Carbon monOxide Mapping Array Project

Jonas G. S. Lunde
PhD fellow
Institute of Theoretical Astrophysics,
University of Oslo

Nils-Ole Stutzer
PhD fellow
Institute of Theoretical Astrophysics,
University of Oslo

Leah Hansen
Master student
Institute of Theoretical Astrophysics,
University of Oslo

Patrick C. Breysse
Assistant Professor of Physics
Southern Methodist University

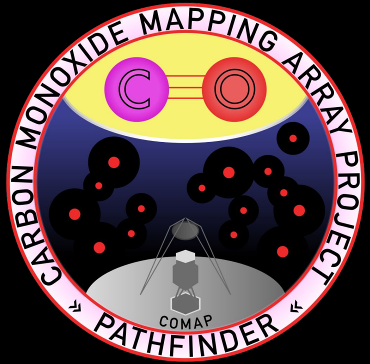
Overview

COMAP-Pathfinder
Jonas:
PCAs and non-linear map cleaning

Nils:
The power spectrum methodology

Leah:
Simulating ground pickup systematics in the time domain

COMAP-wide
Patrick:
A Wide-Area CO Cross-Correlation Survey

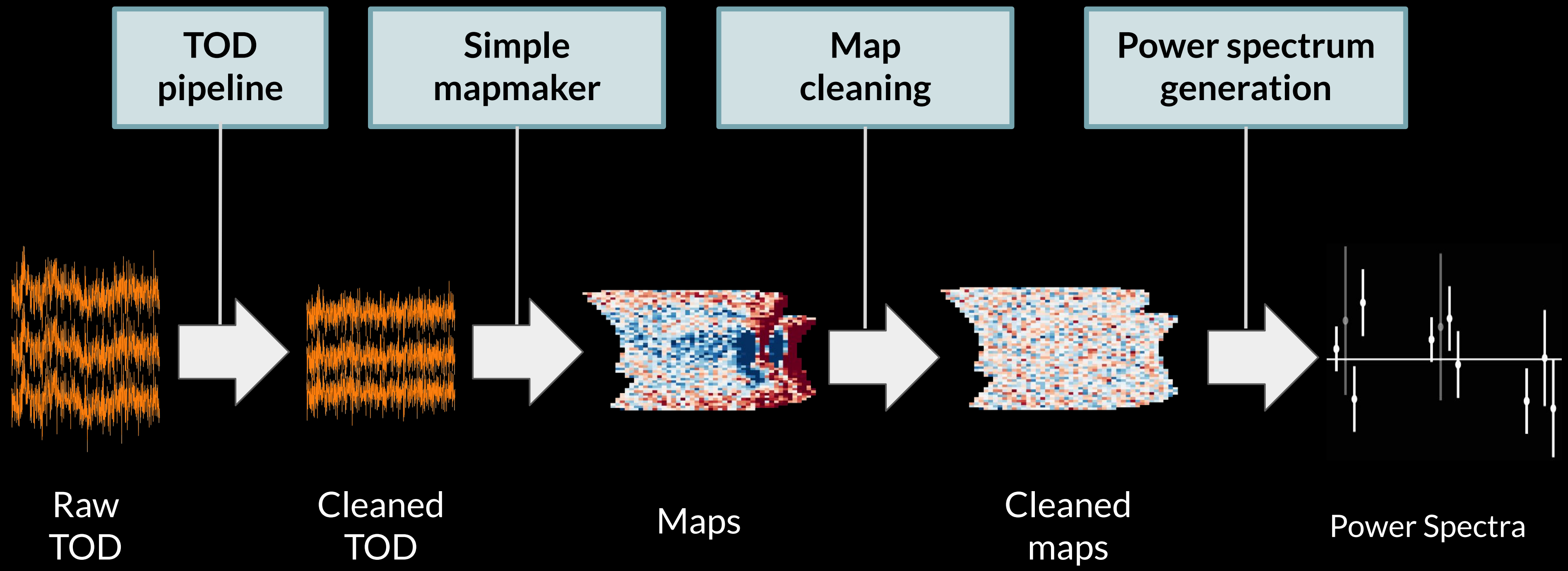


COMAP Pathfinder

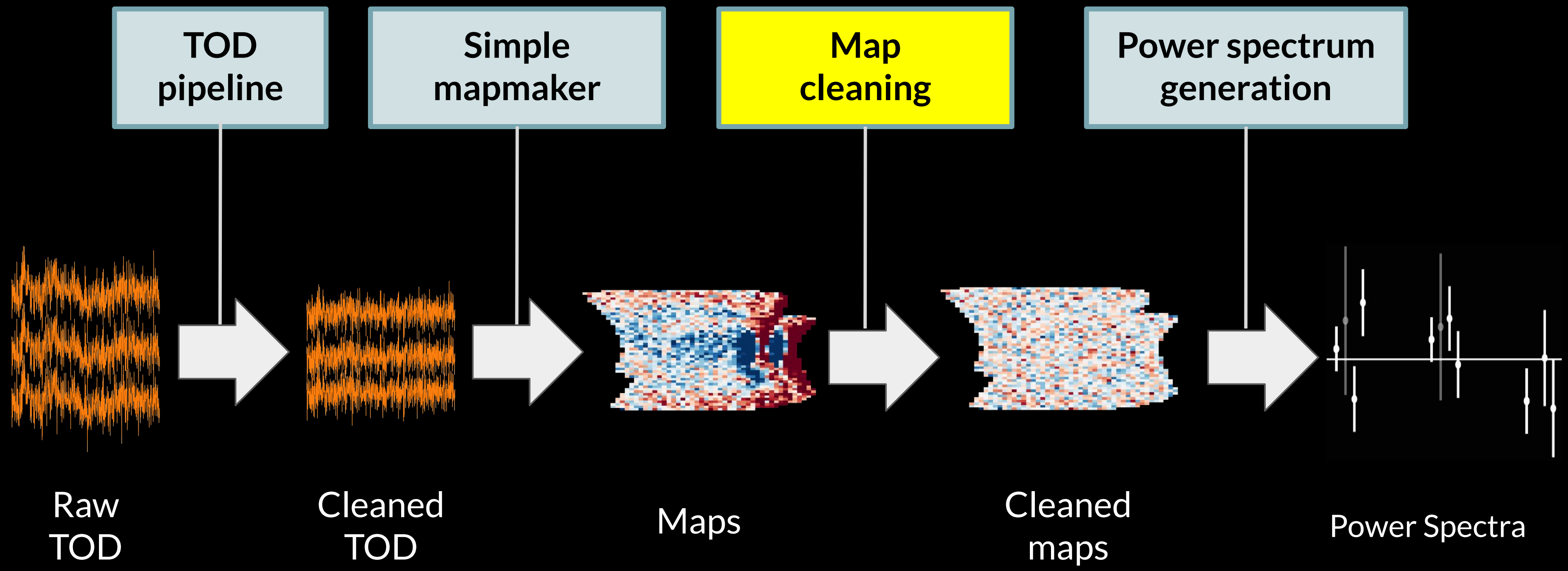
PCAs and non-linear map cleaning

Jonas G. S. Lunde
PhD fellow
Institute of Theoretical Astrophysics,
University of Oslo

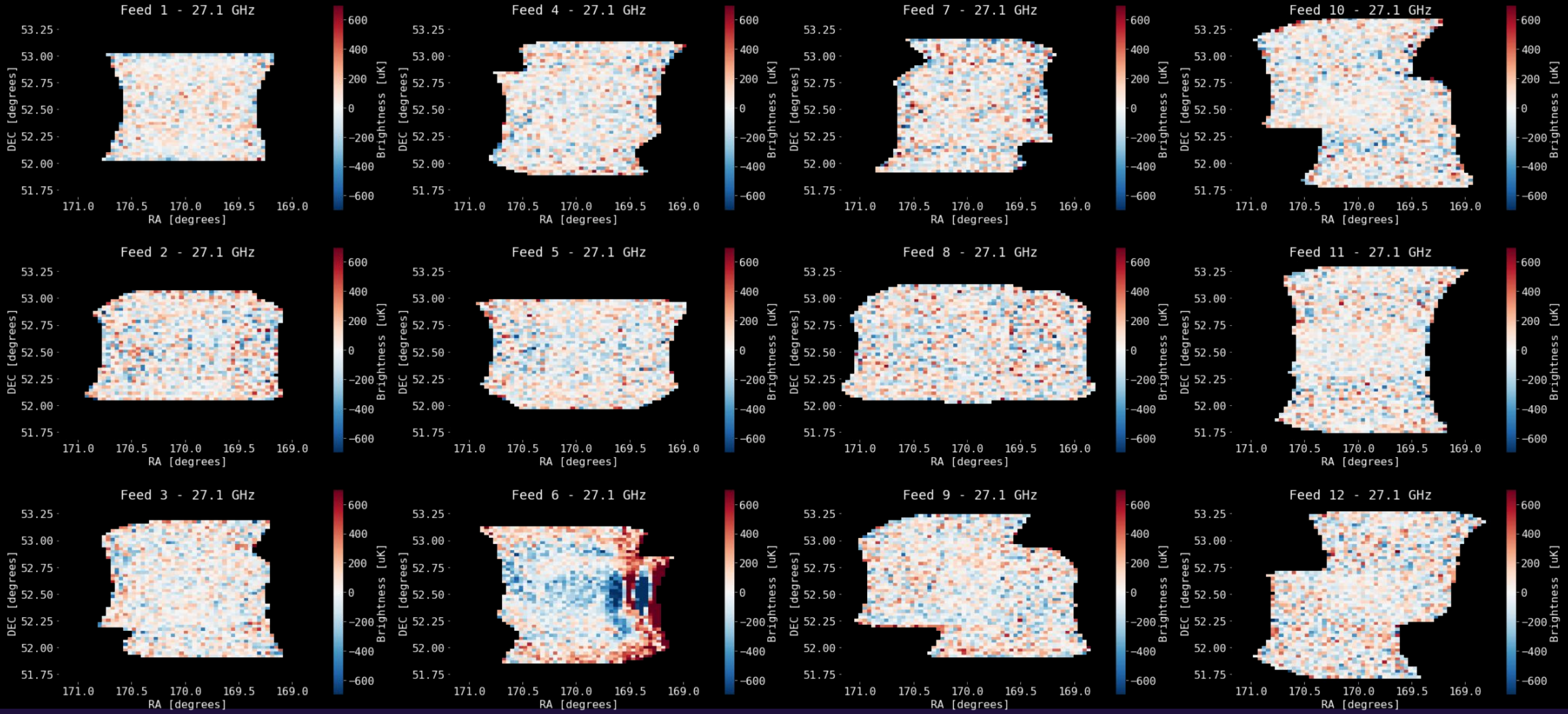
The COMAP data pipeline



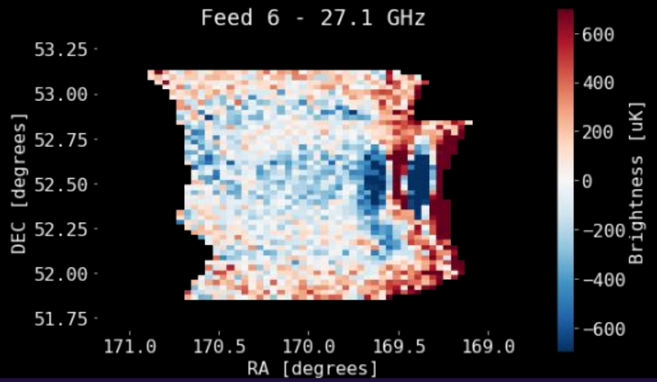
The COMAP data pipeline



With increased sensitivity comes new systematics

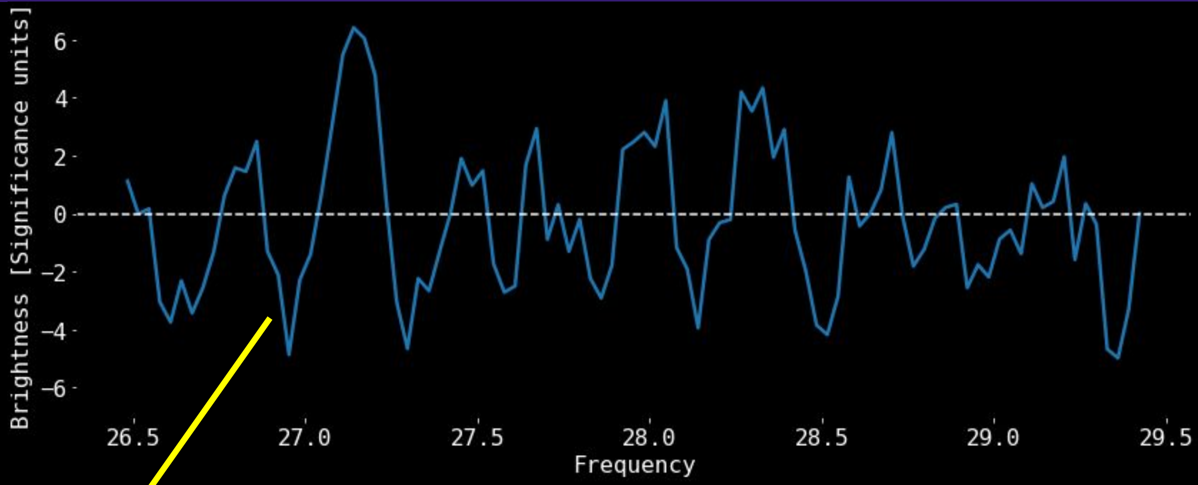


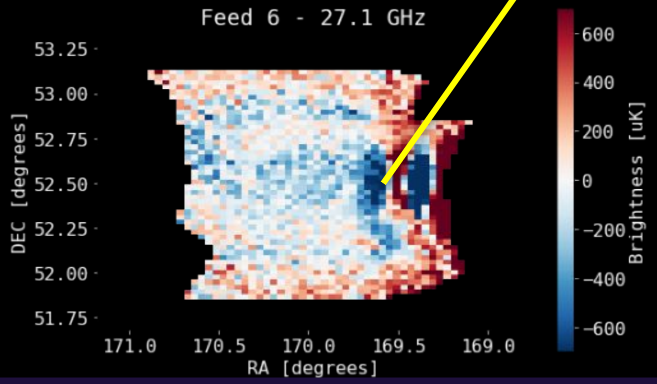
With increased sensitivity comes new systematics



With increased sensitivity comes new systematics

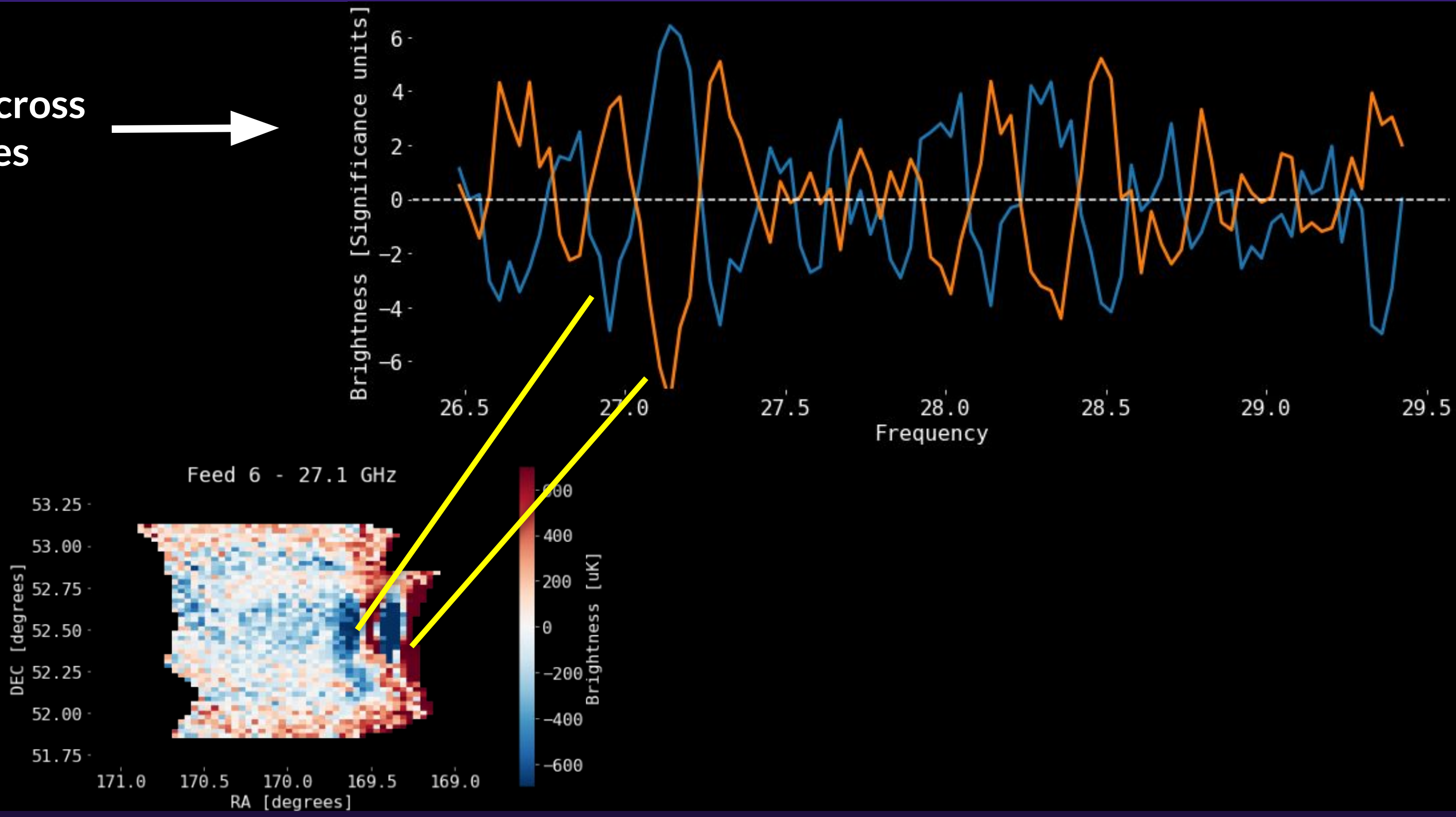
Single pixel across frequencies





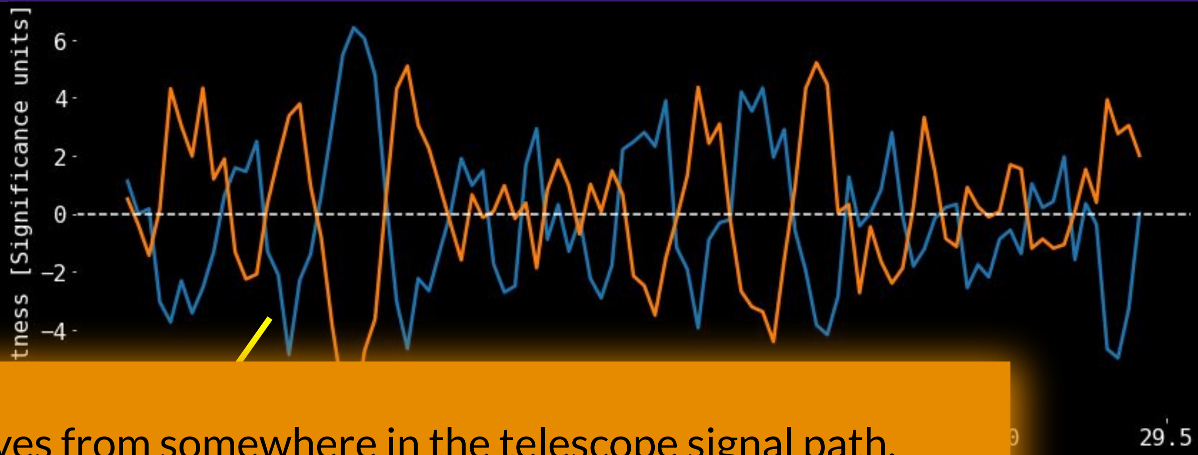
With increased sensitivity comes new systematics

Single pixel across frequencies

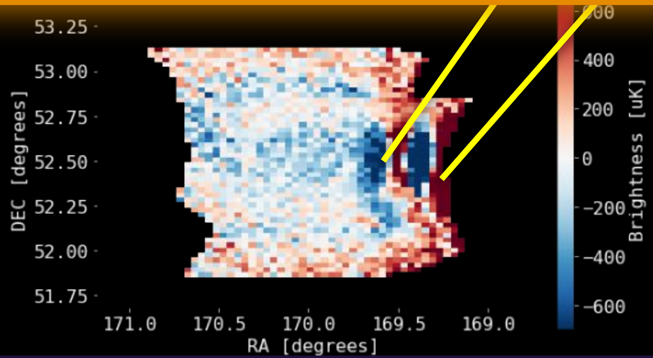


With increased sensitivity comes new systematics

Single pixel across frequencies

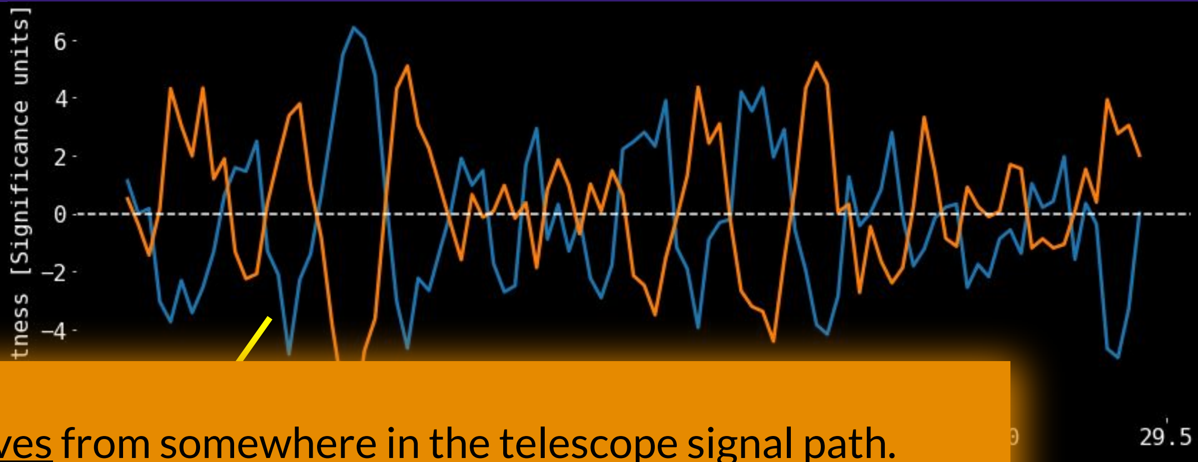


- These are standing waves from somewhere in the telescope signal path.



With increased sensitivity comes new systematics

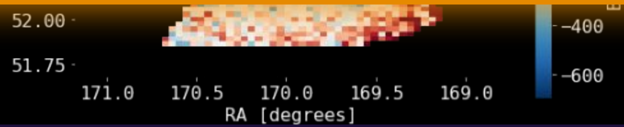
Single pixel across frequencies



- These are standing waves from somewhere in the telescope signal path.

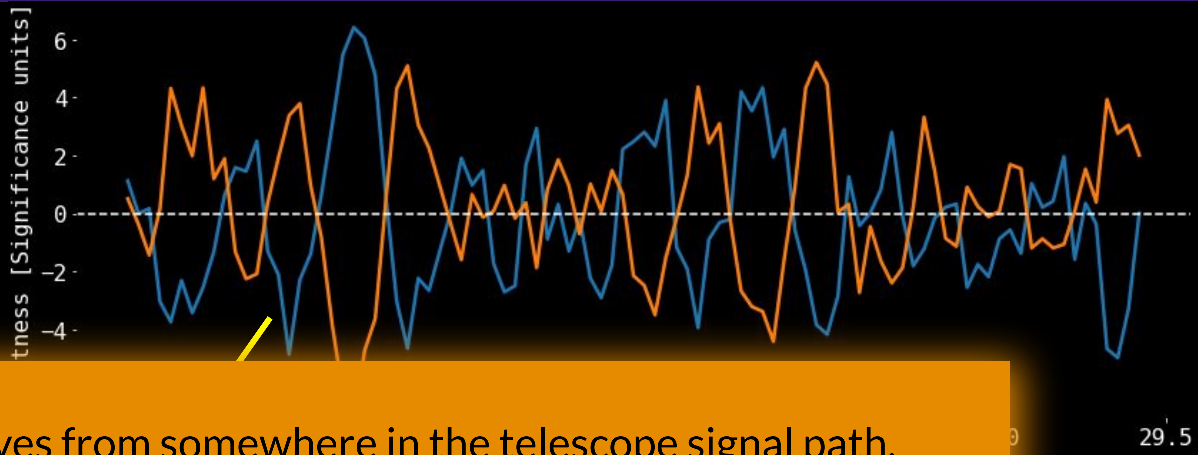
- Because of COMAPs coherent detector design these standing waves will behave very predictably, on the form:

  "*Some function of pixels*" × "*Some function of frequency*"



With increased sensitivity comes new systematics

Single pixel across frequencies



- These are standing waves from somewhere in the telescope signal path.

