Data-Driven Foreground Removal for Line Intensity Mapping

Hannah Fronenberg

Phd Candidate @ McGill → Limbo (currently) → Postdoc @ KICP UChicago

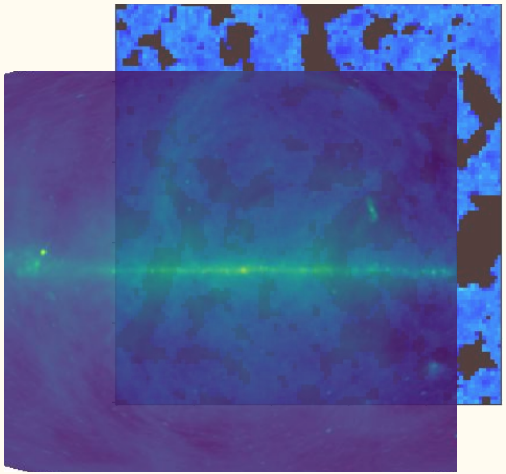
LIM25
June 5, 2025

email: hannah.fronenberg@mail.mcgill.ca
website: www.hannahfro.com

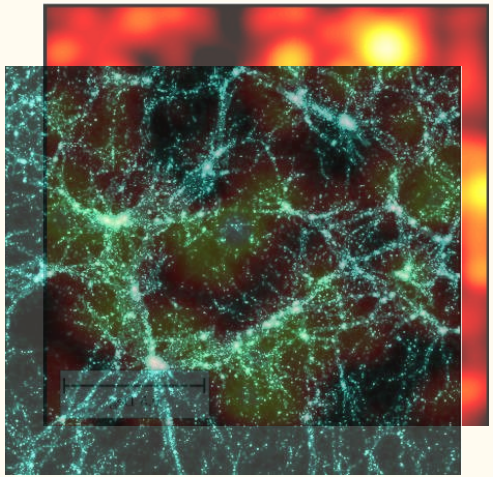
McGill | TSI Trottier Space Institute at McGill | Institut spatial Trottier de McGill

Fièrement propulsée par les FRQ
Québec

Continuum Foreground Contamination in LIM



Galactic Synchrotron



CIB

Zodiacal Light

Existing Foreground Removal Strategies

Foreground Separation

Separate out the foregrounds from the signal.

E.g. Principal component analysis (PCA), singular value decomposition (SVD).

Foreground Avoidance

Throw out sections of data that are heavily contaminated, signal included!

E.g. Wedge cut in 21 cm cosmology, high-pass filtering.

Existing Foreground Removal Strategies

Foreground Subtraction

E.g. Principal component analysis (PCA), singular value decomposition (SVD) Carucci+2024,

Foreground Avoidance

included!

E.g. Wedge cut in 21 cm cosmology

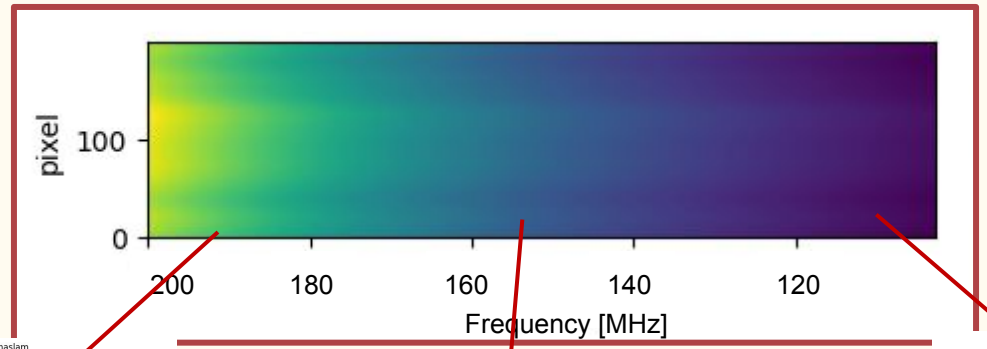
Many of these techniques are not as effective in the presence of complex observational effects.

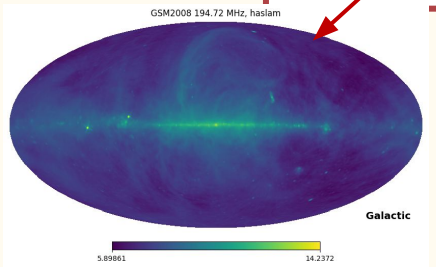
“The foregrounds are correlated across the band and the signal is not.”

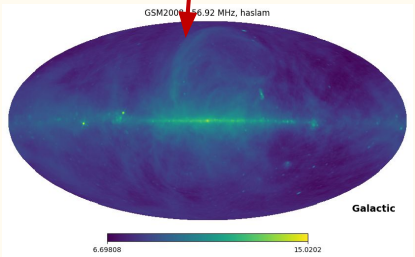
—Haochen Wang (MIT) @ LIM24

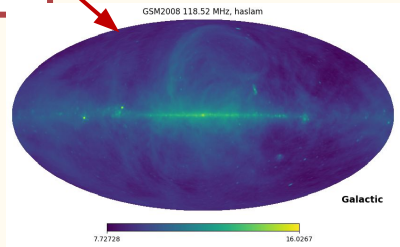
Signal vs. Foreground Correlations

Do these slices look similar? Is synchrotron emission correlated across the band?



