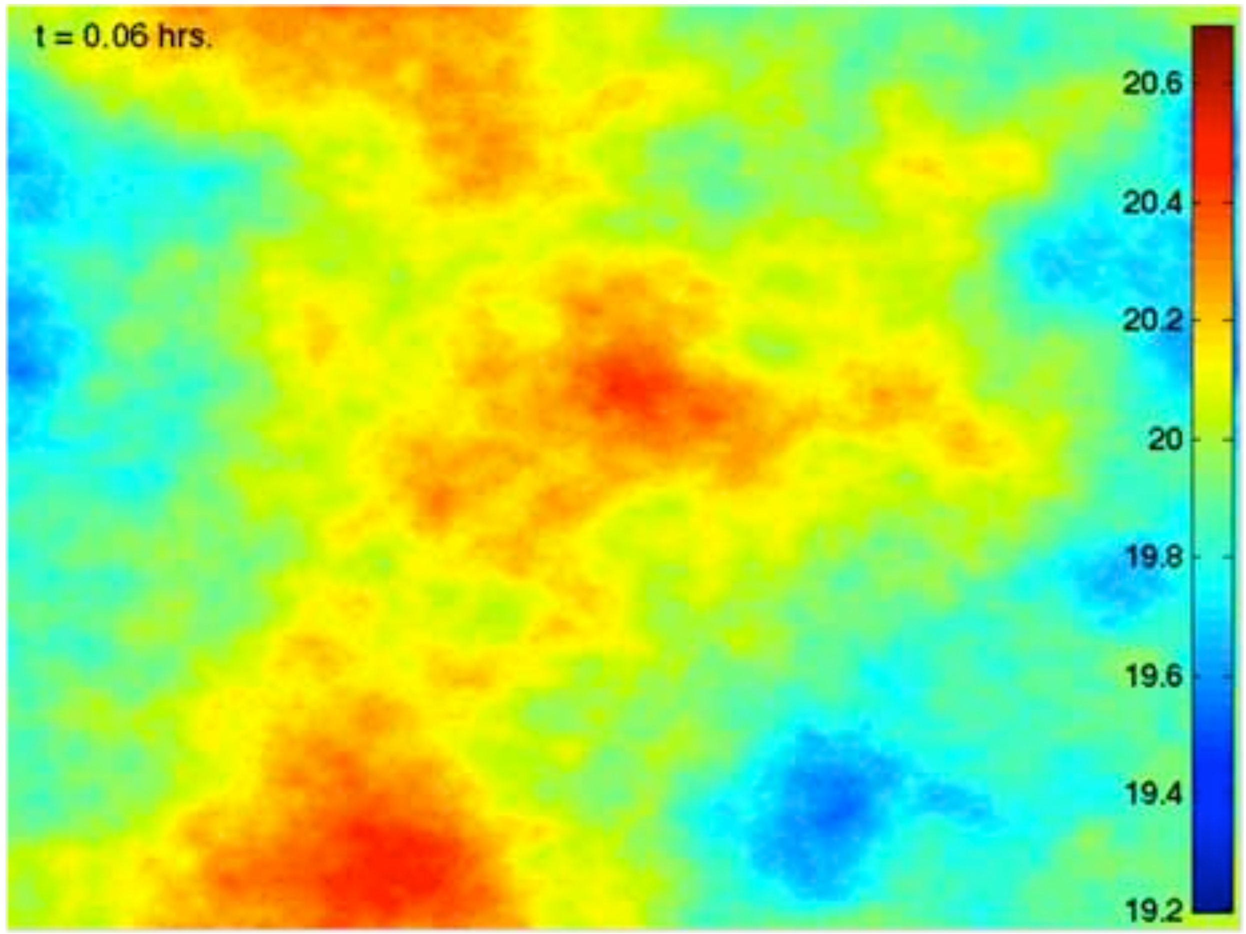
200 Mpc 21-cm signal cubes from R. Thomas

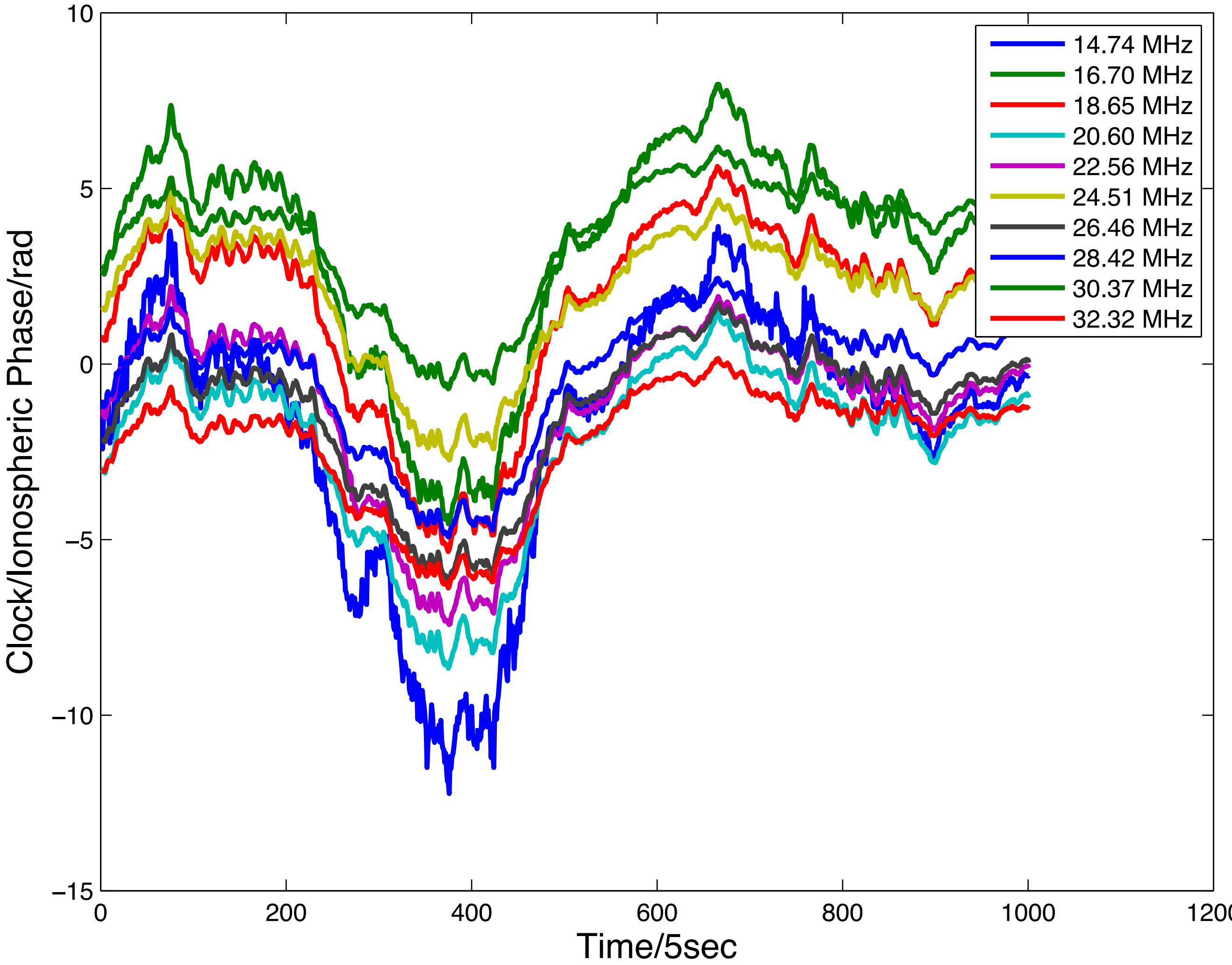


End To End Pipeline: Prediction



Ionosphere

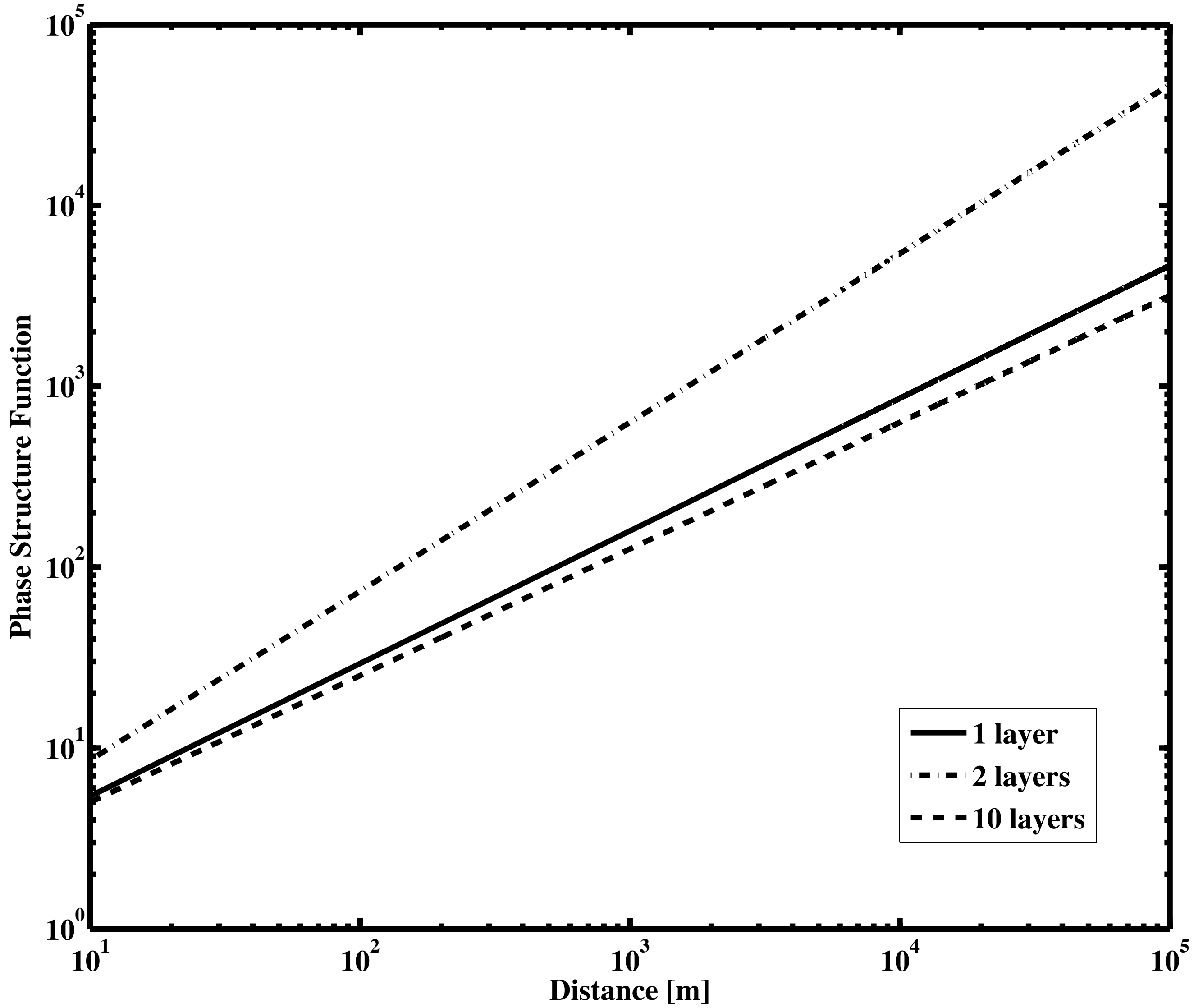




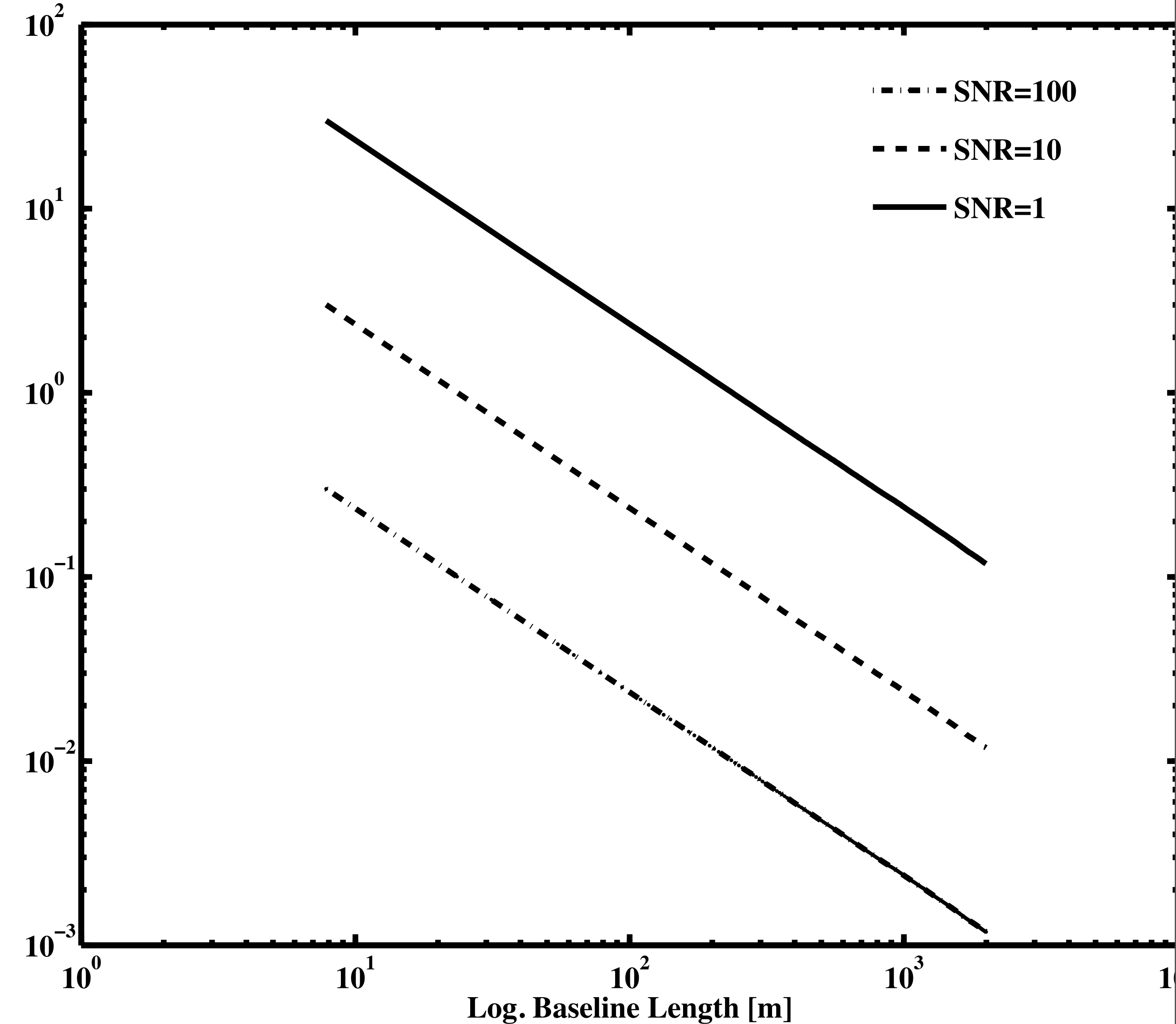
- Superposition of TIDs and Kolmogorov turbulence (cascading)
- Large scale TID behavior from Spoelstra and Velthoven