- Because of COMAPs coherent detector design these standing waves will behave very predictably, on the form:

  "*Some function of pixels*" × "*Some function of frequency*"

- This is the data model assumed by a Principal Component Analysis (PCA)!

Map-cleaning with PCA (principal component analysis)

We first tried PCA, which is equivalent to solving:

$$\min : \sum_{\nu,p} (D_{\nu,p} - a_\nu m_p^T)^2$$

Map-cleaning with PCA (principal component analysis)

We first tried PCA, which is equivalent to solving:

(in practice you solve a SVD, but I find this conceptually easier)

$$\min : \sum_{\nu,p} (D_{\nu,p} - a_\nu m_p^T)^2$$

Map-cleaning with PCA (principal component analysis)

We first tried PCA, which is equivalent to solving:

$$\min : \sum_{\nu,p} (D_{\nu,p} - a_\nu m_p^T)^2$$

Our maps

Some frequency vector

Some pixel vector

Map-cleaning with PCA (principal component analysis)

We first tried PCA, which is equivalent to solving:

$$\min : \sum_{\nu,p} (D_{\nu,p} - a_\nu m_p^T)^2$$

Our maps
(noisy!)
(non-uniformly noisy)

Some
frequency
vector

Some
pixel
vector

Generalizing the PCA to allow noise-weighting

Where can we put a $\sigma^2_{\nu,p}$?

$$\min : \sum_{\nu,p} (D_{\nu,p} - a_\nu m_p^T)^2$$

Generalizing the PCA to allow noise-weighting

The "correct" thing to be solving for:

$$\min : \sum_{\nu,p} \frac{\left(D_{\nu,p} - a_\nu m_p^T\right)^2}{\sigma_{\nu,p}^2}$$

Generalizing the PCA to allow noise-weighting

The “correct” thing to be solving for:

$$\min : \sum_{\nu,p} \frac{\left(D_{\nu,p} - a_\nu m_p^T\right)^2}{\sigma_{\nu,p}^2}$$

This is actually the maximum likelihood solution to the data model:

$$D_{\nu,p} = a_\nu m_p^T + n_{\nu,p}$$

Generalizing the PCA to allow noise-weighting

The “correct” thing to be solving for:

$$\min : \sum_{\nu,p} \frac{\left(D_{\nu,p} - a_\nu m_p^T\right)^2}{\sigma^2_{\nu,p}}$$

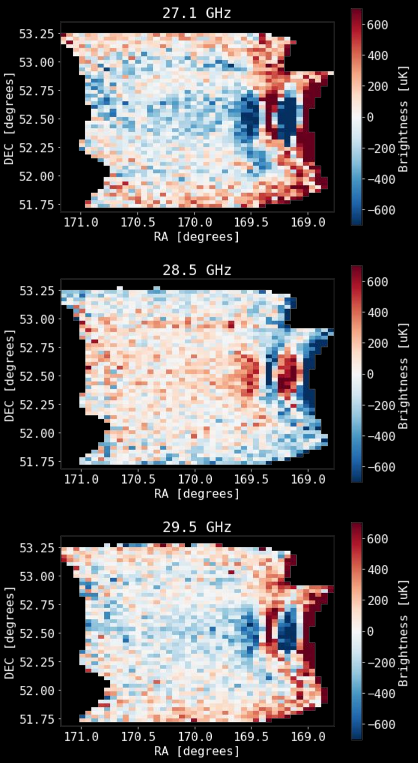
(This is a nonlinear problem that is more costly to solve than a PCA)

This is actually the maximum likelihood solution to the data model:

$$D_{\nu,p} = a_\nu m_p^T + n_{\nu,p}$$

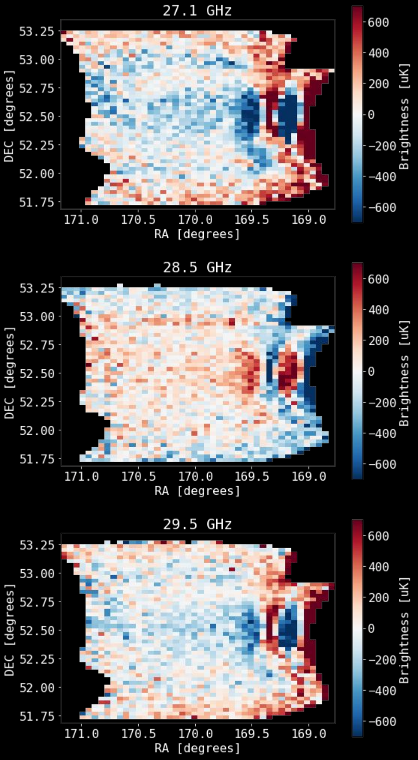
This works very well!

Unfiltered maps

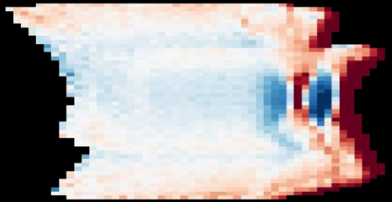


This works very well!

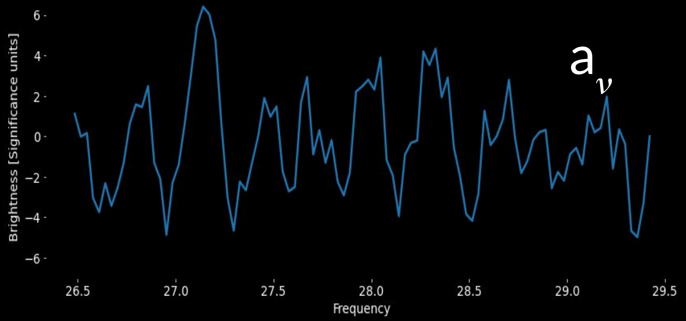
Unfiltered maps



PCA Template

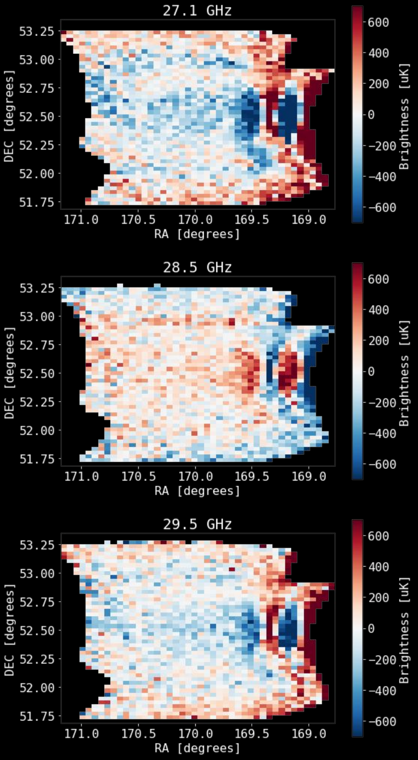


x

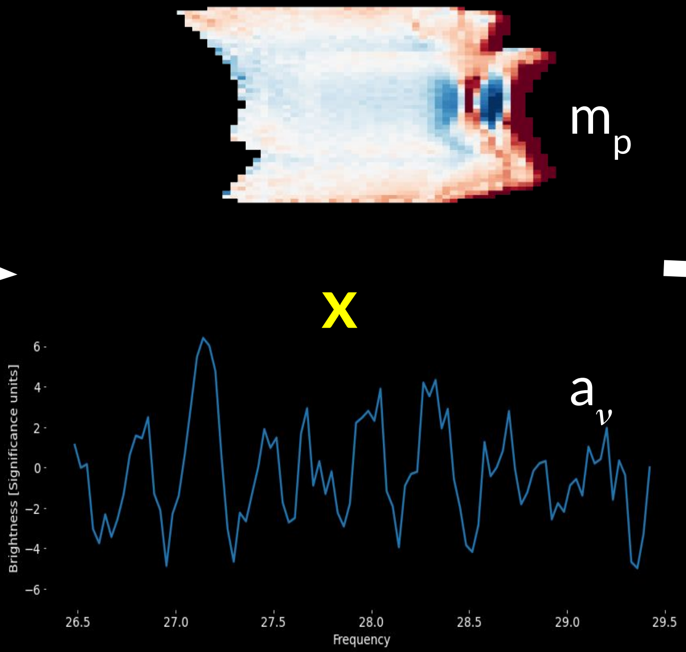


This works very well!

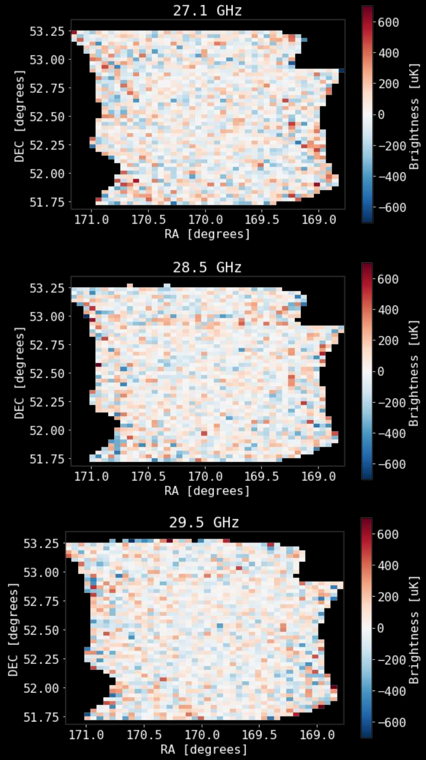
Unfiltered maps



PCA Template



Filtered maps

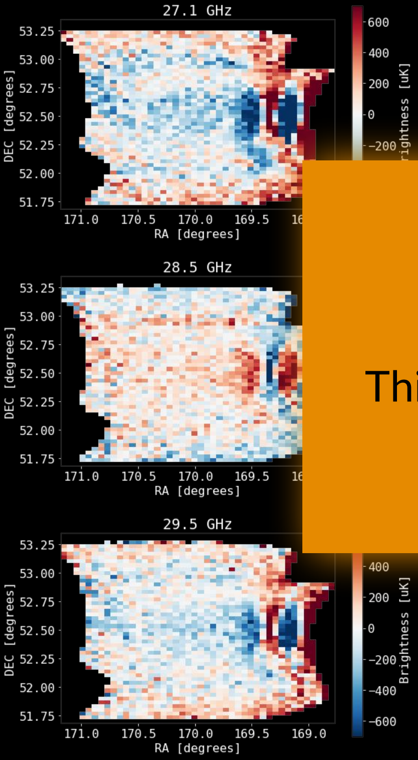


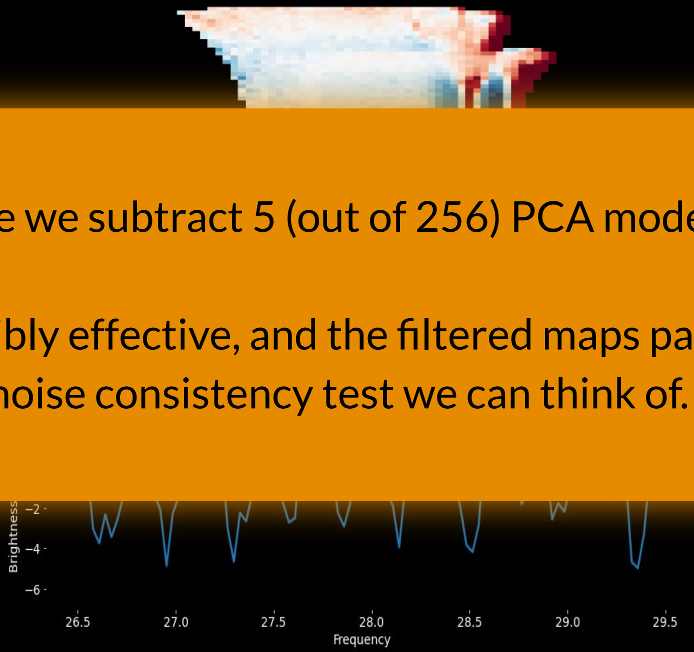
This works very well!

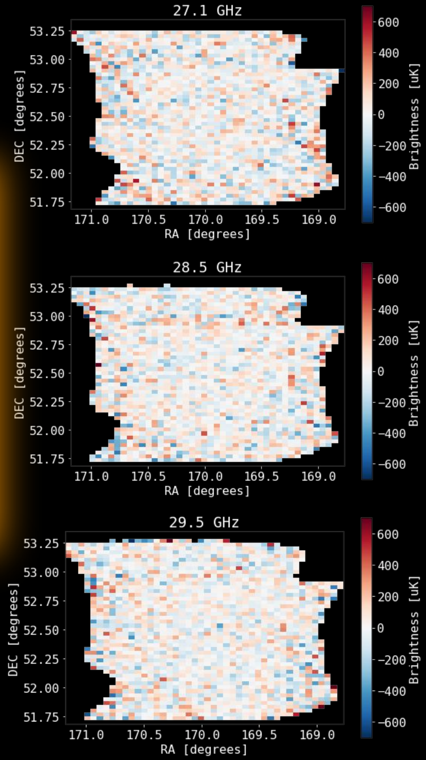
Unfiltered maps

PCA Template

Filtered maps





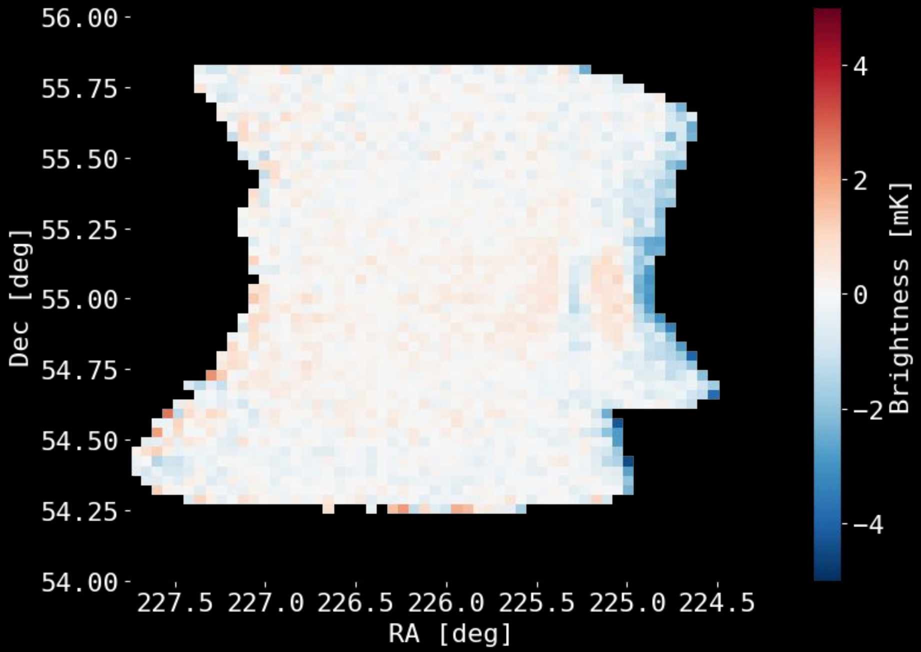


In practice we subtract 5 (out of 256) PCA modes.

This is incredibly effective, and the filtered maps pass any white-noise consistency test we can think of.

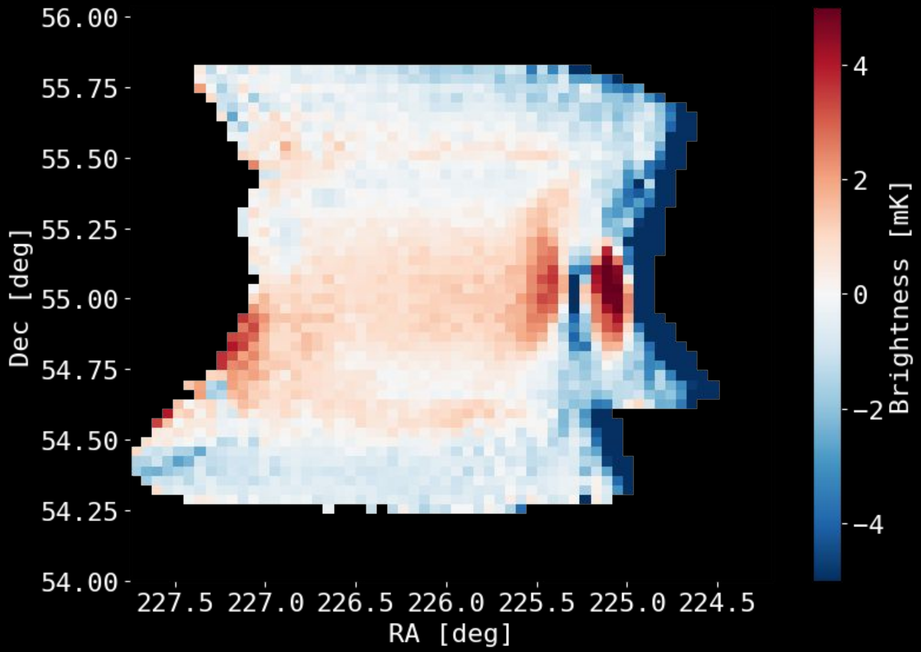
MCMC sampling the full posterior

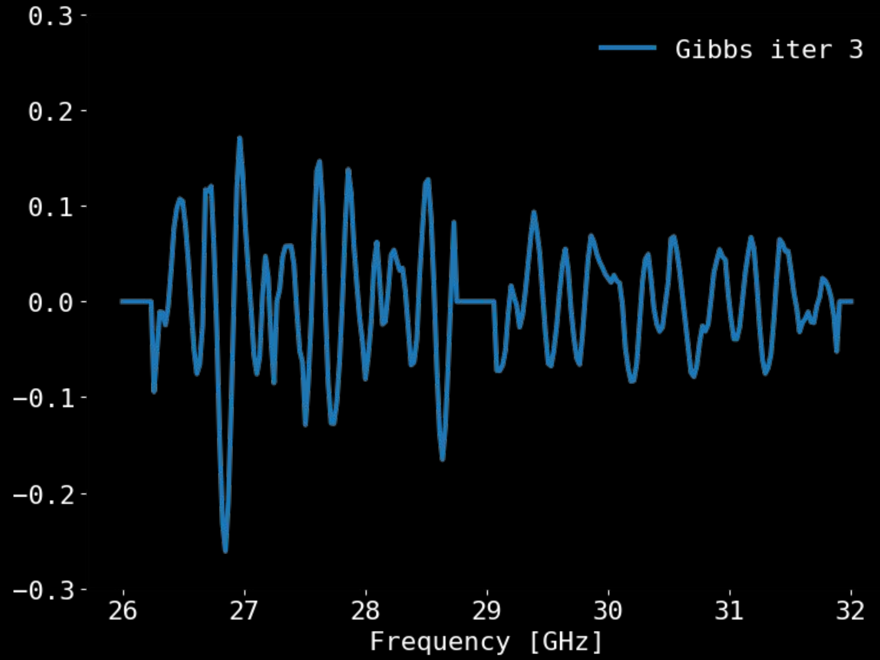
Gibbs sampling from $\sim e^{-\frac{1}{2}(D_{\nu,p}-a_\nu m_p^T)^2/\sigma_{\nu,p}^2}$



MCMC sampling the full posterior

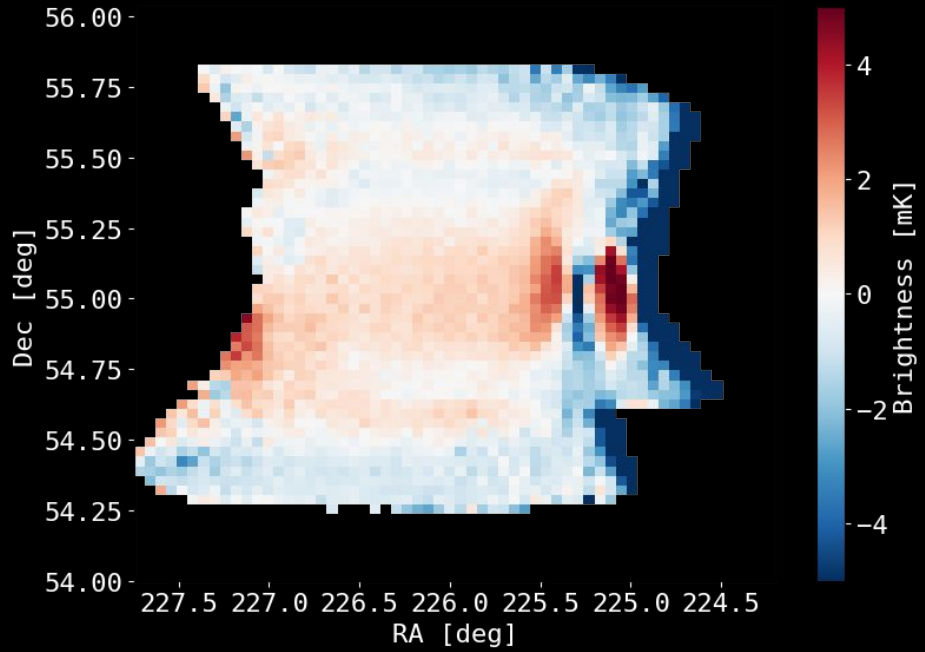
Gibbs sampling from $\sim e^{-\frac{1}{2}(D_{\nu,p} - a_\nu m_p^T)^2 / \sigma_{\nu,p}^2}$

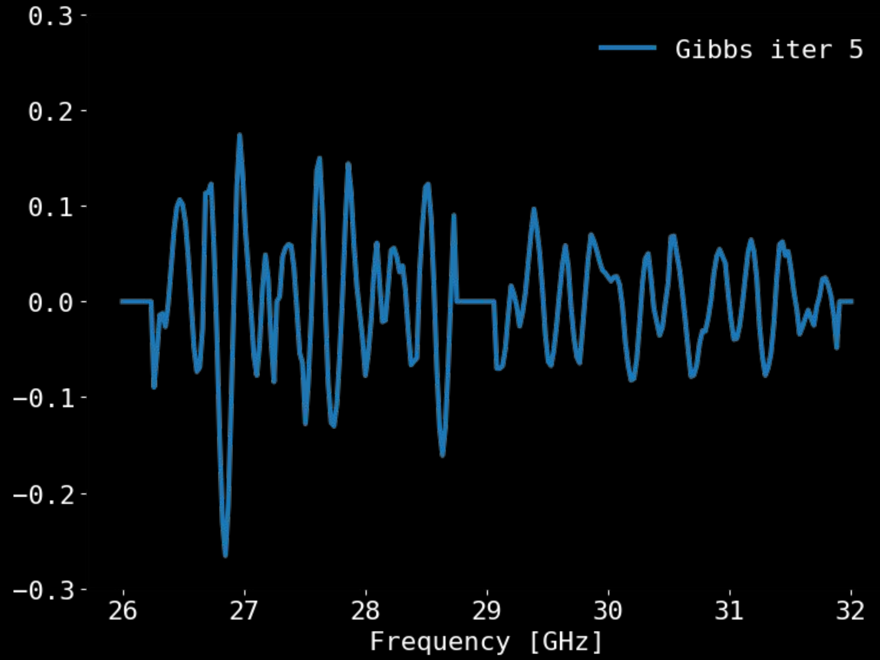




MCMC sampling the full posterior

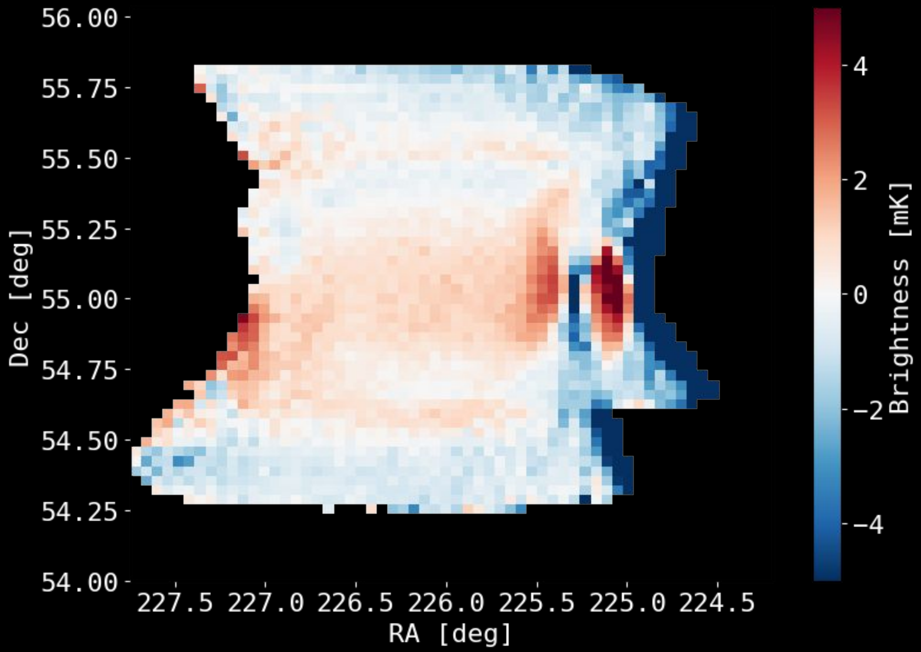
Gibbs sampling from $\sim e^{-\frac{1}{2}(D_{\nu,p}-a_\nu m_p^T)^2/\sigma_{\nu,p}^2}$

