Separating the global 21-cm spectrum from foregrounds with the DARE instrument using a novel SVD approach

David Rapetti (CU Boulder/NASA Ames) & Keith Tauscher (CU Boulder)

Outline

1. Take away result: current 21-cm signal constraints using end-to-end DARE simulations

2. Challenges of extracting the 21-cm spectrum from large foregrounds

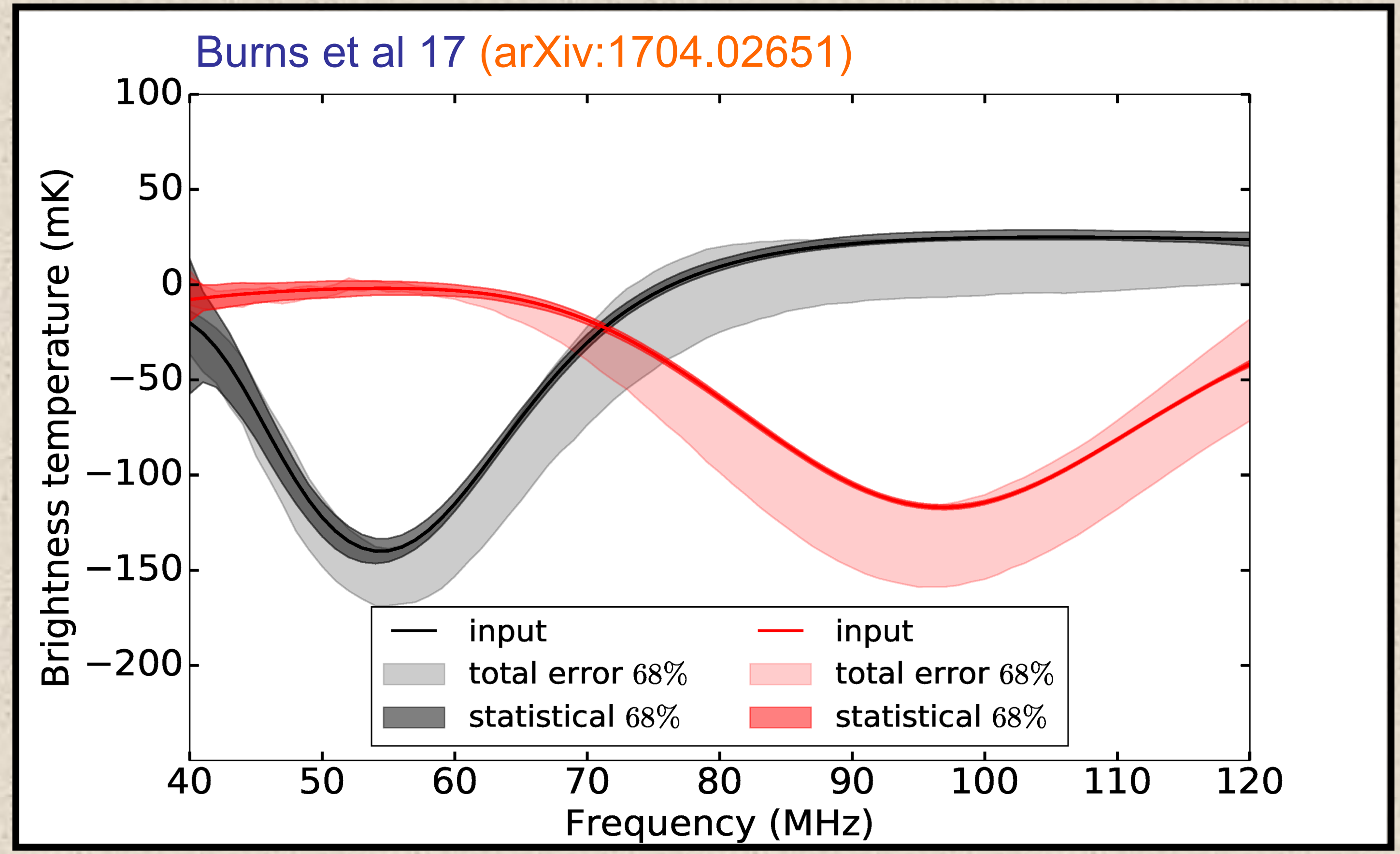
3. Comparing previous approaches to the current

4. Using an SVD approach in an MCMC pipeline

5. Results: modes, covariances, signal extraction

6. Further advances: ongoing work

Current pipeline results using DARE instrument



- Extracted spectra for models with Pop II (red) and Pop III (black) stars
- Dark bands represent thermal noise from the sky (800 hours integration)
- Light bands represent total uncertainty

Biggest challenges of measuring global signal

- Unavoidable (beam-averaged) foregrounds which are $> 10^4$ times larger than signal
  - Requires precise calibration
- Beam chromaticity mixes spatial and spectral structure of foreground

