Component Separation for HI Intensity Mapping with GNILC

Department of Physics and Astronomy

September 17, 2025

Supervisors: Prof. Mario Santos & Dr. Karin

Outline

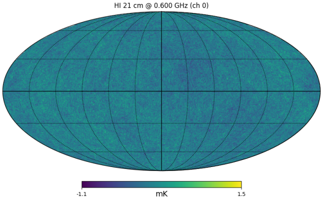
- Introduction
- Simulation - Cosmological signal
- Simulation - Foregrounds
- GNILC method
- Next steps

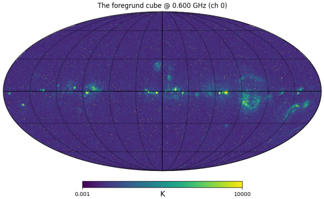
Introduction

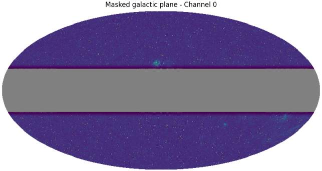
- We want to understand the distribution of matter in the universe
- HI line is a powerful tracer
- Challenge: Foregrounds are orders of magnitude brighter than HI signal - Foreground cleaning is essential
- Goal: Recover the HI signal angular power spectrum C_ℓ^{HI} after foreground removal

Map Simulation

- We use simulations to test GNILC - component separation method before real data
- They allow us to run GNILC under controlled conditions
- Simulations are represented as a 3D cube:
  - Sky position (RA, Dec) $\rightarrow$ 2D maps
  - Frequency ν $\rightarrow$ redshift z (line-of-sight)







HI Power Spectrum

- We want to construct 2D angular power spectrum of HI intensity over some frequency

Brightness temperature $\delta T_b(\hat{n})$

$$\delta T_b(\hat{n}) = \int dz \; W(z) \; \delta_{\mathrm{HI}}(r(z)\hat{n}, z)$$

where,

- $\delta_{HI} = \Delta\rho_{HI}/\rho_{HI}$ is the fractional HI density fluctuation and $W(z)$ is the window function - uniform in redshift z range

- Taking the inverse Fourier Transform of δ_{HI} in k-space,

$$\delta_{HI}(r,z) = \int \frac{d^3k}{(2\pi)^3} \; \delta^*(k,z) \; e^{ikr} \qquad (1)$$

- With comoving position r along direction $\hat{n}$ and redshift z

HI Power Spectrum

- The sky appears to be a sphere - We construct δT_{HI} in spherical Harmonics $Y_{\ell m}$
- From equation 1, We expand the e^{ikr} using the plane-wave spherical expansion,

$$e^{ikr} = 4\pi \, \Sigma_{\ell m} \, i^{\ell} \, j_{\ell} \, (kr) \; Y_{\ell m}(k) \; Y^{*}_{\ell m}(r) \tag{2}$$

Where, j_{ℓ} is the spherical Bessel function

- Then substituting this into δT_{HI},

$$\delta T_b(\hat{n}) = 4\pi \sum_{\ell m} i^{\ell} \int dz \, W(z) \times \int \frac{d^3\mathbf{k}}{(2\pi)^3} \, \delta_{\rm HI}(\mathbf{k}, z) \, j_{\ell}(kr(z)) \, Y_{\ell m}(\hat{\mathbf{k}}) \; Y^{*}_{\ell m}(\hat{n}) \tag{3}$$

HI Power Spectrum

- Expanding the δT_{HI} in spherical Harmonics $Y_{\ell m}$,

$$\delta T_b = \sum_{\ell m} a_{\ell m} Y_{\ell m} \tag{4}$$

- Explicitly, we can identify the harmonic coefficients $a_{\ell m}$ from equation 3,

$$a_{\ell m} = 4\pi i^{\ell} \int dz\, W(z) \int \frac{d^3\mathbf{k}}{(2\pi)^3} \hat{\delta}_{\mathrm{HI}}(\mathbf{k}, z)\, j_{\ell}(kr(z))\, Y^{*}_{\ell m}(\hat{\mathbf{k}}) \tag{5}$$