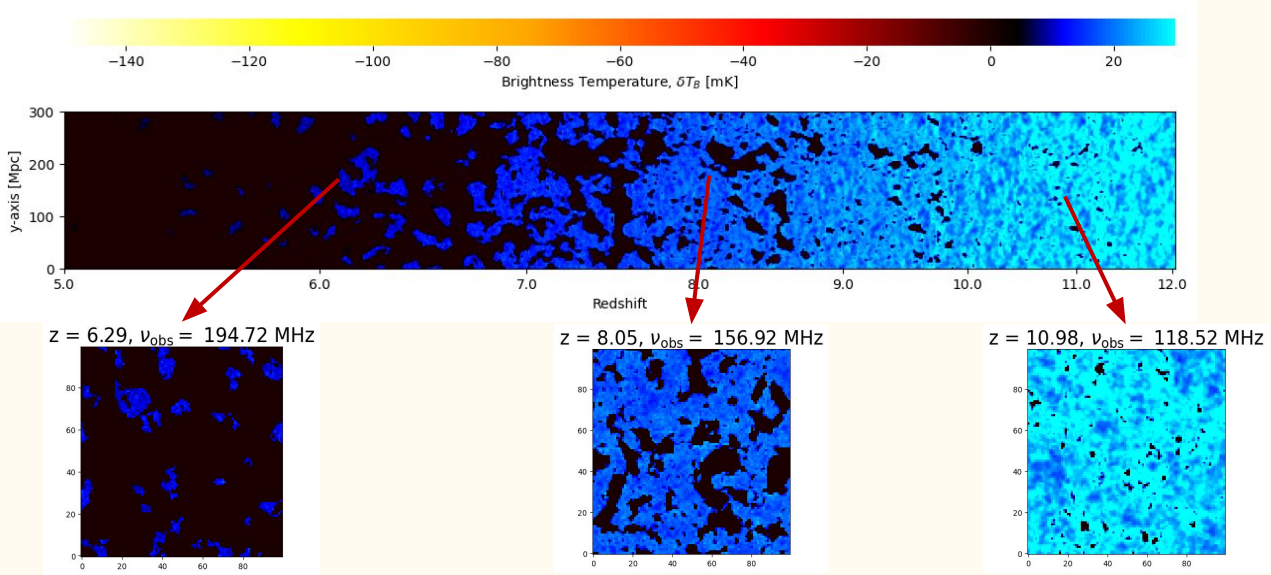




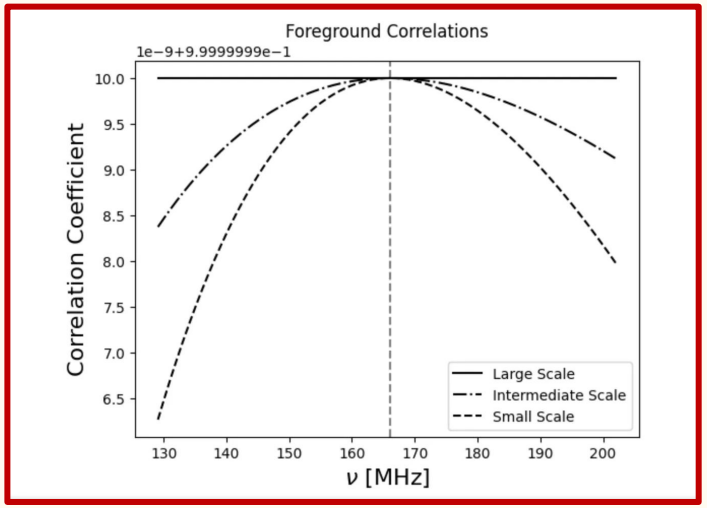


Signal vs. Foreground Correlations

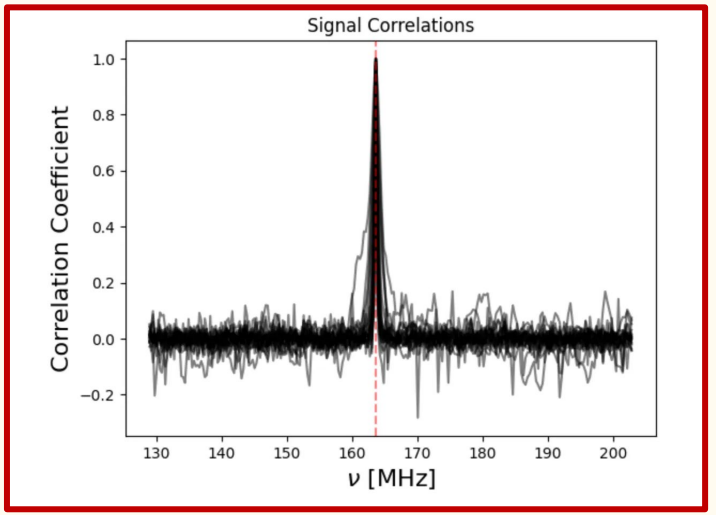
Do these slices look similar? Is the signal correlated across the band?



Broadband Foreground vs. Signal Correlations

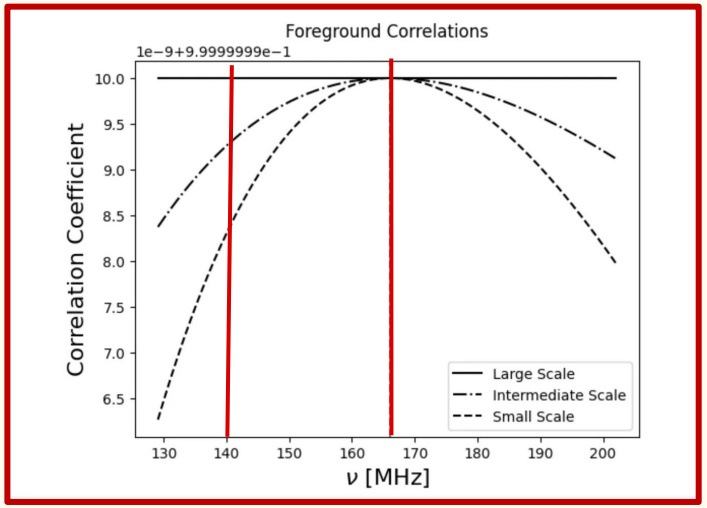


Foregrounds are correlated in widely separated frequency channels

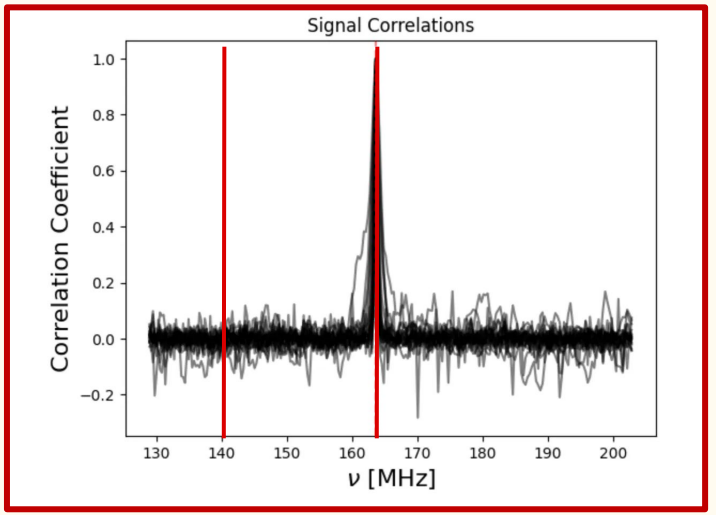


Line emission is only correlated in nearby channels

Broadband Foreground vs. Signal Correlations



Foregrounds are correlated in widely separated frequency channels



Line emission is only correlated in nearby channels

Cross-correlations yield foreground-only power

For simplicity, let's consider a toy model where we are computing the power spectrum of just 2 slices of data from different frequency channels.