$$\phi_{ij} = 2A\sin(k(x_i - x_j)/2)\sin[\omega t - k(x_i + x_j)/2]$$

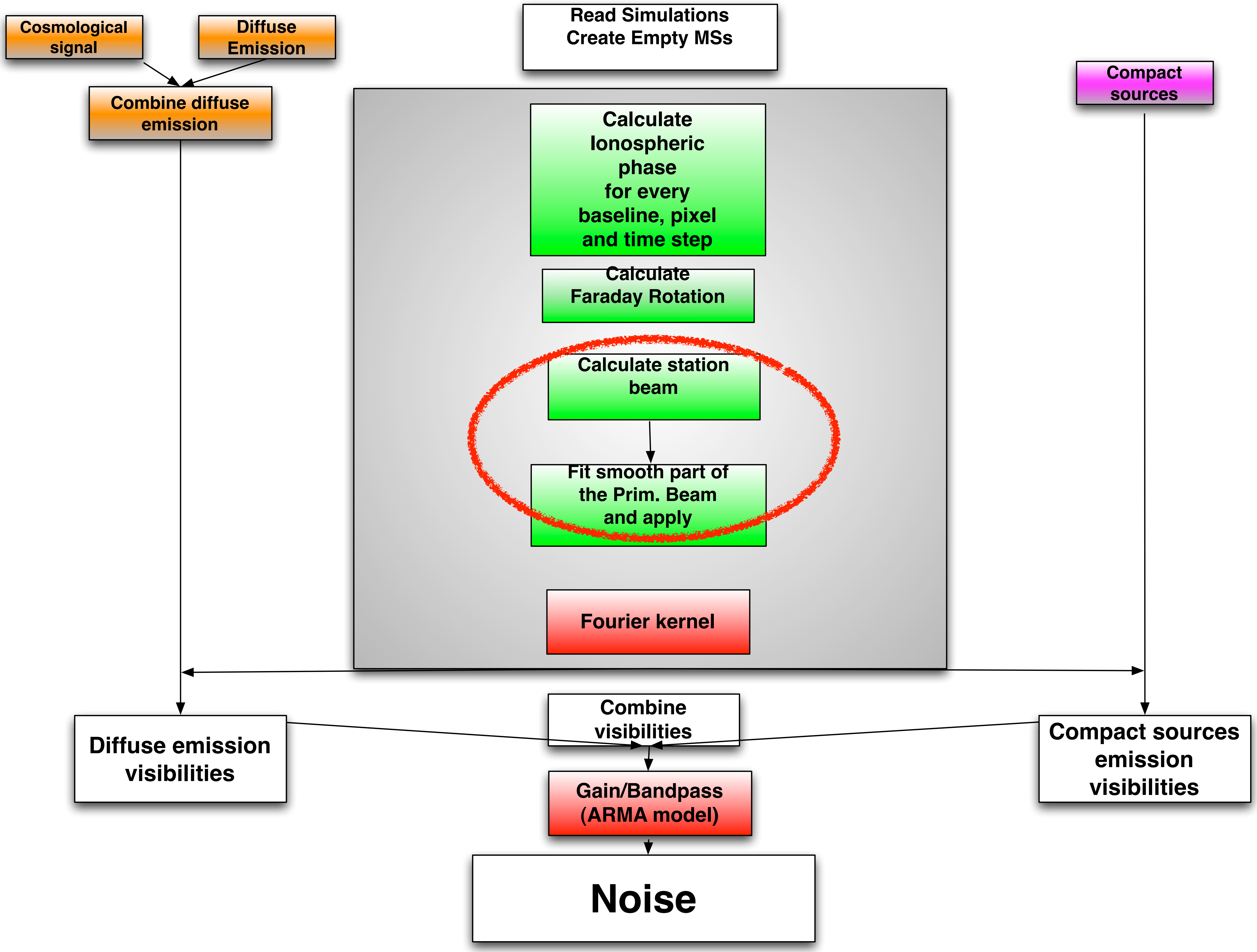
Phase 2-Structure function



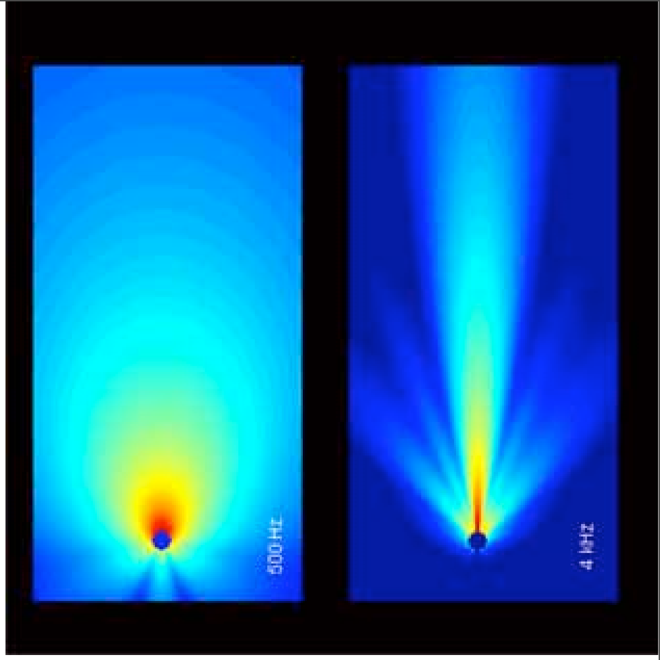
Position Error



End To End Pipeline: Prediction



Beam shapes

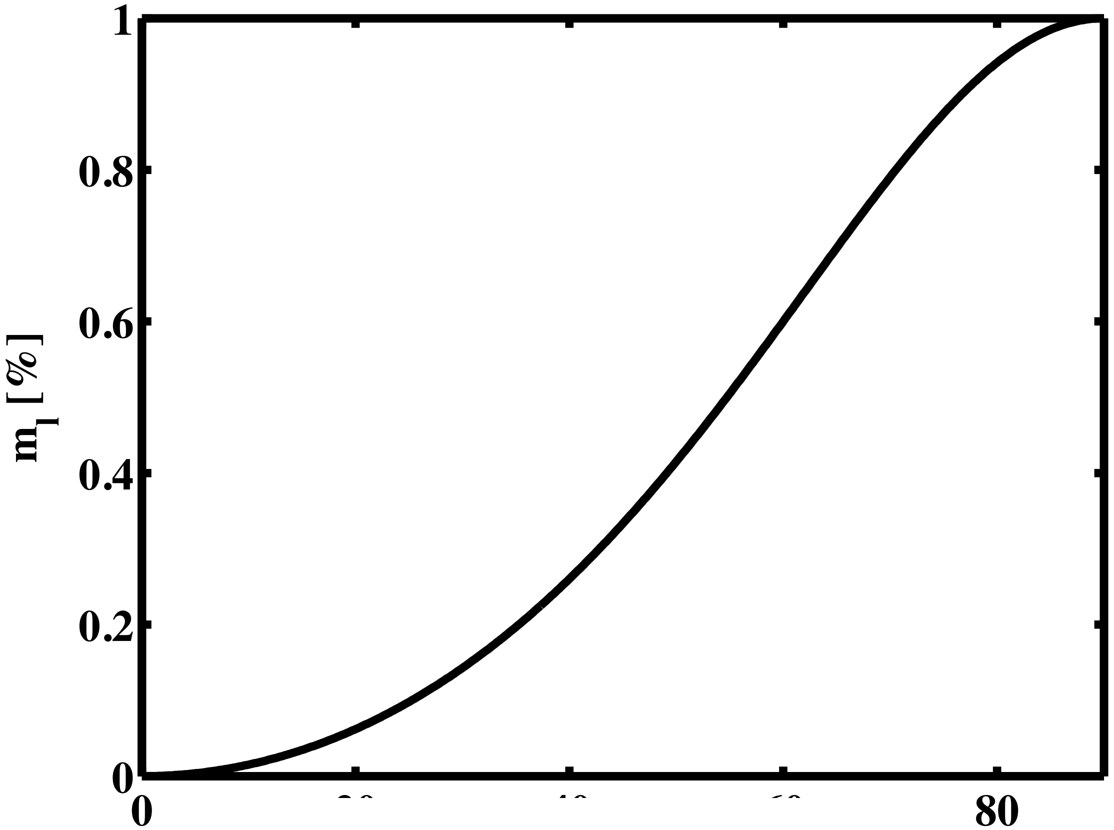


Sarod Yatawatta, "LOFAR beamshapes and their use in calibration and imaging", ASTRON Tech. Report, 2007

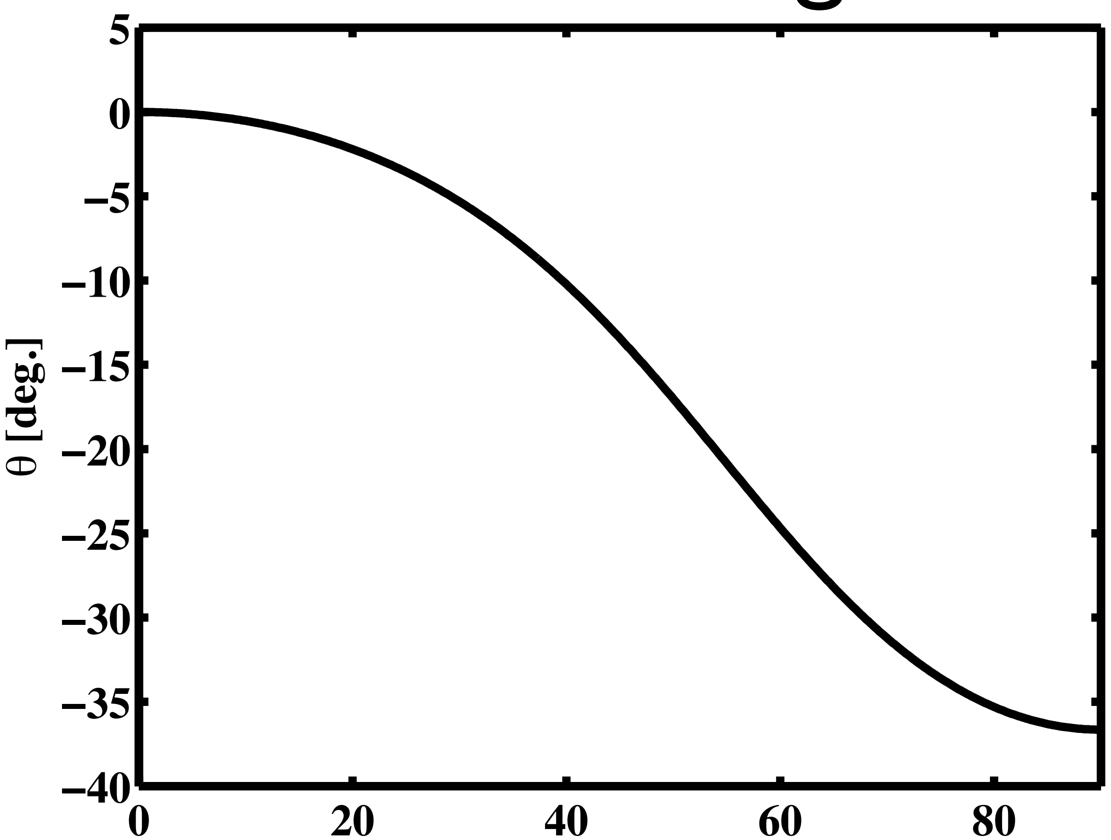
Polarized response due to dipole projection

- Narrowband beamformer

$$\begin{bmatrix} \begin{array}{c} \mathbf{E}_{\theta,p}(\gamma,\beta)\mathbf{E}^\star_{\theta,q}(\gamma,\beta) \\ +\mathbf{E}_{\phi,p}(\gamma,\beta)\mathbf{E}^\star_{\phi,q}(\gamma,\beta) \end{array} & \begin{array}{c} \mathbf{E}_{\theta,p}(\gamma,\beta)\mathbf{E}^\star_{\theta,q}(\gamma,\beta-\pi/2) \\ +\mathbf{E}_{\phi,p}(\gamma,\beta)\mathbf{E}^\star_{\phi,q}(\gamma,\beta-\pi/2) \end{array} \\ & \\ \begin{array}{c} \mathbf{E}_{\theta,p}(\gamma,\beta-\pi/2)\mathbf{E}^\star_{\theta,q}(\gamma,\beta) \\ +\mathbf{E}_{\phi,p}(\gamma,\beta-\pi/2)\mathbf{E}^\star_{\phi,q}(\gamma,\beta) \end{array} & \begin{array}{c} \mathbf{E}_{\theta,p}(\gamma,\beta-\pi/2)\mathbf{E}^\star_{\theta,q}(\gamma,\beta-\pi/2) \\ +\mathbf{E}_{\phi,p}(\gamma,\beta-\pi/2)\mathbf{E}^\star_{\phi,q}(\gamma,\beta-\pi/2) \end{array} \end{bmatrix}$$

