Fitting and Comparing Galactic Foreground Models for Unbiased 21-cm Cosmology

Joshua J. Hibbard, David Rapetti, Jack O. Burns, Nivedita Mahesh, and Neil Bassett

Image credit: JWST NIR-cam

2 Main Sources of Bias and Error in 21-cm Signal Extraction

1. Overlap between FG-space and Signal-space

- Inevitable as these vector spaces are not, in general, orthogonal.
- Main focus of pipelines.
- Decrease overlap by utilizing LST-dependence of FG and polarization.

THIS WORK

2. Inadequate FG Models

- Models which cannot fit the FG down to the noise level IN THE SIGNAL will introduce ADDITIONAL bias and error, regardless of (1).
- Requires testing the ability of FG models to fit realistic FG-only spectra.
- Less attention, if any, in the literature.

joshua.hibbard@colorado.edu

The Flowchart of Despair

Unmodelled FG + Overlap with Signal Model → Biased Signal Extraction

↓

Increase number of FG parameters/complexity

↓

Complex, HD Parameter FG + Overlap with Signal Model → Large Signal Errors

joshua.hibbard@colorado.edu

The FG Test

STEP ONE → STEP TWO → STEP THREE

- Generate hyper-realistic mock spectra (NO 21-cm SIGNAL)

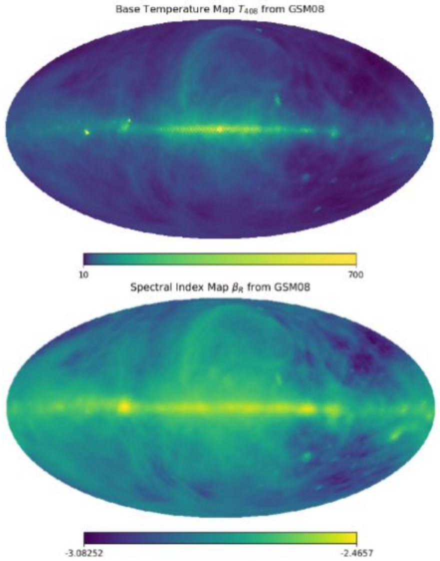
- Fit the spectra with seven different commonly employed FG models

- Compare which model(s) produces the best fits and remove all FG power down to the noise level

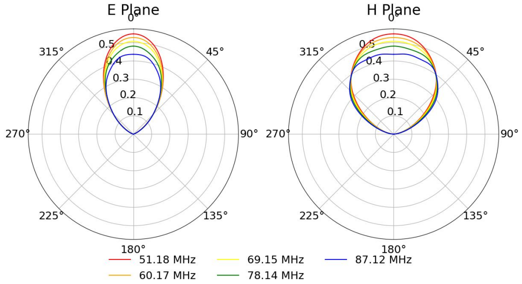
joshua.hibbard@colorado.edu

Foreground Tests: Mock Spectra Simulations

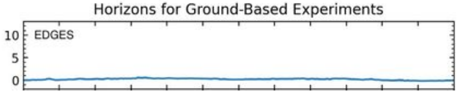
Figures from Hibbard et al. 2023, under review



Intrinsic FG



Beam Simulation,
51-87 MHz band



Horizon, pointing

$$f_n(t - t_i) = \frac{1}{2N_n\delta t}\sum_{i=1}^{N_n} I(|t - t_i| < \delta t)$$

