Forward Modeling HIRAX Observations Using Fast Halo-Model-Based Simulations of Cosmological HI

Pascal Hitz
ETHZ Cosmology Group*

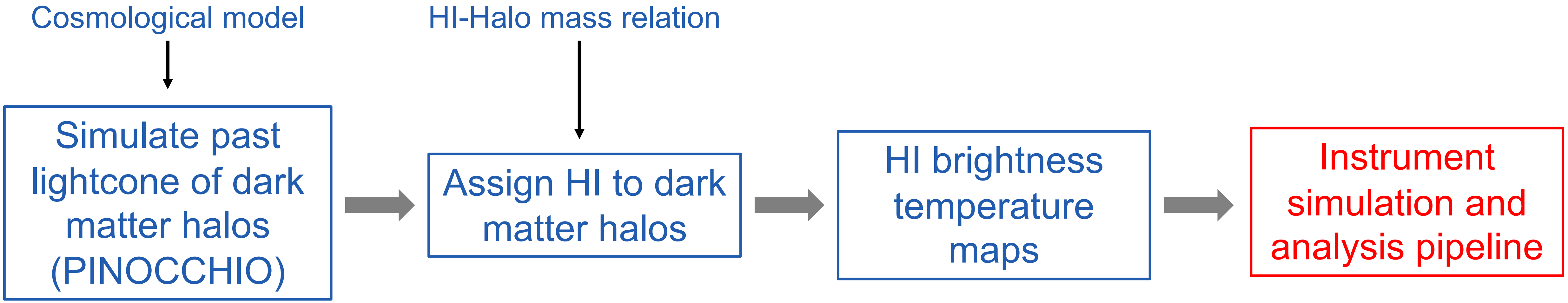
Cosmology in the Alps 2026, Les Diablerets, 19.03.2026

*Alexandre Refregier, Devin Crichton, John Hennig

Image Credit: SARAO, Heywood et al. (2022) / J.C. Muñoz-Mateos

Overview

- Fast and large volume simulations of the neutral hydrogen (HI) distribution for post-reionization ($z \lesssim 6$) 21 cm intensity mapping experiments
  - Assumption: Most HI resides inside of dark matter halos
- Test instrument simulation and analysis pipeline of HIRAX



Sky signal simulation
Application to telescope

Hitz et al. 2025

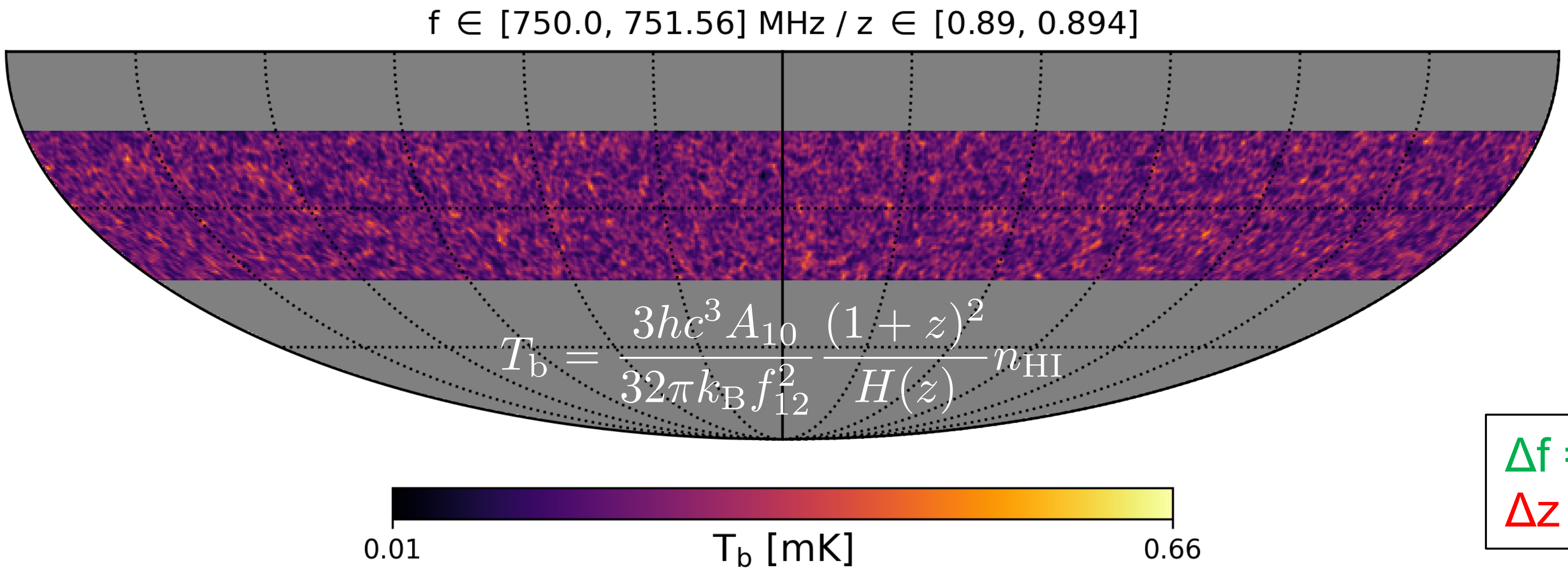
https://arxiv.org/abs/2410.01694

Sky Signal Simulation

- PINOCCHIO Simulation (Monaco et al. 2002, 2013, Taffoni et al. 2002, Munari et al. 2017)
  - 1 Gpc/h box size
  - 6700^3 simulation particles
  - ≥ 10 particles per halo ↔ ≥ 4.3×10^9 $M_\odot$
  - Lightcone settings:
    - Frequency range: 700 – 800 MHz ↔ Redshift 0.77 – 1.03
    - Declinations between −15° and −45°
    - 40 box replications

  → 2 – 3% HI mass missing (1 Gpc/h box size; 6700^3 simulation particles; ≥ 10 particles per halo)

- HI-Halo mass relation (Padmanabhan et al. 2017)



