HI stacking with MeerKLASS

Challenges, Results and Prospects for SKA-Mid

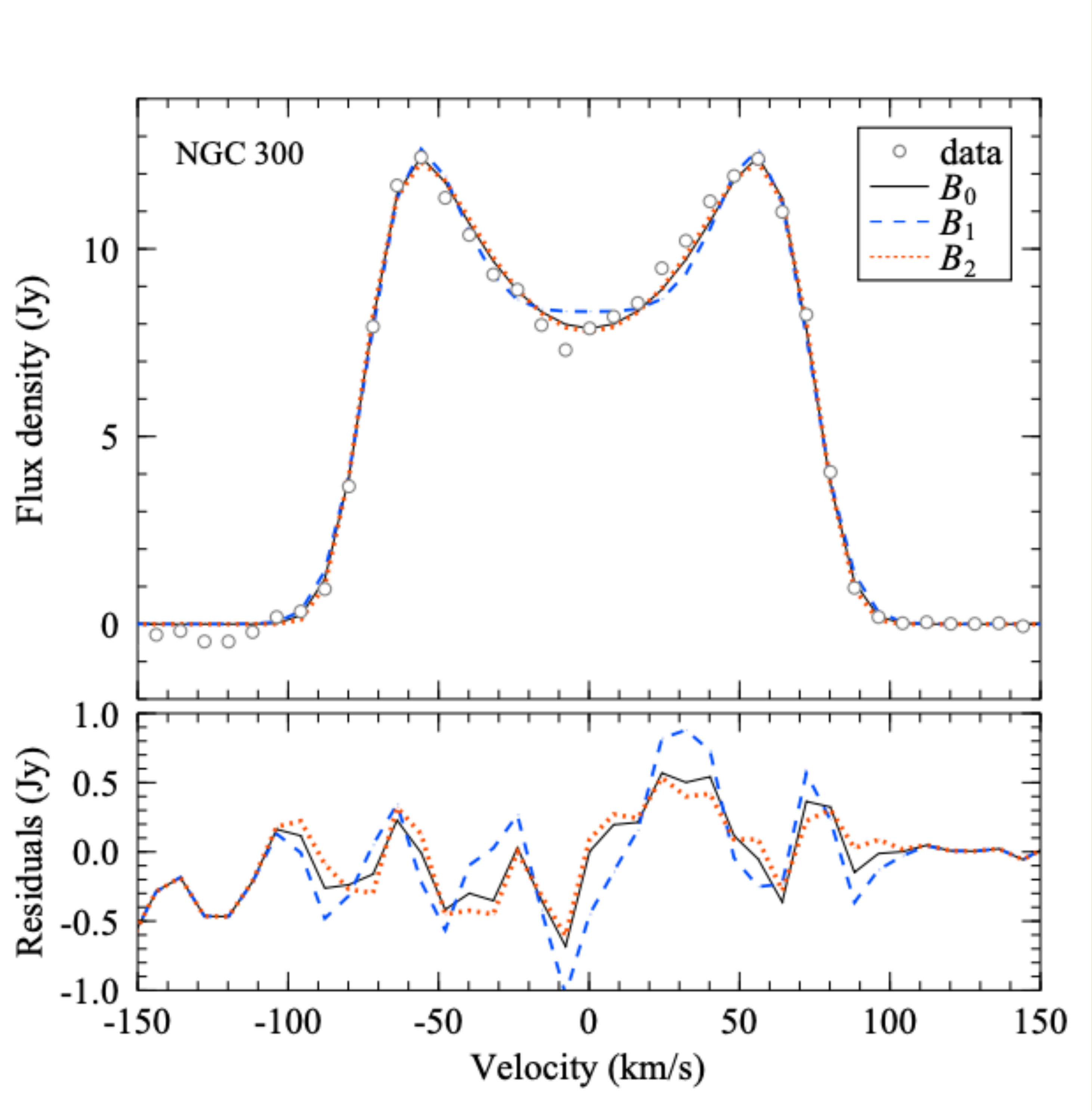
陈兆庭 (Chen, Zhaoting),
with Alkistis Pourtsidou, Steve Cunnington and Laura Wolz

SKAO Cosmology SWG 2024
05/11/2024

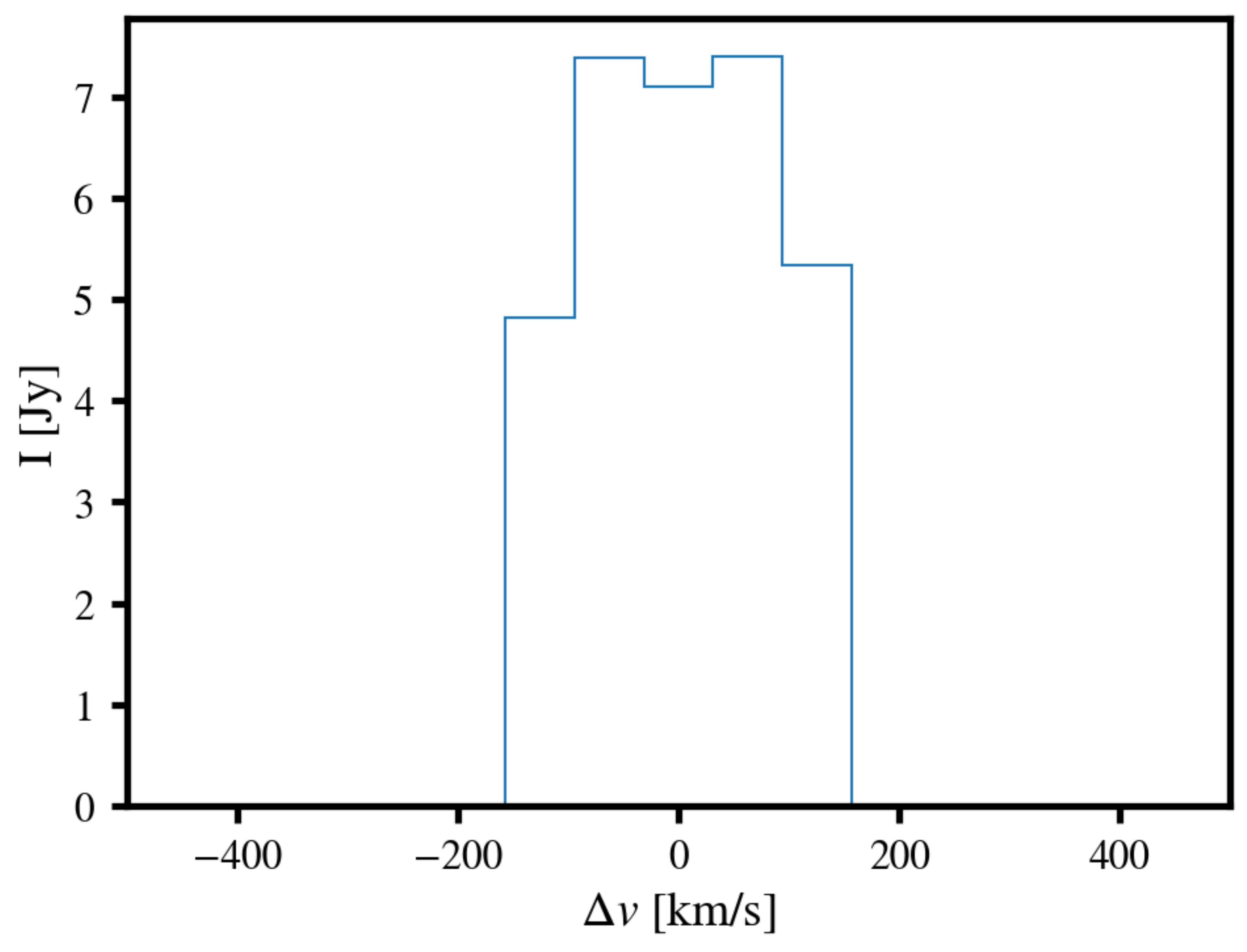
The emission line profile of HI galaxy

- In 4k mode, the MeerKAT L-band receiver has a velocity resolution of ~50km/s at z~0 (~65km/s at z~0.4).

Westmeier et al., 1311.5308



* Illustrative



An ideal stacking for single dish IM

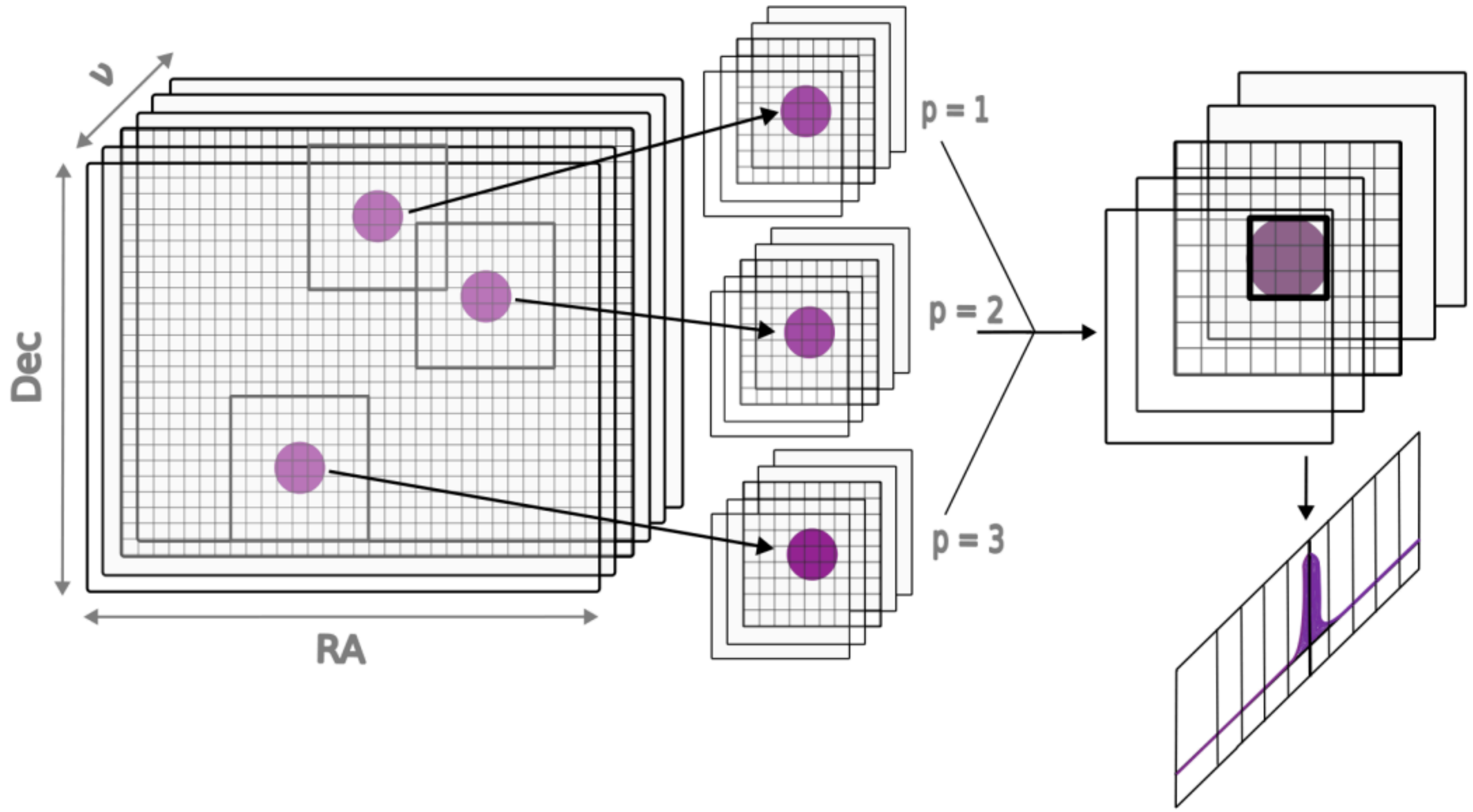
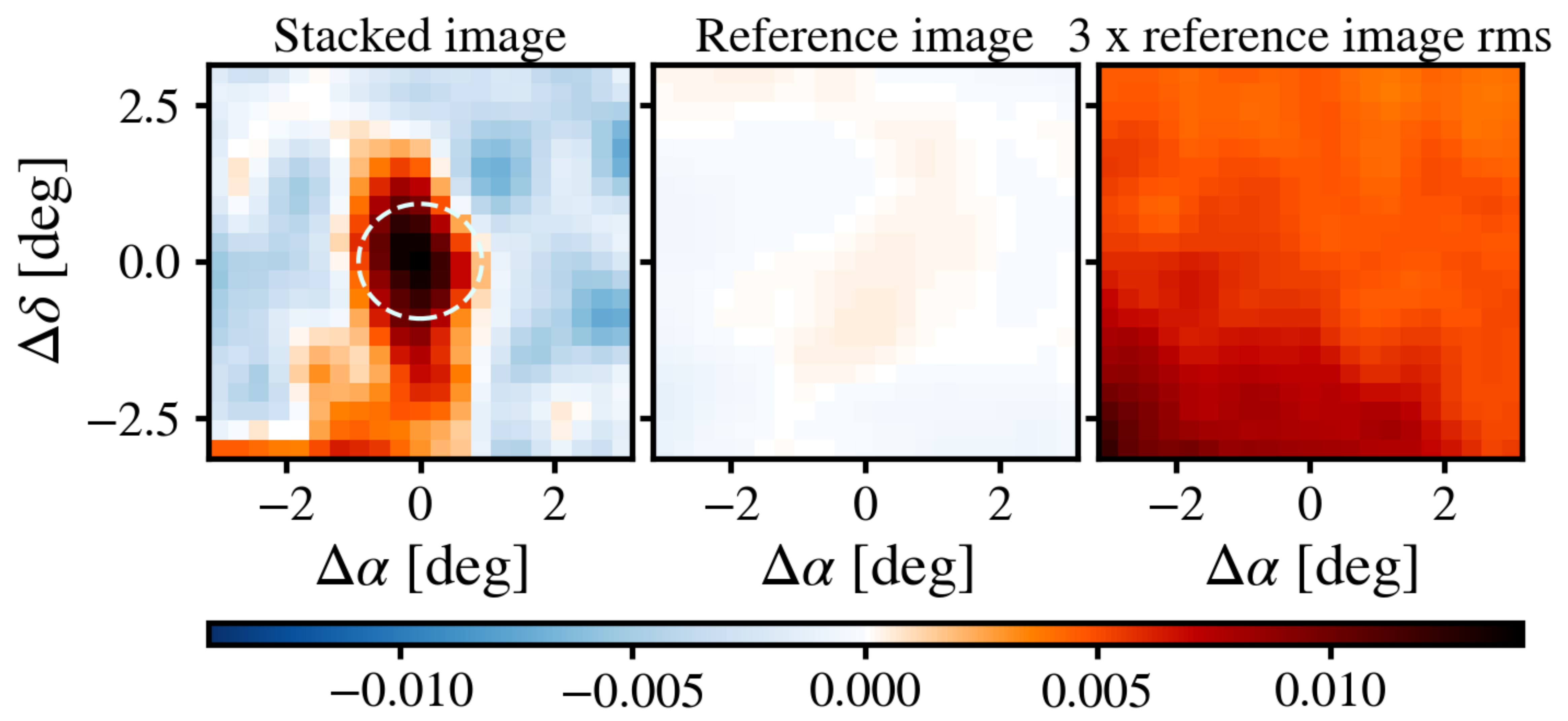


Figure 3. A diagram displaying the methodology used for this stacking analysis. From the full three-dimensional COMAP data cube (left), smaller 3-D cubelets are cut out centered on the position of each eBOSS object (center). These cubelets are then averaged into a single 3-D stack (top right), which is used to determine the average COMAP spectrum of the eBOSS objects (bottom right).