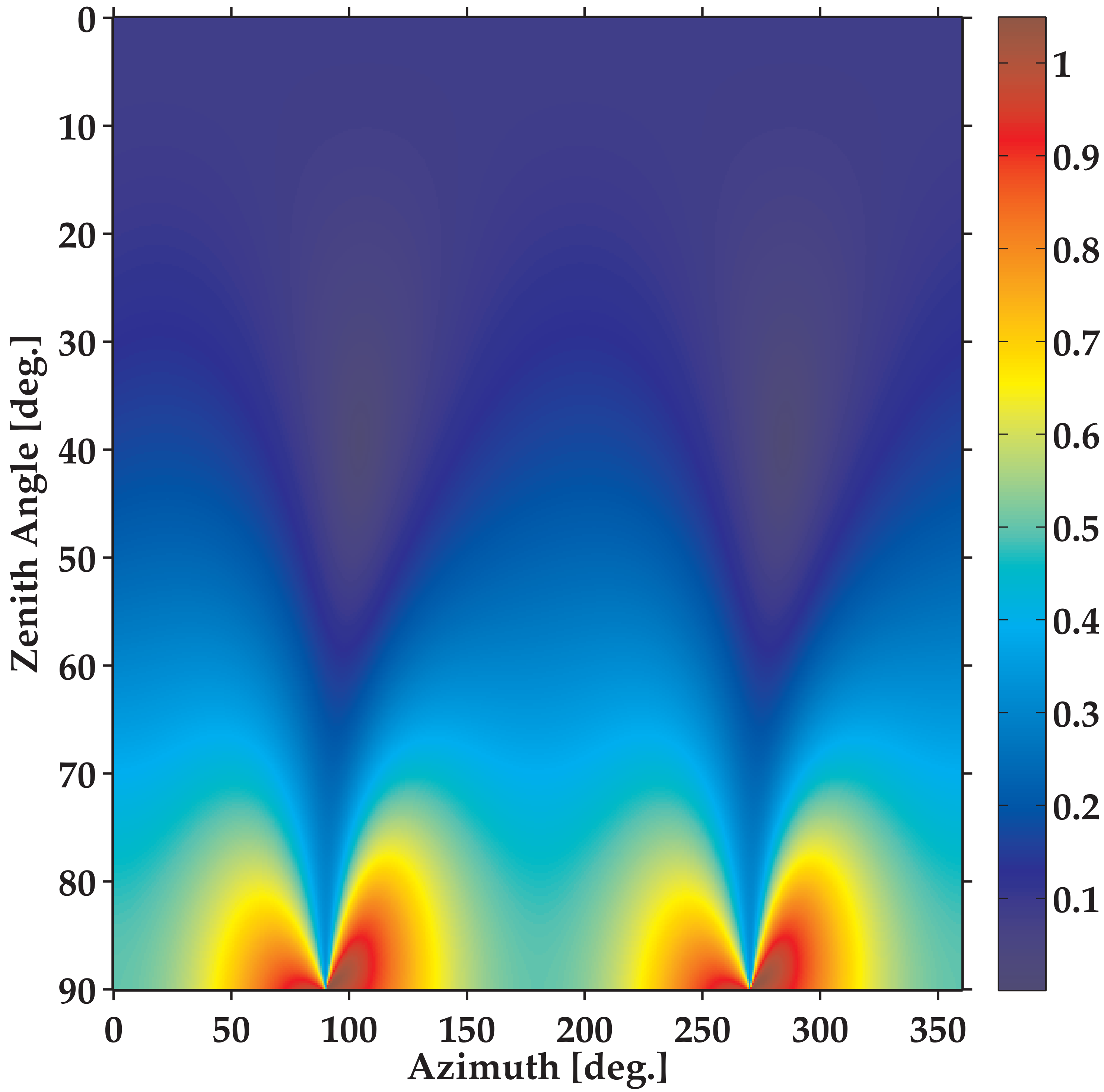


Zenith Angle



Zenith Angle

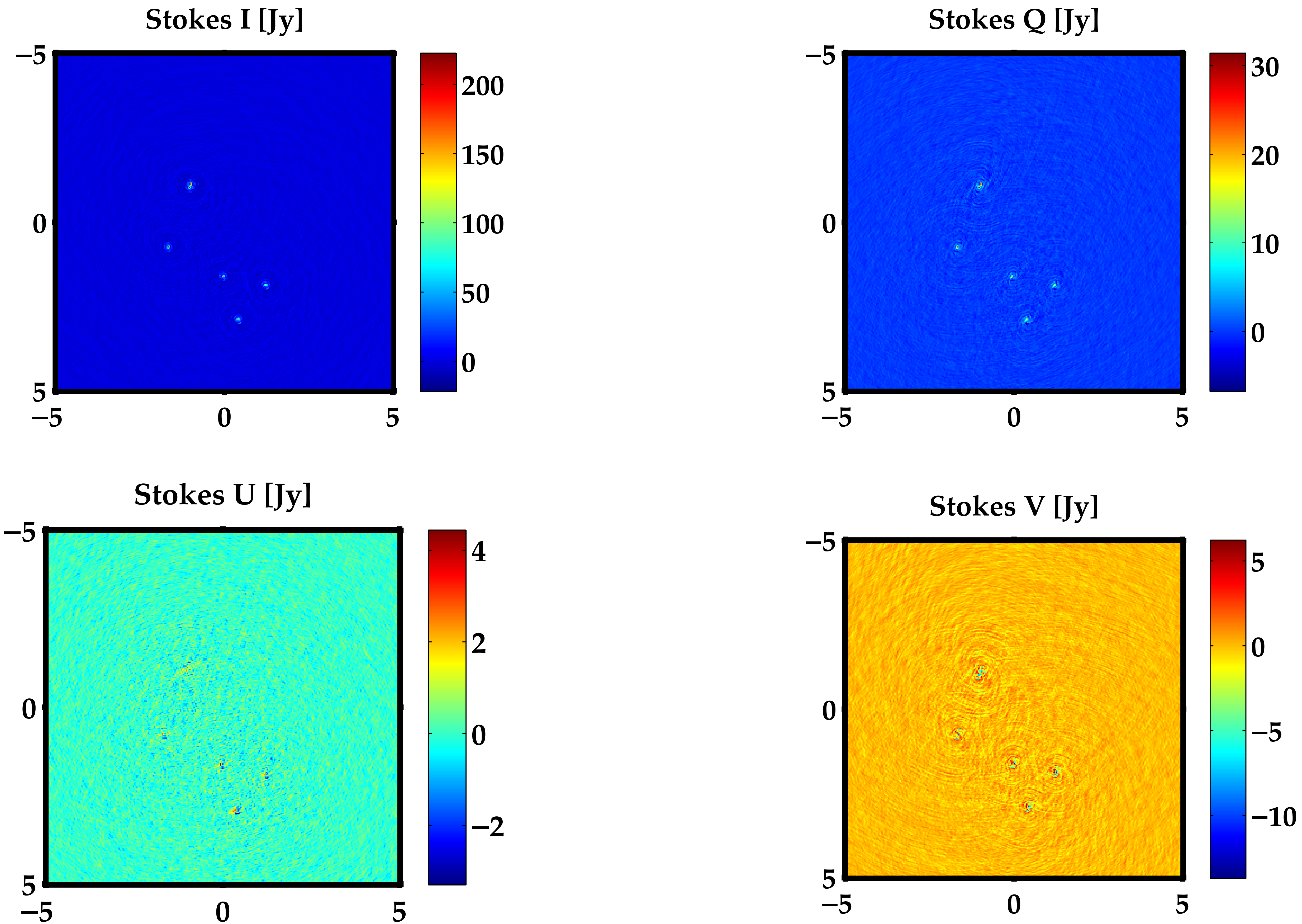
Instrumental polarization



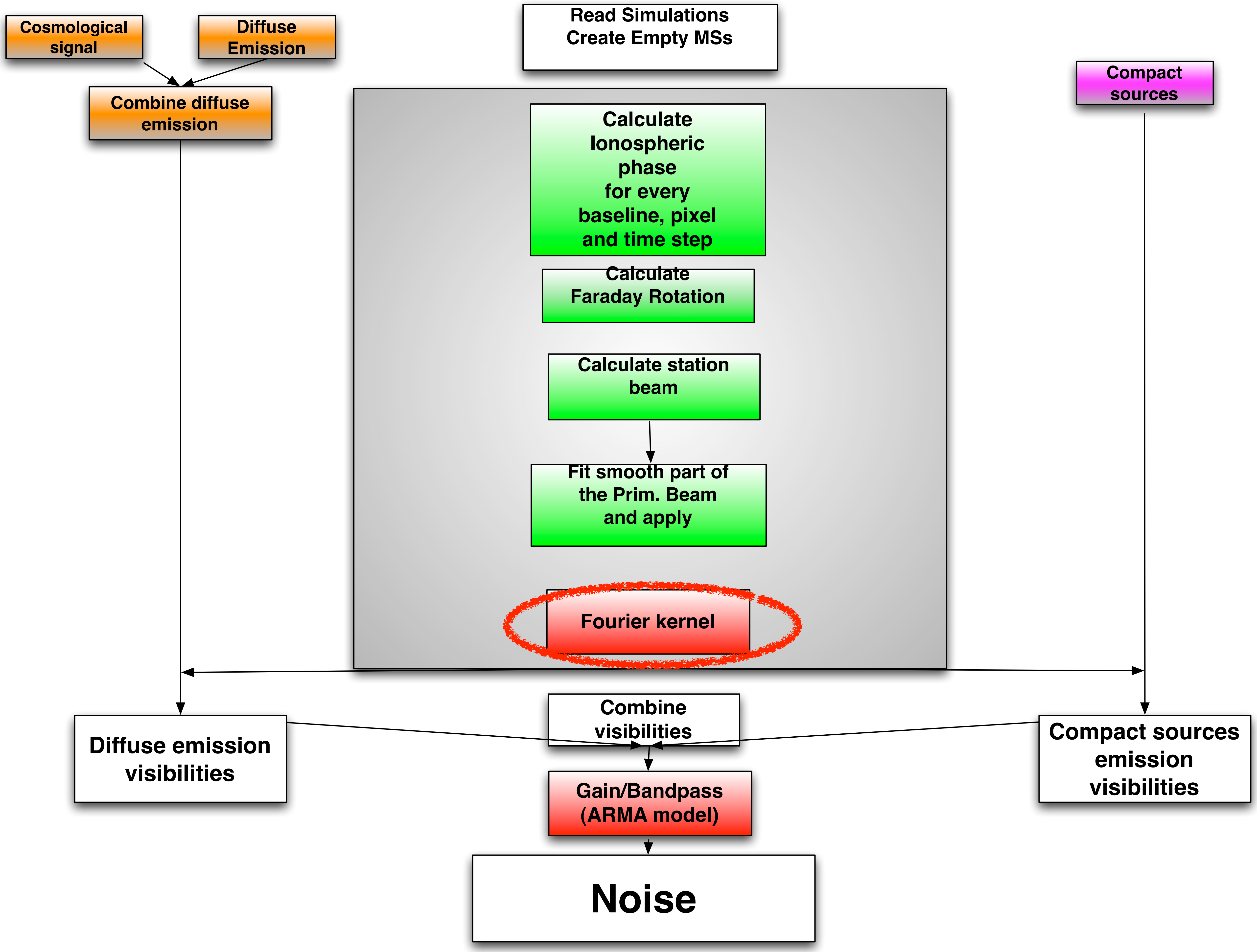
see also Bhatnagar, Carozzi

Beam polarization distortion

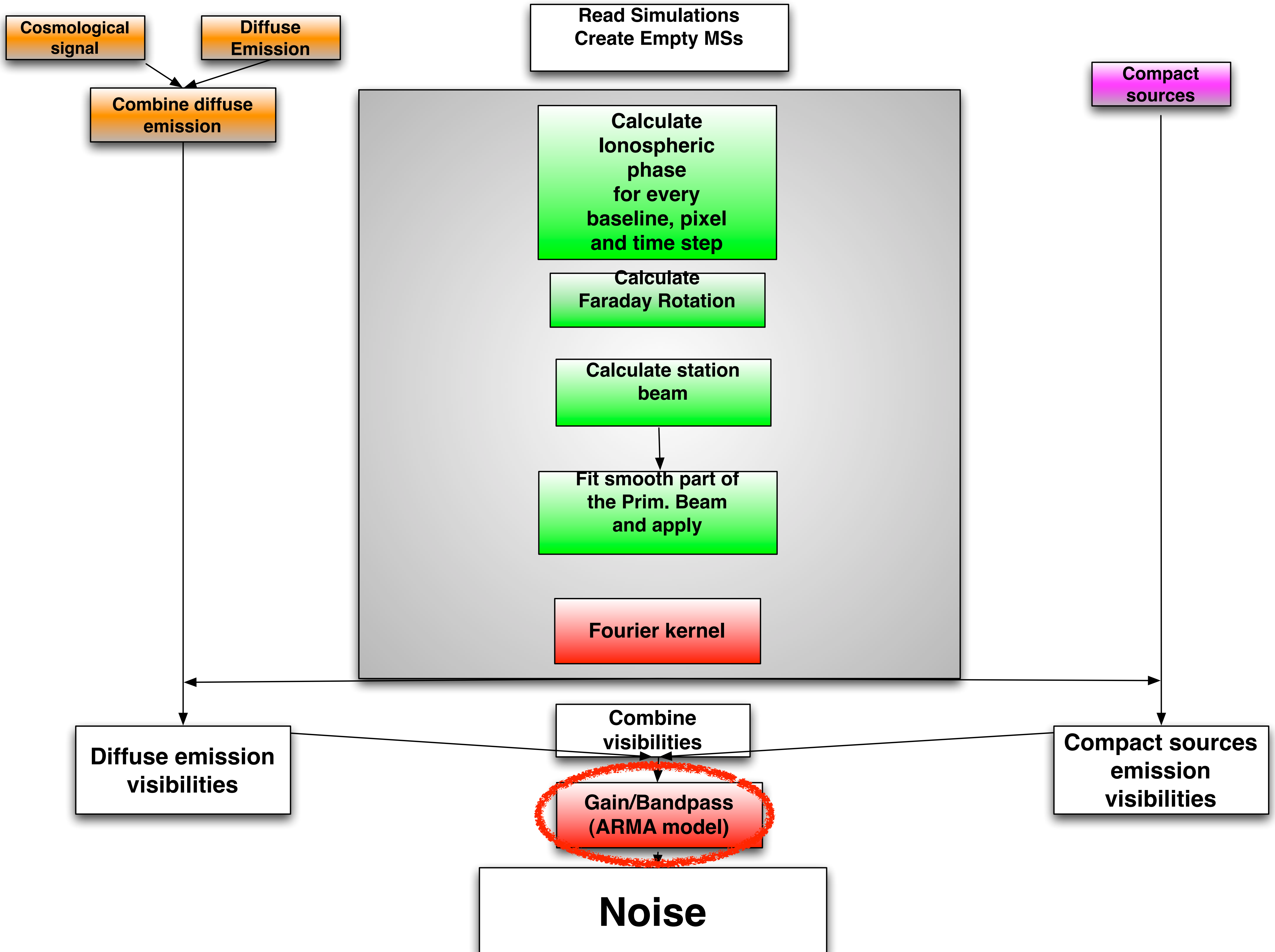
Initially non-polarized sources



End To End Pipeline: Prediction



End To End Pipeline: Prediction



Gain modeling

- **A**uto **R**egressive **M**oving **A**verage are used to describe stationary time series
- Can be generated by passing white noise through a recursive (AR) and non recursive (MA) filter.
- ARMA is appropriate when a system is a function of a series of unobserved shocks (the MA part, ie temperature fluctuations) as well as its own behavior (i.e. clock jitter).

$$X_t = c + \sum_{i=1}^{p} \varphi_i X_{t-i} + \varepsilon_t.$$

$$X_t = \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}$$

$$X_t = c + \varepsilon_t + \sum_{i=1}^{p} \varphi_i X_{t-i} + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i}.$$

(ARMA)

- Likelihood

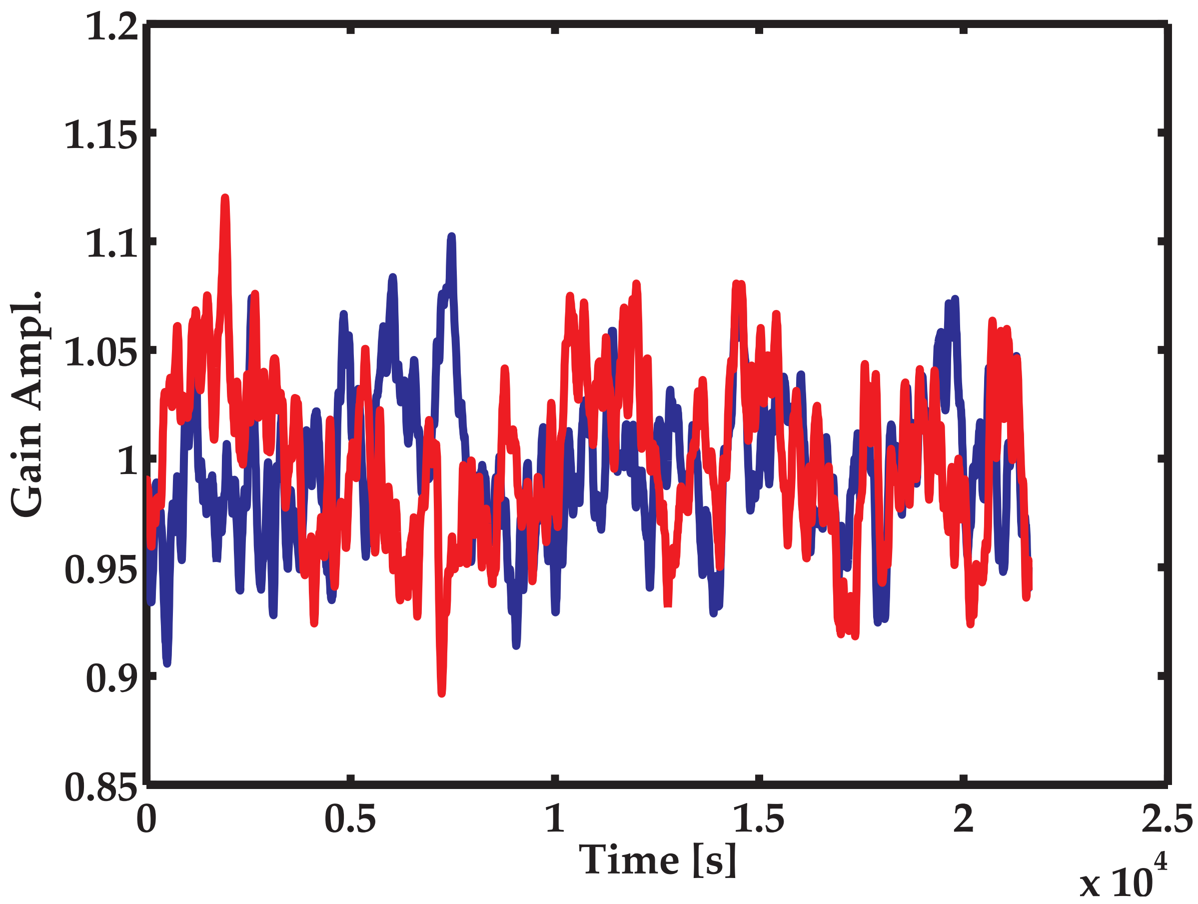
$$L(\mu, \phi, \sigma_w^2) = (2\pi\sigma_w^2)^{-n/2}\sqrt{1-\phi^2}\exp\left[-\frac{S(\mu,\phi)}{2\sigma_w^2}\right],$$