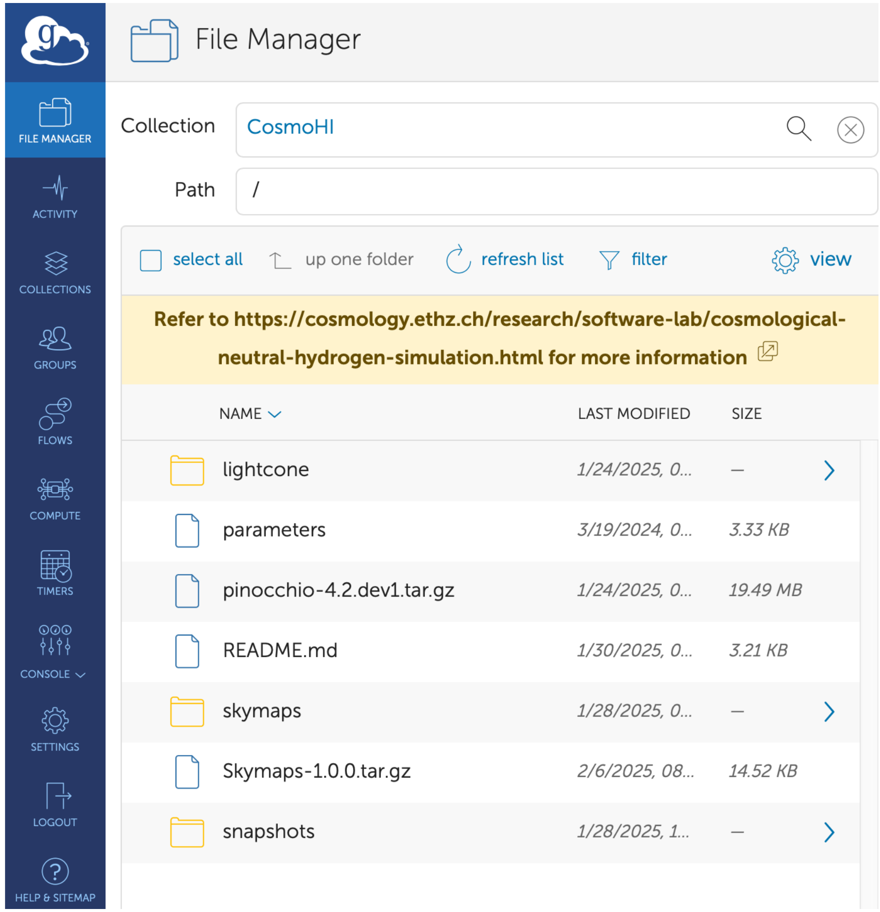
$$T_{\mathrm{b}} = \frac{3hc^3 A_{10}}{32\pi k_{\mathrm{B}} f_{12}^2} \frac{(1+z)^2}{H(z)} n_{\mathrm{HI}}$$

Δf = 1.56 MHz
Δz ~ 0.004

Data Availability

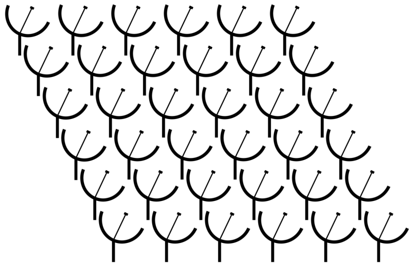
- Data documentation on our group website https://cosmology.ethz.ch/research/software-lab/data-products/cosmological-neutral-hydrogen-simulation.html

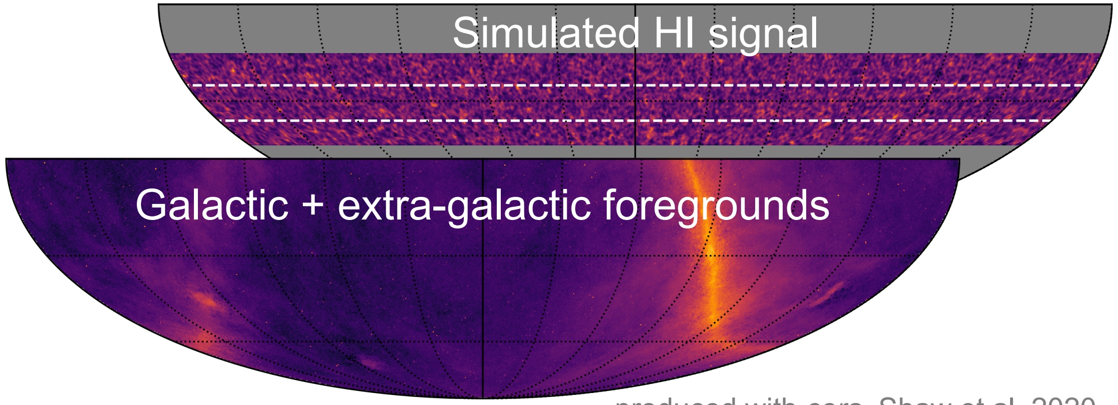
- Fast data transfer via Globus network
  - **./snapshots:** DM halo box snapshot catalogues at z = 0.8, 0.9 and 1
  - **./lightcone:** DM halo past light cone catalogue from z = 0.77 – 1.03
  - **./skymaps:** DM mass and HI mass / brightness temperature HEALPix maps



Simulated HIRAX Observation

Scaled HIRAX array configuration	
Number of dishes:	36 (6 × 6 grid)
Operating mode:	Drift-scan (single pointing)
Dish diameter:	6 m
Dish separation:	6.5 m EW, 8.5 m NS
Primary Beam:	Gaussian
Latitude:	−30.7°
Longitude:	+21.6°





produced with *cora,* Shaw et al. 2020

Construct instrument model

Generate synthetic data: Visibilities

Add Gaussian noise

Foreground filtering

Generate ringmaps

m-mode Formalism Shaw et al. 2015

T: Sky signal ***B$_{ij}$^{T}***: Beam transfer functions (Instrument model) ***n$_{ij}$***: Noise

$$V_{ij,}(\varphi,\nu) = \langle F_i F_j^* \rangle = \int \left[B_{ij}^T(\hat{\boldsymbol{n}};\varphi,\nu) T(\hat{\boldsymbol{n}},\nu) + Polarisations \right] d^2\hat{n} + n_{ij}(\varphi,\nu)$$

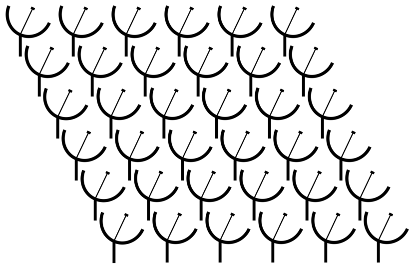
Spherical harmonics transform

$$V_{ij}(\varphi,\nu) = \sum_{l,m} [B_{ij;lm}^T(\varphi) a_{lm}^T(\nu) + Polarisations] + n_{ij}(\varphi,\nu)$$

$$V_{ij;m} \equiv \int \frac{d\varphi}{2\pi} V_{ij}(\varphi,\nu) e^{-im\varphi}$$

m-mode $$V_{ij;m}(\nu) = \sum_{l} [B_{ij;lm}^T(\nu) a_{lm}^T(\nu) + Polarisations] + n_{ij;m}(\nu)$$

