Making EDGES Bayesian

Steven Murray, for the **EDGES** team.

21cm Global Workshop, Montreal, Oct 2019.

Outline

- A Bayesian computational framework for EDGES
- Dual-Band analysis
- Bayesian Calibration

Overall Motivation

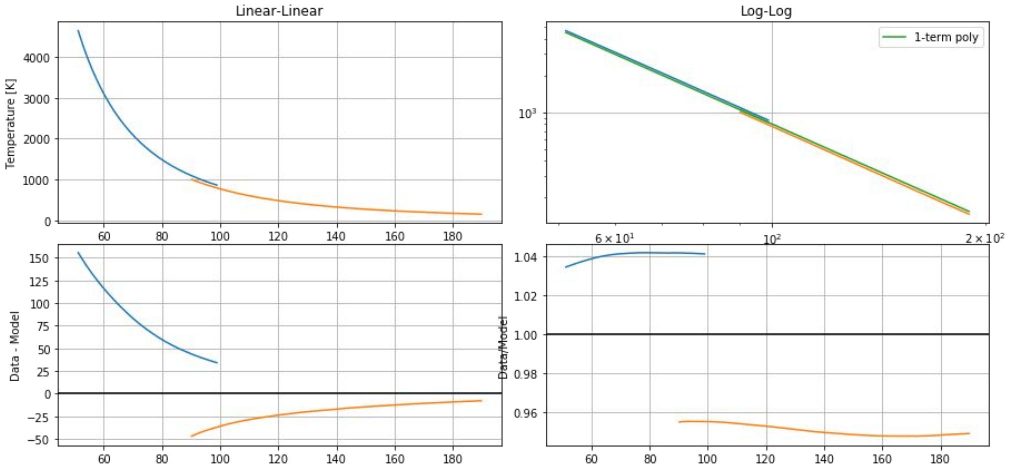
- EDGES 2018 result is surprising.
- Rigorous confirmation of derived absorption feature is very important.
- There are a number of ways in which the data analysis component of the EDGES results can be improved (cf. HIlls et al. 2019, Singh+Subrahmanyan 2019 etc.).

A Bayesian Model-Maker for EDGES

- New Bayesian models (likelihoods, priors, derived parameters…) implemented using https://github.com/steven-murray/yabf
- Defines a Python ←→ YAML "standard" for specifying computable hierarchical Bayesian models.
- All current models for EDGES implemented in https://github.com/edges-collab/edges_estimate

Dual Band Analysis

- Desire to use multiple antennas' data in simultaneous constraint.
- High-band data could possibly provide better handle on foregrounds.
- Difficulty arises due to differences in absolute calibration between instruments