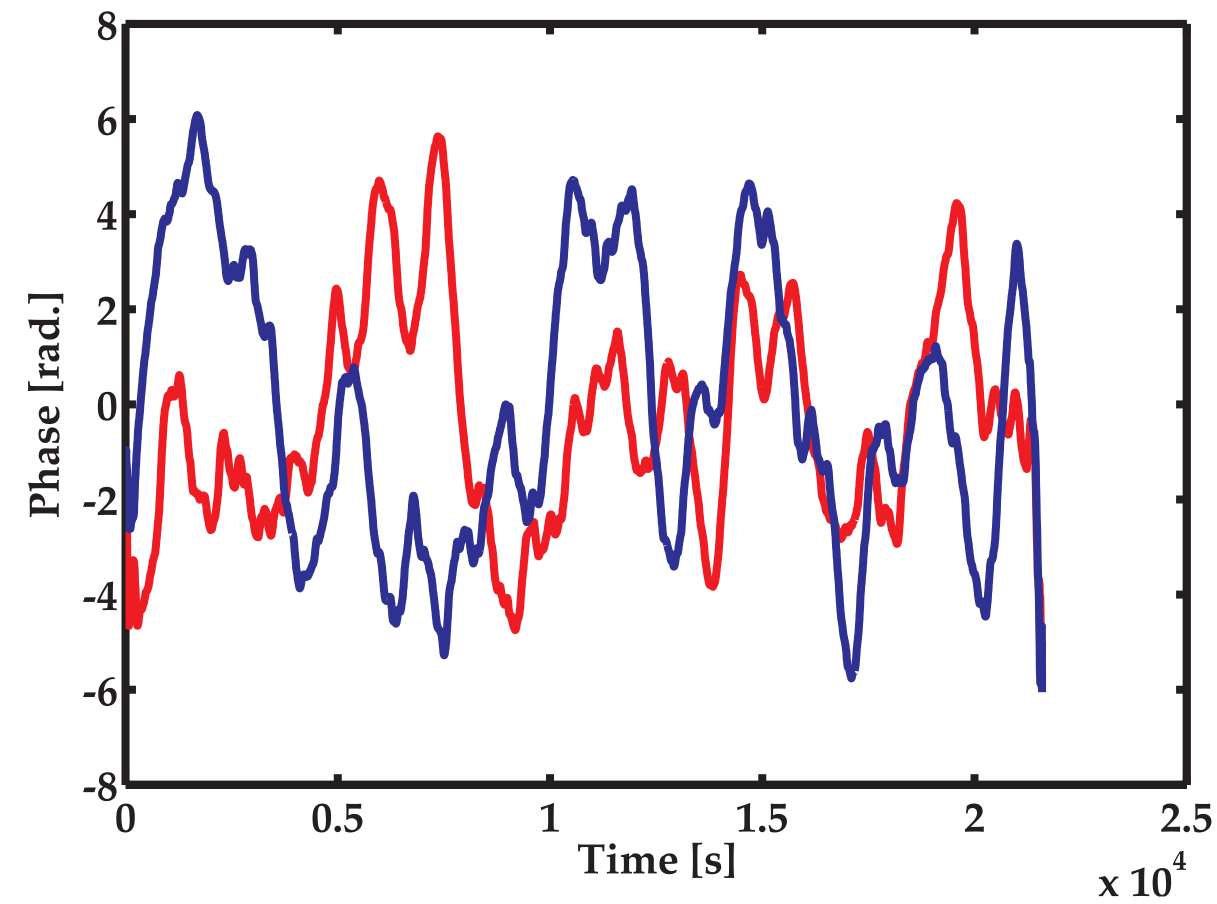
where

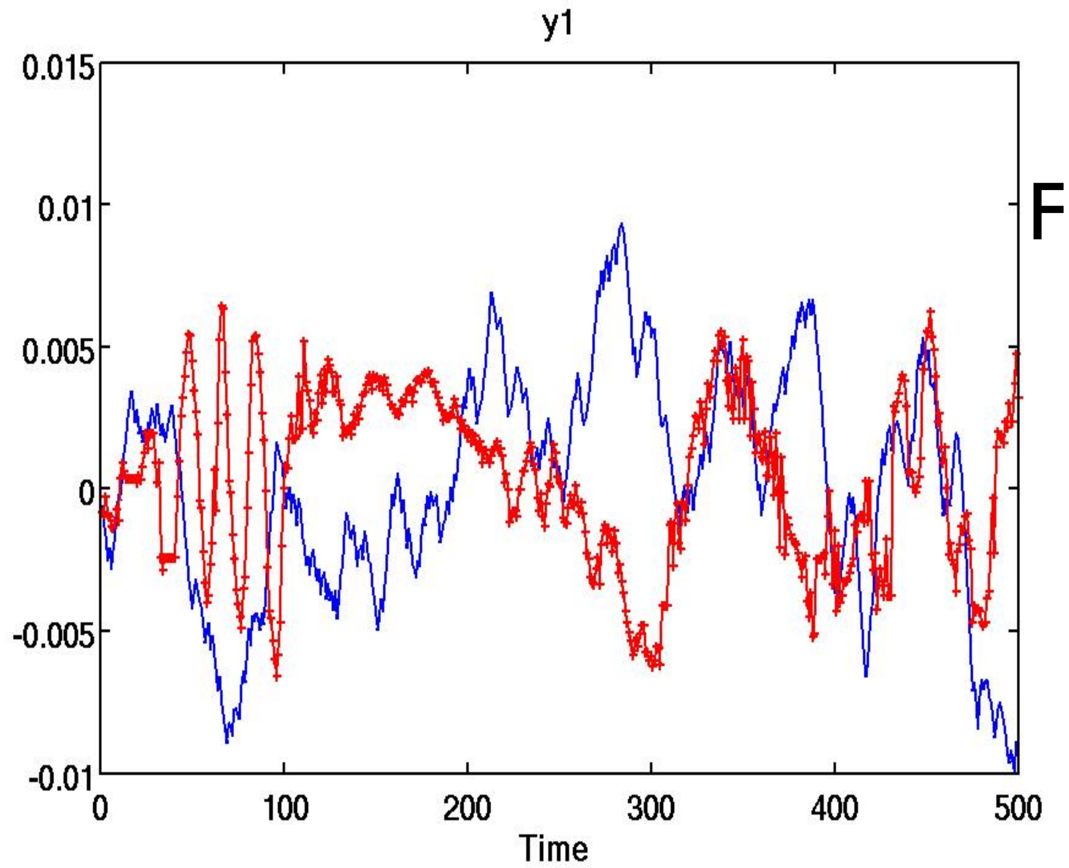
$$S(\mu,\phi) = (1-\phi^2)(x_1-\mu)^2 + \sum_{t=2}^{n}\left[(x_t-\mu) - \phi(x_{t-1}-\mu)\right]^2.$$

- (un)conditional least squares
- State space methods
- Brure force – use observed cov. Instead of theoretical

$$\begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} \gamma_0 & \gamma_{-1} & \gamma_{-2} & \cdots \\ \gamma_1 & \gamma_0 & \gamma_{-1} & \cdots \\ \gamma_2 & \gamma_1 & \gamma_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \varphi_3 \\ \vdots \end{bmatrix}$$







Real data:
Flagged areas + residual ionospheric errors

Model selection

- Aikake's final prediction error provides a measure of model quality in the situation where differenet datasets are available.

$$FPE = V\left(\frac{1+d/N}{1-d/N}\right)$$

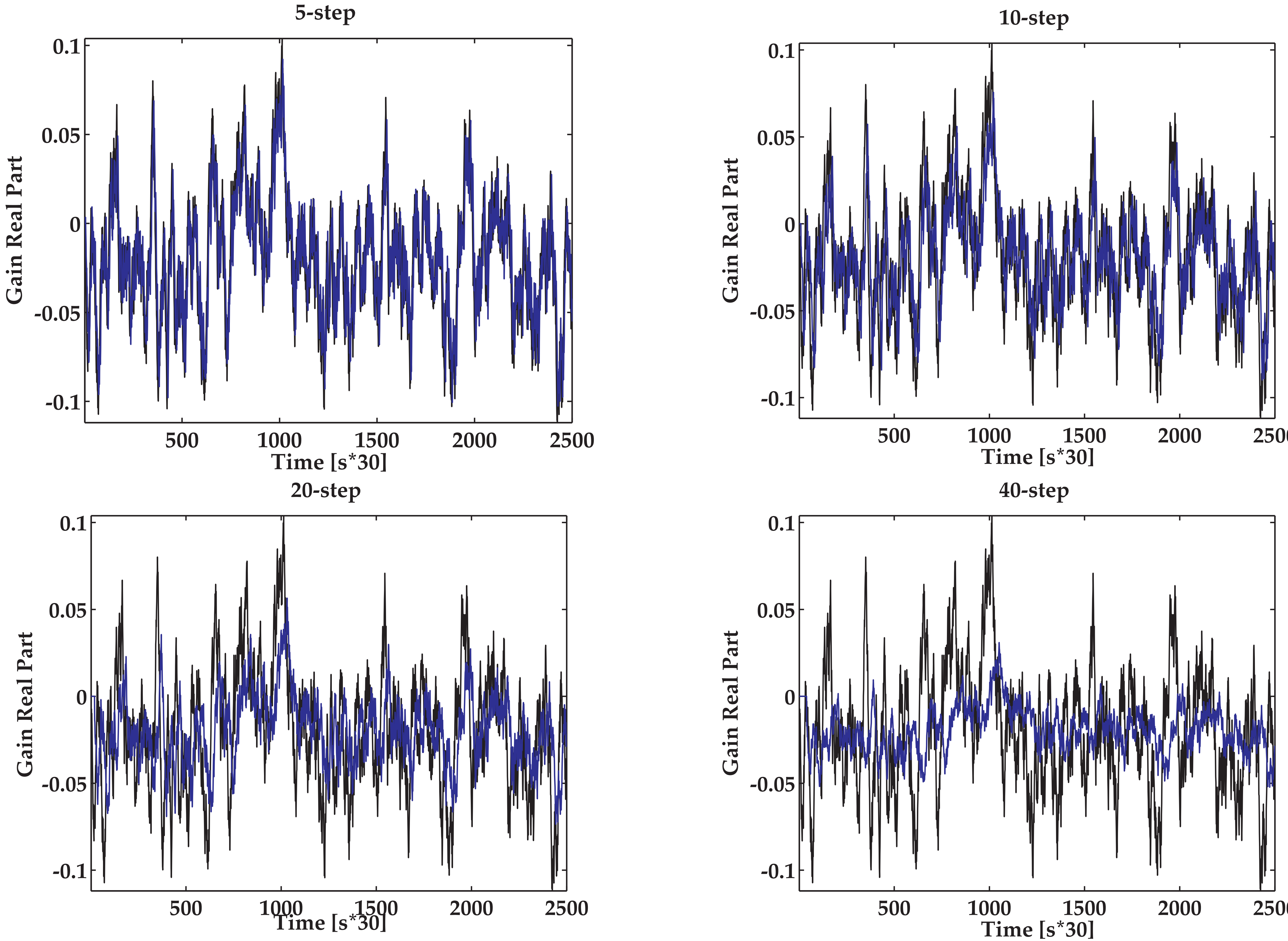
$$V = \det\left(\frac{1}{N}\sum_{1}^{N}\varepsilon(t,\theta_N)\left(\varepsilon(t,\theta_N)\right)^T\right)$$

V is the loss function: a function that maps an outcome into a real number representing the cost of that outcome. In general it depends on the difference between the true or desired value.

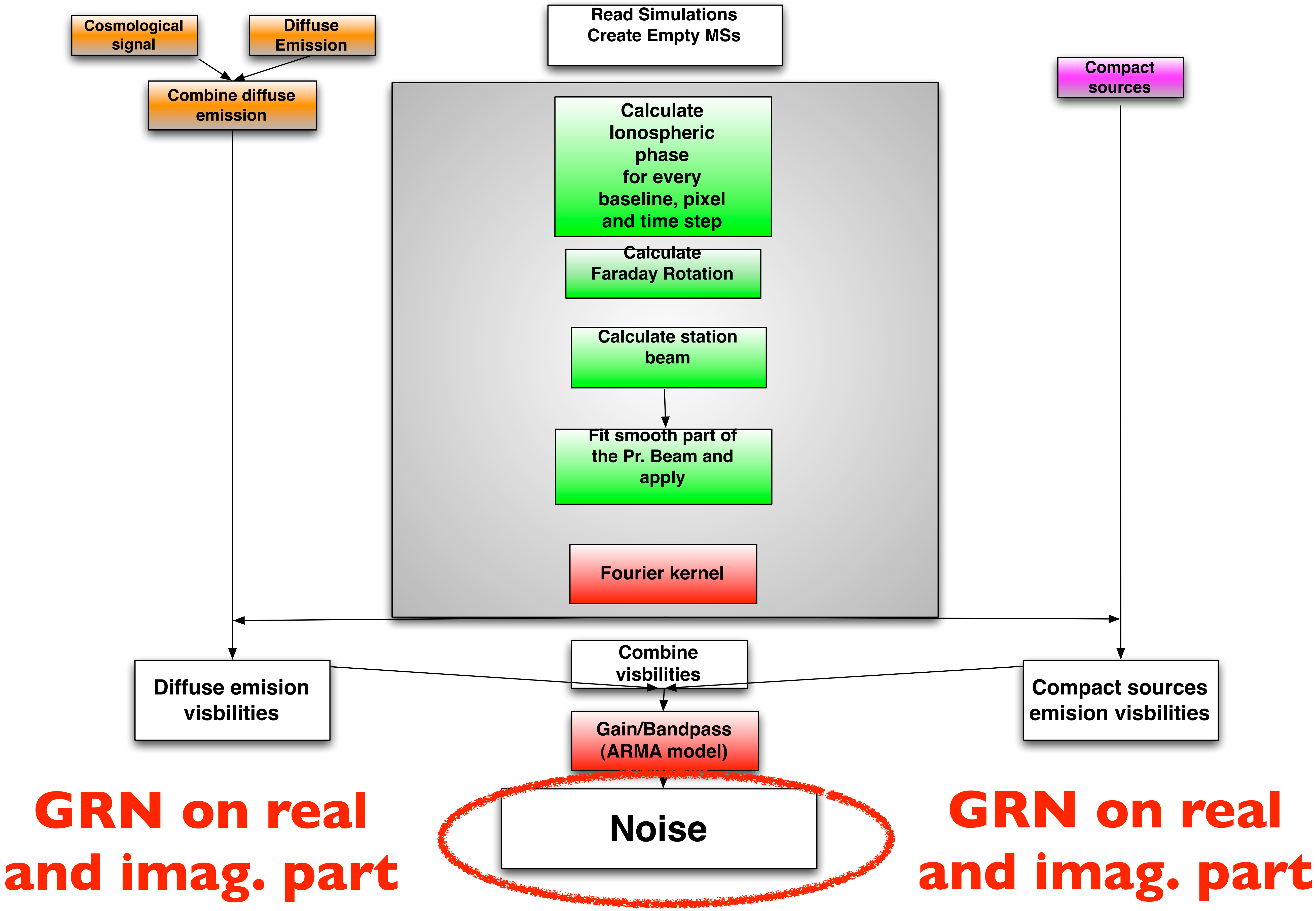
Frequentist: Calculate the expected value w.r.t. The PDF of the observed data

$$V = \det\left(\frac{1}{N}\sum_{1}^{N}\varepsilon(t,\theta_N)\left(\varepsilon(t,\theta_N)\right)^T\right)$$