$T_{\nu_1} = s_{\nu_1} + f_{\nu_1} \qquad T_{\nu_2} = s_{\nu_2} + f_{\nu_2}$ Observed sky in position space

$$P_{12}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_2(k)\rangle = \langle \tilde{s}_1\tilde{s}_2\rangle + \langle \tilde{s}_1\tilde{f}_2\rangle + \langle \tilde{f}_1\tilde{s}_2\rangle + \langle \tilde{f}_1\tilde{f}_2\rangle$$

Cross-correlations yield foreground-only power

For simplicity, let's consider a toy model where we are computing the power spectrum of just 2 slices of data from different frequency channels.

$$T_{\nu_1} = s_{\nu_1} + f_{\nu_1} \qquad T_{\nu_2} = s_{\nu_2} + f_{\nu_2}$$ Observed sky in position space

$$P_{12}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_2(k) \rangle = \cancel{\langle \tilde{s}_1 \tilde{s}_2 \rangle} + \cancel{\langle \tilde{s}_1 \tilde{f}_2 \rangle} + \cancel{\langle \tilde{f}_1 \tilde{s}_2 \rangle} + \langle \tilde{f}_1 \tilde{f}_2 \rangle$$

Cross-correlations yield foreground-only power

For simplicity, let's consider a toy model where we are computing the power spectrum of just 2 slices of data from different frequency channels.

$$T_{\nu_1} = s_{\nu_1} + f_{\nu_1} \qquad T_{\nu_2} = s_{\nu_2} + f_{\nu_2}$$ Observed sky in position space

$$P_{12}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_2(k) \rangle = \cancel{\langle \tilde{s}_1 \tilde{s}_2 \rangle} + \cancel{\langle \tilde{s}_1 \tilde{f}_2 \rangle} + \cancel{\langle \tilde{f}_1 \tilde{s}_2 \rangle} + \langle \tilde{f}_1 \tilde{f}_2 \rangle$$

$$= \langle \tilde{f}_1 \tilde{f}_2 \rangle$$

Cross-correlations yield foreground-only power

For simplicity, let's consider a toy model where we are computing the power spectrum of just 2 slices of data from different frequency channels.

$$T_{\nu_1} = s_{\nu_1} + f_{\nu_1} \qquad T_{\nu_2} = s_{\nu_2} + f_{\nu_2}$$

Observed sky in position space

$$P_{12}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_2(k) \rangle = \cancel{\langle \tilde{s}_1 \tilde{s}_2 \rangle} + \cancel{\langle \tilde{s}_1 \tilde{f}_2 \rangle} + \cancel{\langle \tilde{f}_1 \tilde{s}_2 \rangle} + \langle \tilde{f}_1 \tilde{f}_2 \rangle$$

$$= \langle \tilde{f}_1 \tilde{f}_2 \rangle$$

The cross-spectrum of widely separated frequency channels yields foreground-only information

If only we had some estimator that could get you an auto-spectrum from cross-spectra...

Auto-Spectrum from Cross-Spectra

Sarkar+ in prep, McBride & Liu 2023, Beane+2019

B19 Estimator

$$P_{ii}(k) = \frac{P_{ij}(k) P_{ik}(k)}{P_{jk}(k)}$$

B19 In Action

$$P_{ii}(k) = \frac{P_{ij}(k)P_{ik}(k)}{P_{jk}(k)}$$

Recall

$$P_{ij}(k) = \langle \tilde{f}_i \tilde{f}_j \rangle$$

$$P_{ii}(k) = P_{ii}^{\mathrm{FG}}(k)$$

B19 In Action

$$P_{ii}(k) = \frac{P_{ij}(k)P_{ik}(k)}{P_{jk}(k)}$$

Recall

$$P_{ij}(k) = \langle \tilde{f}_i \tilde{f}_j \rangle$$

$$P_{ii}(k) = P_{ii}^{\mathrm{FG}}(k)$$

$$P_{ii}^{\mathrm{obs}}(k) = P_{ii}^{s}(k) + P_{ii}^{FG}(k)$$

B19 In Action

$$P_{ii}(k) = \frac{P_{ij}(k)P_{ik}(k)}{P_{jk}(k)}$$

Recall

$$P_{ij}(k) = \langle \tilde{f}_i \tilde{f}_j \rangle$$

$$P_{ii}(k) = P_{ii}^{\mathrm{FG}}(k)$$

$$P_{ii}^{\mathrm{obs}}(k) = P_{ii}^{s}(k) + P_{ii}^{FG}(k) - P_{ii}^{FG,B19}(k)$$

B19 In Action

$$P_{ii}(k) = \frac{P_{ij}(k)P_{ik}(k)}{P_{jk}(k)} \xrightarrow{\text{Recall}} P_{ij}(k) = \langle \tilde{f}_i \tilde{f}_j \rangle$$

$$P_{ii}(k) = P_{ii}^{\mathrm{FG}}(k)$$

$$P_{ii}^{\mathrm{obs}}(k) = P_{ii}^{s}(k) + P_{ii}^{FG}(k) - P_{ii}^{FG,B19}(k) = P_{ii}^{s}(k)$$

B19 In Action

$$P_{ii}(k) = \frac{P_{ij}(k)P_{ik}(k)}{P_{jk}(k)}$$

Recall

$$P_{ij}(k) = \langle \tilde{f}_i \tilde{f}_j \rangle$$

$$P_{ii}(k) = P_{ii}^{\mathrm{FG}}(k)$$

$$P_{ii}^{\mathrm{obs}}(k) = P_{ii}^{s}(k) + P_{ii}^{FG}(k) - P_{ii}^{FG,B19}(k) = P_{ii}^{s}(k)$$

This estimator gives us, from the data itself, the foreground power spectrum at frequency “i” which we can then subtract off from the total measured power in order to reveal the underlying signal

The Benefits:

1. Relies on very few assumptions about the nature of the foregrounds
2. Recovers the foreground power spectrum *as seen by the instrument* (beam & systematics are naturally folded in!)
3. Recovers *all* foreground power spectrum cross-terms (e.g. signal-foreground and noise-foreground residuals)

The Benefits:

1. Relies on very few assumptions about the nature of the foregrounds
2. Recovers the foreground power spectrum *as seen by the instrument* (beam & systematics are naturally folded in!)
3. Recovers *all* foreground power spectrum cross-terms (e.g. signal-foreground and noise-foreground residuals)

Relies on very few assumptions about the nature of the foregrounds.

The foregrounds are:

1. The result of broadband emitters and are therefore correlated across the observing band. This doesn't work for interlopers!
2. Well described by

$$\tilde{T}_i(k) = \beta_i \tilde{\delta}_{\mathrm{FG}}(k)$$

Relies on very few assumptions about the nature of the foregrounds.

The foregrounds are:

1. The result of broadband emitters and are therefore correlated across the observing band. This doesn't work for interlopers!
2. Well described by

$$\tilde{T}_i(k) = \beta_i \tilde{\delta}_{\mathrm{FG}}(k)$$

Frequency dependent amplitude

Foreground mass distribution

Relies on very few assumptions about the nature of the foregrounds.

The foregrounds are:

1. The result of broadband emitters and are therefore correlated across the observing band. This doesn't work for
2. Well described by

$$\tilde{T}_i(k) = \beta_i \tilde{\delta}_{\mathrm{FG}}(k)$$

Frequency dependent amplitude

Foreground mass distribution

No assumption of spectral smoothness of the foregrounds! They could in principle be very weird!