Past approach

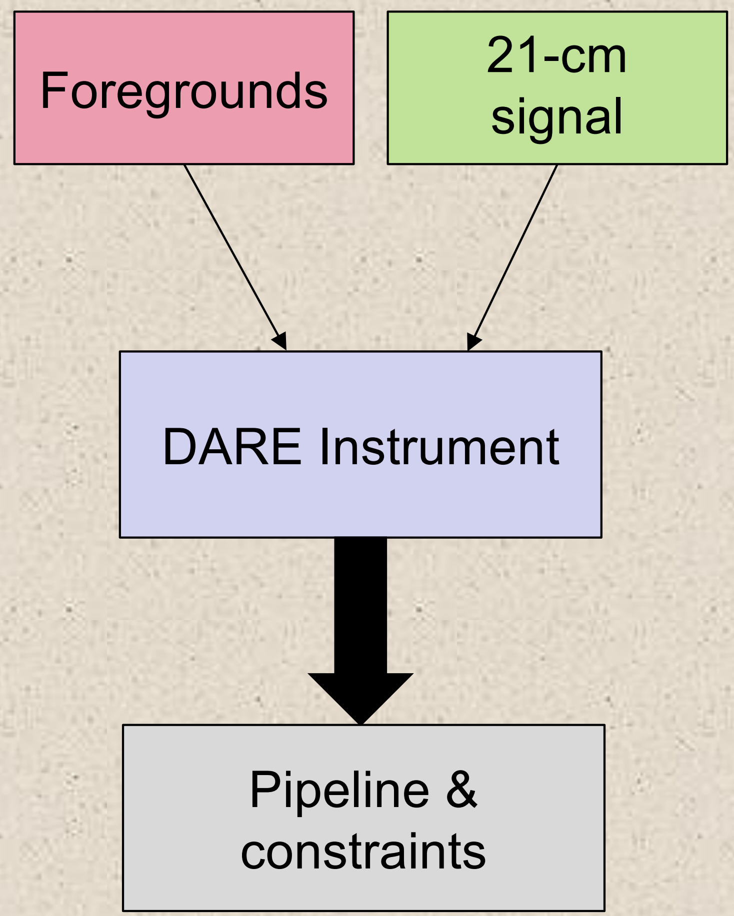
(Harker et al. 2012, 2016)

- Polynomials used to fit foregrounds
- Fourier series used to fit instrumental effects
- In both cases above, the basis functions overlap significantly with the signal
  - Extreme knowledge of foregrounds and instrument were required

New pipeline's extraction approach

- Pipeline calculates main modes of spectral variation in the data via an algorithm which learns from simulations based on lab and sky measurements.

- Foregrounds and instrument modes calculated this way (i.e. adaptively) are less likely to overlap with the signal than polynomials and Fourier series.



Singular Value Decomposition (SVD)

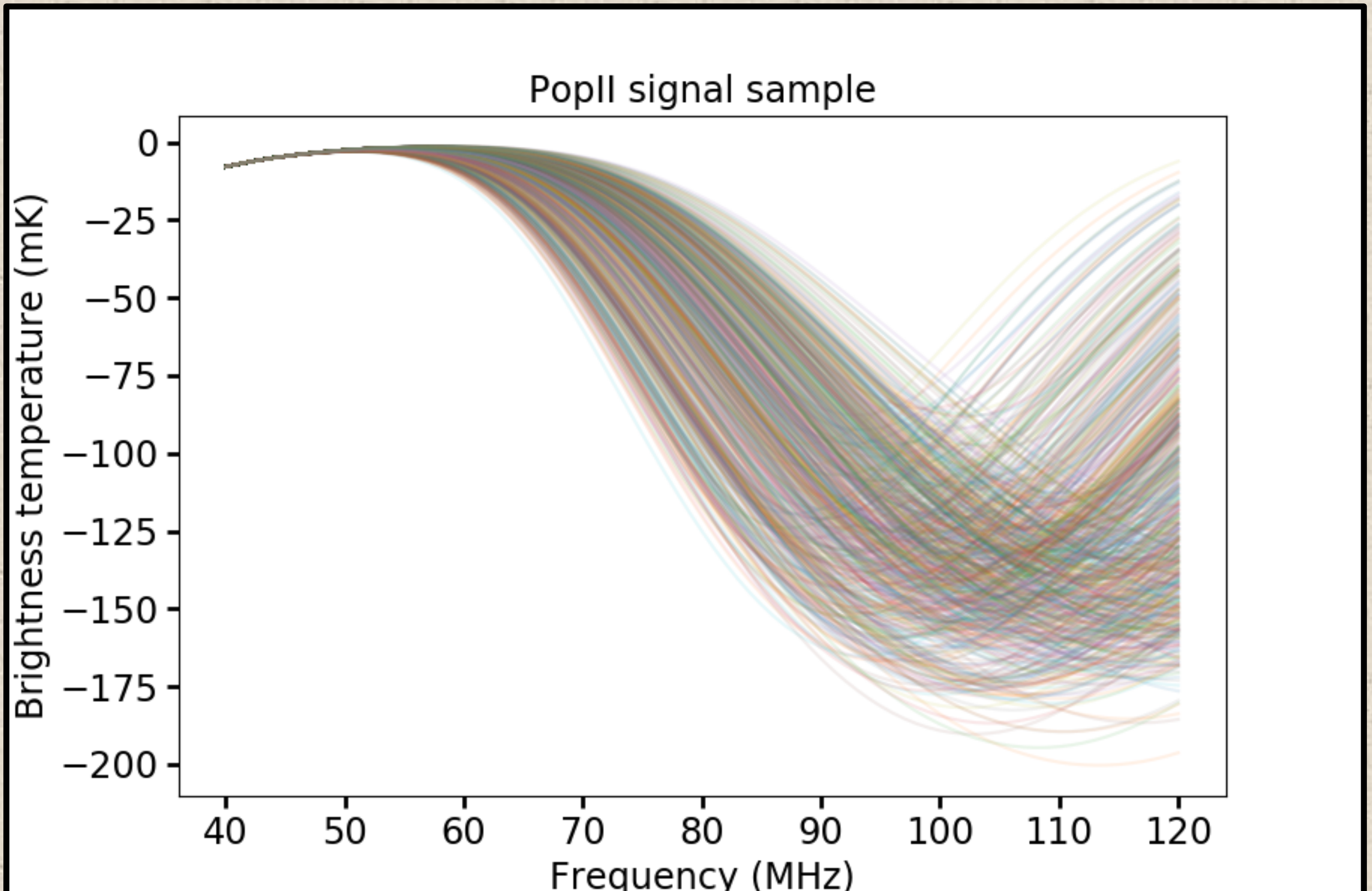
$$\underbrace{\boldsymbol{M}} = \underbrace{\boldsymbol{U}}\boldsymbol{\Sigma}\boldsymbol{V}^T$$