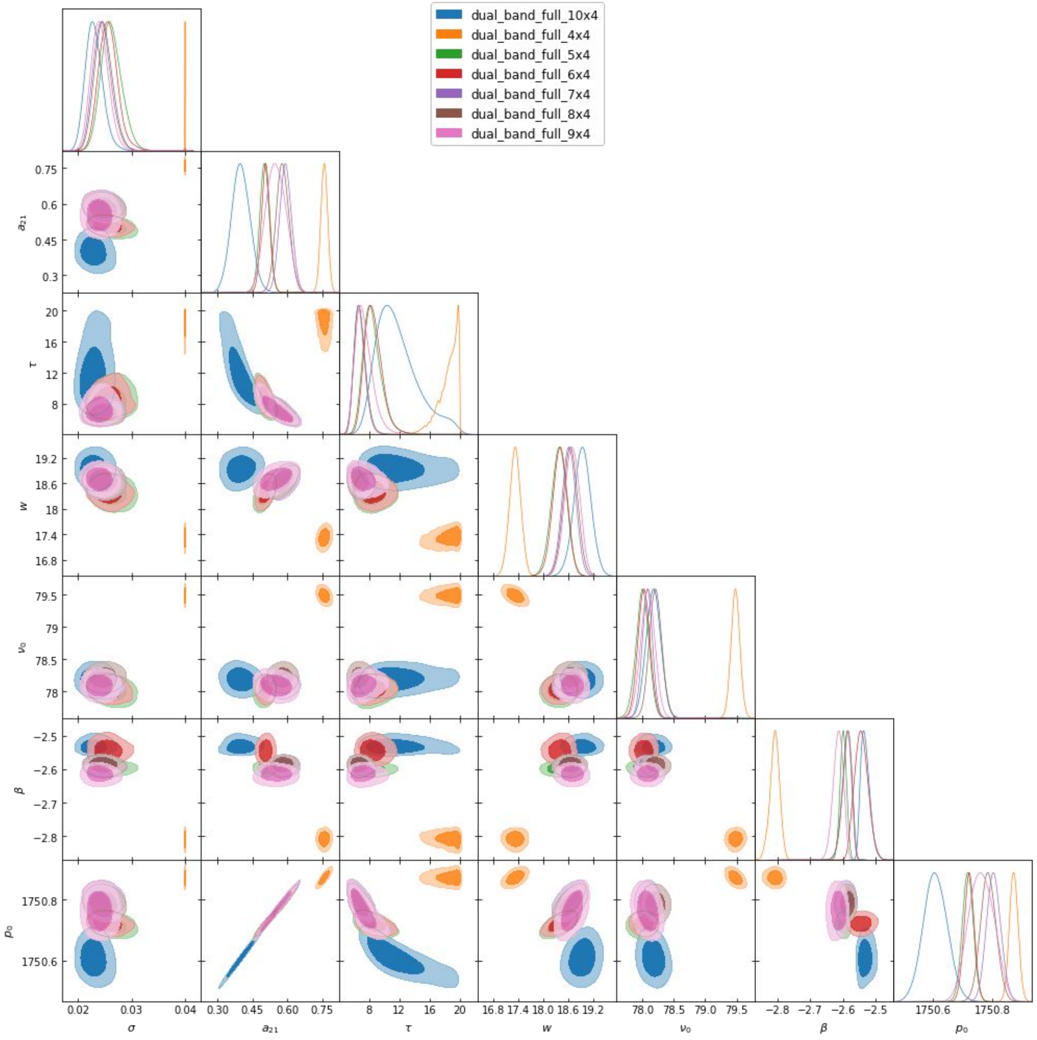
Dual Band Analysis

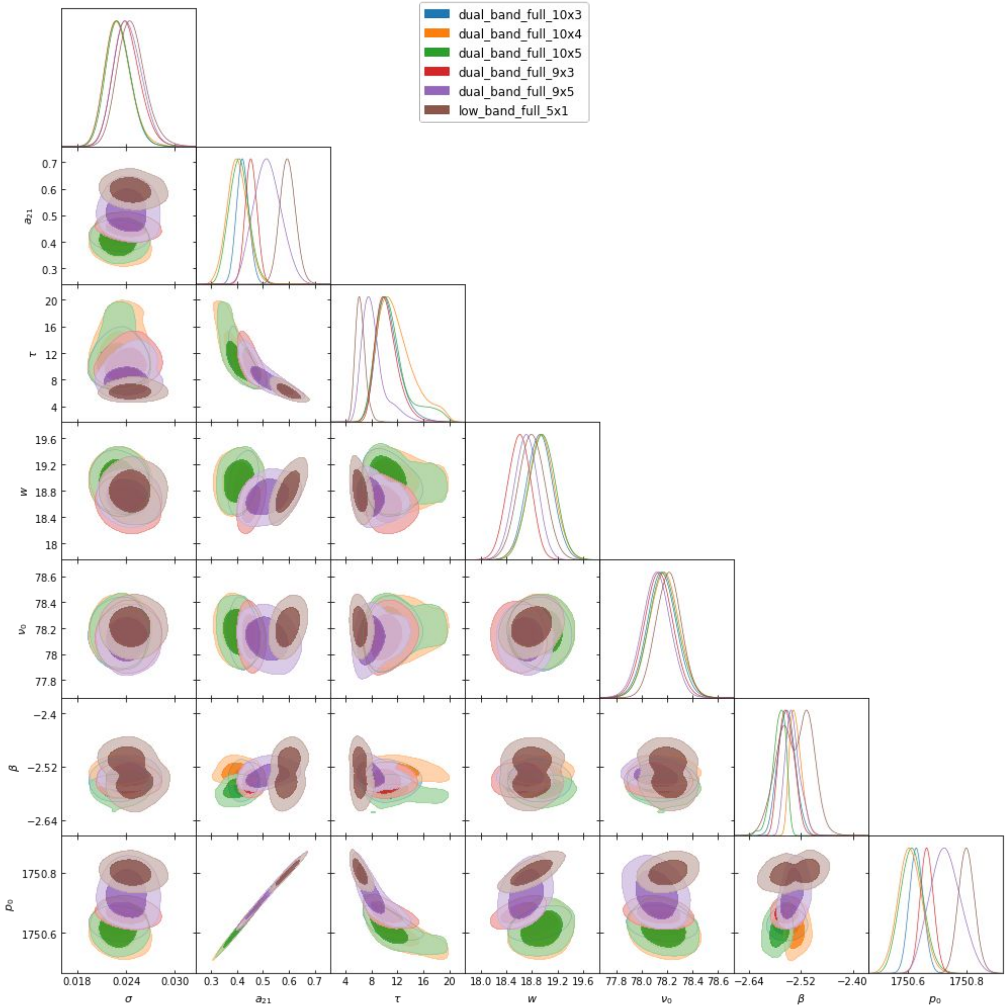
- We use a flexible model to capture the relative miscalibration:

$$T_{\text{high}}(\nu) = f^{\beta} \sum_{i=0}^{N_{\text{fg}}-1} p_i (\ln f)^i \sum_{j=0}^{N_{bias}-1} b_j (\ln f)^j$$