The Benefits:

1. Relies on very few assumptions about the nature of the foregrounds
2. Recovers the foreground power spectrum *as seen by the instrument* (beam & systematics are naturally folded in!)
3. Recovers *all* foreground power spectrum cross-terms (e.g. signal-foreground and noise-foreground residuals)

Recovers the foreground power spectrum *as seen by the instrument.*

$$\tilde{T}_i^{\rm obs}(k) = \tilde{B}_i(k)\tilde{T}_i^{\rm true}(k)$$

$$P_{ii}(k) = \frac{P_{ij}(k)P_{ik}(k)}{P_{jk}(k)}$$

Recovers the foreground power spectrum *as seen by the instrument.*

$$\tilde{T}_i^{\rm obs}(k) = \tilde{B}_i(k)\tilde{T}_i^{\rm true}(k)$$

$$P_{ii}(k) = \frac{\tilde{B}_i(k)\tilde{B}_j(k)P_{ij}^{\rm true}(k)\tilde{B}_i(k)\tilde{B}_k(k)P_{ik}^{\rm true}(k)}{\tilde{B}_j(k)\tilde{B}_k(k)P_{jk}^{\rm true}(k)}$$

Recovers the foreground power spectrum *as seen by the instrument.*

$$\tilde{T}_i^{\rm obs}(k) = \tilde{B}_i(k)\tilde{T}_i^{\rm true}(k)$$

$$P_{ii}(k) = \tilde{B}_i(k)\tilde{B}_i(k)P_{ii}^{\rm true}(k)$$

Recovers the foreground power spectrum *as seen by the instrument.*

$$\tilde{T}_i^{\rm obs}(k) = \tilde{B}_i(k)\tilde{T}_i^{\rm true}(k)$$

$$P_{ii}(k) = \tilde{B}_i(k)\tilde{B}_i(k)P_{ii}^{\rm true}(k)$$

The same argument holds if you have some frequency-dependent systematic that is linear in Fourier space.

$$\tilde{T}_i^{\rm obs}(k) = \tilde{S}_i(k)\tilde{B}_i(k)\tilde{T}_i^{\rm true}(k)$$

The Benefits:

1. Relies on very few assumptions about the nature of the foregrounds
2. Recovers the foreground power spectrum *as seen by the instrument* (beam & systematics are naturally folded in!)
3. Recovers *all* foreground power spectrum cross-terms (e.g. signal-foreground and noise-foreground residuals)

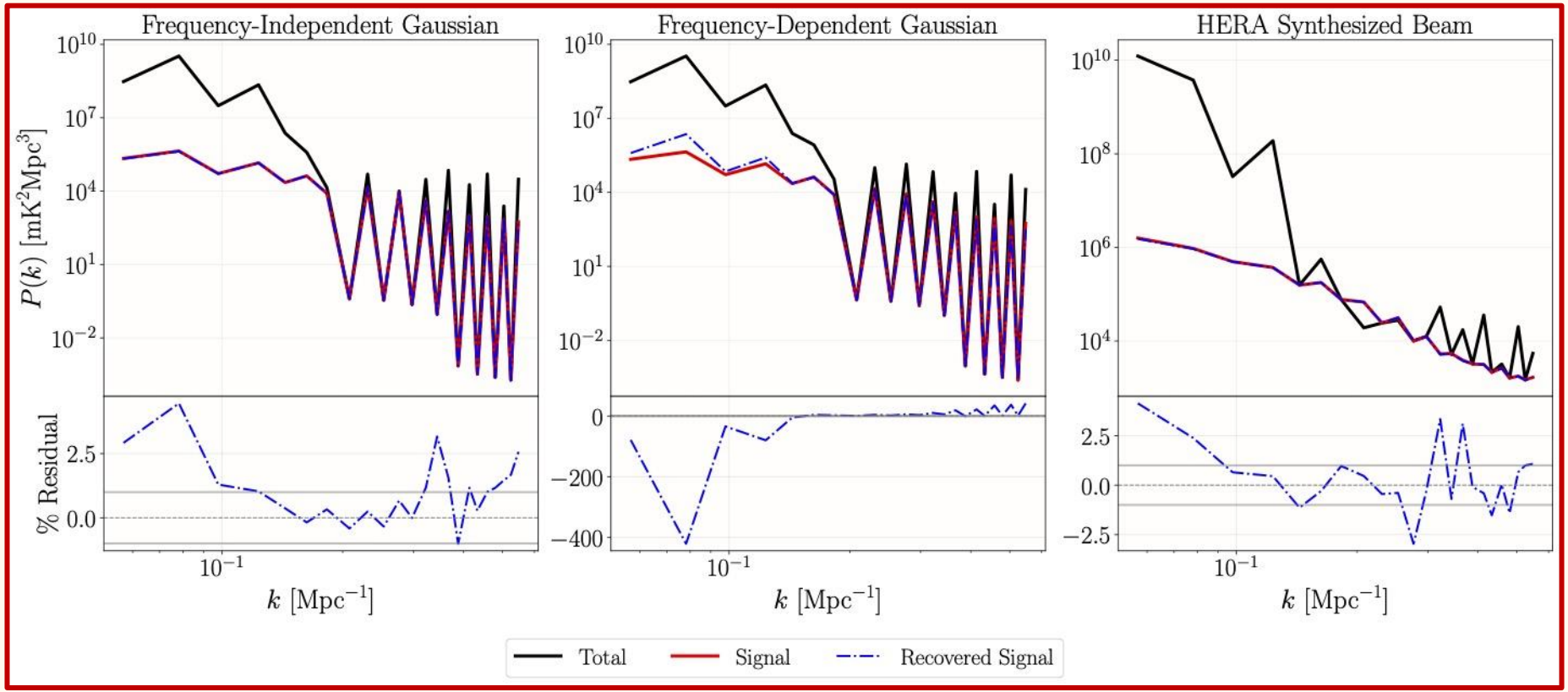
Simulating different instrument response

Frequency-independent beam (achromatic)

Frequency-dependent beam (chromatic)