Discrete time-sampling

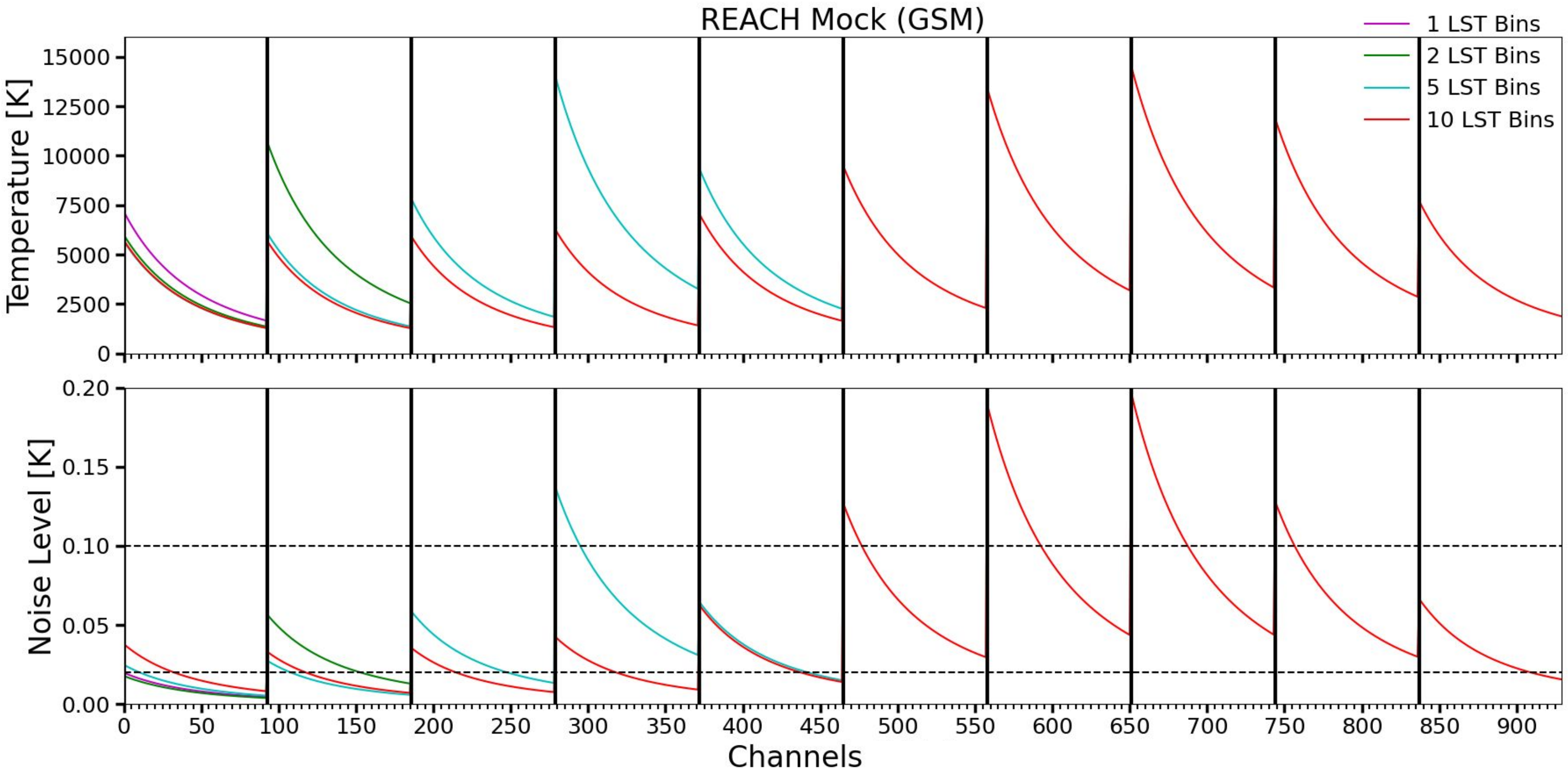
$$\sigma_n^2 = \frac{T_n^2(\nu)}{D_w},$$

Radiometer noise

joshua.hibbard@colorado.edu

Mock Spectra

Figures from Hibbard et al. 2023, under review



joshua.hibbard@colorado.edu

Phenomenological models

- Add more terms to increase complexity
- Good at modelling high-order ripple effects (unknowns)
- Difficult to include time-dependence, other Stokes parameters
- Unclear if they capture all kinds of systematics effectively without forward-modelling.

Forward Models

- Require simulations of the beam, intrinsic FG, local environment, etc.
- Can be computationally expensive, requiring physics simulations and models of all effects
- Can include time dependence, other Stokes parameters
- Easy to model and include other “physical” effects
- Break degeneracies with signal model

joshua.hibbard@colorado.edu

Phenomenological

Polynomial

$$\left(\frac{\nu}{\nu_0}\right)^{-2.5}\sum_{k=0}^{N_{py}} a_k \left[\ln\frac{\nu}{\nu_0}\right]^k \quad (10)$$

$$\left(\frac{\nu}{\nu_0}\right)^{-2.5}\sum_{k=0}^{N_{py}} a_k \left(\frac{\nu}{\nu_0}\right)^k \quad (11)$$

$$\left(\frac{\nu}{\nu_0}\right)^{-2.5}\sum_{k=0}^{2} a_k \left(\ln\frac{\nu}{\nu_0}\right)^k + a_3\left(\frac{\nu}{\nu_0}\right)^{-4.5} + a_4\left(\frac{\nu}{\nu_0}\right)^{-2} \quad (12)$$

(Pritchard and Loeb 2012, Bernardi et al. 2016, Mozdzen et al. 2016, Bowman et al. 2018, Singh et al. 2021, etc.)

- Polynomials, modulated by a constant power-law representing the "average" galactic spectral index.
- 5-6 terms, traditionally.