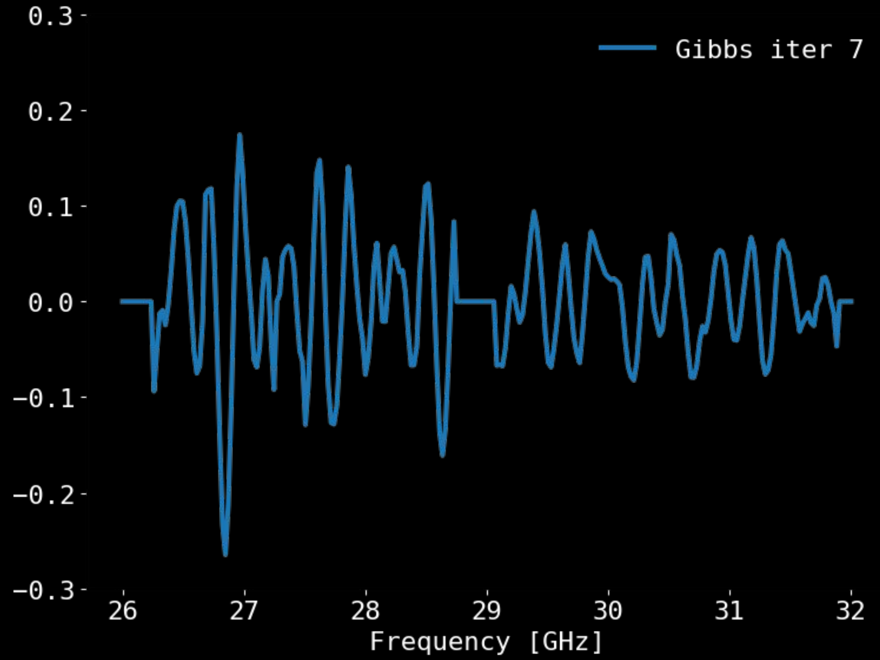




MCMC sampling the full posterior

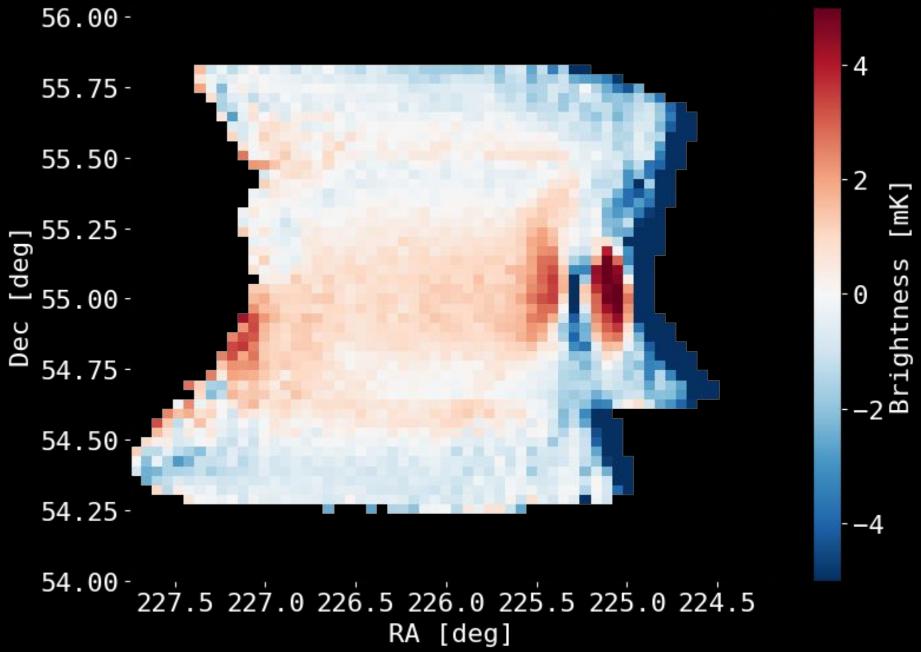
Gibbs sampling from $\sim e^{-\frac{1}{2}(D_{\nu,p} - a_\nu m_p^T)^2/\sigma^2_{\nu,p}}$

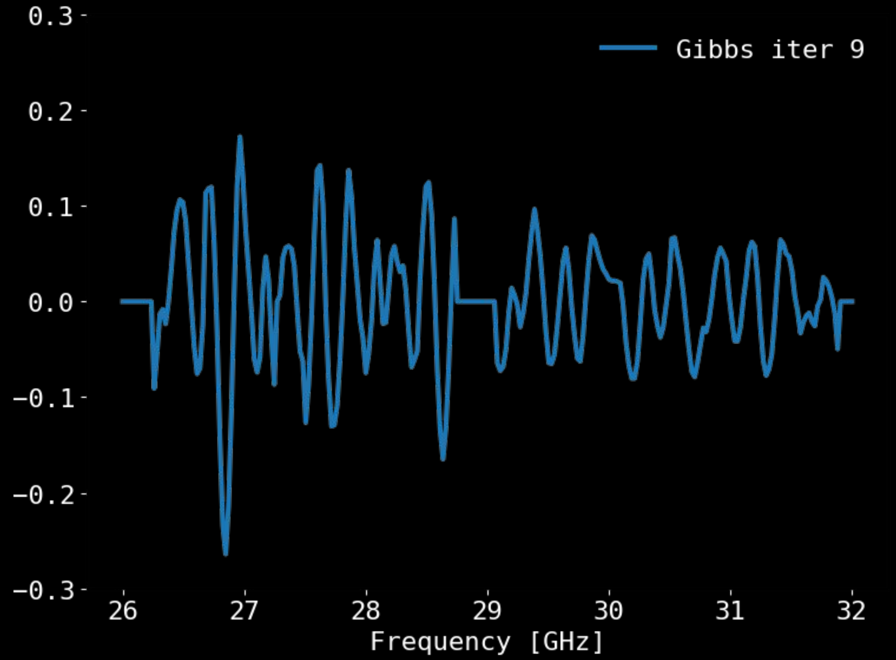




MCMC sampling the full posterior

Gibbs sampling from $\sim e^{-\frac{1}{2}(D_{\nu,p}-a_\nu m_p^T)^2/\sigma_{\nu,p}^2}$





Transfer function estimation

- COMAP uses a transfer function to correct signal lost to filtering.

$$P^{\text{true}}(\mathbf{k}) = P^{\text{est}}(\mathbf{k}) \cdot T(\mathbf{k})$$

Transfer function estimation

- COMAP uses a transfer function to correct signal lost to filtering.

$$P^{\mathrm{true}}(\mathbf{k}) = P^{\mathrm{est}}(\mathbf{k}) \cdot T(\mathbf{k})$$

- This is estimated by passing simulated data through the pipeline.

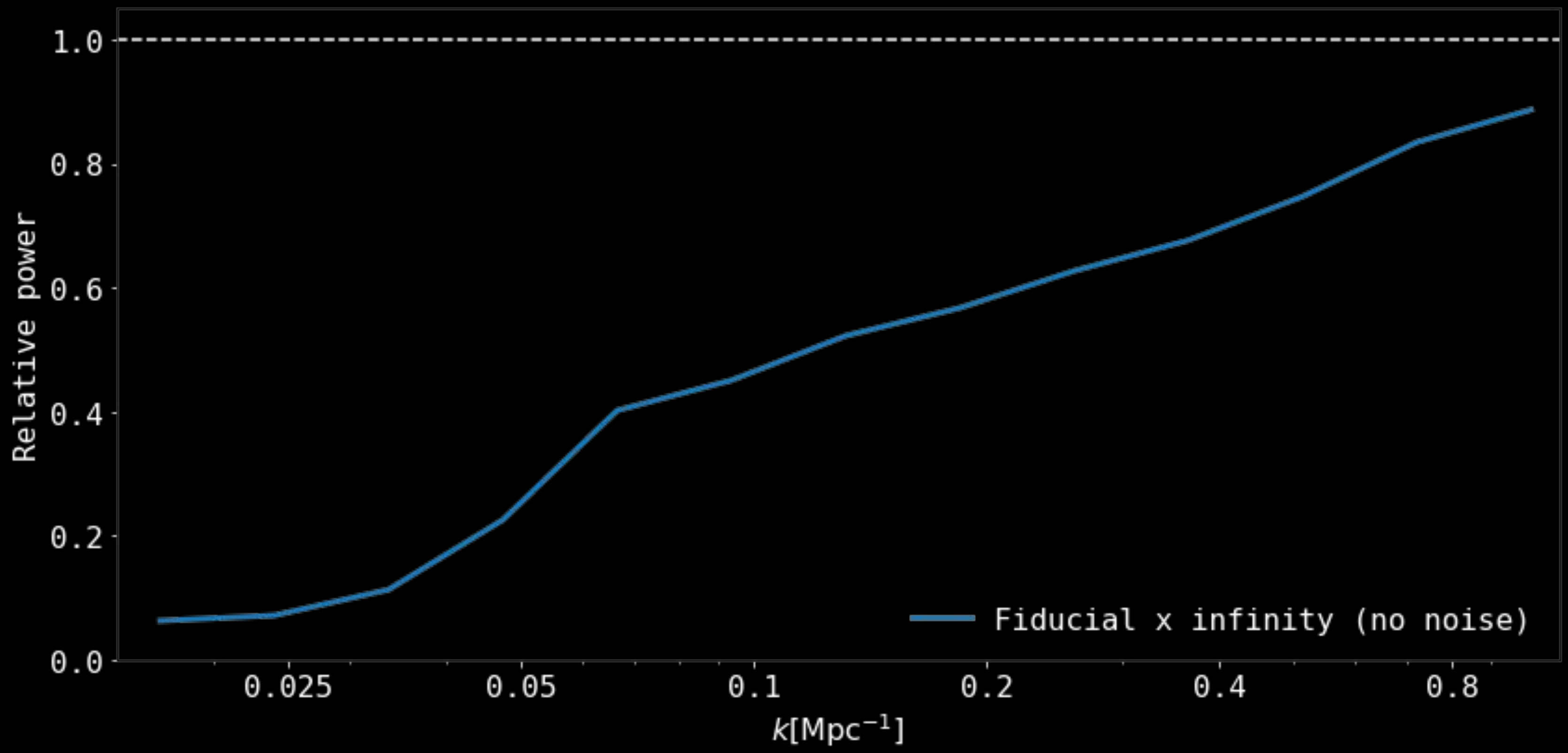
Transfer function estimation

- COMAP uses a transfer function to correct signal lost to filtering.

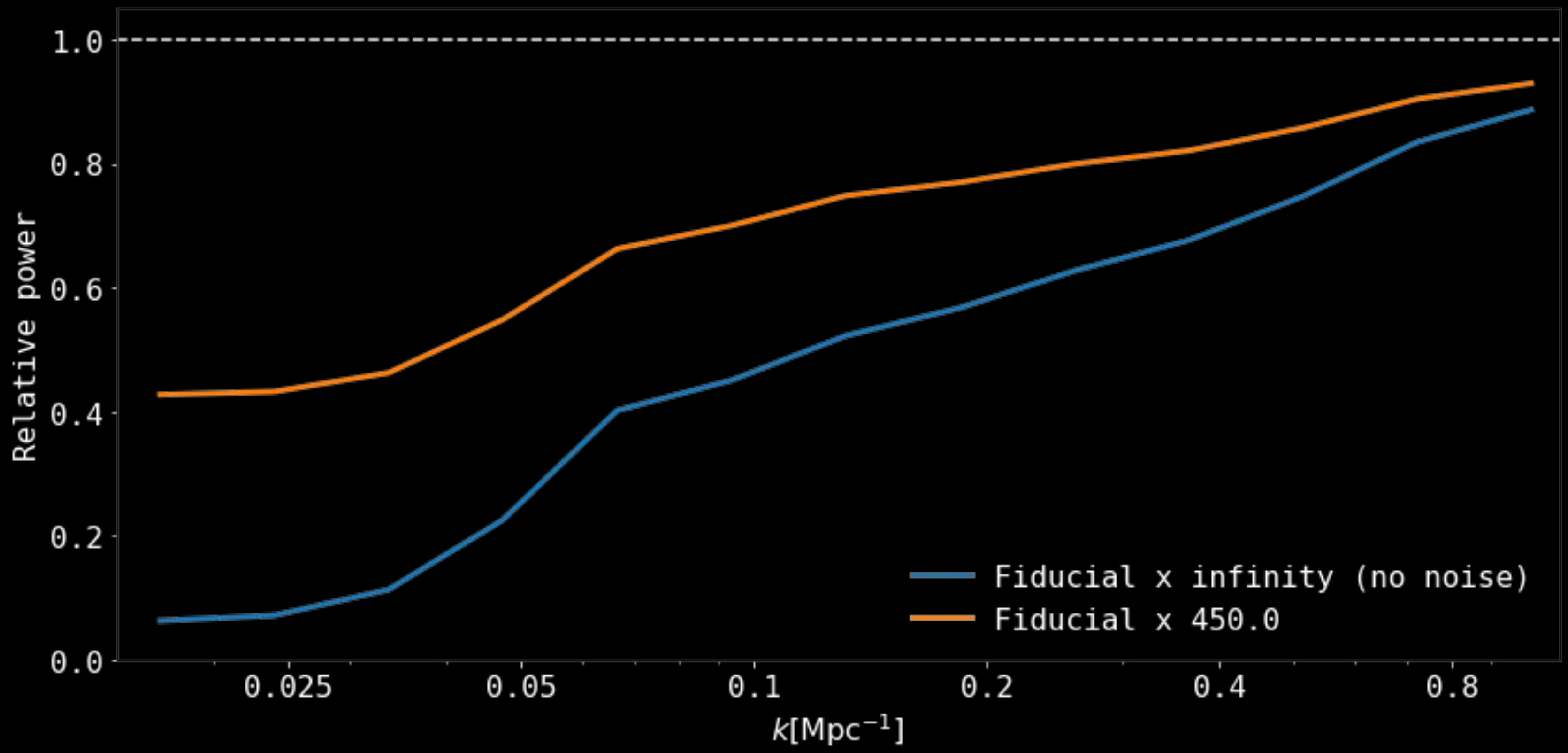
$$P^{\text{true}}(\mathbf{k}) = P^{\text{est}}(\mathbf{k}) \cdot T(\mathbf{k})$$

- This is estimated by passing simulated data through the pipeline.
- The PCA is a non-linear operation on the data. This makes the PCA transfer function depend on the signal-to-noise of the CO signal.

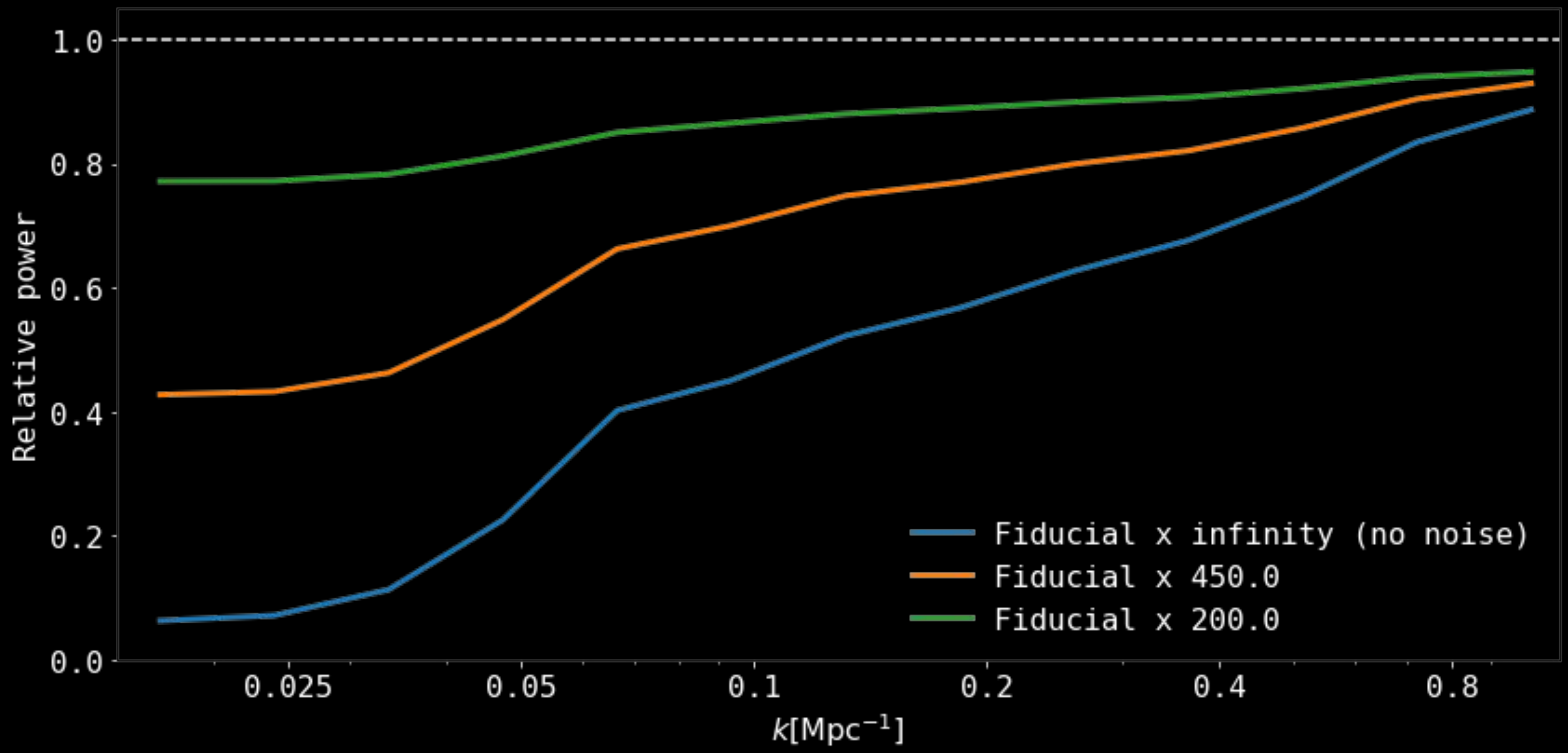
PCA transfer functions for different signal strengths



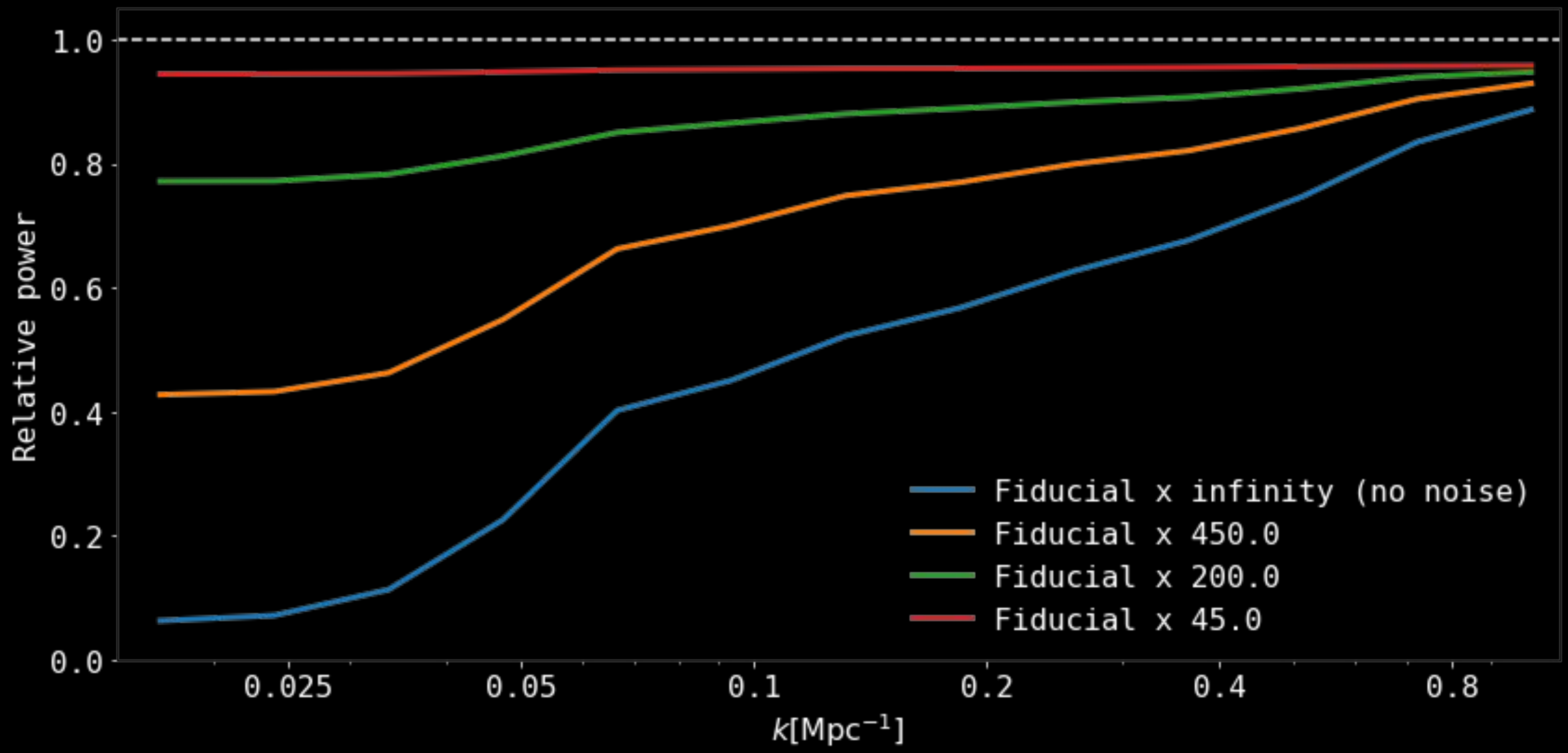
PCA transfer functions for different signal strengths



