Validation of Software Pipelines: the HERA Experience

(with an emphasis on simulating and understanding systematic effects)

James Aguirre

University of Pennsylvania

And the HERA Collaboration

Outline

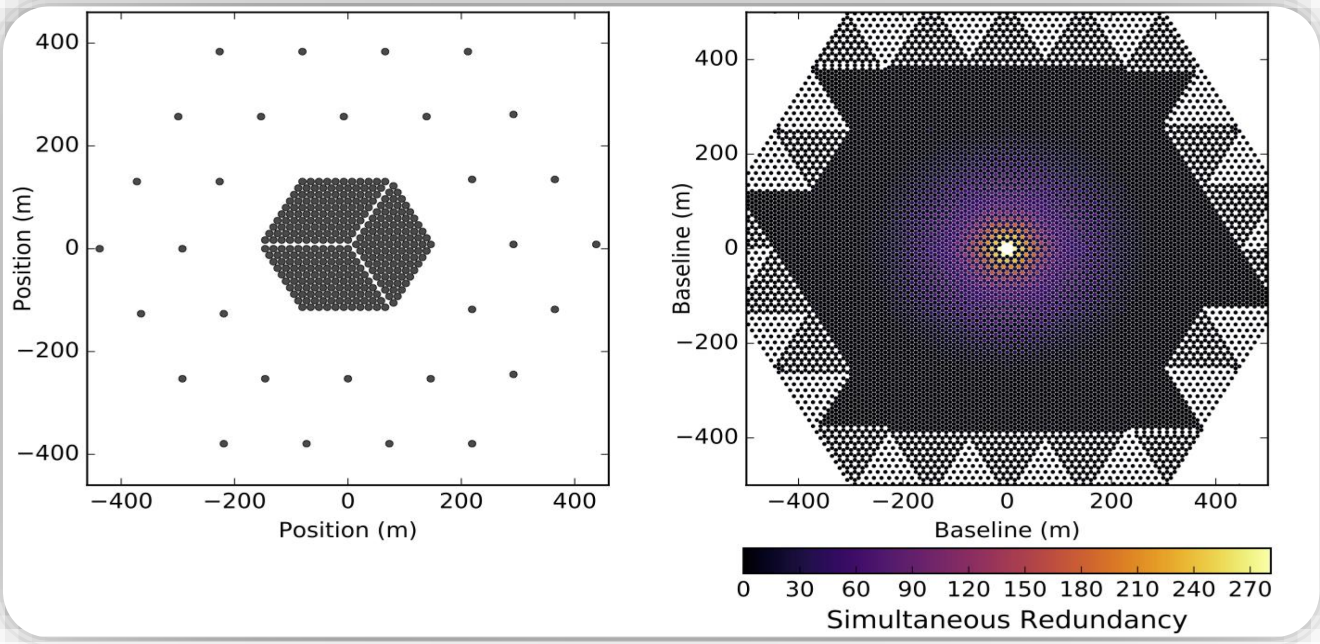
- Brief overview of HERA
- The HERA “validation” (= full instrument simulation) approach
- Simulation: the essentials
- Simulation: what are these systematics of which you speak?
- Open and unaddressed questions
- Some parting thoughts

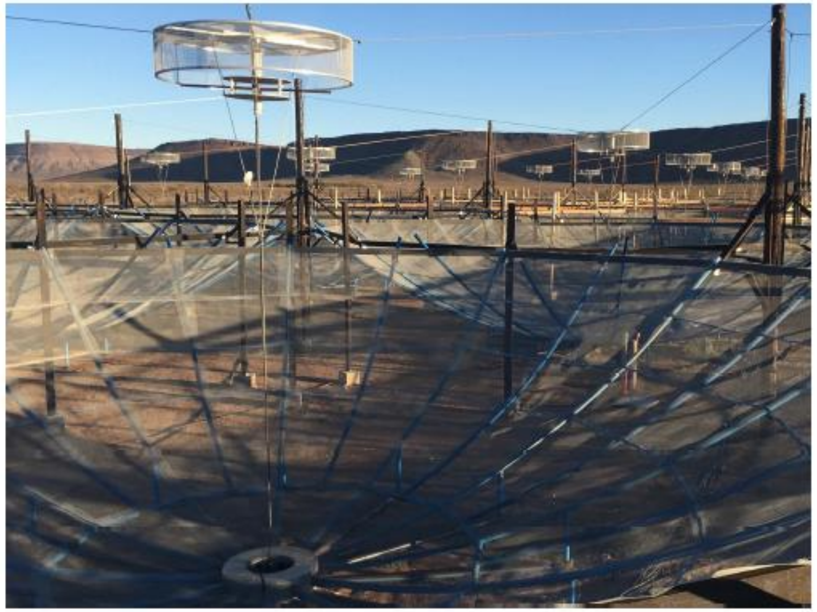
Hydrogen Epoch Reionization Array (HERA)

- A dedicated experiment to observe large scale structures before and during the epoch of reionization, 50 – 250 MHz ($5 < z < 27$)
- Located at the SKA / MeerKAT site in South Africa.
- Second generation instrument. Phase I re-used antenna feeds, analog, and digital electronics from the Precision Array for Probing the Epoch of Reionization (PAPER). Phase II has all new feeds, analog and digital chains
- The 14-m non-tracking (zenith pointed) parabolic antennas in a hexagonally closely packed configuration; core has a maximum baseline length of ~300 m, longest outrigger baselines will be ~1 km
- The array will consist of 350 antennas; about 2/3 complete

Instrument Design Specification	*Observational Performance*
Element Diameter: 14 m	**Field of View:** 9°
Minimium Baseline: 14.6 m	**Largest Scale:** 7.8°
Maximum Core Baseline: 292 m	**Core Synthesized Beam:** 25′
Maximum Outrigger Baseline: 876 m	**Outrigger Synthesized Beam:** 11′
EOR Frequency Band: 100–200 MHz	**Redshift Range:** $6.1 < z < 13.2$
Extended Frequency Range: 50–250 MHz	**Redshift Range:** $4.7 < z < 27.4$
Frequency Resolution: 97.8 kHz	**LoS Comoving Resolution:** 1.7 Mpc (at $z = 8.5$)
Survey Area: $\sim 1440\ \mathrm{deg}^2$	**Comoving Survey Volume:** $\sim 150\ \mathrm{Gpc}^3$
$\mathbf{T_{sys}}$: $100 + 120(\nu/150\ \mathrm{MHz})^{-2.55}$ K	**Sensitivity after 100 hrs:** 50 μJy beam^{-1}

DeBoer et al 2017

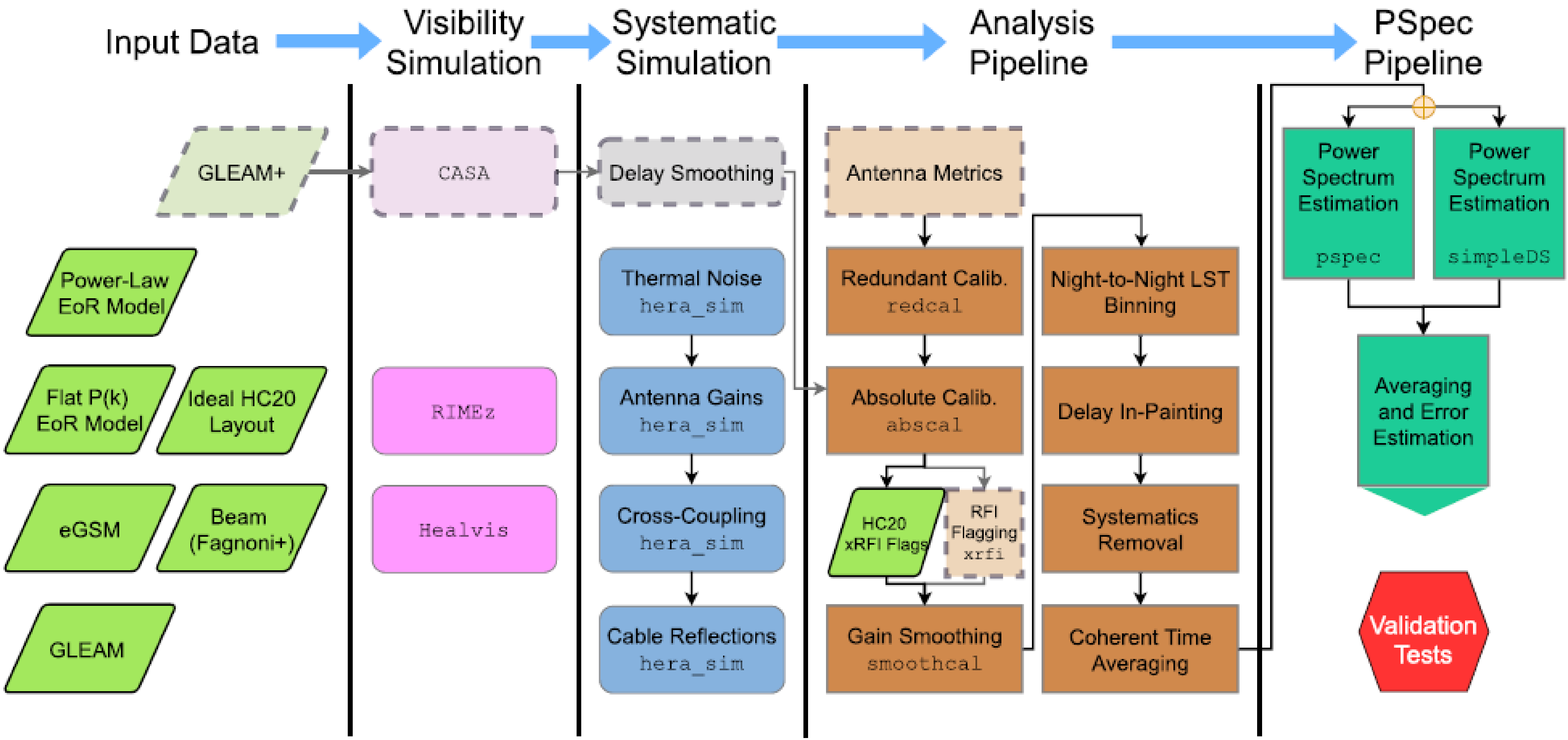




(a)

(b)

Basic outline of end-to-end simulation



Open source and unit-tested code

- Made an effort to make code and tests publicly available
  https://github.com/HERA-Team/hera-validation
- Keep code to high standards and unit test
- Have many users for exercising functionality, finding bugs

NAME	URL (https://github.com/)	DESCRIPTION
PIPELINE		
hera_cal	hera-team/hera_cal	Redundant and sky-based calibration routines.
hera_pspec	hera-team/hera_pspec	Robust foreground-avoidance power spectrum (and covariance) estimation.
SIMULATION		
RIMEz	upenneor/rimez	Fast and accurate visibility calculation implementing multiple methods for different source and beam function definitions.
spin1_beam_model	upenneor/spin1_beam_model	Harmonic space decomposition of the HERA primary beam.
healvis	rasg-affiliates/healvis	Fast visibility simulation based on HEALPIX discretization.
pyuvsim	RadioAstronomySoftwareGroup/pyuvsim	Accurate visibility simulation of point sources with very limited approximations.
gcfg	zacharymartinot/redshifted_gaussian_fields	Consistently simulate cosmological Gaussian fields over the full sky
hera_sim	hera-team/hera_sim	Add HERA-specific instrumental systematics to visibilities.

