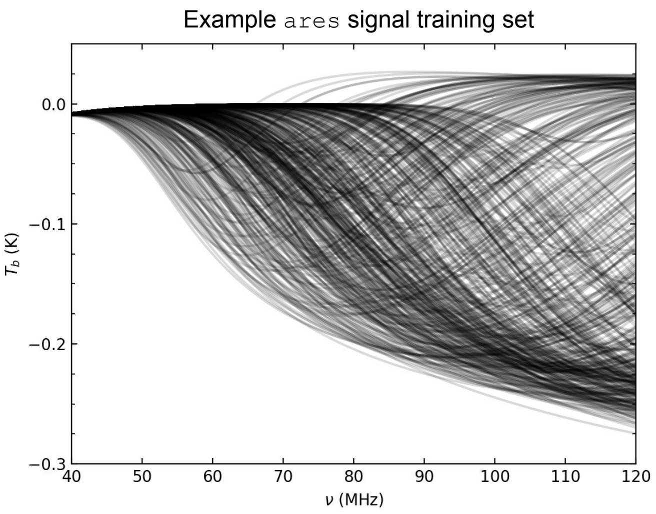
Ensuring Robustness in Training Set Based Global 21-cm Analysis

Neil Bassett
University of Colorado Boulder

In collaboration with
David Rapetti, Jack Burns, Keith Tauscher, and Joshua Hibbard

Training Set-Based Pipeline

Step 1: Create training sets for each component of data

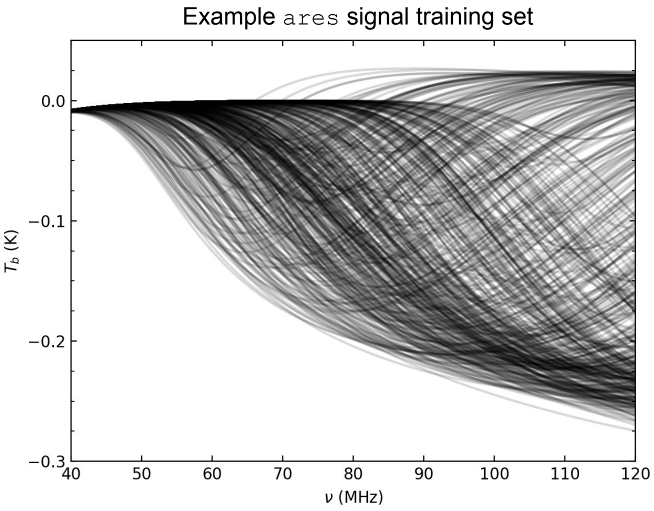


Training sets require knowledge of how components may vary rather than precise, absolute knowledge of their exact form

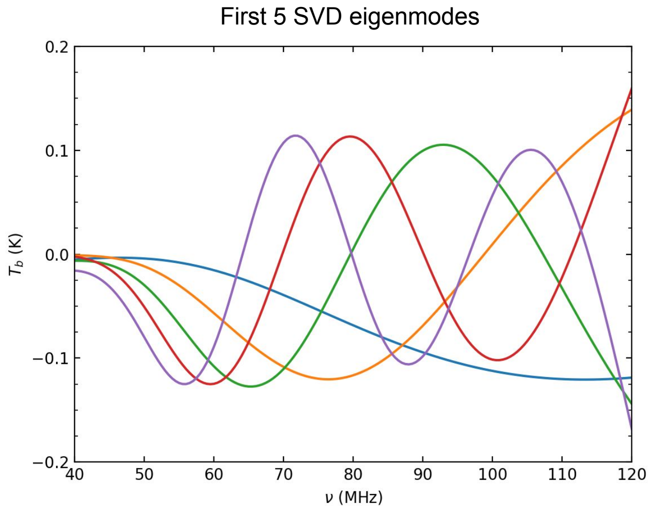
Linear portion of pipeline detailed in Tauscher et al. 2018a (Paper I)

Training Set-Based Pipeline

Step 2: Perform singular value decomposition (SVD) on training sets



$$\boldsymbol{B} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$$



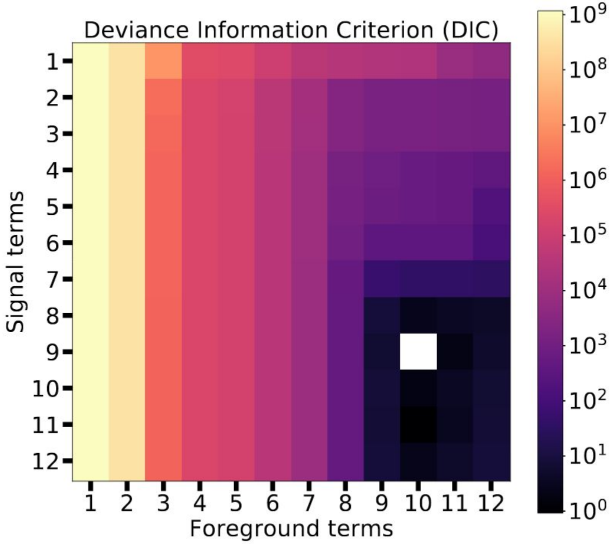
Linear portion of pipeline detailed in Tauscher et al. 2018a (Paper I)

Training Set-Based Pipeline

Step 3: Minimize information criterion to determine number of modes to use

$$\mathrm{DIC} = \boldsymbol{\delta}^T \boldsymbol{C}^{-1} \boldsymbol{\delta} + 2N_p$$

Information criterion chooses number of parameters large enough to describe data, but small enough to provide reasonable uncertainties.



Linear portion of pipeline detailed in Tauscher et al. 2018a (Paper I)

How do we have confidence that our training sets are sufficient to extract the 21-cm signal both accurately and precisely?

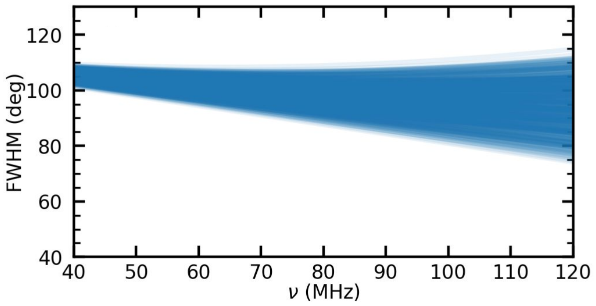
Simulating Observations

Data = (Beam ∗ Sky) + 21-cm Signal + noise

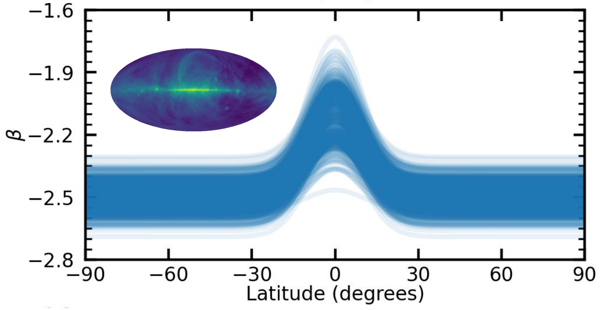
"Foreground"