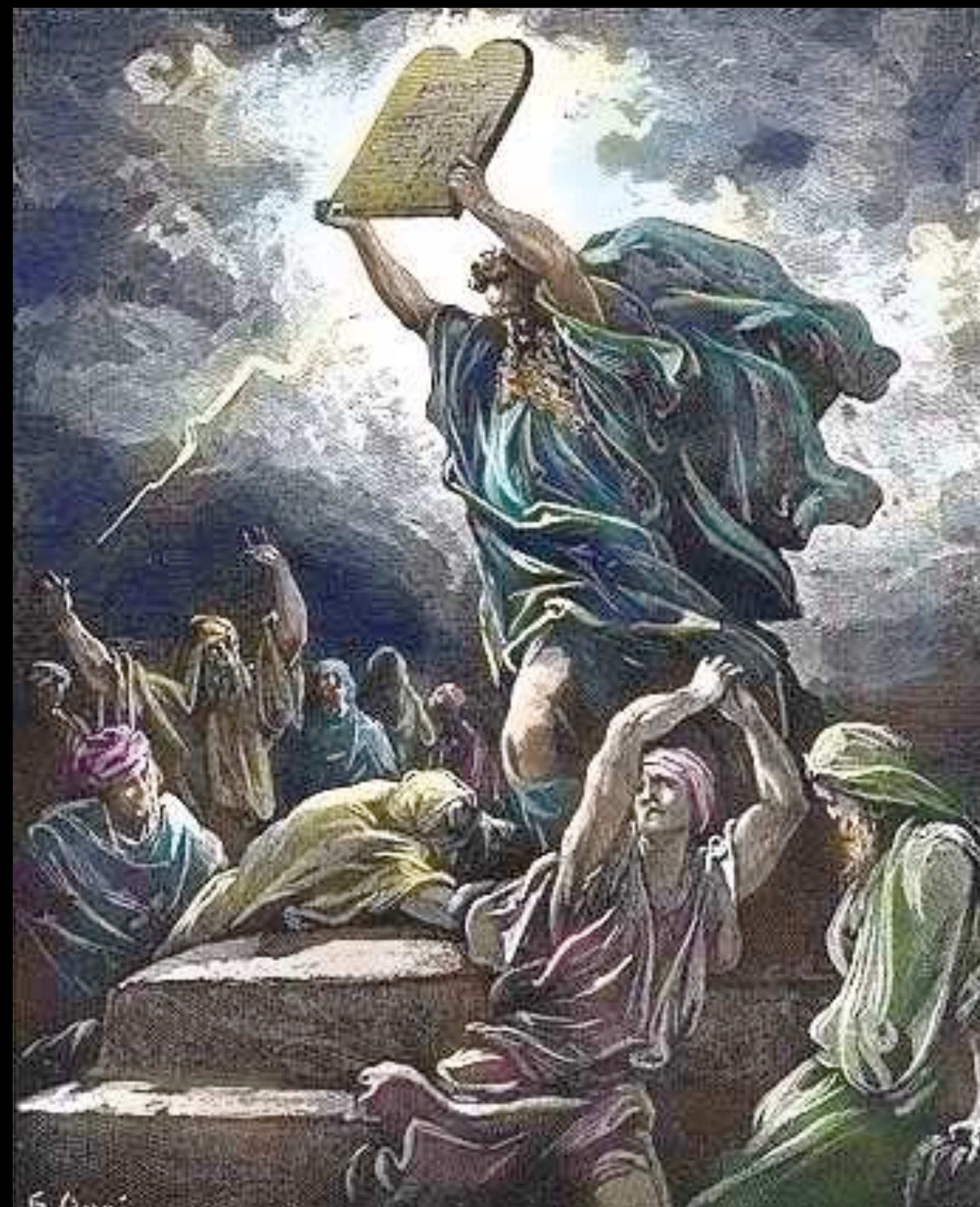
Prediction: 2.5, 5, 10 and 20 minutes



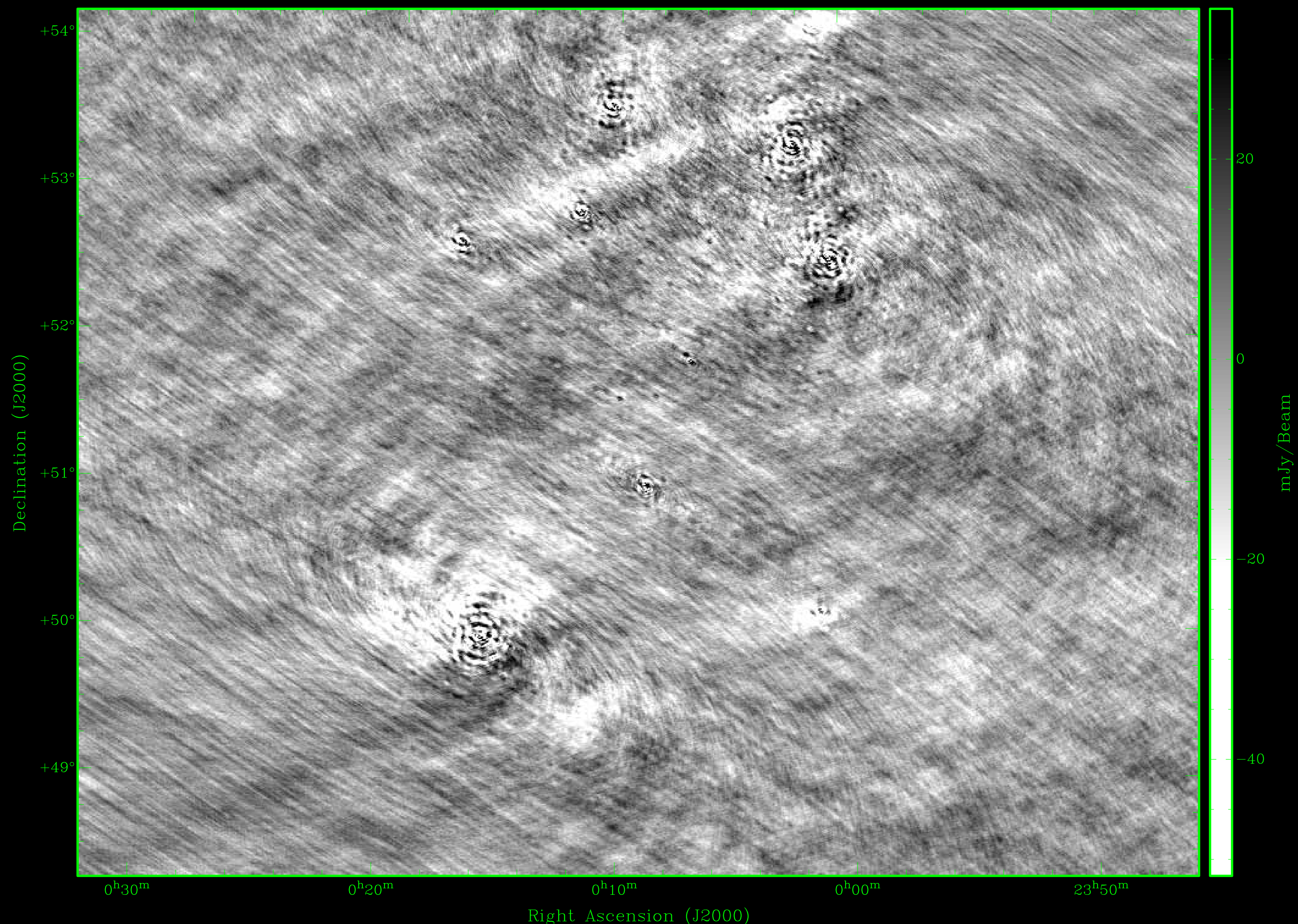
End To End Pipeline: Prediction

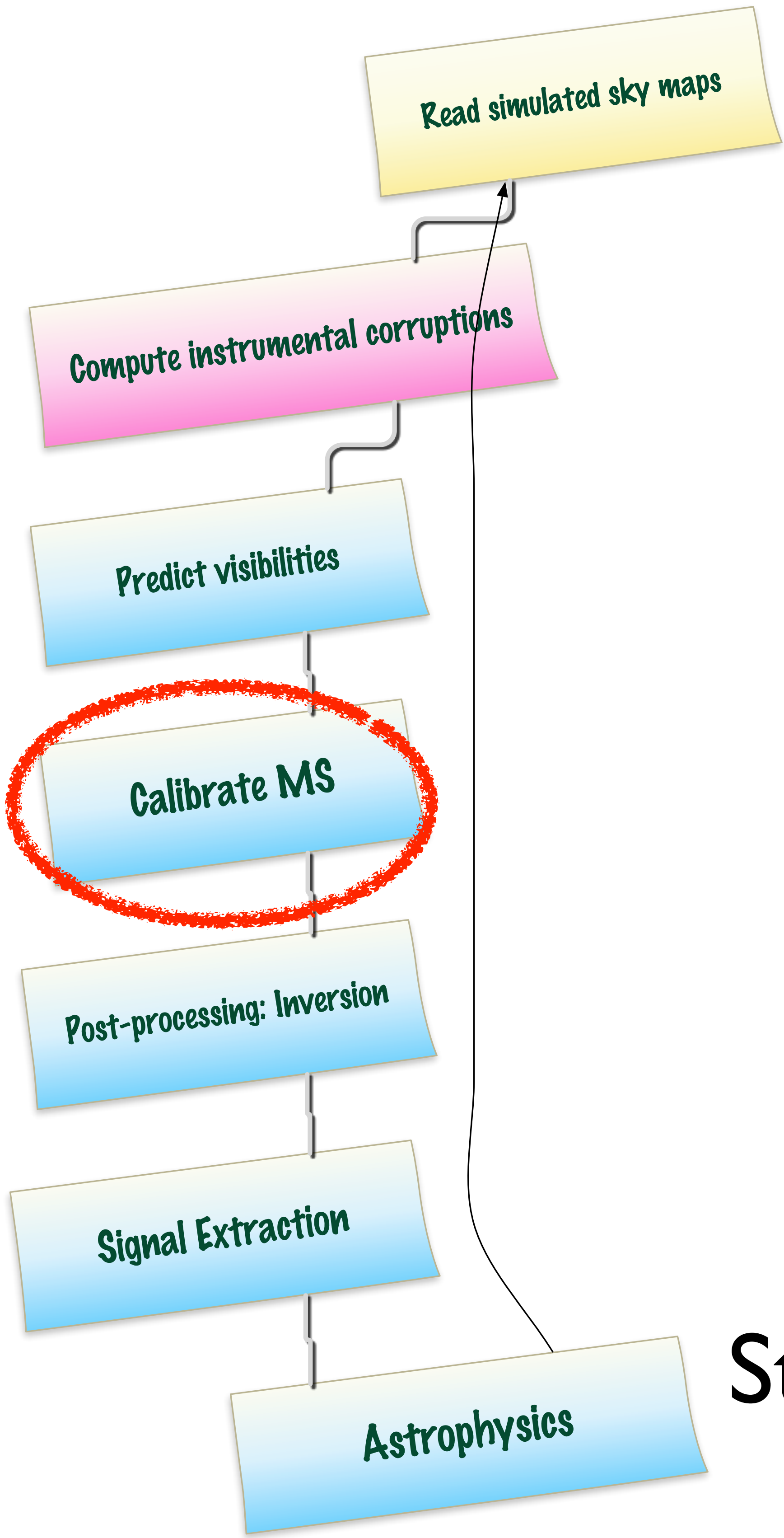


CASA Tables



The result


+54°
+53°
+52°
+51°
+50°
+49°
Declination (J2000)
0h30m
0h20m
0h10m
0h00m
23h50m
Right Ascension (J2000)
20
0
−20
−40
mJy/Beam



Step 1: Sky realizations

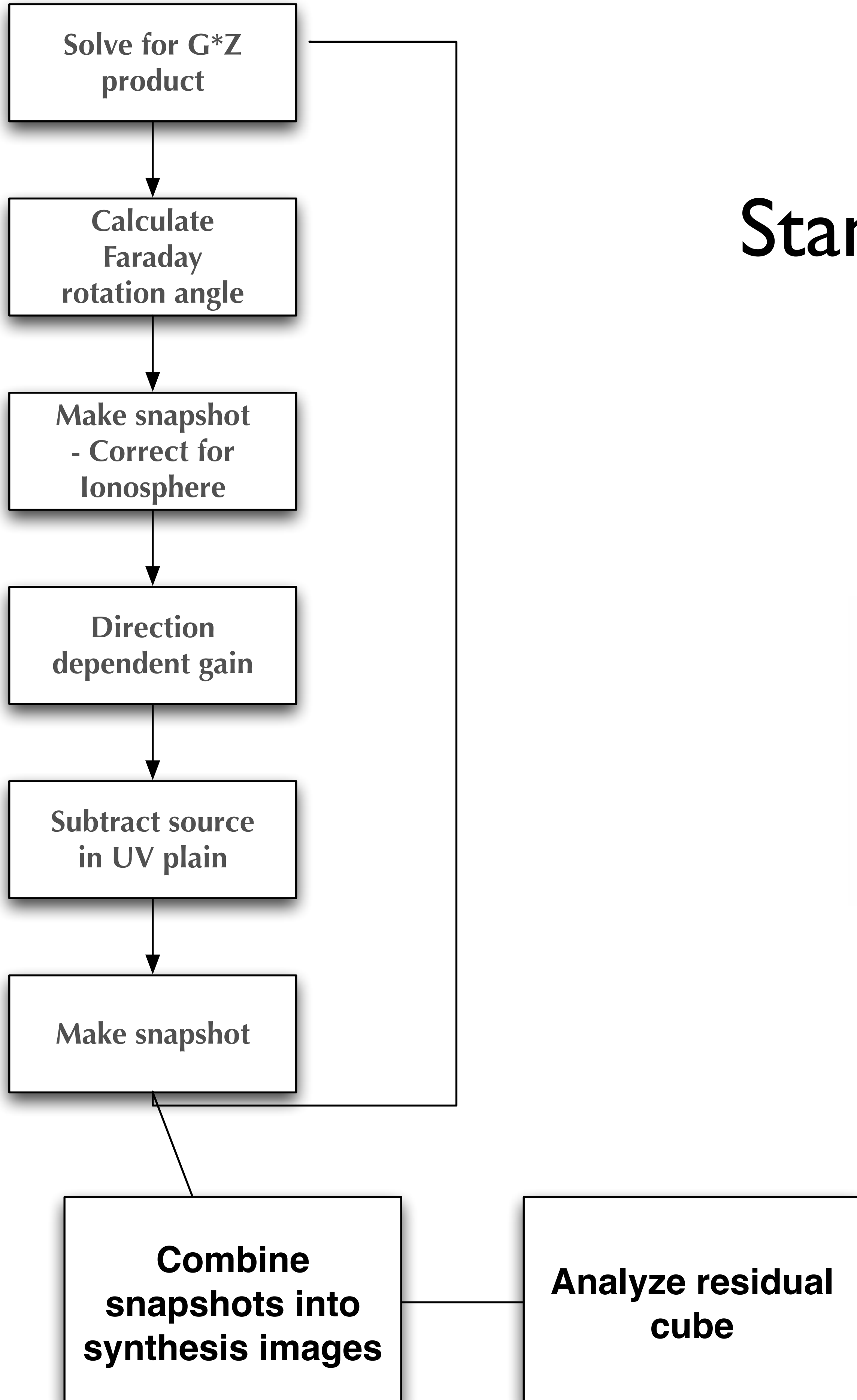
Step 2: Predict Visibilities

Step 3: Self-cal

Step 4: Inversion

Step 5: Extraction and interpretation

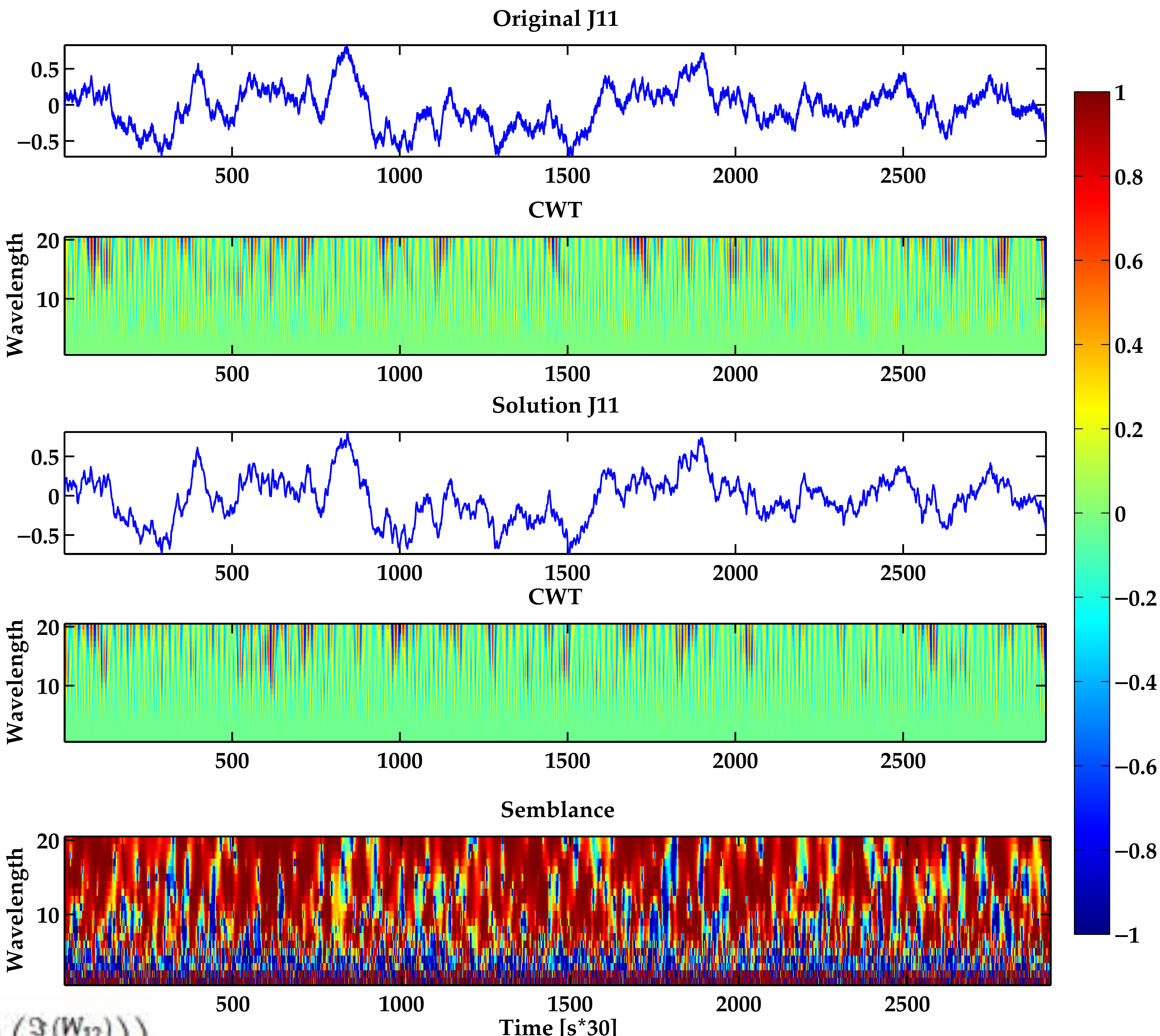
Standard calibration pipeline



Standardized Moments of Diffrences		
Error Level [%]	0.1	1
Mean	6.9×10^{-4}	6.7×10^{-2}
Variance	3.3×10^{-3}	0.09
Skewness	7.0×10^{-2}	0.38
Kurtosis	2.94	3.07

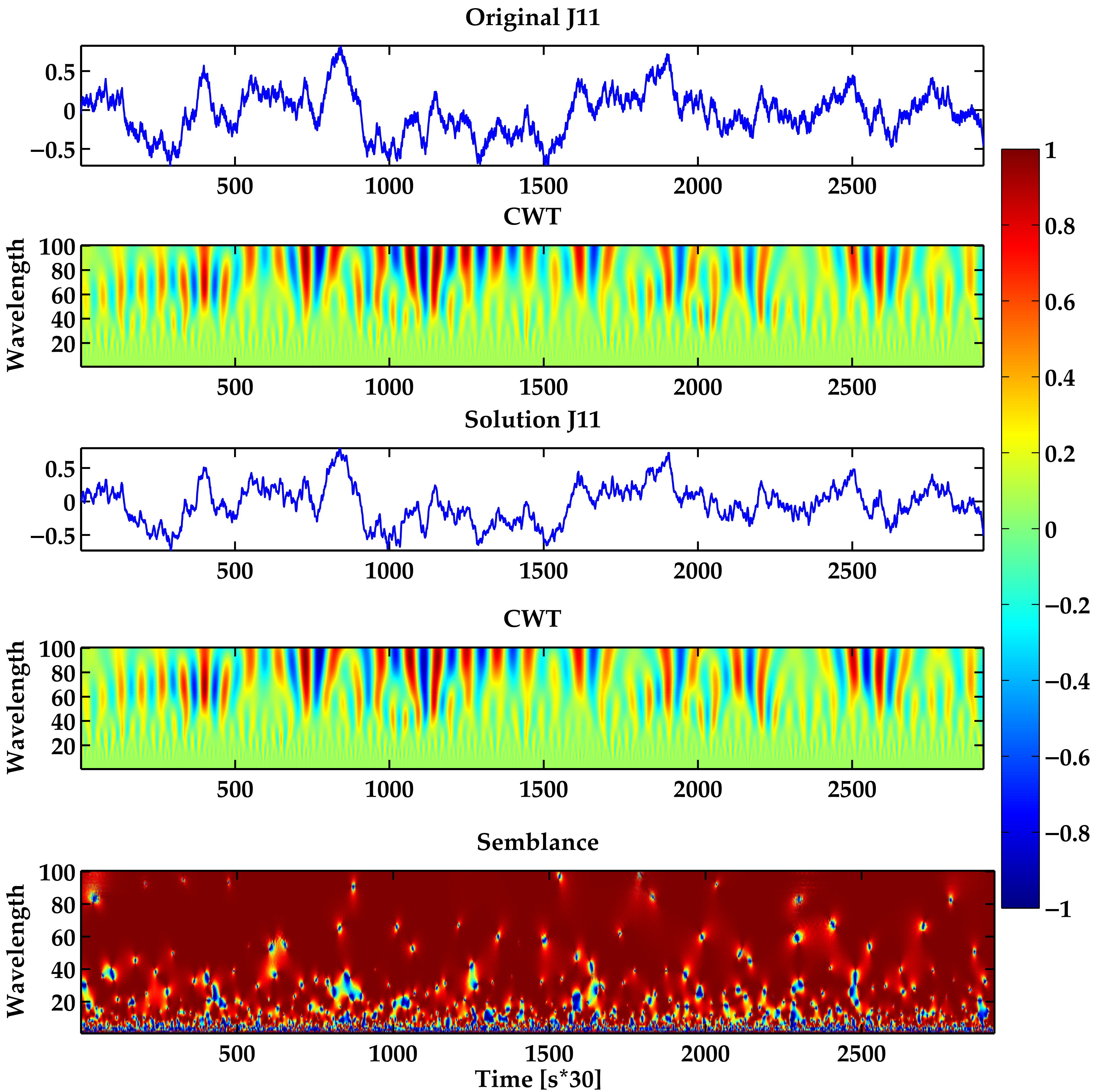
Comparison between original Jones matrix and solution