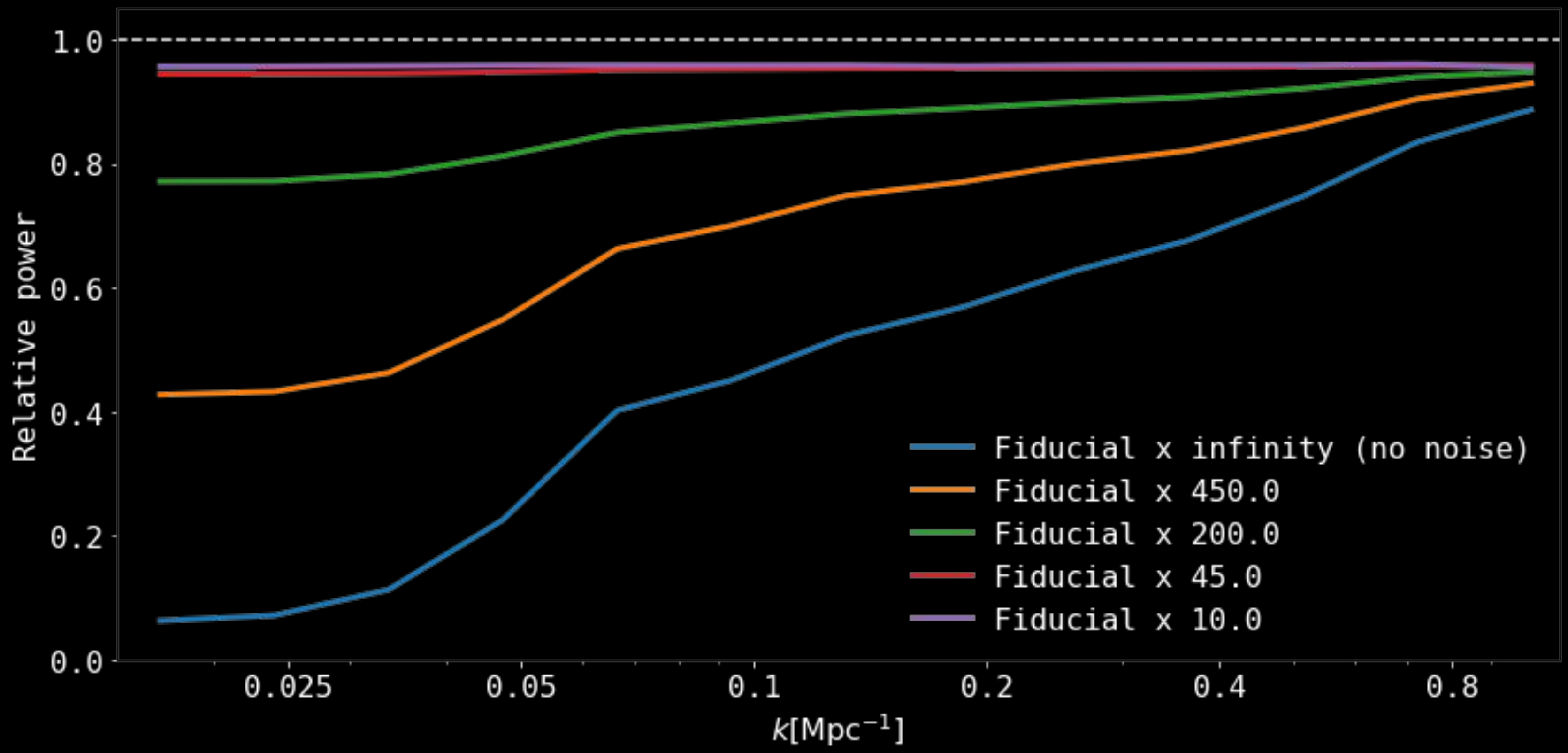
PCA transfer functions for different signal strengths



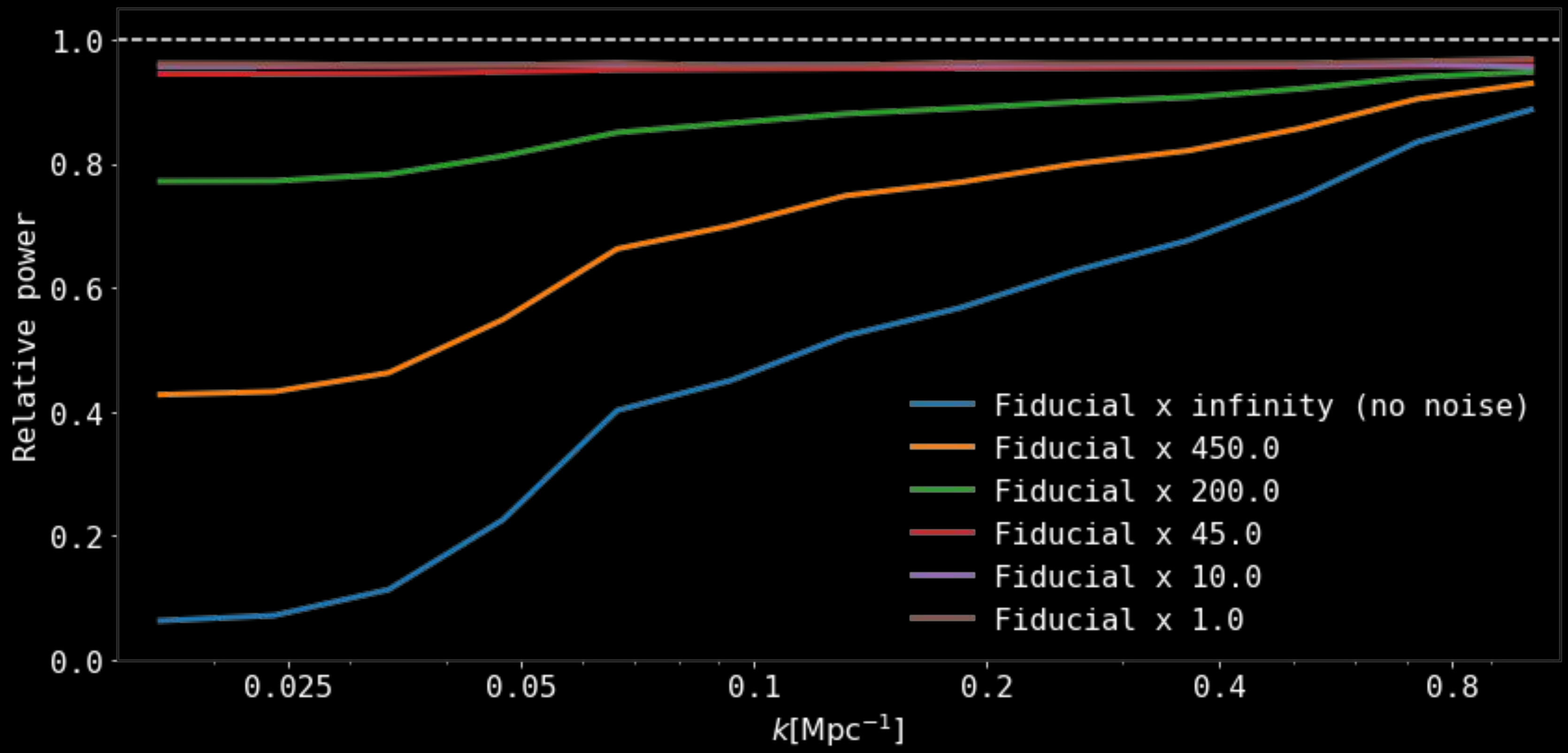
PCA transfer functions for different signal strengths



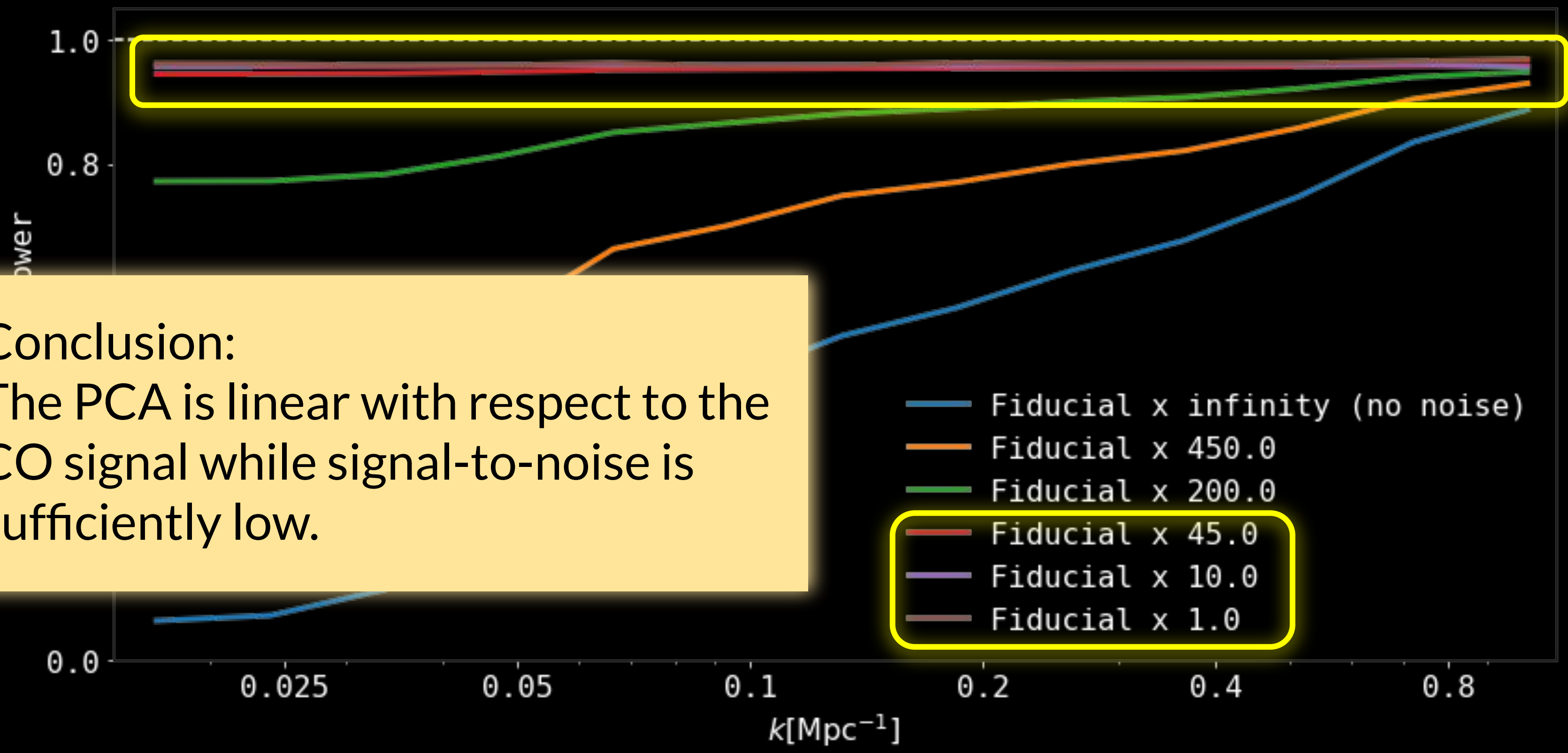
PCA transfer functions for different signal strengths



PCA transfer functions for different signal strengths



PCA transfer functions for different signal strengths



Conclusion:
The PCA is linear with respect to the CO signal while signal-to-noise is sufficiently low.

Summary

1. The raw COMAP maps are contaminated by standing wave systematics.

2. These are very effectively filtered with a generalized, noise-weighted PCA.
This involves solving the non-linear optimization problem:
$min\ (D_{v,p} - a_v m_p^T)^2/\sigma_{v,p}^2$

3. This is a non-linear operation on the data, meaning that any transfer function will in theory be a function of the signal model and amplitude.

4. However, we find that the PCA behaves linearly with respect to any sufficiently weak signal, which is the case for COMAP.

COMAP Pathfinder analysis
CO(1–0) power spectrum analysis

Nils-Ole Stutzer
Final month PhD fellow at
Institute of Theoretical Astrophysics,
University of Oslo

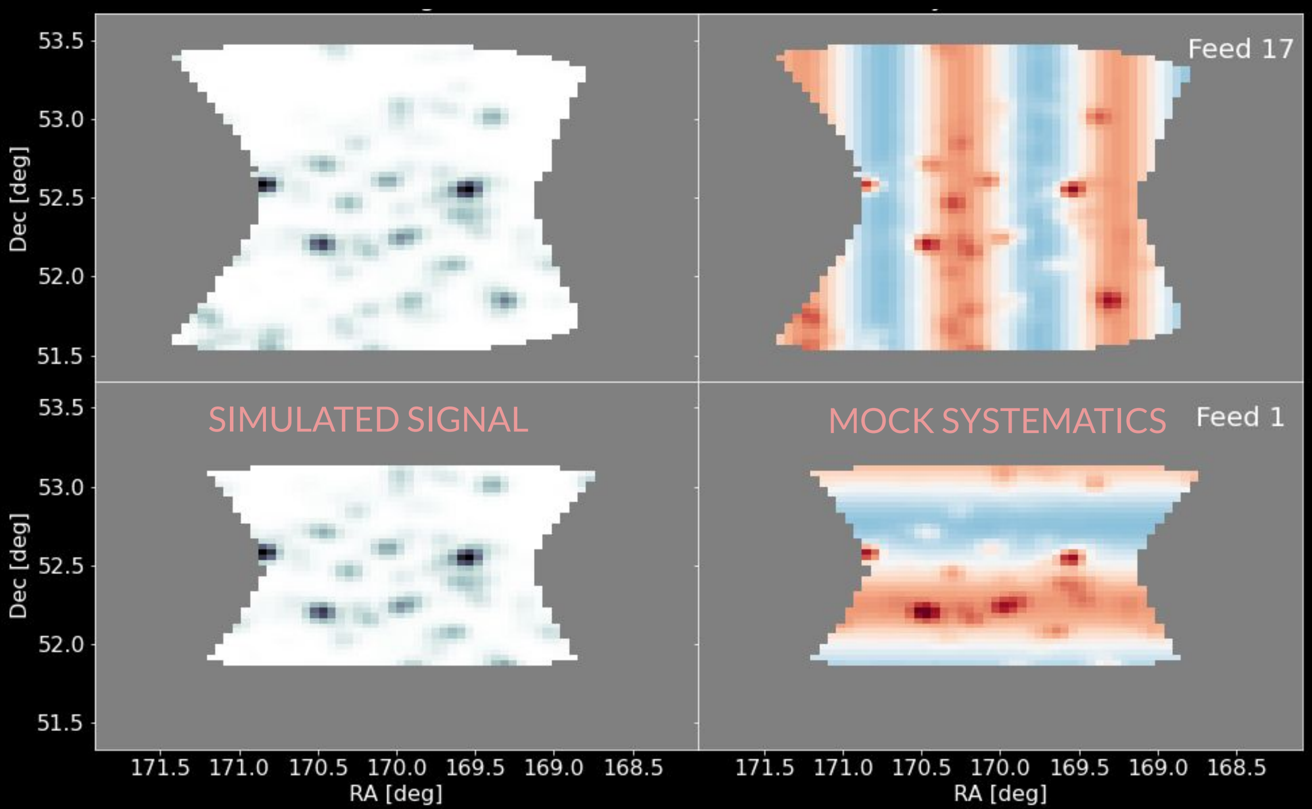
A second flavor of the power spectrum

Each feed (group) sees the same sky signal, but largely independent systematic effects.

Can isolate signal and cancel residual systematics with cross-power spectrum

$$C(k) \propto \langle \mathrm{Re}\{f_i^*(k) f_j(k)\}\rangle$$

$$\langle C(k)\rangle = P_{\mathrm{CO}}(k) + \cancel{C_{\mathrm{common}}(k)}$$



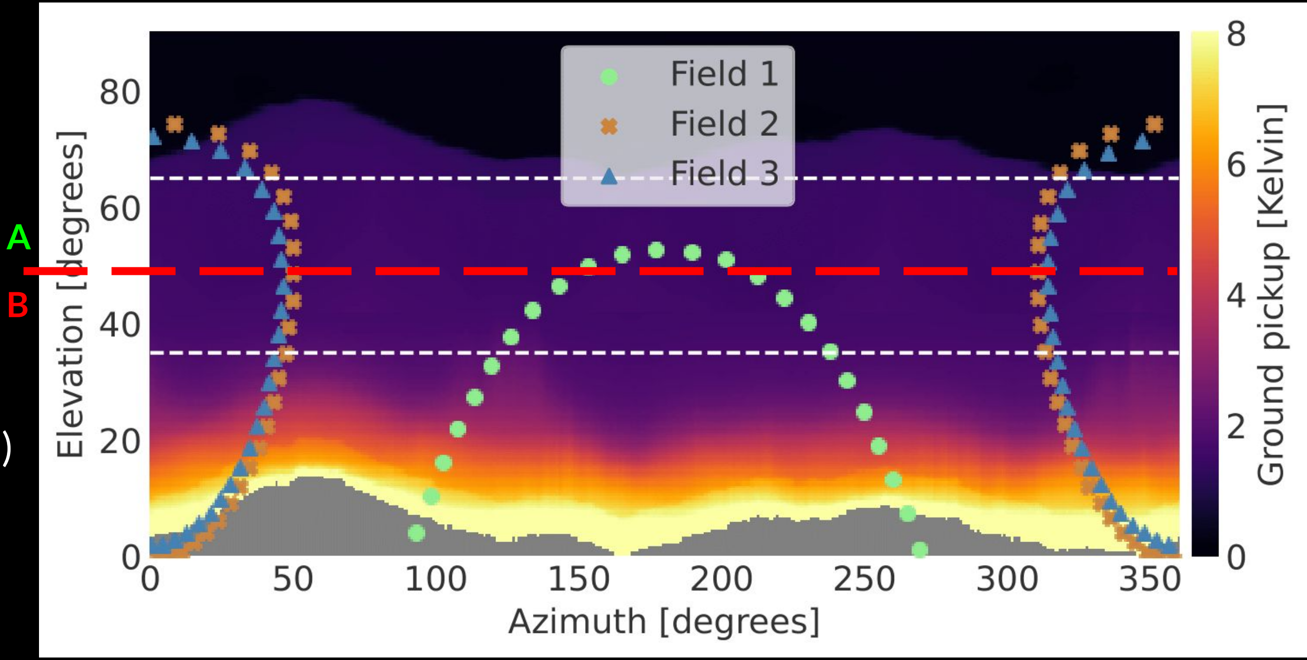
Methodology: How do we do it?

Data splits to remove additional systematics from ground pickup

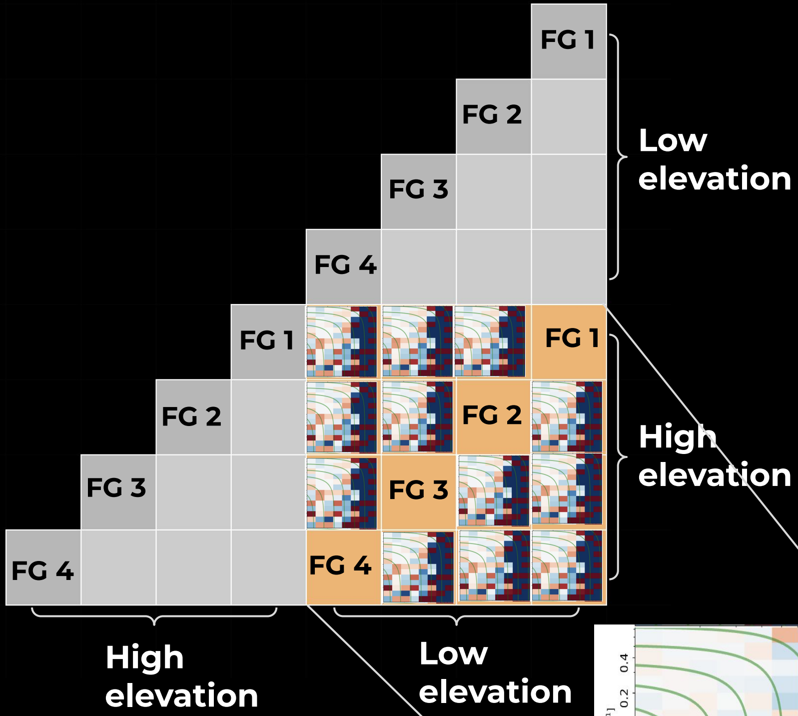
Split data into N parts:

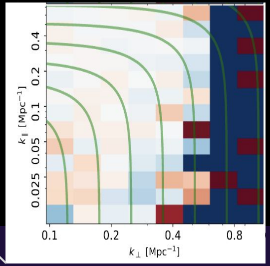
E.g. N = 2 elevation parts;
→ Part A above median elevation
→ Part B below median elevation

(Easy to generalize to N > 2
and other variables than elevation)

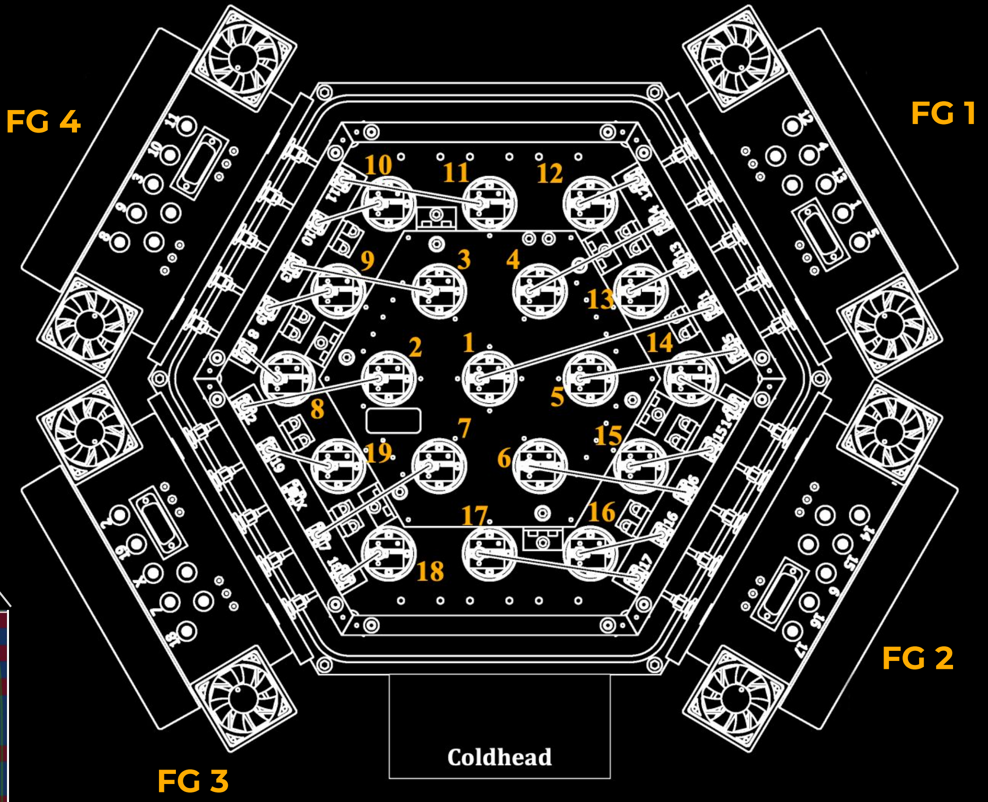


Feed(-Group) Pseudo Cross-Power Spectrum (FPXS/FGPXS)



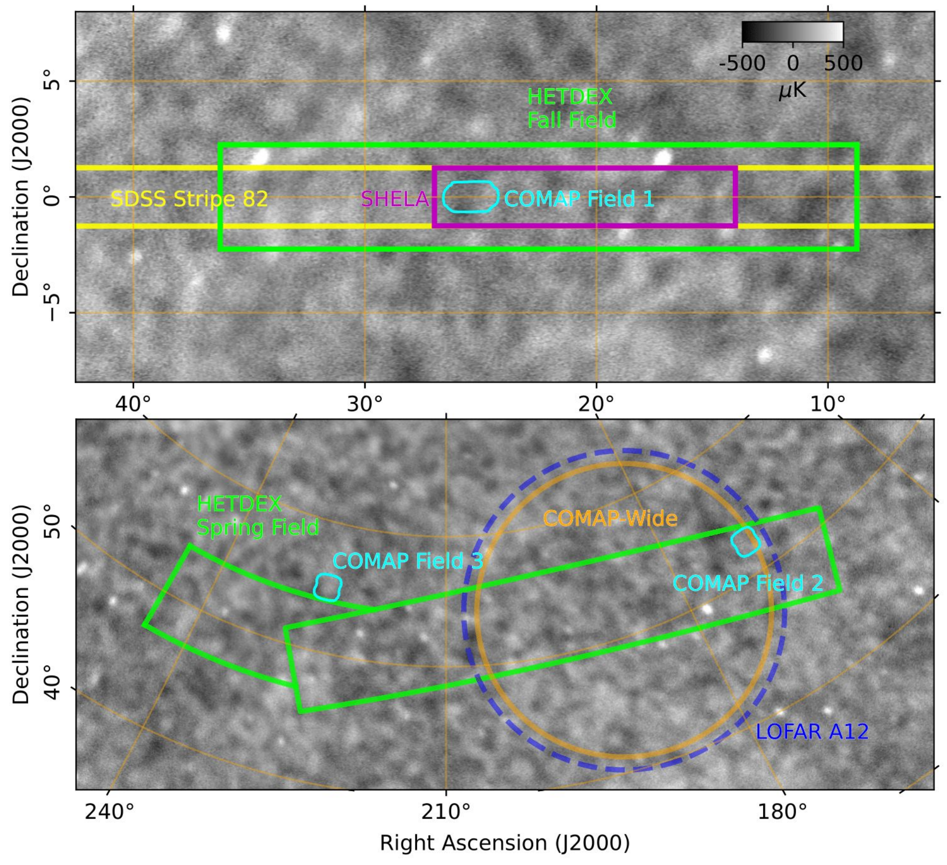


$$C(\mathbf{k}) = \frac{\sum_{A_i \neq B_j} \frac{C_{A_iB_j}(\mathbf{k})}{\sigma^2_{C_{A_iB_j}}(\mathbf{k})}}{\sum_{A_i \neq B_j} \frac{1}{\sigma^2_{C_{A_iB_j}}(\mathbf{k})}}$$