$$\sigma(\nu) = \frac{T_{\rm sys}}{\sqrt{\Delta\nu \Delta t}}$$

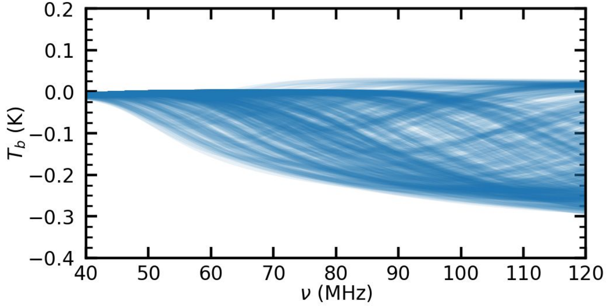
Beam



Sky



Signal



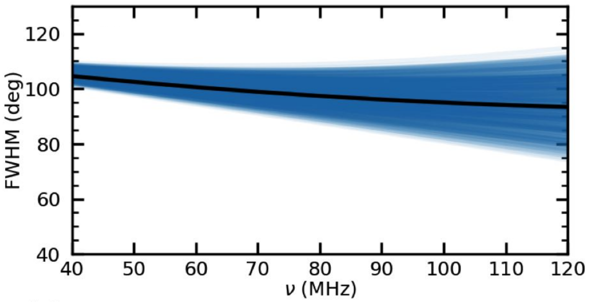
Simulating Observations

Data = (Beam ∗ Sky) + 21-cm Signal + noise

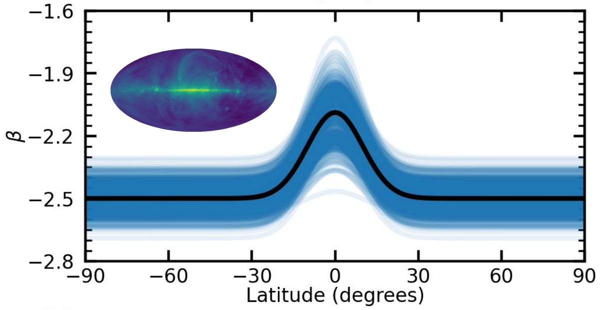
"Foreground"

$$\sigma(\nu) = \frac{T_{\rm sys}}{\sqrt{\Delta\nu\Delta t}}$$

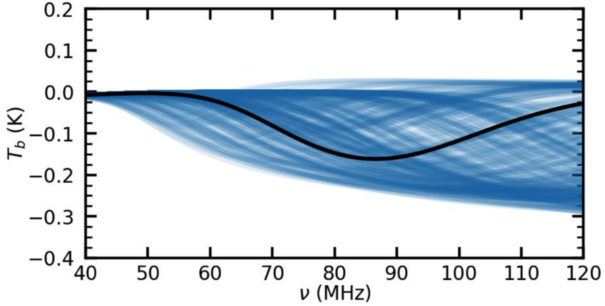
Beam



Sky

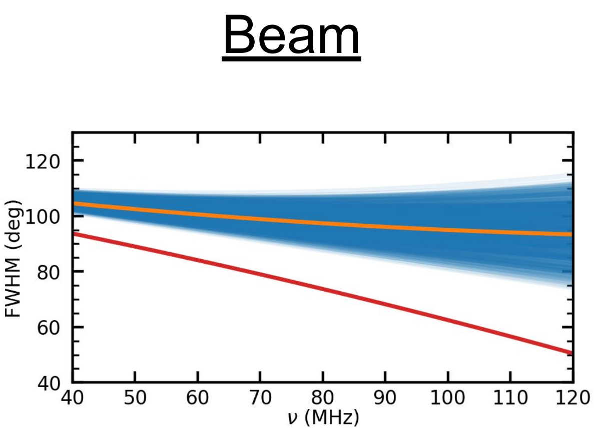


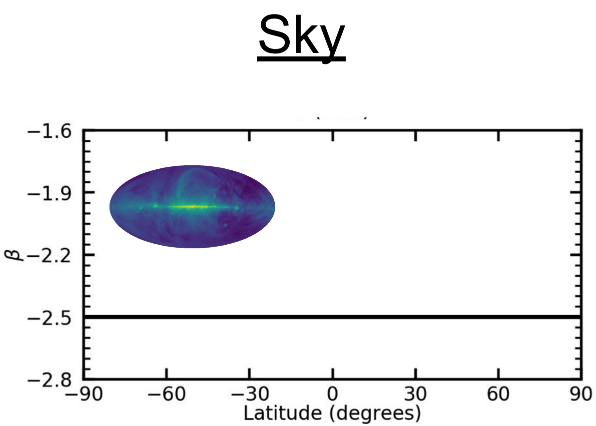
Signal

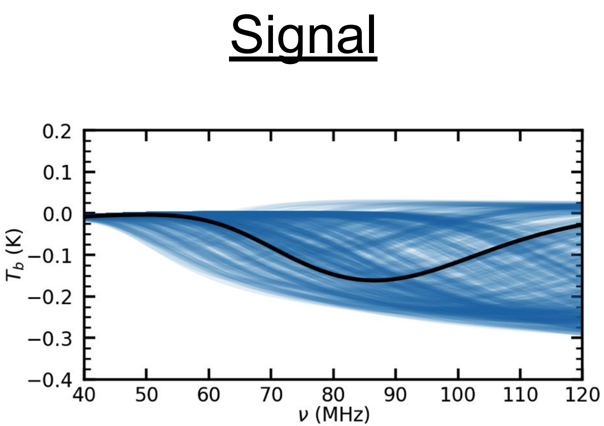


But what happens if one or more of the training sets is wrong?