A.1. Code Standards

HERA has adopted a set of high-standard open-source software practices that encourage transparency, reproducibility, interoperability, and peer-verification. All systems-level HERA code is hosted open-source on a single GitHub organization.[39] A set of well-defined software standards is applicable across the organization, encouraging a certain degree of homogeneity between project-level packages. Among these standards are

1. Documentation: Python code is self-documented (i.e., includes "docstrings"[40] for all public modules, functions, classes, and methods), using a uniform docstring format (typically NUMPYDOC). Extra tutorials and examples are also encouraged.
2. Testing: all systems-level HERA packages are thoroughly unit-tested,[41] and kept at >95% code coverage.[42] Testing is performed continuously via an online Continuous Integration provider (e.g., Travis or Github Actions).
3. Formatting: all code is PEP8 compliant[43] (often enforced by the use of external tools such as BLACK[44] PRE-COMMIT[45]), making each package more homogeneous (important when there are many contributors to the repository) and easy to read. This is important for transparency both within and without the collaboration.
4. Review: each package uses the GitHub flow[46] as a software delivery workflow. In brief, in this workflow the "master"[47] branch is considered protected and is disabled for direct code changes on GitHub. This requires new code additions (and bug fixes) to be developed in a branch that is "not master" and a formal "pull request" (PR) to be created and accepted before merging back into the protected "master" branch. All repositories have an option enabled in which PRs must be first reviewed and

Visibility simulation: some woes

- Seems to be necessary to use custom software for wide-field, fully polarized transit arrays (or at least, not CASA)
- Particular care has to be taken to deal with high dynamic range between foregrounds and EoR, and to ensure no artifacts in delay and fringe-rate transforms
- Actual source catalogs have limited coverage (which can produce artifacts in time (LST)) and depth, there are … idiosyncrasies in the diffuse models (and not much information for $\ell > 100$)
- Simulation of different beam patterns can be very time consuming
- Interesting questions about the horizon
- Full simulation of large arrays is now really a supercomputer problem; we have been optimizing to use GPUs at XSEDE, for example

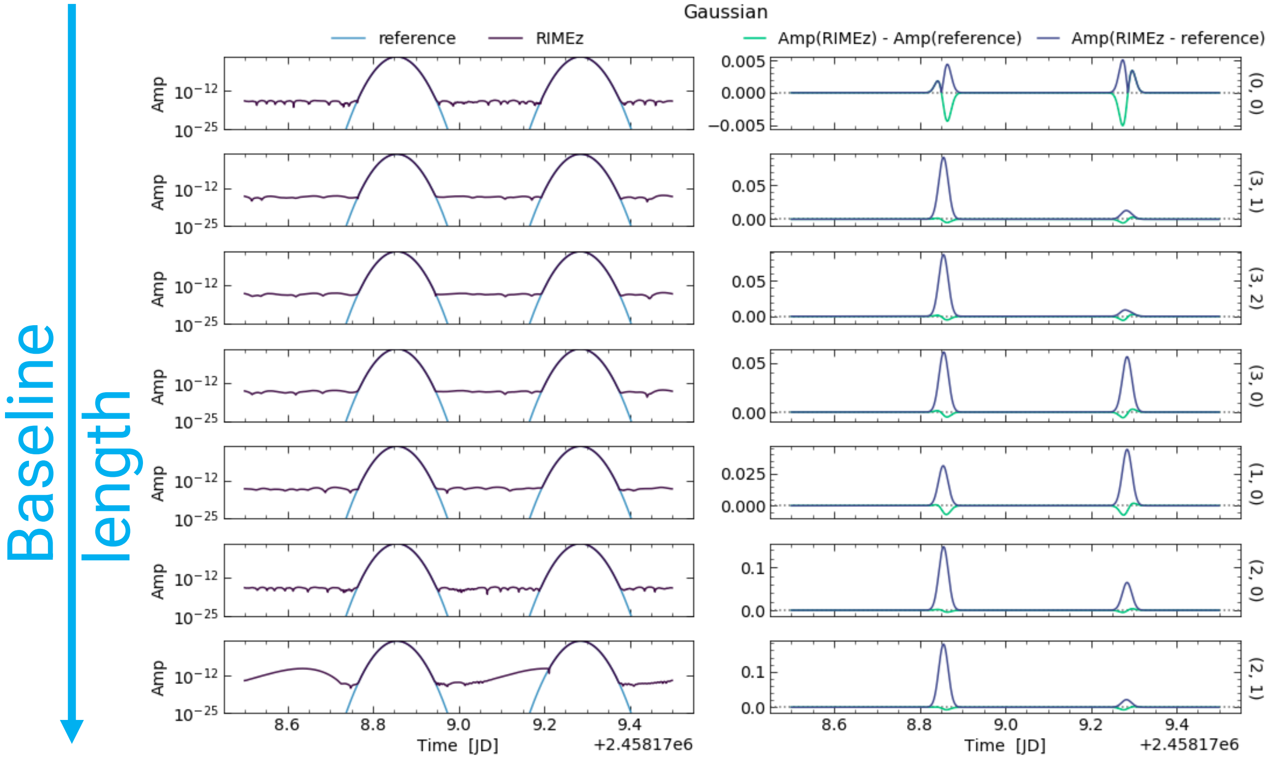
Pyuvsim used as “simulator of reference”

RIMEz: fast m-mode sim by Zachary Martinot (UPenn grad)

Validation: comparison by Lily Whitler (ASU Undergrad)

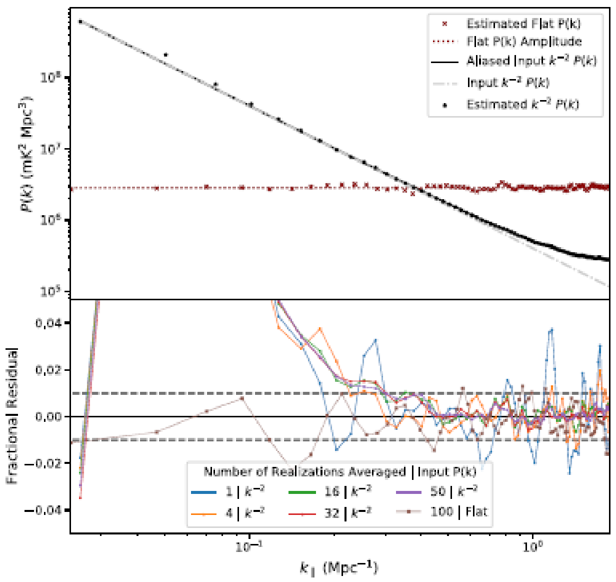
Now using vis_cpu / vis_gpu; paper in the works with extensive tests and comparisons

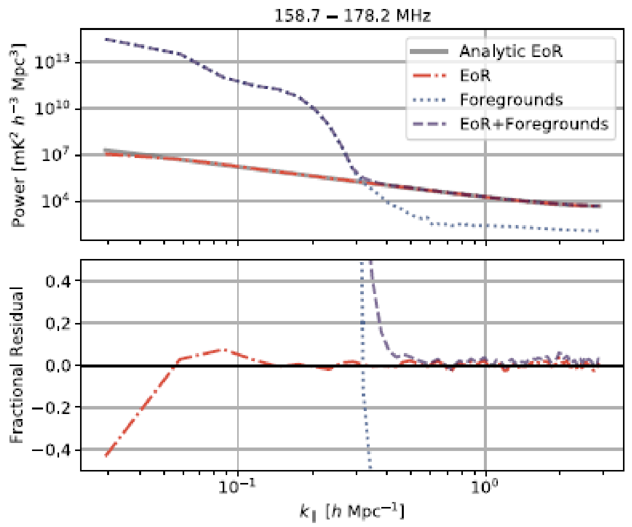
(see [Lanman et al 2019](#) and https://github.com/RadioAstronomySoftwareGroup/pyuvsim)



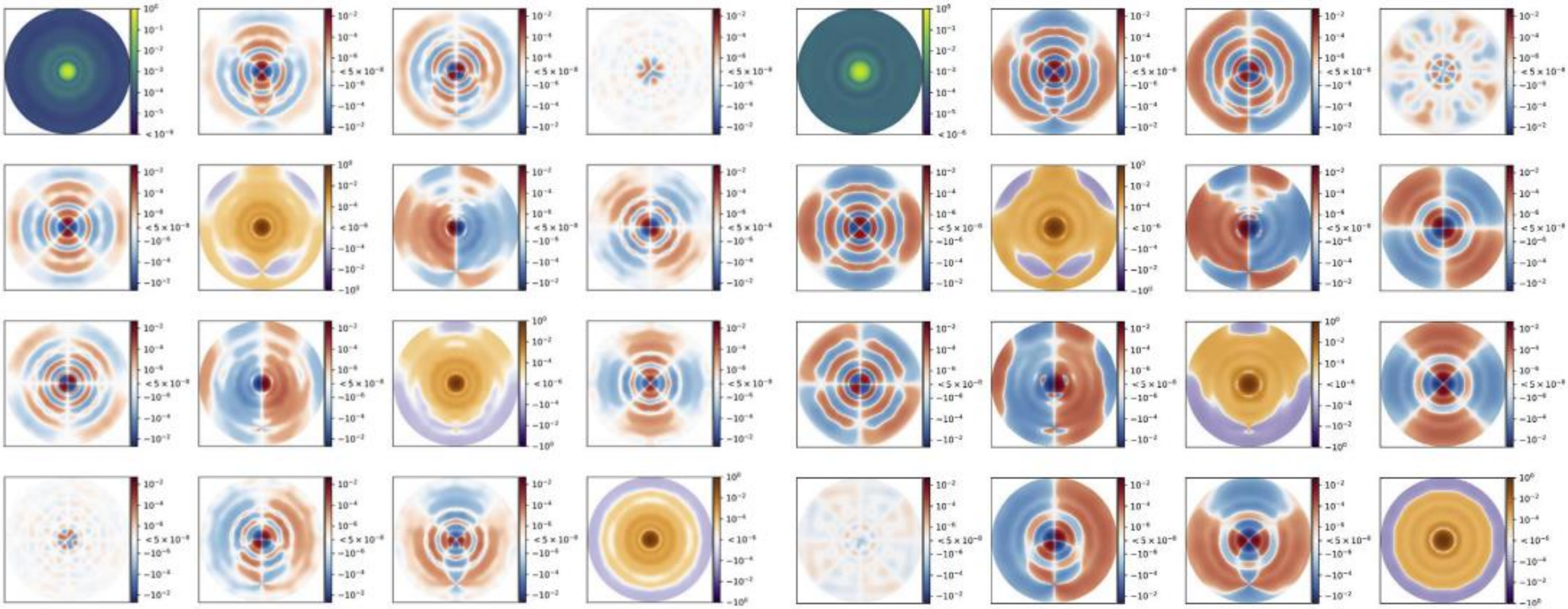
Sky models (see also https://github.com/RadioAstronomySoftwareGroup/pyradiosky)

- Signal: full-sky (gaussian) EoR with specified power spectrum (see Zachary Martinot thesis)
- Diffuse foregrounds: GSM (or variants)
- Point sources:
    - GLEAM (and extensions)
    - Synthetic models with similar dN/dS to cover the whole sky