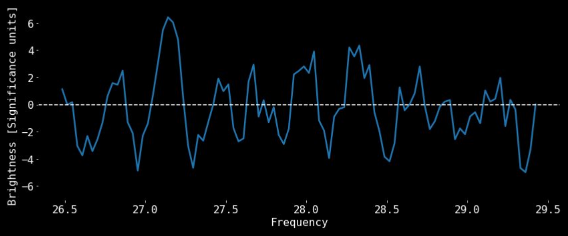


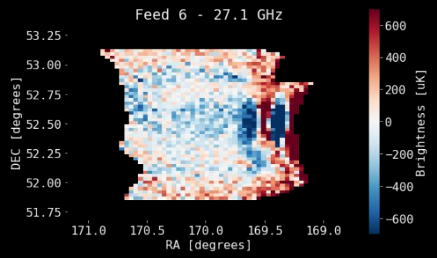
COMAP-Pathfinder X COMAP-Wide – Using two telescopes jointly

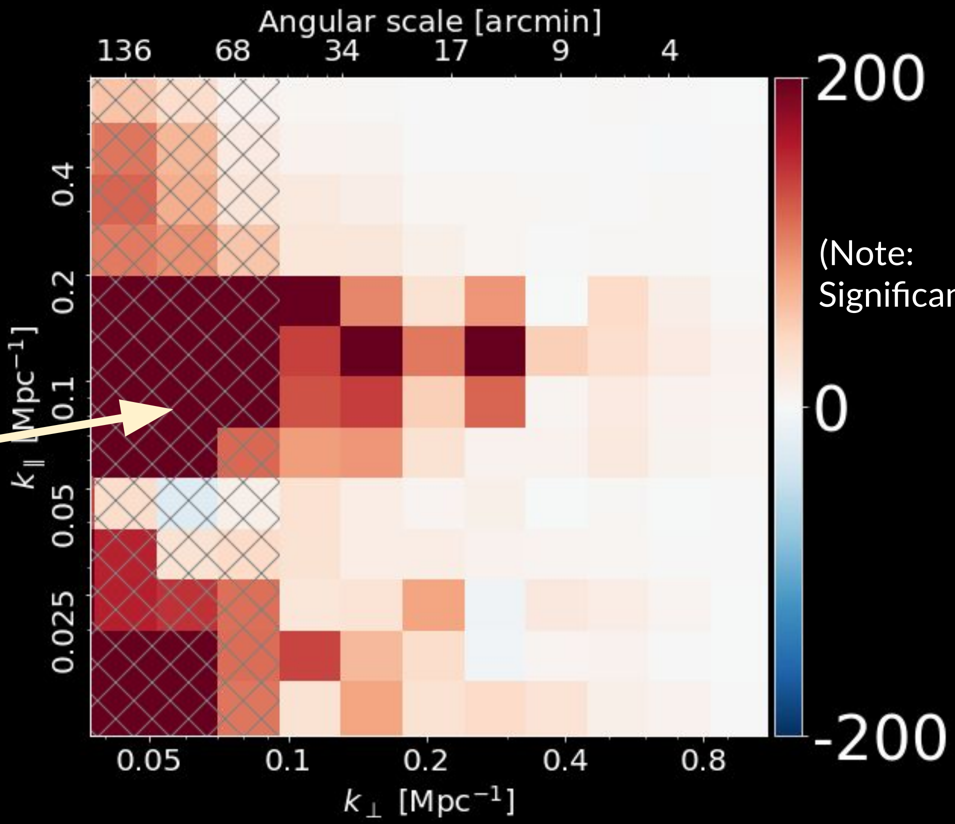
- COMAP-Wide → (near) copy COMAP-Pathfinder telescope
- Standing waves → the biggest systematic effect right now
- Can cross-correlate between two telescopes as well as feeds
- We could optimize the scanning strategy for each telescope to alternate between Pathfinder and Wide observations.



Auto-feed-group spectrum

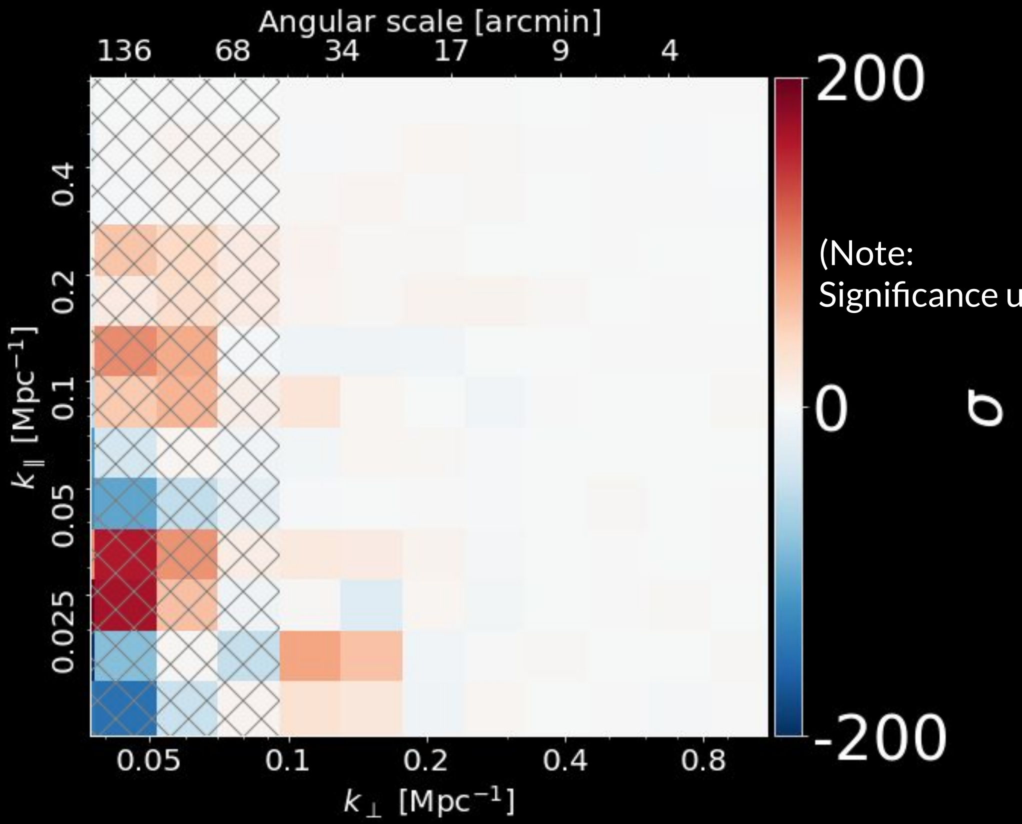






(Note:
Significance units!)

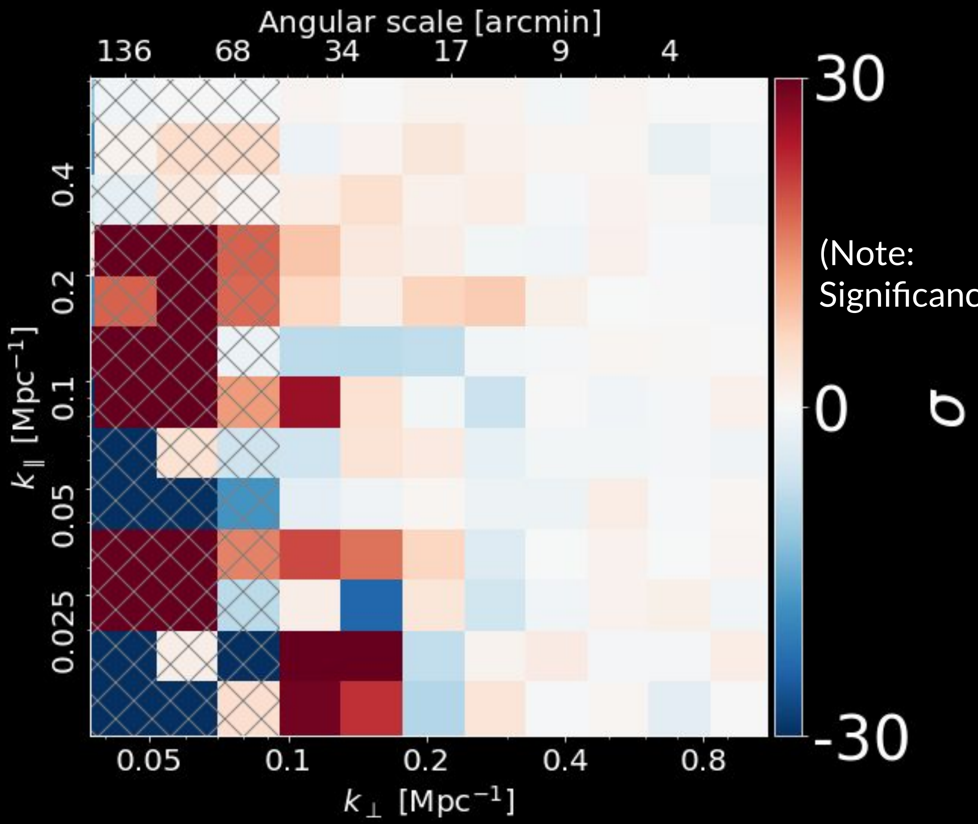
Cross-feed-group spectrum



(Note:
Significance units!)

Cross-feed-group spectrum

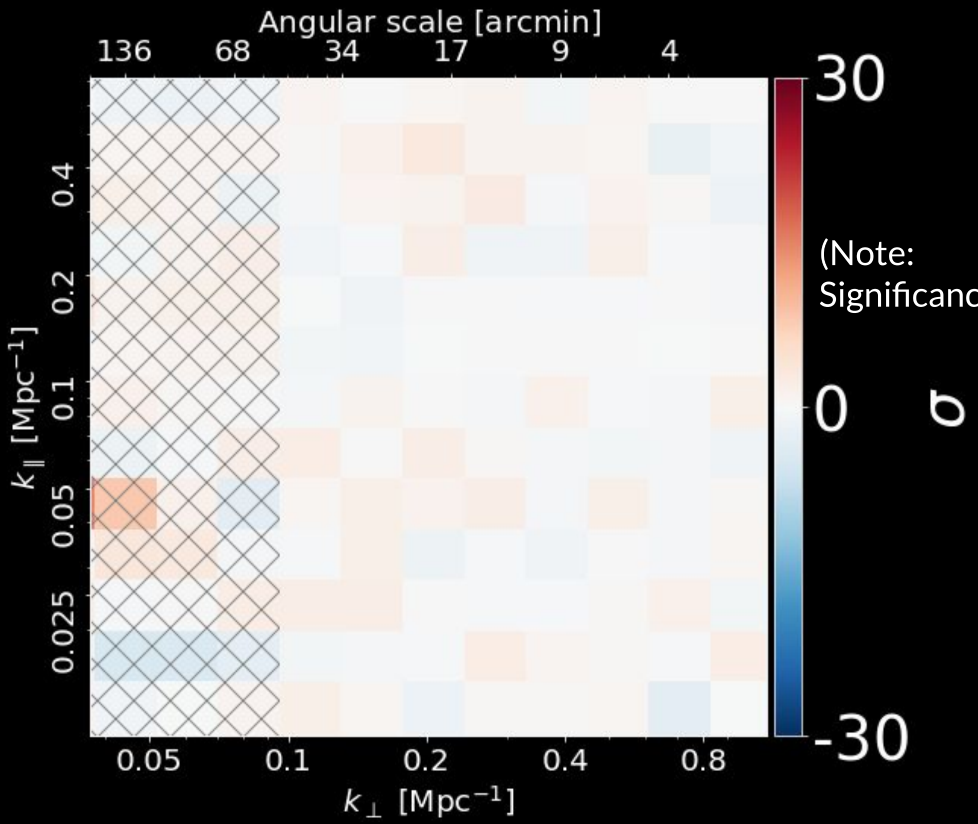
(zoomed in on the color scale)



(Note: Significance units!)

Cross-feed-group spectrum + map-PCA

(zoomed in on the color scale)

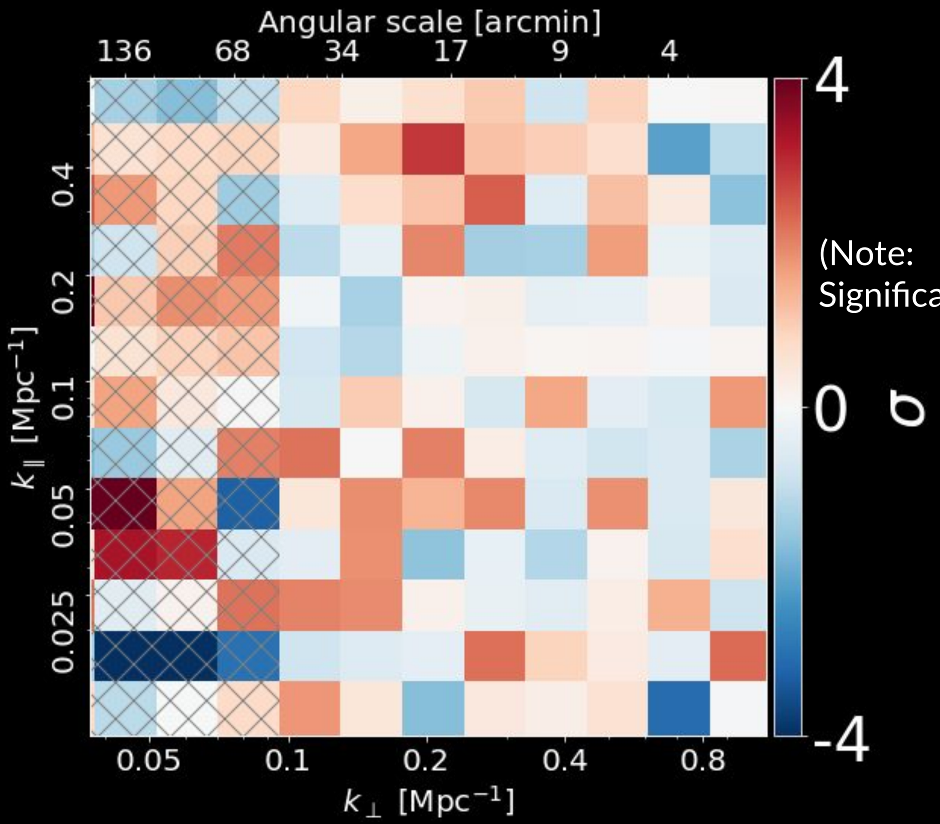


(Note: Significance units!)

Cross-feed-group spectrum + map-PCA

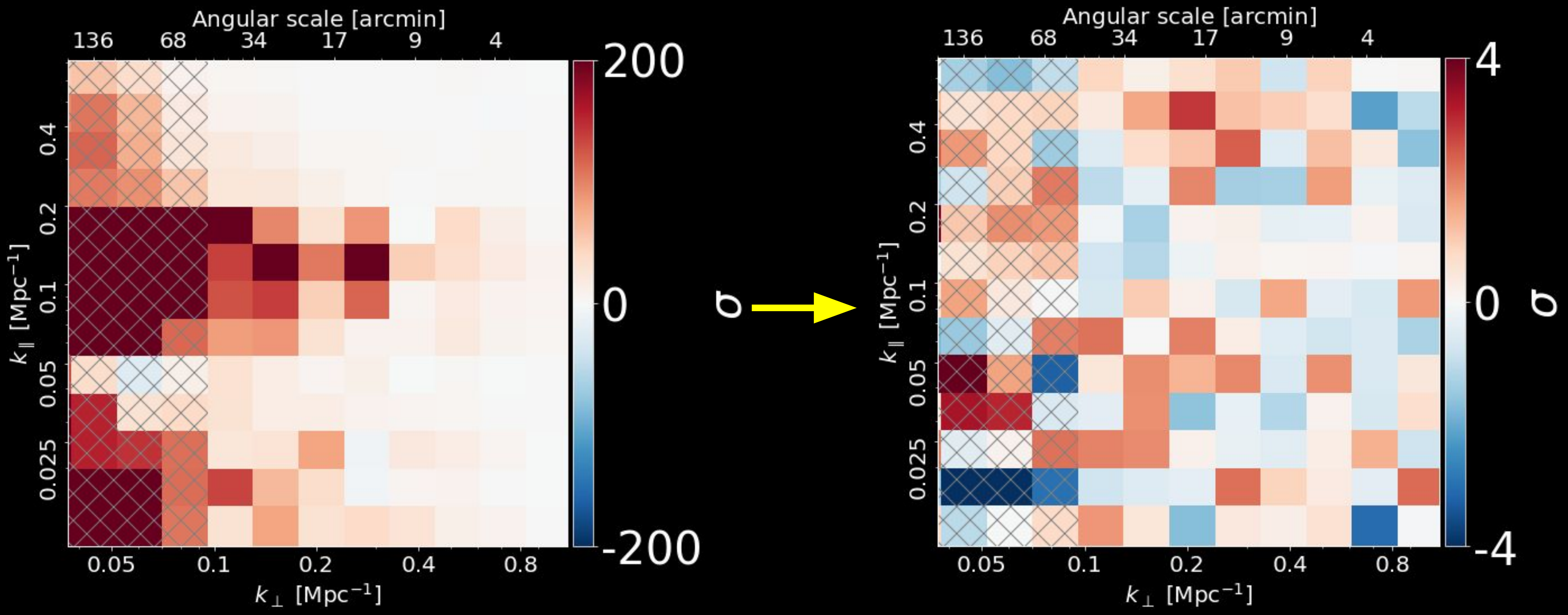
(zoomed in on the color scale)

(zoomed in on the color scale even more)

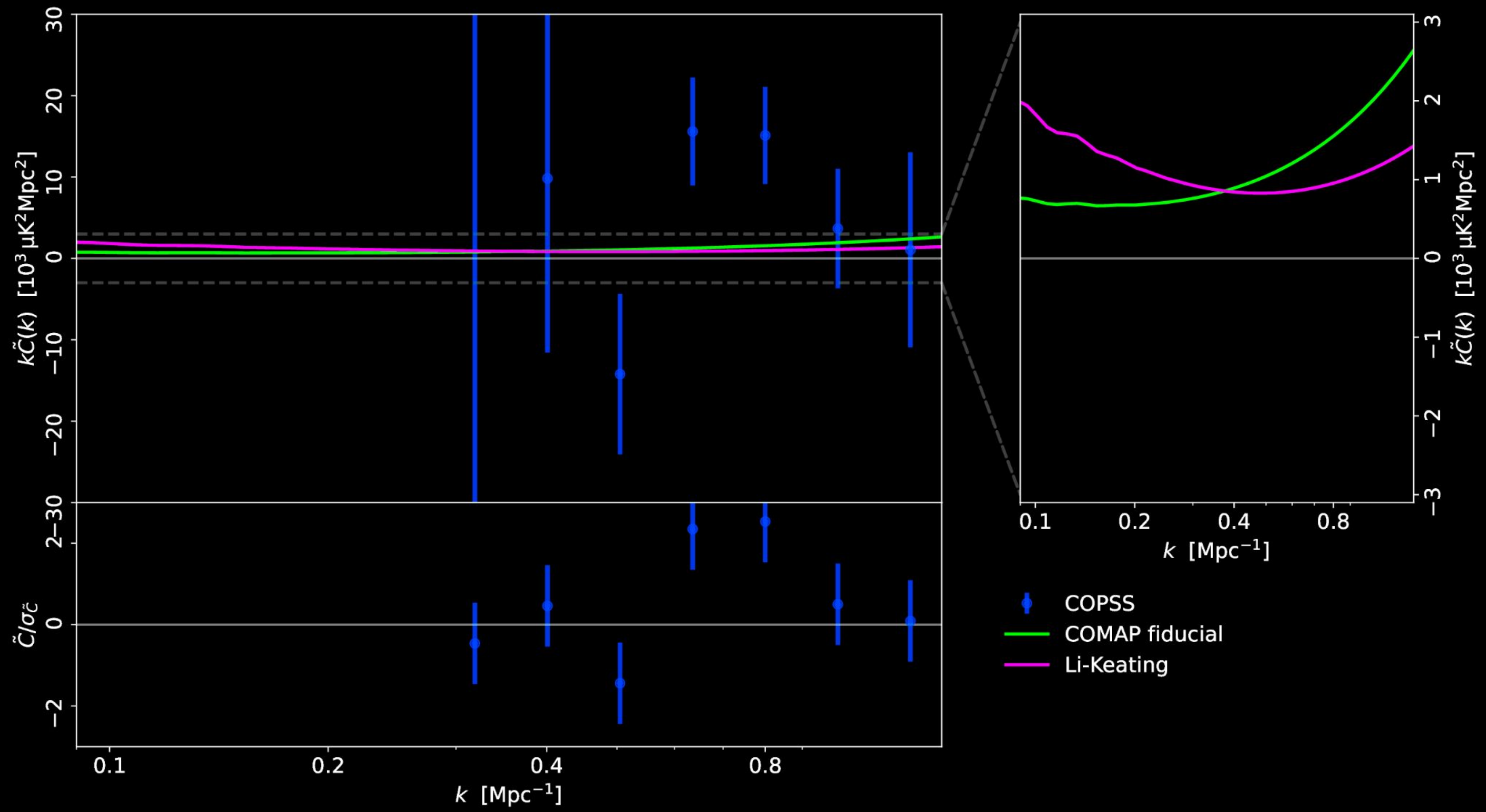


(Note: Significance units!)