Training Set: $(N_{channel} \times N_{curves})$

Ordered basis functions: $(N_{channel} \times N_{channel})$

- SVD orders the orthogonal modes of the N_{curves} curves of the training set, $\boldsymbol{M}$, by importance
- $\boldsymbol{\Sigma}$ is a diagonal matrix containing the importance of the modes (square root of eigenvalues of $\boldsymbol{M}\boldsymbol{M}^T$)

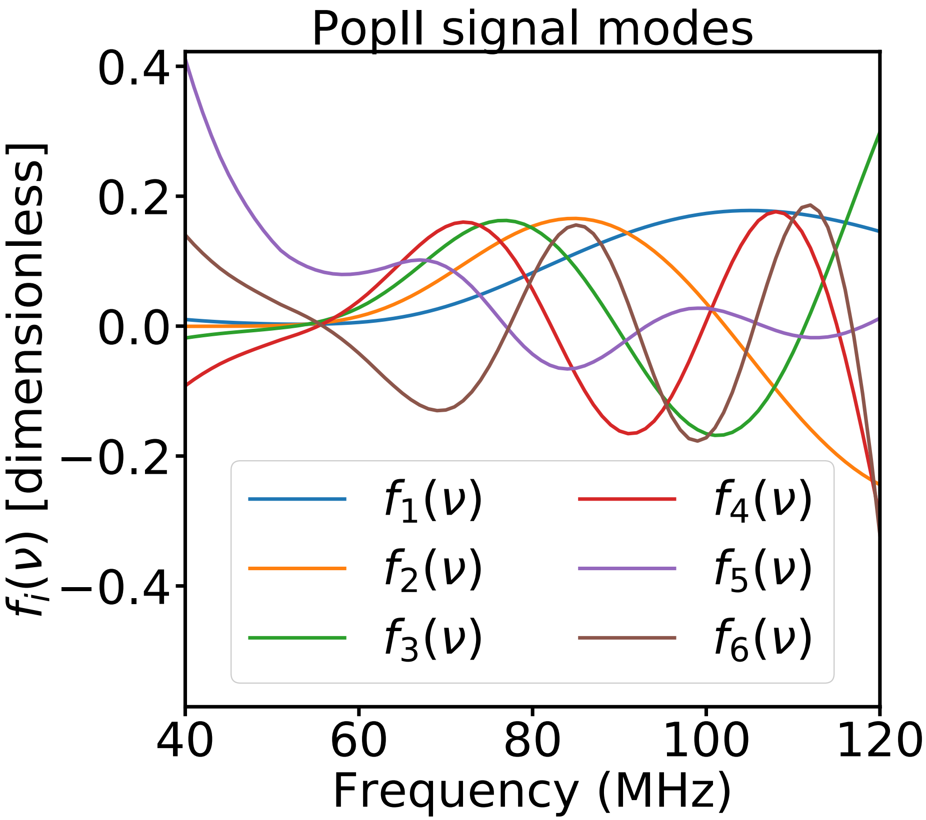
Training set

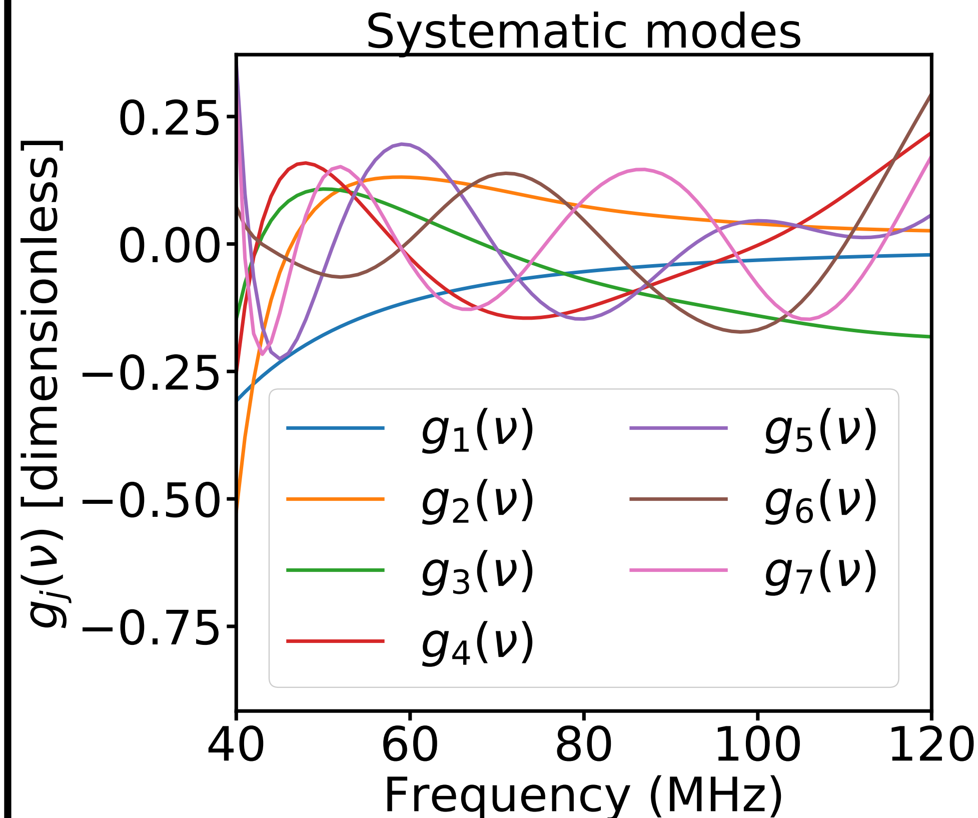


- Derived by randomly sampling the parameter space surveyed in Mirocha et al 17 with the addition of two parameters describing UV and X-ray photon production efficiency in minihalos.

SVD modeling of the signal and systematics

Burns et al 17 (arXiv:1704.02651)