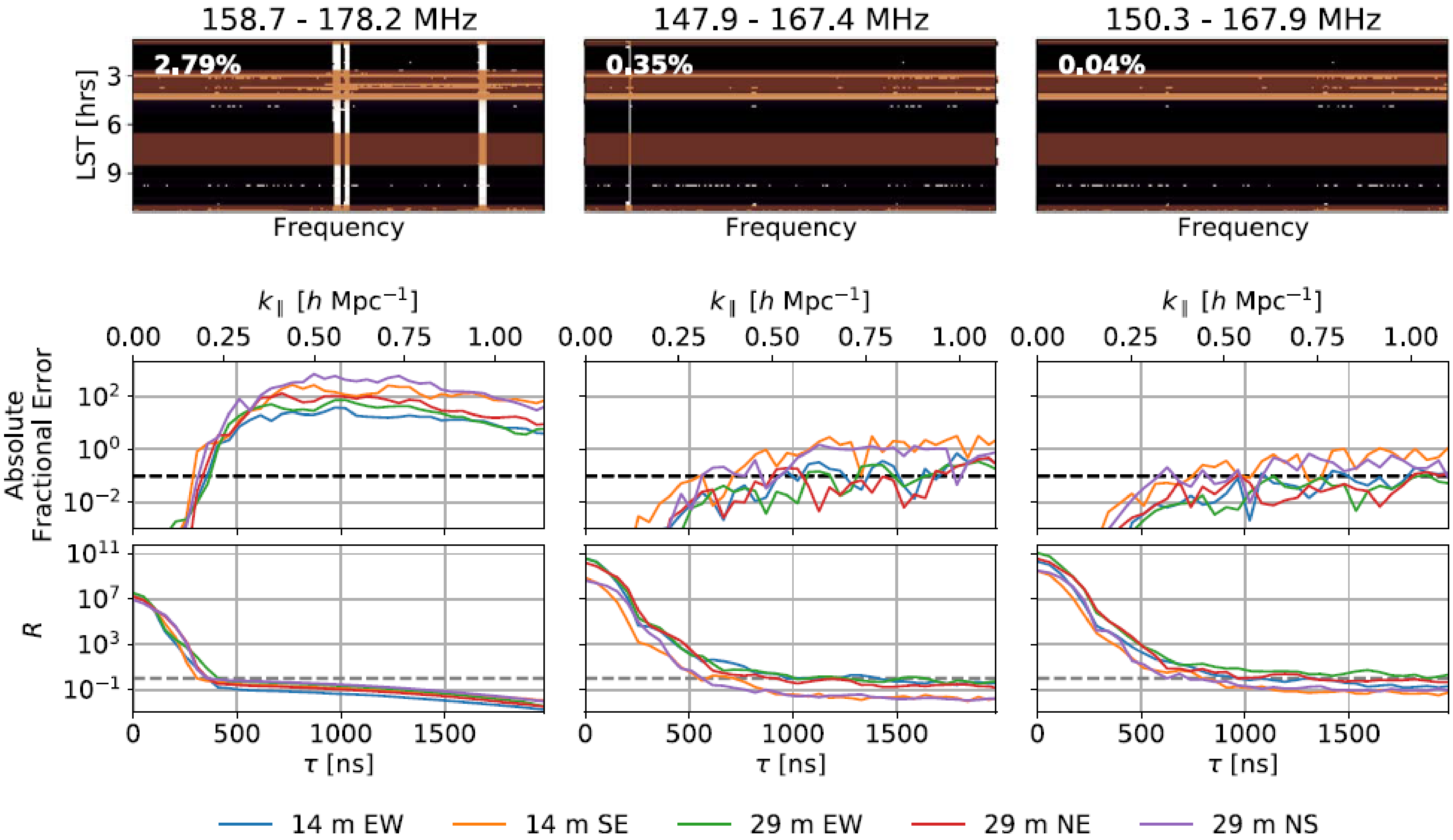




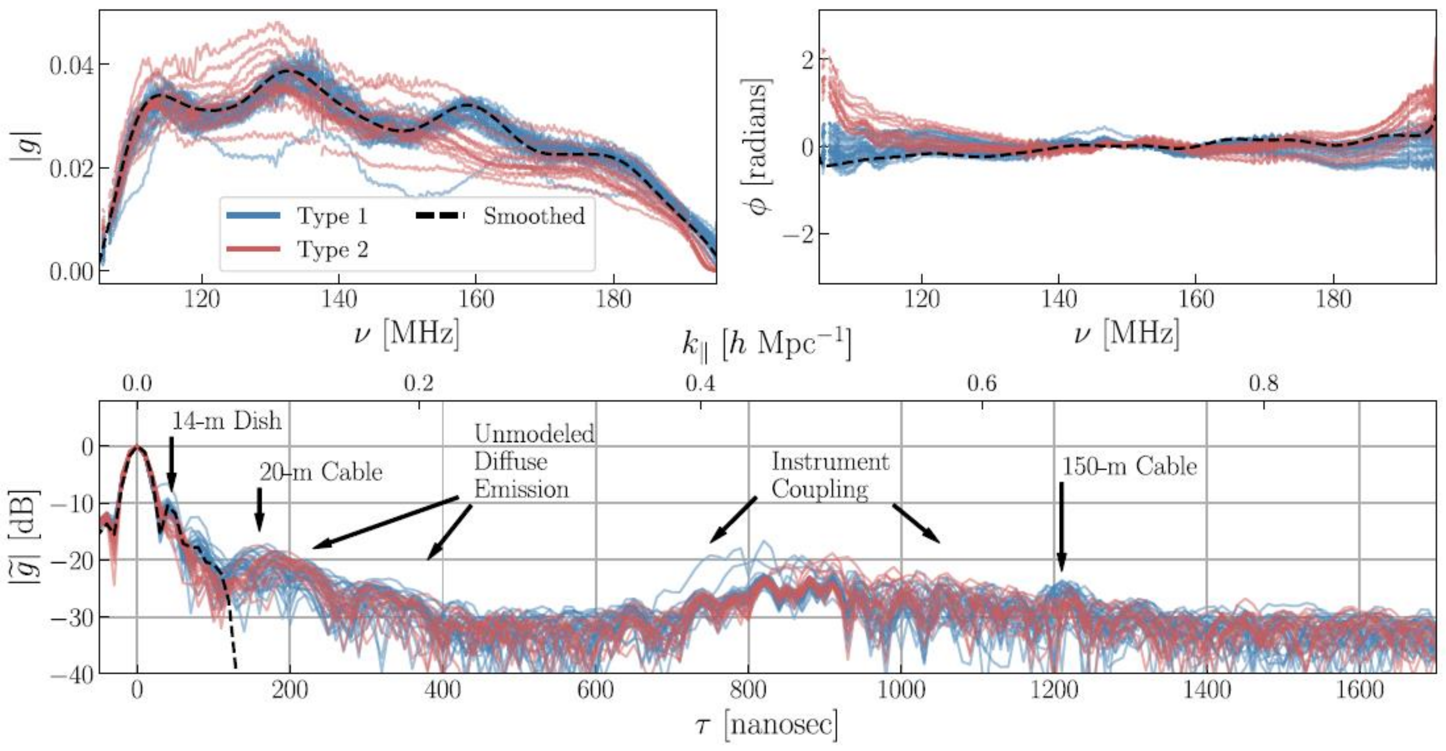
Primary beam



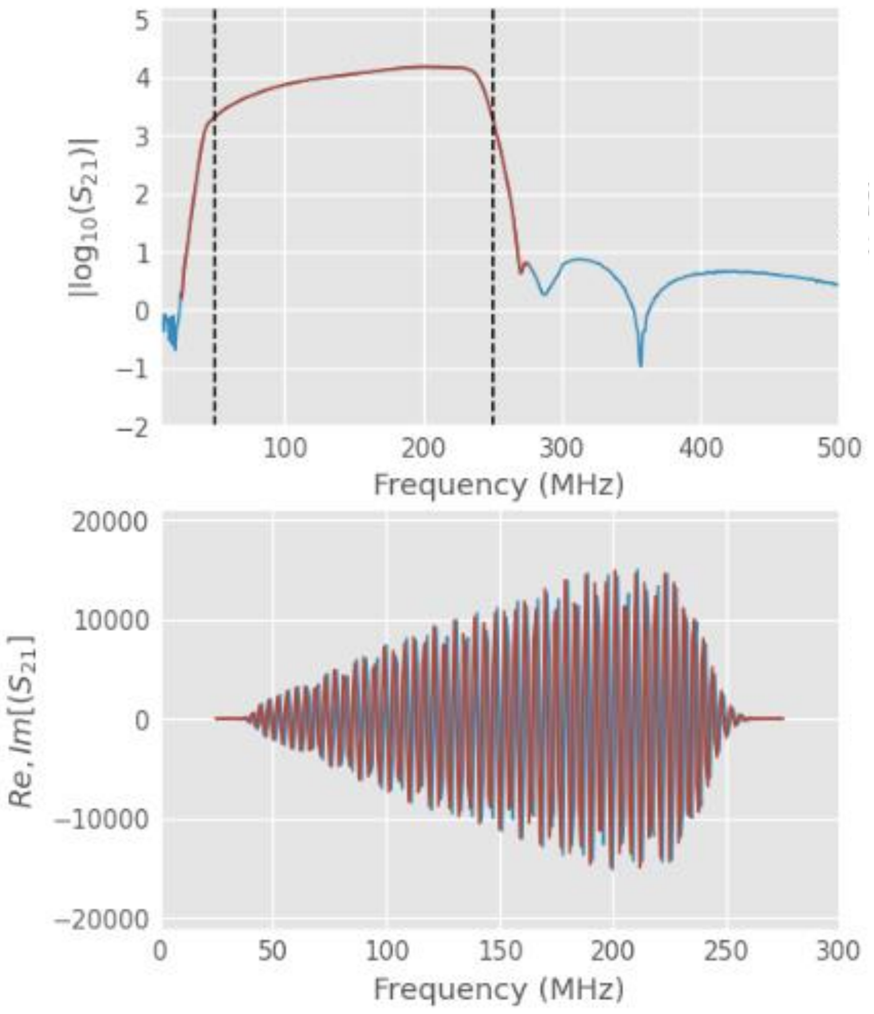
The effect of flagging and gaps



Bandpasses

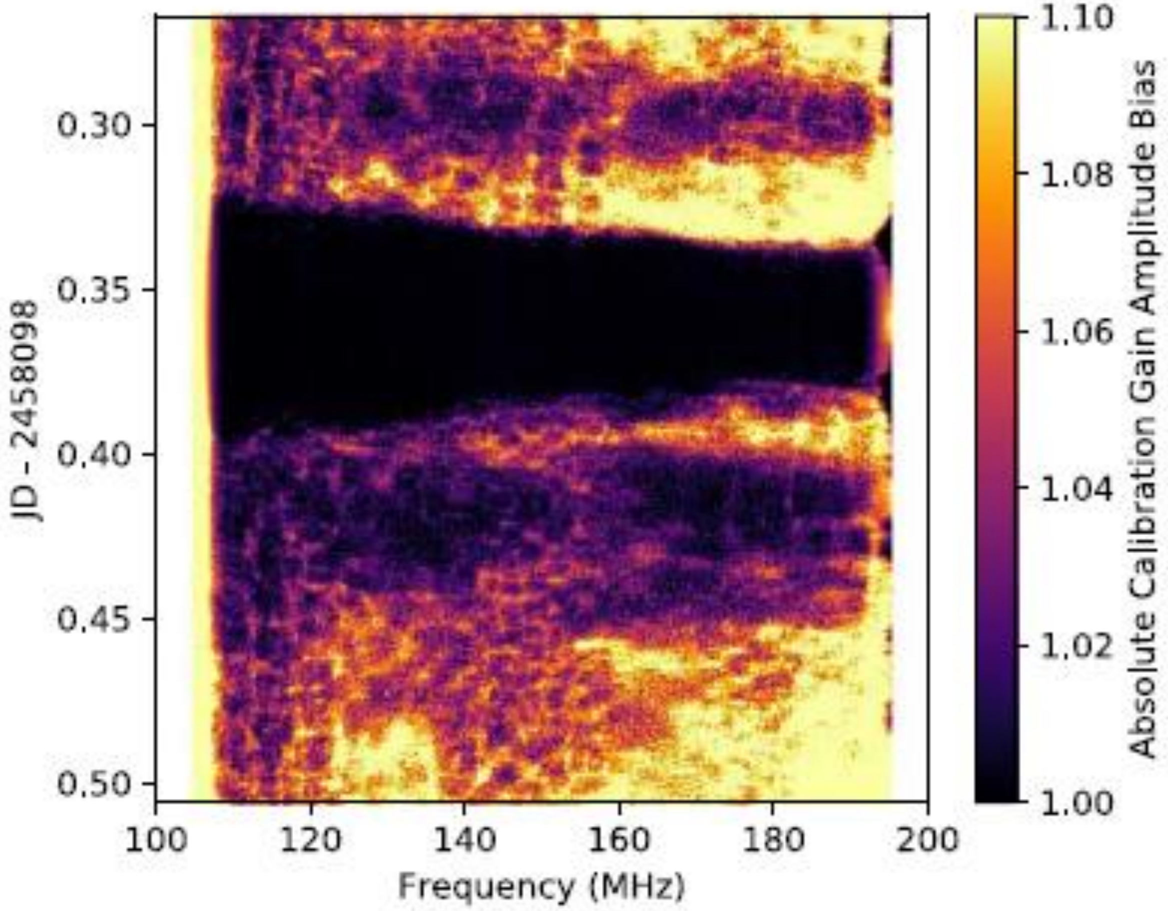


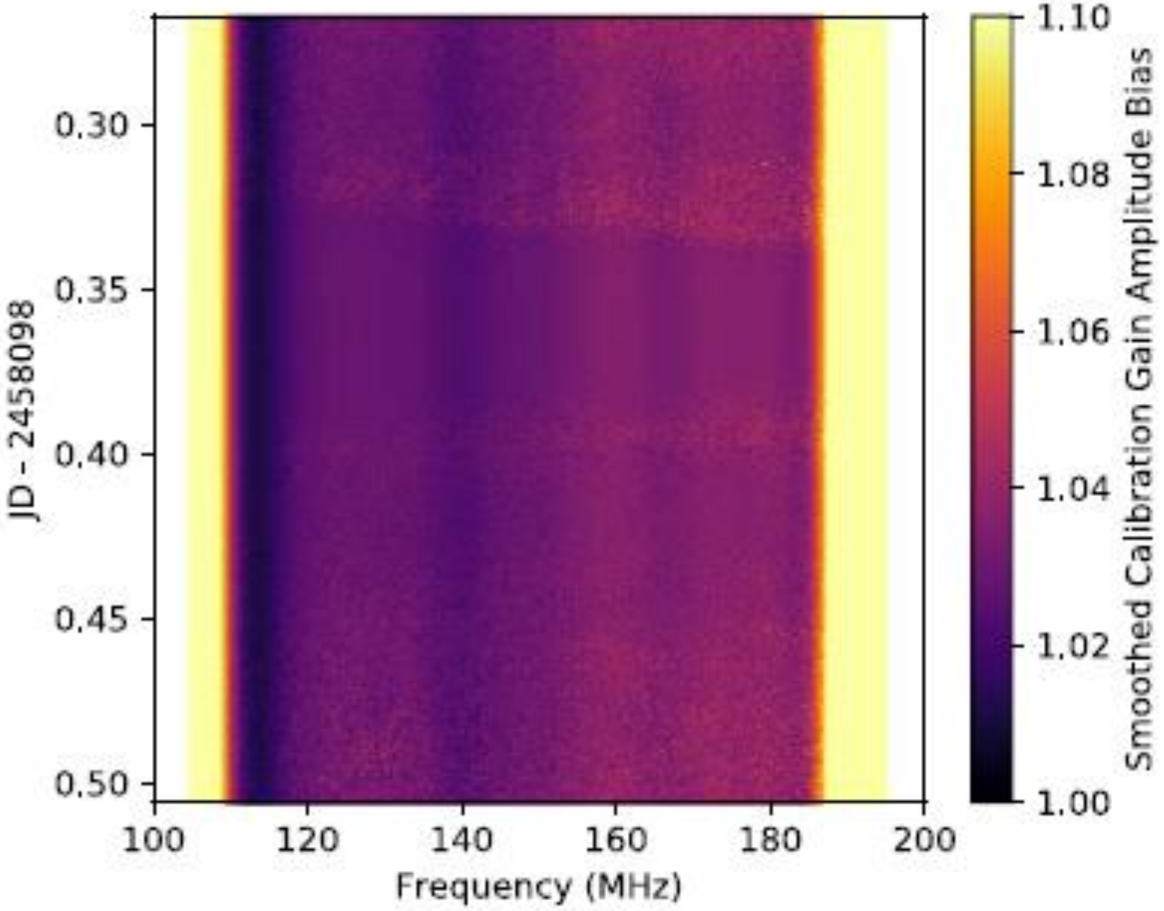
Kern et al 2020 (calibration)