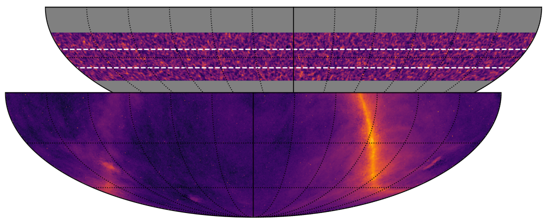
Simulated HIRAX Observation



$B^T_{ij;lm}$

$V_{ij;m} = \sum_l [B^T_{ij;lm} a^T_{lm}] + n_{ij;m}$

a^T_{lm}



Inverse Fourier transform in φ-direction

$V_{ij}(\varphi, \nu)$

Fourier transform to delay space τ

$\tilde{V}_{ij}(\varphi, \tau)$

High-pass delay filter

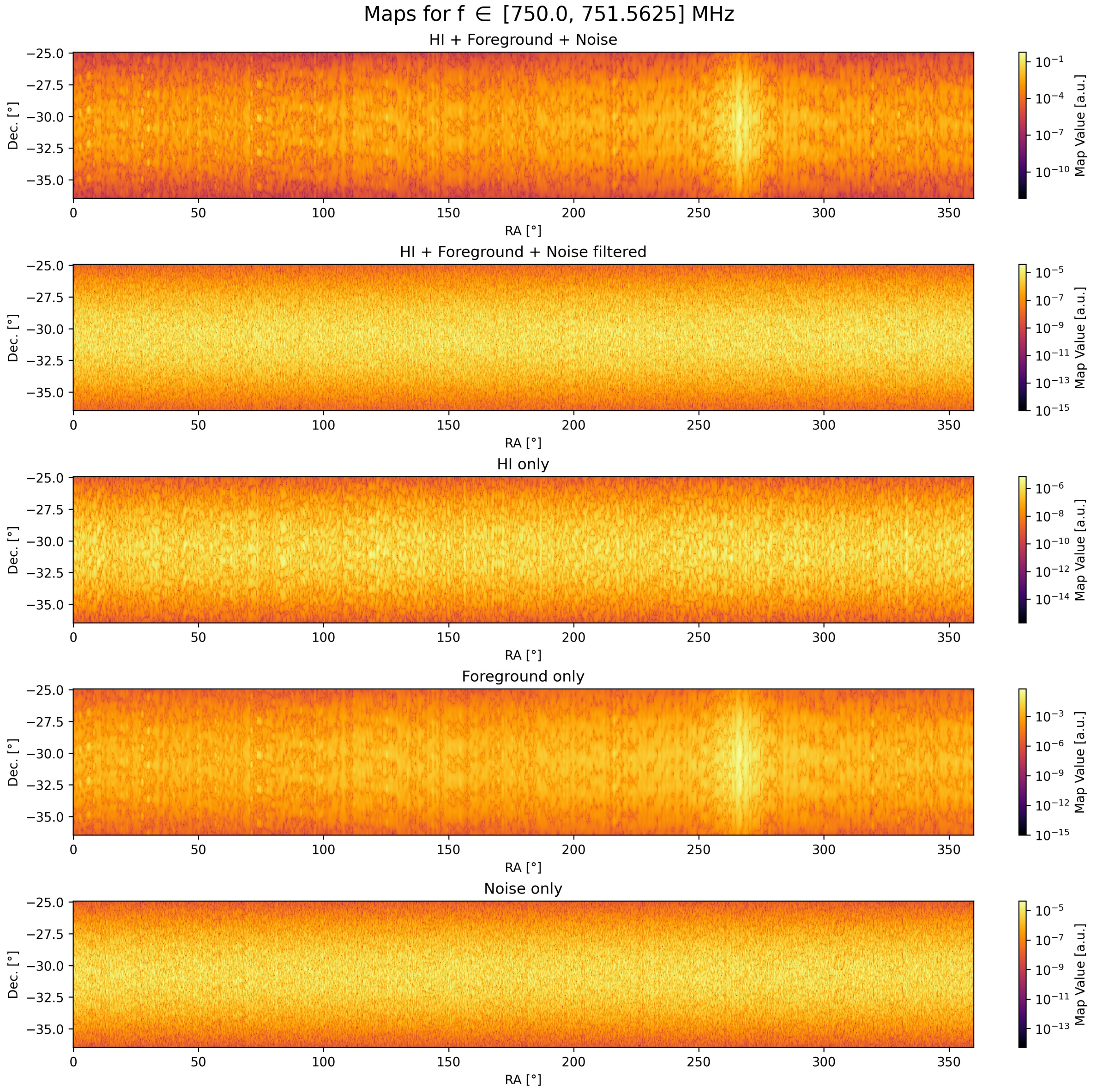
NS + EW Beamforming

Ringmaps

draco, Shaw et al. 2020

driftscan, Shaw et al. 2020

Simulated HIRAX Observation: Ringmaps



HI + Foregrounds + Noise (120 day observation)

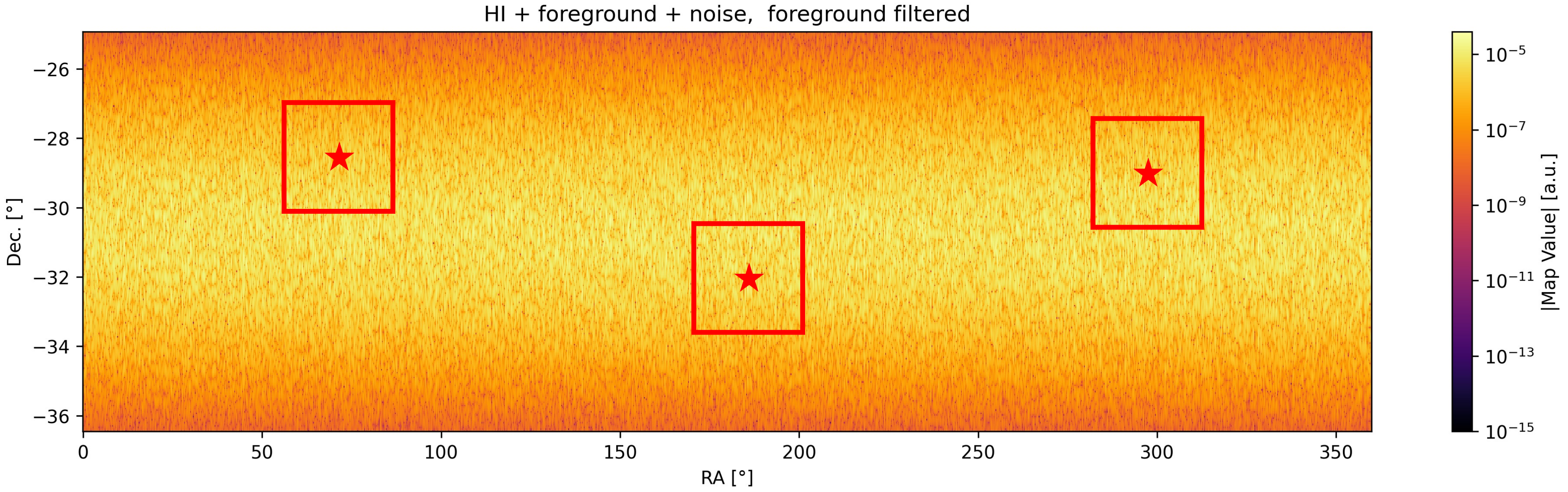
HI + Foregrounds + Noise, delay filtered (τ < 12 ns)

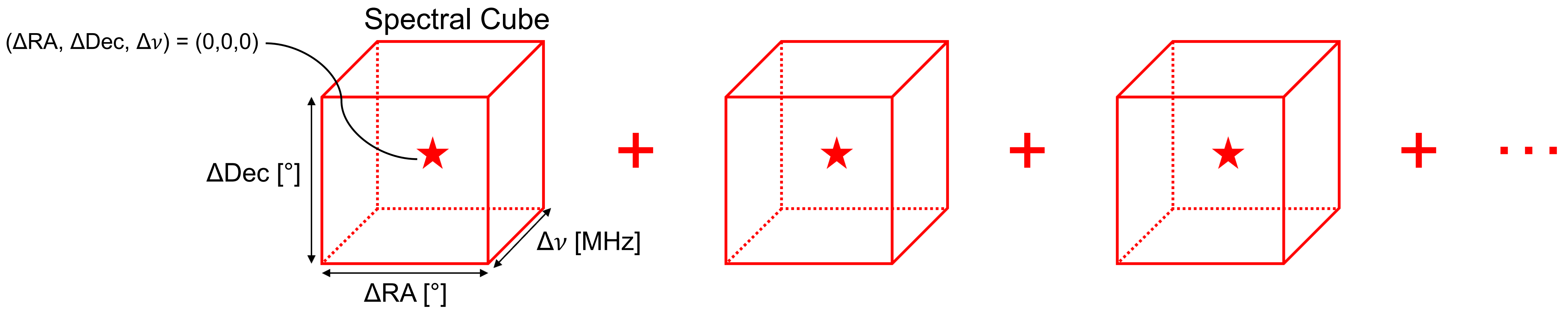
HI only

Foreground only

Noise only (corresponding to 120 days of observation)

