HI INTENSITY MAPPING: EXTRACTING THE 21 CM SIGNAL IN THE PRESENCE OF FOREGROUNDS

L. C. OLIVARI
Jodrell Bank Centre for Astrophysics, Alan Turing Building, School of Physics and Astronomy, The University of Manchester, Oxford Road, Manchester, M13 9PL, U.K.

HI intensity mapping is a new observational technique to map fluctuations in the large-scale structure of matter using the 21 cm emission line of atomic hydrogen (HI). Sensitive HI intensity mapping experiments have the potential to detect Baryon Acoustic Oscillations (BAO) at low redshifts ($z \leq 1$), which can be used to constrain the properties of dark energy. Observations of the HI signal will be contaminated by instrumental noise and by astrophysical foregrounds, such as Galactic synchrotron emission, which is at least four orders of magnitude brighter than the HI signal. We study the ability of the Generalized Needlet Internal Linear Combination (GNILC) method to subtract radio foregrounds and to recover the cosmological HI signal. For simulated radio observations including HI emission, Galactic synchrotron, Galactic free-free, radio sources and 0.05 mK thermal noise, we find that the GNILC method can reconstruct the HI power spectrum for multipoles $30 < \ell < 150$ with 6% accuracy on 50% of the sky for a redshift $z \sim 0.25$.

1 Introduction

Intensity mapping is the study of the large-scale fluctuations in the intensity of a given spectral line, in particular the 21 cm line of the neutral hydrogen, emitted by a number of unresolved objects. It has been suggested that the HI intensity field can be used to measure the power spectrum as a function of redshift [1]. The main advantage of HI intensity mapping, compared to the optical surveys of galaxies, is that a large sky volume is achieved within a relatively short observing time. This type of radio survey can then be used to measure the Baryon Acoustic Oscillations (BAO), which can be used to constraint the properties of the dark energy, the mysterious component that drives the accelerated expansion of the Universe.

The success of any HI intensity mapping experiment strongly depends on our ability to subtract the astrophysical contaminations that will be present in the observed HI signal. At $\nu = 1$ GHz, the most relevant foregrounds are the Galactic emission, mostly synchrotron radiation, and the background emission of extra-galactic point sources. These emissions are at least four orders of magnitude larger, $T_b \sim 10$ K, than the HI signal, $T_b \sim 1$ mK. The high spectral resolution offered by any HI intensity mapping experiment allows us to use the frequency information of the observed data. As the foregrounds spectra are expected to be smooth, we can approximate them by a power-law in the frequency range of interest. This property can then be used to separate the HI signal from any other signal correlated in frequency [2].

In this work, we describe the application of the Generalized Needlet Internal Linear Combination (GNILC) [3,4] as a non-parametric component separation technique for HI intensity mapping experiments. In general, the GNILC method can extract the emission of a multidimensional component (spatially correlated components) from the observed data. Here we review the performance of the GNILC method for a single-dish HI intensity mapping experiment in low

redshifts and make a brief discussion on how component separation connects to cosmology.

2 GNILC

The GNILC method can be divided into two main steps. First, using a prior on the HI power spectrum, we determine the local signal to noise ratio and perform a constrained Principal Component Analysis (PCA) of the data to determine the effective dimension of the HI subspace. Second, we perform a multimensional ILC filter within the HI subspace and reconstruct the HI signal. In the constrained PCA step, the number of principal components of the observation covariance matrix is estimated locally both in space and in angular scale, which is done by using a wavelet (needlet) decomposition of the observations. We also use a statistical information criterion (AIC) [5] to make the selection of the principal components of the observation covariance matrix.

The steps of the GNILC method can be summarized as follows:

1) To isolate the different ranges of angular scales, we first define a set of needlet windows in harmonic space. These needlet windows work as band-pass filters. The spherical harmonic coefficients $a_{\ell m}$ of the observed frequency maps are then band-pass filtered by the needlet windows. We then produce one observed map for each needlet scale j.

2) For each needlet scale j, we compute the data covariance matrix, at pixel p, of a pair of frequencies a and b as

$$\hat{\mathbf{R}}_{ab}(p) = \sum_{p' \in \mathcal{D}(p)} \mathbf{x}_a(p') \mathbf{x}_b^T(p'), \tag{1}$$

where $\mathcal{D}$ is a domain of pixels centred around the pixel p.

3) For each needlet scale j, we also compute the HI emission covariance matrix by using HI maps $\mathbf{y}$ simulated from a theoretical prior on the HI angular power spectrum, exactly like Eq. 1.

4) We diagonalize the transformed data covariance matrix, $\hat{\mathbf{R}}_{\rm HI}^{-1/2} \hat{\mathbf{R}} \hat{\mathbf{R}}_{\rm HI}^{-1/2}$, as

$$\hat{\mathbf{R}}_{\rm HI}^{-1/2} \hat{\mathbf{R}}(m) \hat{\mathbf{R}}_{\rm HI}^{-1/2} = \mathbf{U}_N \mathbf{D}_N \mathbf{U}_N^T + \mathbf{U}_S \mathbf{U}_S^T, \tag{2}$$

where D_N collects the m largest eigenvalues as selected by the minimization of AIC [4], $\mathbf{U}_N$ the corresponding eigenvectors, and $\mathbf{U}_S$ the $(n_{\rm ch} - m)$ eigenvectors related to the HI emission subspace.