Traditional Goodness-of-fit Metrics

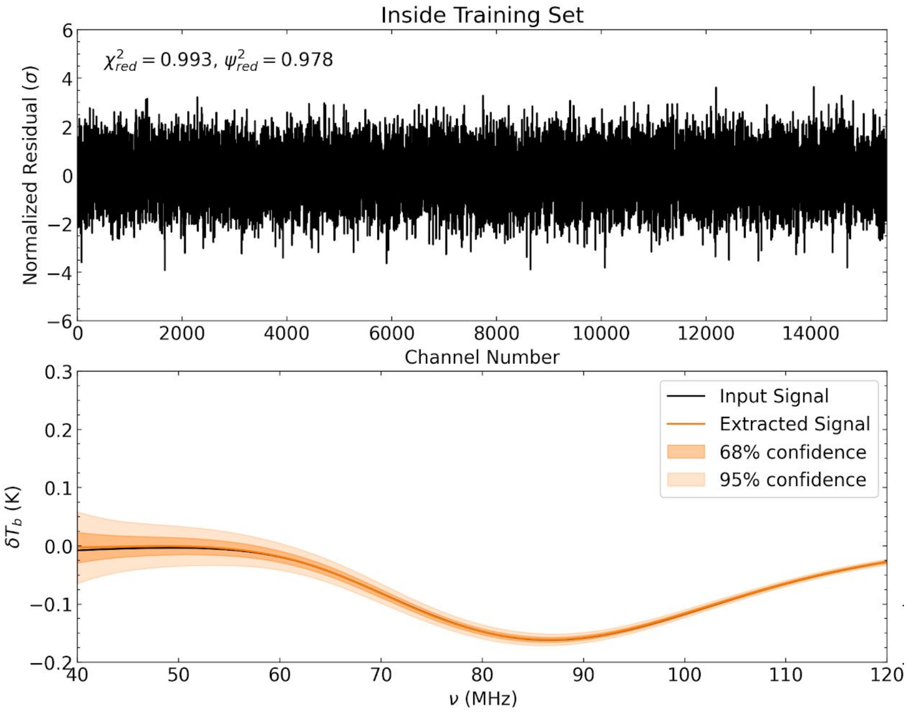




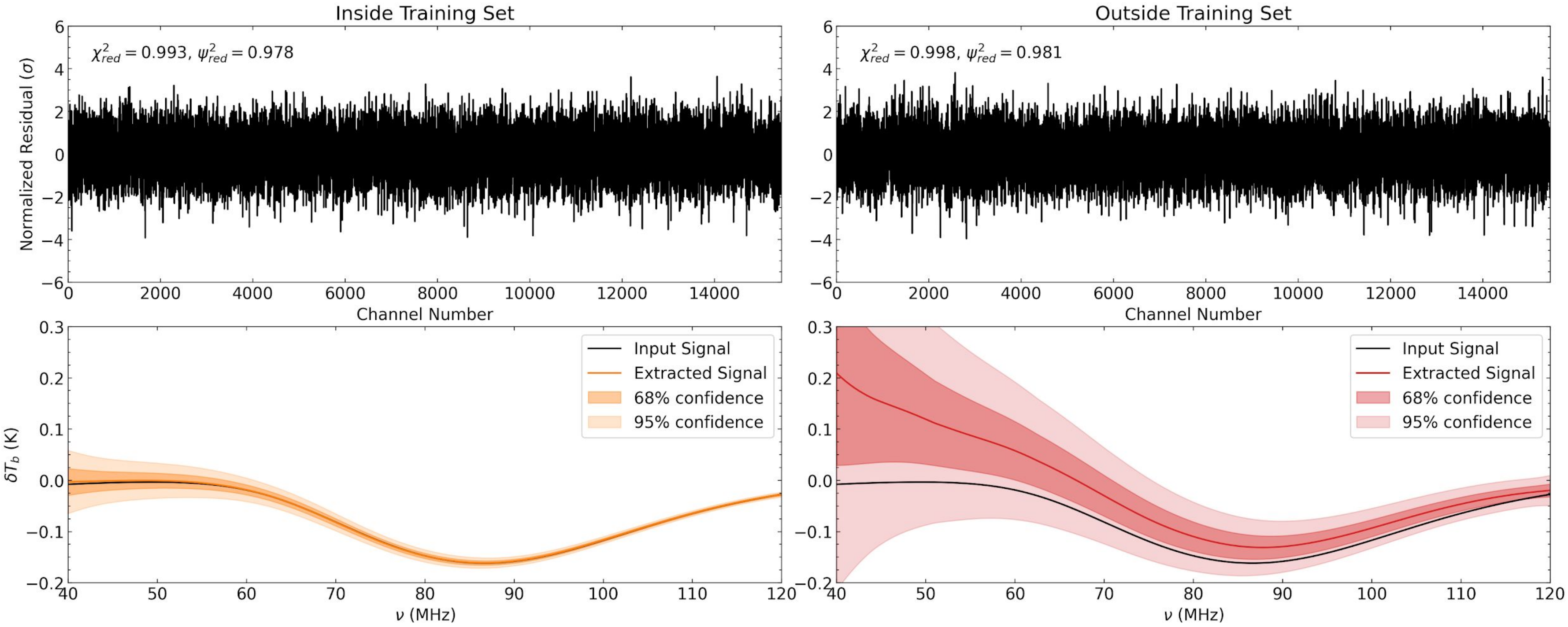


- When analyzing observations, how do we determine how well our model fits our data?
- The χ^2 statistic is often used to measure goodness-of-fit based on the magnitude of the residual.
- The newly-introduced ψ^2 statistic (Tauscher et al. 2018b) is sensitive to channel-to-channel correlations in the residual and may better recognize unmodeled, wide-band features close to the noise level.

Traditional Goodness-of-fit Metrics

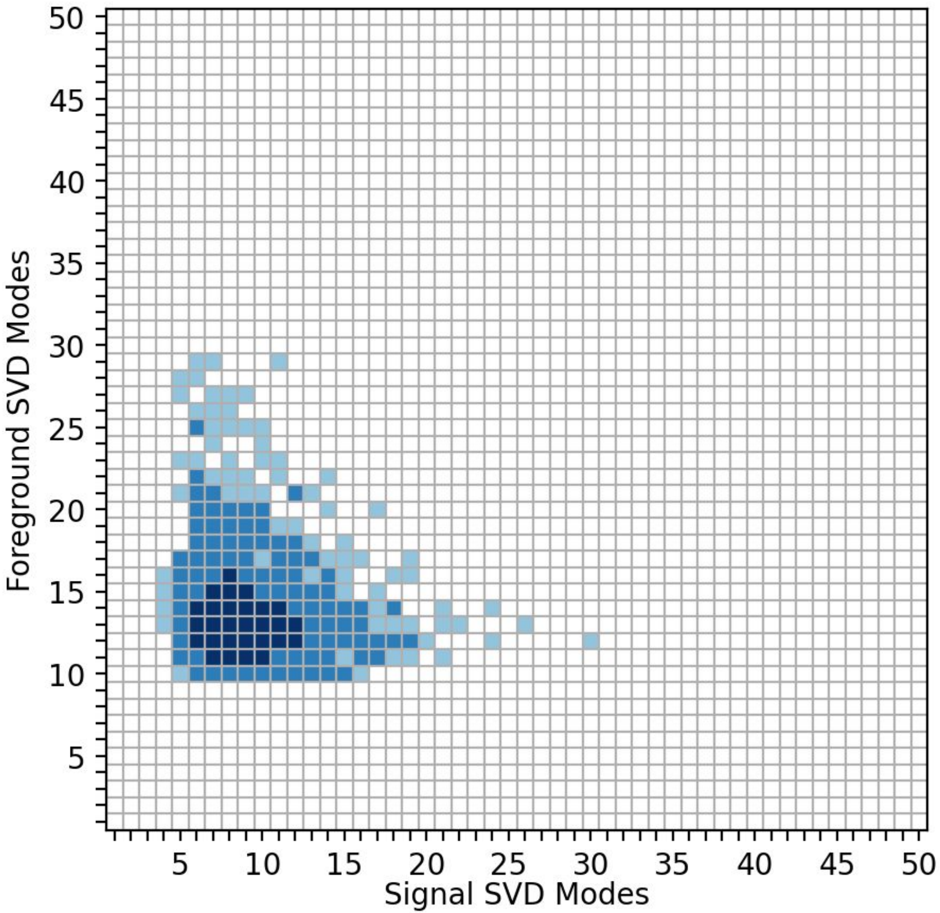


Traditional Goodness-of-fit Metrics



In this example, χ^2 and ψ^2 are unable to detect when the signal extraction is suboptimal!