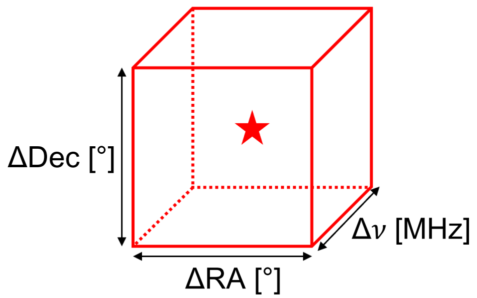
Stacking Signal on Halo Positions



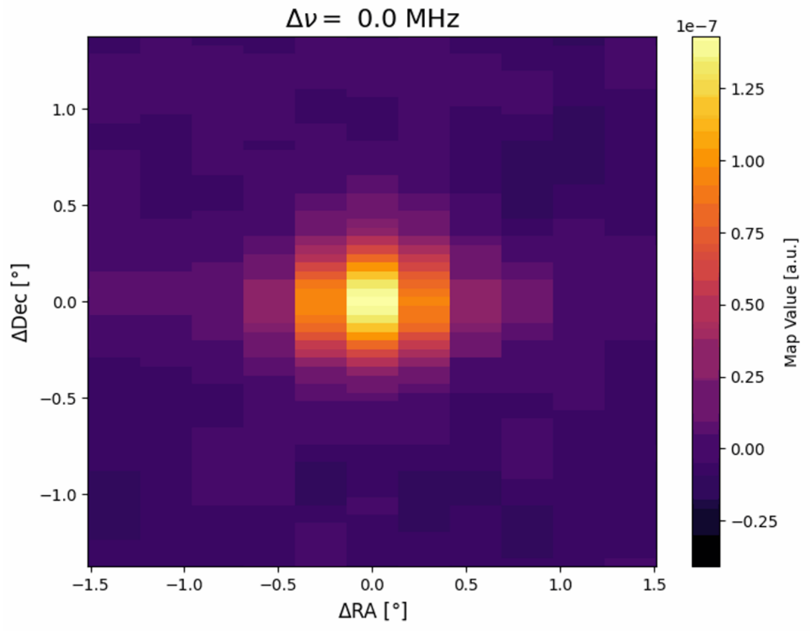


Stacking Signal

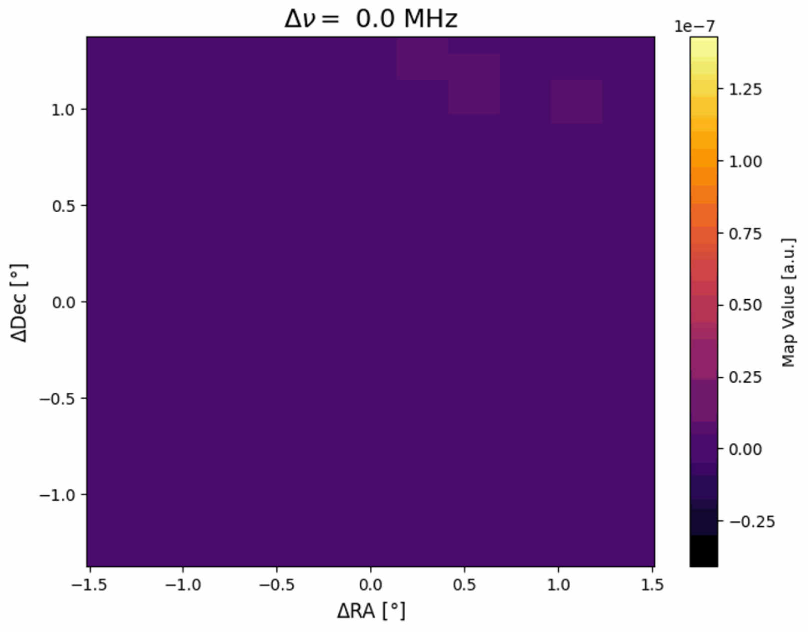
Stacking on $6{\times}10^6$ halos with M > $2{\times}10^{12}$ $M_\odot$