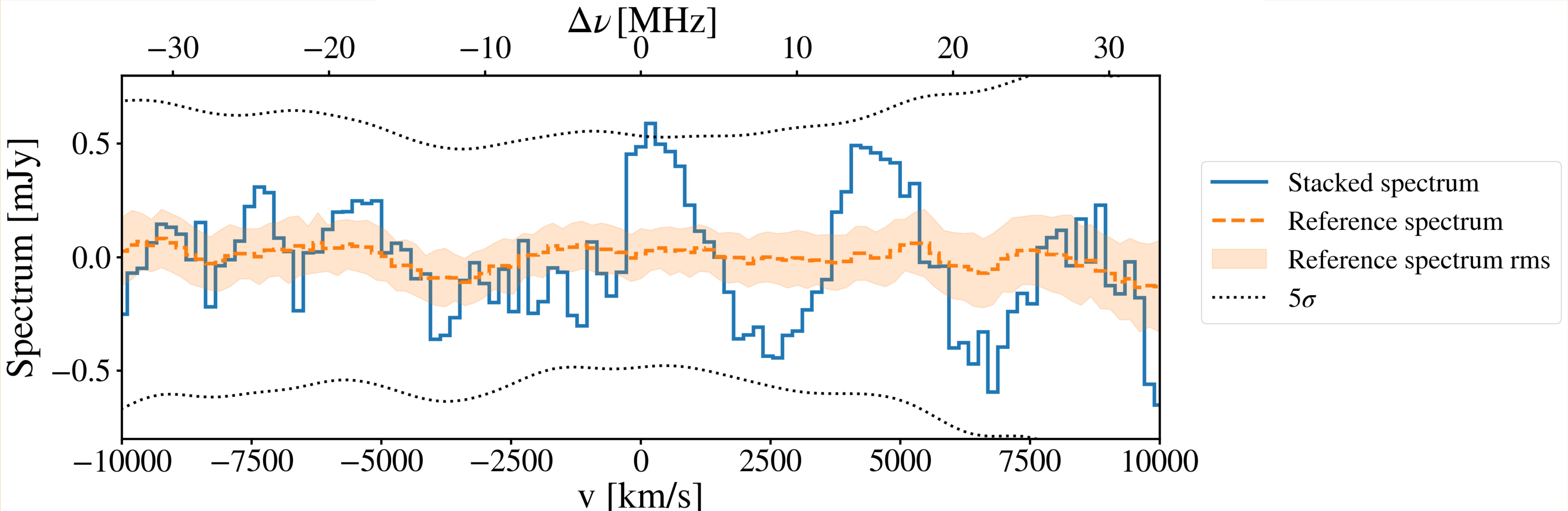
COMAP, 2304.09832

The reality

~ 8 sigma



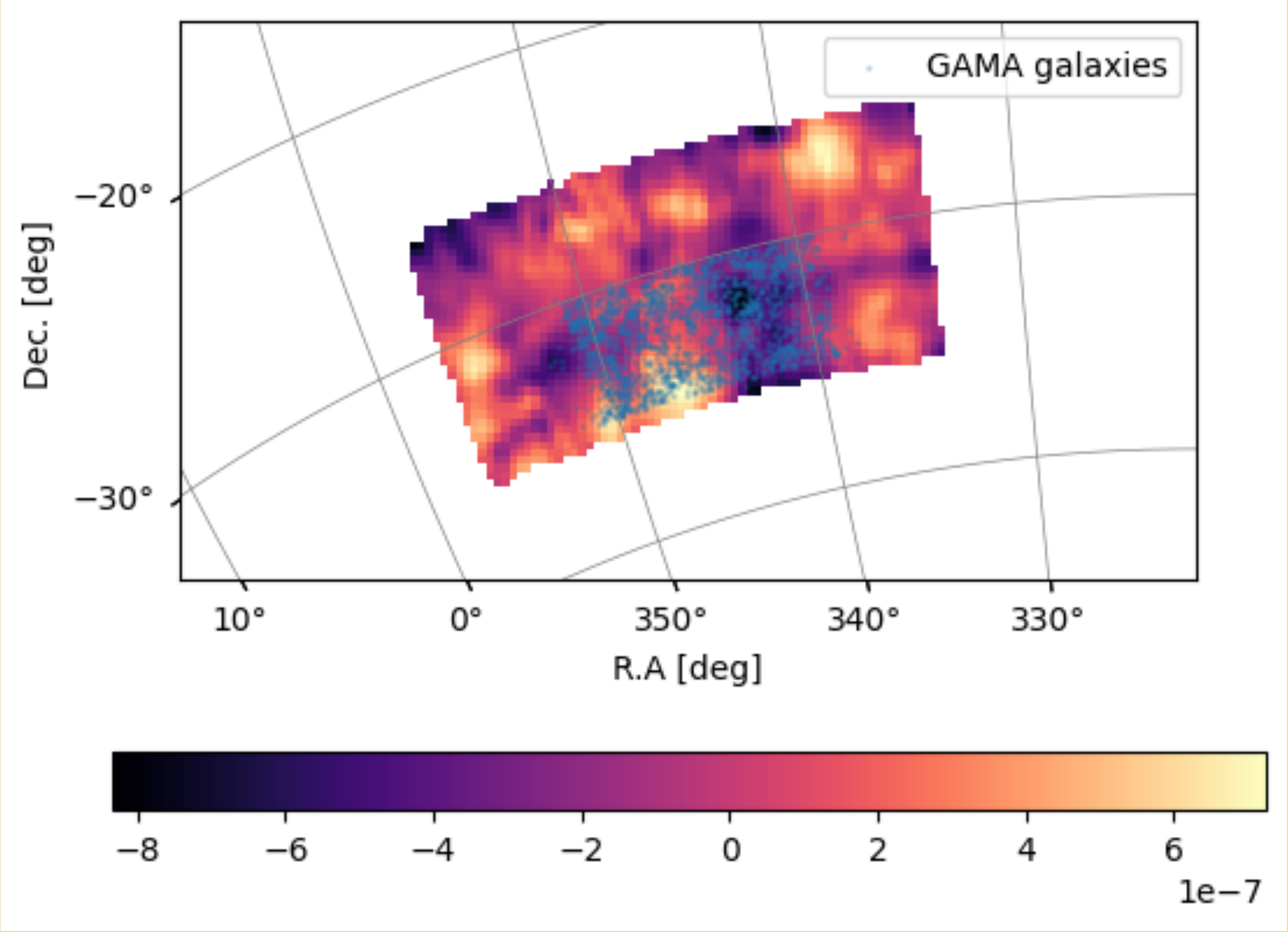
~ 6 sigma



The MeerKLASS collbartion,
2407.21626

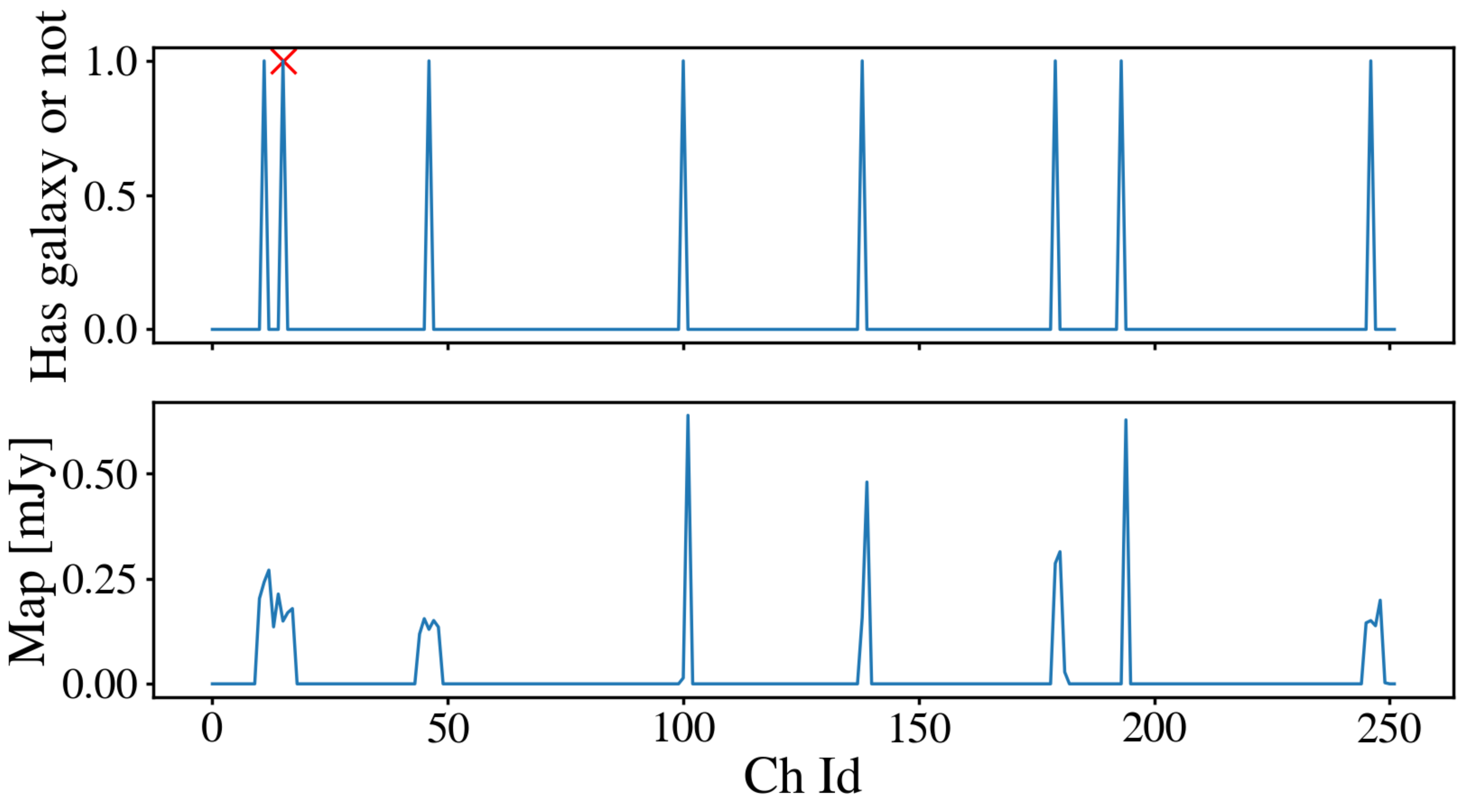
Problem 1: Too many galaxies, too few pixels

- In our data (L-band deep-field x GAMA galaxies), the area covered is only ~60 sqdeg with 2269 galaxies, resulting in ~40 galaxies per sqdeg. The MeerKAT beam is ~1 deg and our scanning strategy produces maps with 0.3 deg size pixel.



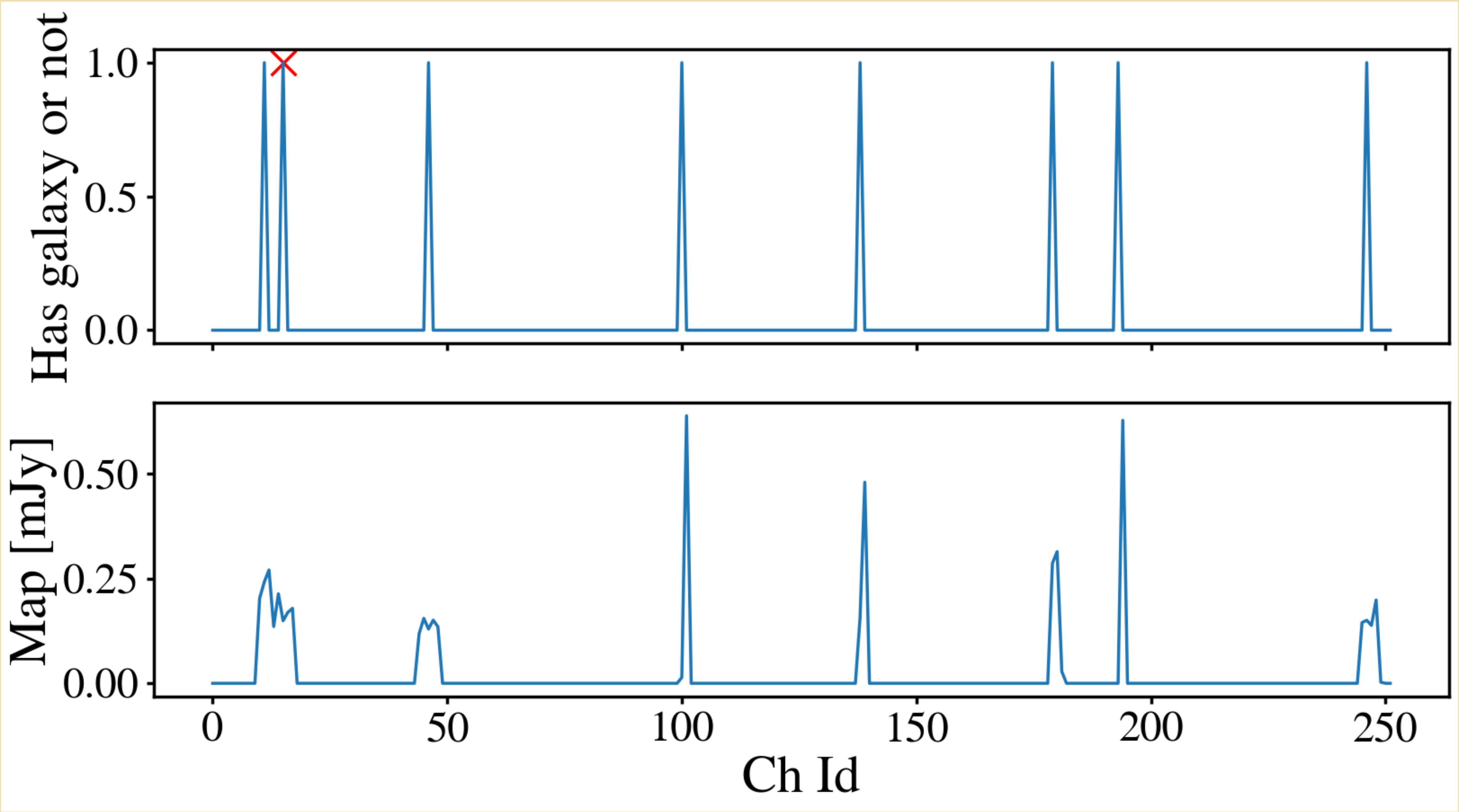
Problem 1: Too many galaxies, too few pixels

- Even without considering the beam, along the line-of-sight of a map pixel there are multiple galaxies.

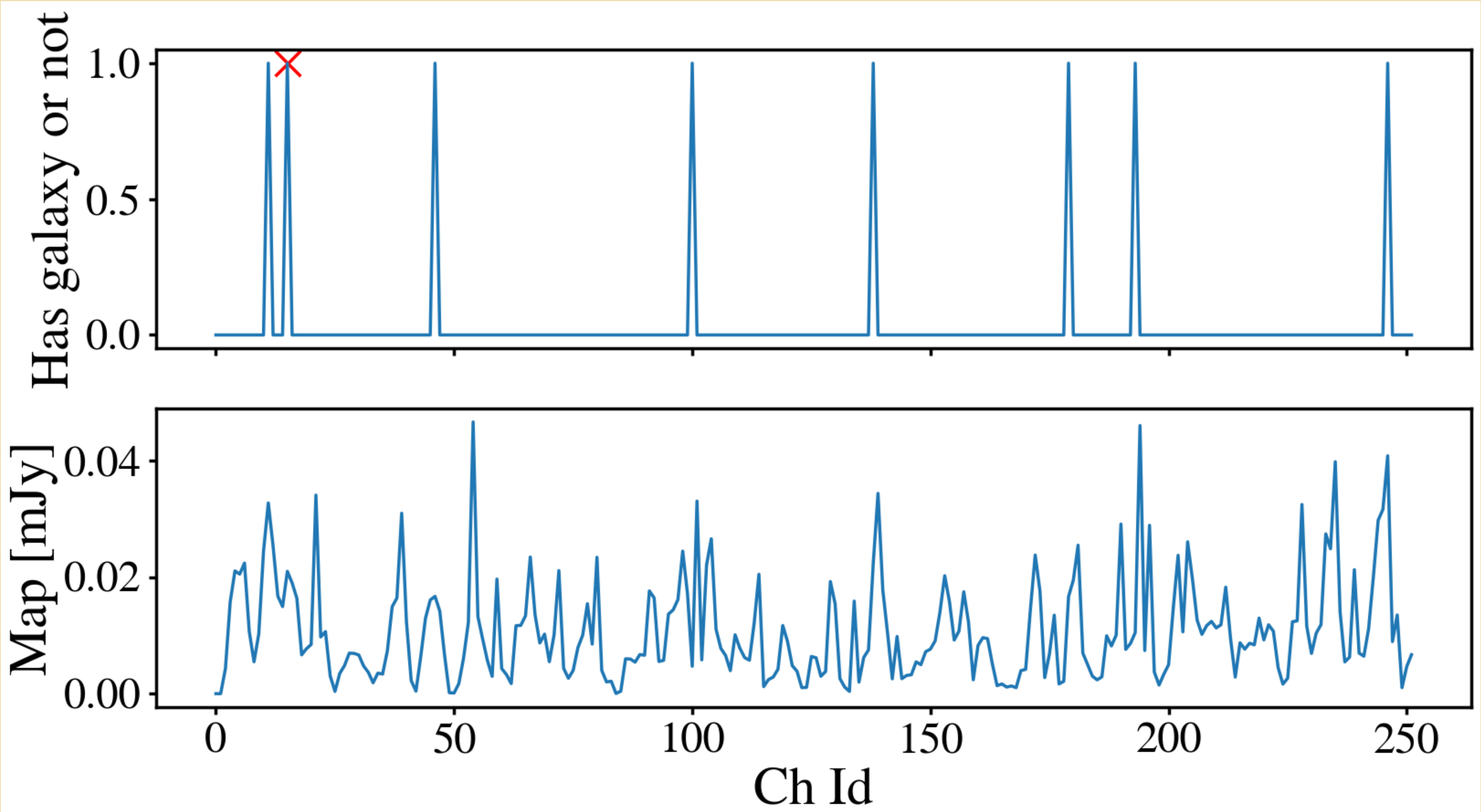


Problem 2: Scattering by the beam

- Including the 1deg beam then further creates double counting in angular space:



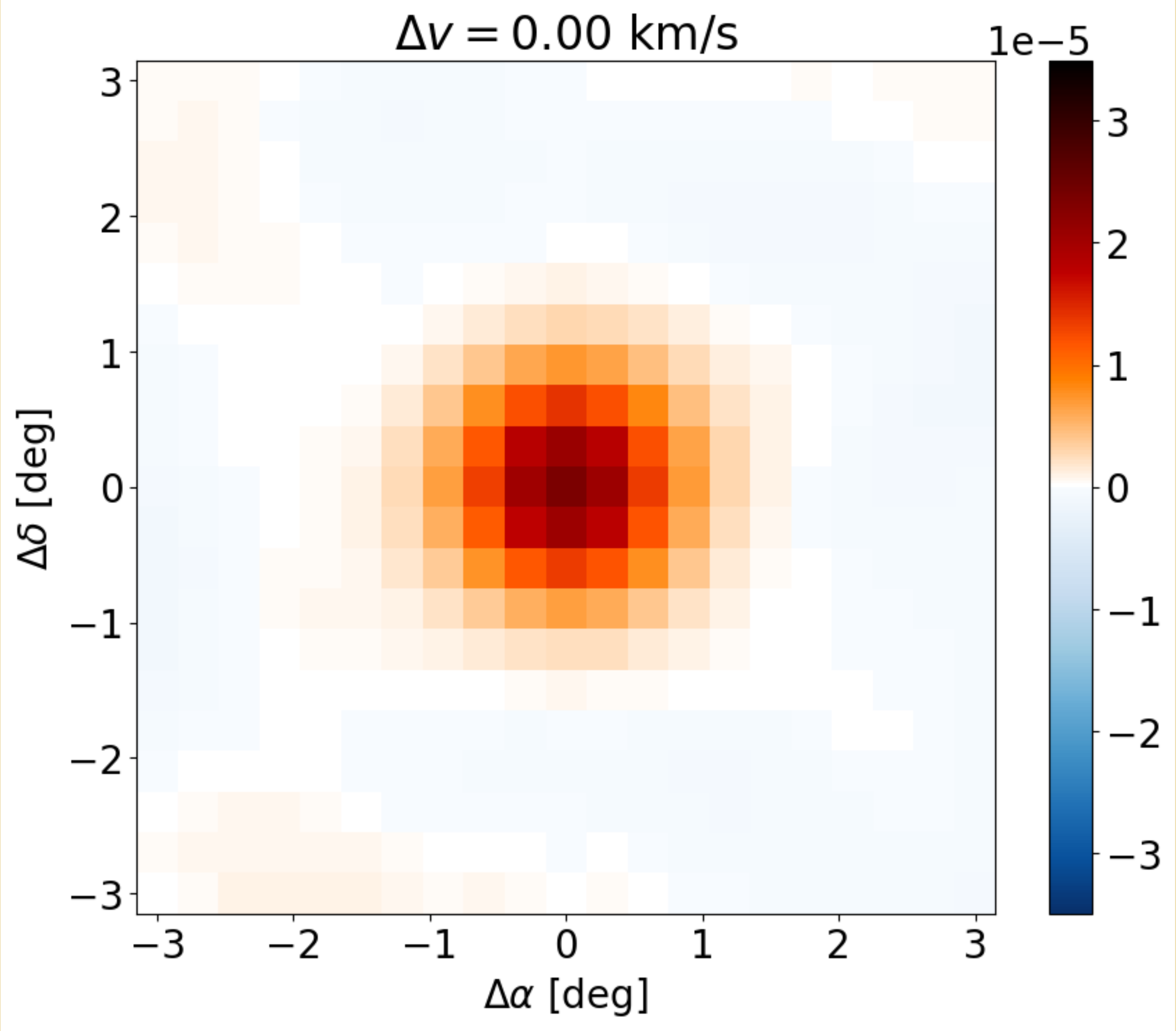
With Beam



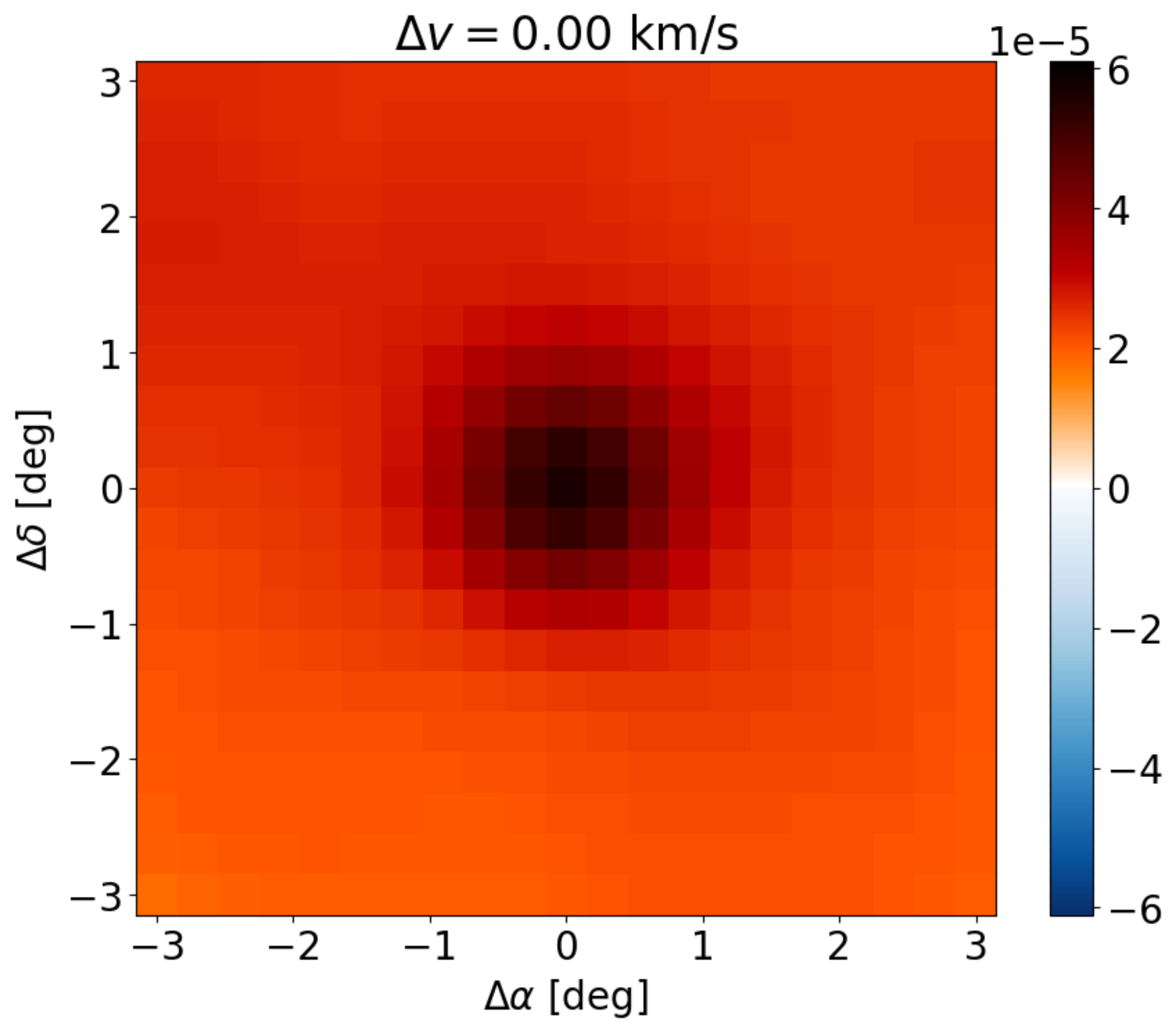
Problem 2: Scattering by the beam

- Including the 1deg beam then further creates double counting in angular space:

Ideal



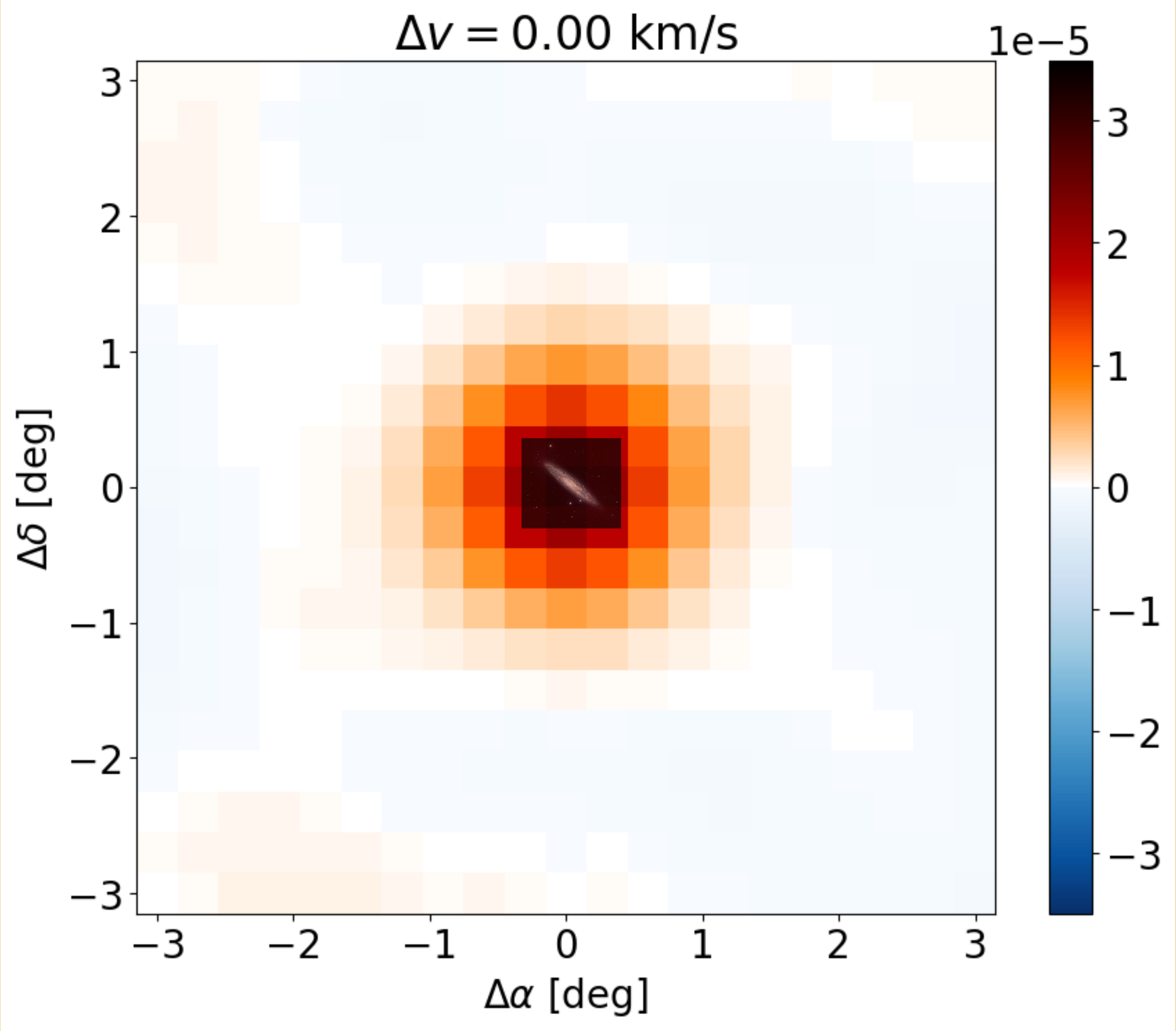
With double counting



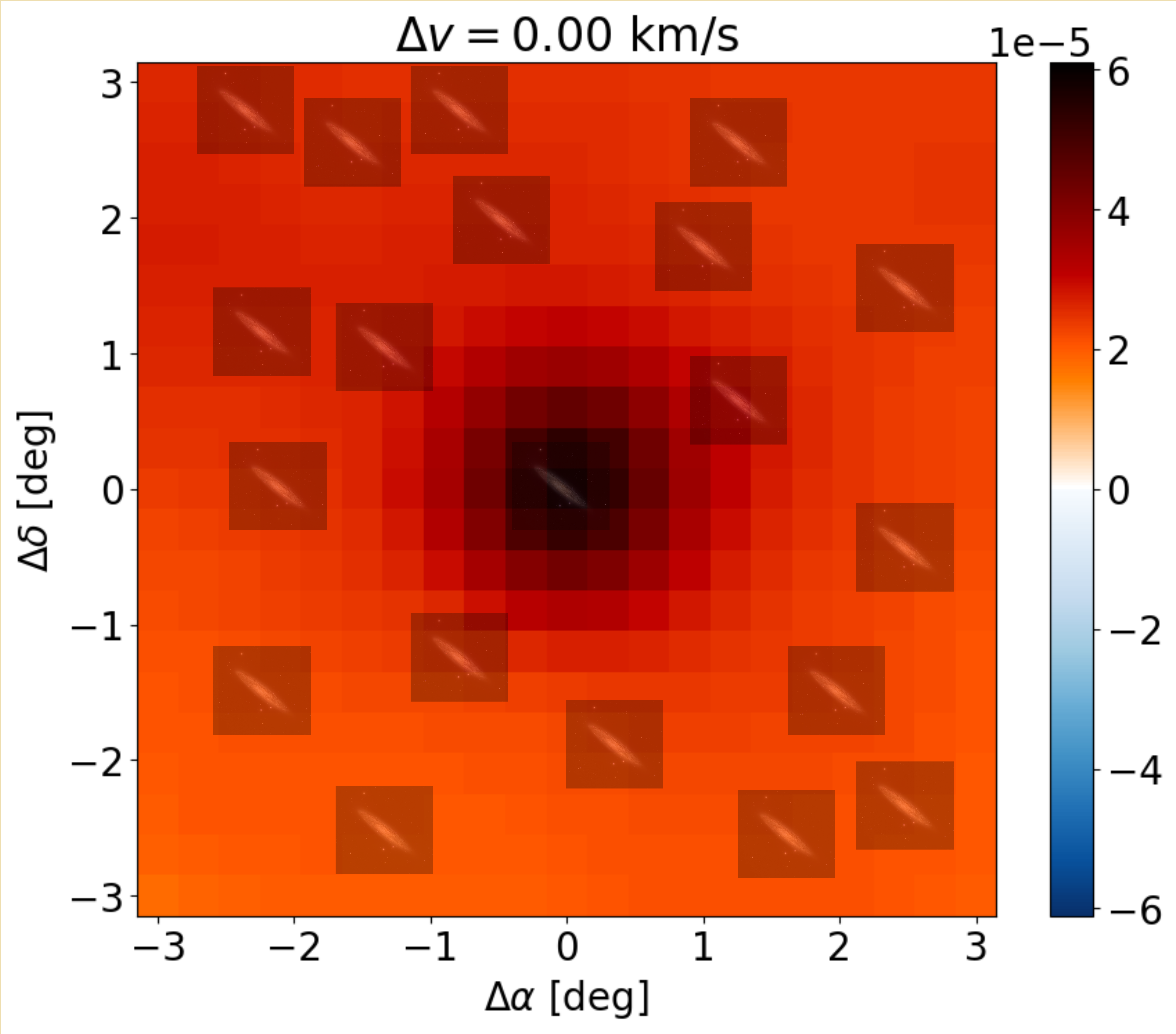
Problem 2: Scattering by the beam

- Including the 1deg beam then further creates double counting in angular space:

Ideal

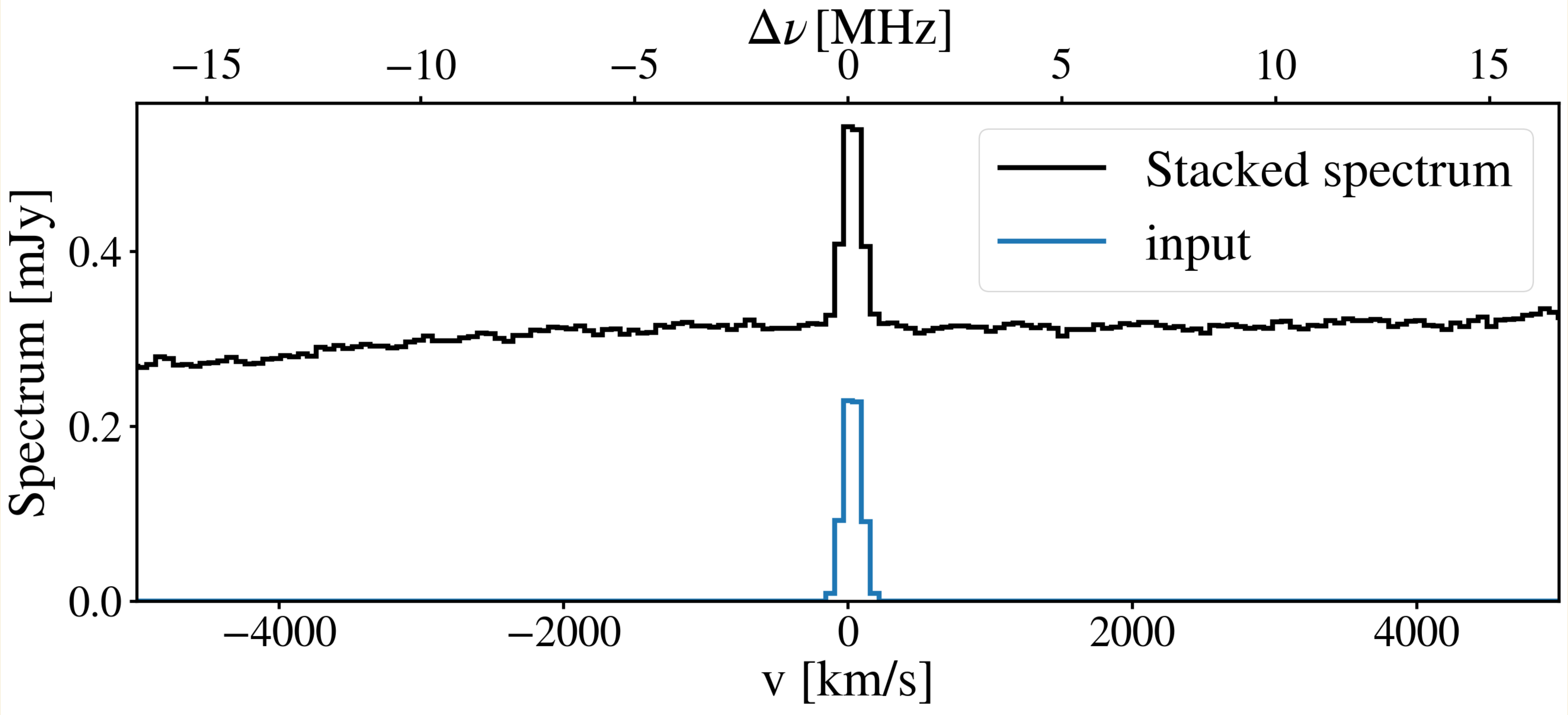


With double counting



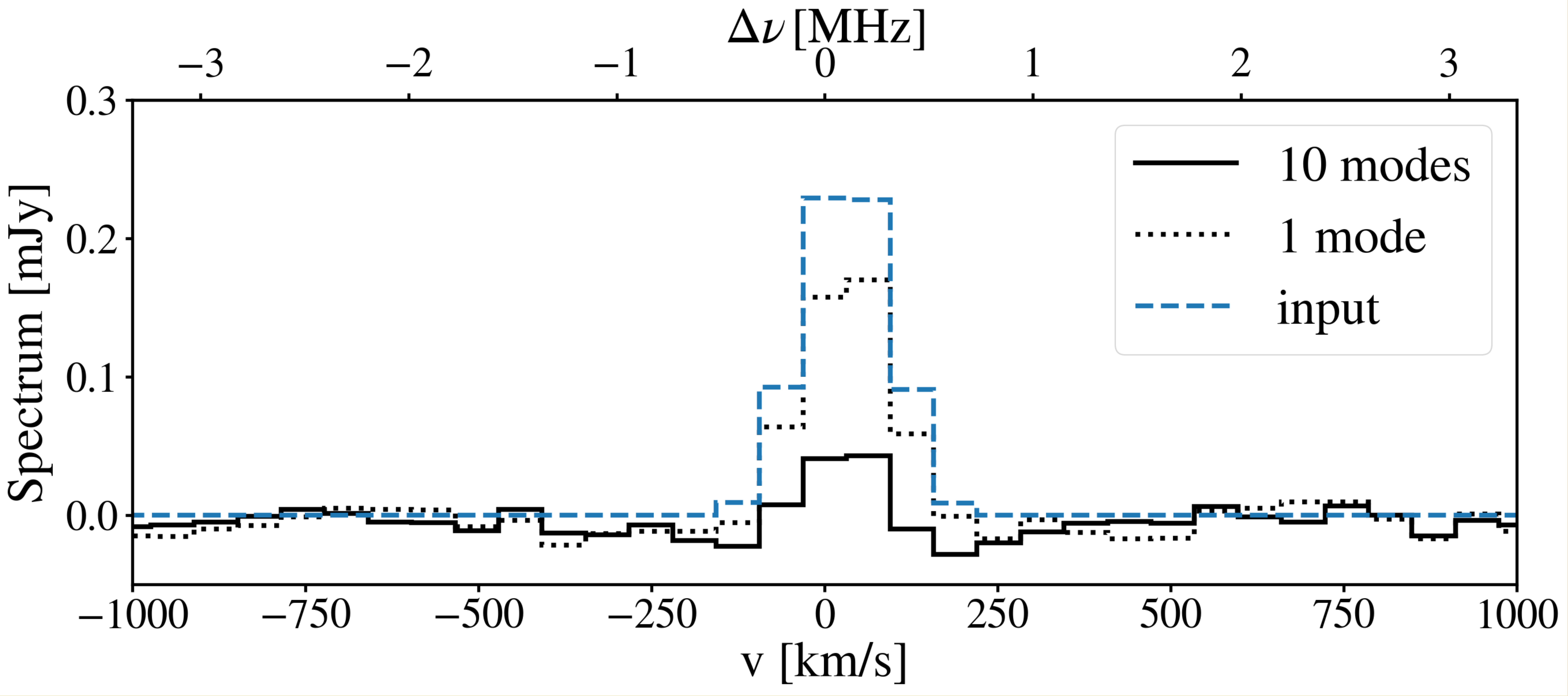
Problem 2: Scattering by the beam

- The double-counting of galaxies effectively creates a plateau:



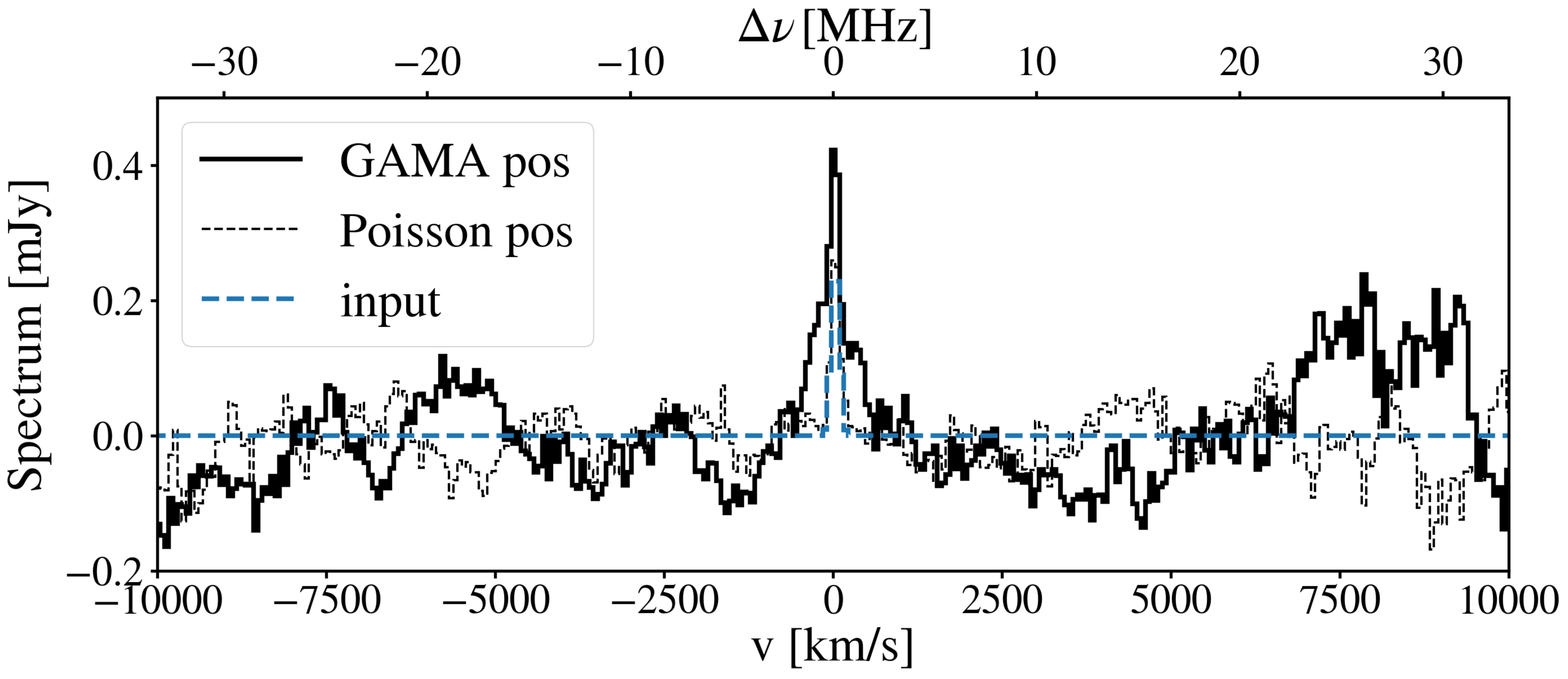
Solution: good ol' PCA

- PCA removes the plateau at the expenses of signal loss and effects from residual systematics



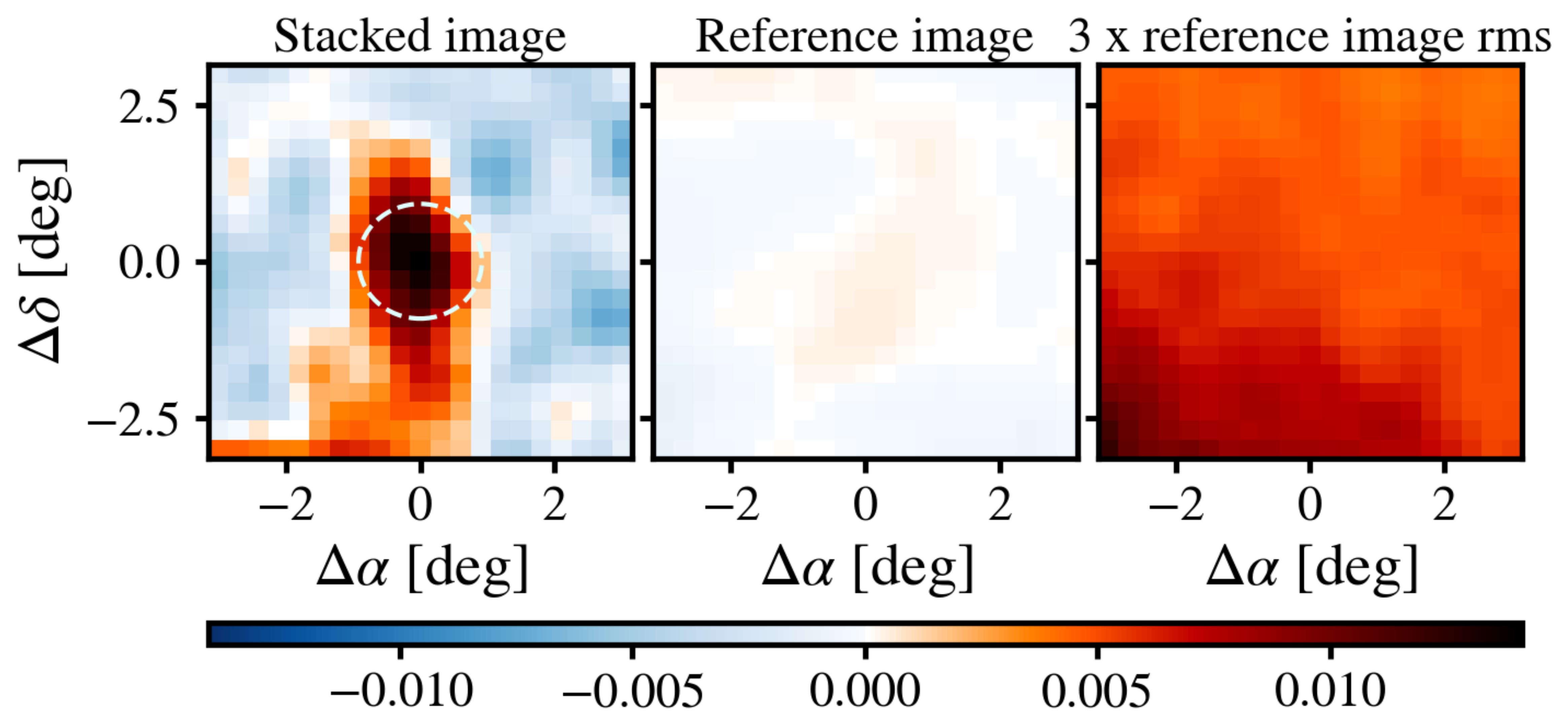
Problem 3: Clustering amplification of signal

- Including the ~1deg beam corresponds to ~15Mpc scale. The stacked signal includes a clustering component.
- Requires forward modelling of the signal.