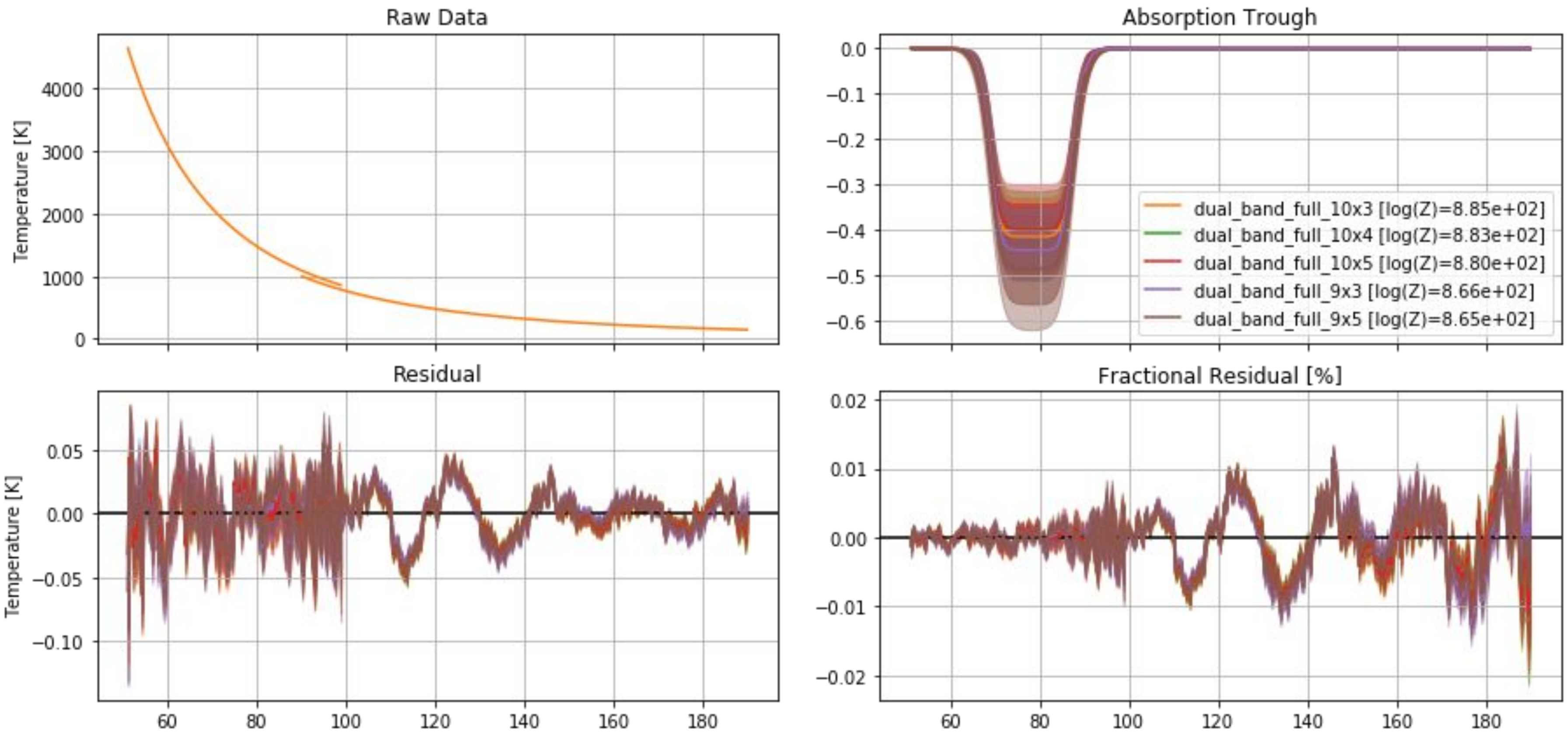
Equivalent to Low Band

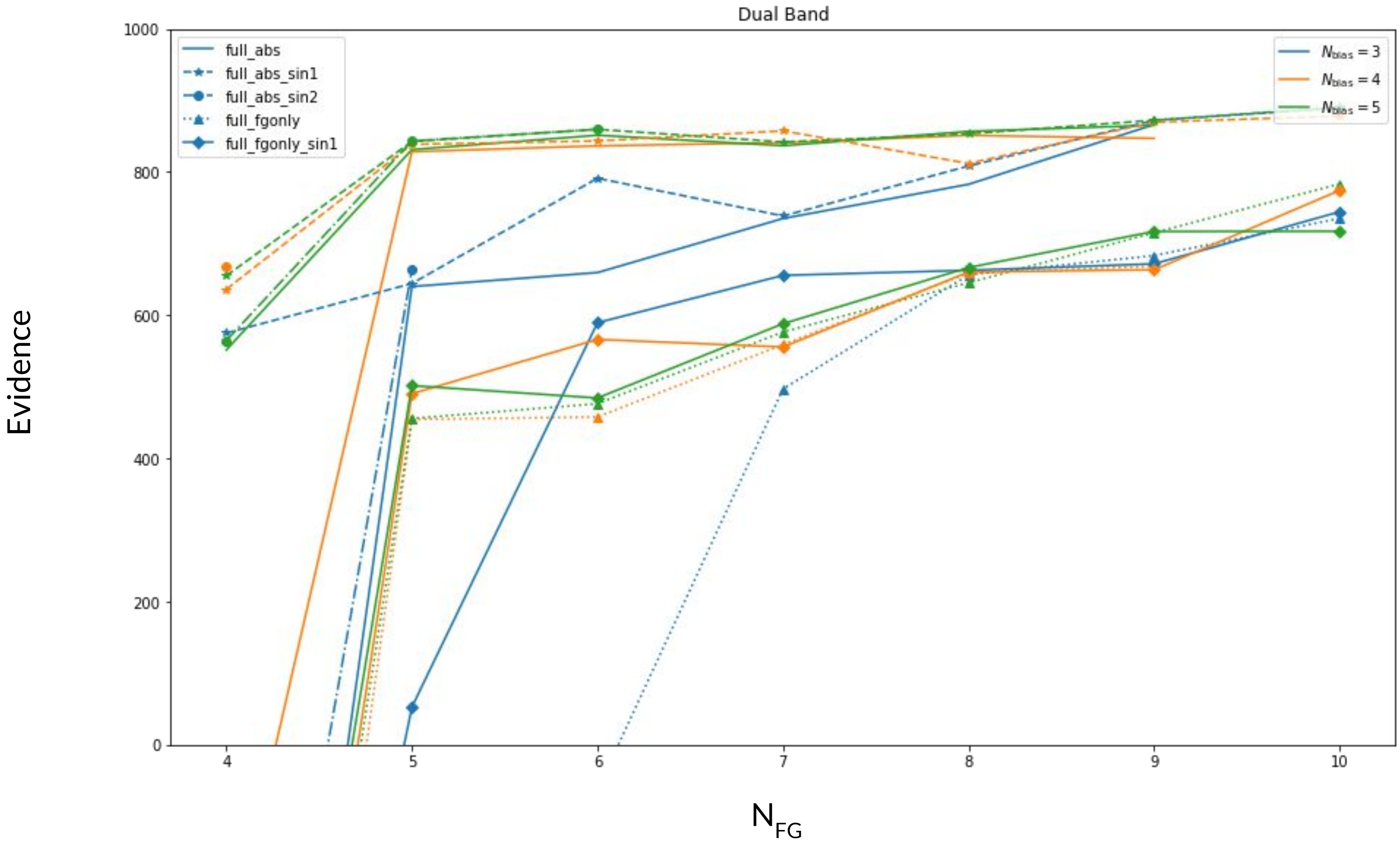
High Band relative “bias”


dual_band_full_10x4
dual_band_full_4x4
dual_band_full_5x4
dual_band_full_6x4
dual_band_full_7x4
dual_band_full_8x4
dual_band_full_9x4
σ
a_{21}
τ
w
ν_0
β
p_0