Maximally-Smooth

$$\frac{d^m \mathcal{M}_{MDP}}{d\nu^m} \leq 0$$

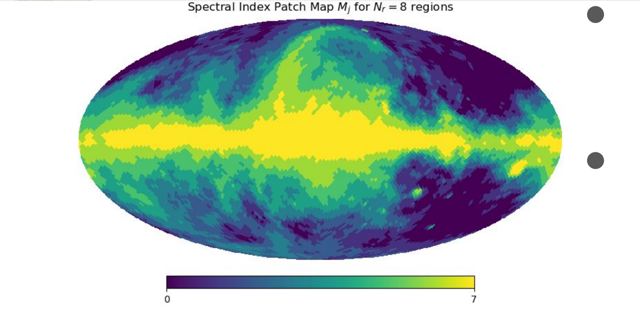
$$\frac{d^m \log_{10}(\mathcal{M}_{MLLP})}{d\log_{10}(\nu)^m} \leq 0$$

(Rao et al. 2017, Bevins et al. 2021, Singh et al. 2021)

- Polynomials with derivatives set to have no inflections or "ripples."
- Supposed to account for smoothly varying foregrounds without accidentally picking up any signal power.

Forward-Models

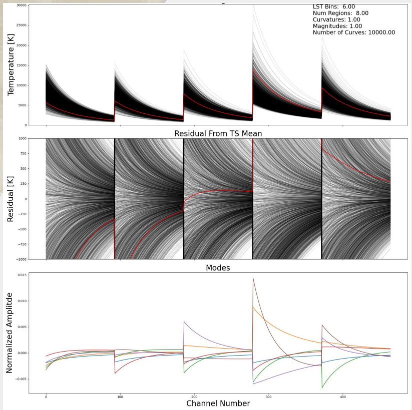
Nonlinear



(Anstey et al. 2021, etc., REACH Collab.)

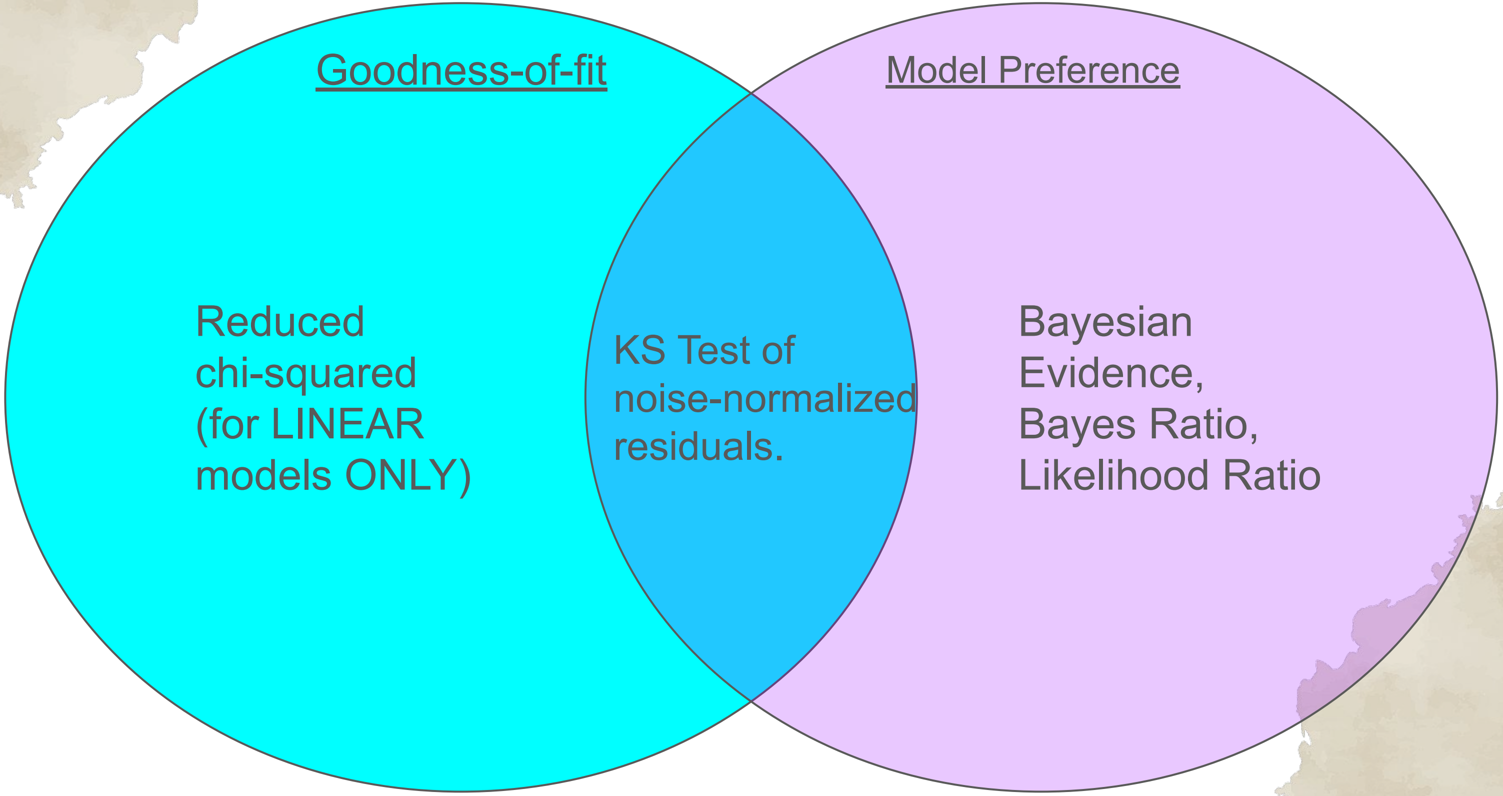
- Split sky into N_r equal-perce ntile regions
- Assign index, magnitude, and curvature to each region