0.1%
Error
Level

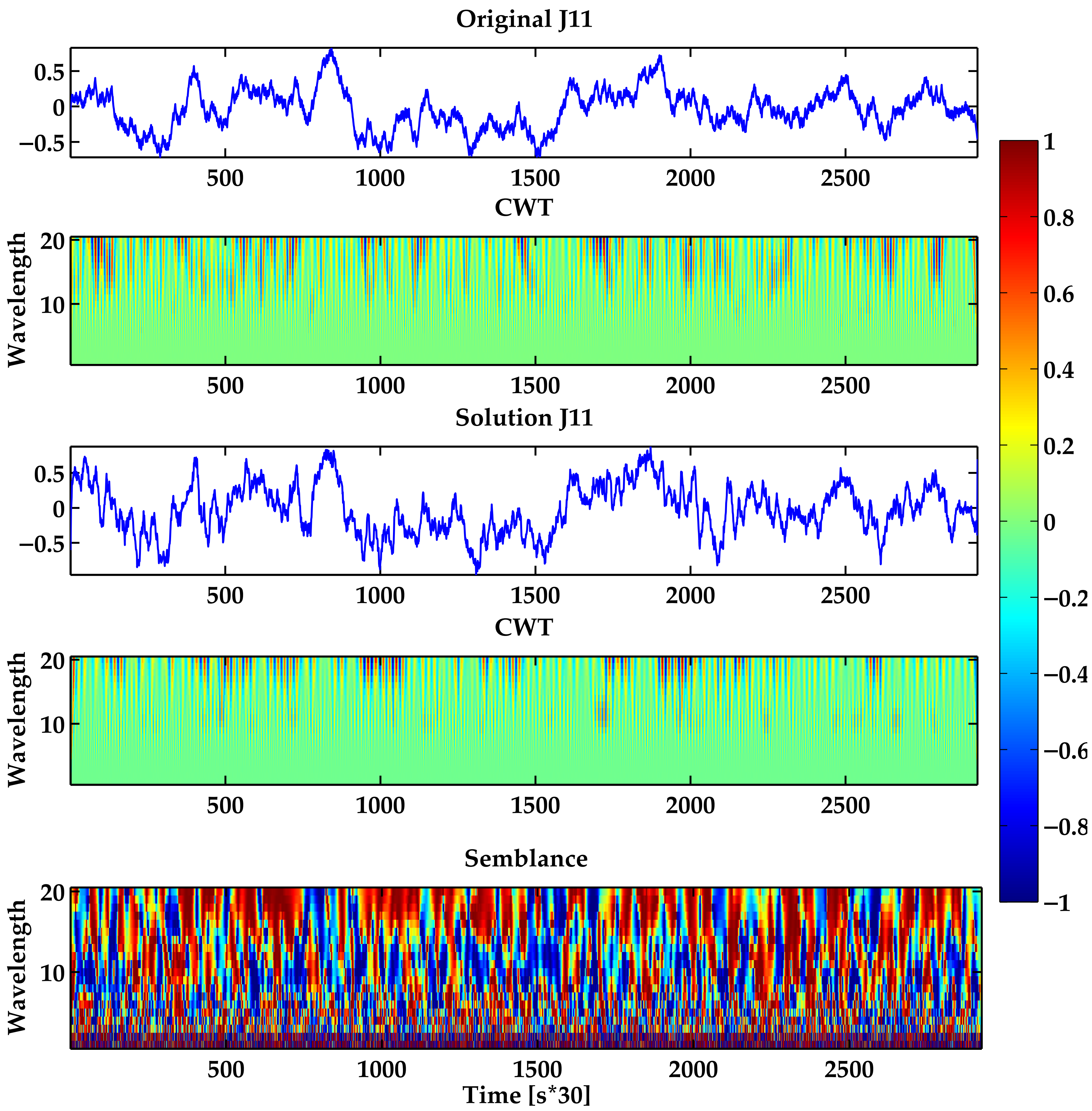


$$S = \cos^n\left(\tan^{-1}\left(\frac{\Im(W_{12})}{\Re(W_{12})}\right)\right)$$

0.1% Error Level

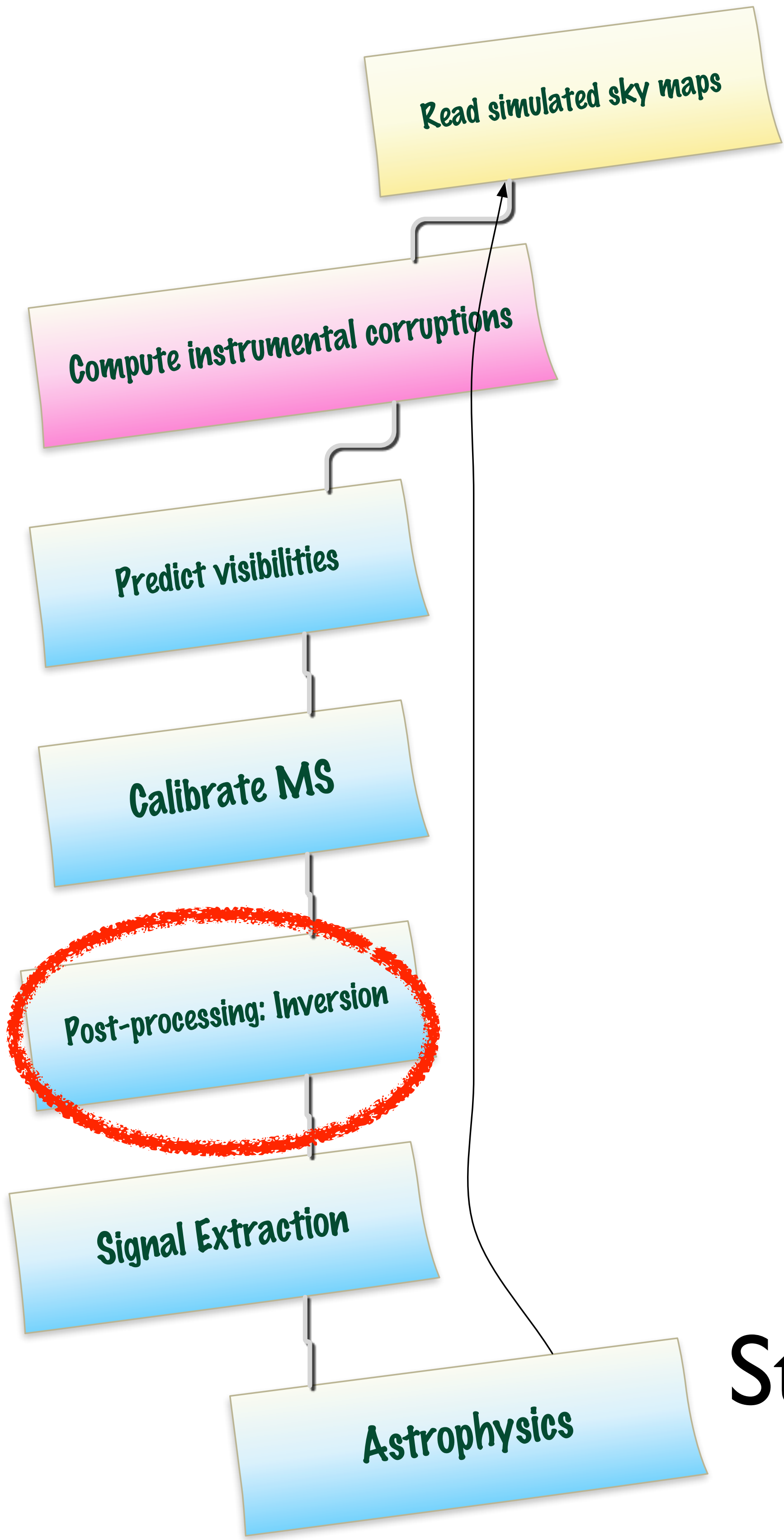


1% Error Level



Things to be done

- Include RFI effects in the simulation
- Include a more complex GSM
- Effects of bandwidth smearing
- Longer baselines
- Murual-coupling



Step 1: Sky realizations

Step 2: Predict Visibilities (Part I)

Step 3: Self-cal

Step 4: Inversion (Part II)

Step 5: Extraction and interpretation

The data model

$$\mathbf{v} = \mathbf{A}(\mathbf{p})\mathbf{s} + \mathbf{n} \qquad \mathbf{s} = \begin{pmatrix} \mathbf{s}_{\mathrm{GSM}} \\ \mathbf{s}_{\mathrm{LSM}} \\ \mathbf{s}_{\mathbf{grid}} \end{pmatrix}$$

- Each visibility is a linear combination of the sky components
- Selfcal solves for s through CLEANing and p through a calibration step (usually cLLS)

(see also work of Boonstra, Wijnholds, Leshem)

We can write the data model as:

$$\mathbf{V}_k = \left(\bar{\mathbf{A}}_k \circ \mathbf{A}_k\right) vecdiag(\mathbf{b})$$

$$A_+ = \begin{bmatrix} \bar{\mathbf{A}}_1 \circ \mathbf{A}_1 \\ \vdots \\ \bar{\mathbf{A}}_k \circ \mathbf{A}_k \end{bmatrix}$$

The the map is given by:

$$\mathbf{b} = \mathbf{A}_+^{PI}\mathbf{v}_+ = \left(\mathbf{A}_{+,j}^{\dagger}\mathbf{A}_{+,i}\right)^{-1}\mathbf{A}_{+,j}^{\dagger} vec(v_k)$$

The factor $\mathbf{A}_{+,j}^{\dagger} vec(v_k)$ corresponds to the dirty map and the matrix $\left(\mathbf{A}_{+,j}^{\dagger}\mathbf{A}_{+,i}\right)^{-1}$ is the deconvolution step.

This approach will lead to a ML solution but is computationally very expensive. CLEAN is less optimal but faster.

$$\mathbf{b} = \left(\mathbf{A}_{+,j}^{\dagger}\mathbf{C}^{-1}\mathbf{A}_{+,i}\right)^{-1}\mathbf{A}_{+,j}^{\dagger}\mathbf{C}^{-1} vec(v_k)$$

$$\mathbf{b} = \left(\mathbf{A}^{\dagger}_{+,j}\mathbf{C}^{-1}\mathbf{A}_{+,i}\right)^{-1}\mathbf{A}^{\dagger}_{+,j}\mathbf{C}^{-1}vec(v_k)$$

Counting (beam-forming)

Projection of data into pixels

Using priors (also obtained by Wiener filtering, Zaroubi et al., 1995 :

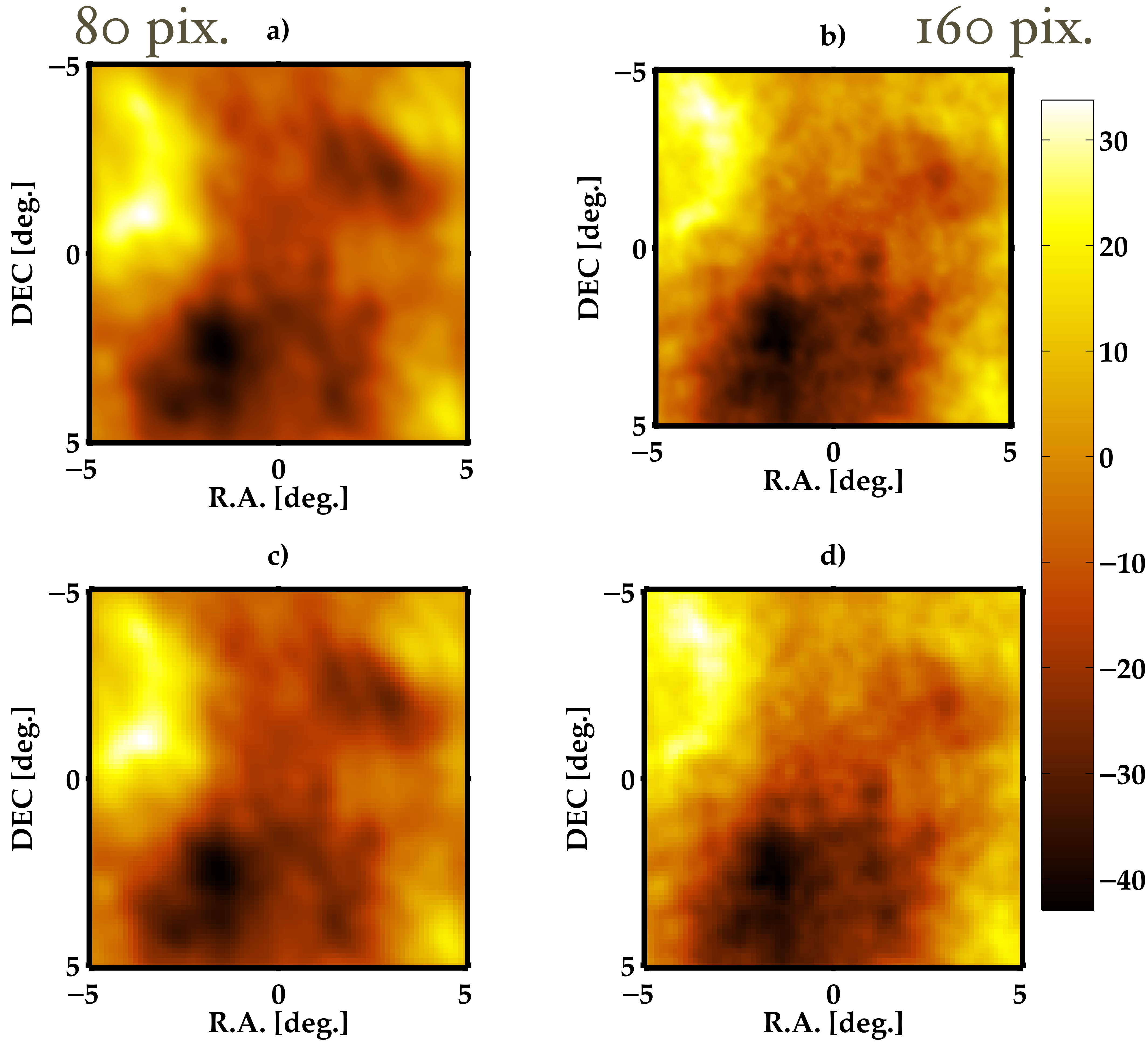
$$\mathbf{b} = \left(\mathbf{C}_S^{-1} + \mathbf{A}^{\dagger}_{+,j}\mathbf{C}^{-1}\mathbf{A}_{+,i}\right)^{-1}\mathbf{A}^{\dagger}_{+,j}\mathbf{C}^{-1}vec(v_k)$$

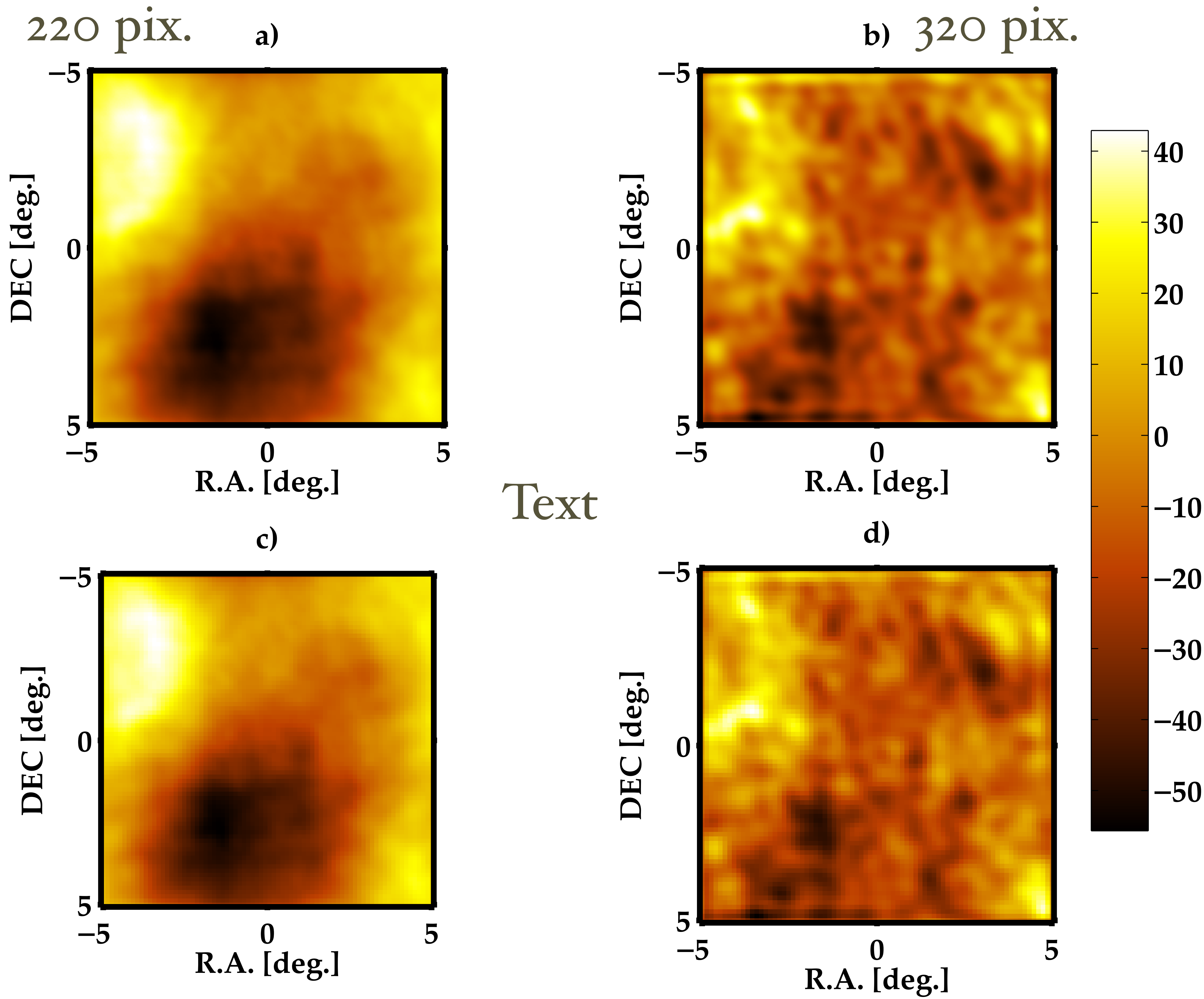
Deconvolution condition number/ Regularization

- The deconvolution matrix gives an estimate of the redistribution of noise on the image plain.
- It also measures how well-posed the inverse problem is.
- For the LOFAR EoR KSP, the image plain is oversampled by a factor 2-4.
- Regularization method as well as strength affect the final result.