HERA synthesized beam

Results – Simulating different instrument response

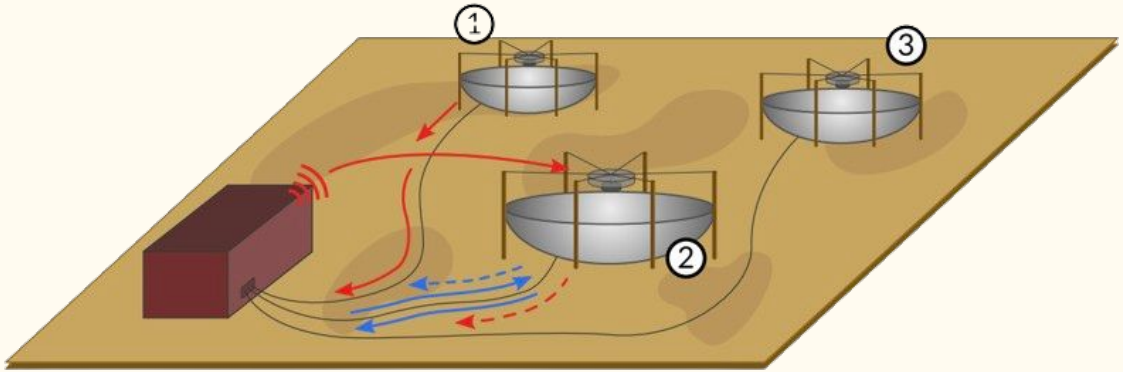


Adding Systematics: Cable Reflections

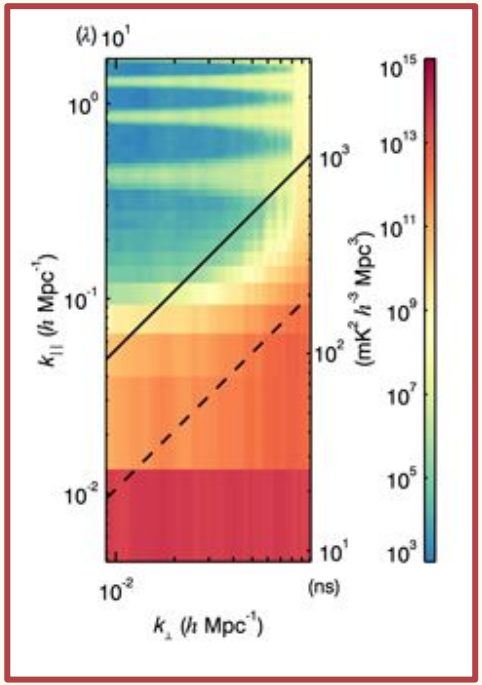
Suppression

time-delay

$$T^{\rm obs}(\hat{\mathbf{n}},\nu) = T(\hat{\mathbf{n}},\nu)\,(1+\epsilon\cos(2\pi\nu\tau))$$

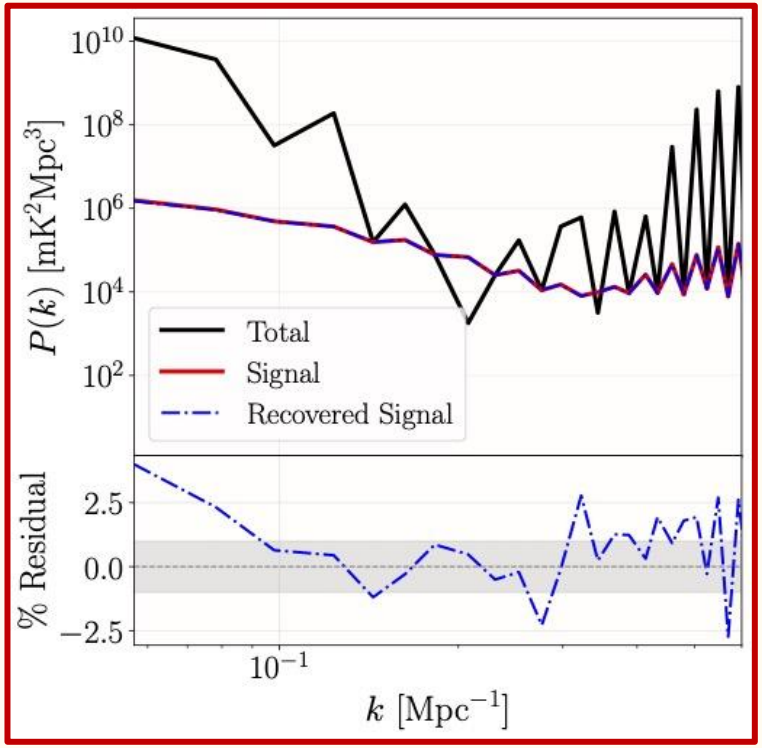


Murphy+2023



Barry+2019

Results – Simulating cable reflections + HERA synthesized beam



$\tau = 1000$ ns
Barry+2019

Is this too good to be true?
What is the breaking point of this estimator?

Relies on very few assumptions about the nature of the foregrounds.

The foregrounds are:

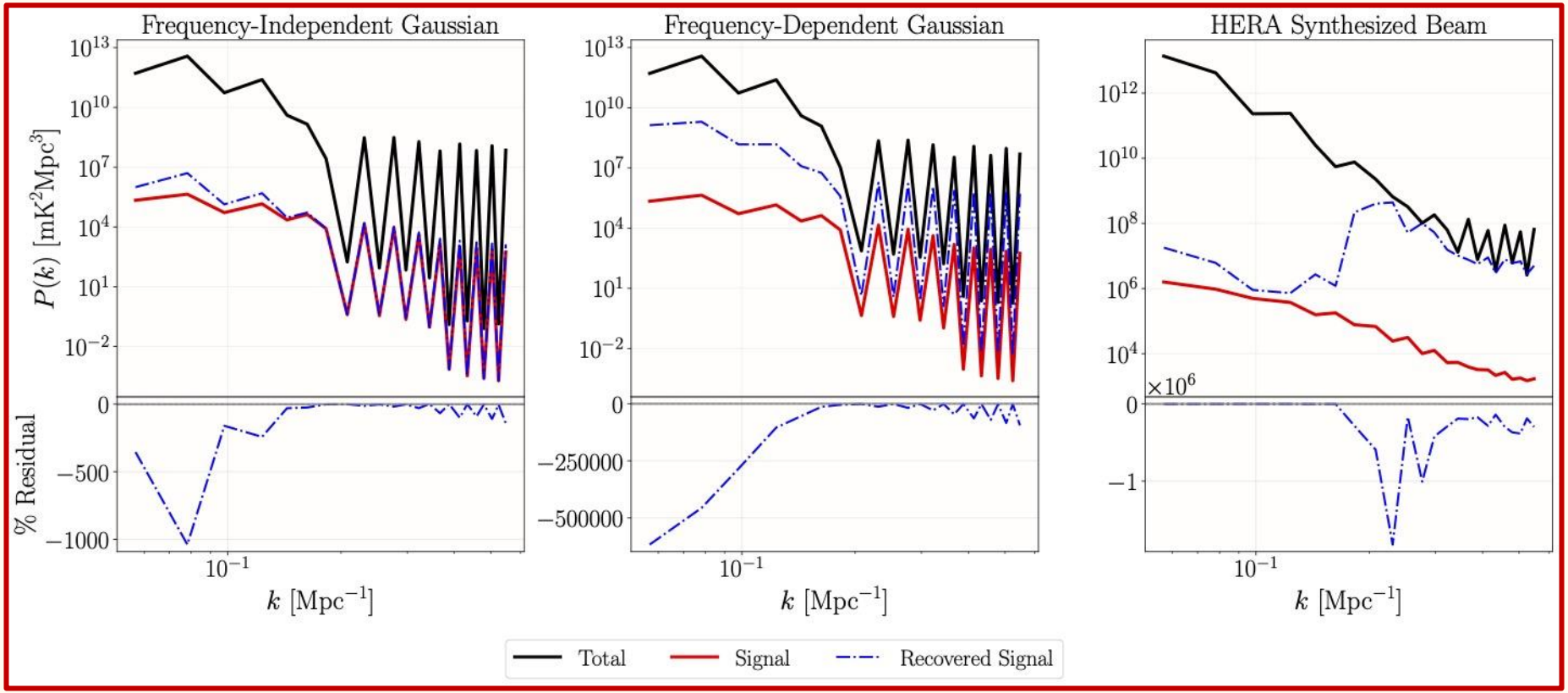
1. The result of broadband emitters and are therefore correlated across the observing band. This doesn't work for interlopers!
2. Well described by