Linear

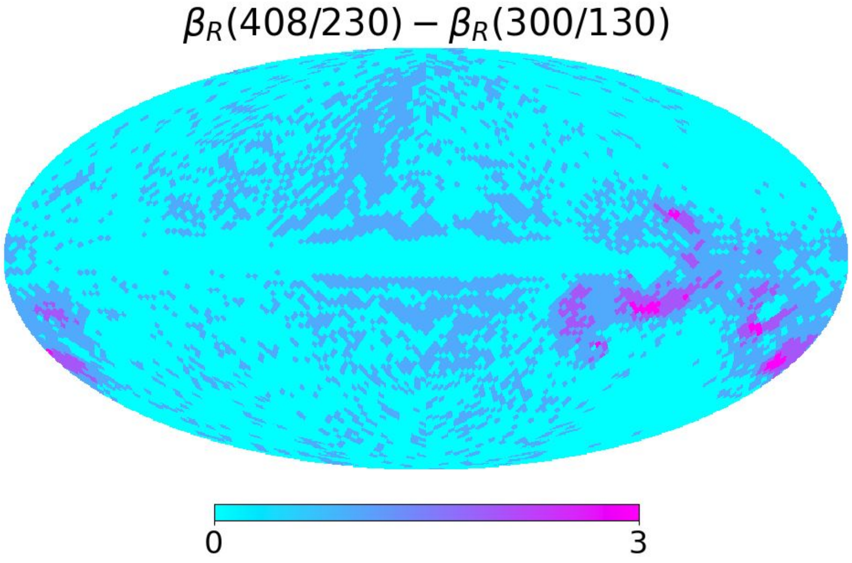


- Generate ten thousand example spectra using nonlinear forward-model and large parameter ranges
- Eigenmodes (spectra) from SVD.
- "...ought to rebrand as SBI, because that's what it is..." - W. Handley

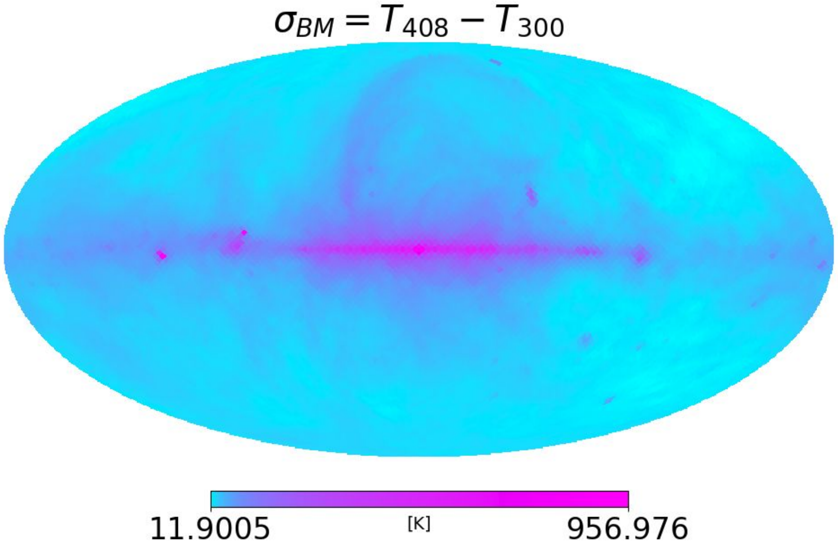
(Pylinex paper series, Tauscher et al. 2018, etc.)