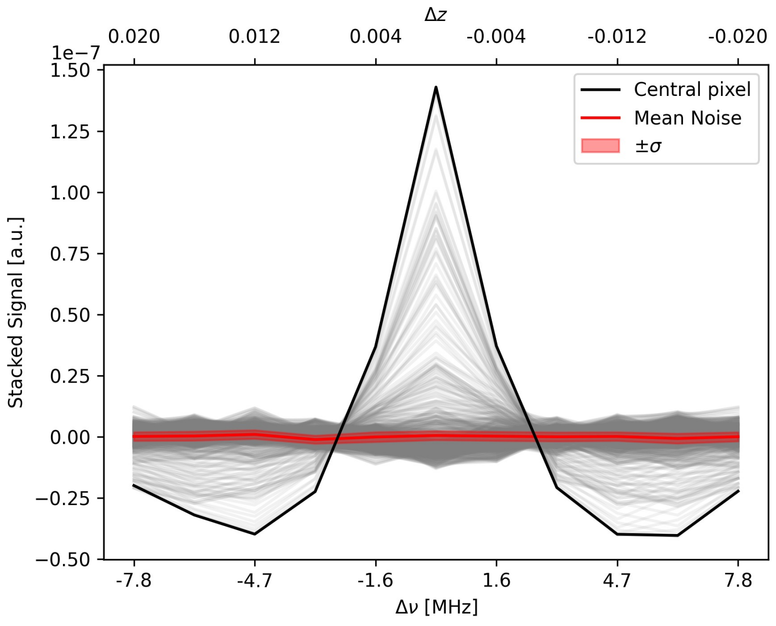


Stacking on halo positions



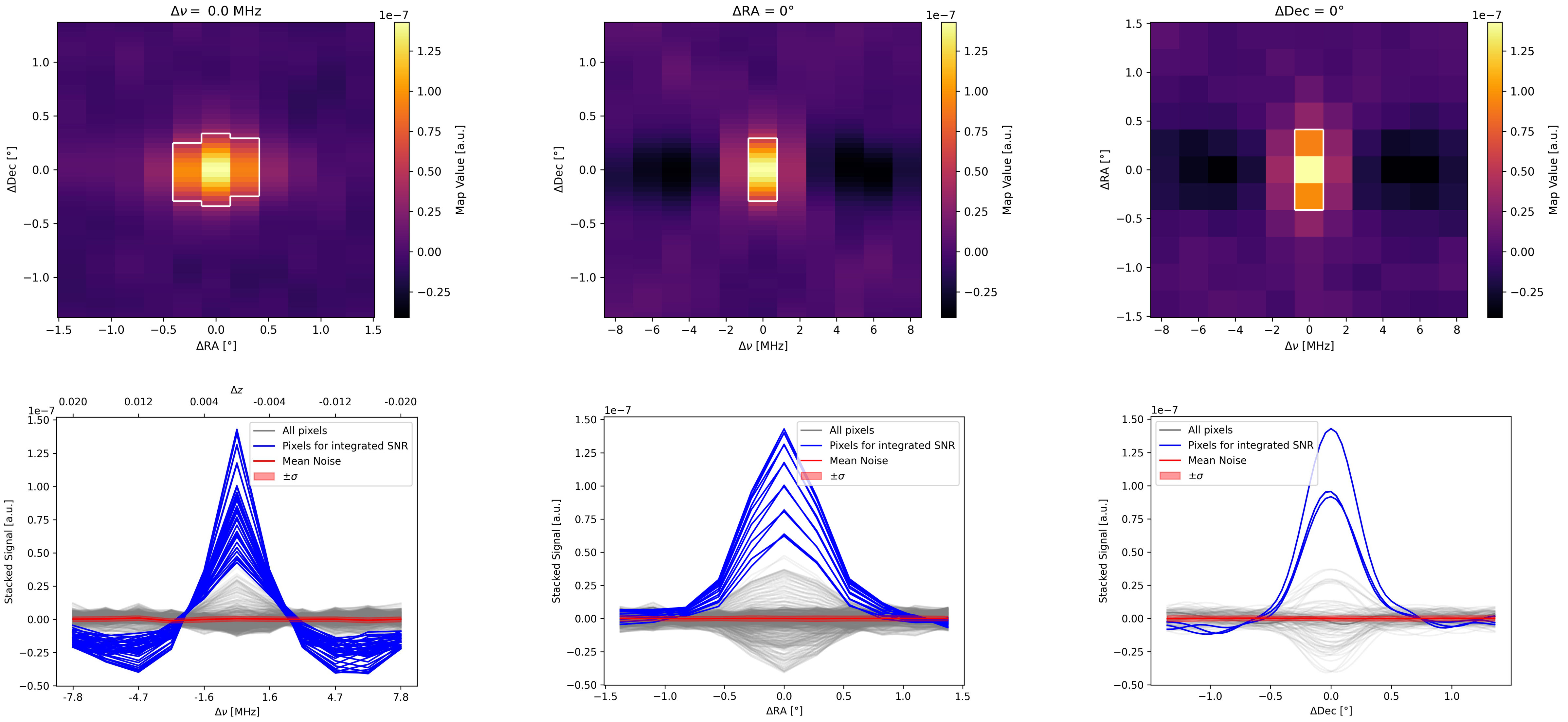
Stacking on random positions
(used for noise estimation)





Stacking Signal

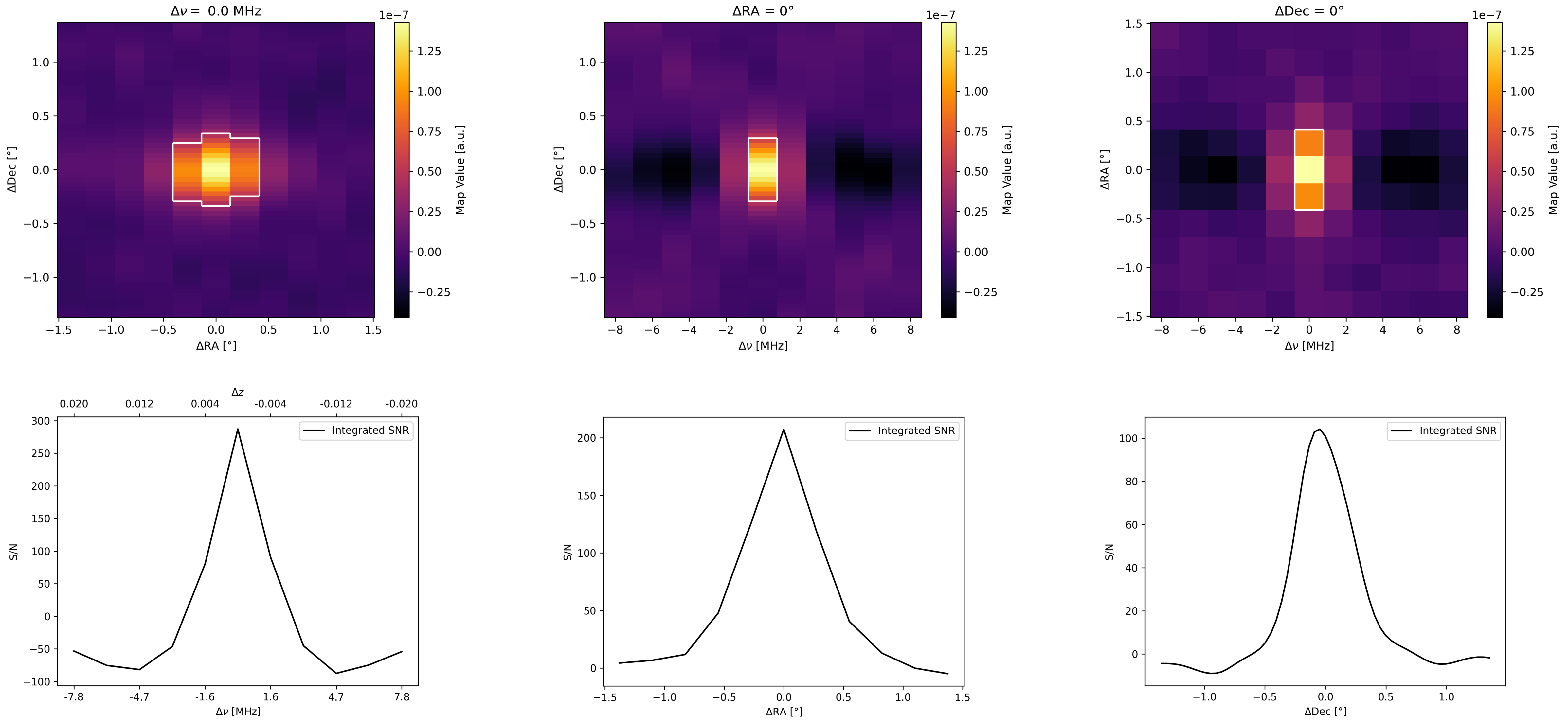
Stacking on 6×10^6 halos with $M > 2\times10^{12}\ M_\odot$