$$T_i(k) = \beta_i \tilde{\delta}_{\rm FG}(k)$$

Frequency dependent amplitude

Foreground mass distribution

Multi-component, inseparable foreground model



Conclusion

Stay tuned later this month for HF 2025 in prep.

- Intra-dataset cross-correlations may be an effective tool (of many!) for removing the broadband contaminants of line intensity maps.
- This may be more robust to observational effects such as spatially varying and achromatic beam response, and frequency-dependent instrument systematics.
- Open questions:
  - Signal loss?
  - Can an analogous technique be implemented at the map level instead?

Q for you:
What sorts of effects do *you* need your foreground removal need to be robust to?

email: hannah.fronenberg@mail.mcgill.ca
website: www.hannahfro.com

McGill | Trottier Space Institute at McGill | Institut spatial Trottier de McGill

BONUS SLIDES

Foreground Angular Power Spectrum Estimator

$$C_l^f(\nu_i, \nu_i') = \frac{C_l^f(\nu_i, \nu_j) C_l^f(\nu_i', \nu_k)}{C_l^f(\nu_j, \nu_k)}$$

ν, ν' are indexing over a small chunk of data where you wish to estimate the cylindrically (2D) or spherically (1D) averaged power spectrum. → The target channels

ν_j, ν_k are two channels far away from ν, ν' that we are using to clean. → The cleaning channels

Recovers *all* foreground power spectrum cross-terms.

$$P_{12}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_2(k)\rangle \;\; = \langle \tilde{s}_1\tilde{s}_2\rangle + \langle \tilde{s}_1\tilde{f}_2\rangle + \langle \tilde{f}_1\tilde{s}_2\rangle + \langle \tilde{f}_1\tilde{f}_2\rangle$$

Recovers *all* foreground power spectrum cross-terms.

$$P_{12}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_2(k)\rangle = \langle \tilde{s}_1\tilde{s}_2\rangle + \langle \tilde{s}_1\tilde{f}_2\rangle + \langle \tilde{f}_1\tilde{s}_2\rangle + \langle \tilde{f}_1\tilde{f}_2\rangle$$

$$P_{12}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_2(k)\rangle = \cancel{\langle \tilde{s}_1\tilde{s}_2\rangle} + \cancel{\langle \tilde{s}_1\tilde{f}_2\rangle} + \cancel{\langle \tilde{f}_1\tilde{s}_2\rangle} + \langle \tilde{f}_1\tilde{f}_2\rangle$$

Recovers *all* foreground power spectrum cross-terms.

$$P_{12}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_2(k)\rangle = \langle \tilde{s}_1\tilde{s}_2\rangle + \langle \tilde{s}_1\tilde{f}_2\rangle + \langle \tilde{f}_1\tilde{s}_2\rangle + \langle \tilde{f}_1\tilde{f}_2\rangle$$

$$P_{12}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_2(k)\rangle = \cancel{\langle \tilde{s}_1\tilde{s}_2\rangle} + \cancel{\langle \tilde{s}_1\tilde{f}_2\rangle} + \cancel{\langle \tilde{f}_1\tilde{s}_2\rangle} + \langle \tilde{f}_1\tilde{f}_2\rangle$$

Not true! Since we always observe a finite number a Fourier modes on the sky, we get spurious correlations that stay in the data as residual contamination.

Recovers *all* foreground power spectrum cross-terms.

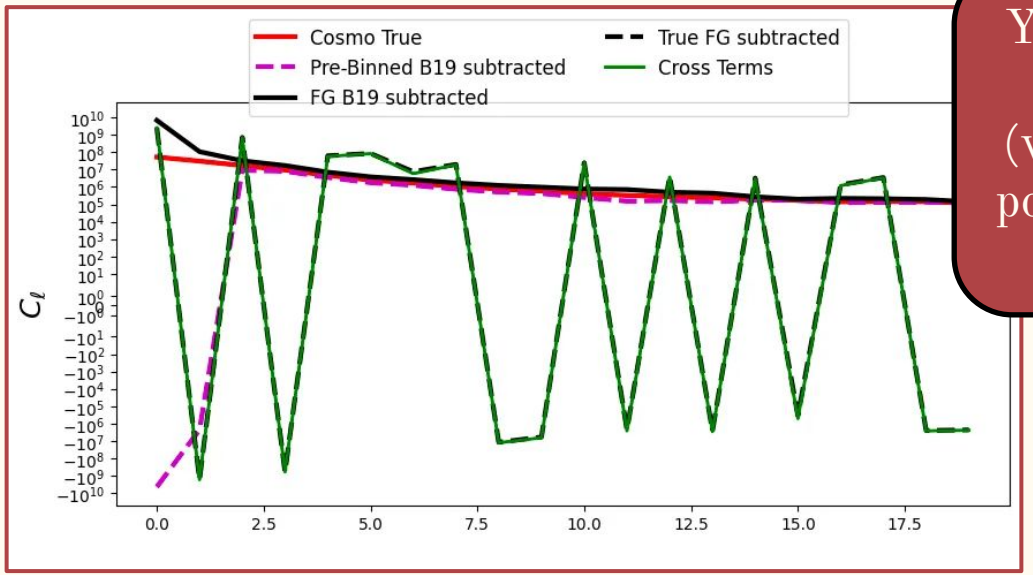
This is why spectral-based foreground removal has been largely unpopular, we can't ever model the auto-spectrum residual cross-terms

$$P_{11}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_1(k)\rangle = \langle \tilde{s}_1\tilde{s}_1\rangle + \langle \tilde{s}_1\tilde{f}_1\rangle + \langle \tilde{f}_1\tilde{s}_1\rangle + \langle \tilde{f}_1\tilde{f}_1\rangle$$

Even if you magically knew this foreground power spectrum *exactly*, you could never model the cross-terms. This would require you to know the exact realization of the signal and the foregrounds of our universe, not just their statistical properties.

Recovers *all* foreground power spectrum cross-terms.