Goodness-of-fit
Model Preference
Reduced chi-squared (for LINEAR models ONLY)
KS Test of noise-normalized residuals.
Bayesian Evidence, Bayes Ratio, Likelihood Ratio

Testing Model Spatial Inputs (Maps)



Spatial “Error” in input spectral index patch map (PM).

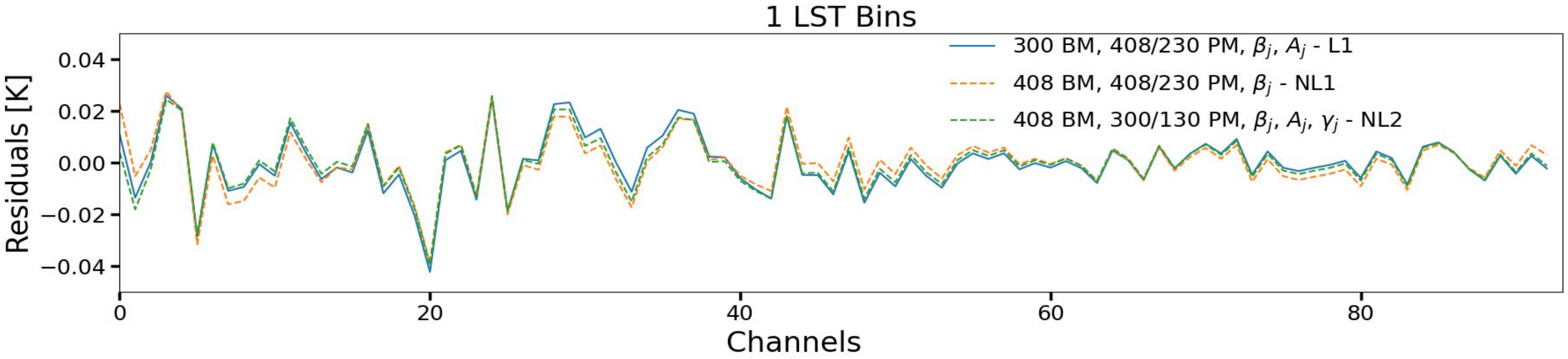


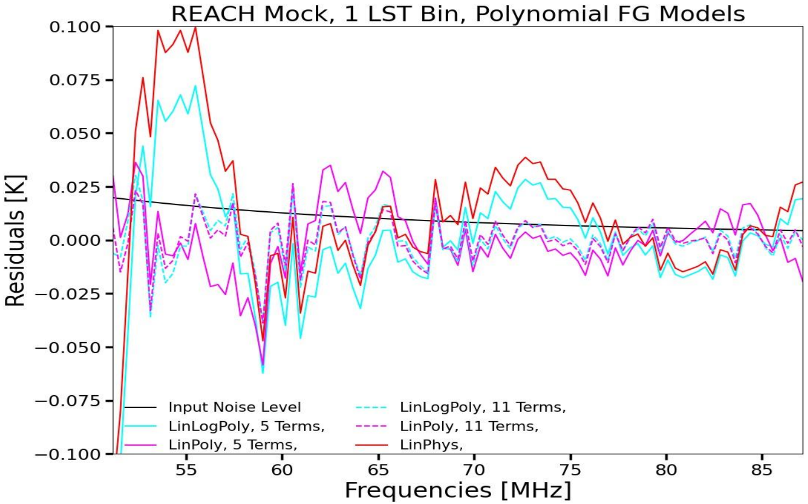
Spatial “Error” in input base temperature map (BM).