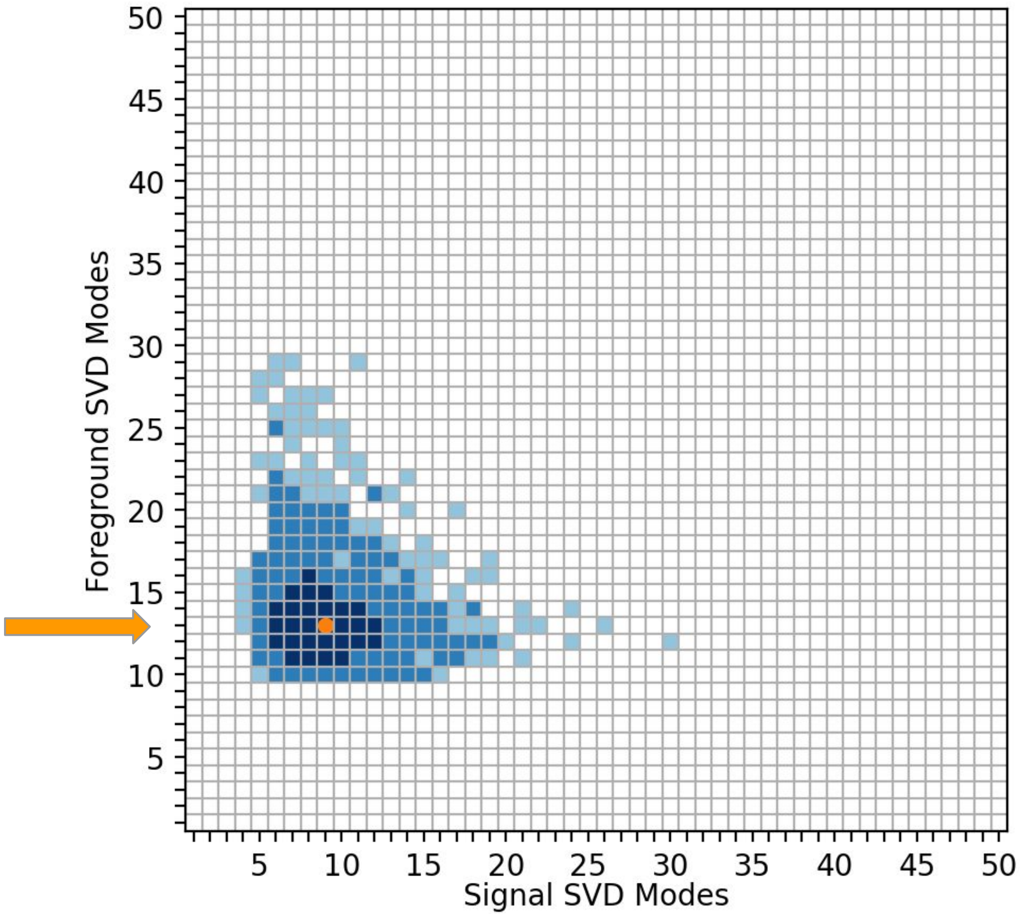
Numbers of SVD Modes



The blue distributions contain 68, 95, and 99.7 percent of 5000 fits with data realizations made directly from the training set.

Numbers of SVD Modes

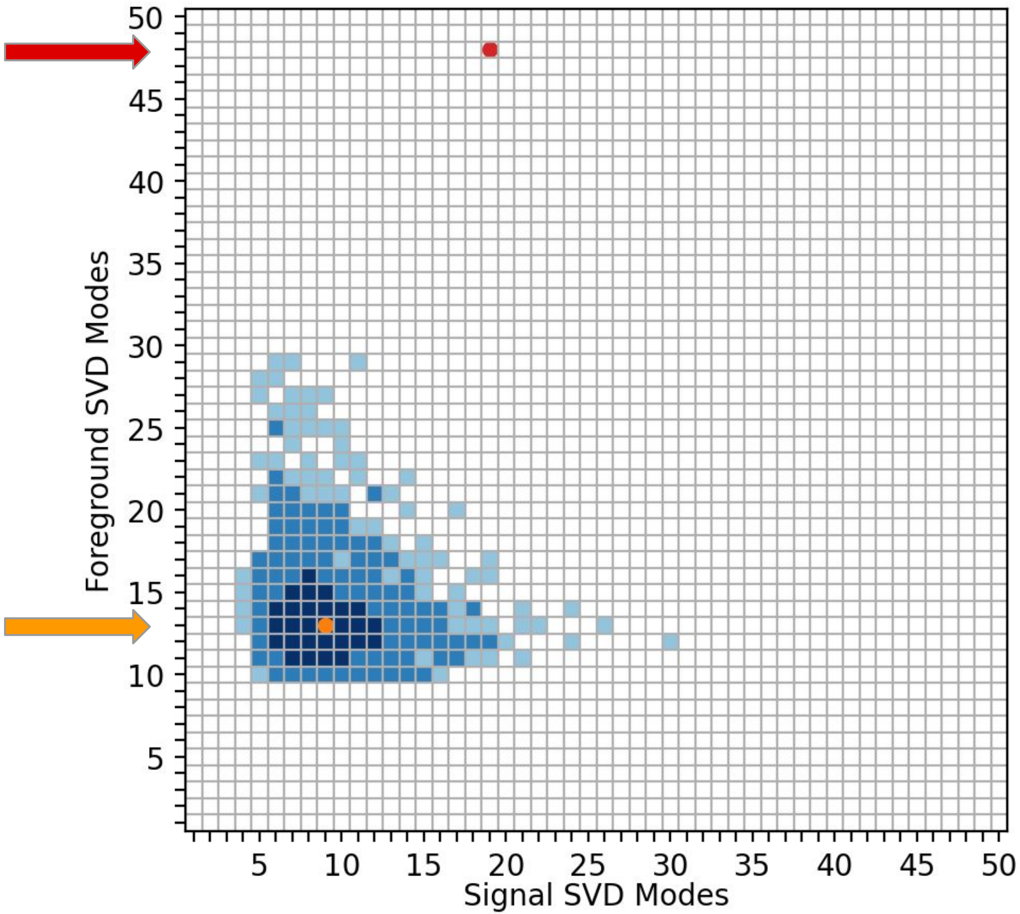
The fit to the data with the beam inside the training set falls well within the training set distribution



The blue distributions contain 68, 95, and 99.7 percent of 5000 fits with data realizations made directly from the training set.

Numbers of SVD Modes

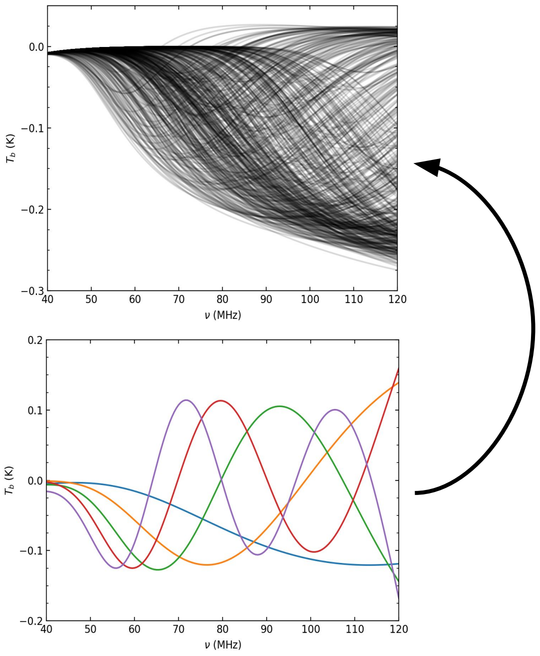
The fit to the data with the beam outside the training set is far from the training set distribution.