Measured from HERA Phase II amplifier electronics

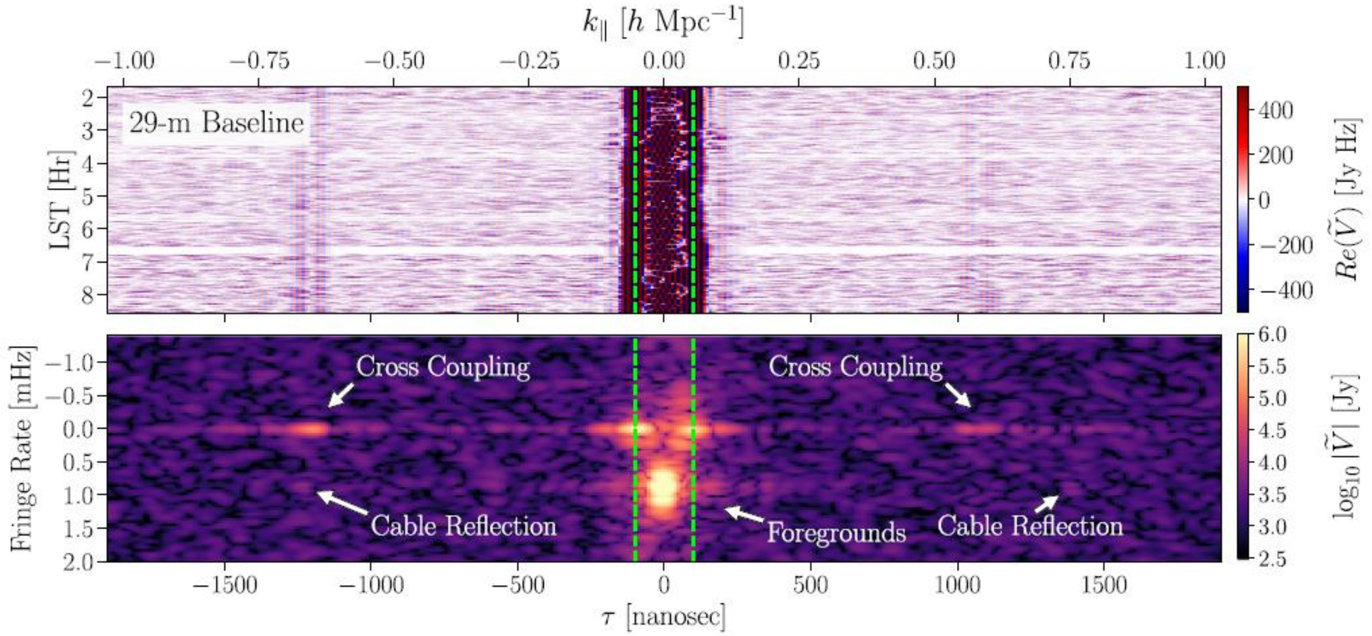
Biases due to calibration errors and gain smoothing

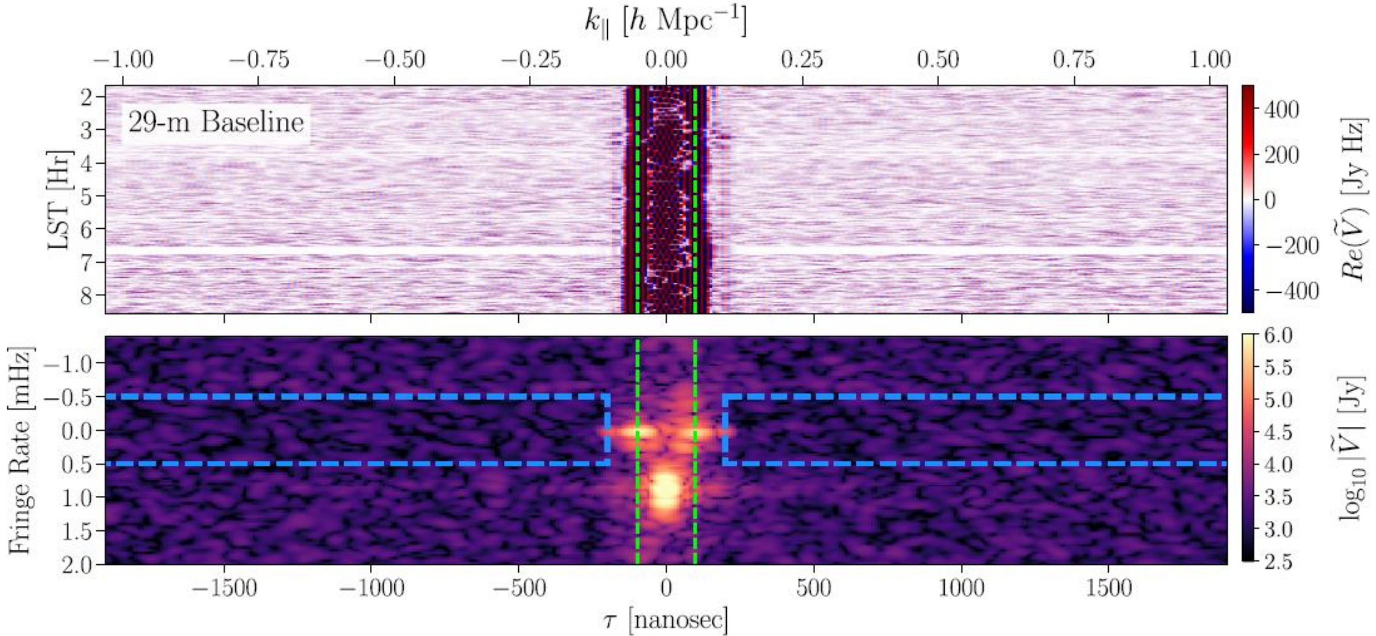




Subtraction of systematics

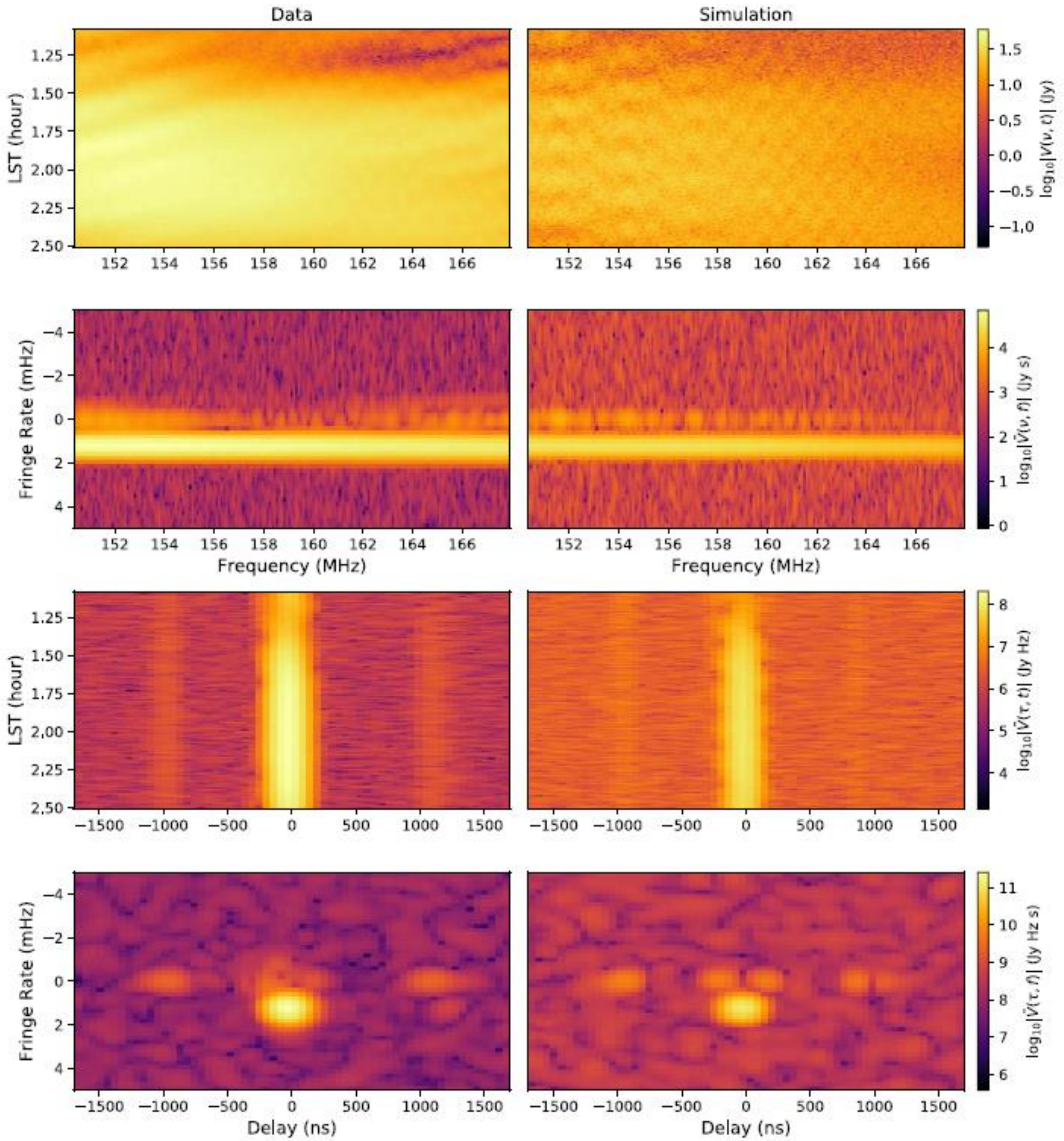
- Done with a semi-empirical model (a more physical model was worked after this work was published)
- More detail on this later in the talk
- Subtractions and filters are potentially lossy to the desired signal; this is the reason for extensive testing!