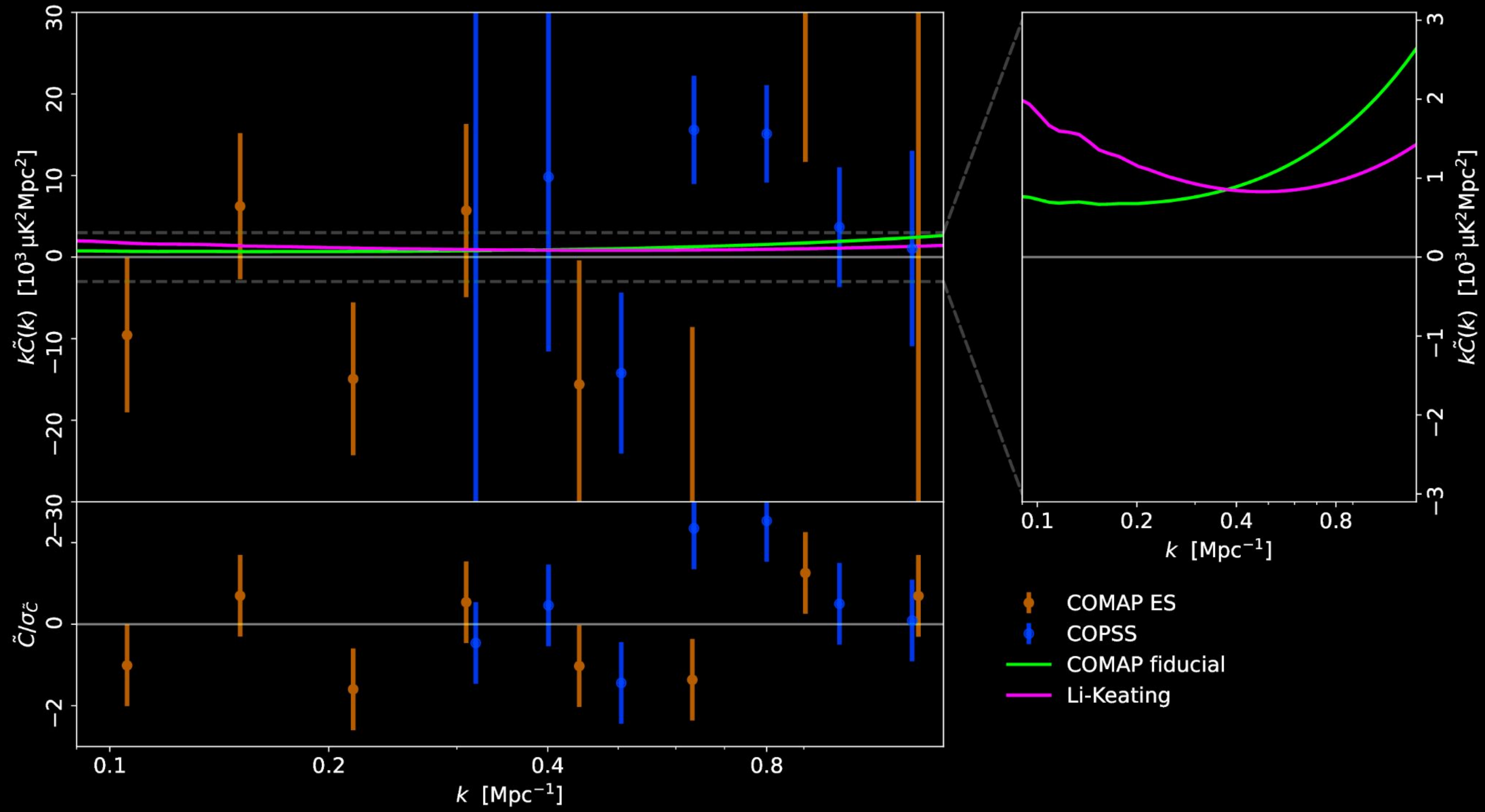
The side-by-side



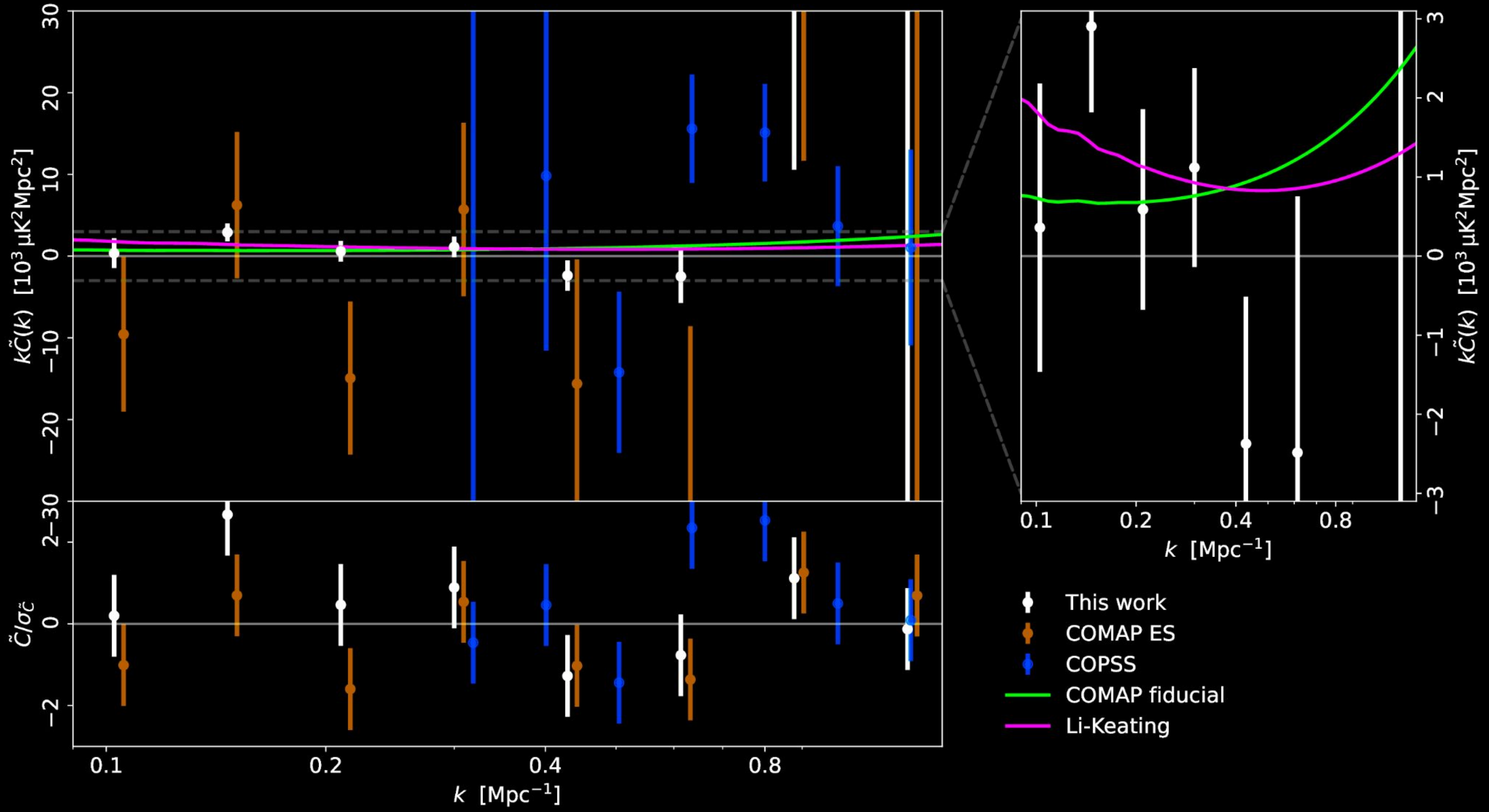
Results — Spherically-averaged power spectrum



Results — Spherically-averaged power spectrum



Results — Spherically-averaged power spectrum



Simulating ground pickup systematics in the time domain

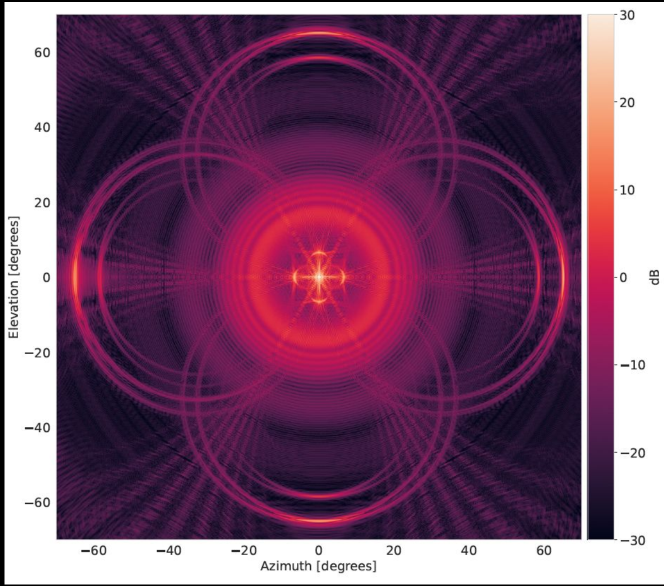
Leah E. Hansen
Masters student at
Institute of Theoretical Astrophysics,
University of Oslo

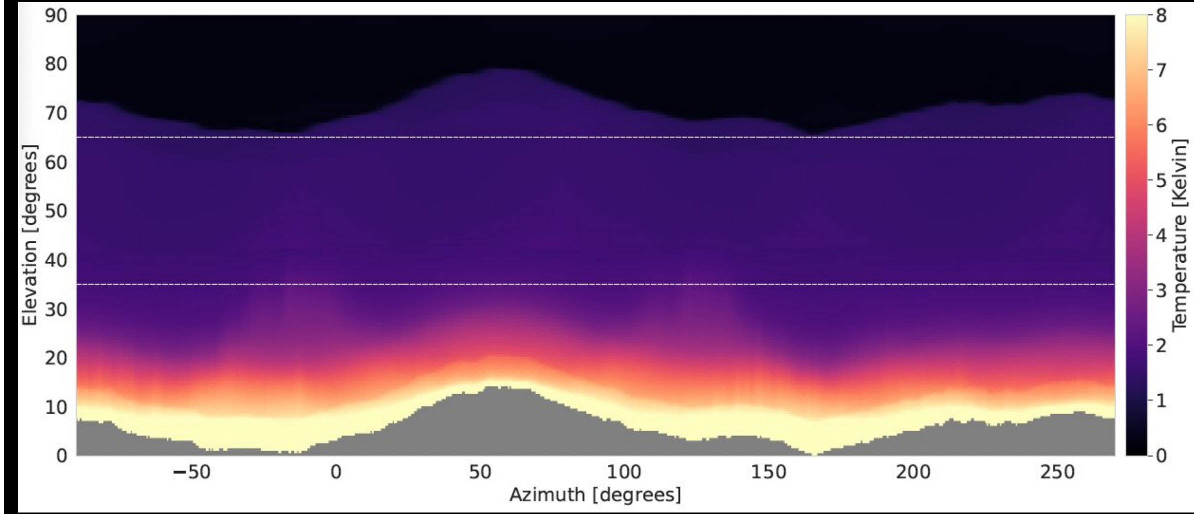
Towards an end-to-end COMAP simulation

Motivation for end-to-end simulation:

- Better understanding of interactions between systematics and pipeline
- Inform choices of cross power spectrum variables
- More robust transfer function analysis

Simulating the end-to-end pipeline: Ground pickup

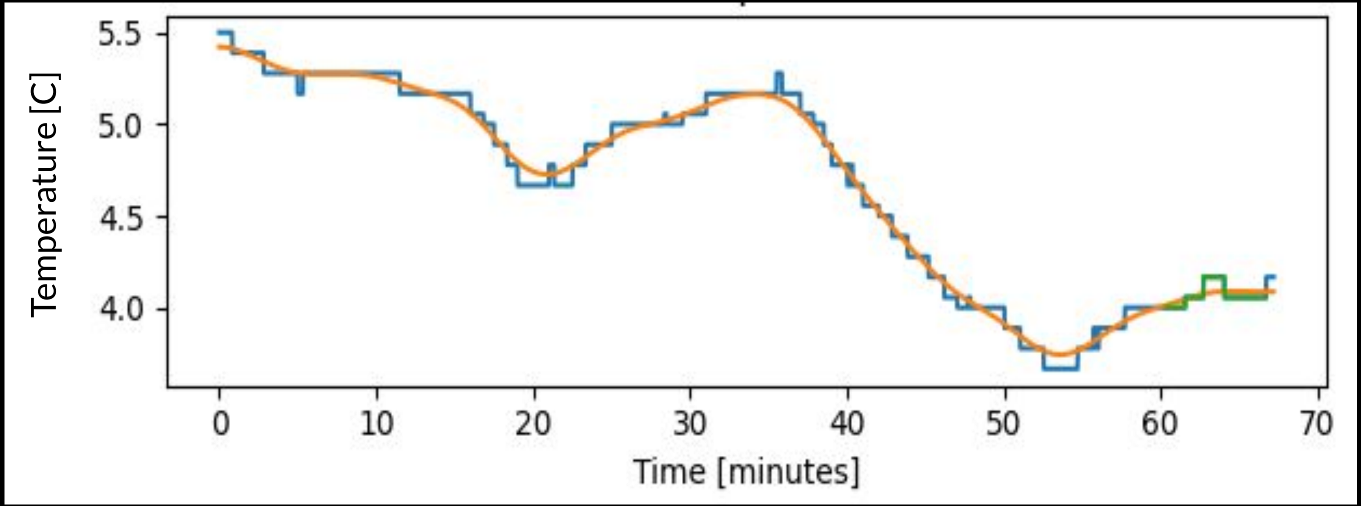




Full TOD data model:

$$d(\boldsymbol{v}, t) = G(\boldsymbol{v}) \times [Tsys(\boldsymbol{v}) + Tground(\boldsymbol{v}, t) + Tnoise(\boldsymbol{v}, t)]$$

Adding more realism to the ground pickup simulation



Added time variation to our model from air temperature data at OVRO.

Simulating ground pickup: component summary

Simulated beam

+

Real mountain profile around telescope

+

“Real” ground temperature variations

+

Real telescope pointing

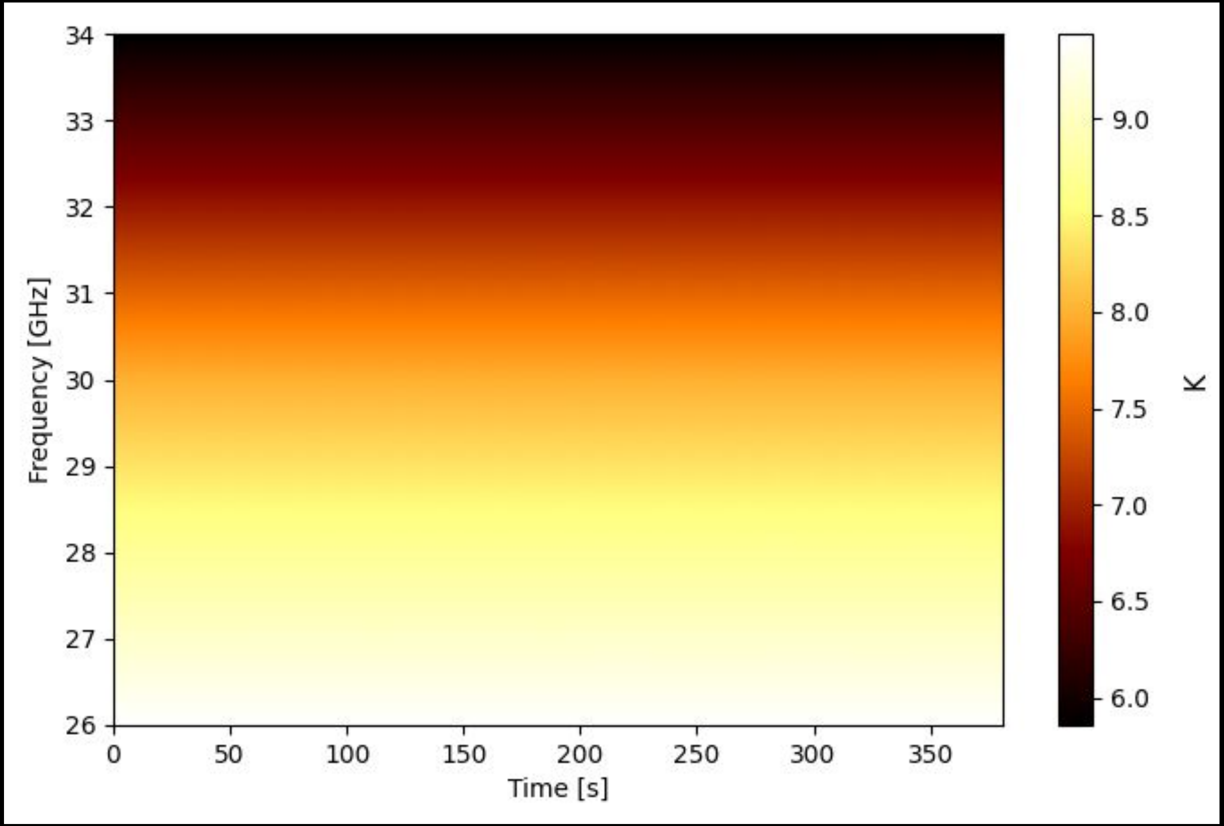
+

Real gain and system temperature

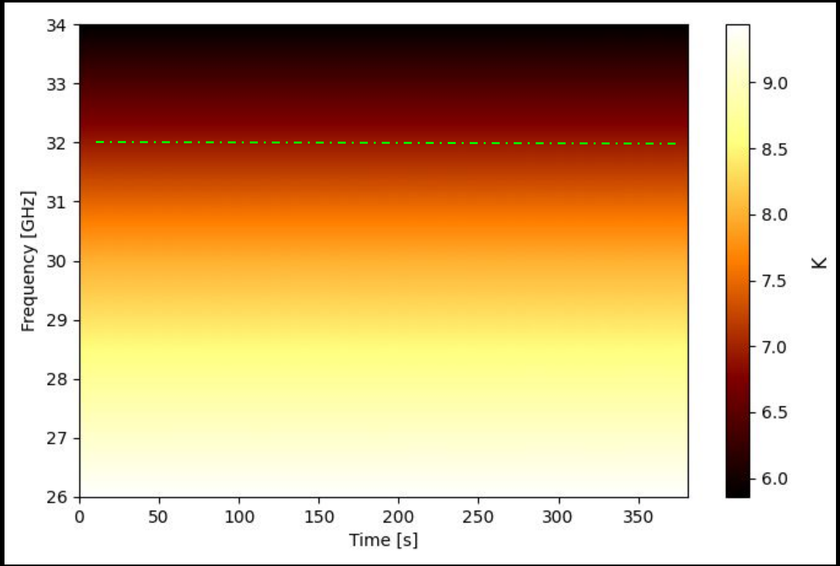
→

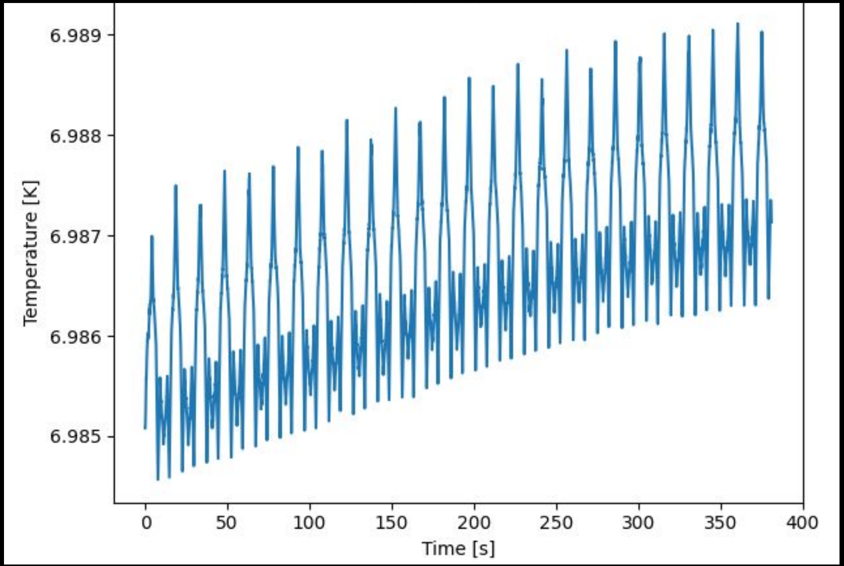
Simulated ground pickup TOD

Results: The ground pickup in time and frequency space



Results: The ground pickup in time and frequency space

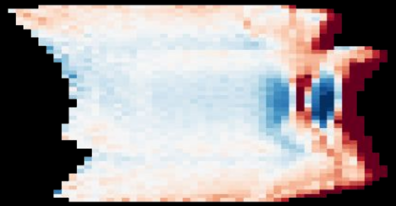


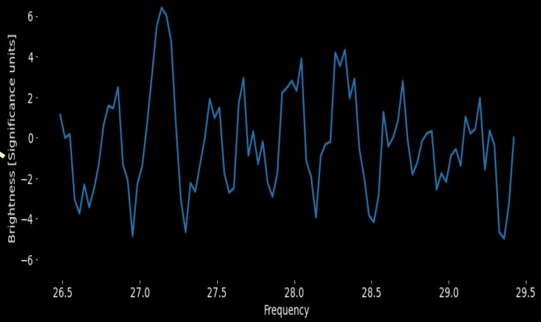


Future outlook

- Pipeline filtering and power spectrum analysis
- Standing wave simulation
- Pipeline and power spectrum analysis of this
- ++ More systematics

Thank you!





COMAP-*wide*

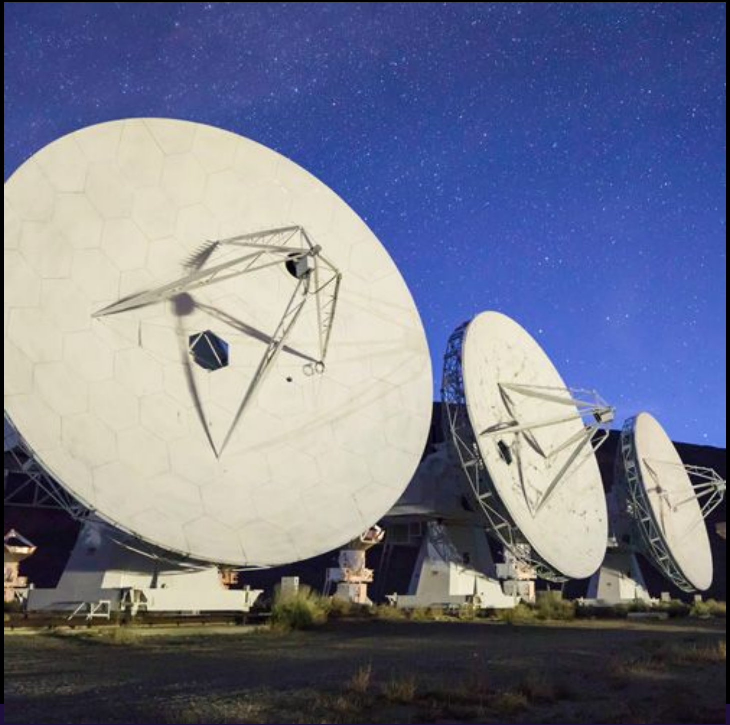
A Wide-Area CO Cross-Correlation Survey

Patrick C. Breysse
Southern Methodist University

COMAP-*wide*

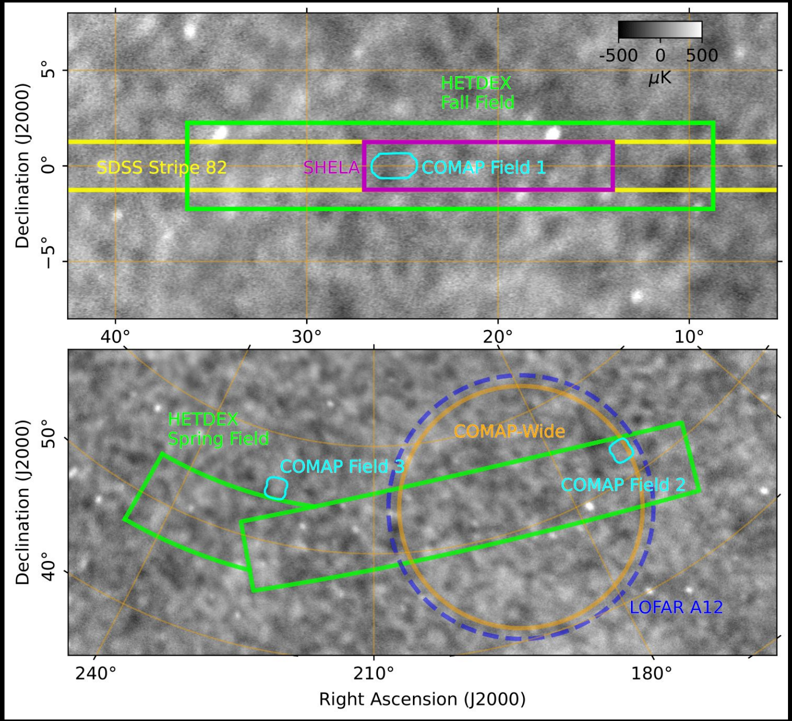
Duplicate* COMAP Pathfinder for wide-field intensity mapping

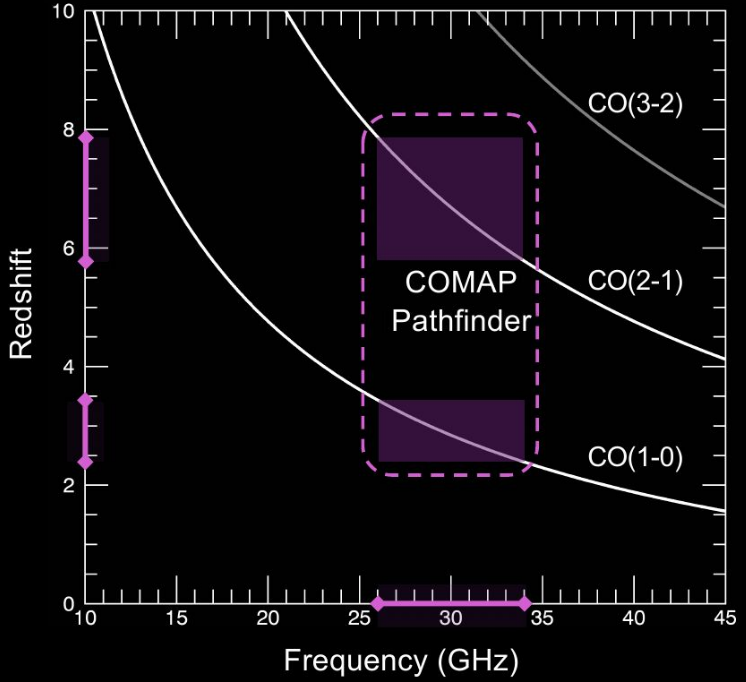
- 10.4 m antenna at Owens Valley Radio Observatory
- 19 feeds covering 26-34 GHz
- 2 year observing campaign on ~400 deg^2
- Science Targets: Cross-Correlations
  - CO x 21 cm at z~7
  - CO x Spectroscopic Galaxies at z~3



COMAP-*wide*

Survey area chosen to use both CO lines accessible at 30 GHz





Why go Wide?