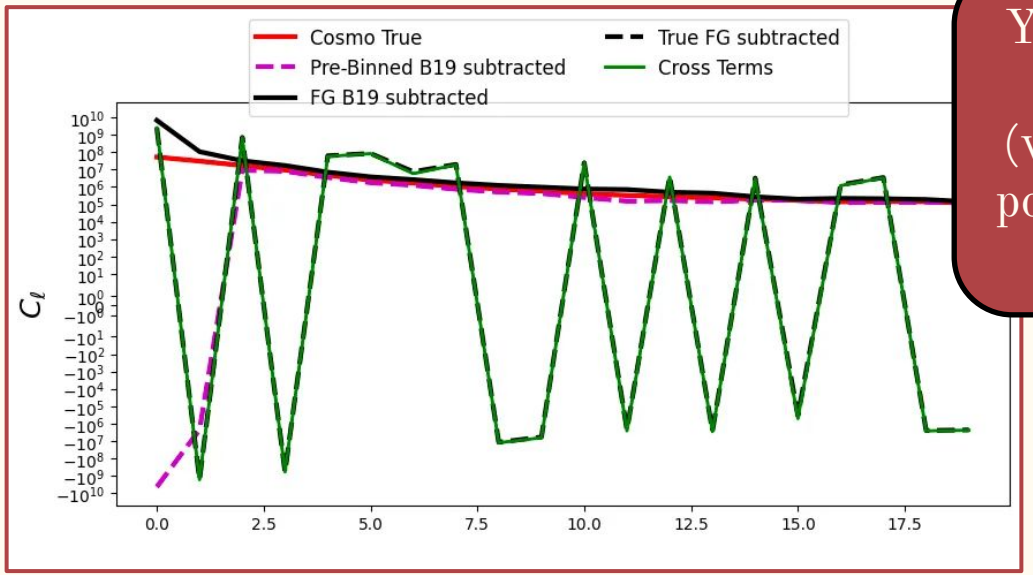
$$P_{11}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_1(k)\rangle = \langle \tilde{s}_1\tilde{s}_1\rangle + \langle \tilde{s}_1\tilde{f}_1\rangle + \langle \tilde{f}_1\tilde{s}_1\rangle + \cancel{\langle \tilde{f}_1\tilde{f}_1\rangle}$$



You would measure the cross-terms (with a small signal power contribution)! Useless!

Recovers *all* foreground power spectrum cross-terms.

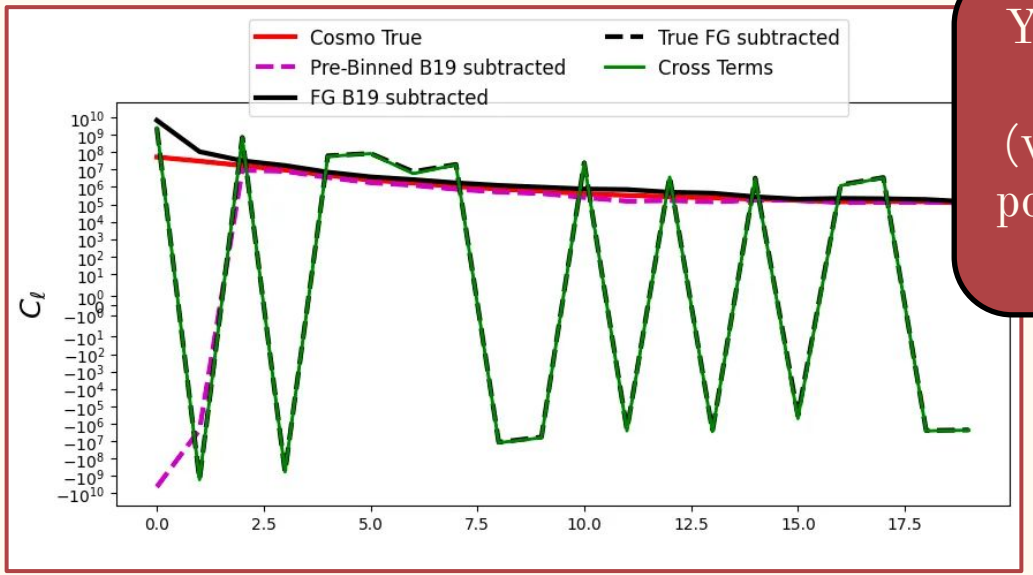
$$P_{11}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_1(k)\rangle = \langle \tilde{s}_1\tilde{s}_1\rangle + \langle \tilde{s}_1\tilde{f}_1\rangle + \langle \tilde{f}_1\tilde{s}_1\rangle + \langle \tilde{f}_1\tilde{f}_1\rangle$$



You would measure the cross-terms (with a small signal power contribution)! Useless!

Recovers *all* foreground power spectrum cross-terms.

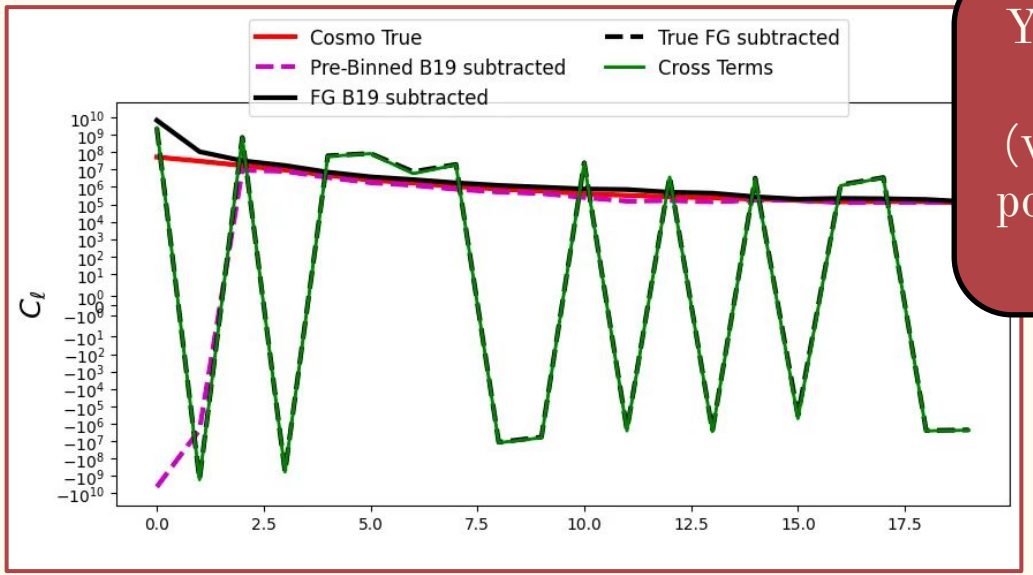
$$P_{11}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_1(k)\rangle = \langle \tilde{s}_1\tilde{s}_1\rangle + \langle \tilde{s}_1\tilde{f}_1\rangle + \langle \tilde{f}_1\tilde{s}_1\rangle + \langle \tilde{f}_1\tilde{f}_1\rangle$$



You would measure the cross-terms (with a small signal power contribution)! Useless!