- Finally the angular power spectrum C_ℓ can be estimated by average over m for fixed ℓ

$C'_\ell s$

$$C_\ell = \frac{1}{2\ell + 1} \sum_{\ell m} |a_{\ell m}|^2$$

Angular Power Spectrum C_ℓ

- From previous equation, we want to calculate $\langle a_{\ell m} a^*_{\ell' m'} \rangle$
- Considering the Dirac delta and the completeness relation,

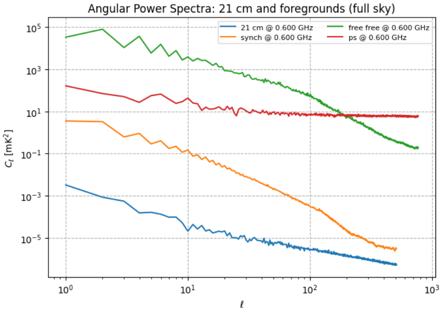
$$\langle \delta_{\rm HI}(\mathbf{k}, z)\, \delta^*_{\rm HI}(\mathbf{k}', z) \rangle = (2\pi)^3\, \delta^3(\mathbf{k} - \mathbf{k}')\, P_{\rm c}(k), \qquad (6)$$

- Intuitively, modes with $k = k'$ are correlated; and $k \neq k'$ their correlation vanishes
- Using equation (6) we get the exact angular power spectrum,

$$C_\ell = \frac{2}{\pi} \int k^2\, dk\ P_c(k) \left[\int dz\, W(z)\, D(z)\, j_\ell(kr(z)) \right]^2 \qquad (7)$$

- This equation is computationally demanding (for $\ell > 10^3$), in this project we compute $C_{\ell m}$ from exact equation ($\ell < 10^3$)
- The Limber approximation is valid for large ($\ell = \pi/\theta$) values

C_ℓ from simulations



- Astrophysical components: Galactic synchrotron, Galactic free-free and extragalactic point sources
- Before testing GNILC, we add components and cosmological signal to have one single cube

GNILC Method

Definition

We denote the observed sky as $x_i(p)$ in both pixel p and frequency i space, and is the sum of HI signal $s_i(p)$ and foreground componets $f_i(p)$,

$$x(p) = s(p) \ + \ f(p) \tag{8}$$

All in $N_{ch} \times 1$ column vector

- From these sky observations $x(p)$, with GNILC we want to decompose them using needlets (type of spherical wavelet)
- We compute the covariance matrices $R_x = \langle \mathrm{x(p)} \ x(p)^T \rangle$,

$$R(p) \ = \ R_s(p) + R_f(p)$$

- Where $R_s(p) = \langle \mathrm{s(p)} \ s(p)^T \rangle$ the covariance matrix of the HI signal

- For General case, the $s(p)$, can be expressed as a linear combination of $(N_{chan} - m)$ independent templates t,

$$s = St$$

- Where S is the mixing matrix $N_{chan} \times \ (N_{chan} - m)$
- We can also find $R_s \ = \ \langle s\, s^T \rangle \ = \ \langle S\, R_t\, S^T \rangle$
- Using the internal Linear combination (ILC), we want to estimate s by linear operation,

$$\hat{s} \ = \ Wx \tag{9}$$

- Where, W weight matrix which Keeps the HI signal while minimizing the total variance $->$ Key

Next steps:

- Computing the covariance matrices
- Construct GNILC weight matrix
- perform a ‘constrained’ Principal Component Analysis (PCA)

References

- Olivari, L. C. (2018). *Approach to probe the large-scale structure of the Universe.*
- Olivari, L. C (2016). *(GNILC) - Generalized Needlet Internal Linear Combination*
- De Caro, B., Carucci, et all. (2025). eGNILC approach for HI intensity mapping.
- Dai, X., Ma, Y. (2025). *Expanded Generalized Needlet Internal Linear Combination (eGNILC).*