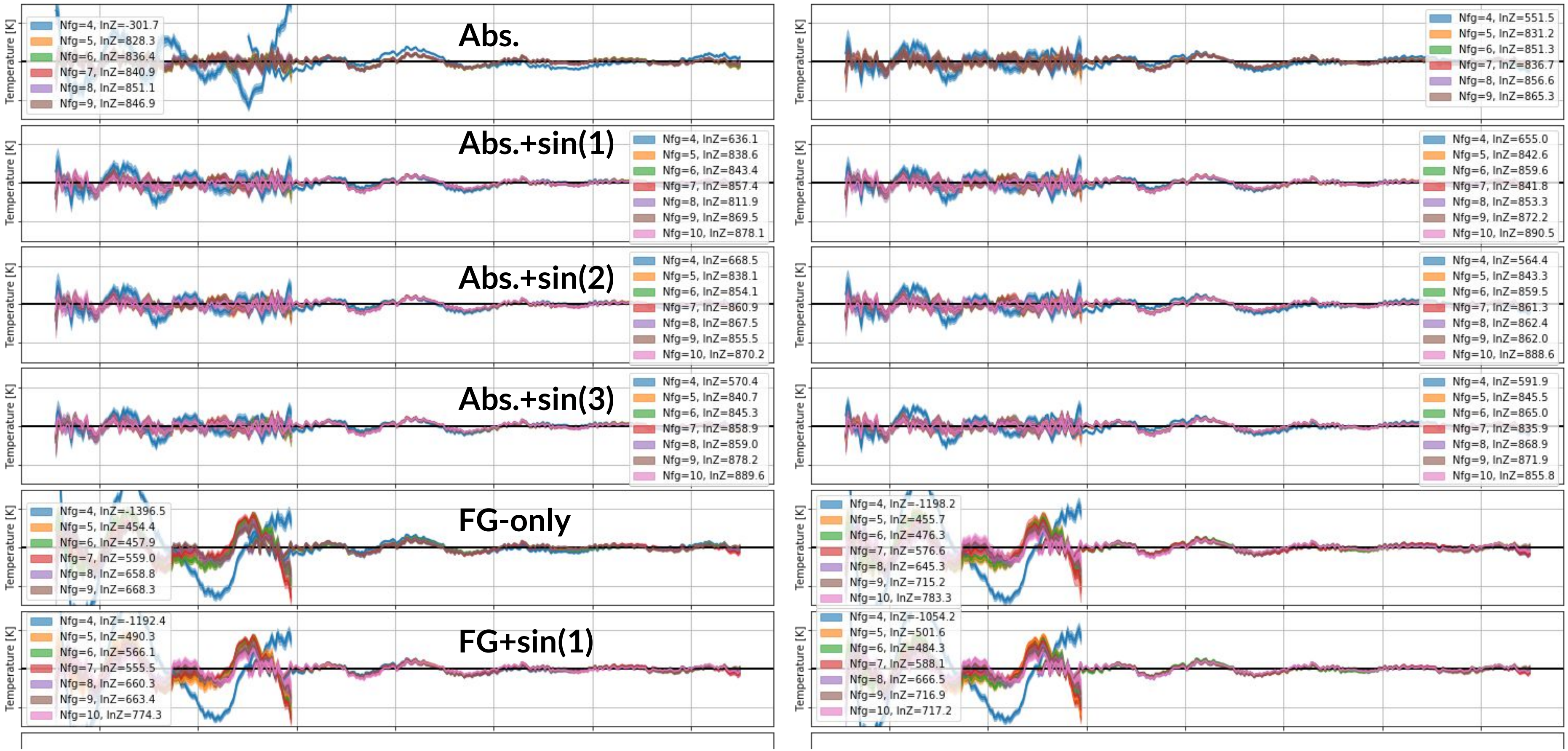


dual_band_full_10x3
dual_band_full_10x4
dual_band_full_10x5
dual_band_full_9x3
dual_band_full_9x5
low_band_full_5x1
σ
a_{21}
τ
w
ν_0
β
p_0
0.018
0.024
0.030
0.3
0.4
0.5
0.6
0.7
4
8
12
16
20
18.0
18.4
18.8
19.2
19.6
18
77.8
78.0
78.2
78.4
78.6
78
−2.64
−2.52
−2.40
−2.4
1750.6
1750.8


Raw Data
Temperature [K]
4000
3000
2000
1000
0
Absorption Trough
0.0
−0.1
−0.2
−0.3
−0.4
−0.5
−0.6
dual_band_full_10x3 [log(Z)=8.85e+02]
dual_band_full_10x4 [log(Z)=8.83e+02]
dual_band_full_10x5 [log(Z)=8.80e+02]
dual_band_full_9x3 [log(Z)=8.66e+02]
dual_band_full_9x5 [log(Z)=8.65e+02]
Residual
Temperature [K]
0.05
0.00
−0.05
−0.10
60
80
100
120
140
160
180
Fractional Residual [%]
0.02
0.01
0.00
−0.01
−0.02
60
80
100
120
140
160
180