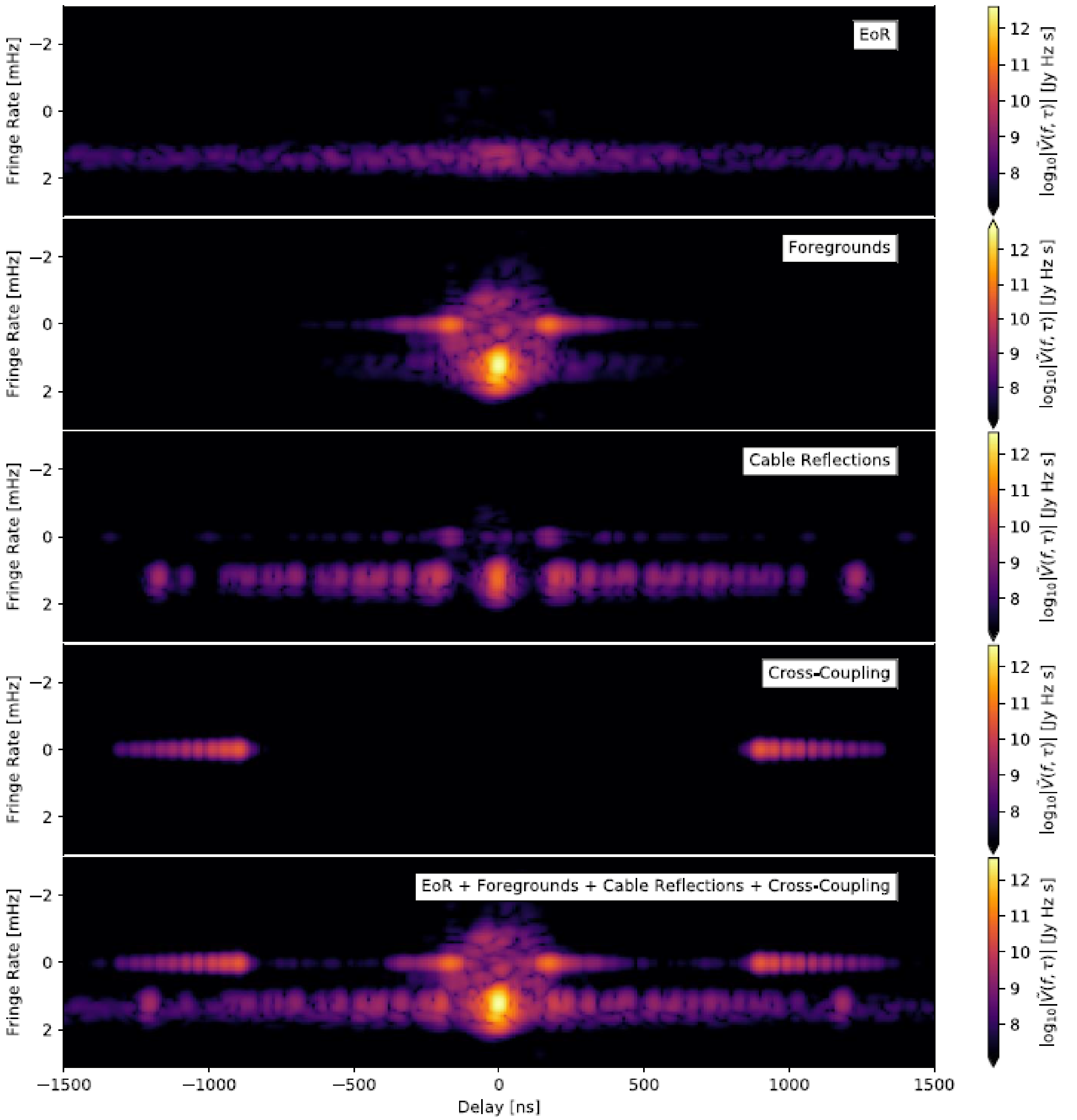
Visibility waterfalls

- Generally good qualitative agreement with the data
- Perennial problem that sky models do not represent the visibilities in detail (calibration schemes are still not fully sky-based)

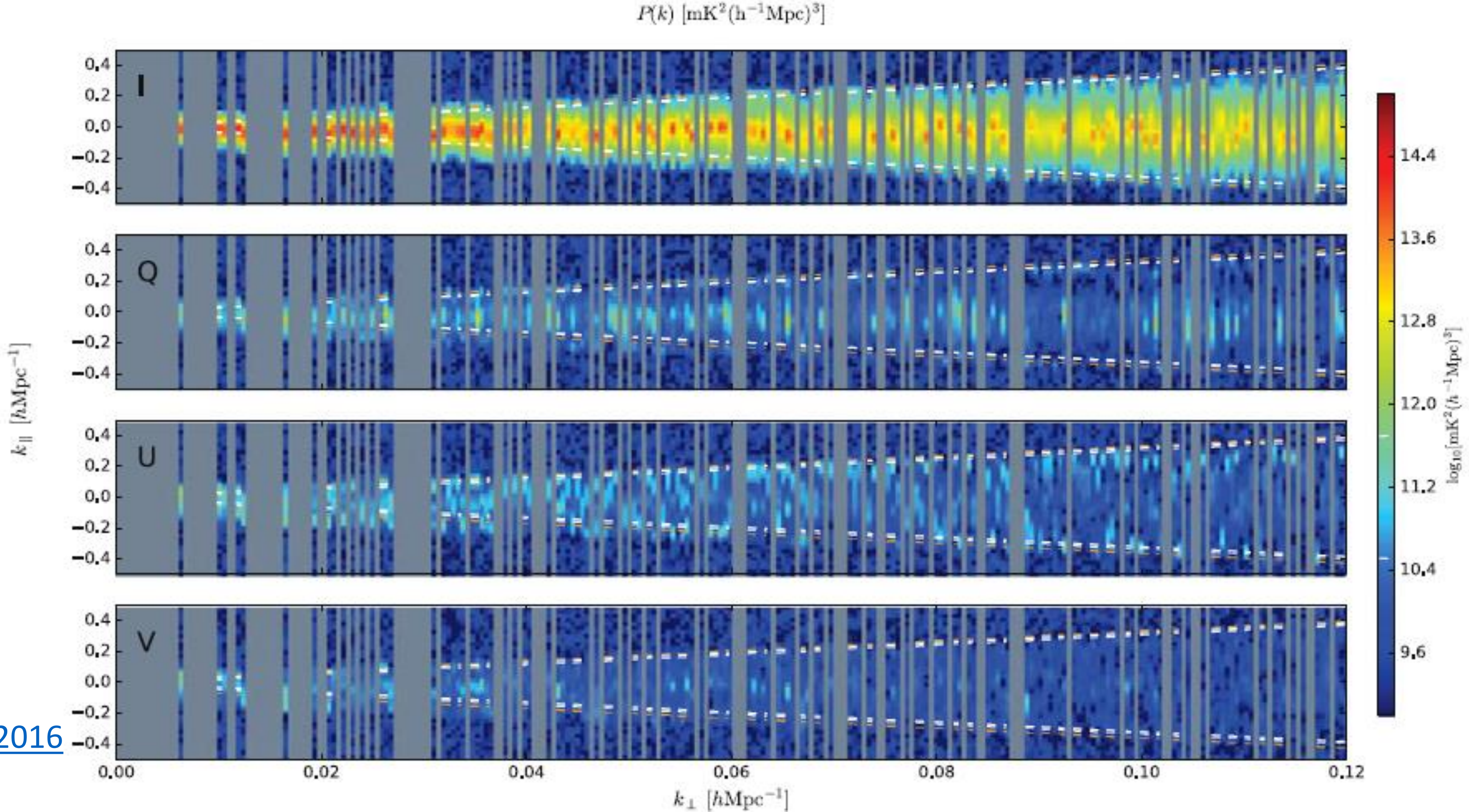


Visualizing signal and systematic effects

- Viewing in different spaces helps check for simulation artifacts
- Can also suggest appropriate bases for unwanted signal



(Aside: other forms of visualization)

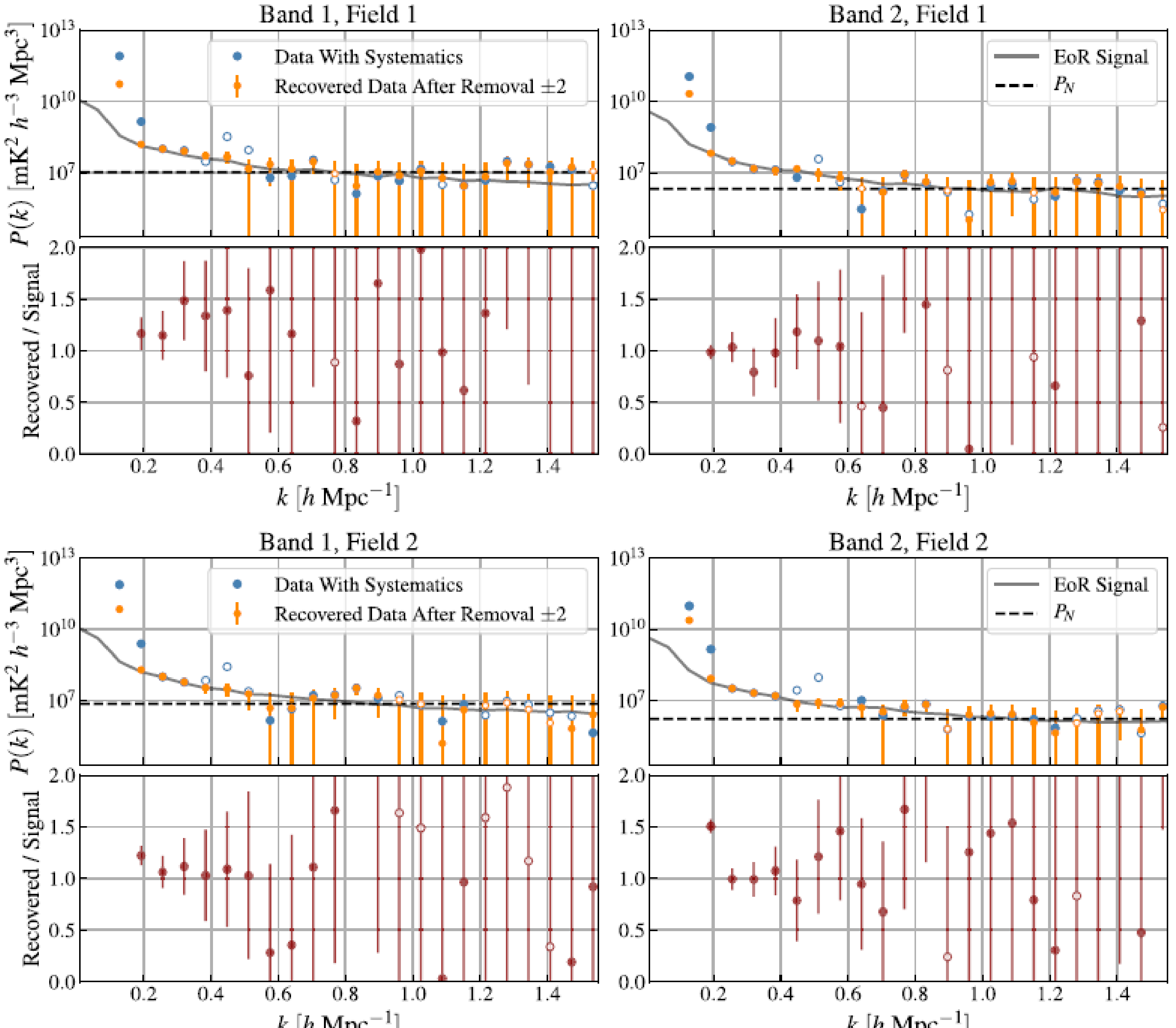


Kohn et al 2016

End-to-end results

Key findings:

- Noise matches models and expectations
- Systematics subtraction does not affect recovered EoR
- EoR is recovered at the expected level (no obvious bias or loss)



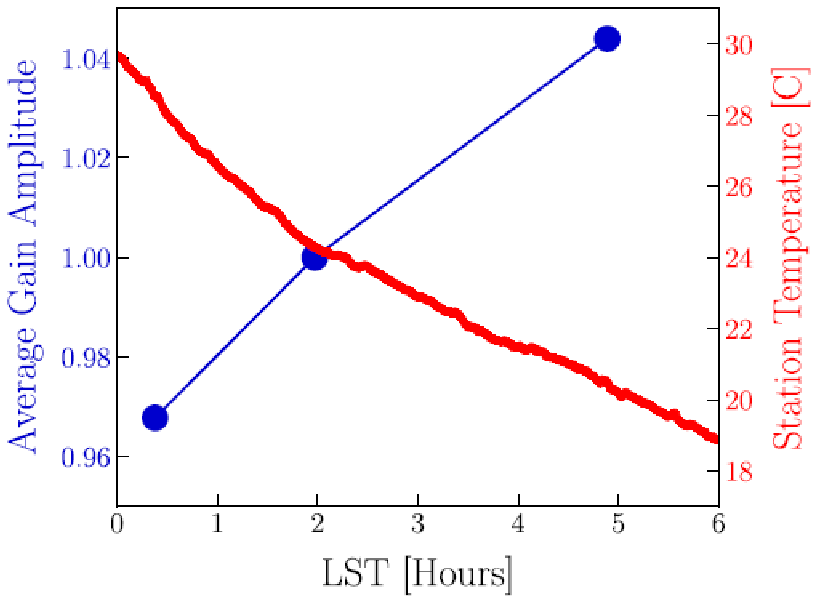
What didn't we check?