Autocorrelation Errors - Knox Criterion

For a measurement of the auto spectrum,

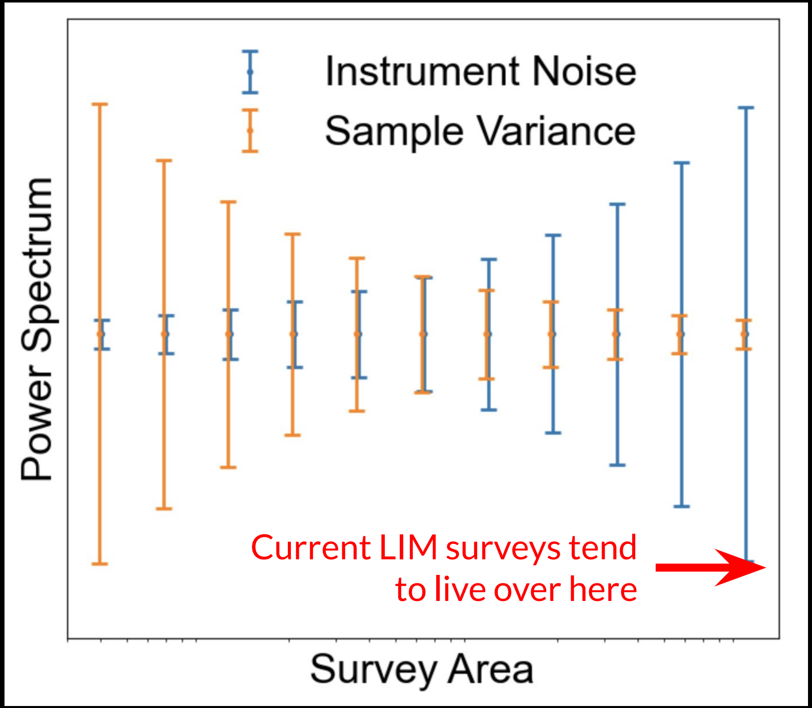
$$P_{\text{auto}}(k) = \langle T_{\text{CO}} T^{*}_{\text{CO}} \rangle$$

Map noise scales like,

$$\sigma_{\text{pix}} \propto t_{\text{pix}}^{-1/2} \propto \Omega_{\text{surv}}^{1/2}$$

So the error on the power spectrum scales like

$$\sigma_{\text{auto}} \propto \sigma_{\text{pix}}^{2} / N_{\text{modes}}^{1/2} \propto \Omega_{\text{surv}}^{1/2}$$



Why go wide?

Autocorrelation Errors - Knox Criterion

For a measurement of the auto spectrum,

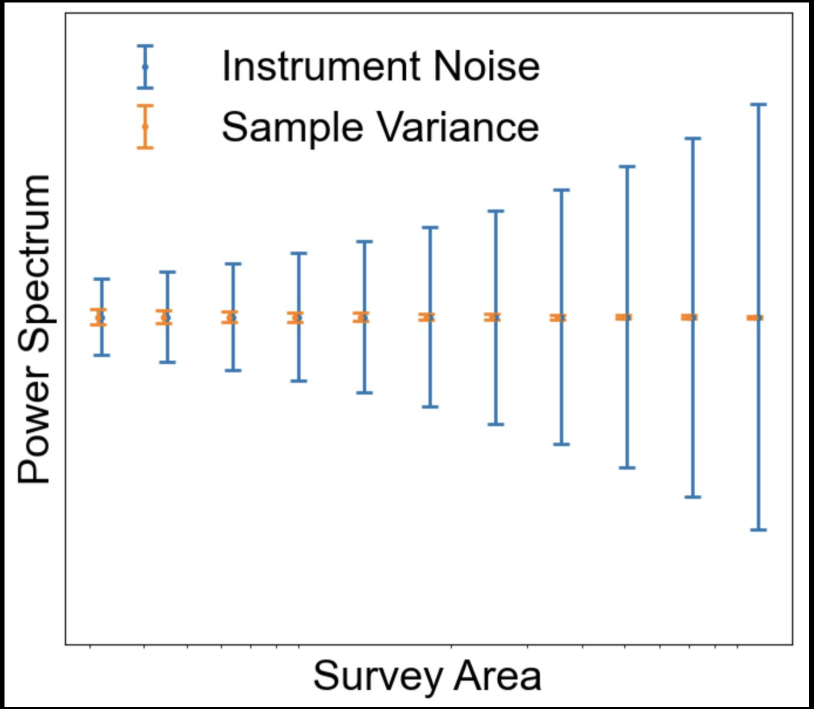
$$P_{\mathrm{auto}}(k) = \langle T_{\mathrm{CO}} T^*_{\mathrm{CO}} \rangle$$

Map noise scales like,

$$\sigma_{\mathrm{pix}} \propto t_{\mathrm{pix}}^{-1/2} \propto \Omega_{\mathrm{surv}}^{1/2}$$

So the error on the power spectrum scales like

$$\sigma_{\mathrm{auto}} \propto \sigma_{\mathrm{pix}}^2 / N_{\mathrm{modes}}^{1/2} \propto \Omega_{\mathrm{surv}}^{1/2}$$



Why go wide?

Cross-Correlation Errors

For a measurement of the cross-spectrum,

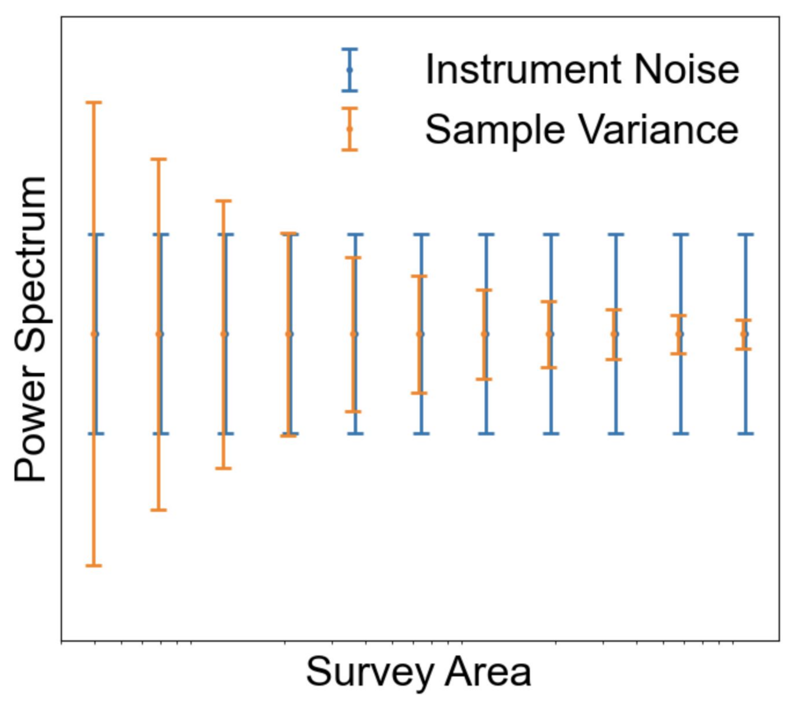
$$P_{\mathrm{cross}}(k) = \langle T_{\mathrm{CO}} X^* \rangle$$

CO map noise still scales like

$$\sigma_{\mathrm{pix}} \propto t_{\mathrm{pix}}^{-1/2} \propto \Omega_{\mathrm{surv}}^{1/2}$$

But, *if the cross-correlation data set is fixed*, cross-spectrum error now scales like

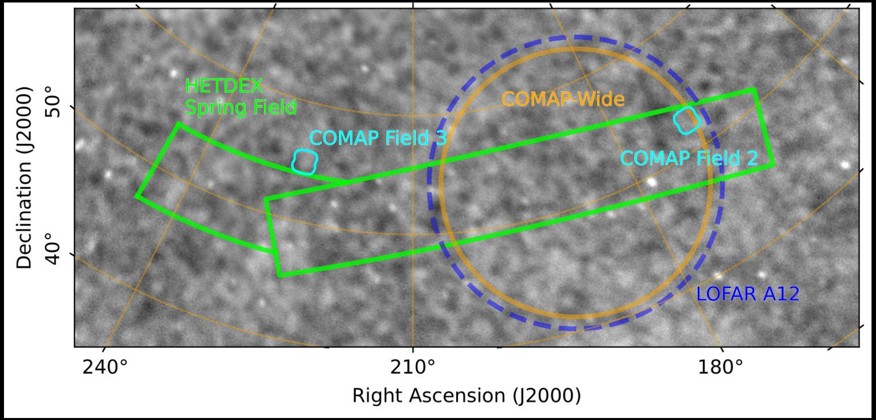
$$\sigma_{\mathrm{cross}} \propto \sigma_{\mathrm{pix}} / N_{\mathrm{modes}}^{1/2} \propto \Omega_{\mathrm{surv}}^{0}$$

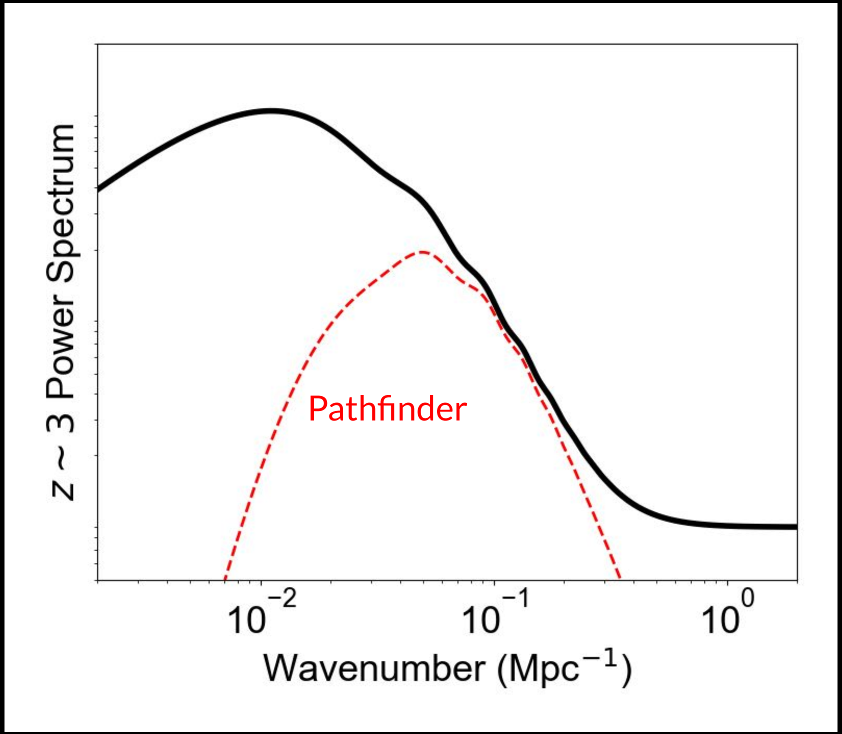


COMAP-*wide* at z~3

Wider area means larger-scale, more linear modes are accessible, at no loss of sensitivity*
(*in the white noise limit)

Systematics will be different, e.g. turnaround and start-of-scan effects

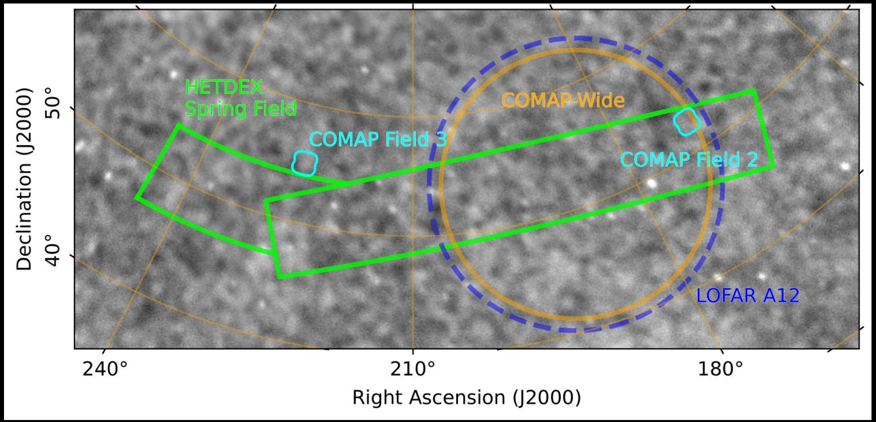


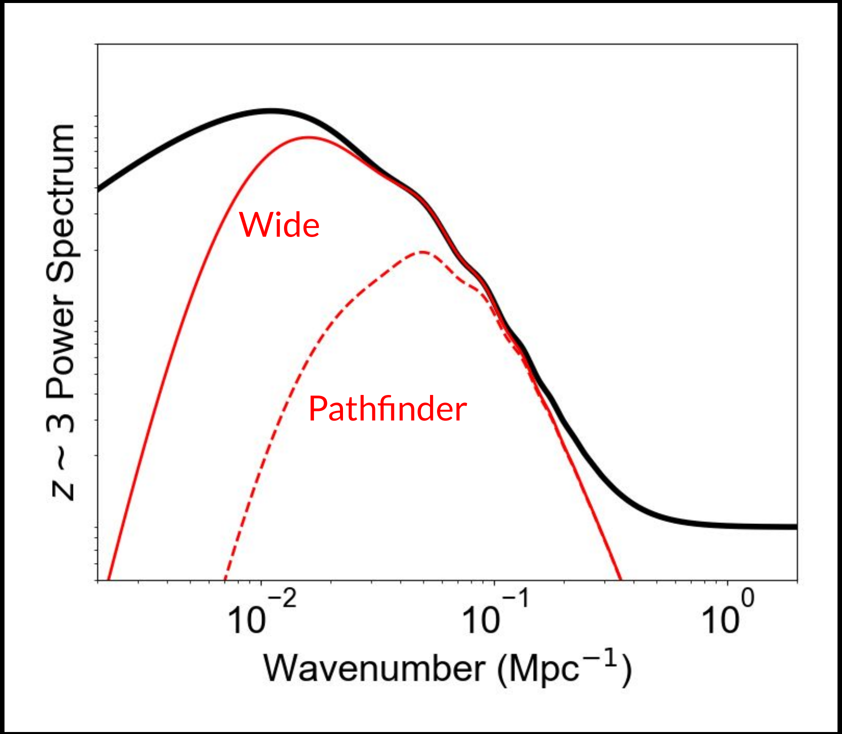


COMAP-*wide* at z~3

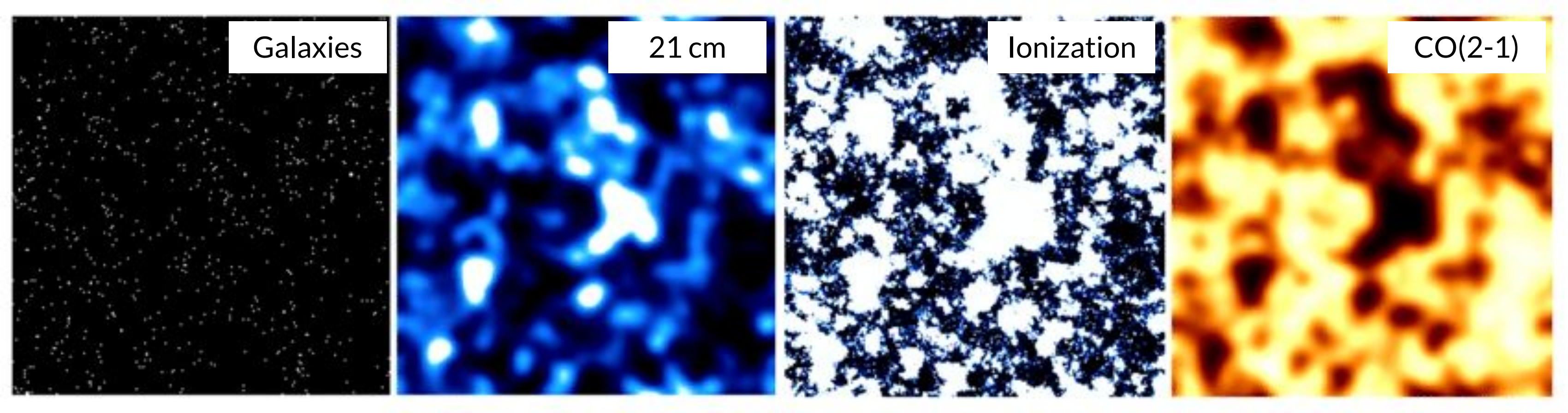
Wider area means larger-scale, more linear modes are accessible, at no loss of sensitivity*
(*in the white noise limit)

Systematics will be different, e.g. turnaround and start-of-scan effects





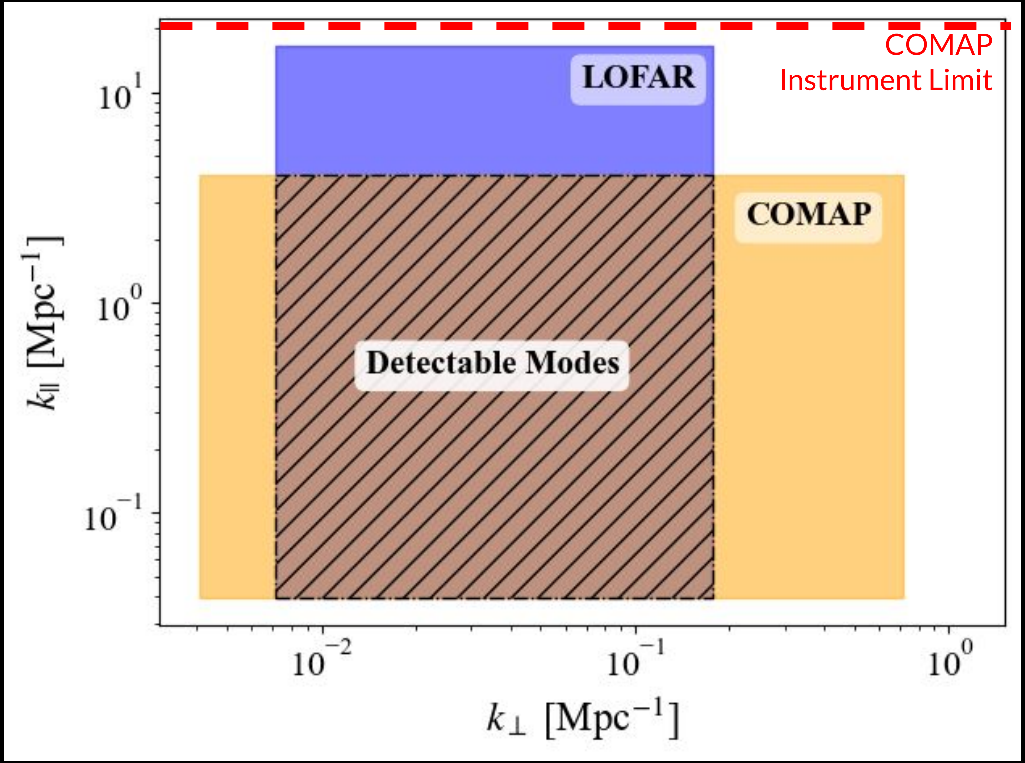
COMAP-*wide* at EoR



Cross-correlating CO(2-1) and 21 cm during reionization lets us search for the characteristic *anticorrelation* between the two lines

(Credit: Adam Lidz)

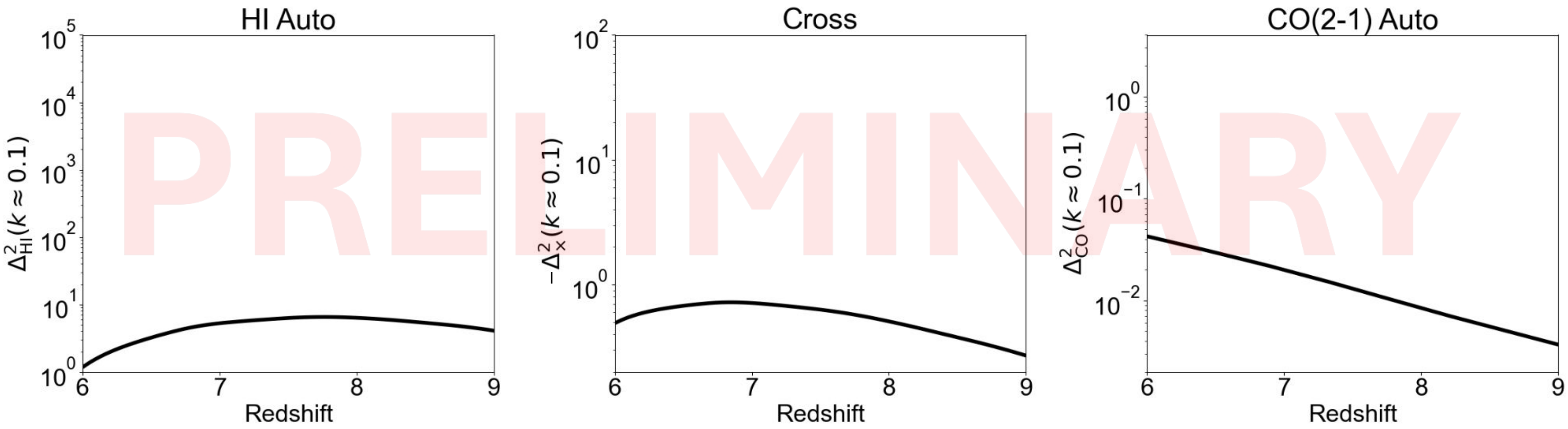
COMAP LOFAR Mode Overlap



OVRO is in the northern hemisphere, so LOFAR provides the best HI accessibility

COMAP-*wide*, LOFAR are well-matched in terms of mode coverage

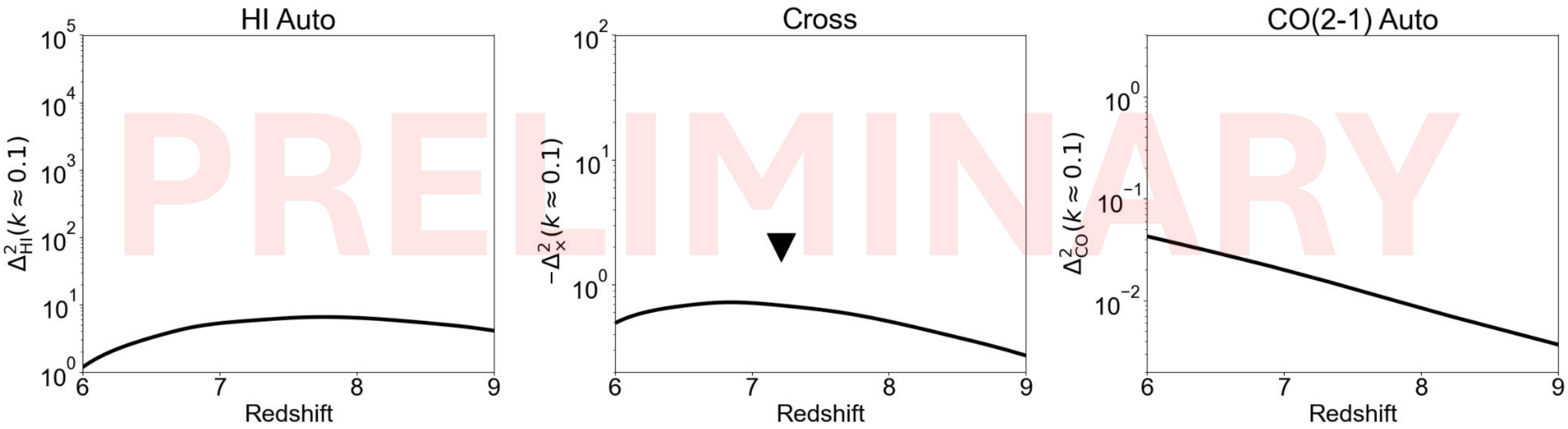
CO-HI Forecasts



21cmFAST (Mesinger+ 2016)