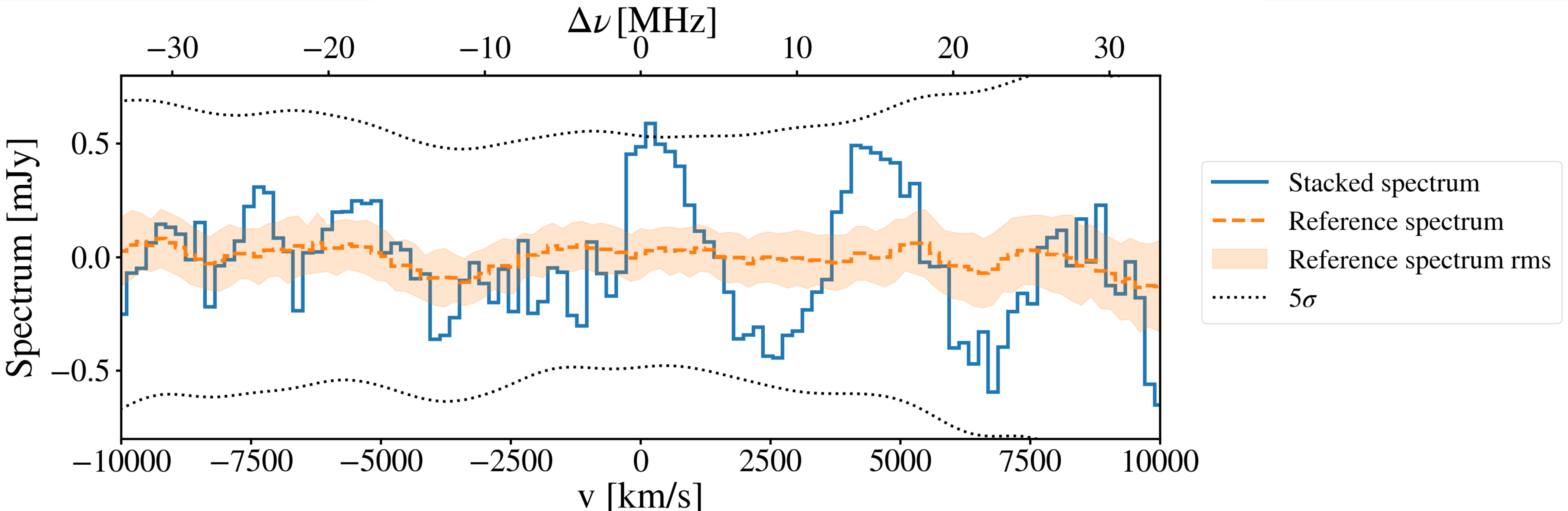


Problem 4: real map has residual systematics

~ 8 sigma

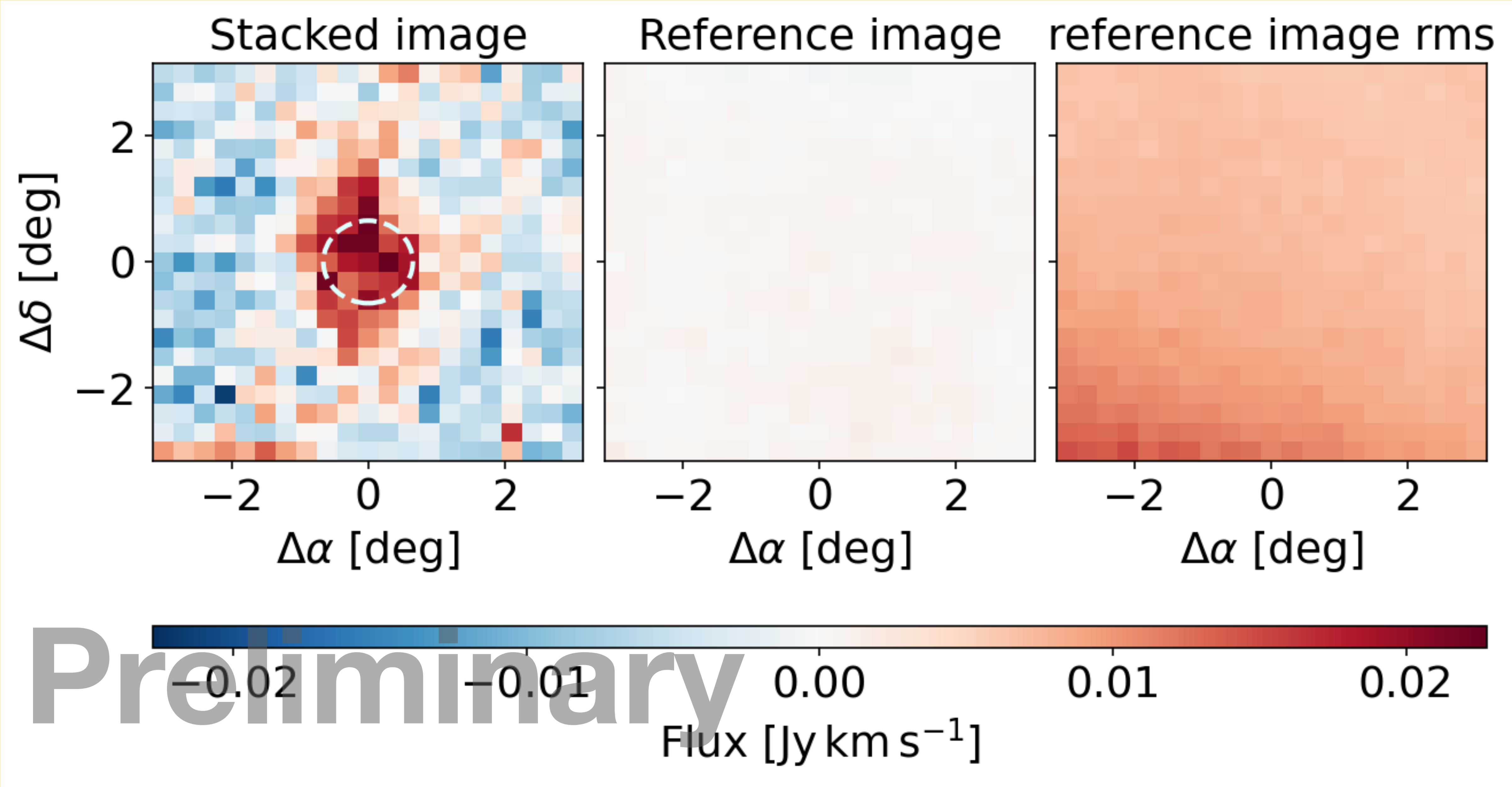


~ 6 sigma

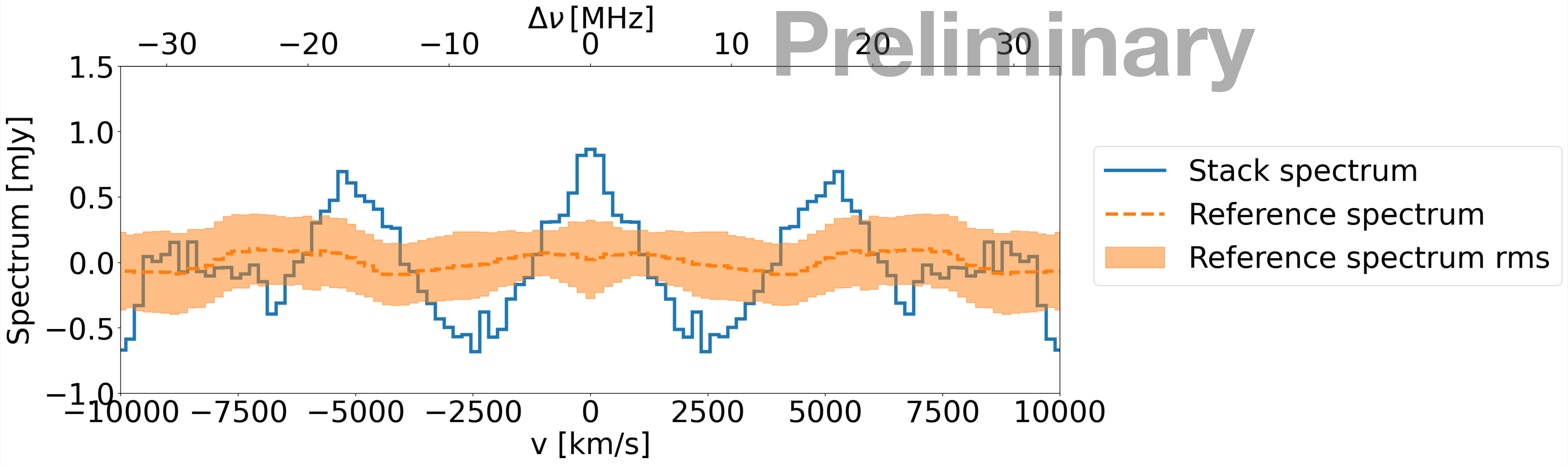


The MeerKLASS collbartion, 2407.21626

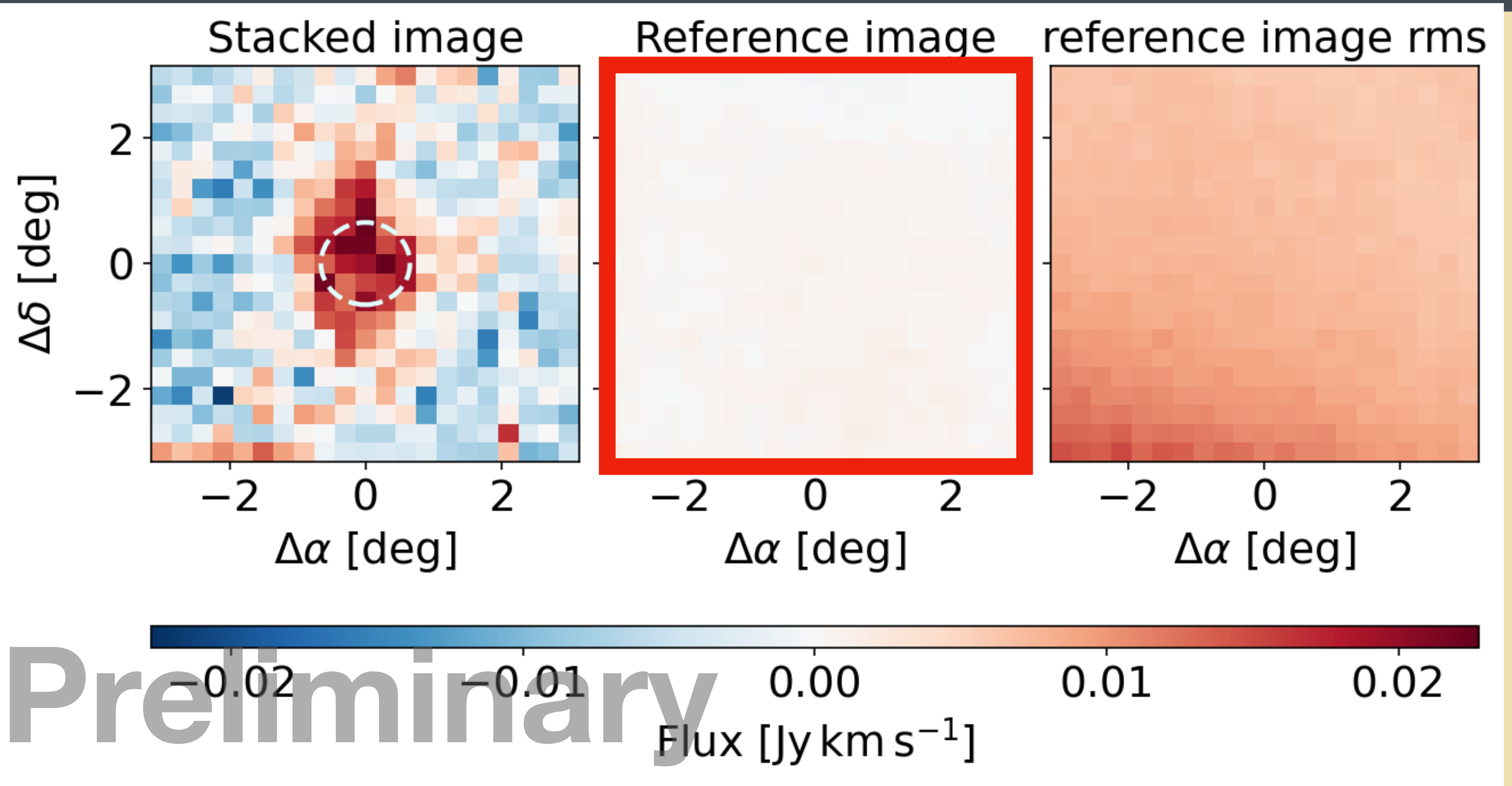
Problem 4: real map has residual systematics



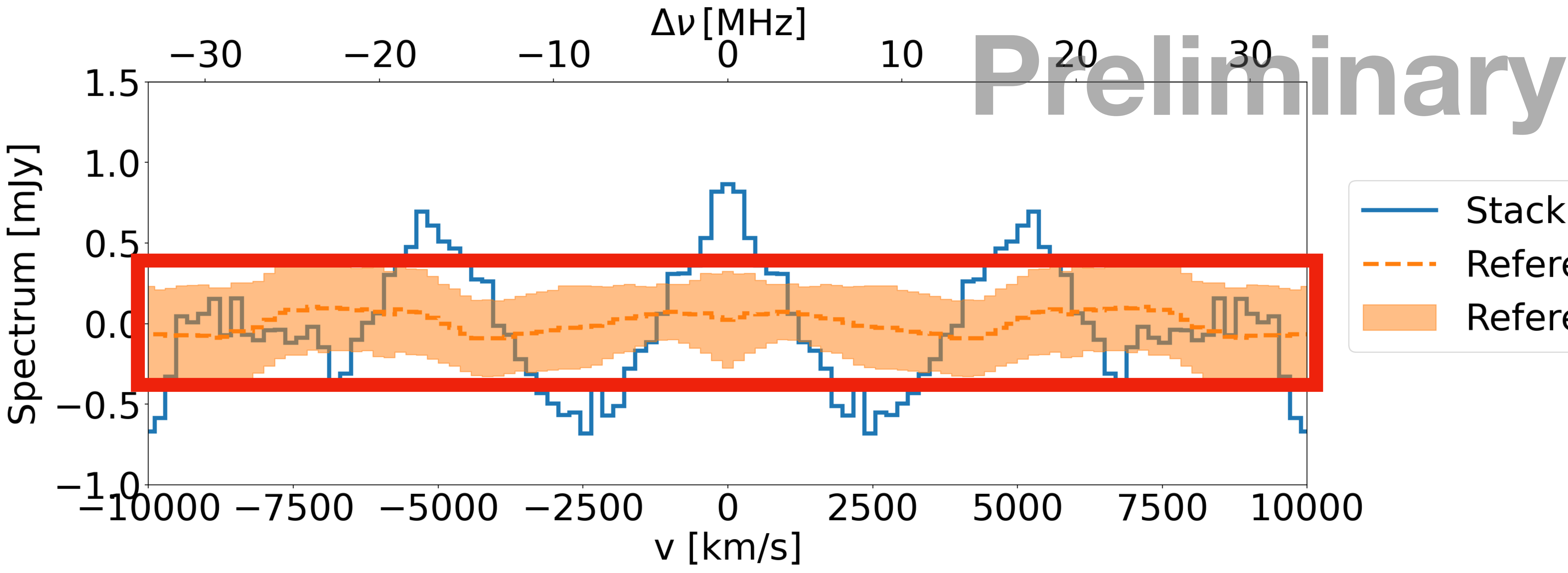
Z. Chen,
The MeerKLASS collbartion, in prep



Hunting down the source of systematics

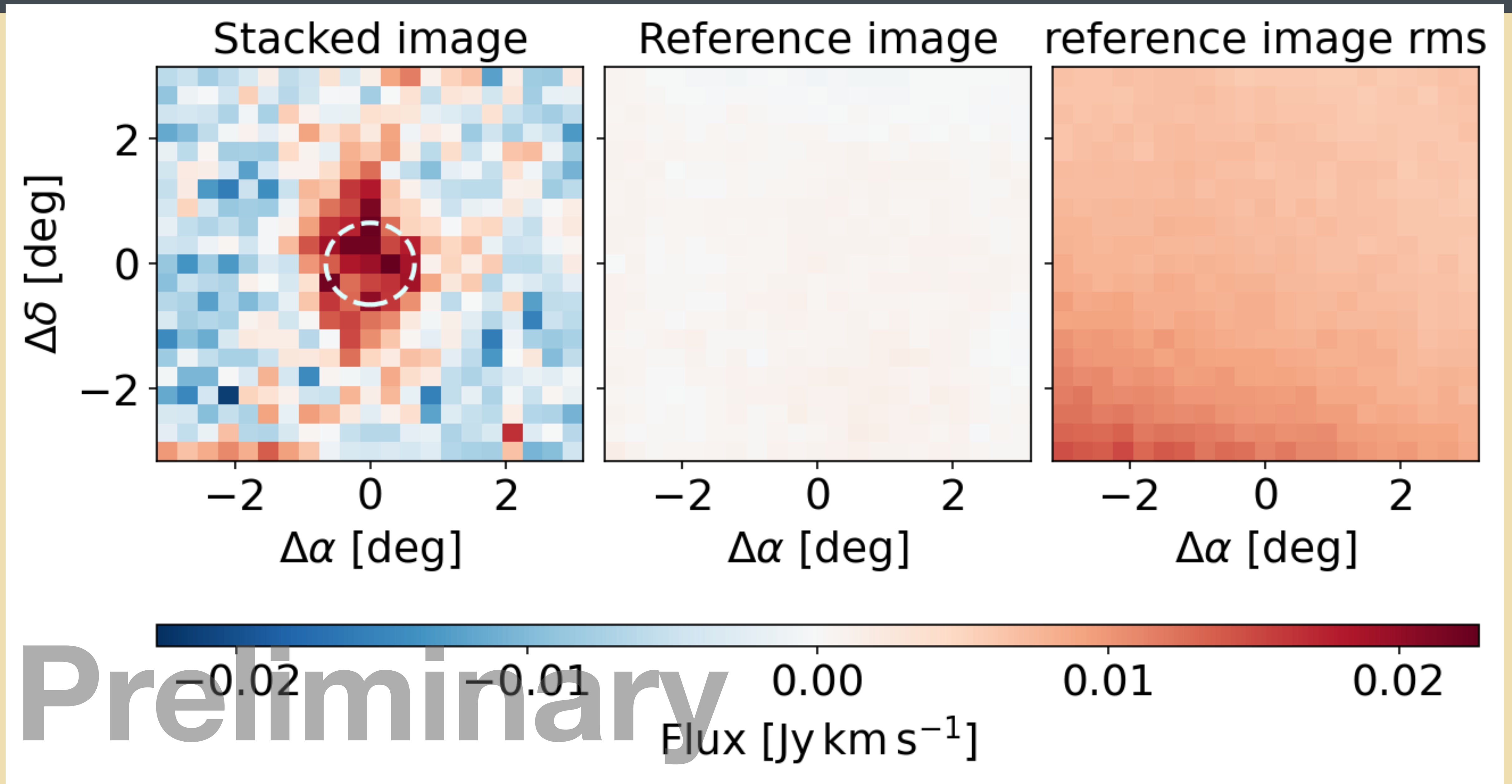


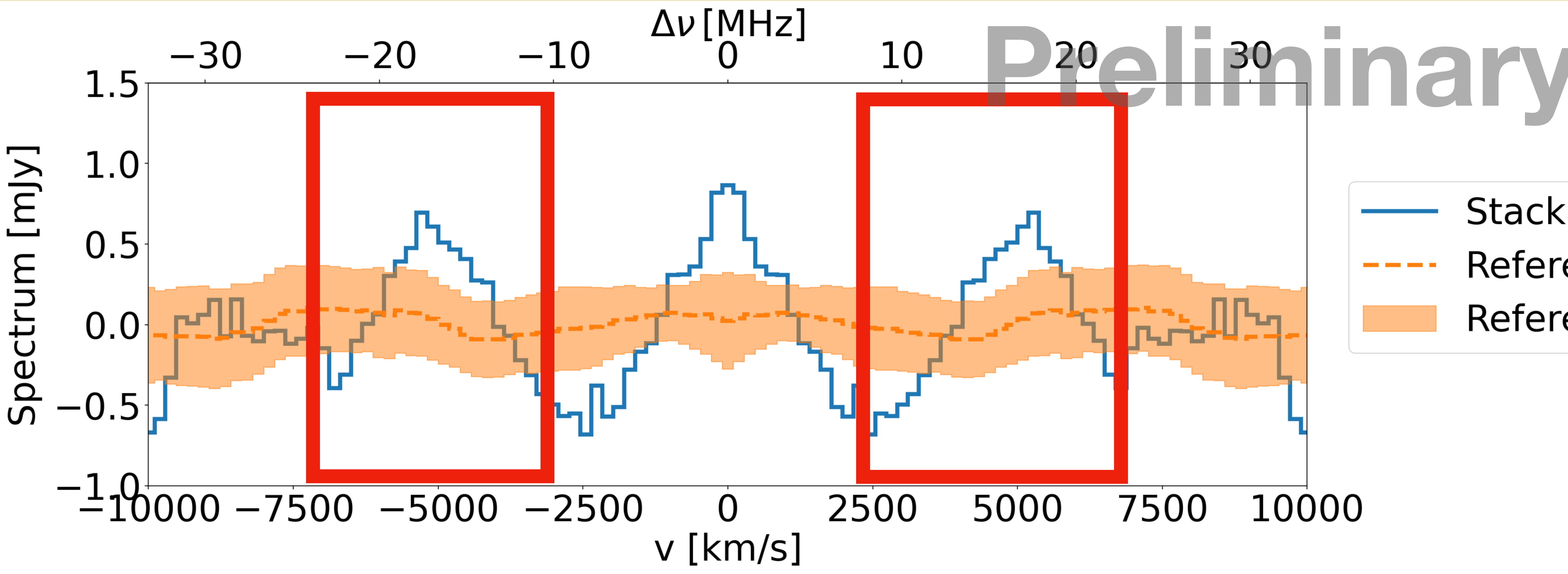
- The reference image and spectrum are calculated by randomly shuffling galaxy positions
- Consistent with zero
- The contamination is not an additional component of the map, but distortion of the HI signal.