ETH zürich

Stacking Signal SNR

Stacking on $6{\times}10^6$ halos with $M > 2{\times}10^{12}\ M_\odot$

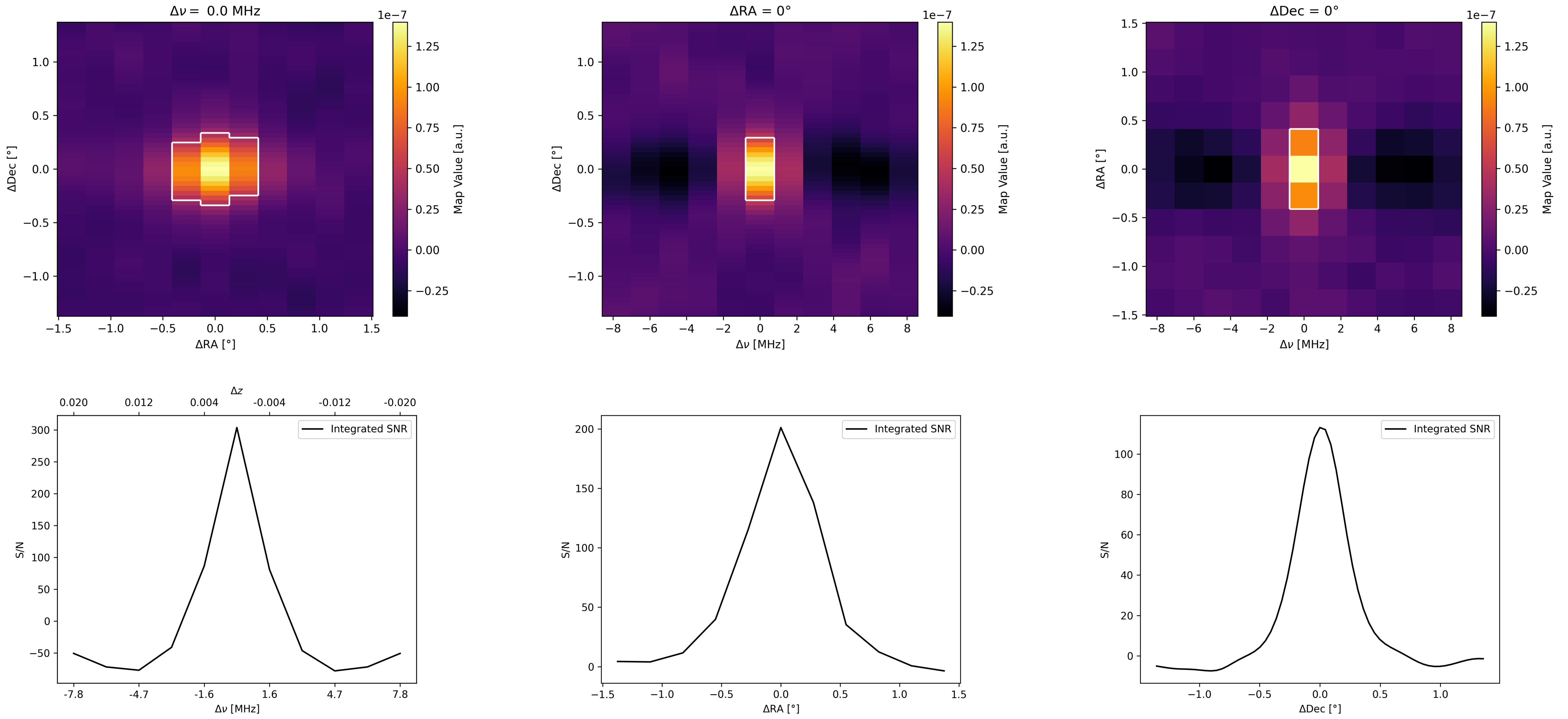


ETH zürich

Stacking Signal SNR

Stacking on 6×10^{6} halos with $M > 2\times10^{12}\ M_{\odot}$

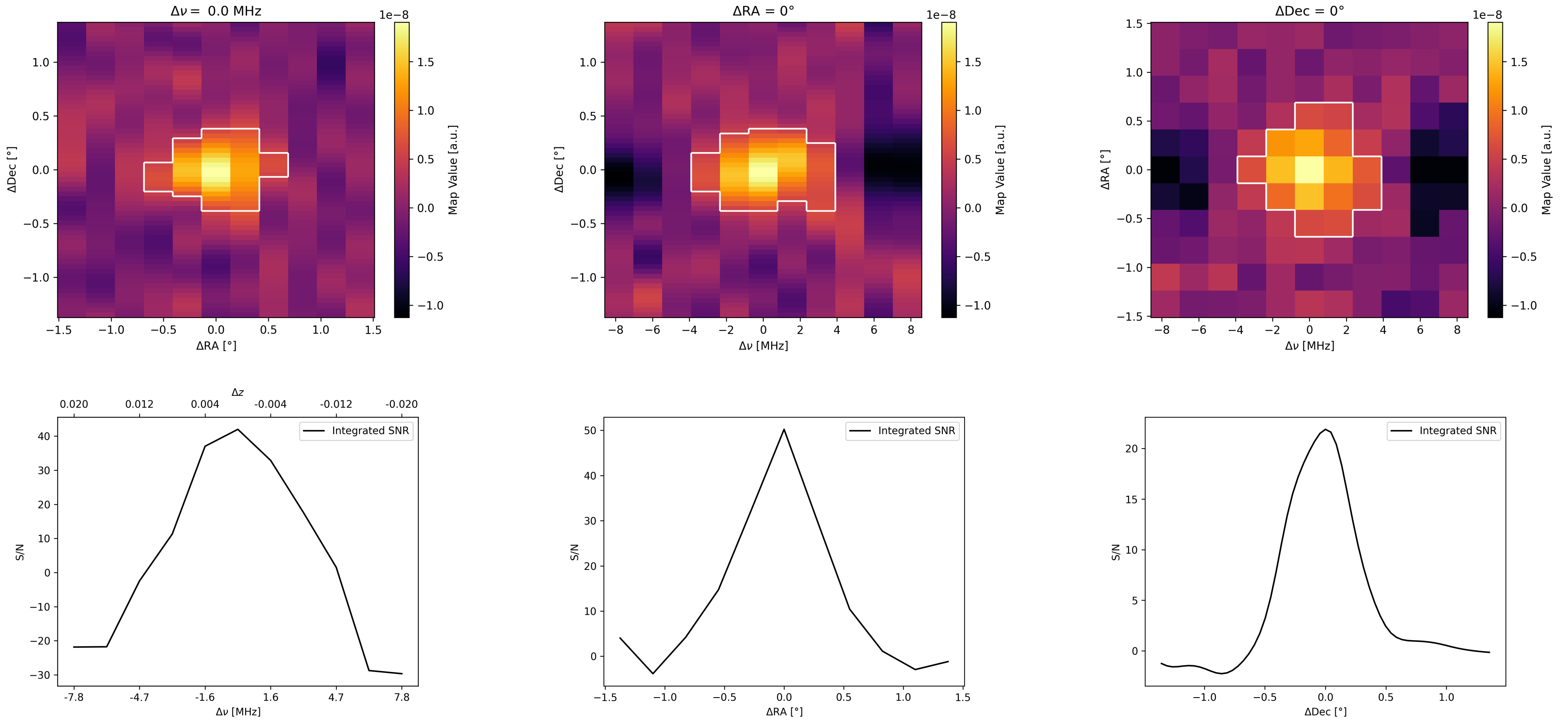
Redshift error: Gaussian with σ=0.0005



Stacking Signal SNR

Stacking on $6{\times}10^6$ halos with $M > 2{\times}10^{12}\,M_\odot$

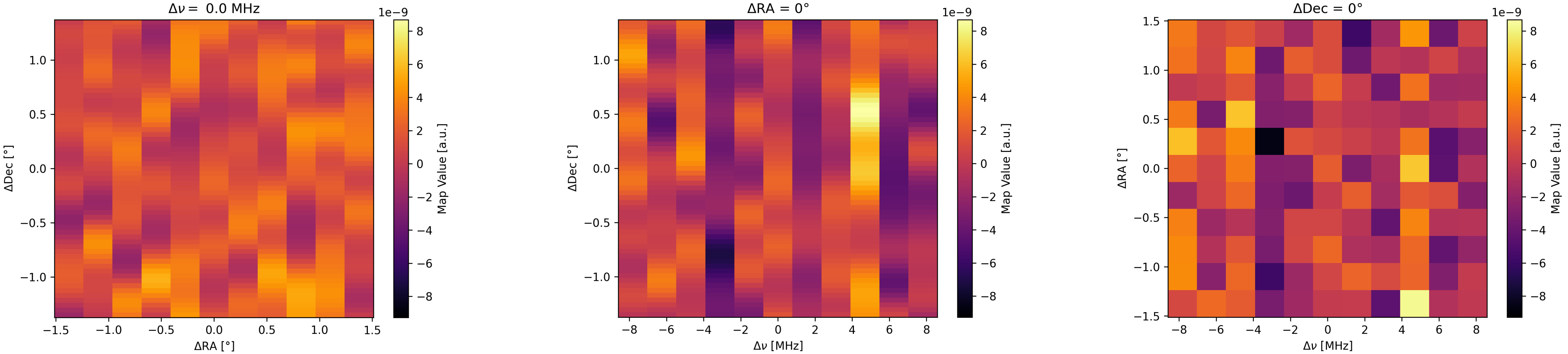
Redshift error: Gaussian with σ=0.01



Stacking Signal SNR

Stacking on 6×10^6 halos with $M > 2\times10^{12}\,M_\odot$

Redshift error: Gaussian with σ=0.05



Summary

Hitz et al. 2025

https://arxiv.org/abs/2410.01694

- Simulation pipeline of HI sky signal for intensity mapping
  - Data publicly available: https://cosmology.ethz.ch/research/software-lab/data-products/cosmological-neutral-hydrogen-simulation.html
  - One-point statistics and power spectrum in good agreement with theoretical expectations from halo model (available in PyCosmo 2.2.0 https://cosmology.ethz.ch/research/software-lab/scientific-software/PyCosmo.html)