- SVD orthonormal modes. The ability to separate the 21-cm signal from DARE's systematics hinges on the ability to distinguish between the signal (f) and systematic (g) modes. Therefore, we want minimal overlap between them.

SVD modeling of the signal and systematics

Likelihood function

$$\ln L(\gamma) = -\frac{1}{2}\sum_{r=1}^{N_r}\sum_{i=1}^{N_\nu}\left[\frac{T_{A,D}^{(r)}(\nu_i) - T_{A,M}^{(r)}(\nu_i,\gamma)}{\sigma_r(\nu_i)}\right]^2$$

Global antenna temperature model

$$T_{A,M}^{(r)}(\nu,\gamma) = \sum_{i=1}^{n}(\gamma_{21})_i f_i(\nu) + \sum_{j=1}^{m}(\gamma_{sys})_j^{(r)} g_j(\nu)$$

D : Data

$(\gamma_{21})_i$: signal pars

$f_i(\nu)$: signal modes

M : Model

$(\gamma_{sys})_j^{(r)}$: systematic pars

$g_j(\nu)$: systematic modes

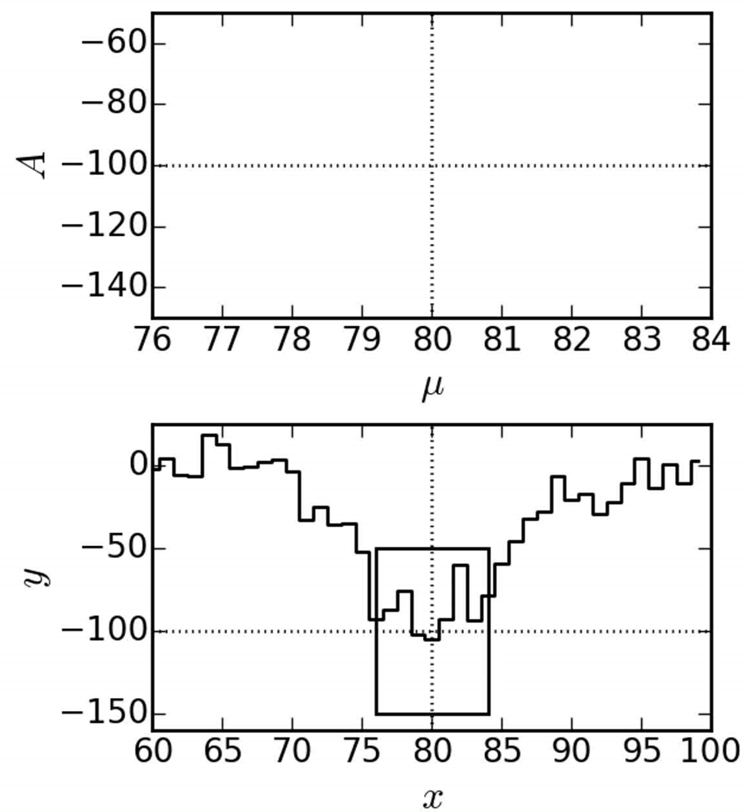
$\sigma_r(\nu) = T_{A,D}^{(r)}(\nu)/\sqrt{\Delta\nu\,\Delta t}$: error