The fit to the data with the beam inside the training set falls well within the training set distribution.

The blue distributions contain 68, 95, and 99.7 percent of 5000 fits with data realizations made directly from the training set.

Simulating Observations With Priors



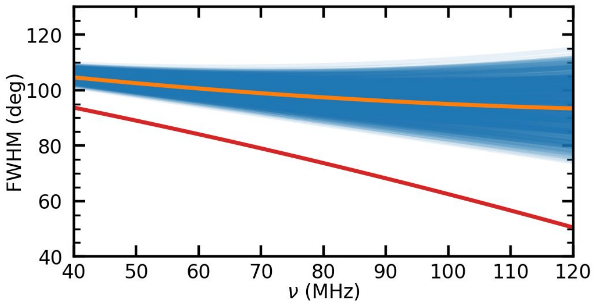
Fitting each training set curve with its own SVD modes produces a mean and covariance for each SVD mode coefficient which we can use to create a Gaussian prior distribution.

Instead of the DIC, we now maximize the Bayesian evidence:

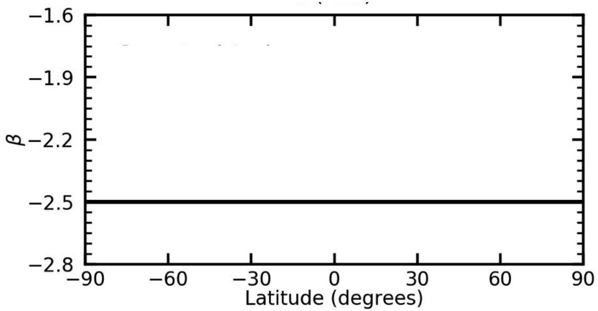
$$\mathcal{Z} \propto \int \mathcal{L}(\boldsymbol{x})\pi(\boldsymbol{x})d^{N_p}\boldsymbol{x}$$