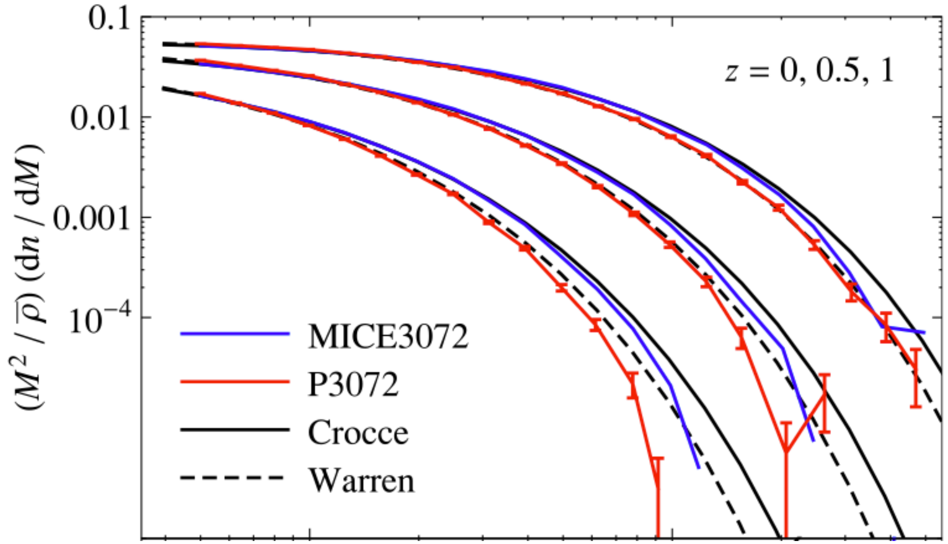
- Simulated HIRAX observation with 6×6 configuration
  - Recovered HI signal from stacking on halo positions

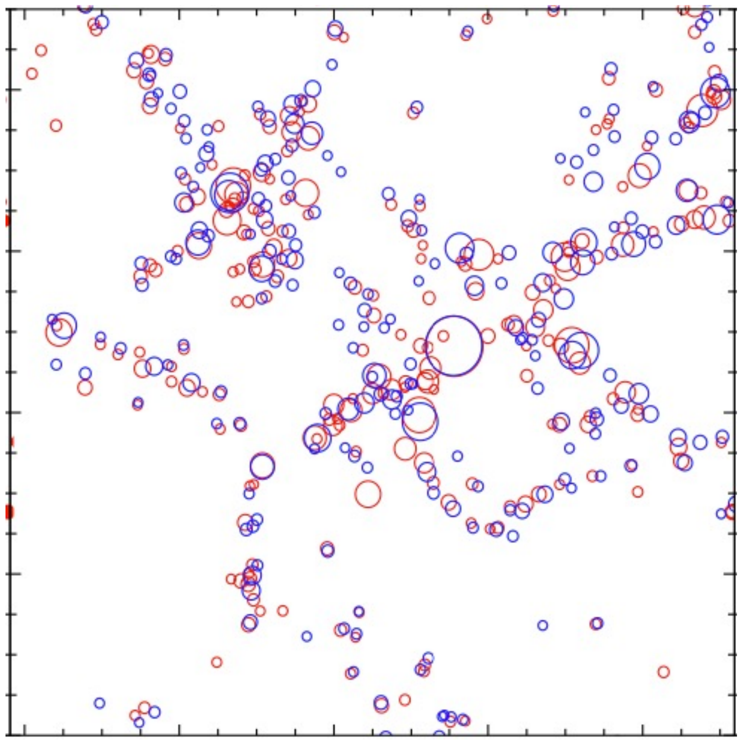
- Ongoing work
  - Increase number of halos in stacking and optimize weighting
  - Stacking on realistic simulated galaxy catalogs
  - Cross-correlation with other cosmological probes from the CosmoGrid simulation suite

Backup Slides

PINOCCHIO: Dark Matter Halo Simulation

- Monaco et al. (2002, 2013), Taffoni et al. (2002), Munari et al. (2017)
- Lagrangian Perturbation Theory
- Collapsed points grouped into halos, hierarchical growth
- Catalogue of dark matter halos
- Much faster than N-body





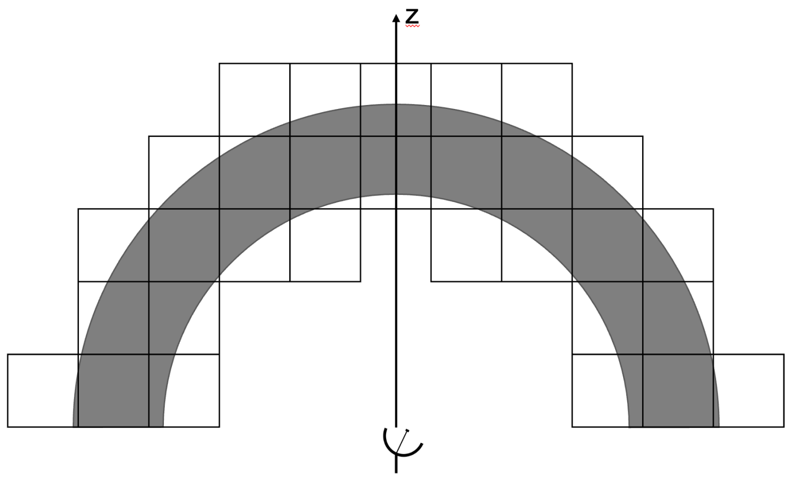
Monaco et al. 2013

Current Setting of DM Simulations

- 1 Gpc/h box size
- 6700^3 simulation particles
- ≥ 10 particles per halo ↔ ≥ $4.3 \times 10^9\ M_{\odot}$