Parameter estimation with MCMC

Markov Chain Monte Carlo (MCMC)

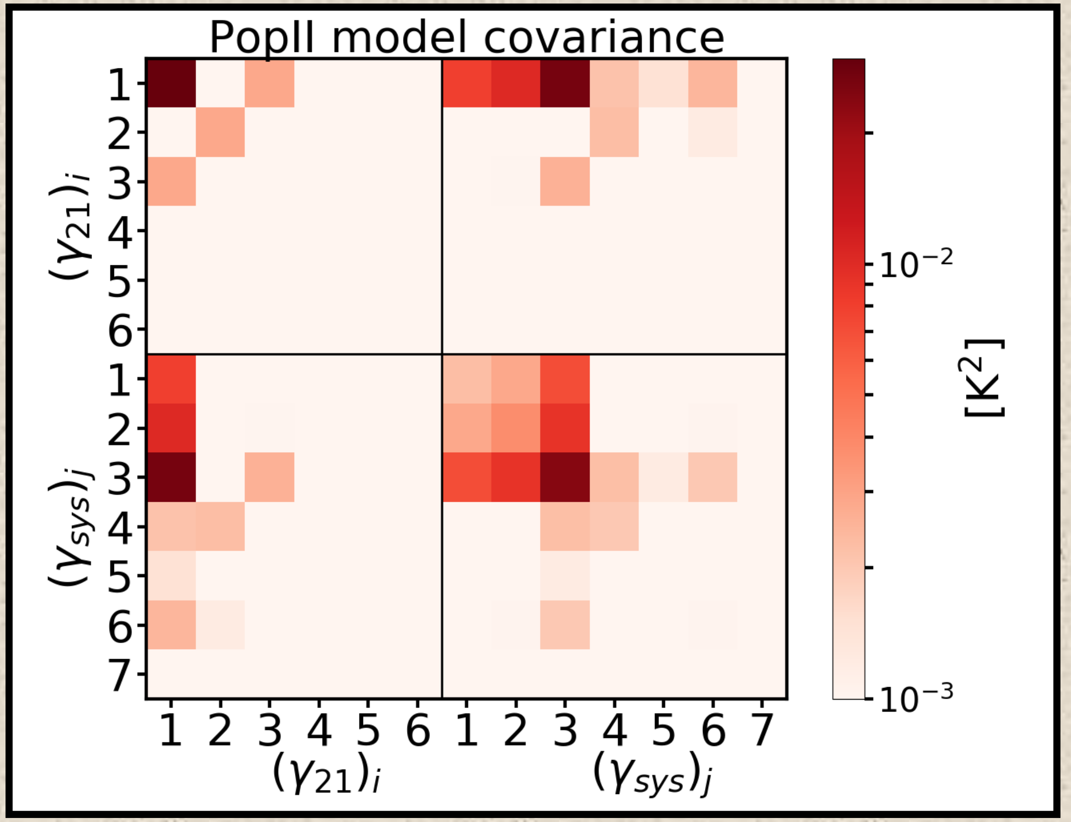
- Explores the parameter space $\gamma=[\gamma_{21},\gamma_{sys}]$ defined by the SVD modes
- Accounts for covariances between all parameters
- Provides robust estimation of posterior parameter distribution