Hunting down the source of systematics

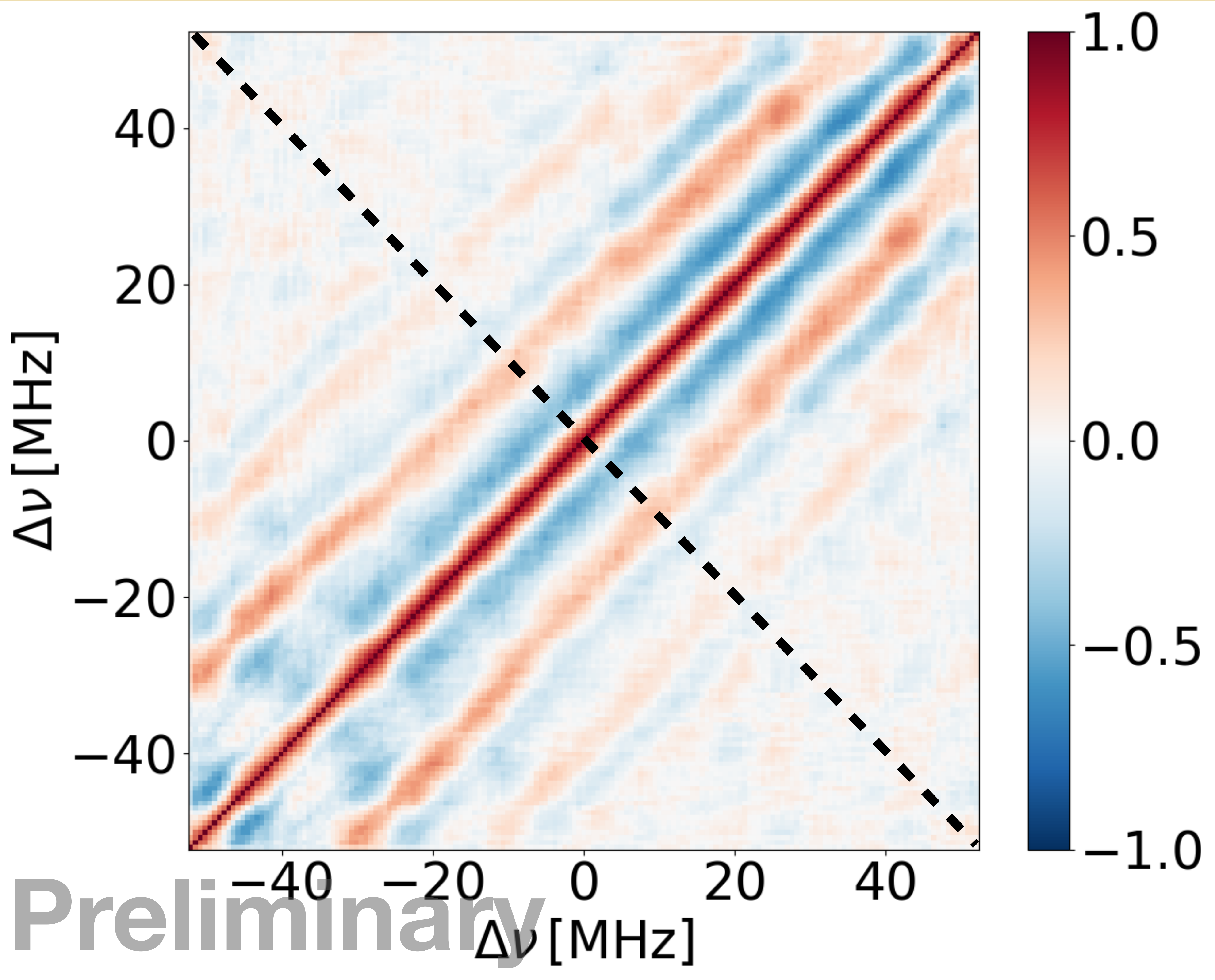
- There is an oscillating structure in frequency and no visible structure in angular space.
- Multiplicative or convolutional effects along the frequencies?





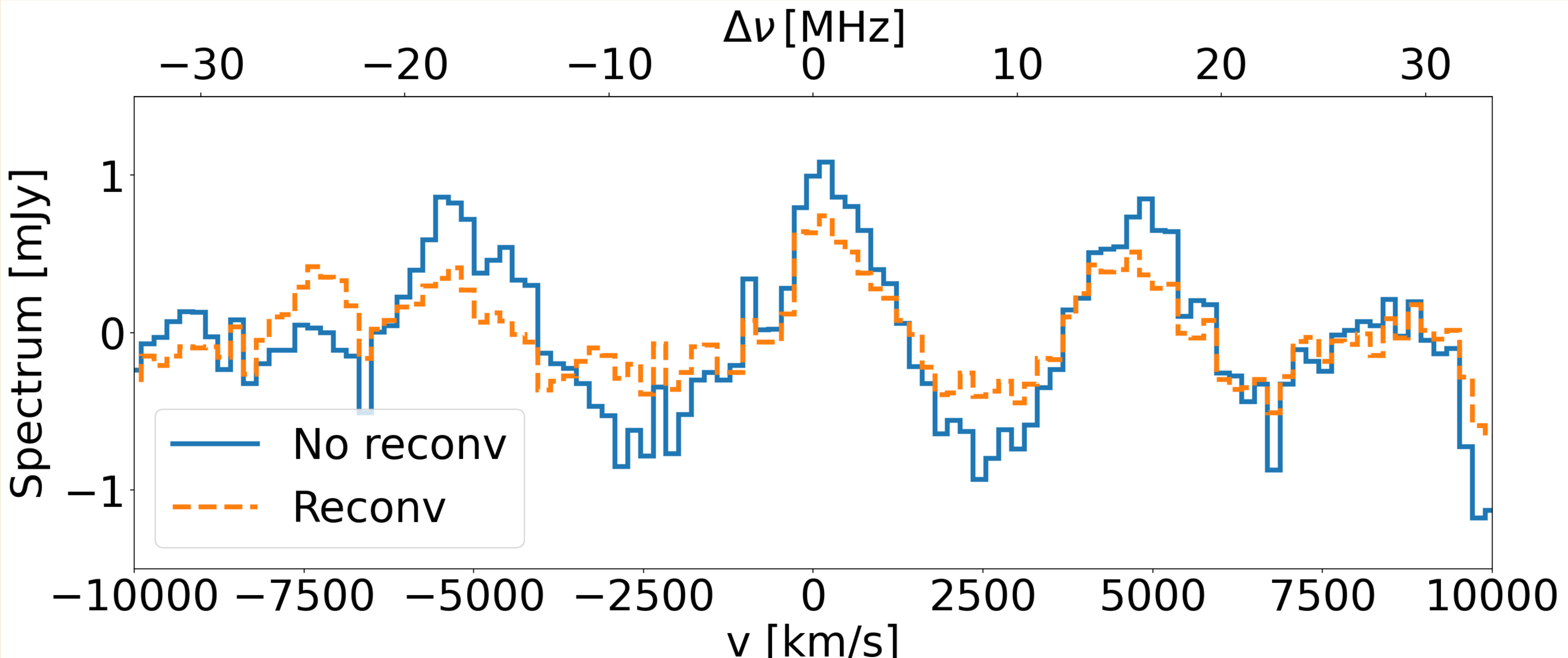
Hunting down the source of systematics

- Using the random shuffling, we can also calculate the covariance of the reference spectrum.
- The correlation matrix shows the oscillation as well
- Suggests that the systematic effect convolves with the signal



Hunting down the source of systematics

- In the L-band deep-field analysis, the maps at each frequency are reconvolved by a common Gaussian beam to a lower resolution.
- If the reconvolution is skipped, the oscillating structure is more visible and at a constant interval.
- The systematics is likely related to the chromaticity of the instrument



Hunting down the source of systematics

- The MeerKAT beam is known to have a rippling structure
- The period of the ripples is ~20MHz, matching the frequency scales of the systematics
- Known to cause effects in foreground cleaning for intensity mapping

Asad et al., 1904.07155

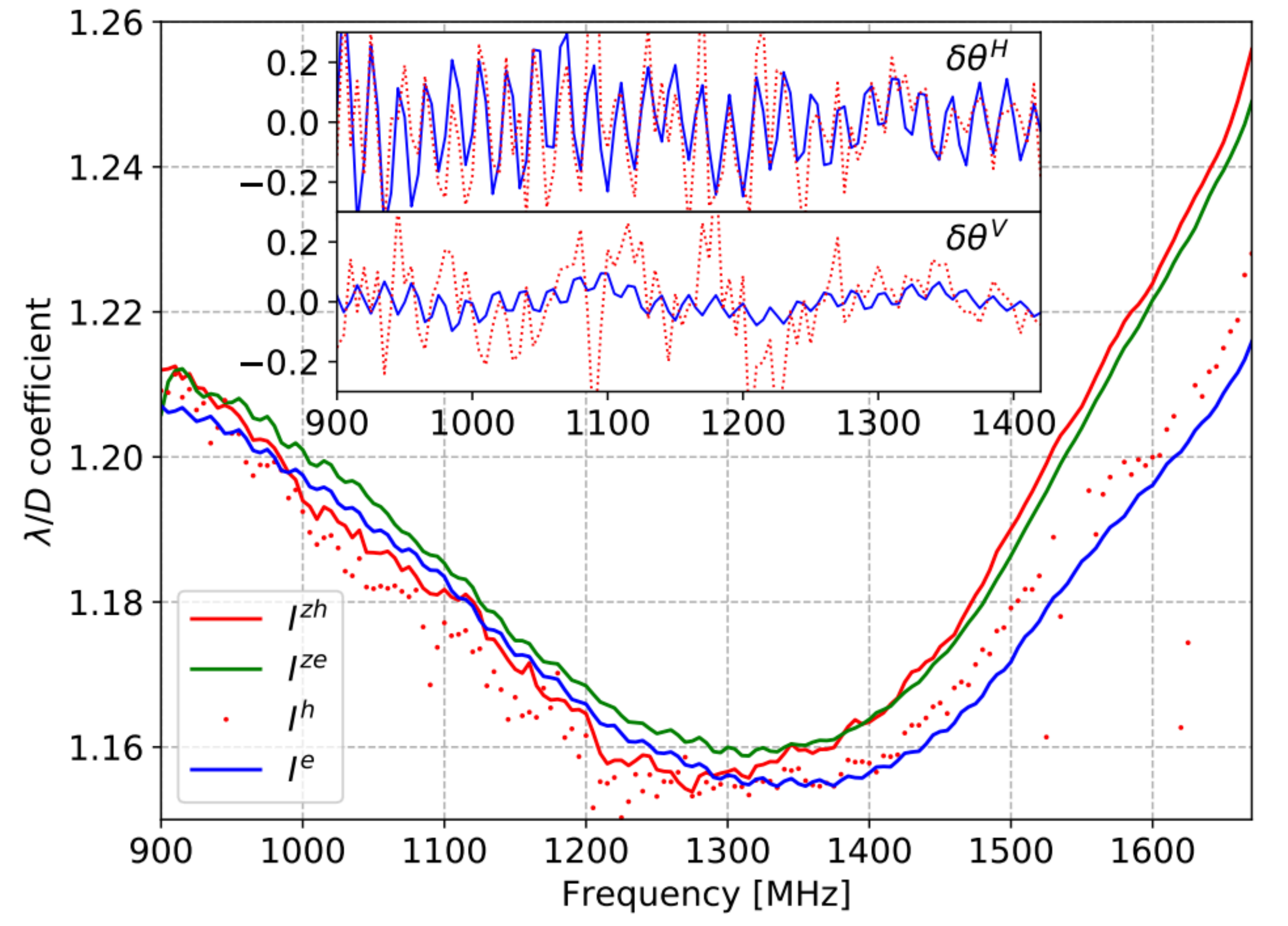
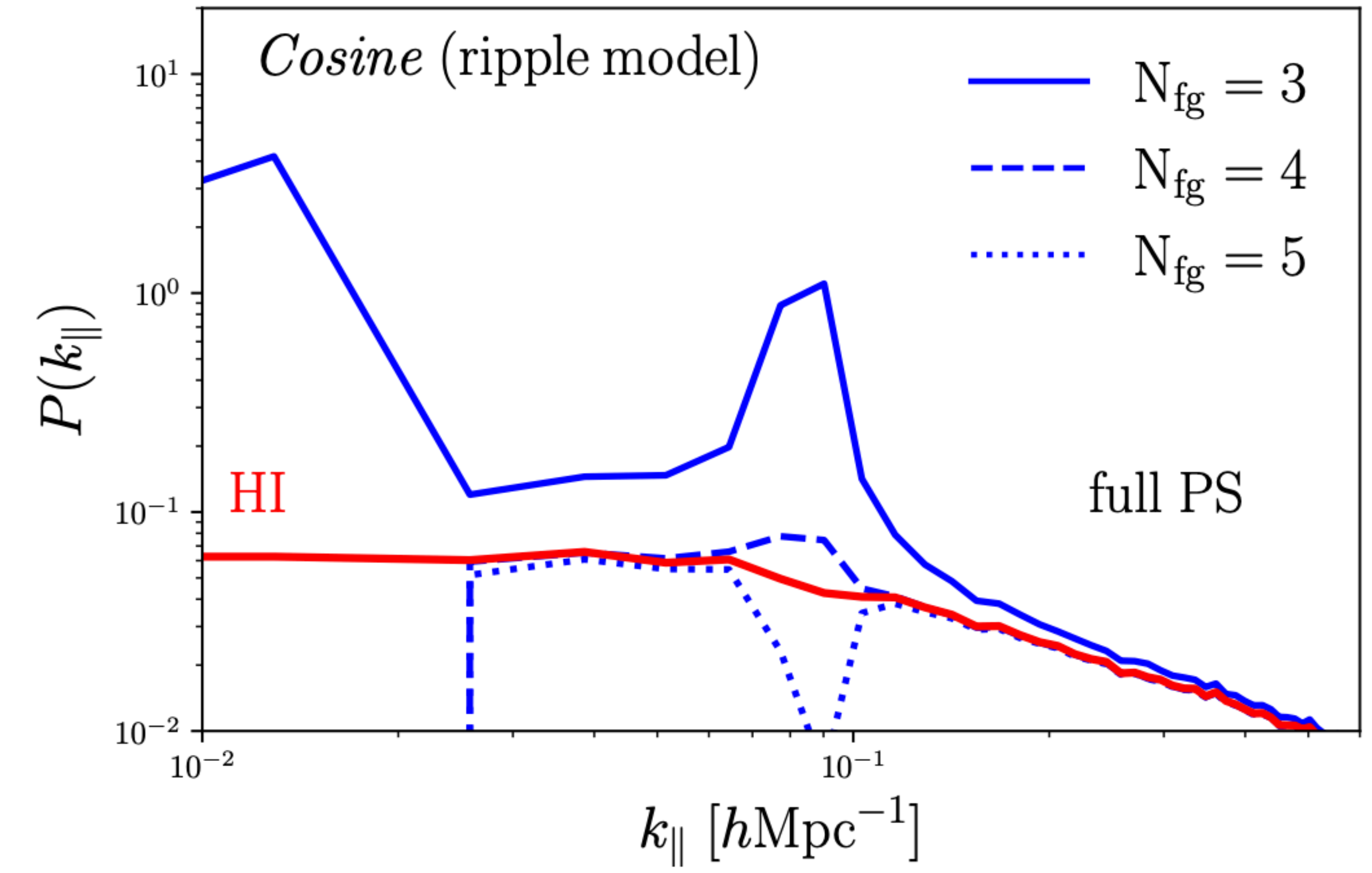


Figure 16. Coefficients of the theoretical beamwidth (λ/D) as a function of frequency for the given AH (red dots) and EM (blue) datasets, and the Zernike-based models created from the AH (red