→ 2 – 3% HI mass missing

- Lightcone settings:
  - Frequency range: 700 – 800 MHz ↔ Redshift 0.77 – 1.03
  - Declinations between −15° and −45°
  - 40 box replications



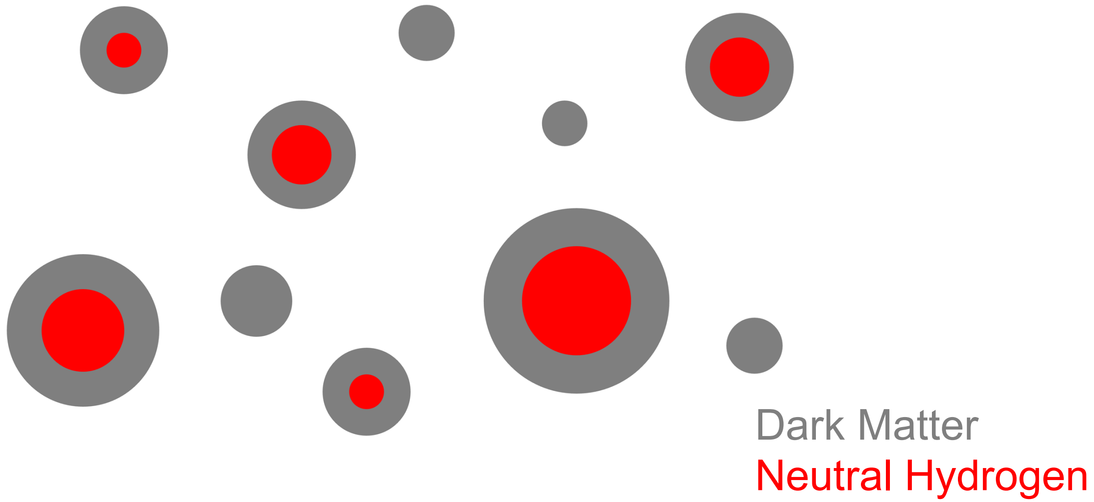
- Ran on Piz Daint with MPI parallelization
  - 2400 nodes with 12 cores each
  - 150 TB RAM, 40’000 CPU h runtime

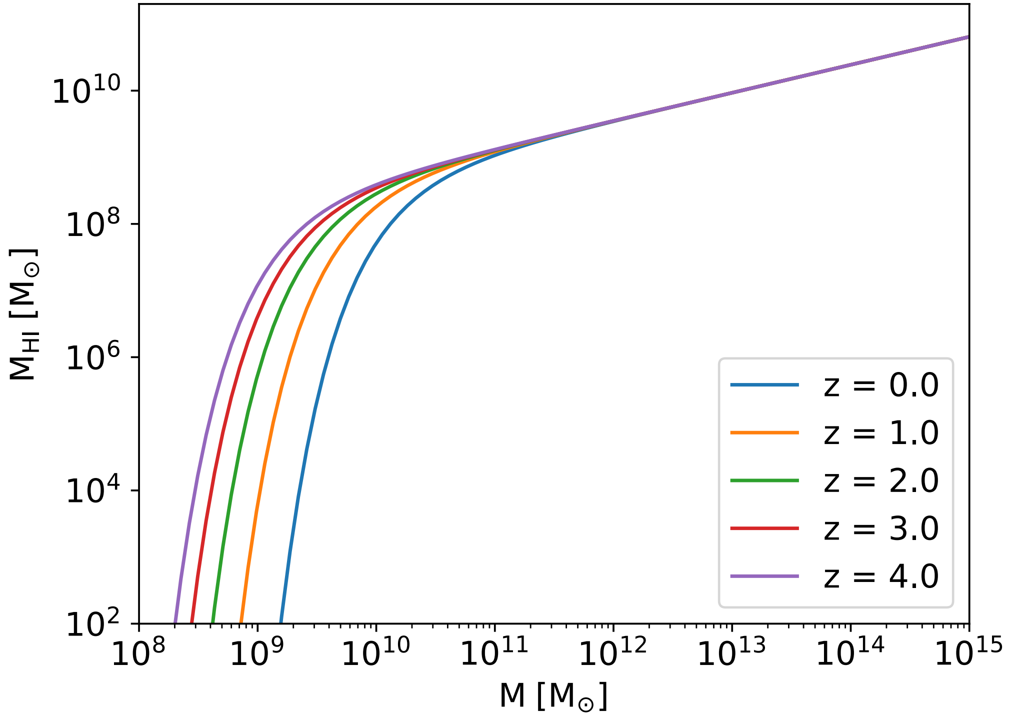
Halo Model for Cosmological HI

HI-halo mass relation fitted to observations:

$$M_{\mathrm{HI}}(M,z) = \alpha f_{\mathrm{H,c}} M \left(\frac{M}{10^{11} h^{-1} \mathrm{M}_\odot}\right)^{\beta} \exp\left[-\left(\frac{v_{\mathrm{c},0}}{v_{\mathrm{c}}(M,z)}\right)^{3}\right]$$

Padmanabhan et al. 2017

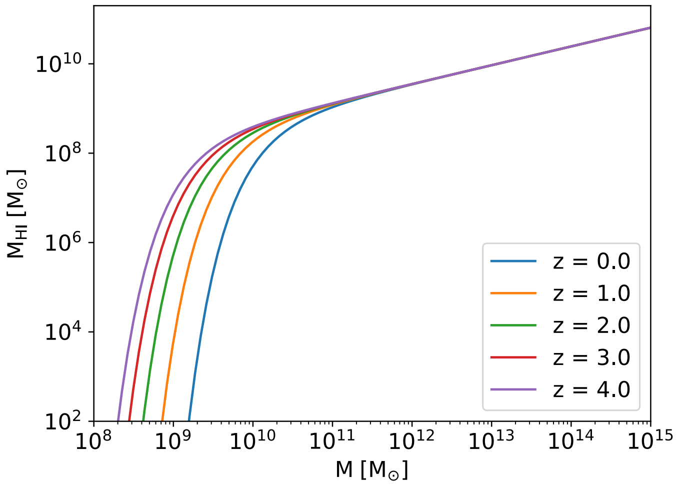




- More massive halos contain more HI
- **But:** Many more small halos than large ones
  → Important not to neglect small halos

- 2 – 3% loss over considered redshift range

HI Mass Loss



- More massive halos contain more HI
- **But:** Many more small halos than large ones

➔ Important not to neglect small halos

- 2 – 3% loss over considered redshift range