Credit: Jordan Mirocha



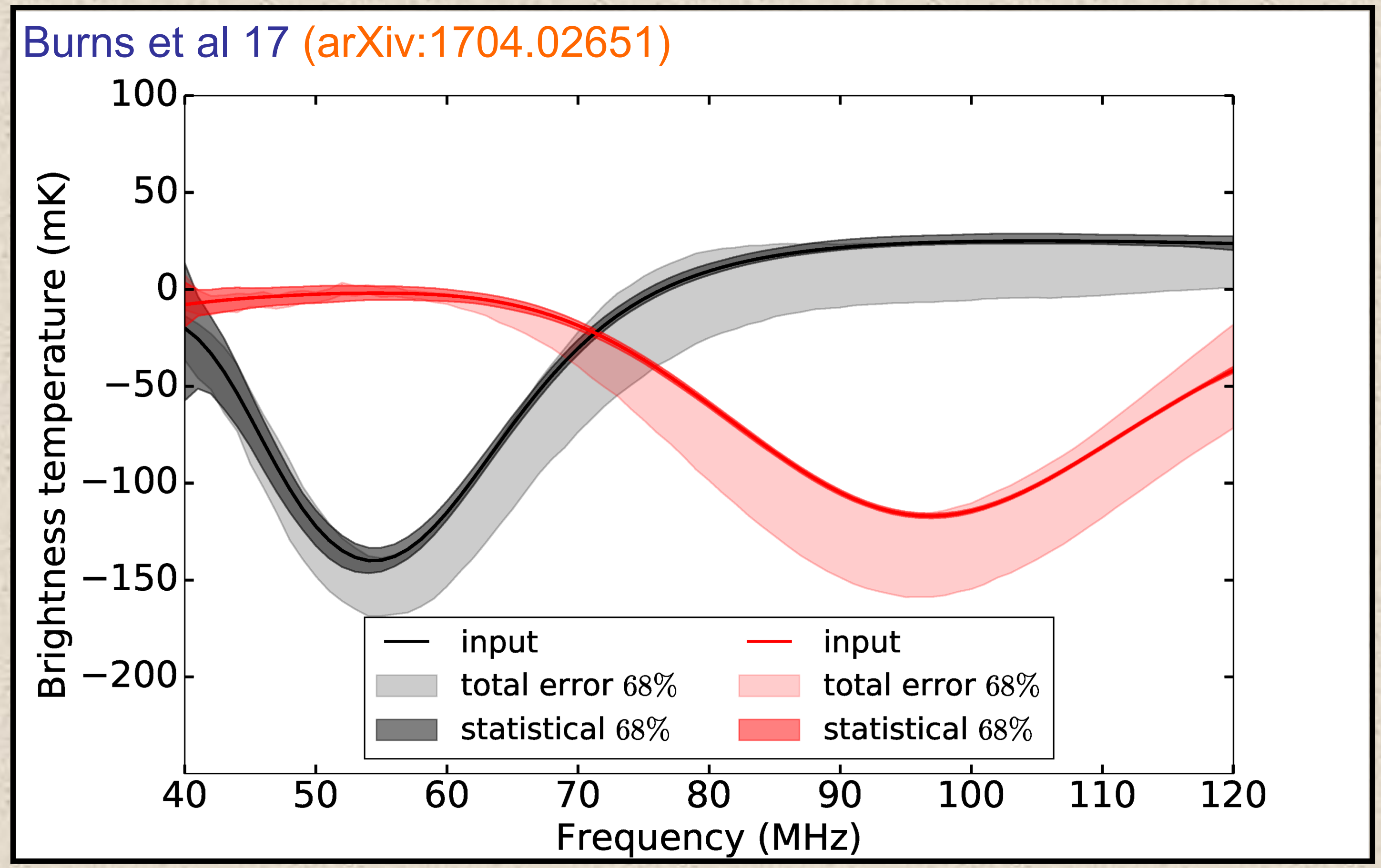
Covariances between the signal and systematics modes

Burns et al 17 (arXiv:1704.02651)



-The vertical and horizontal black lines separate the regions with covariances between signal parameters (top left) and systematic parameters (bottom right). The other two regions are symmetric and show the covariances between signal and systematic parameters.