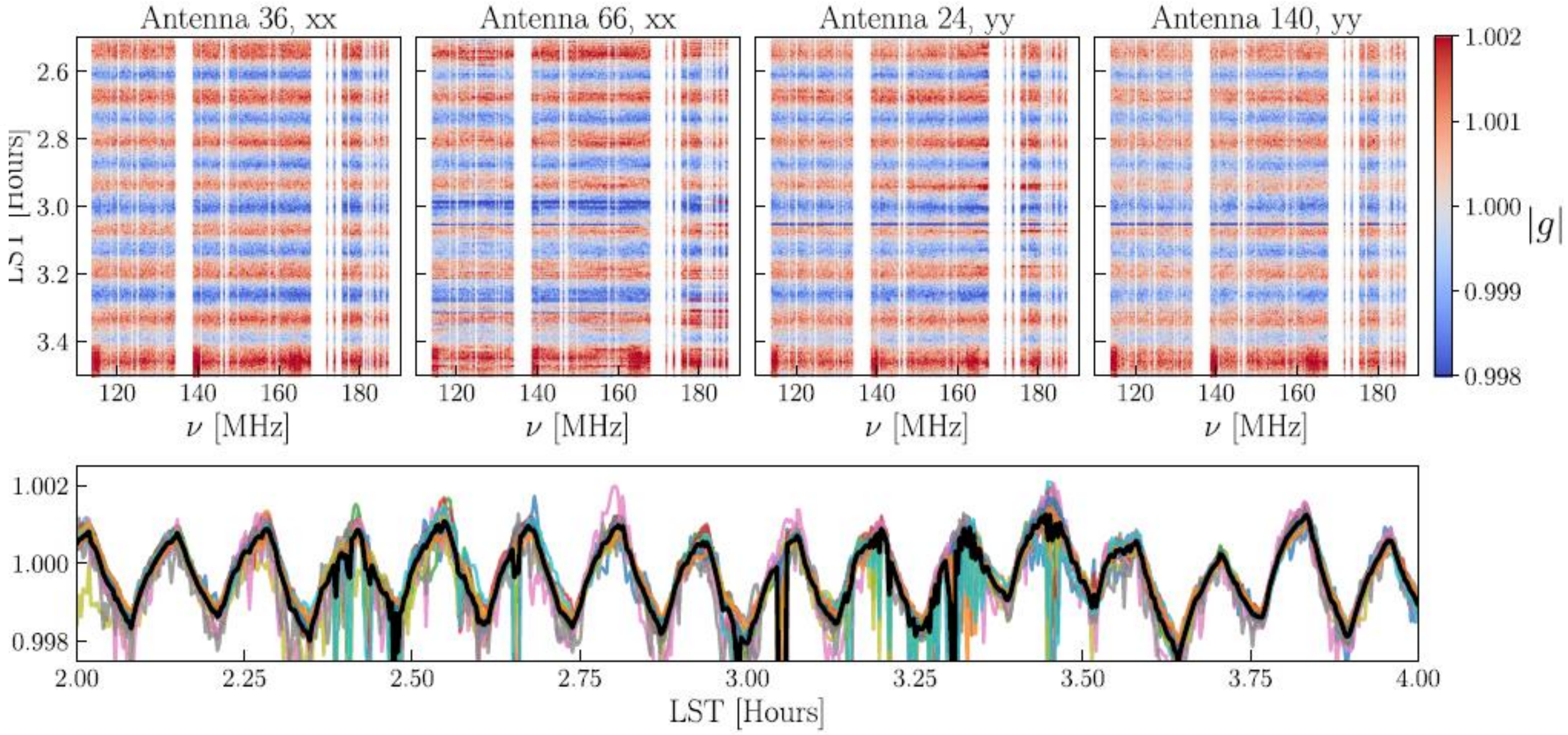
SIMULATED EFFECTS
Known bright point sources
Point-source foregrounds from the GLEAM catalog
Diffuse foregrounds on scales $>3^\circ$
Sky-derived thermal noise
Simple receiver noise
Realistic EM-simulated beam
Direction-independent gains (per feed, time invariant)
Cross-coupling model
Per-antenna cable reflections
Realistic flagging patterns
IGNORED EFFECTS (MAJOR)
Misidentification of dysfunctional antennas
Misidentification of RFI
Generation of `abscal` model from images
Antenna nonredundancy (in primary beams)
Fully realistic antenna cross-coupling
Antenna position errors
IGNORED EFFECTS (MINOR)
Digital (correlator) artifacts
Confused point sources
Fully polarized direction-independent gains
Time variation in direction-independent gains (due to, e.g., temperature variations)
Polarized and/or transient sources
Full suite of possible shapes for $P_{\rm EoR}(k)$
Ionosphere

Temperature-dependent effects

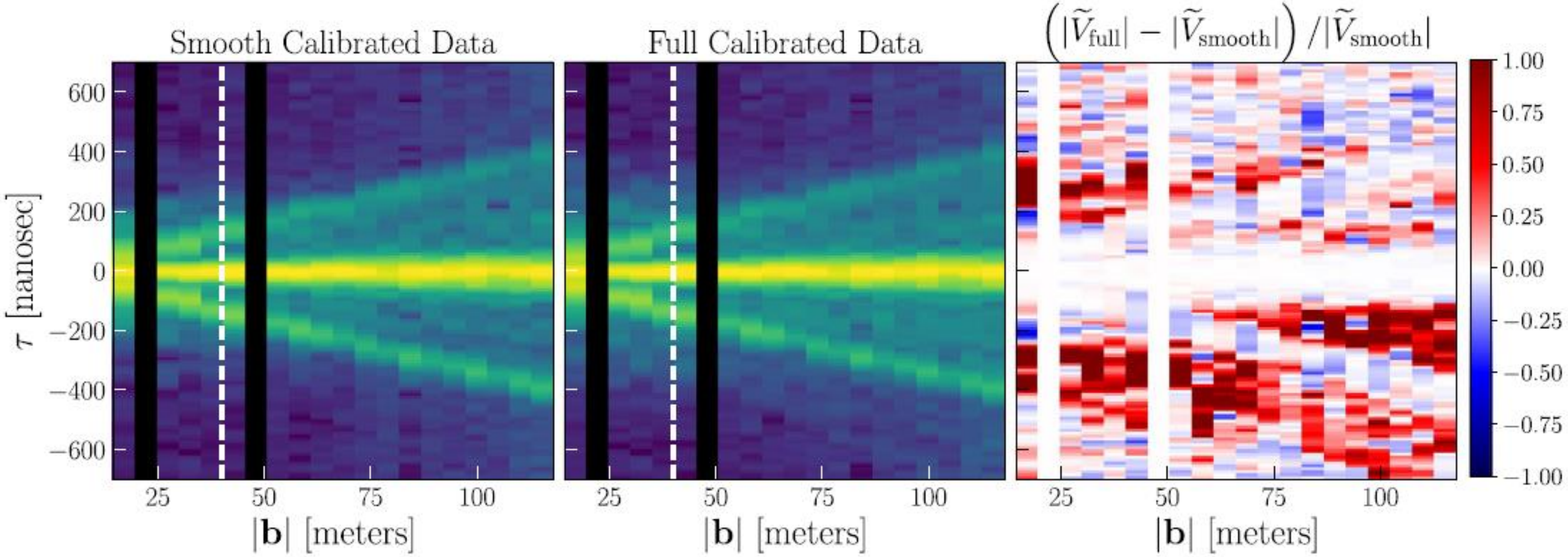
- Obviously, if possible, temperature should be controlled or temperature-insensitive electronics used to minimize these
- Interestingly, using the HERA approachwe found that these could lead to non-trivial *biases* in the calibration if the
- Still an active area of research for mitigation and simulation

[Kern et al 2020 (calibration)]



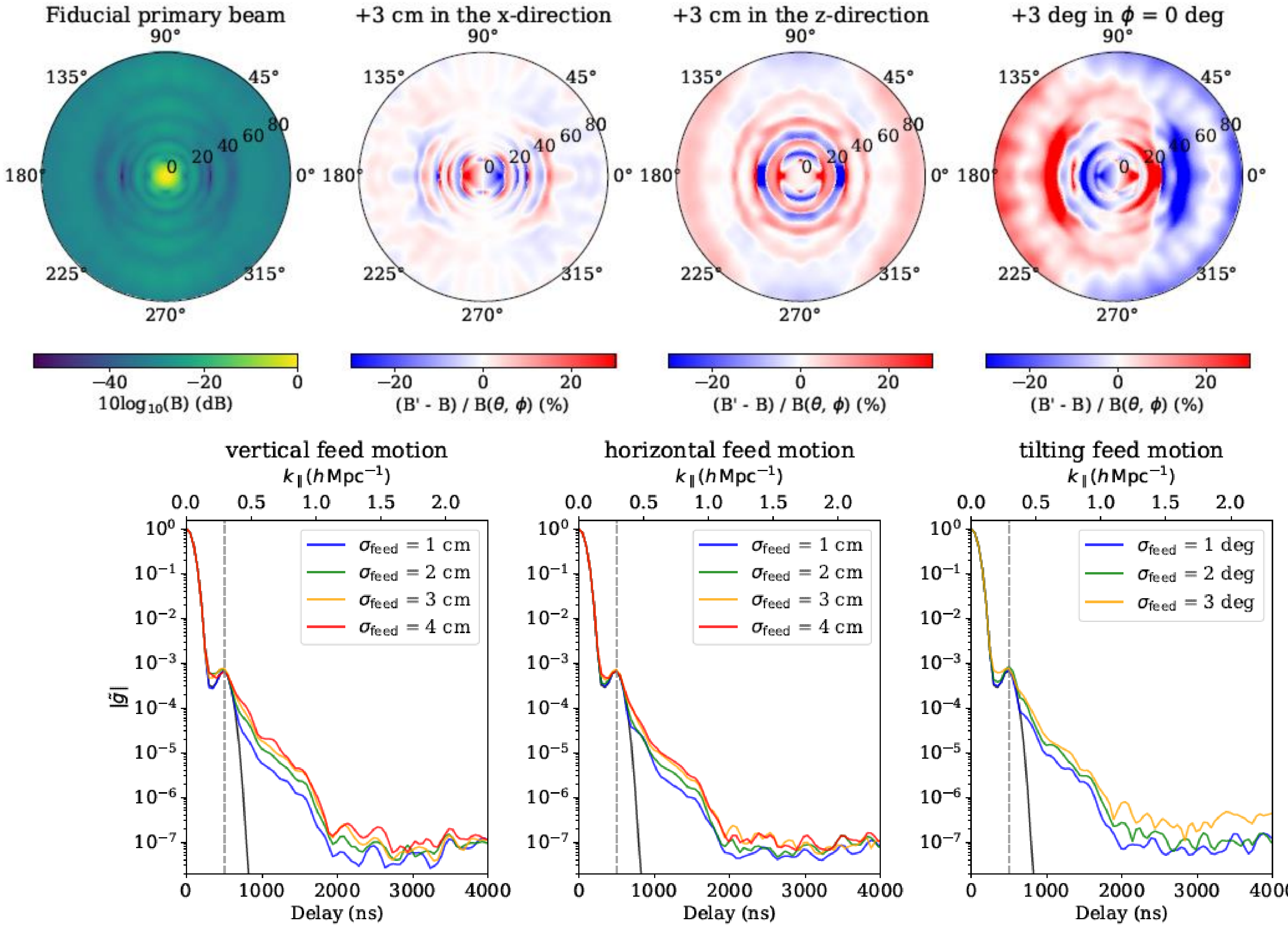


Effect of systematically mis-estimating the gain (due to smoothing)