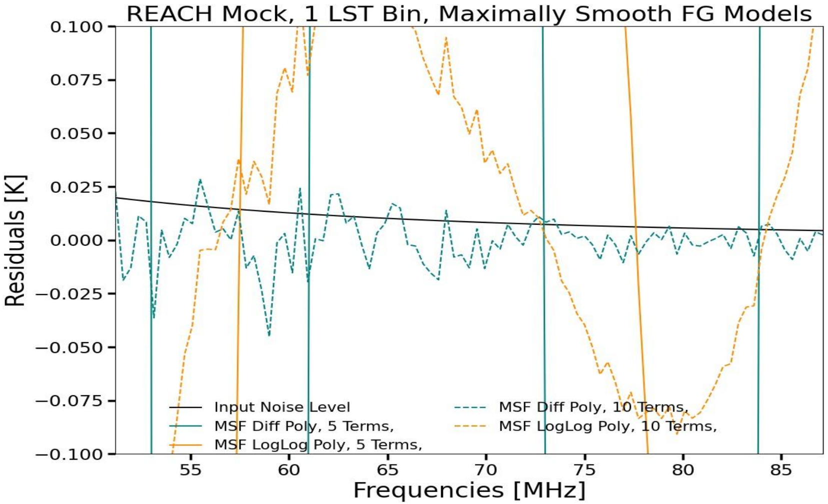
joshua.hibbard@colorado.edu

Comparison - 1 LST Bin

Figures from Hibbard et al. 2023, under review







joshua.hibbard@colorado.edu

1 LST Bin–Statistics
Gold - Pass KS Test, Gray - Fail

Tables from Hibbard et al. 2023, under review

Forward Models

	Model		Linear							Nonlinear					
LSTs	Input	θ_j	N_r	χ^2_{red}	σ	$\ln Z$	N_x	p_{ks}	BF	N_r	χ^2_{red}	$\ln Z$	N_θ	p_{ks}	BF
	IDEAL	β_j	4	0.96	1	-615.71	7	0.94	–	4	1.02	-81.05	4	0.99	NL1
		β_j, A_j	3	1.14	1	-625.24	6	0.97	–	4	1.02	-105.78	8	0.97	–
		β_j, A_j, γ_j	2	1.14	1	-628.66	6	0.90	–	4	1.18	-119.07	12	0.88	–

Polynomial

Model	N_{py}	χ^2_{red}	σ	RMS (mK)	p_{ks}
LinLogPoly	5	6.2	34	31	9.3e-6
	6	1.44	3	13	0.40
	11	1.11	1	11	0.94
LinPoly	5	2.85	12	16	0.04
	6	1.60	4	14	0.33
	11	1.1	1	10	0.95
LinPhys	7	9.04	53	36	1.87e-11

MSF

Model	N_{MSF}	χ^2_{red}	RMS (mK)	p_{ks}
Diff Poly	5	5e5	6e3	1.4e-24
	6	2.81e4	1.36e3	1.56e-21
	10	1.28	12	0.76
	15	1.41	11	0.58
LogLog Poly	5	2.9e4	2.1e4	8.9e-27
	6	3.1e4	2.2e3	8.9e-27
	10	110	103	1.4e-17
	15	3.92	23	0.008

joshua.hibbard@colorado.edu

1 LST Bin - Incorrect Map Inputs?

1	300 BM	β_j	4	0.96	1	-614.26	7	0.93	–	4	1754	-7.8e4	4	1.36e-24	–
		β_j, A_j	3	1.10	1	-623.23	6	0.99	L1	4	1.02	-107.39	8	0.98	–
		β_j, A_j, γ_j	2	1.12	1	-626.91	6	0.95	–	4	1.10	-98.22	12	0.93	–
	(300/130) PM	β_j	4	0.99	1	-613.8	5	0.998	–	4	3.9e10	-1.74e12	4	~ 0	–
		β_j, A_j	3	1.06	1	-618.83	8	0.998	–	4	1.03	-112.26	8	0.94	–
		β_j, A_j, γ_j	2	1.09	1	-623.53	6	0.88	–	4	1.09	-112.71	12	0.999	NL2

Tables from Hibbard et al. 2023, under review

- BM → Wrong Input Base Map to Mock spectra
- PM → Wrong Input Patch Map to Mock spectra

joshua.hibbard@colorado.edu

Comparison - Multiple LST Bins

2	IDEAL	β_j	5	1.00	1	-1129.88	9	0.68	–	8	1.13	-154.24	8	0.81	–
		β_j, A_j	4	1.02	1	-1160.86	13	0.81	–	8	1.03	-335.72	16	0.83	NL3
		β_j, A_j, γ_j	15	1.03	1	-1192.70	18	0.80	–	8	1.53	-307.33	24	0.09	–
	300 BM	β_j	5	1.03	1	-1130.74	13	0.51	–	8	7.83e4	-6.96e6	8	5.6e-45	–
		β_j, A_j	4	1.02	1	-1155.56	13	0.89	L2	8	1.15	-364.17	16	0.82	–
		β_j, A_j, γ_j	4	1.08	1	-1189.71	14	0.78	–	8	1.27	-234.77	24	0.50	–
	(300/130) PM	β_j	6	1.13	1	-1134.67	21	0.71	–	8	2.1e10	-1.82e12	8	~ 0	–
		β_j, A_j	4	1.00	1	-1142.86	13	0.97	L3	8	3.43	-500.73	16	0.002	–
		β_j, A_j, γ_j	4	1.04	1	-1162.65	14	0.96	–	8	1.34	-220.24	24	0.26	–