Simulating Observations with Priors

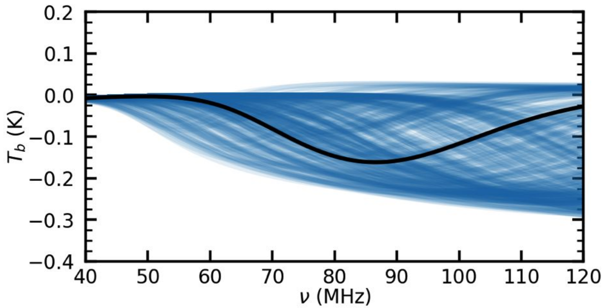
Beam



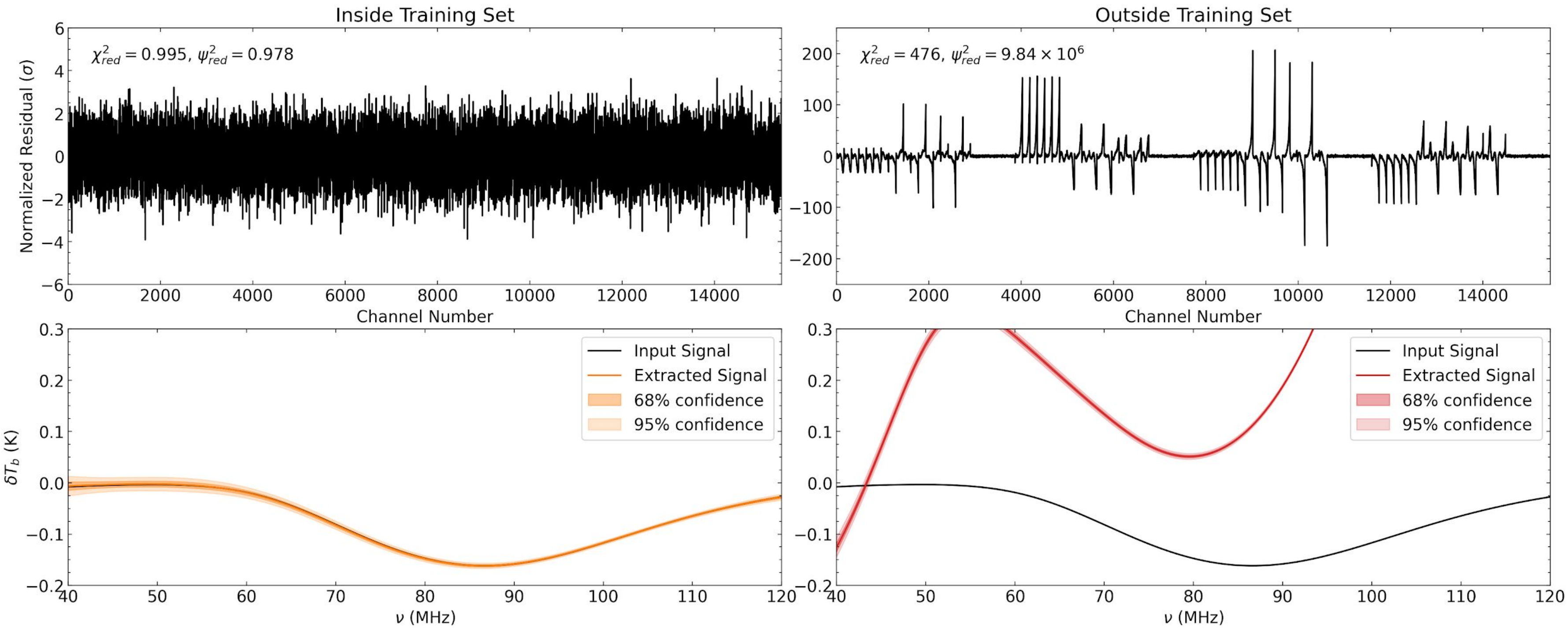
Sky



Signal

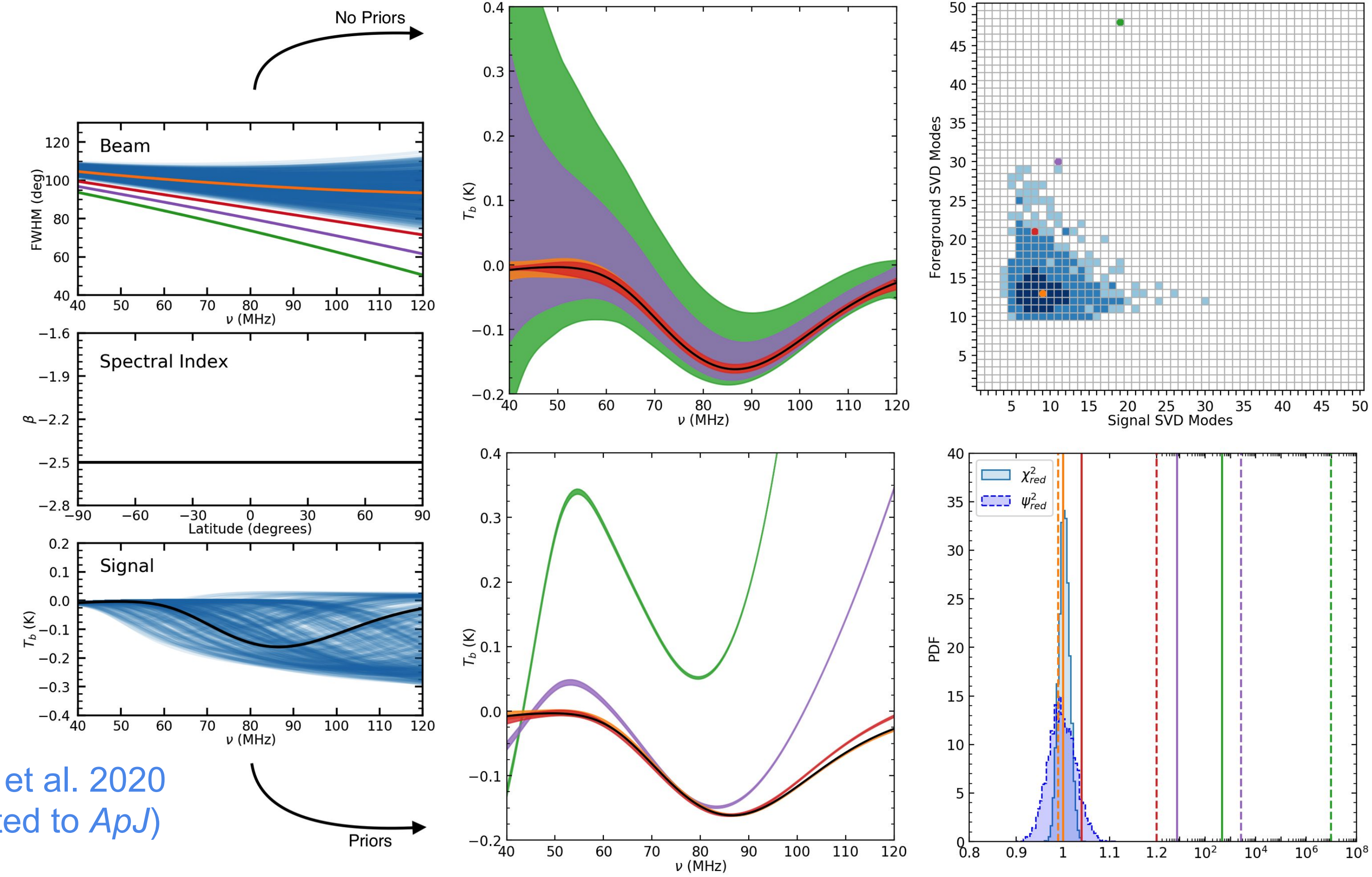


Simulating Observations with Priors



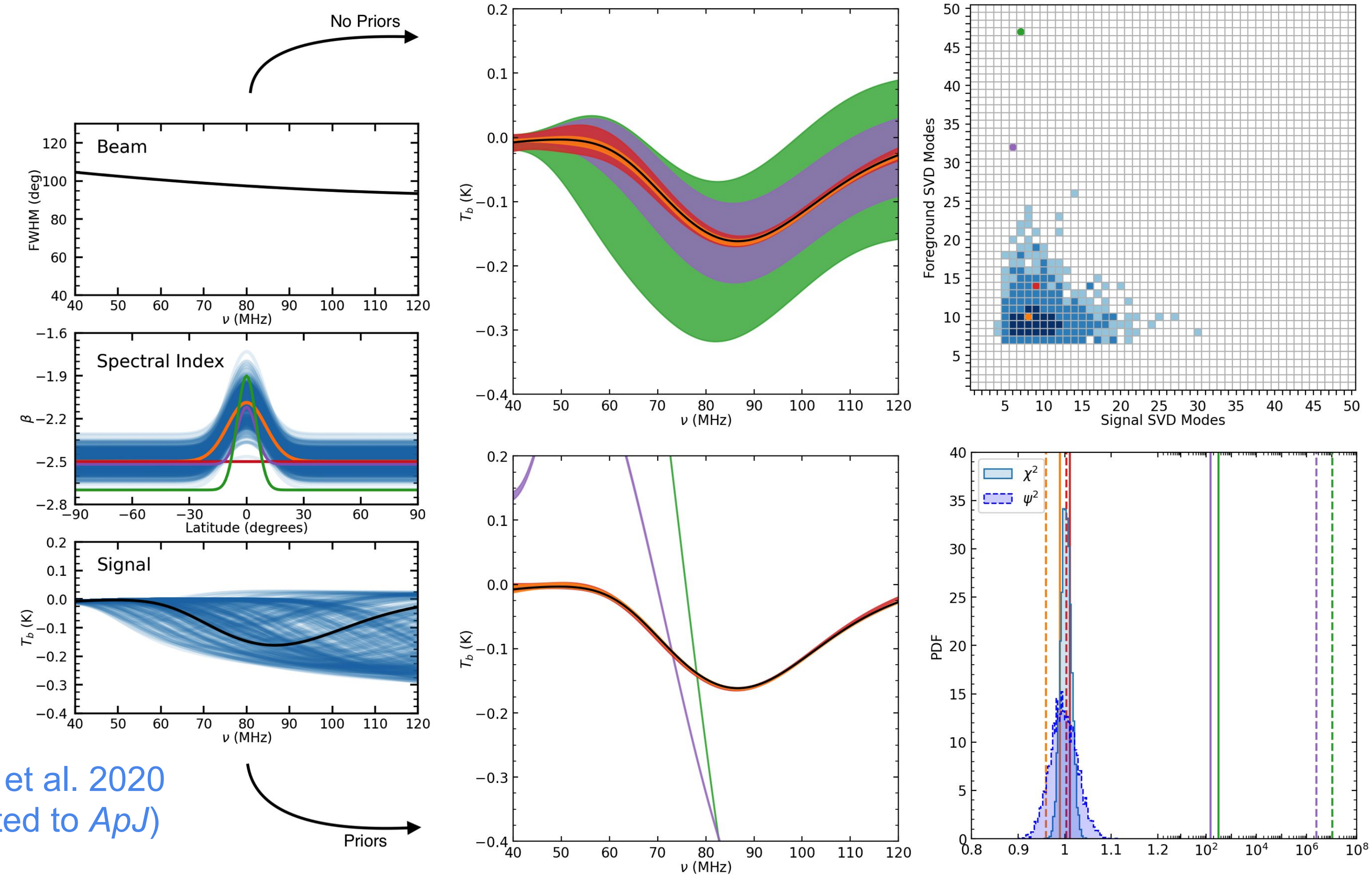
When using priors, χ^2 and ψ^2 are able to detect when the training sets are insufficient.