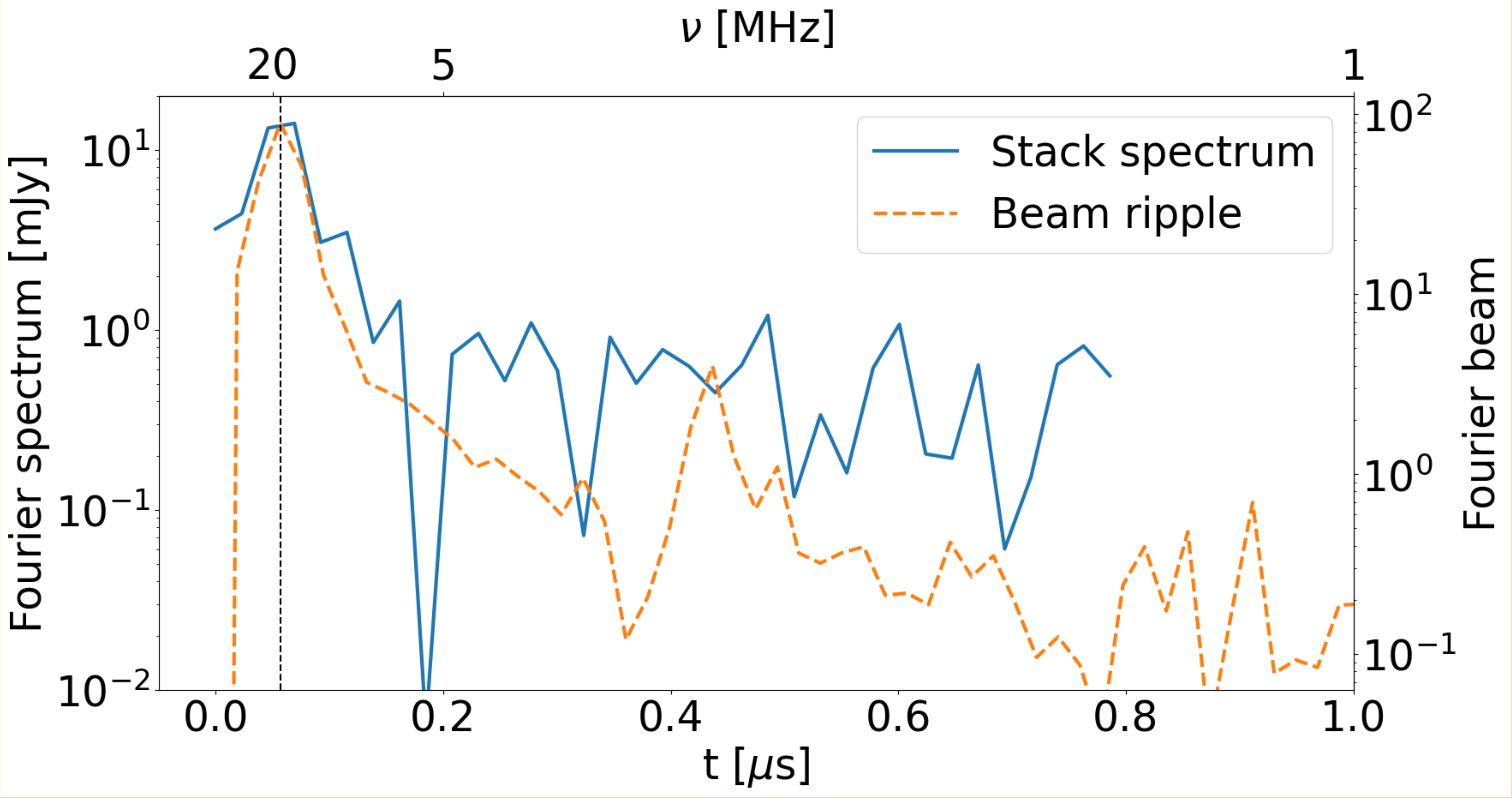
Matshawule et al., 2011.10815

Spinelli et al., 2107.10814



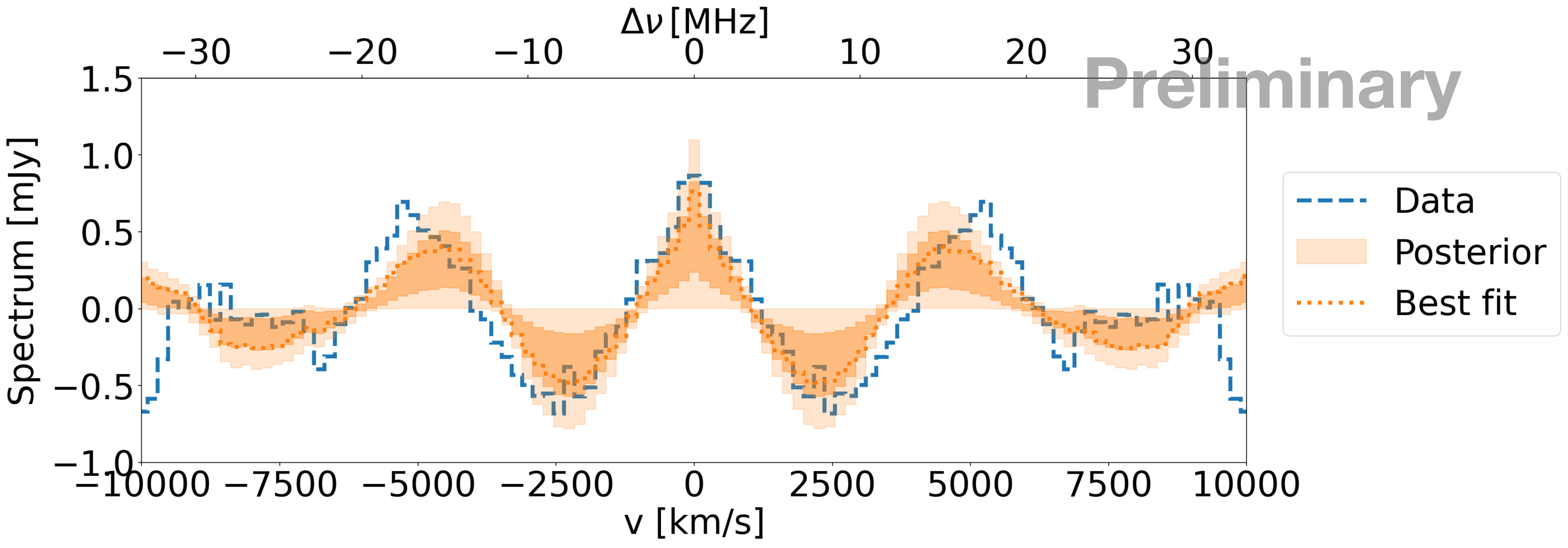
Hunting down the source of systematics

- The Fourier transform of the beam ripple and the Fourier transform of the stacked spectrum match in their peak position
- The systematics in the stacked signal has the same characteristic frequency scale as the beam ripple



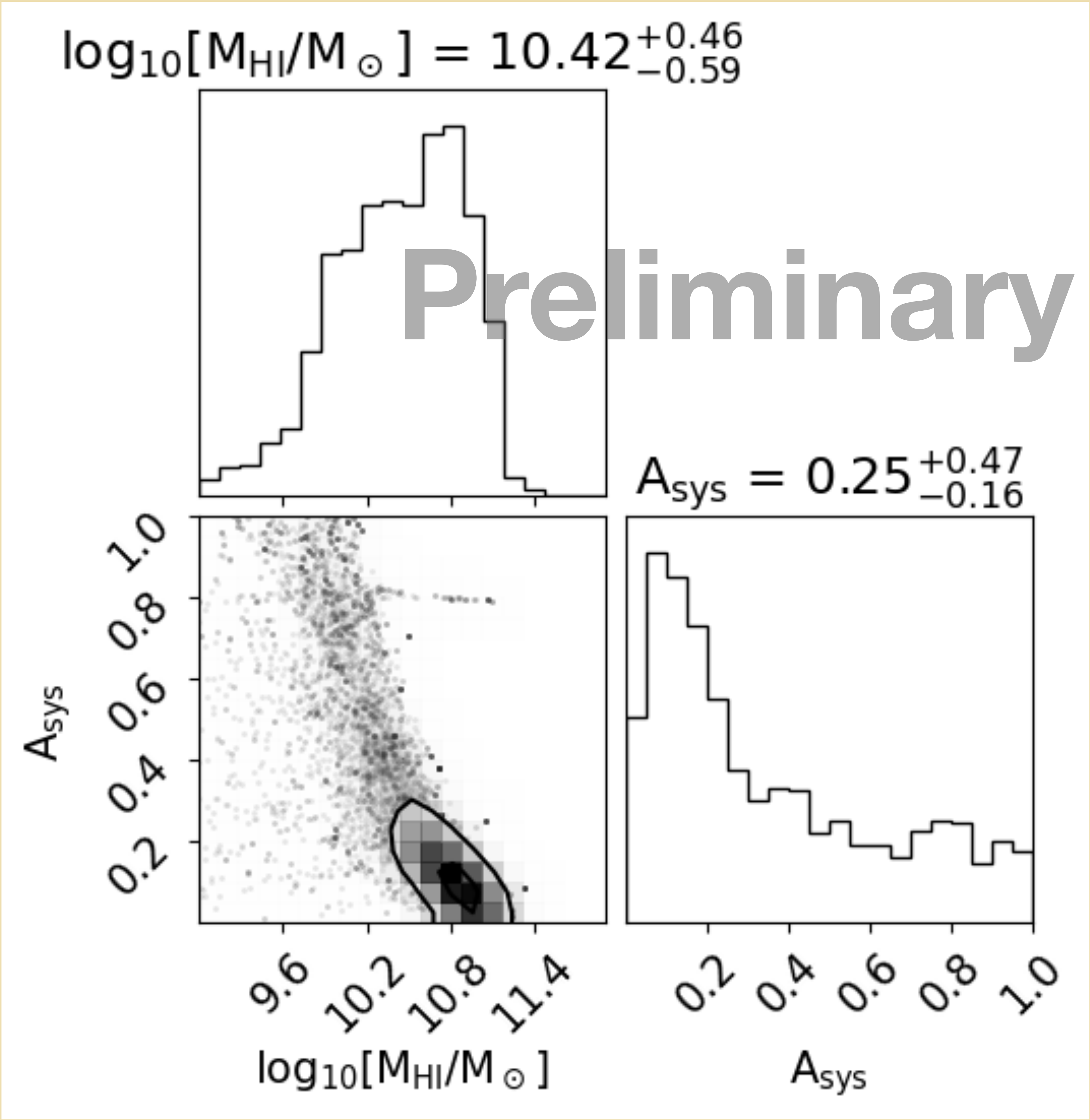
Constraining the systematics

- Simple two-parameter model: amplitude of oscillation and average HI mass of galaxies.
- Fitting suggests a small mismatch of frequency.



Constraining the systematics

- The amplitude of the systematics much larger than the measured beam ripple (A>4% at 95% confidence level).
- Strong degeneracy between the systematics and HI.
- Tentative constraints of HI.
- However, it is likely overestimated due to the incomplete galaxy catalogue and biased due to degeneracy.



Outlook

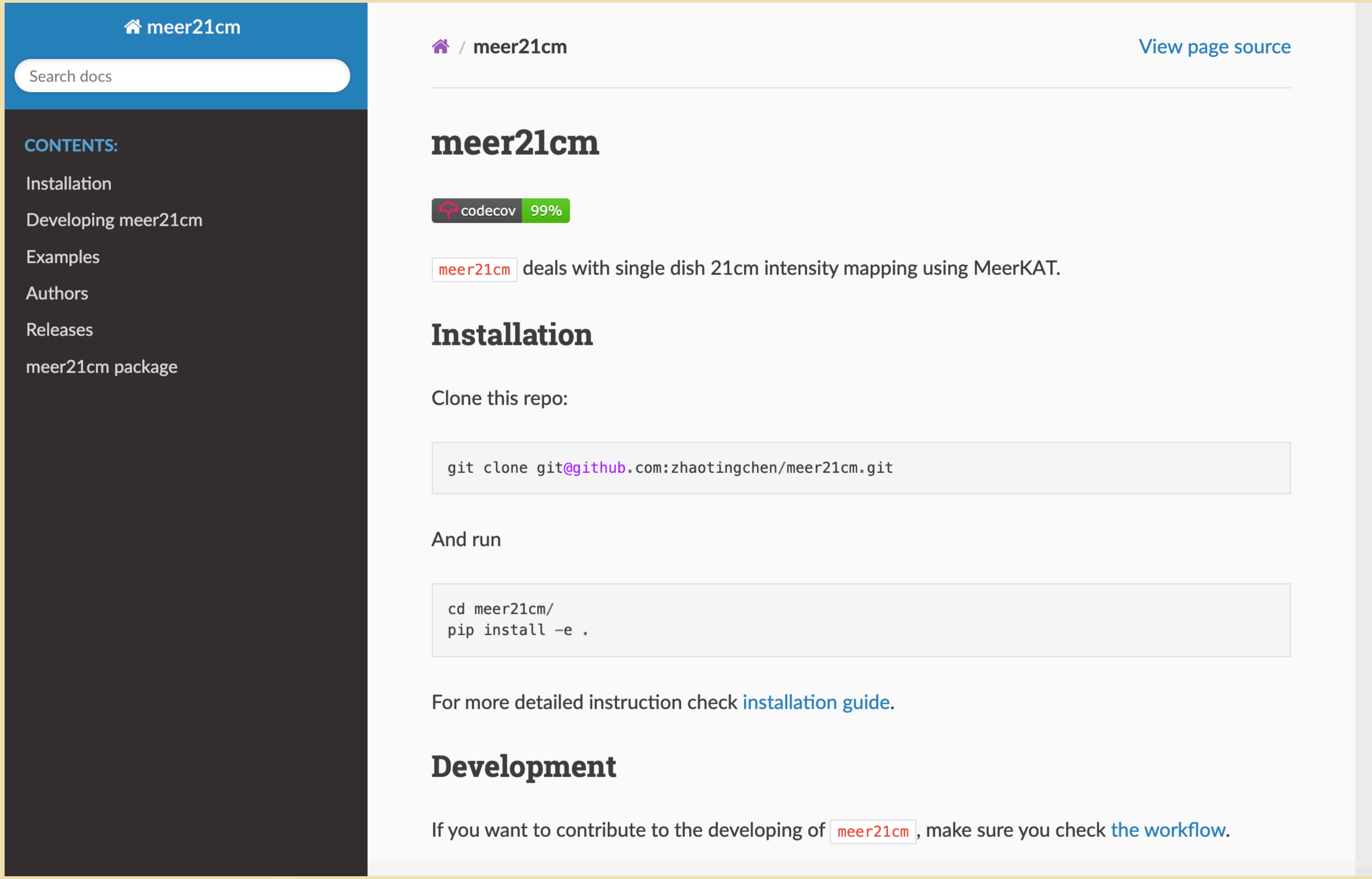
- We can parameterise the beam ripples and incorporate in our forward modelling of the stacked signal
- The stacking measurements can be used to simultaneously marginalise over the HI model and the systematics.
- The fitting results suggest that the systematics is still significant and stands in the way of constraining HI.
- Future MeerKAT UHF and SKAO will fix a few key issues: data quality, overlap with galaxy surveys, signal-to-noise…

Conclusion

- Spectral line stacking using single dish IM survey of MeerKAT/SKA-Mid faces a series of challenges before HI science can be extracted
- The scale of measurements, as well as mixing of systematic and signal, requires forward modelling of the observed spectrum
- Stacking is a very powerful tool for validating detection, analysing systematics, and can be used to simultaneously constrain HI signal and observational systematics
- In the future, MeerKAT and SKA-Mid can be used to measure quantities such as HI density and scaling relations through stacking.

Ad: the meer21cm package

- Used in this work for performing stacking analysis and forward modelling.
- Also working on integrating power spectrum pipeline
  Cunnington & Wolz, 2312.07289
- Fully modulated, easy to use, thoroughly tested and will be documented
- Will be public soon(ish)


meer21cm
Search docs
CONTENTS:
Installation
Developing meer21cm
Examples
Authors
Releases
meer21cm package

/ meer21cm
View page source

meer21cm

codecov 99%

`meer21cm` deals with single dish 21cm intensity mapping using MeerKAT.

Installation

Clone this repo:

```
git clone git@github.com:zhaotingchen/meer21cm.git
```

And run

```
cd meer21cm/
pip install -e .
```

For more detailed instruction check installation guide.

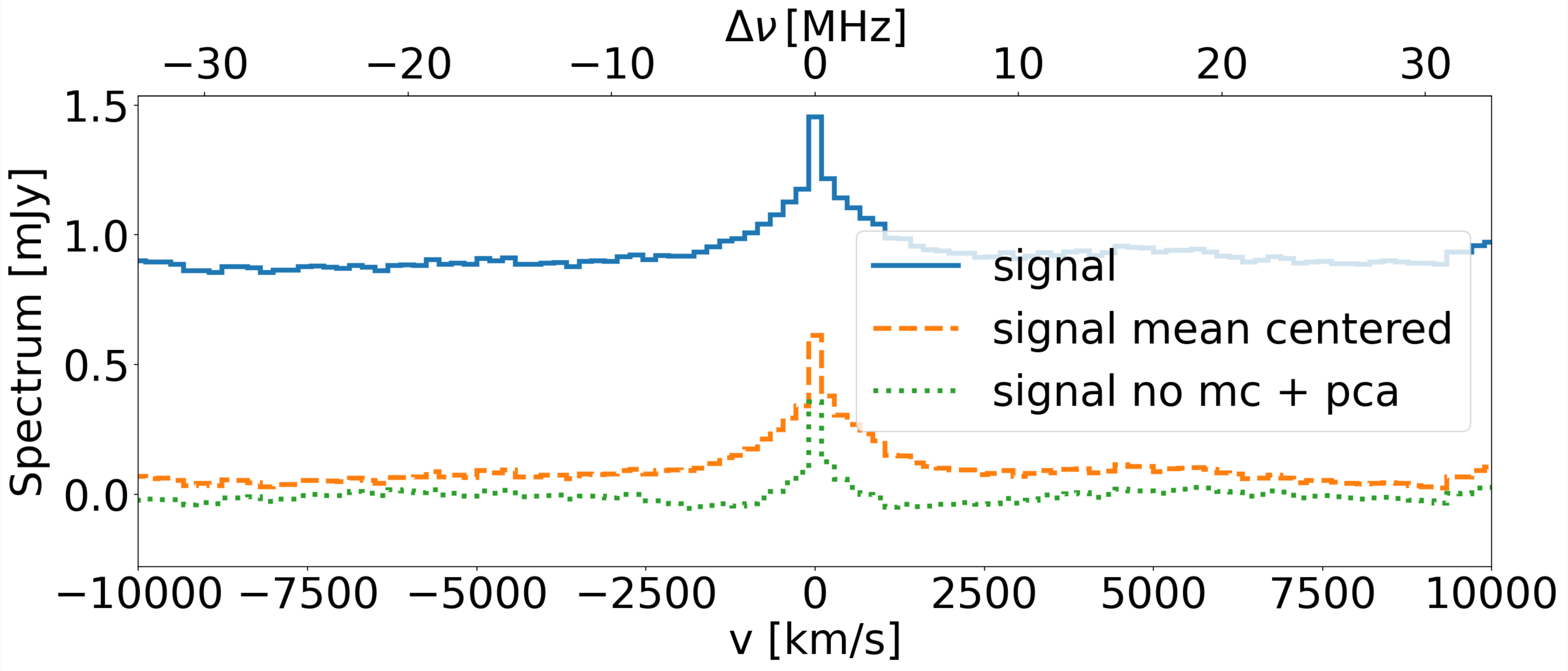
Development

If you want to contribute to the developing of `meer21cm`, make sure you check the workflow.