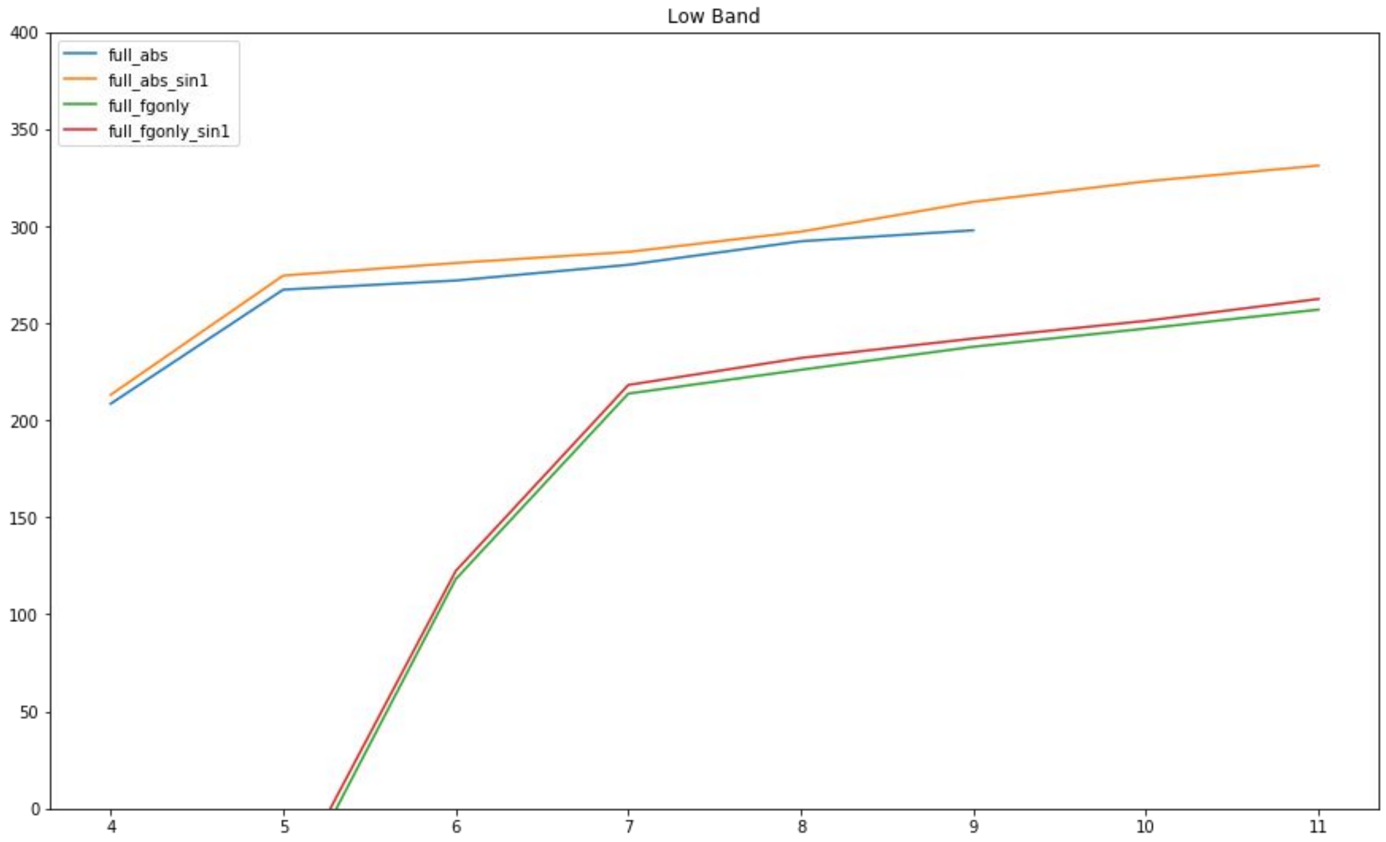


Dual Band
full_abs
full_abs_sin1
full_abs_sin2
full_fgonly
full_fgonly_sin1
$N_{bias} = 3$
$N_{bias} = 4$
$N_{bias} = 5$
1000
800
600
400
200
0
4
5
6
7
8
9
10
Evidence
N_{FG}



High-Band not well-modeled enough to improve constraints!

Low-Band data also shows the same trend



Bayesian Calibration

Motivation:

- Current method uses “best fit” calibration parameters -- no uncertainty.
- Accounts for uncertainties in the calibration, and propagate to final analysis.
- Consistently includes the implicit **covariance** of the calibrated spectrum.
- Can use Bayesian Evidence to compare calibration **and** analysis parameters.

Calibration Framework (Monsalve 2017)

Measured Quantity

Any "source"

$$Q_P \equiv \frac{\mathbf{P}_{\mathrm{ant}} - \mathbf{P}_{\mathrm{L}}}{\mathbf{P}_{\mathrm{L+NS}} - \mathbf{P}_{\mathrm{L}}}$$

Stored every 39 seconds.

Calibration Framework (Monsalve 2017)

$$\mathbf{Q}_P = \frac{1}{C_1 T_{\mathrm{NS}} (1 - |\Gamma_{\mathrm{rec}}|^2)} \Big[T_{\mathrm{ant}} (1 - |\Gamma_{\mathrm{ant}}|^2)|F|^2 + T_{\mathrm{unc}} |\Gamma_{\mathrm{ant}}|^2 |F|^2 + |\Gamma_{\mathrm{ant}}||F| (T_{\cos} \cos\alpha + T_{\sin} \sin\alpha) + (C_2 - T_L) \Big],$$