Beam Examples



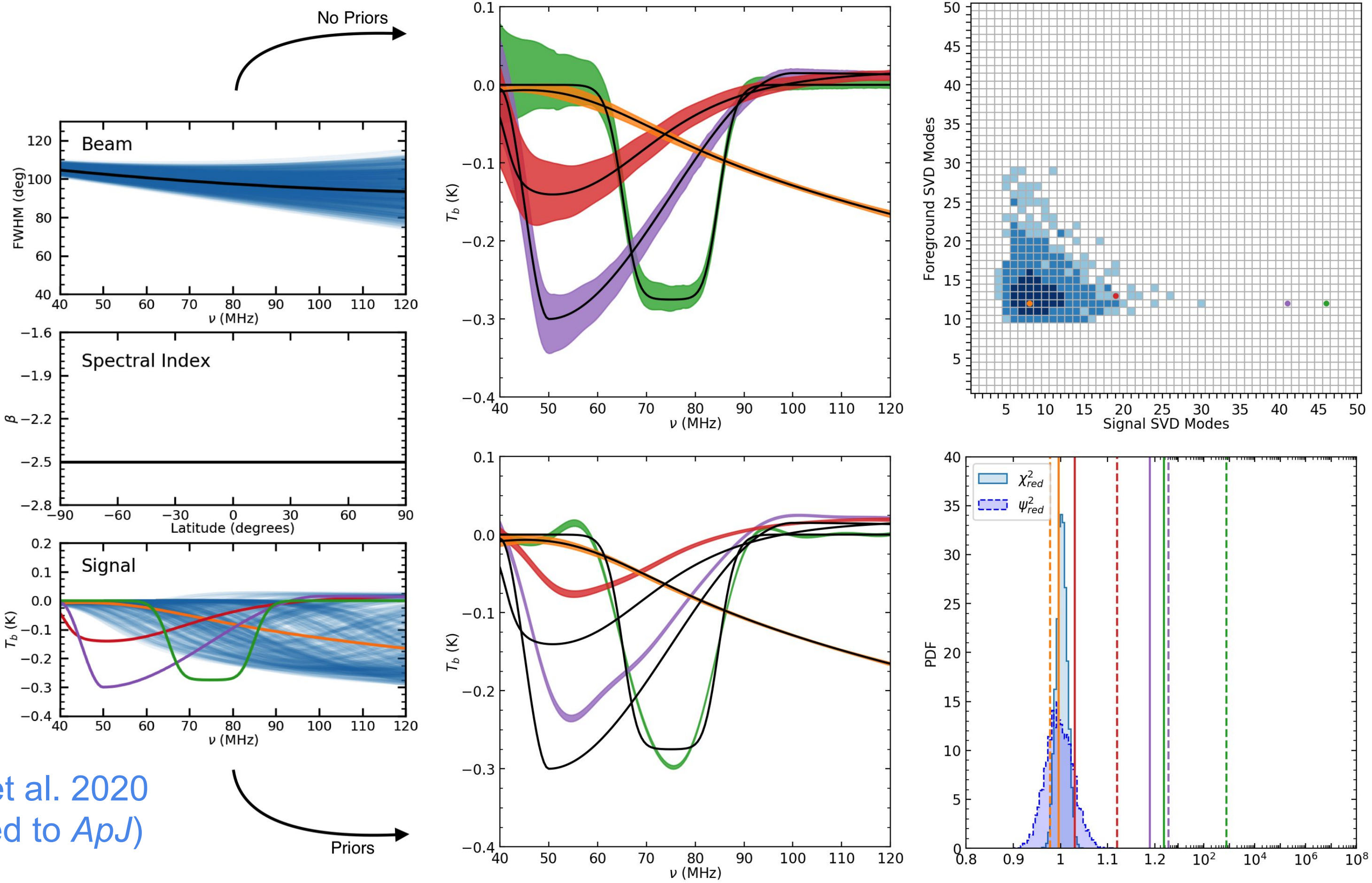
Bassett et al. 2020
(submitted to *ApJ*)

Spectral Index Examples



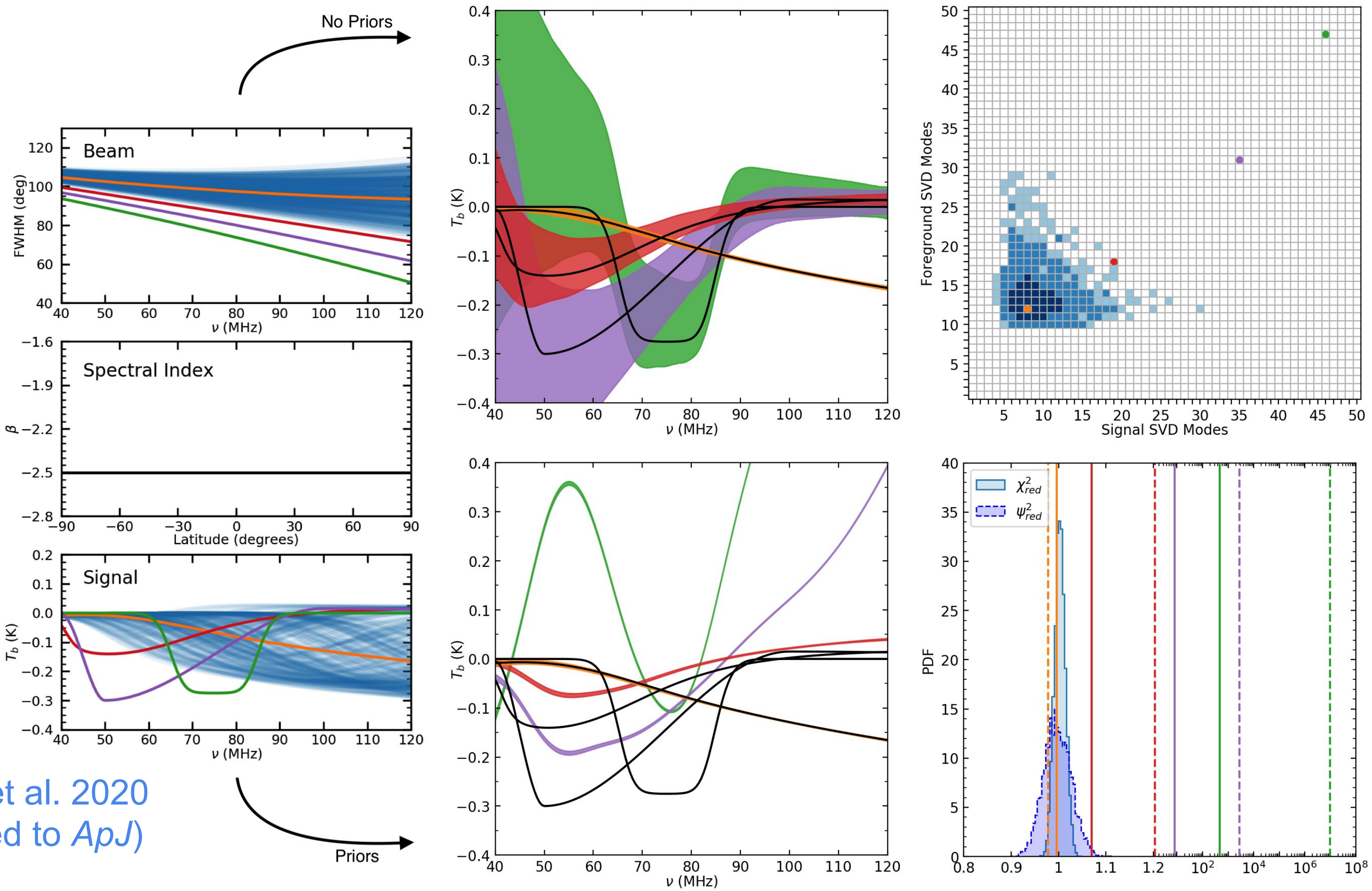
Bassett et al. 2020
(submitted to *ApJ*)