Regularization

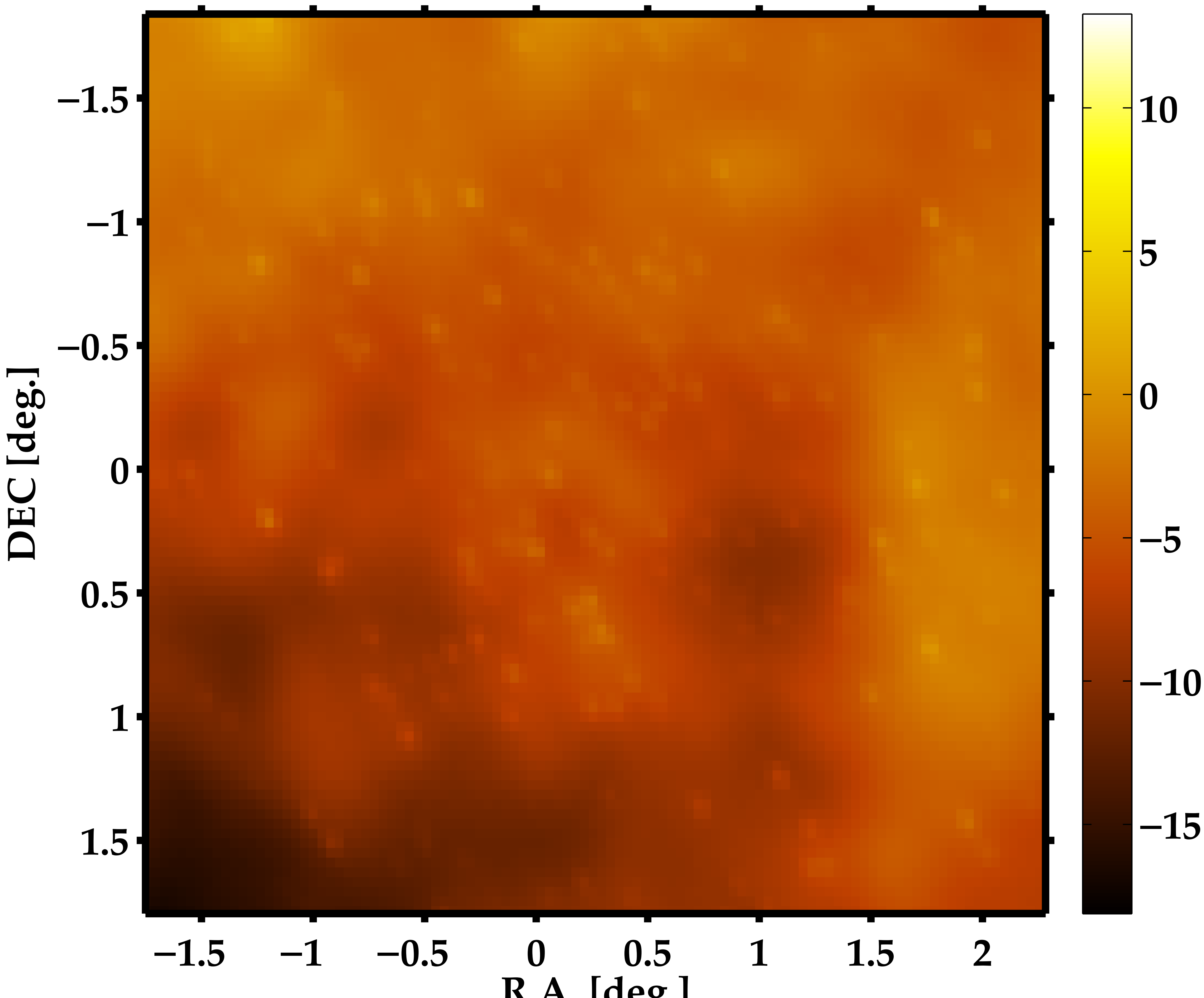
- Two regularization methods used: Tikhonov and local diffusion operators (Vogel & Oman, Fatami et al., Labropoulos et al., in prep.)
- Tikhonov is fast to implement but choice of reg. parameter is difficult.
- Diffusion method needs and iterative approach.


80 pix.
160 pix.
a)
b)
c)
d)
DEC [deg.]
R.A. [deg.]
−5
0
5
30
20
10
0
−10
−20
−30
−40



220 pix.
a)
b)
320 pix.
Text
c)
d)
DEC [deg.]
R.A. [deg.]
−5
0
5
40
30
20
10
0
−10
−20
−30
−40
−50

Regularization error around map peaks



Inversion

- If A has full rank then the Gauss-Markov gives the best linear unbiased estimator given by:

$$(\mathbf{A}^T\mathbf{C}_N^{-1}\mathbf{A}_N + \lambda\mathbf{R}^T\mathbf{R})\mathbf{s} = \mathbf{A}^T\mathbf{C}_N^{-1}\mathbf{v}$$

- Remark: In the case where C= σ^2 I then this is equivalent to deterministic weighted LS problem:

$$\min_{\mathbf{p}\in\mathbb{R}^n} \|\mathbf{A}s - \mathbf{v}\|_{C_N^{-1}}$$

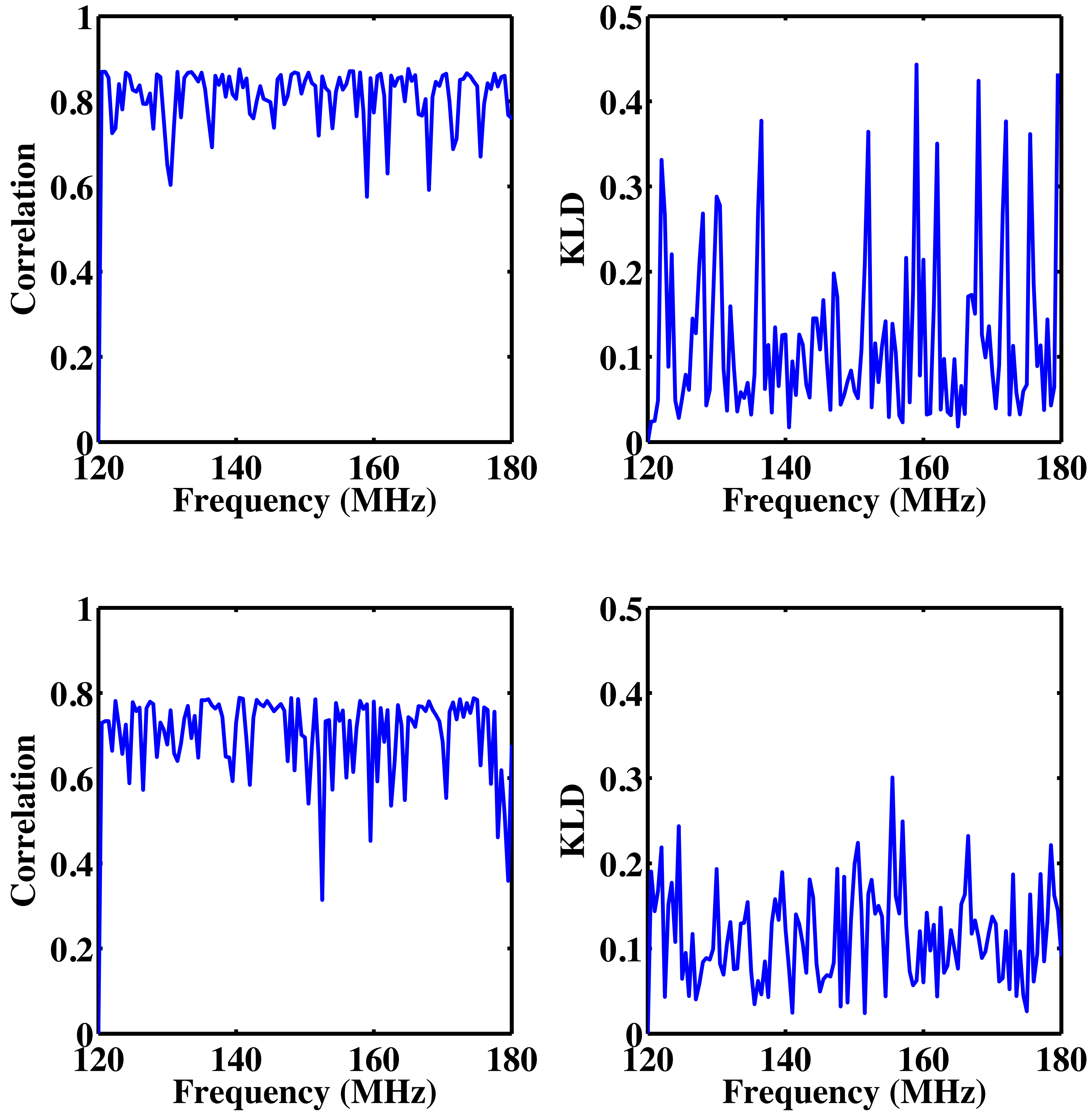
Inversion

- The previous estimator is unsuitable for the solution of ill-conditioned systems

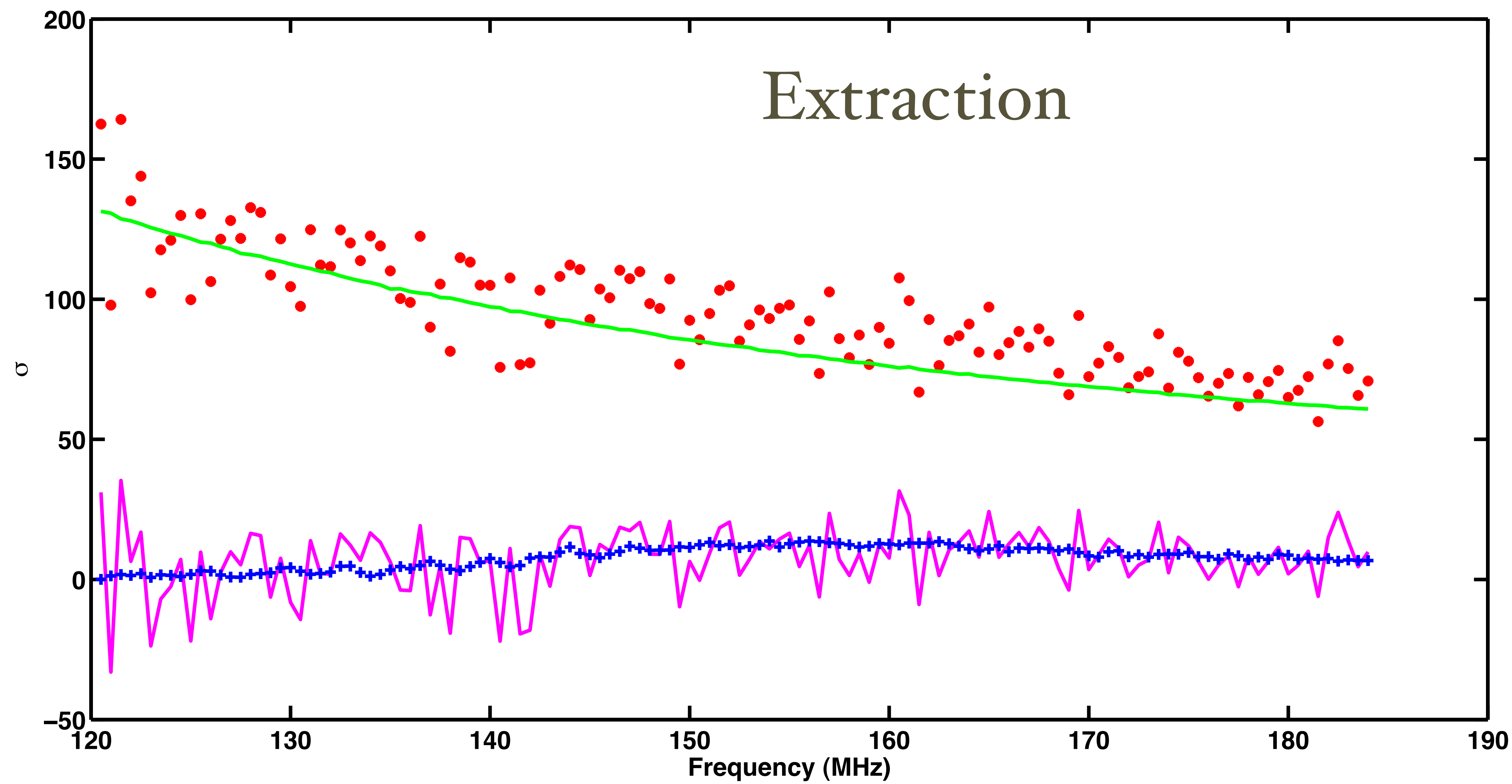
- In that case the best estimator is the Min. Variance estim.:

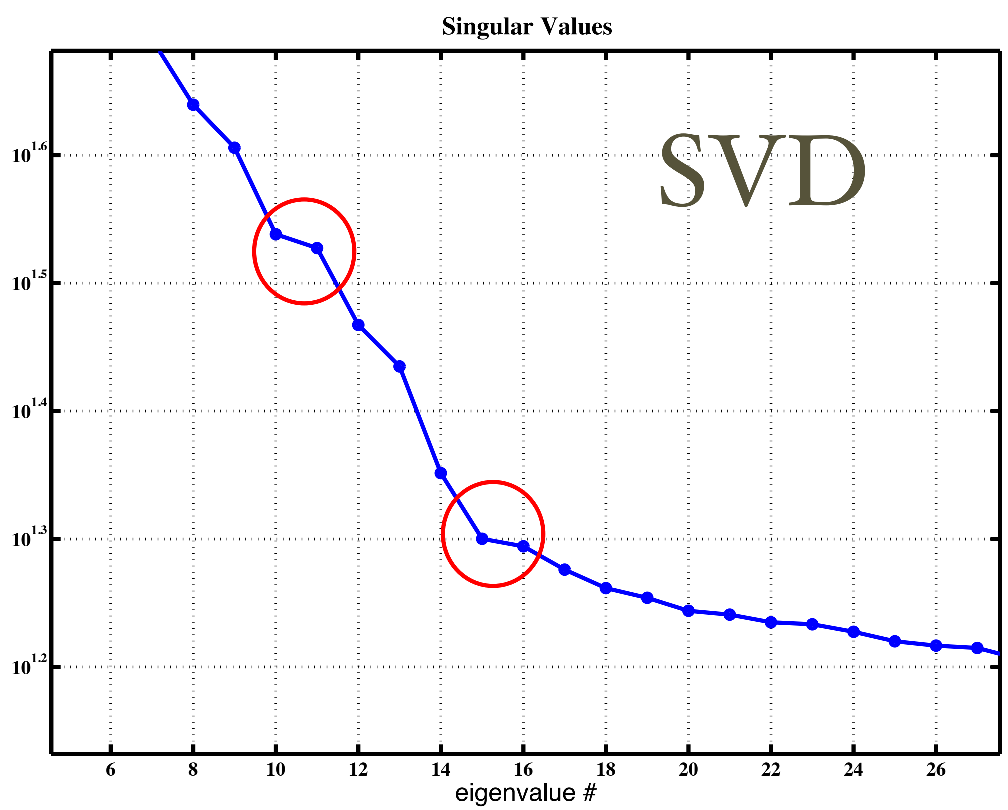
$$\mathbf{C}_{\mathbf{VV}}\mathbf{A}^{T}\left(\mathbf{A}\mathbf{C}_{\mathbf{VV}}^{-1}\mathbf{A}^{T}+\mathbf{C}_{N}\right)^{-1}\mathbf{s}=\left(\mathbf{A}^{T}\mathbf{C}_{N}^{-1}\mathbf{A}+\mathbf{C}_{\mathbf{VV}}^{-1}\right)^{-1}\mathbf{A}^{T}\mathbf{C}_{N}^{-1}\mathbf{s}$$