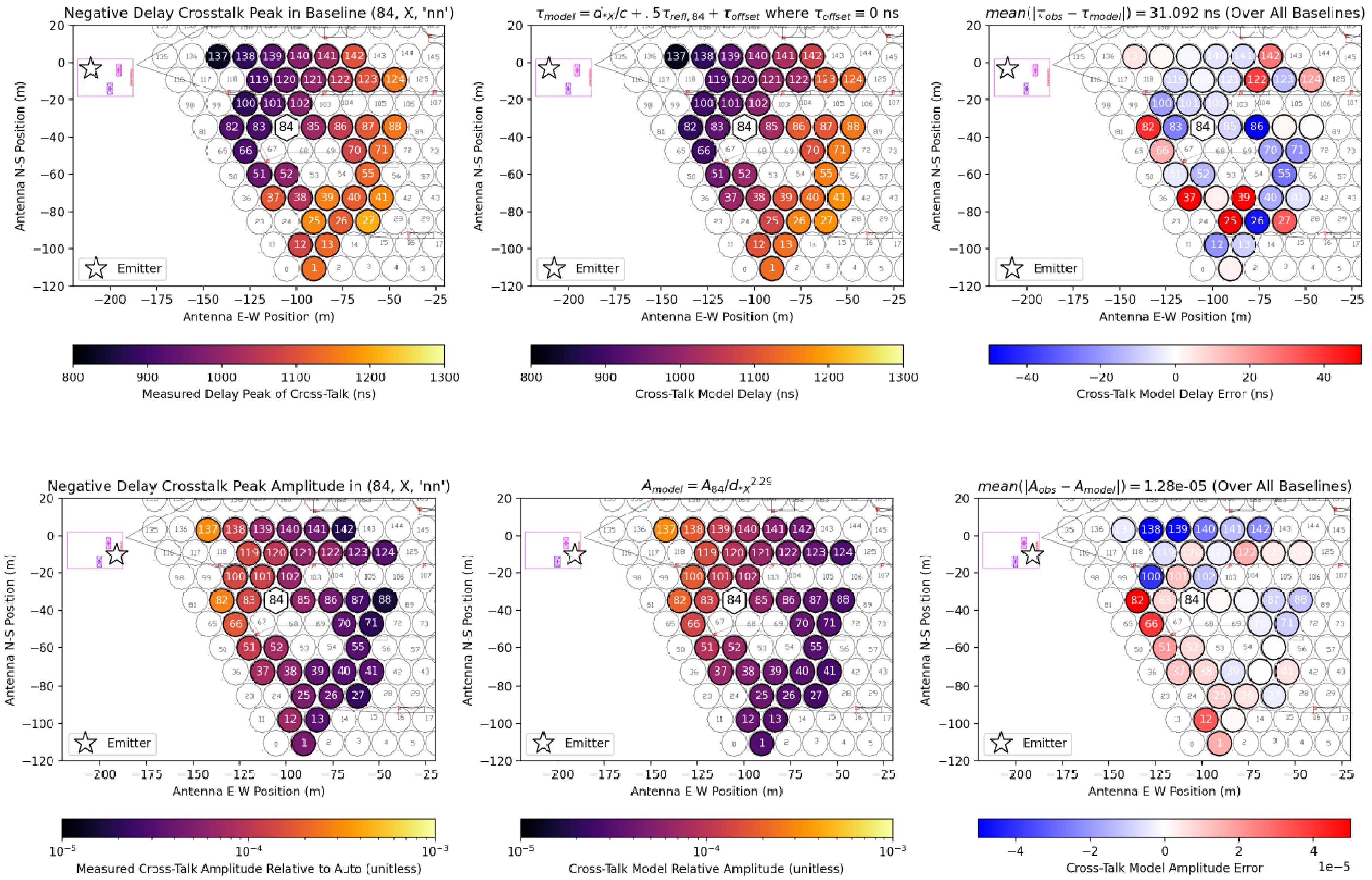
Variation in feed position

- Honggeun Kim et al 2022 (in prep)
- Considers the effect on the primary beam of variations in the feed position

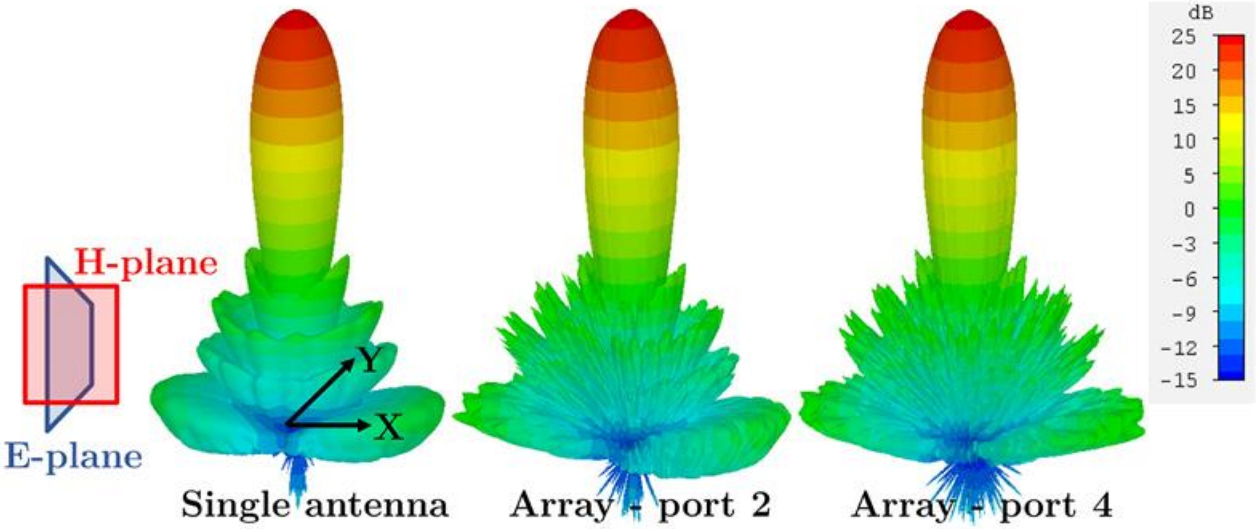


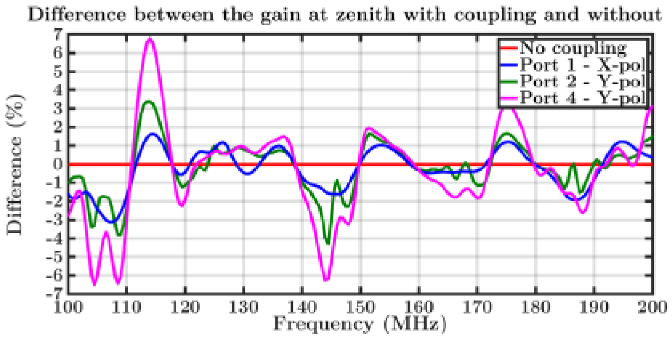
Self-contamination

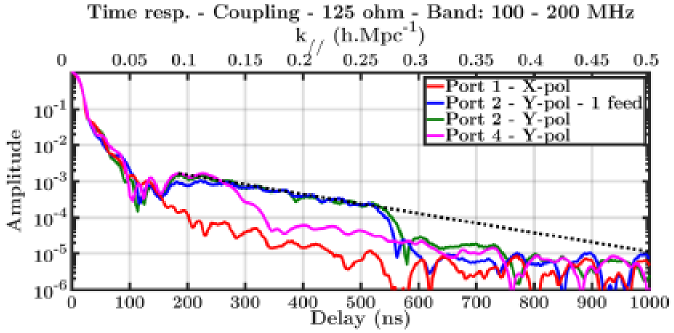
- Broadcasting source near the electronics seems to have been responsible for high delay cross-coupling effect in the Phase I data