Thanks!

Backup: Why PCA removes plateau

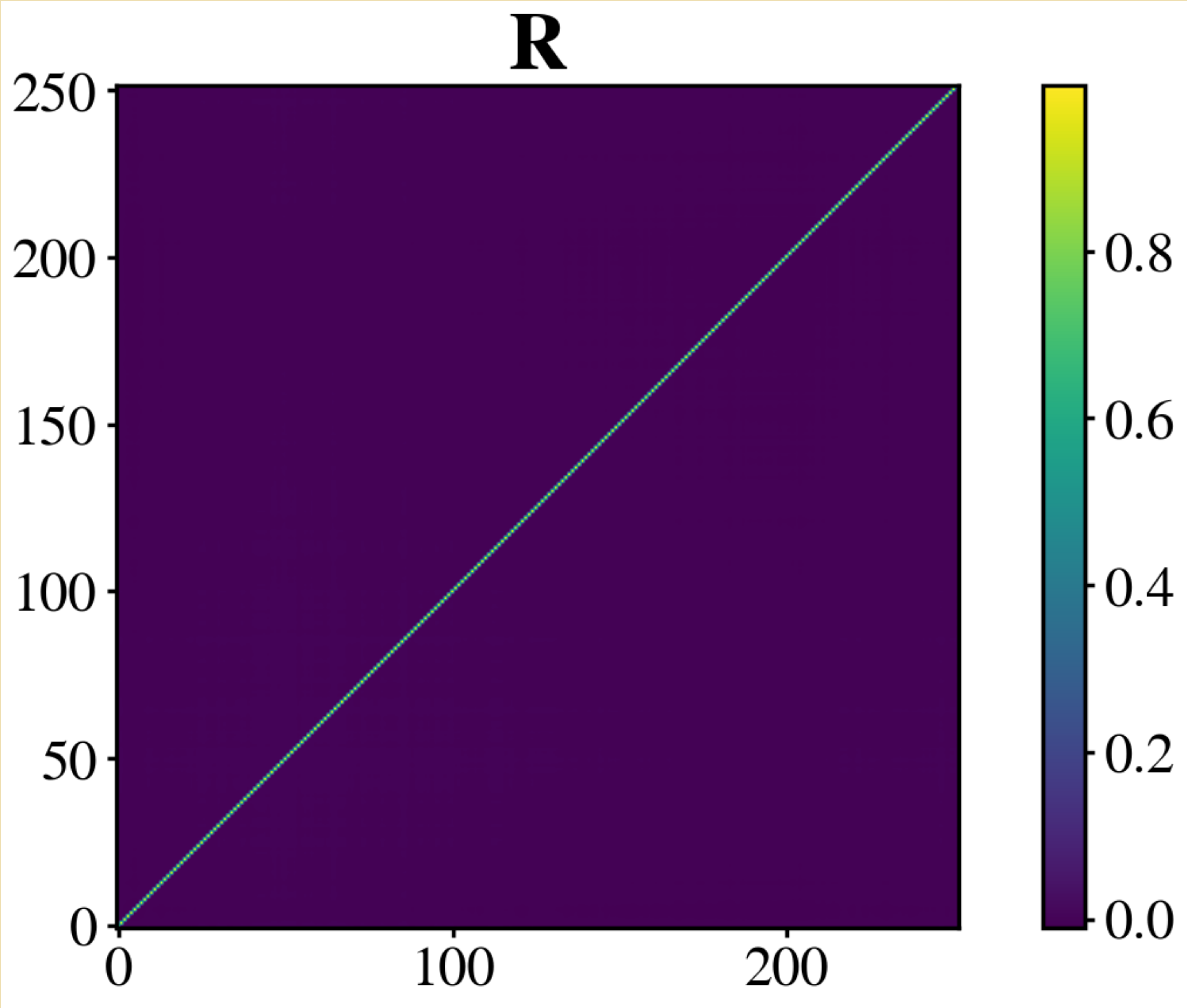
- It is naturally to then conclude that PCA removes the mean so the plateau is gone. But actually even without mean centering:

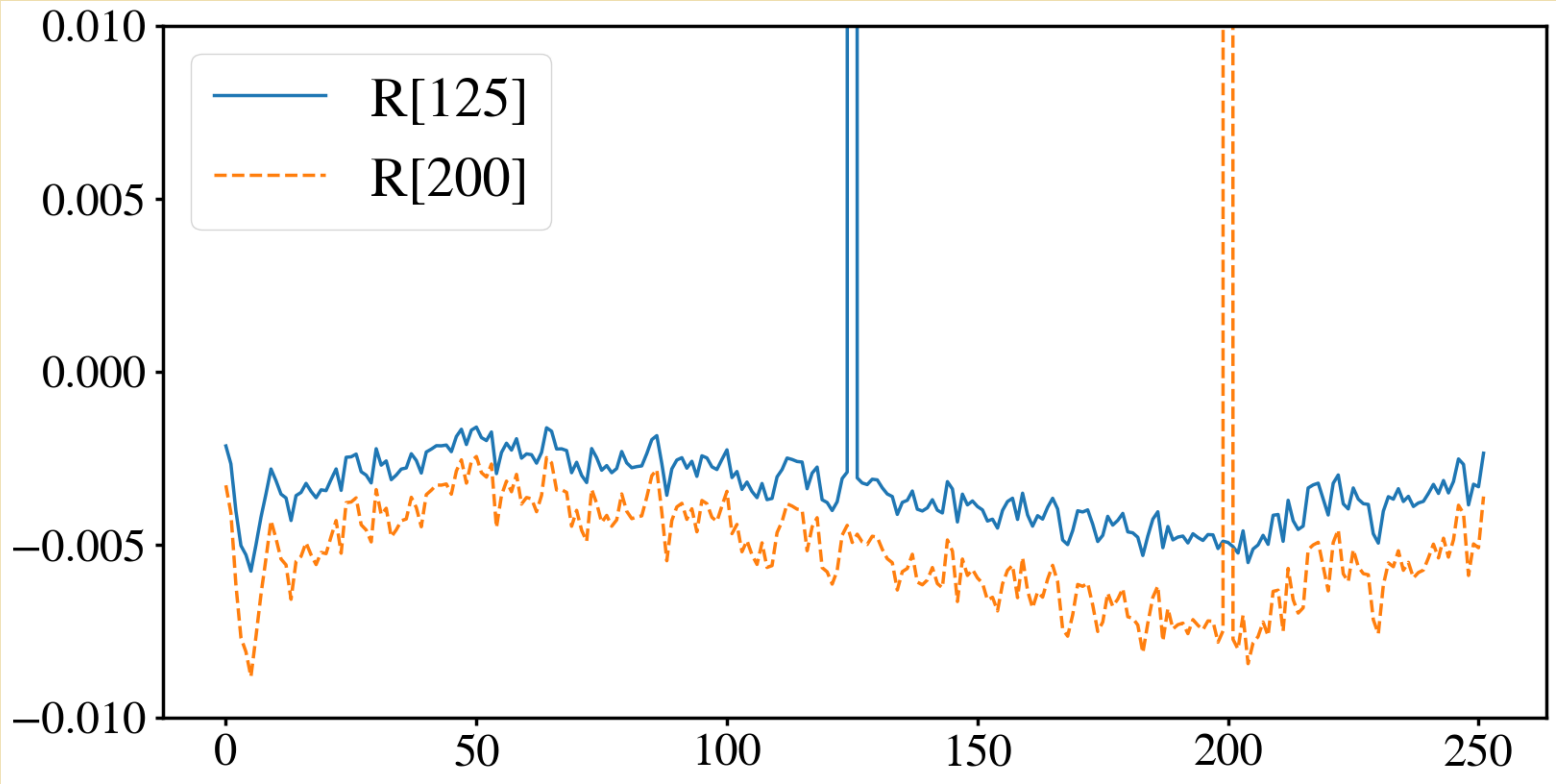


Backup: Why PCA removes plateau

- What PCA does is that it preserves the peak, but also scatters the peak amplitude into negative values across the frequencies:

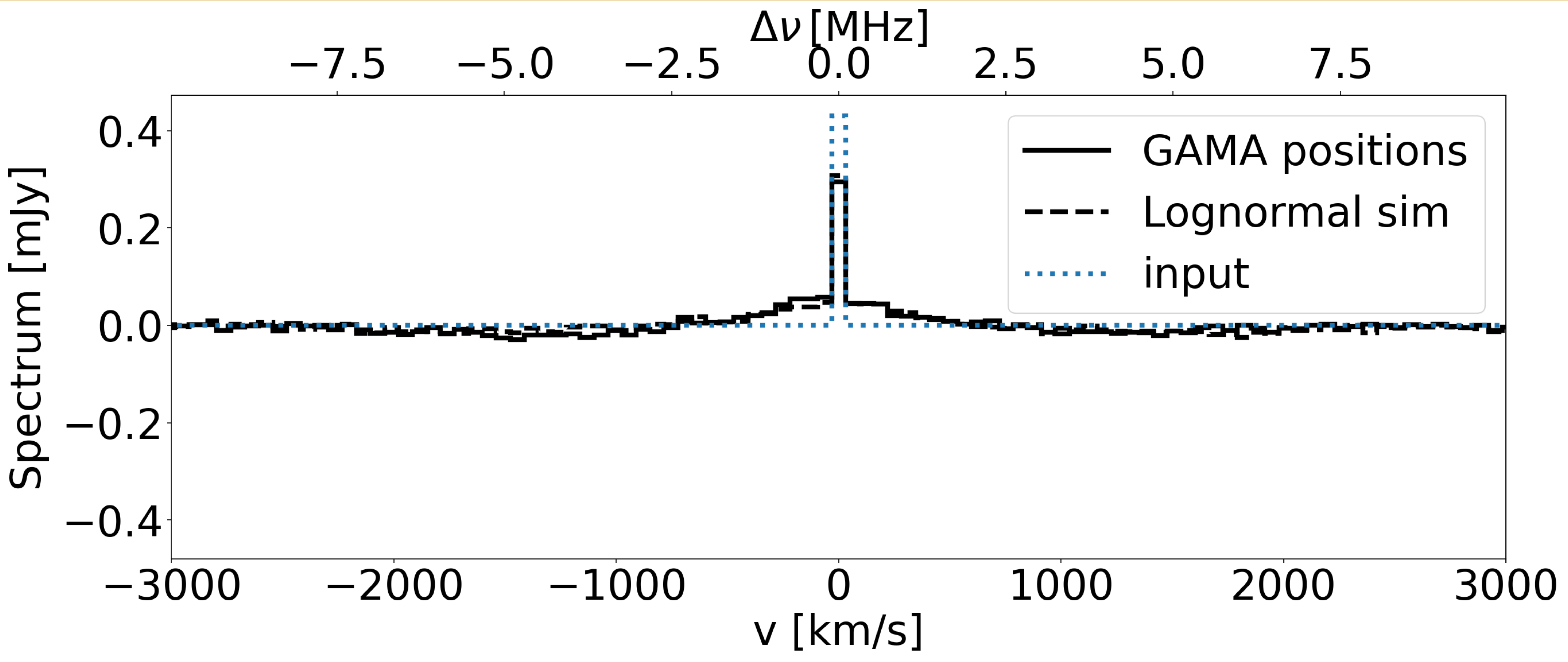
$$\vec{r} = \mathbf{R}\vec{m} = (\mathbf{I} - \mathbf{A}\mathbf{A}^{\mathrm{T}})\vec{m}$$





Backup: matching the clustering amplitude

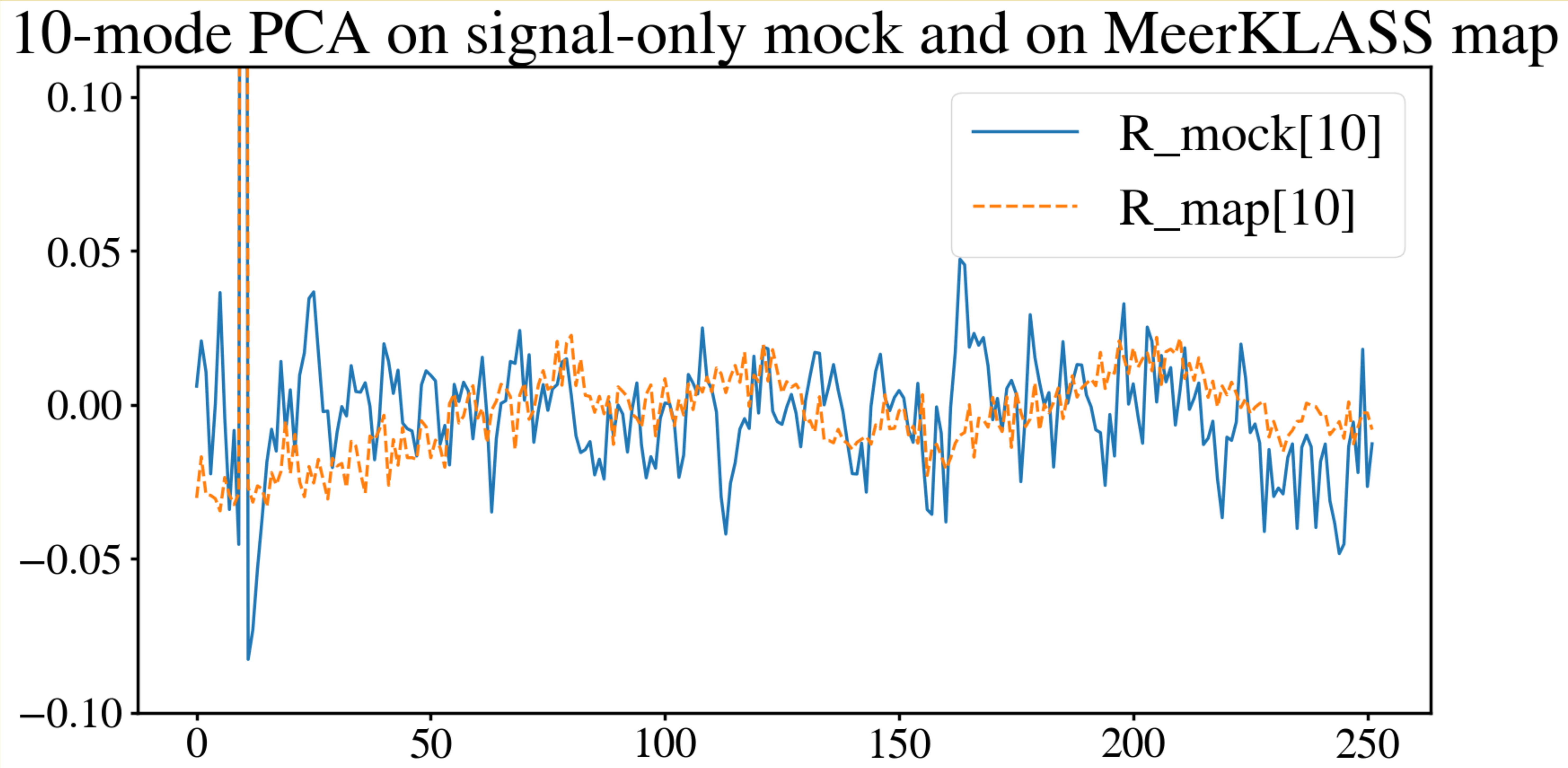
- Simulations using lognormal galaxy mock positions and actual GAMA galaxy positions match quite well.



Backup: correlation comes from PCA

- The thermal noise in the map is not correlated originally.
- The correlation also comes from the fact that PCA matrix is applied to the data
- The oscillation structure is visible in the PCA matrix.

$$\vec{r} = \mathbf{R}\vec{m} = (\mathbf{I} - \mathbf{A}\mathbf{A}^{\mathrm{T}})\vec{m}$$



Backup: parameter fitting

- Need careful treatment of priors for modelling the systematics