Tables from Hibbard et al. 2023, under review

- Nonlinear model performs well for Ideal case, but begins to break down.

joshua.hibbard@colorado.edu

Comparison - Multiple LST Bins

10	IDEAL	β_j	14	1.08	2	-4833.53	36	0.80	–	9	477.92	-2.20e5	9	1.4e-198	–
		β_j, A_j	13	2.58	33	-6783.17	46	1.91e-8	–	9	64.25	-2.9e4	18	7.1e-148	–
		β_j, A_j, γ_j	19	1.04	1	-5059.1	58	0.82	L6	9	27.6	-1.29e4	27	2.5e-107	–
	300 BM	β_j	12	1.40	8	-5004.78	39	0.03	–	9	76973	-3.54e7	9	2.4e-252	–
		β_j, A_j	13	2.39	29	-6696.14	46	1.3e-7	–	9	32.01	-1.48e4	18	1.7e-122	–
		β_j, A_j, γ_j	19	1.07	2	-5183.84	56	0.94	L7	9	19.66	-9134.45	27	8.6e-80	–
	(300/130) PM	β_j	17	2.34	28	-6408.78	54	2.7e-9	–	9	4.0e9	-1.84e12	9	~ 0	–
		β_j, A_j	20	2.02	5	-6060.94	73	1.0e-4	–	9	41717	-1.90e7	18	2.0e-251	–
		β_j, A_j, γ_j	20	1.34	7	-5104.81	79	0.12	L8	9	18109	-8.18e6	27	6.6e-223	–

Tables from Hibbard et al. 2023, under review

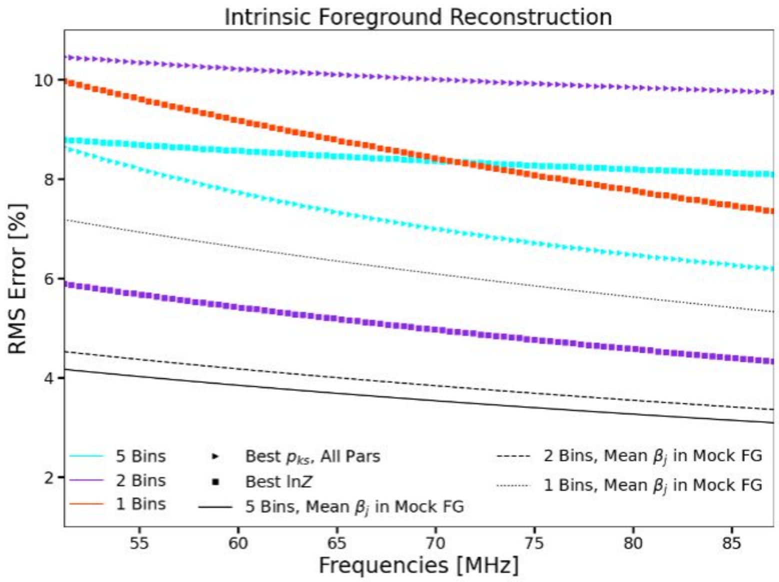
joshua.hibbard@colorado.edu

Conclusions

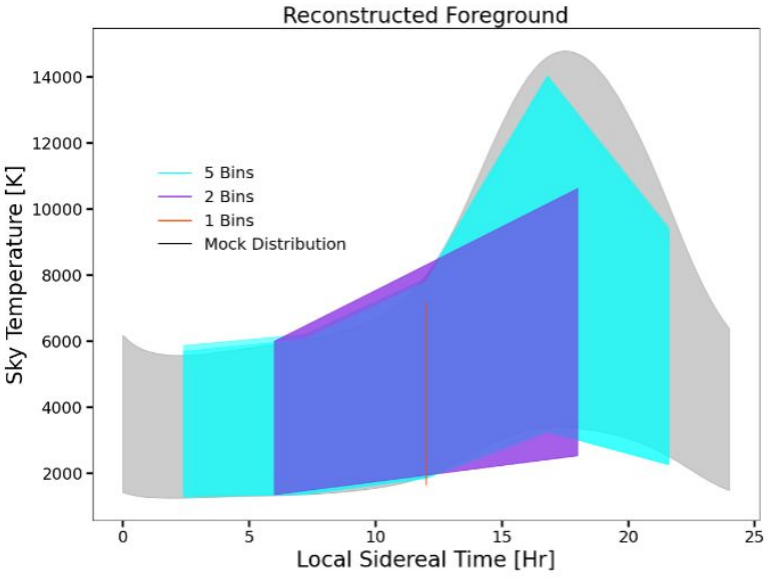
- **Inadequate FG Models are a significant source of bias and error in 21-cm signal extractions.**
- For 1 LST bin, the nonlinear forward-model is preferred (slightly)
  - Linear forward-model also works
  - Polynomials and MSFs require >5 parameters, at least, and some don't work at all.
- For multiple LST bin fits, linear forward-model is highly preferred
- KS-test is a robust way of measuring goodness-of-fit and model preference.

joshua.hibbard@colorado.edu

Nonlinear FG Reconstruction?