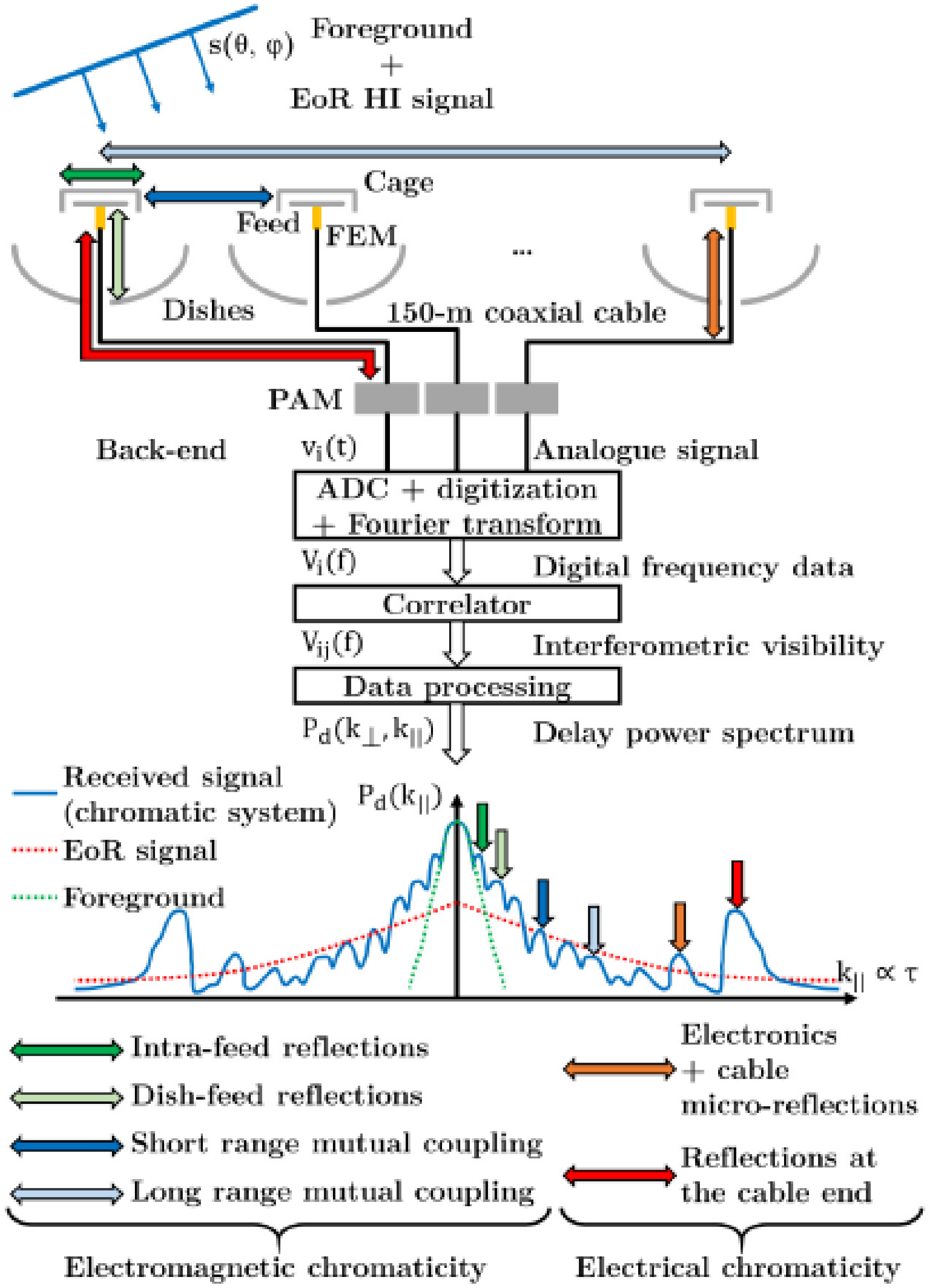


A more systematic approach to systematics





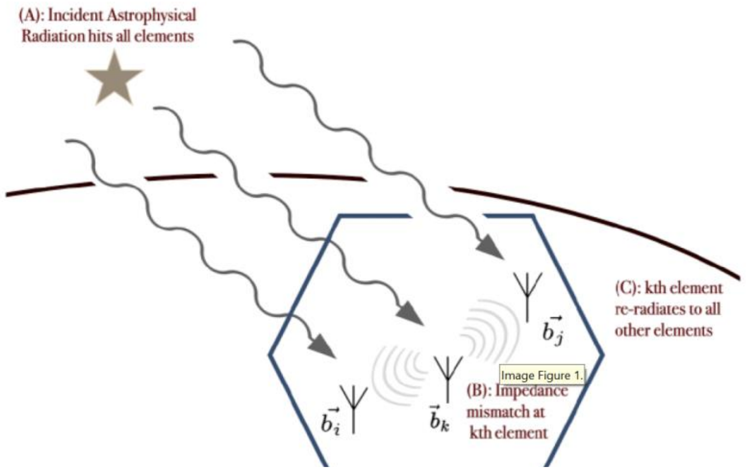


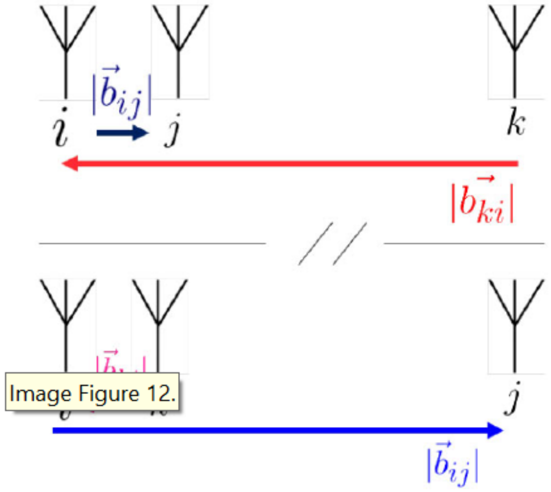


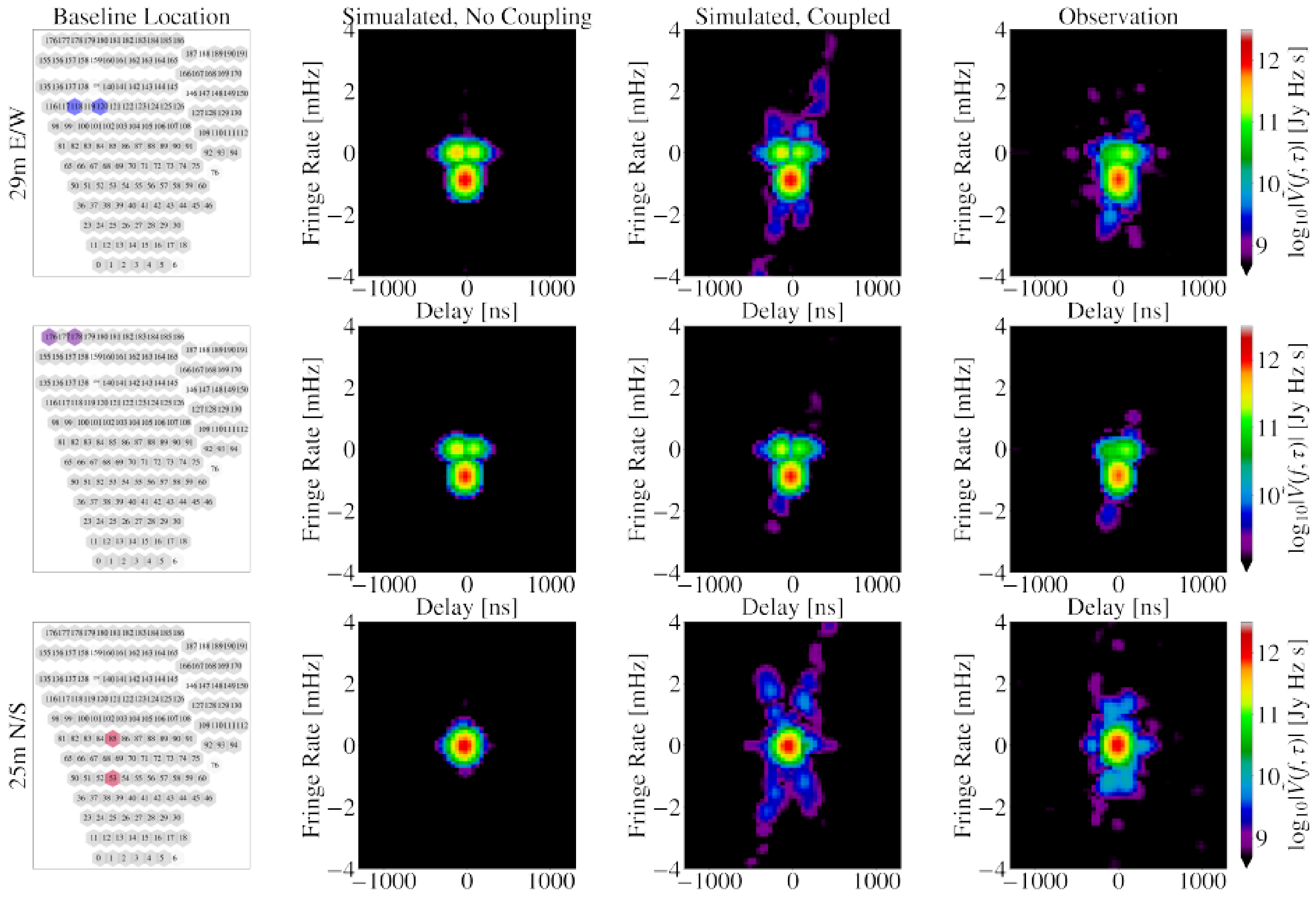
Fagnoni et al 2021a (HERA phase I beams)

A simplified cross-coupling model

$$\mathbf{V}^1_{ij} = \mathbf{V}^0_{ij} + \frac{i\eta_0}{4\lambda}\left(-\sum_{k\neq i} \frac{\Gamma_k}{R_k|\vec{b_{ki}}|} e^{2\pi i \frac{\nu}{c}|\vec{b}_{ki}|}\mathbf{J}_i(\hat{b_{ik}})\mathbf{J}_k(\hat{b_{ki}})\mathbf{V}^0_{kj} + \sum_{k\neq j} \frac{\Gamma_k^*}{R_k|\vec{b_{kj}}|} e^{-2\pi i \frac{\nu}{c}|\vec{b}_{kj}|}\mathbf{V}^0_{ik}\mathbf{J}^y_k(\hat{b_{kj}})\mathbf{J}^y_j(\hat{b_{jk}})\right).$$

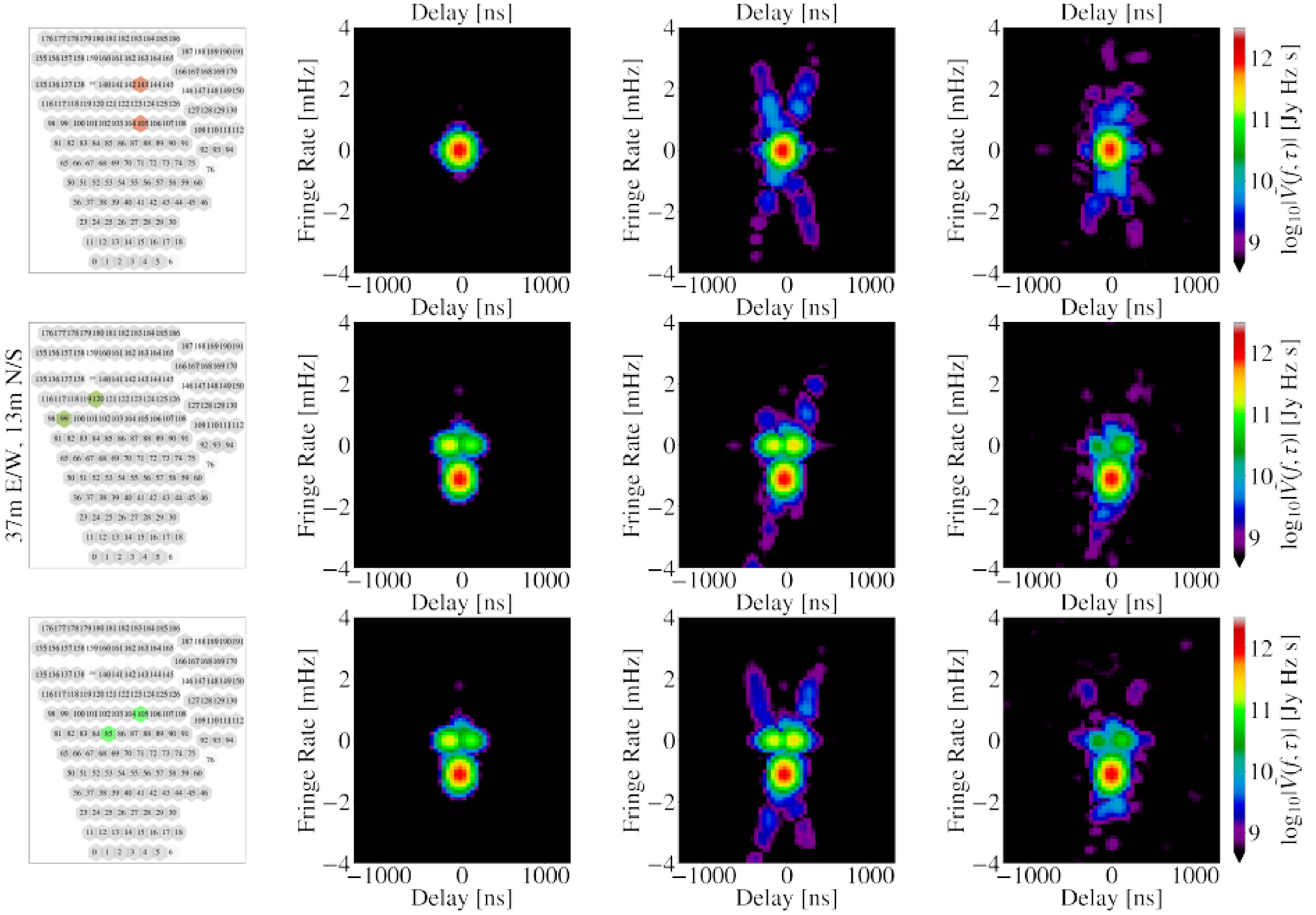






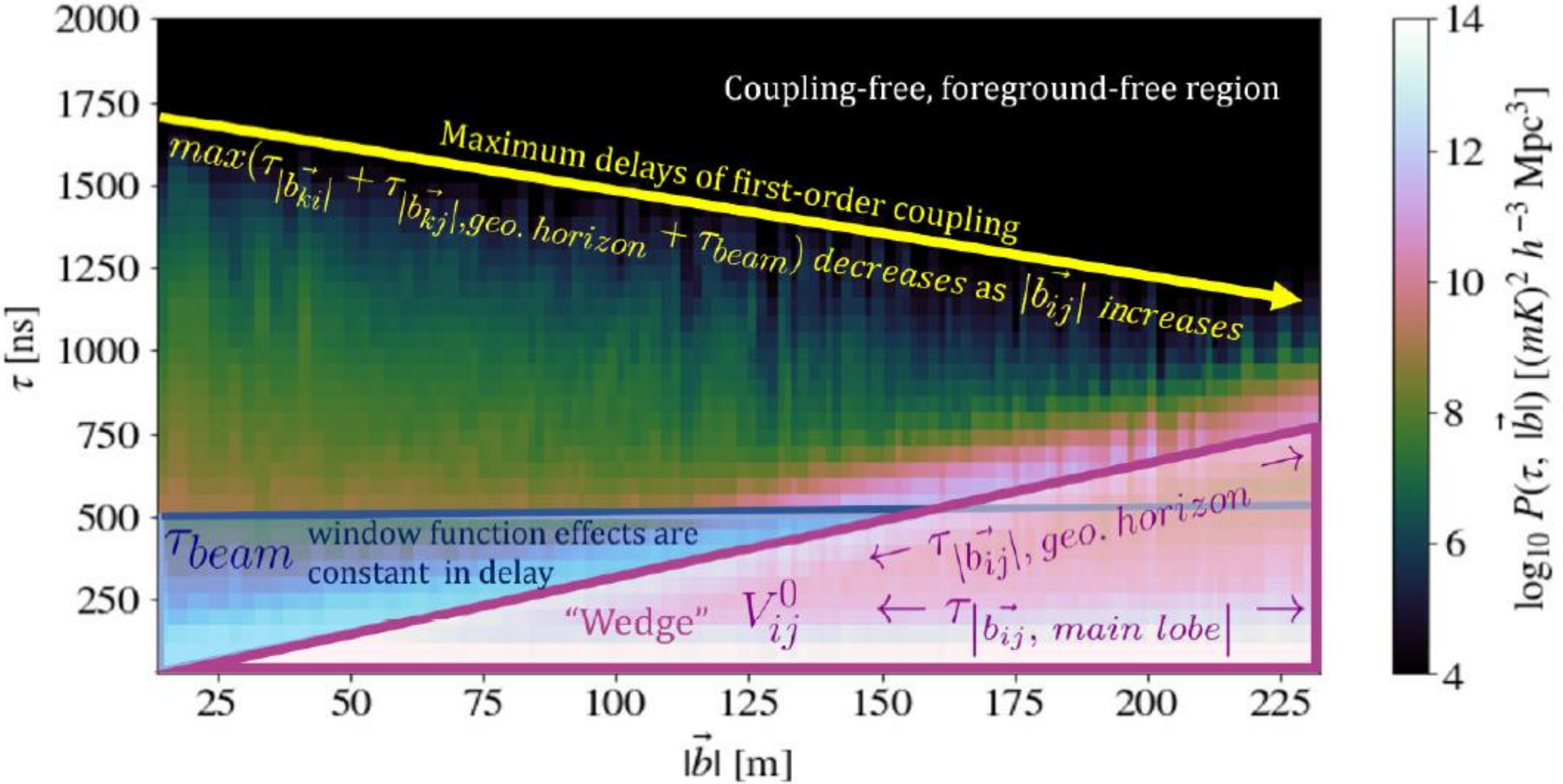
[Josaitis et al 2022a (mutual coupling)]

Does not capture all "embedded beam" effects, but does match the HERA Phase II data qualitatively quite well



Josaitis et al 2022 (in prep); follow-on to [Josaitis et al 2022a (mutual coupling)]

Much of the leaked power from cross-coupling actually comes from fringe-rates that don't lie in the main lobe of the beam



Conclusions and closing thoughts

- End-to-end simulations of as many instrument effects as possible is highly desirable for testing both the analysis pipeline as well as physical understanding of the effects
- Truly including all effects is difficult, and requires an interplay between EM simulations, empirical models, data analysis, and visibility simulation
- Because removal or suppression of systematic effects can be lossy or otherwise distort the desired signal, elimination is best, but progress can still be made by assessing the effects of mitigation strategies on the analysis via simulations

References

DeBoer et al 2017

Aguirre et al 2022 (software validation)

Kern et al 2019 (systematics I)

Kern et al 2020 (systematics II)

Kern et al 2020 (calibration)

HERA Memo #104 (self-contamination)

Josaitis et al 2022a (mutual coupling)

Fagnoni et al 2021a (HERA phase I beams)

Fagnoni et al 2021b (HERA phase II beams)

Kohn et al 2016

Kohn et al 2019 (polarized beams)