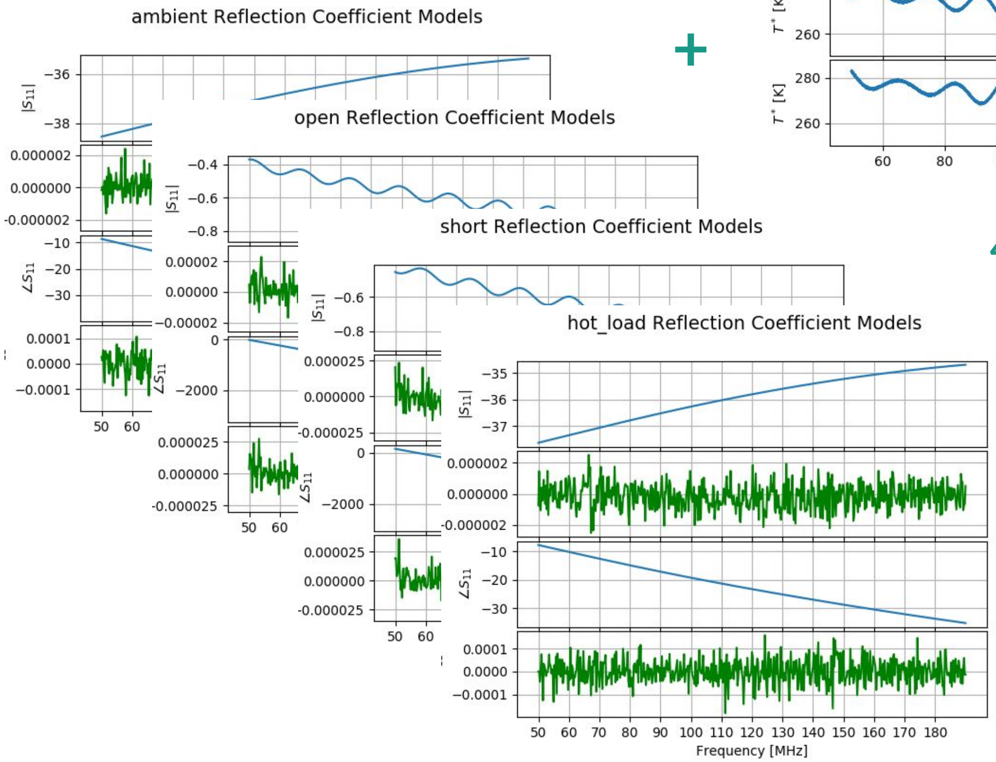
Measured with VNA

Measured with Receiver

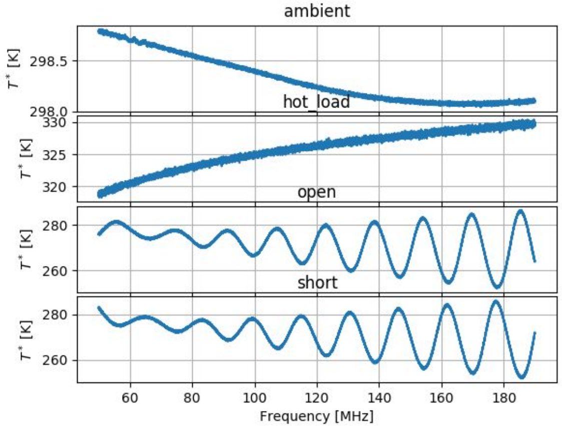
Unknown Calibration Parameters

Unknown

4 S11 Models



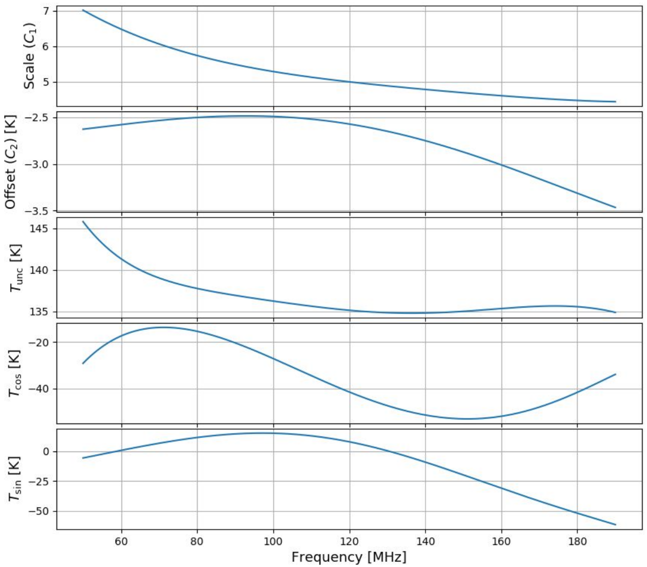
+



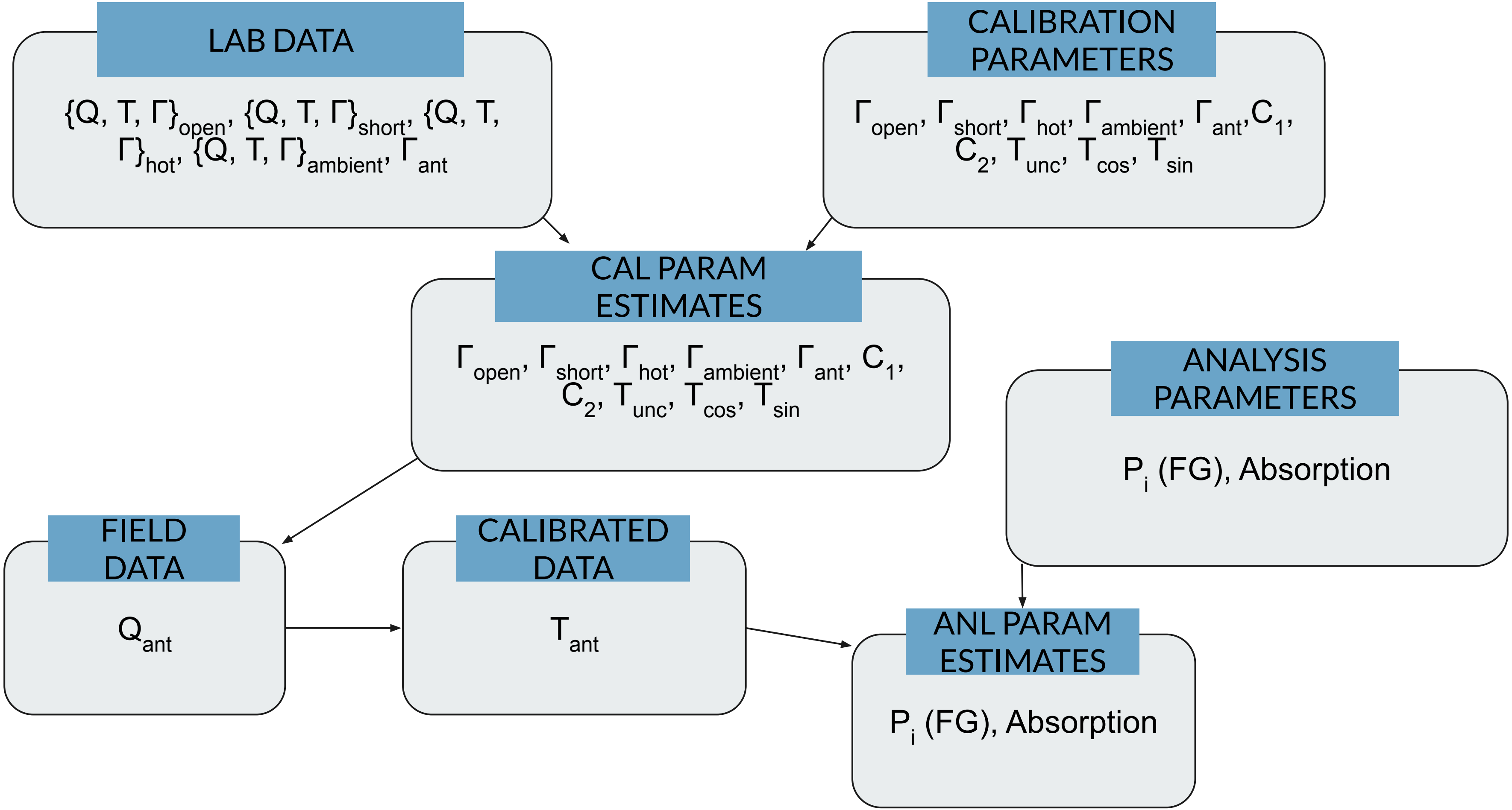
4 Raw Spectra

+ Known Temps

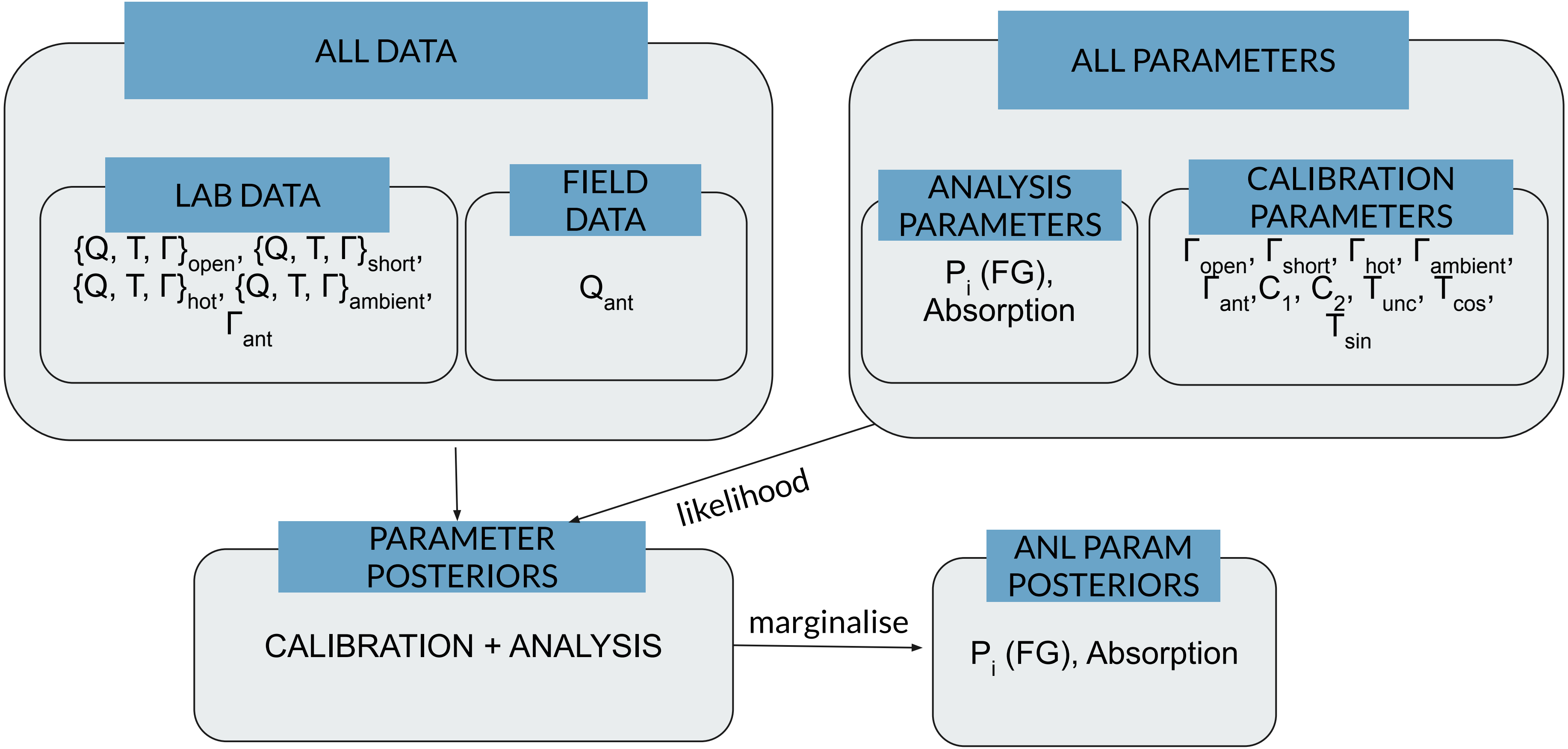
Calibration Solutions



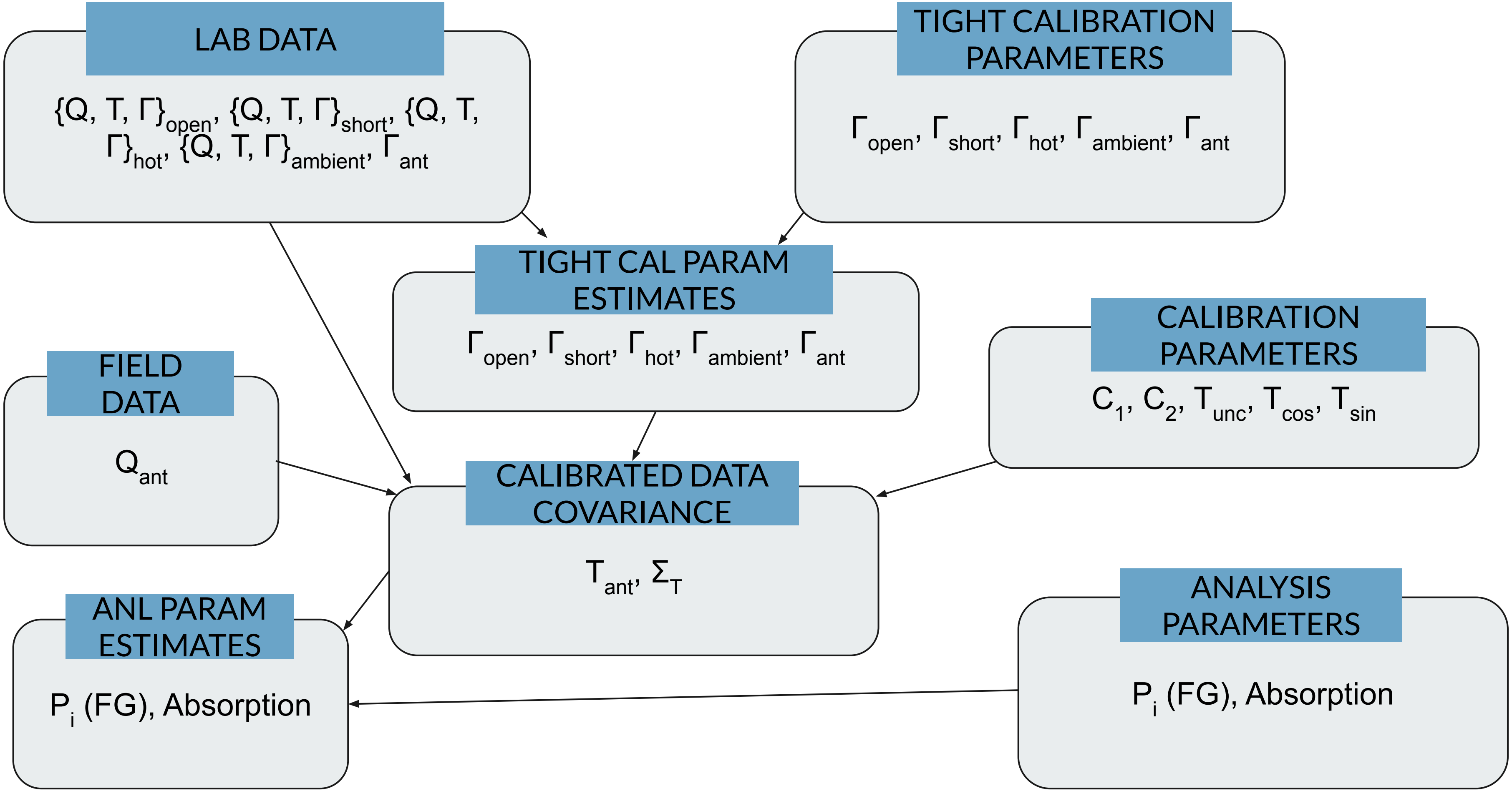
What We Do Now



Fully Bayesian Approach



Conceptually...



Bayesian Calibration Framework

Gaussian likelihood for each "calibrator" relating measured Q_P to model, with extra (optional) simultaneous likelihood for sky data:

$$\ln \mathcal{L} = -\sum_{\nu} \sum_{s=A,H,O,S,sky} \ln \sigma_{s,\nu} + \frac{(\mathbf{Q}_P^{s,\nu} - Q_P^{s,\nu})^2}{2\sigma_{s,\nu}^2}$$