(a) RMS Error of the Best-fit Nonlinear Foreground Models Compared to the Input Intrinsic Foreground.

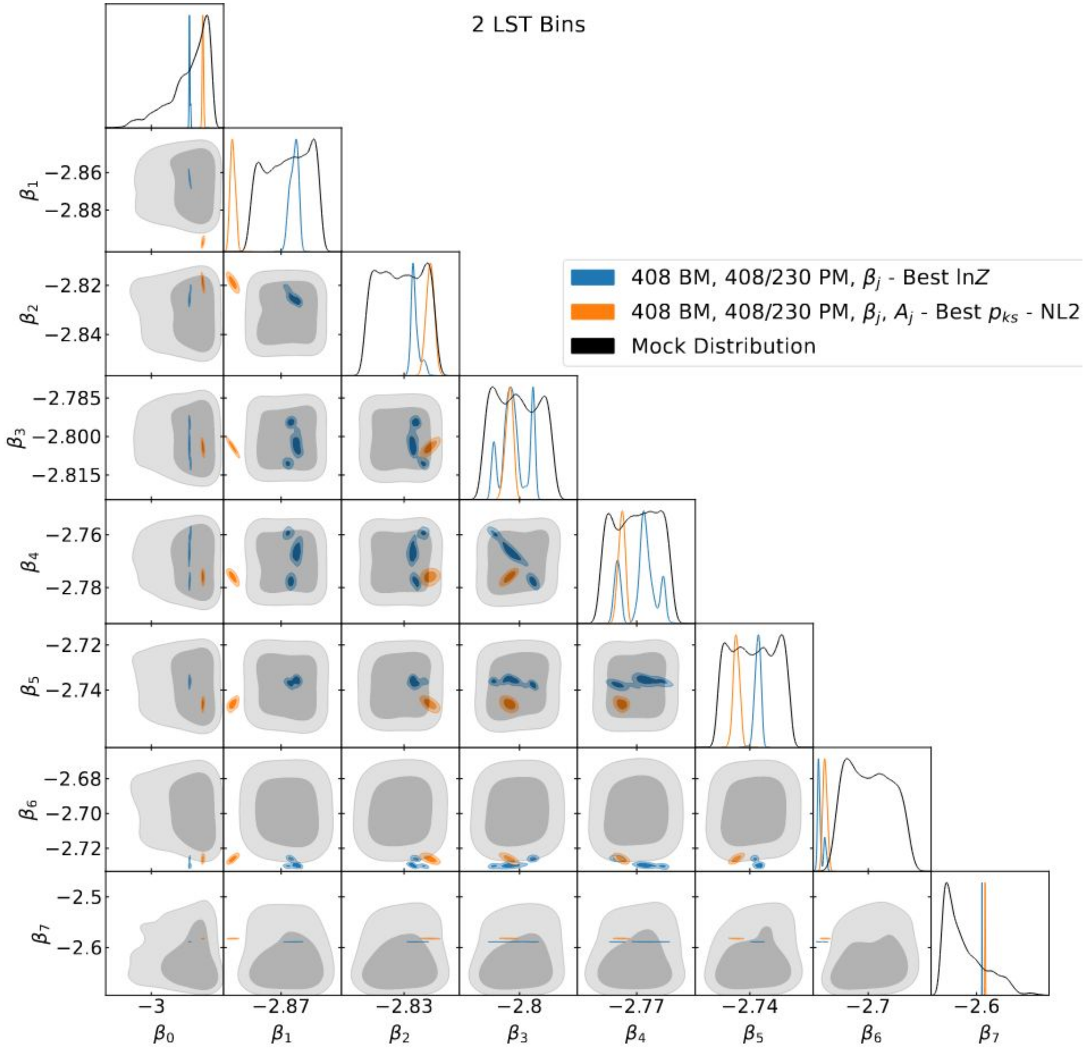


(b) Reconstructed foreground as a function of LST.

Figure from Hibbard et al. 2023, in prep

Even with the corruption of the beam, we can still recover Intrinsic FG's with less than ~10 % error across the band.

joshua.hibbard@colorado.edu



joshua.hibbard@colorado.edu