Recovers *all* foreground power spectrum cross-terms.

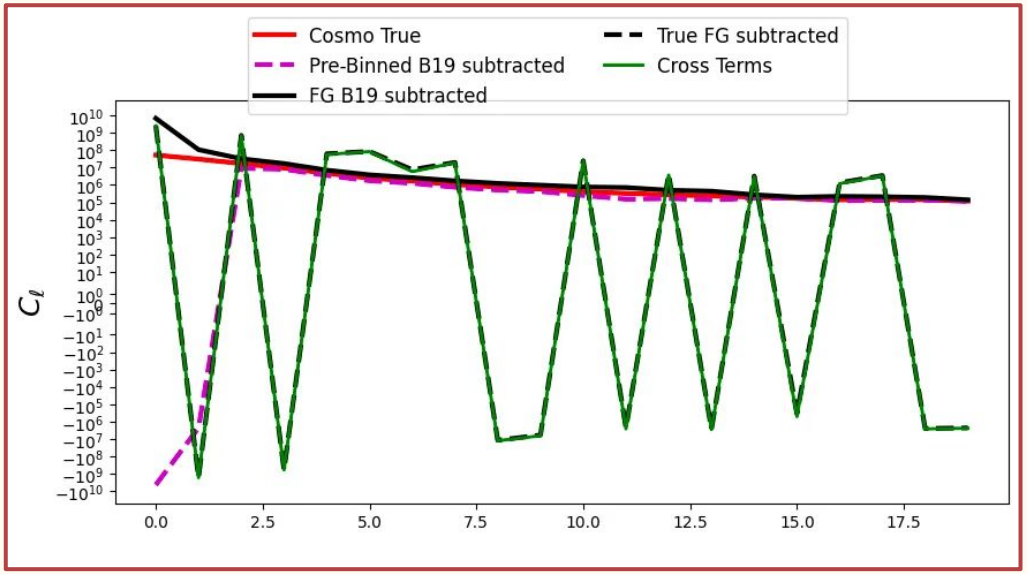
$$P_{11}(k) \propto \langle \tilde{T}_1(k)\tilde{T}_1(k)\rangle = \langle \tilde{s}_1\tilde{s}_1\rangle + \langle \tilde{s}_1\tilde{f}_1\rangle + \langle \tilde{f}_1\tilde{s}_1\rangle + \langle \tilde{f}_1\tilde{f}_1\rangle$$



You would measure the cross-terms (with a small signal power contribution)! Useless!

Recovers *all* foreground power spectrum cross-terms.

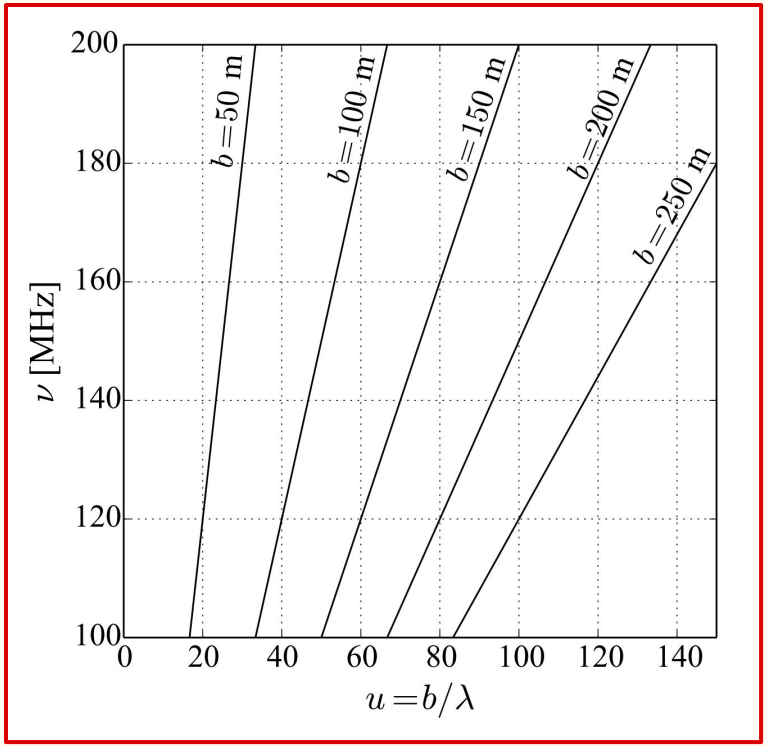
$$P_{ii}(k) = \frac{P_{ij}(k)P_{ik}(k)}{P_{jk}(k)} \longrightarrow P_{ij}(k) = \boxed{\langle \tilde{f}_i \tilde{f}_i \rangle + 2\langle \tilde{s}_i \tilde{f}_i \rangle}$$



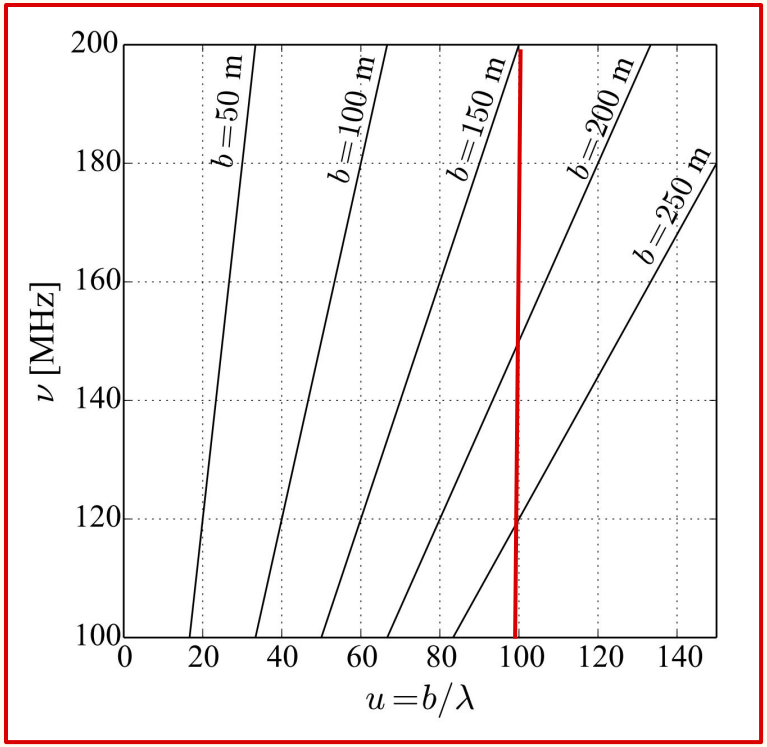
Since you are using the data to build the foreground subtraction model, you get all the cross-terms! This generalizes to other things that could be coupled to the foregrounds (e.g. noise)

$$\langle \tilde{n}_i \tilde{f}_i \rangle$$

uv samples as a function of frequency



uv samples as a function of frequency



uv samples as a function of frequency