- When the noise and th visibilities are independent this is the MAP estimator



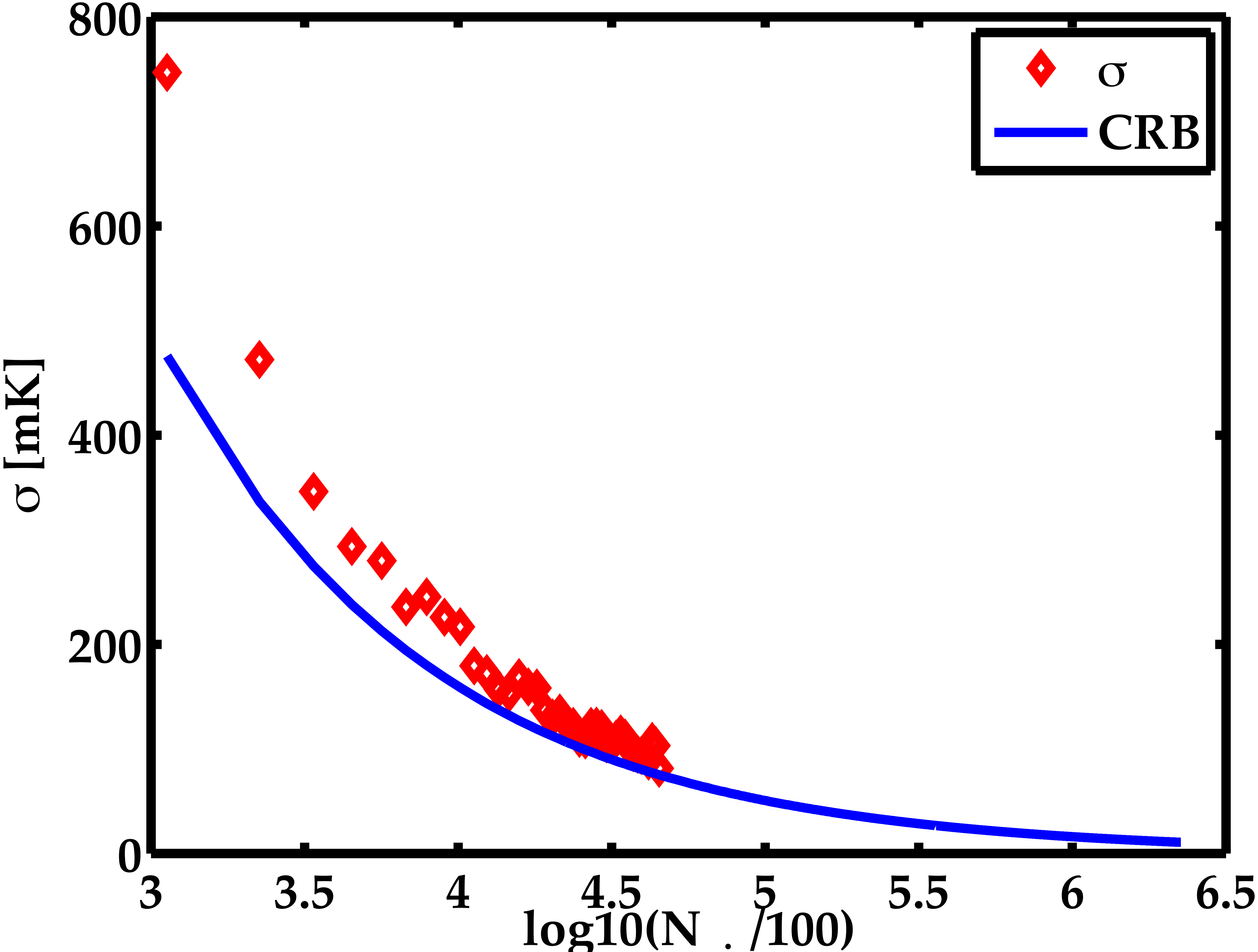
Correlation
Frequency (MHz)
KLD
Frequency (MHz)
Correlation
Frequency (MHz)
KLD
Frequency (MHz)


Extraction
σ
200
150
100
50
0
−50
120
130
140
150
160
170
180
190
Frequency (MHz)

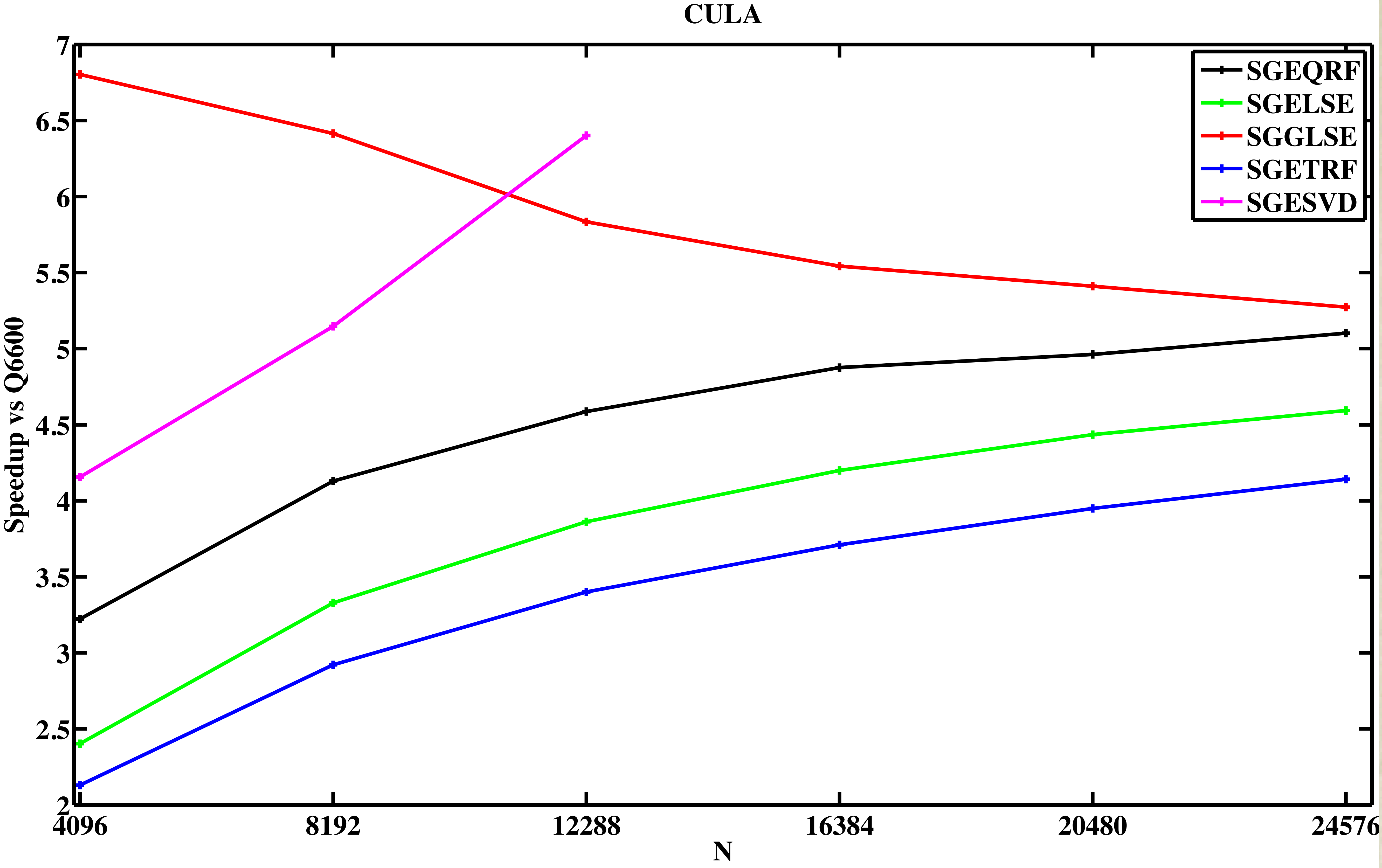

Singular Values
SVD
$10^{1.6}$
$10^{1.5}$
$10^{1.4}$
$10^{1.3}$
$10^{1.2}$
6
8
10
12
14
16
18
20
22
24
26
eigenvalue #

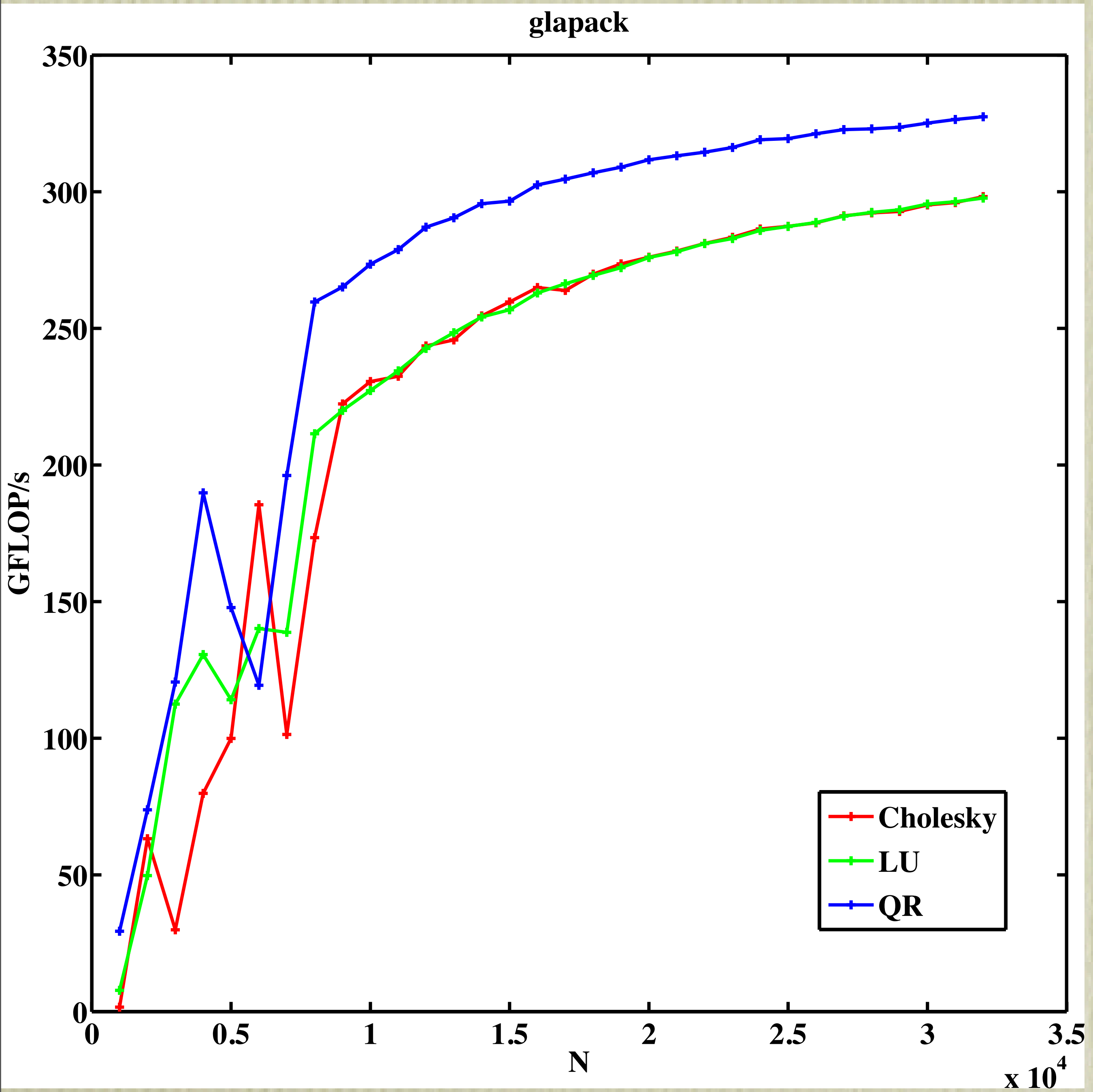

10^{-16}
10^{-12}
10^{-8}
10^{-4}
10^{0}
10^{-16}
10^{-12}
10^{-8}
10^{-4}
10^{0}

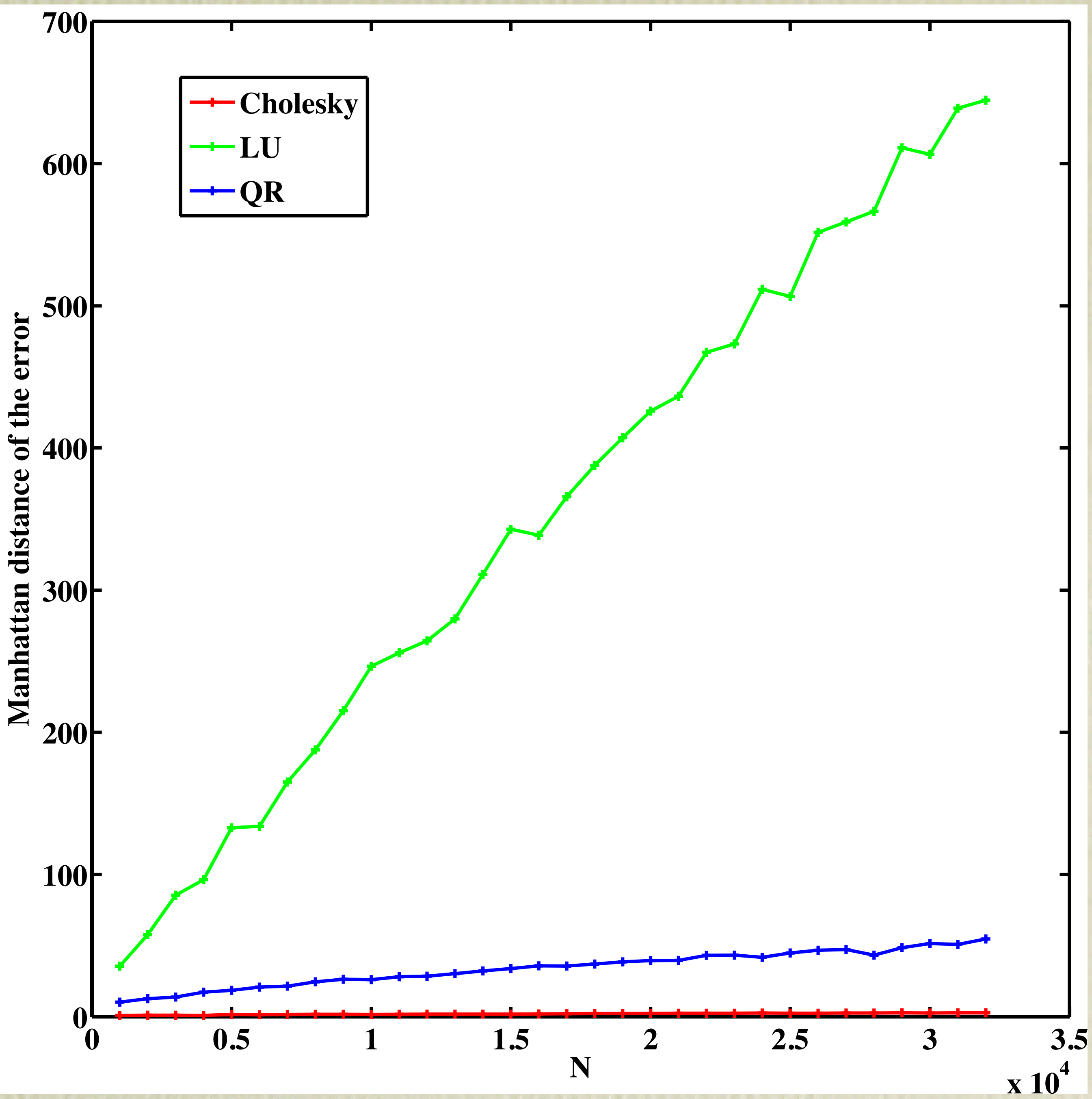
Cramer-Rao Bound



GP-GPU Benchmarking




glapack
GFLOP/s
N
x 10^4
Cholesky
LU
QR
0
50
100
150
200
250
300
350
0
0.5
1
1.5
2
2.5
3
3.5


Manhattan distance of the error
N
x 10^4
Cholesky
LU
QR
0
100
200
300
400
500
600
700
0
0.5
1
1.5
2
2.5
3
3.5

Conclusions:

- Simulations can help us understand interferometric observations but...
- The solutions are degenerate. The full parameter space might be explored
- Time-series analysis can model the underlying dynamics of a system
- ML imaging is optimal and feasible for EoR observations with a compact array