Extracting the signal from the systematics

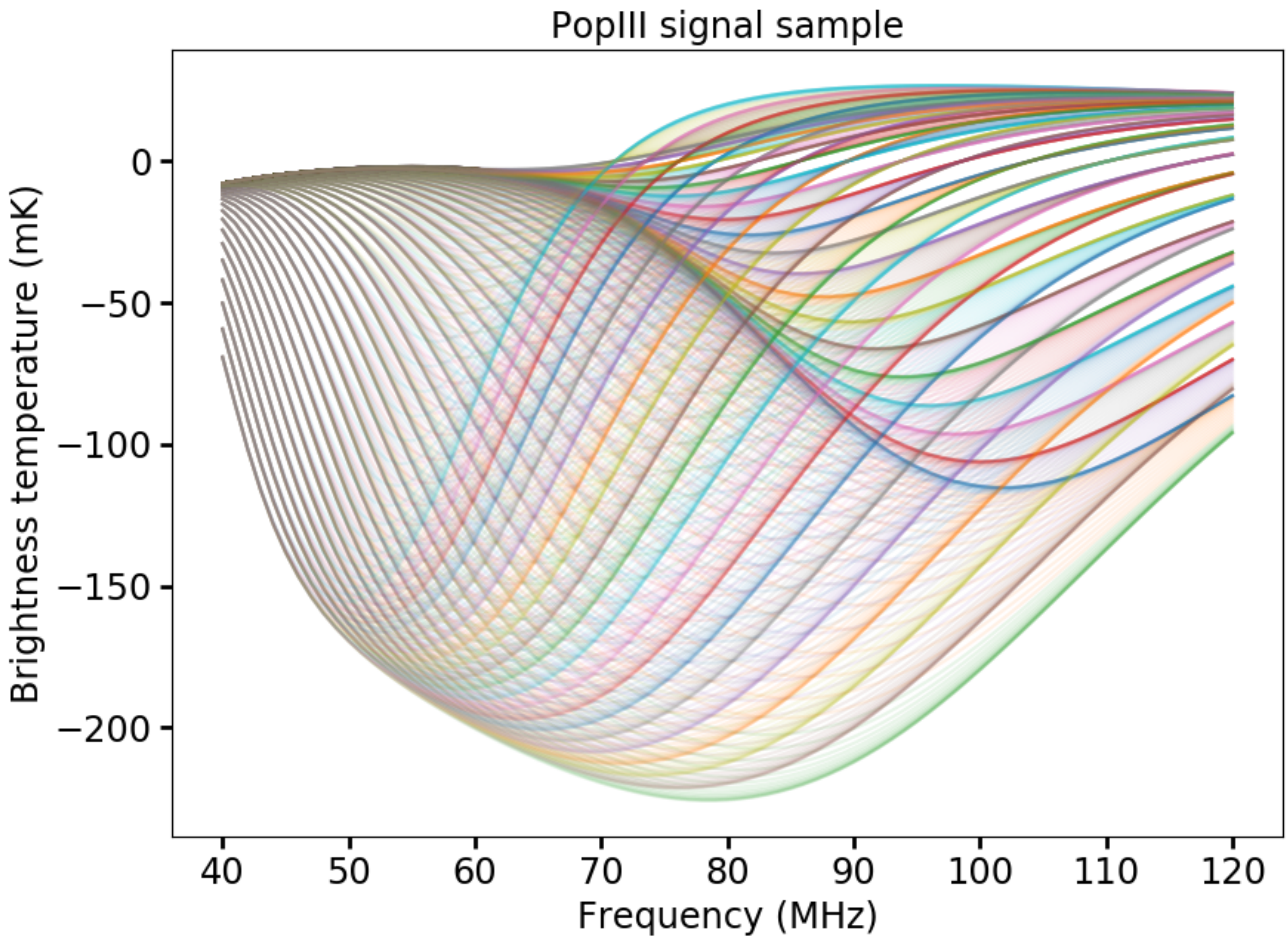


- The extracted 21-cm spectra for models with primordial Pop II (red) and Pop III (black) stars for 800 hours of observation with DARE.
- Dark bands: thermal (statistical) noise from the sky. Lighter bands: total uncertainty, statistical plus systematic effects (instrument and foreground).
- The covariance between SVD signal and systematic modes dominates the total error.

Ongoing work

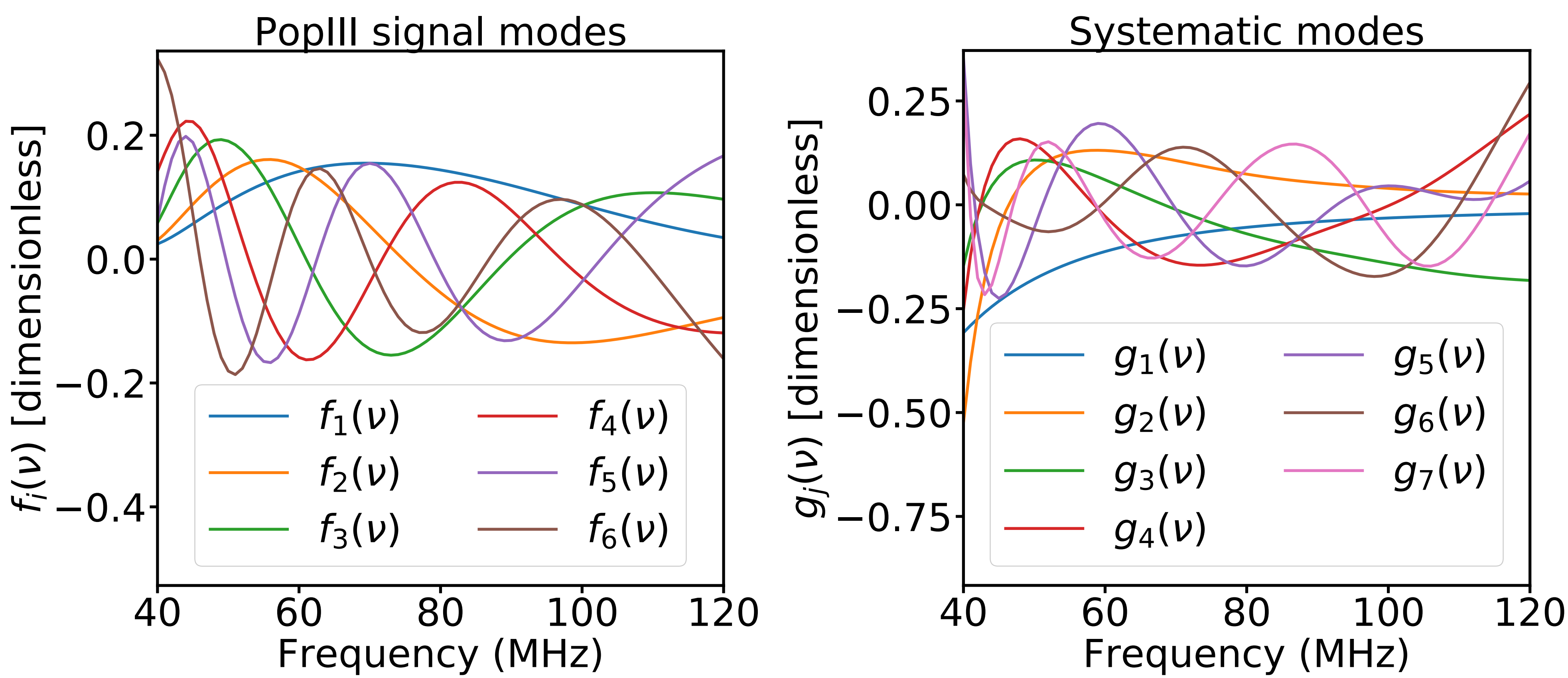
- Incorporating data from rotating dual polarization antennas to take advantage of projection induced polarization caused by the large beam.
- Optimizing basis set and number of parameters:
  - Bayesian evidence
  - Importance of SVD eigenmodes
  - Errors introduced by each eigenmode