Covariance is diagonal in frequency and sources, but model temperatures are ***not***.

Uncertainty Model

What is the variance, as a function of parameters?

Assume each term is uncorrelated Gaussian with noise given by radiometer equation, use approximate variance of ratio of independent variables, and assume distribution is Gaussian, to find

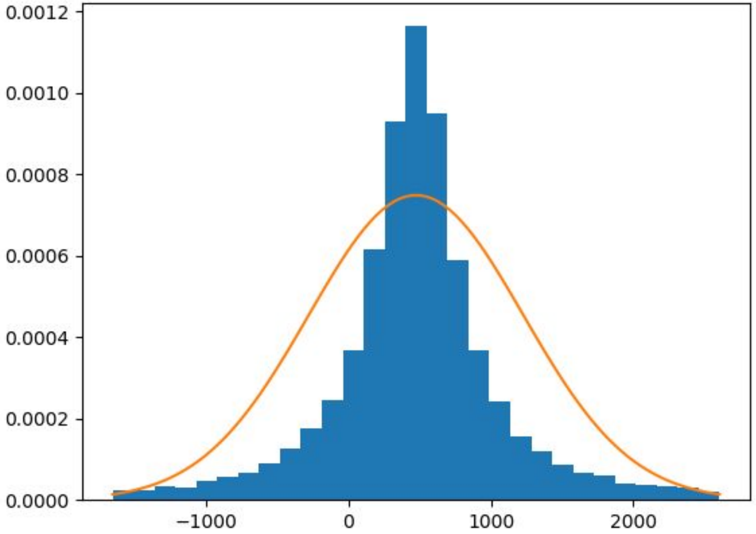
$$\sigma^2 = \frac{Q_P^2}{N_{\mathrm{samp}} \tau \Delta \nu} (1 + \mathcal{Q})$$

$\mathcal{Q} > 0$

$\mathcal{Q} < 1$ if $(T_{\cos}, T_{\sin})$ small.

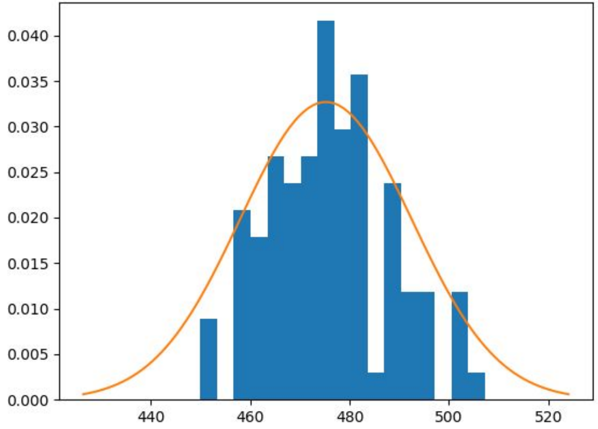
Uncertainty Model

Histogram of Q_P for "ambient" calibrator in one frequency bin.



Non-gaussian distribution means this simple model will be inaccurate. However...

Histogram of **mean** Q_P for "ambient" calibrator in 100 channels.



Sky Model

For simplicity, use parameterized sky model:

$$T'_{\mathrm{sky}} = \int d\theta A(\theta) T_{\mathrm{sky}}(\theta)$$

$$\approx \bar{A} \left(T_{21} + c(\nu, t) T_{\mathrm{FG}} \right)$$

Flattened Gaussian

Chromaticity Correction

Polynomial (eg. LinLog)

Summary

- Progressively moving EDGES analysis to a purely Bayesian framework.
- Simultaneous analysis of multiple bands is not very promising without more robust calibration.
- Work has begun on Bayesian calibration, starting with the most uncertain calibration parameters.
- Future work includes consistent LST-binned analysis.