Li+2016 / Keating+ 2020
Scaling Model

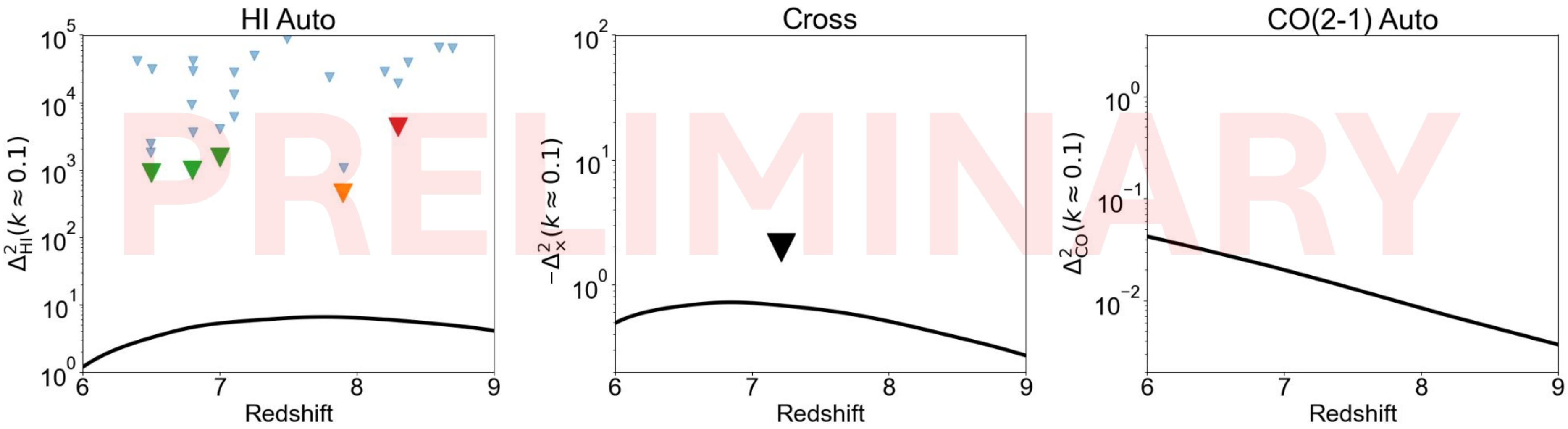
CO-HI Forecasts



Cross-Correlation Upper Limit
from 2 years of COMAP-*wide*

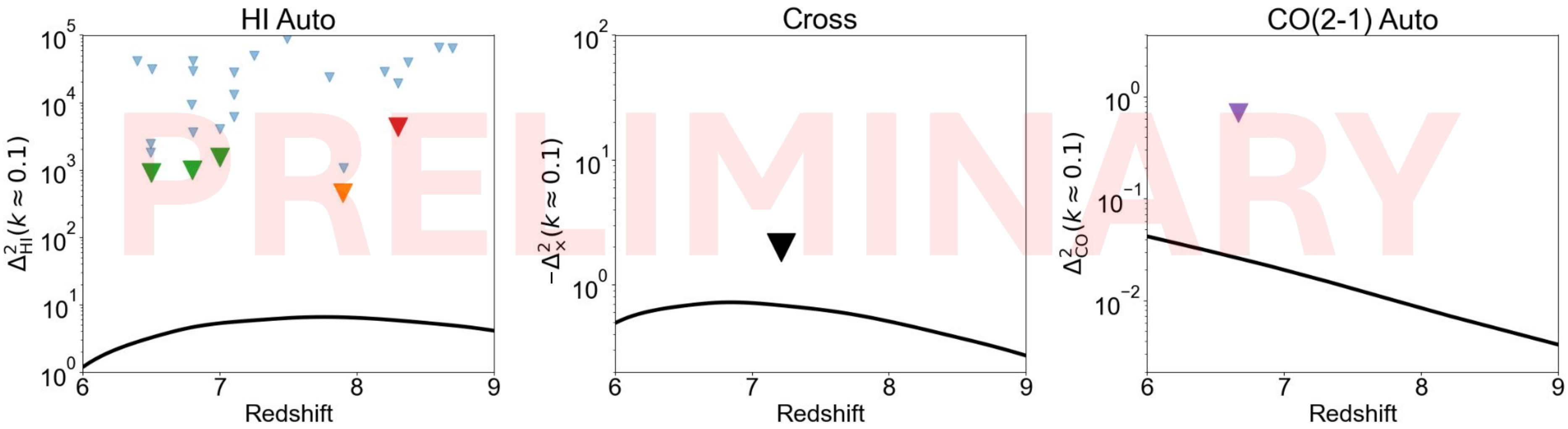
Estimated using LIMstat,
Fronenberg et al. 2025

CO-HI Forecasts



Existing HI limits, including
MWA, LOFAR, and HERA

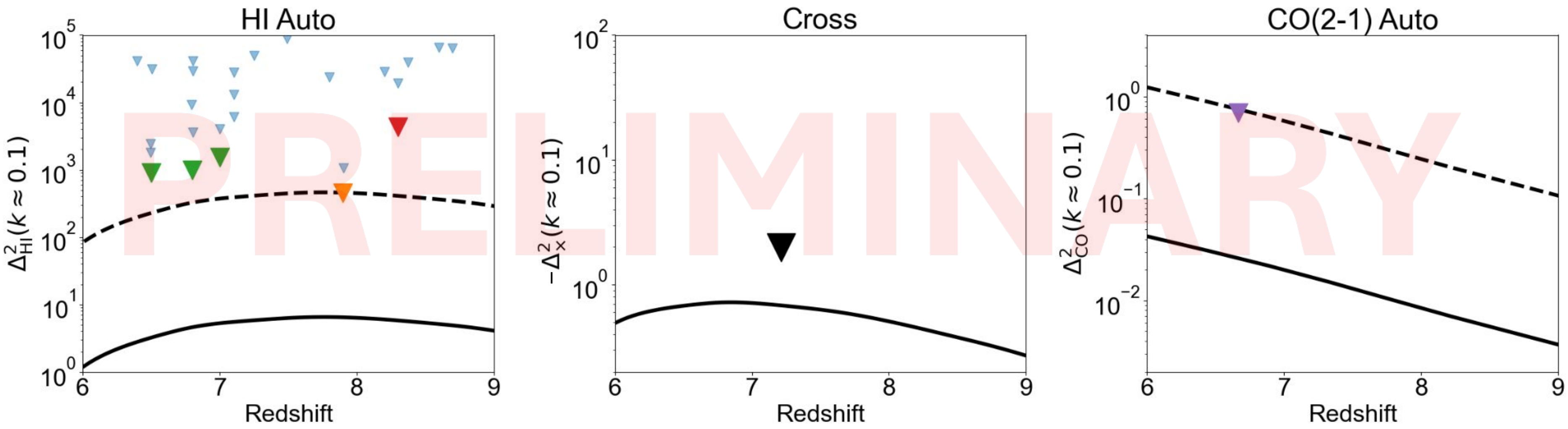
CO-HI Forecasts



Existing HI limits, including MWA, LOFAR, and HERA

CO(2-1) upper limit from COMAP Season 2

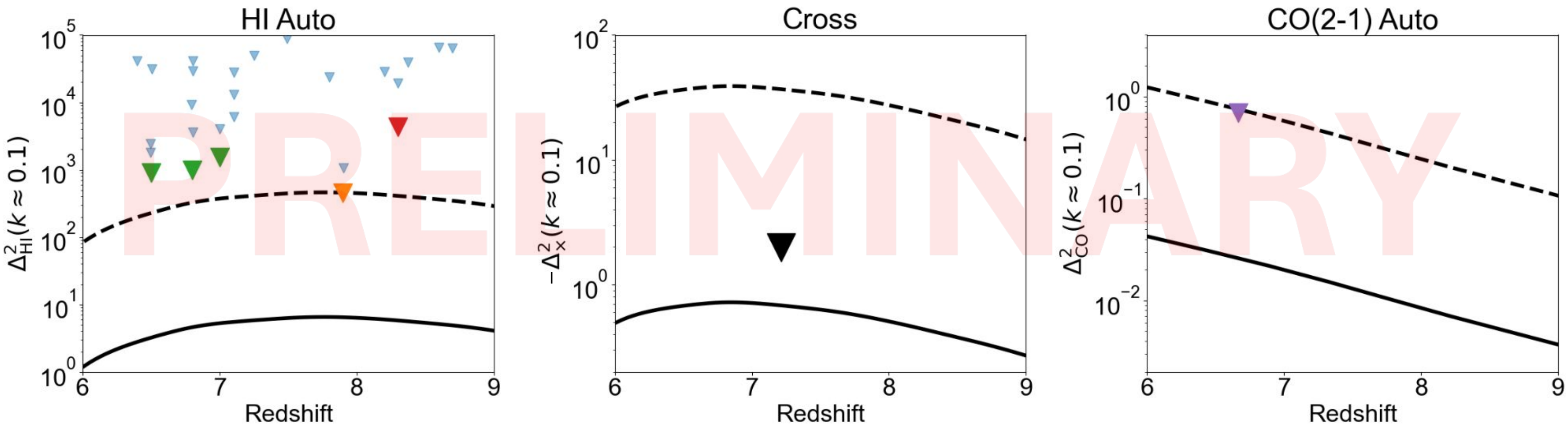
CO-HI Forecasts



Existing HI limits, including
MWA, LOFAR, and HERA

CO(2-1) upper limit from
COMAP Season 2

CO-HI Forecasts

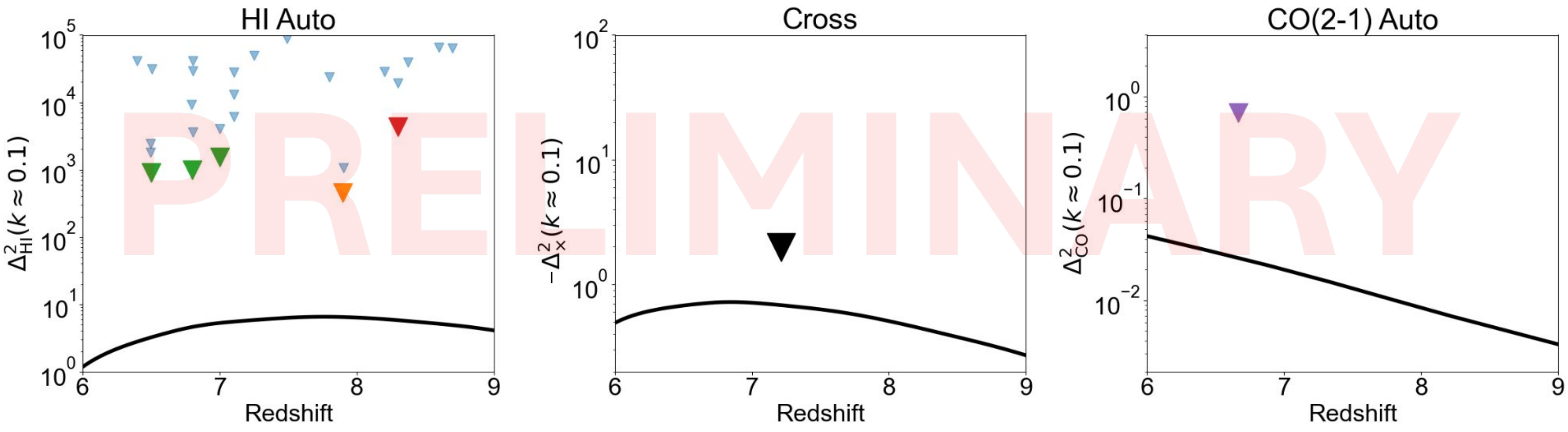


Existing HI limits, including MWA, LOFAR, and HERA

Brightest model allowed by current data

CO(2-1) upper limit from COMAP Season 2

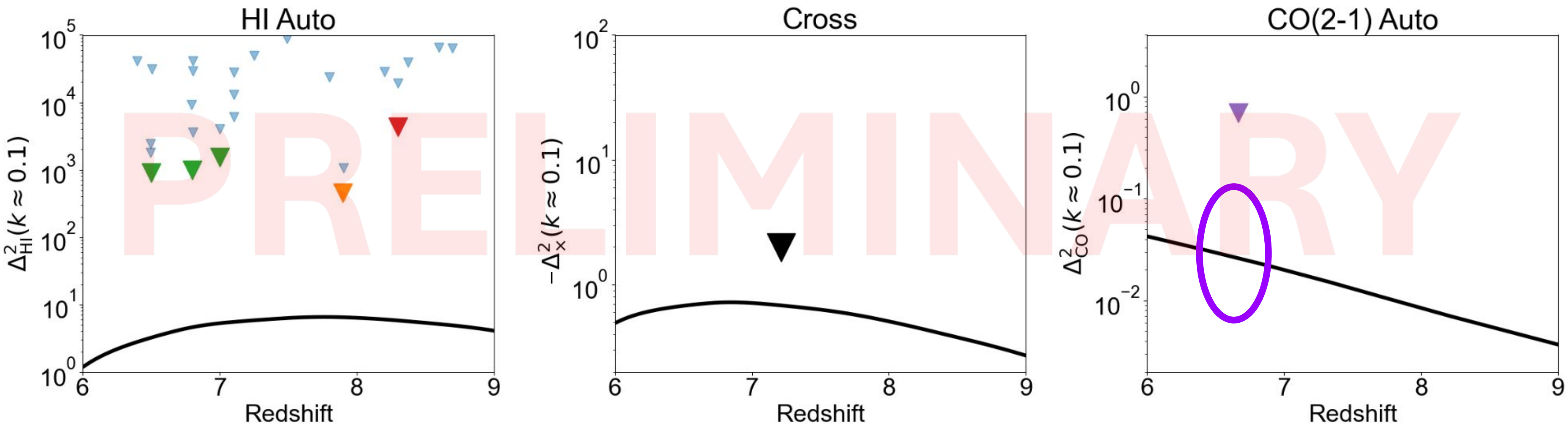
CO-HI Forecasts



Existing HI limits, including MWA, LOFAR, and HERA

CO(2-1) upper limit from COMAP Season 2

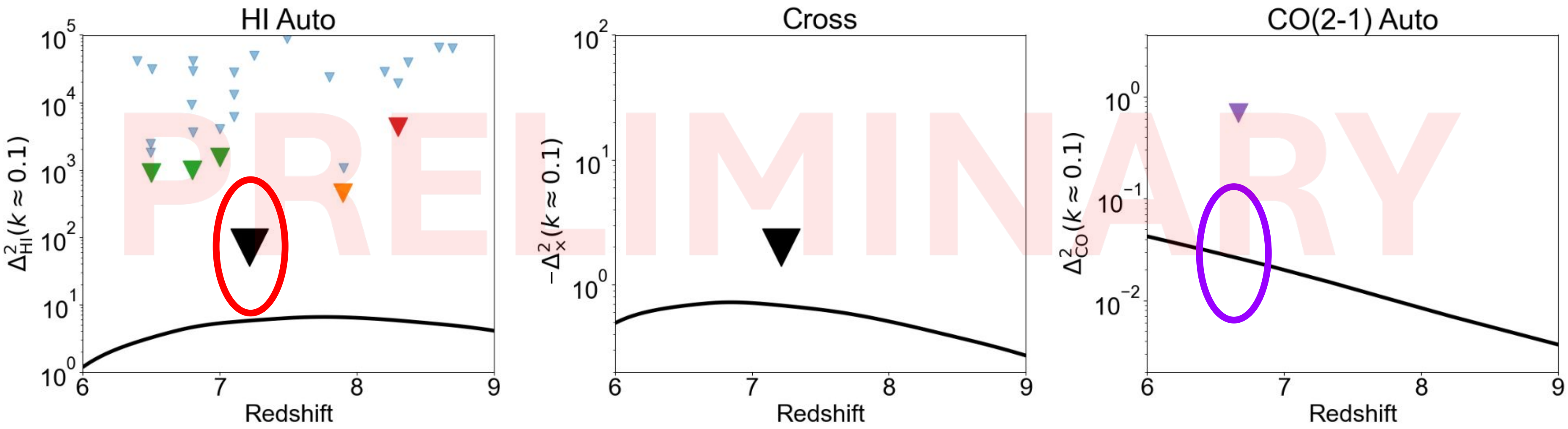
CO-HI Forecasts



If we have a measurement of the CO(2-1) signal, we can infer a measurement of the HI auto power

(Such a measurement could be obtained from COMAP-EoR deep fields)

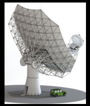
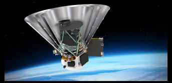
CO-HI Forecasts



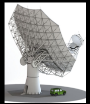
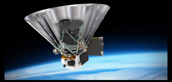
If we have a measurement of the CO(2-1) signal, we can infer a measurement of the HI auto power

(Such a measurement could be obtained from COMAP-EoR deep fields)

LIM Lines at Reionization

	21 cm	CO(1-0)	CO(2-1)	CII	Lyα
21 cm	 LOFAR/HERA/ MWA/etc.				
CO(1-0)		 COMAP EoR			
CO(2-1)			 COMAP Pathfinder?		
CII				TIME/FYST/ CONCERTO	
Lyα					 SPHEREx

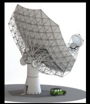
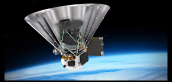
LIM Lines at Reionization

	21 cm	CO(1-0)	CO(2-1)	CII	Lyα
21 cm	 LOFAR/HERA/ MWA/etc.				
CO(1-0)		 COMAP EoR			
CO(2-1)			 COMAP Pathfinder?		
CII				TIME/FYST/ CONCERTO	
Lyα					 SPHEREx

Reionization is hard!

(both theoretically and experimentally)

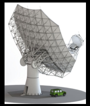
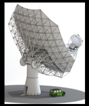
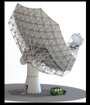
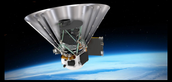
LIM Lines at Reionization

	21 cm	CO(1-0)	CO(2-1)	CII	Lyα
21 cm	 LOFAR/HERA/ MWA/etc.				
CO(1-0)		 COMAP EoR			
CO(2-1)			 COMAP Pathfinder?		
CII				TIME/FYST/ CONCERTO	
Lyα					 SPHEREx

Reionization is hard!
(both theoretically and experimentally)

Cross-correlations can help us solve it

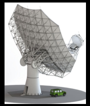
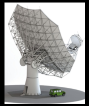
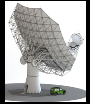
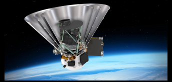
LIM Lines at Reionization

	21 cm	CO(1-0)	CO(2-1)	CII	Lyα
21 cm	 LOFAR/HERA/MWA/etc.		 COMAP-*wide*		
CO(1-0)		 COMAP EoR	 COMAP-EoR		
CO(2-1)	 COMAP-*wide*	 COMAP-EoR	 COMAP Pathfinder?		
CII				TIME/FYST/CONCERTO	
Lyα					 SPHEREx

Reionization is hard!
(both theoretically and experimentally)

Cross-correlations can help us solve it

LIM Lines at Reionization

	21 cm	CO(1-0)	CO(2-1)	CII	Lyα
21 cm	 LOFAR/HERA/ MWA/etc.		 COMAP-*wide*		???
CO(1-0)		 COMAP EoR	 COMAP-EoR		???
CO(2-1)	 COMAP-*wide*	 COMAP-EoR	 COMAP Pathfinder?		???
CII				TIME/FYST/ CONCERTO	???
Lyα	???	???	???	???	 SPHEREx

Reionization is hard!
(both theoretically and experimentally)

Cross-correlations can help us solve it

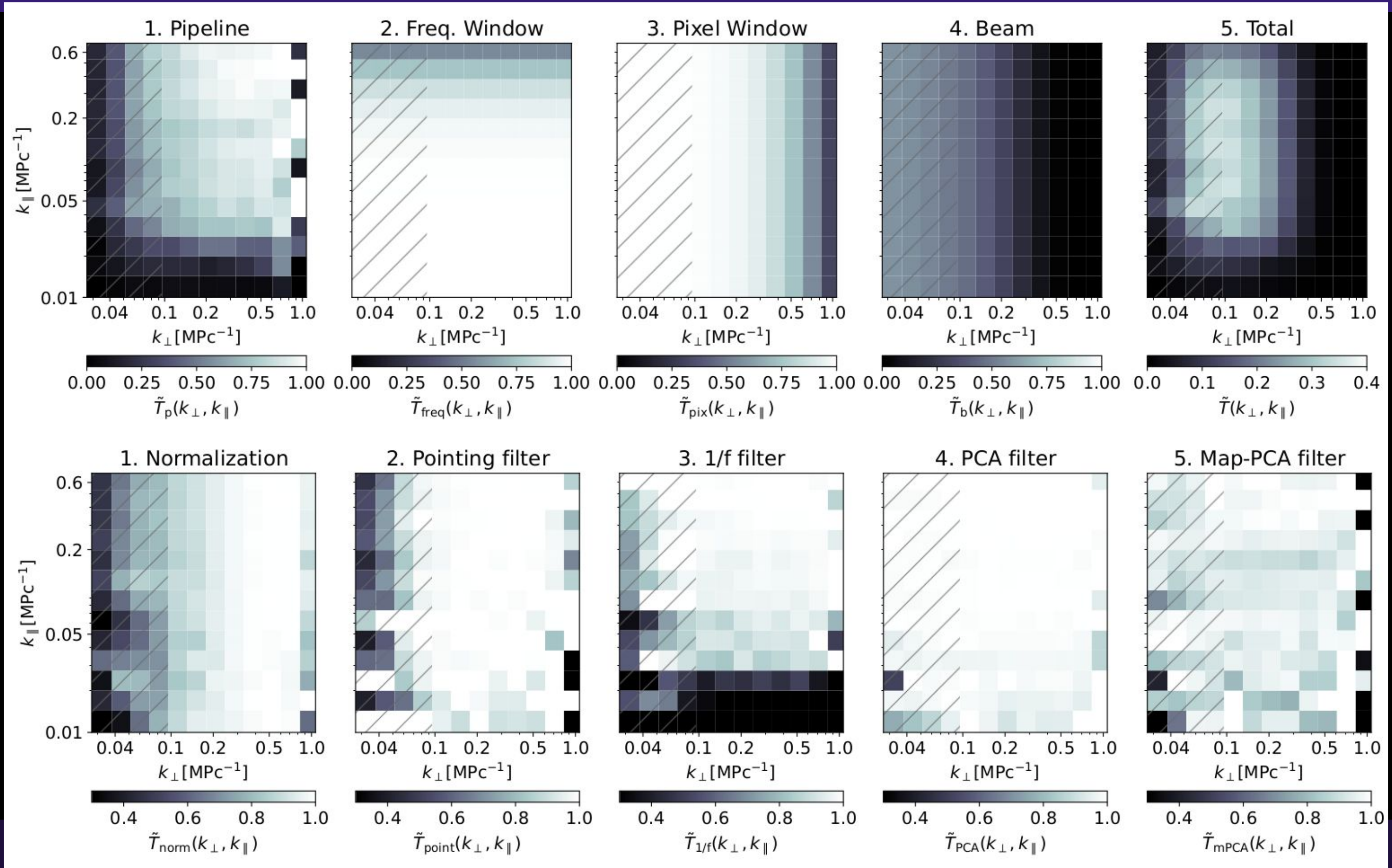
Funding

- National Science Foundation
    - Grants Nos. 1517108, 1517288, 1517598, 1518282, 1910999, 2206834
- Keck Institute for Space Studies
    - "The First Billion Years: A Technical Development Program for Spectral Line Observations."
- Research Council of Norway
    - Research Council of Norway through grants 251328 and 274990
- European Research Council:
    - Commander: EU ERC-AdG agreement No. 101141621 – PI H. K. Eriksen - 2024-2029
    - Cosmoglobe: EU ERC-CoG Grant agreement No. 819478 - PI I. K. Wehus - 2019-2024
    - Bits2cosmology: EU ERC-CoG Grant agreement No. 772253 - PI H. K. Eriksen - 2018-2023

This work has received EU funding from the European Research Council (ERC) and Research and Innovation Staff Exchange (RISE) under the Horizon 2020 Research and Innovation Programme and the Research Council of Norway. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Backup slides

Transfer functions



Map voxel significance