Signal Examples



Bassett et al. 2020
(submitted to *ApJ*)

Beam + Signal Examples

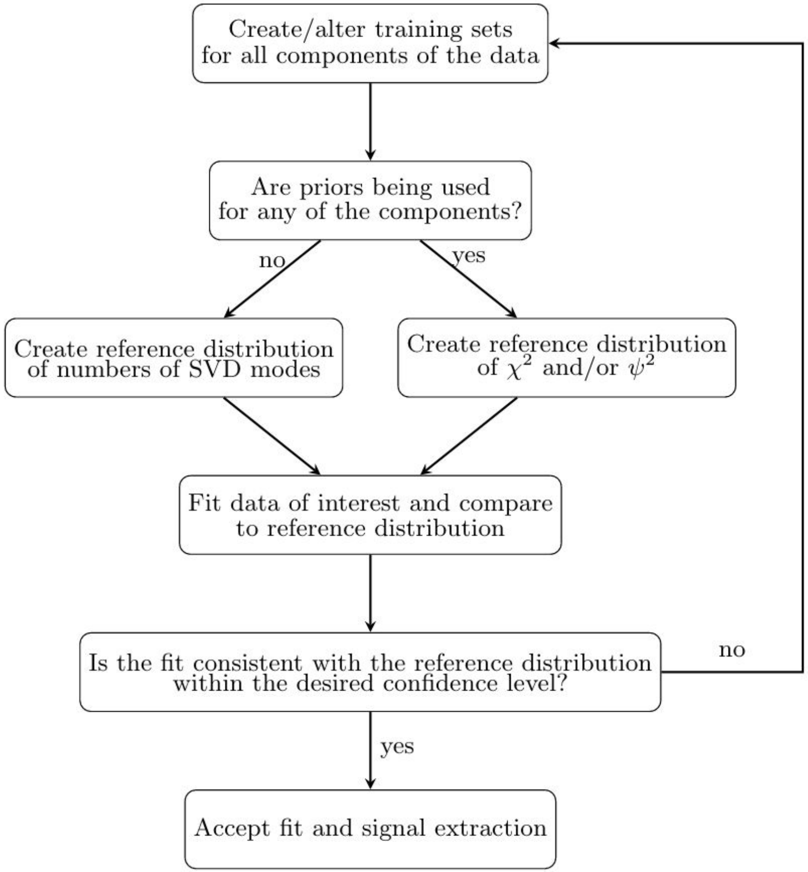


Bassett et al. 2020
(submitted to *ApJ*)

How do we have confidence that our training sets are sufficient to extract the 21-cm signal both accurately and precisely?

- Traditional goodness-of-fit statistics such as chi-squared evaluate the fit to the full data, but may not detect if each component of the data is being fit correctly.
- When no prior distribution is used, the number of SVD modes chosen to fit each component may indicate when one or more of the training sets is insufficient **even when we do not know the true form of each component**.
  - The number of SVD modes may also provide information about which training set should be altered to improve signal extraction.
- When using priors, χ^2 and ψ^2 are able to discern when training sets are inadequate due to the increased restriction of the prior.
- The ability to identify when our training sets are insufficient will lend credibility to any results from our pipeline when analyzing real data.

Test Strategy for Assessing Training Sets

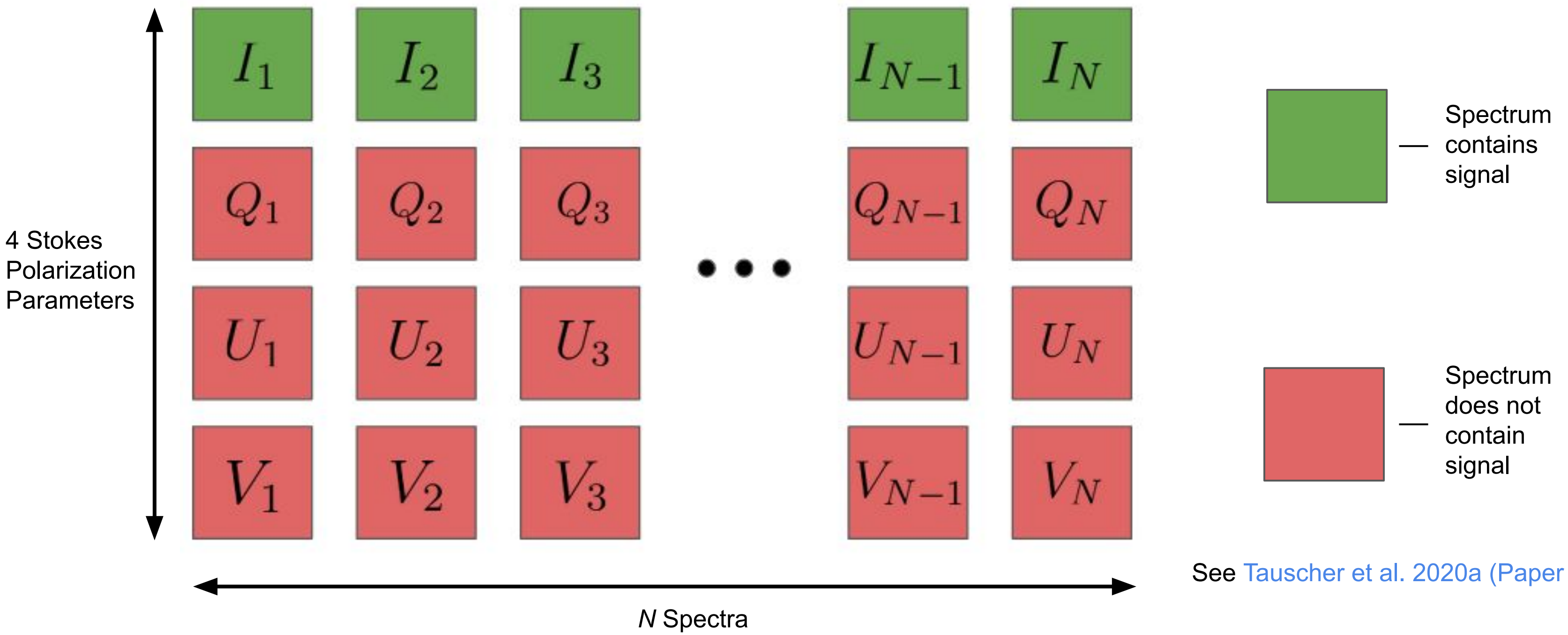


Bassett et al. 2020
(submitted to *ApJ*)

Observation Strategy

Uncertainty of signal extraction depends primarily on overlap between foreground and signal models. This overlap can be decreased by using:

1. Polarization
2. Many Correlated Spectra



See Tauscher et al. 2020a (Paper III)