Backup slides

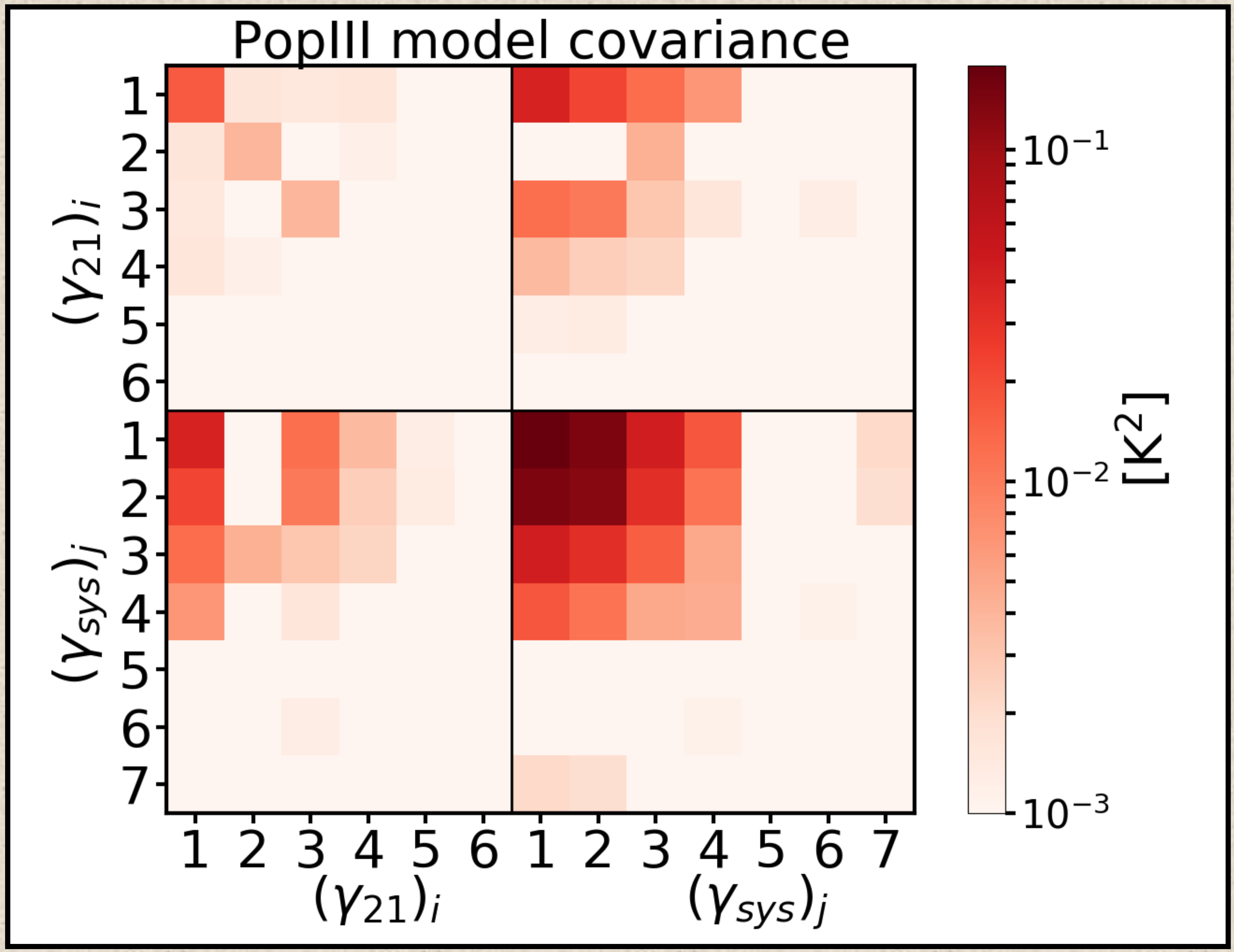

PopIII signal sample
Brightness temperature (mK)
0
−50
−100
−150
−200
40
50
60
70
80
90
100
110
120
Frequency (MHz)

SVD modeling of the signal and systematics

Burns et al 17 (arXiv:1704.02651)



- SVD orthonormal modes. The ability to separate the 21-cm signal from DARE's systematics hinges on the ability to distinguish between the signal (f) and systematic (g) modes. Therefore, we want minimal overlap between them.


PopIII model covariance
$(\gamma_{21})_i$
$(\gamma_{sys})_j$
1 2 3 4 5 6
1 2 3 4 5 6 7
10^{-1}
10^{-2}
10^{-3}
[K^2]

Constraints on extrema frequencies and physical parameters