Experimental Methods & Results: Foreground cleaning for HI using GNILC

Department of Physics and Astronomy

November 6, 2025

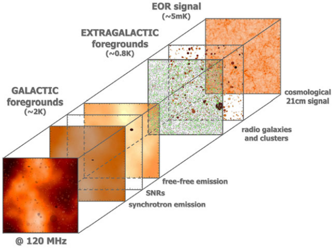
Supervisors: Prof. Mario Santos & Dr. Karin Fornazier

Outline

- Project Overview and Motivation
- Data and Simulations
- GNILC Workflow
- Constrained PCA and Weight Matrix
- HI Signal Recovery and Validation
- Summary

Overview

- Recovery of HI signal from dominant Galactic and Extragalactic foregrounds
- Why is this important?
- To probe the early universe and large-scale structure of the universe

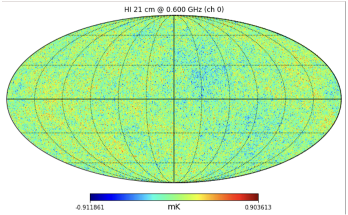


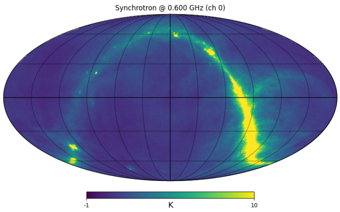
- GNILC separates the HI component from foregrounds using local covariance analysis in needlet space

Data and Simulations

- Simulated foreground maps contain:
  - HI cosmological signal
  - Galactic synchrotron
  - Extragalactic point sources

- Maps are combined into a data cube of dimension $N_{\mathrm{ch}} \times N_{\mathrm{pix}}$.
- Resolution: nside = 256 & Frequency range: UHF





Experimental methods

- We simulated $(252, 786432)$ maps for HI signal and foregrounds
- Software and libraries:
  - Python packages and Healpy for map visualization
  - GNILC implementation for HI intensity mapping (Olivari, 2016) - written in python
- Number of channels: reduced from $N_{chan} = 252$ to $N_{\mathrm{ch}} = 20$

GNILC Workflow

1. Input maps: Foreground maps & prior map
2. Decompose sky maps into needlet space (Olivari, 2016)
3. Compute covariance matrices:

$$R_x(p) = \langle x(p)x^T(p) \rangle$$

4. Uses Prior (HI) to compute signal-to-noise ratio
5. Perform constrained PCA to determine dimensions of signal subspace
6. Apply ILC weights:

$$\hat{s}(p) = W(p)x(p)$$

Component Separation

- Using simulated data in both frequency i & pixel p:

$$x_i(p) = s_i(p) + n_i(p),$$

- Where $s_i(p)$ is the HI signal and n_i astrophysical foregrounds plus the instrumental noise

- Total covariance $R(p) \; = \; R_{HI}(p) \; + \; R_n(p)$

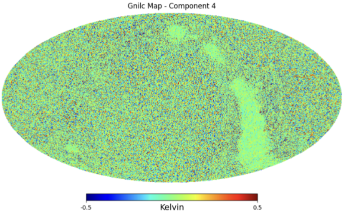
- The estimated signal, $\hat{s}$ by linear operation

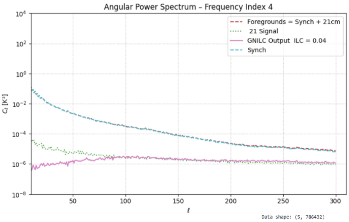
$$\hat{s} = W \cdot x$$

- W - weight matrix that offers unit response to the HI signal while it minimizes the total variance of the foregrounds

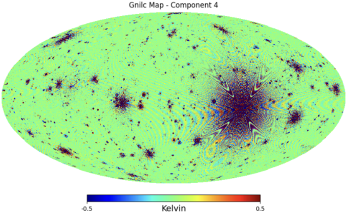
Results

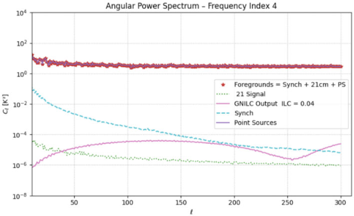
- The recovered HI signal is compared with the simulated true HI signal (Prior): 1. First results, Synchrotron + HI signal





- If we include point sources (stars, quasars, AGNS, etc) :



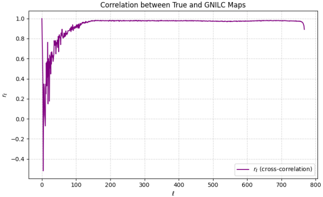


Validation

- Compare recovered and true HI signal, correlation in pixel space:

$$r_p = \frac{Cov\ (Rec, True)}{\sigma_{True}\ \sigma_{Rec}} = 0.956$$

- Correlation coefficient r_ℓ between recovered and true HI:



- RMS - Root Mean Square:

$$RMS(\Delta T) = \frac{1}{N}\sqrt{\Sigma_i (T_{\rm rec} - T_{\rm true})^2} = 0.26K$$

Conclusion & Acknowledgments

- We successfully tested GNILC and recovered the HI signal (with no point sources)
- Improvements can be made in the case of point sources
- Identified key parameters:

Parameter	Formula	Value
z	$z = 1420 \,/\, \nu_{\text{obs}} - 1$	$0.4 - 1.4$
θ_{FWHM}	$\theta = 1.22\,\lambda \,/\, D$	$\sim 1.55^\circ$
ℓ	$\ell \approx \pi \,/\, \theta$	$30 - 300$
SNR	$\text{SNR} = \sigma_{\text{HI}} \,/\, \sigma_{\text{n}}$	$3.689 \gtrsim 1$

- This implies that GNILC can be used to perform foreground cleaning

References

- Olivari, L. C. (2016). *Generalized Needlet Internal Linear Combination.*
- Olivari, L. C. (2018). *Approach to probe large-scale structure.*
- Dai, X., & Ma, Y. (2025). *Expanded GNILC for multi-frequency CMB/HI separation.*

Thank You!

Questions?