5) For each needlet scale j, we apply an $(n_{\rm ch} - m)$-dimensional ILC filter to the observed data,

$$\hat{\mathbf{s}}^{(j)} = \hat{\mathbf{S}} (\hat{\mathbf{S}}^T \hat{\mathbf{R}}^{-1} \hat{\mathbf{S}})^{-1} \hat{\mathbf{S}}^T \hat{\mathbf{R}}^{-1} \mathbf{x}^{(j)}, \tag{3}$$

where the estimated mixing matrix is given by $\hat{\mathbf{S}} = \hat{\mathbf{R}}_{\rm HI}^{1/2} \mathbf{U}_S$.

6) Finally, we synthesize the reconstructed needlet HI maps $\hat{\mathbf{s}}^{(j)}$ as follows: the maps are transformed to spherical harmonic space, their harmonic coefficients are again band-pass filtered by the respective needlet window and the filtered harmonic coefficients are transformed back to maps in real space. This operation gives one reconstructed HI maps per needlet scale. These maps are then added to give, for each frequency channel, the complete reconstructed HI map.

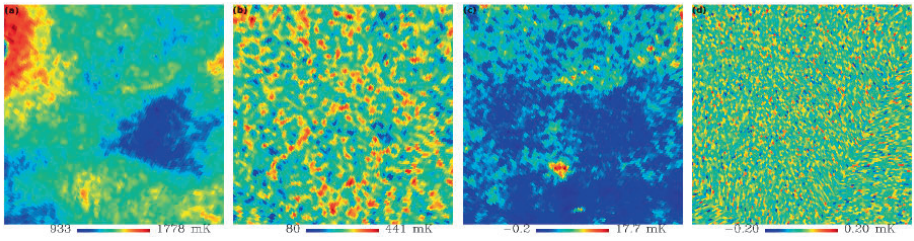

Figure 1 – Maps of four simulated signals covering a $50^\circ \times 50^\circ$ patch of the sky: (a) Galactic synchrotron, (b) extragalactic point sources, (c) Galactic free-free, (d) and thermal noise. The maps are centred at Galactic coordinates $(120^\circ, -60^\circ)$ and observed at 1117.5 MHz (redshift $z \sim 0.3$).

3 Simulation

To test our component separation method we perform a simulation of the observed sky that is expected for a standard HI intensity mapping experiment. For our simulation, we consider the bandwidth of the BINGO experiment [6]: 960 to 1260 MHz. This bandwidth corresponds to a redshift range of 0.13 to 0.48. The frequency channels are chosen to be equally spaced in the given bandwidth. The number of frequency channels is chosen to be 20.

Our astrophysical sky is given by the HI emission, synchrotron radiation with spatially variable spectral index, extragalactic radio point sources, free-free radiation, and thermal noise with an amplitude of 0.05 mK per pixel. Figure 1 shows some maps of the components of the simulated sky at 1117.5 MHz. The temperature scale is in mK.

We use a Galactic mask to cover the area of the sky where the synchrotron emission of our Galaxy is brightest. Our Galactic mask is given by the product of two masks: a $\pm 20^\circ$ latitude mask and a mask designed to cover the Galactic emission with brightness temperatures larger than the threshold of 30 K at 408 MHz. This mask gives 51% of sky coverage.

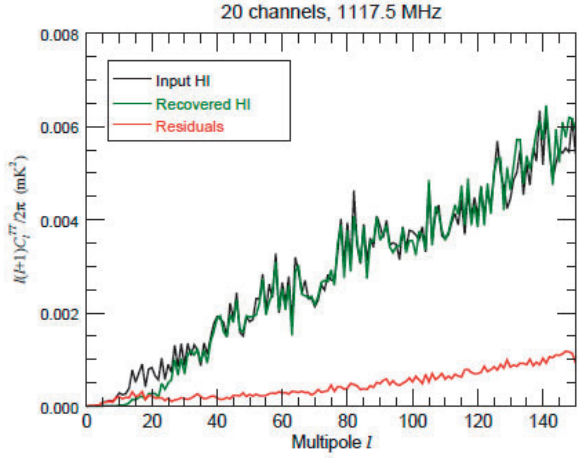

Figure 2 – Result on simulation (HI, synchrotron with spatially variable spectral index, point sources, free-free, and thermal noise): power spectra, $\ell(\ell+1)C_\ell/2\pi$, of the simulated HI signal (black), the recovered HI signal (green), and the residual signal (red) at frequency 1117.5 MHz with 20 frequency channels.

After applying the GNILC method to our simulated maps, we obtain an average normalized absolute difference between the input HI power spectrum, C_ℓ^S, and the reconstructed HI power

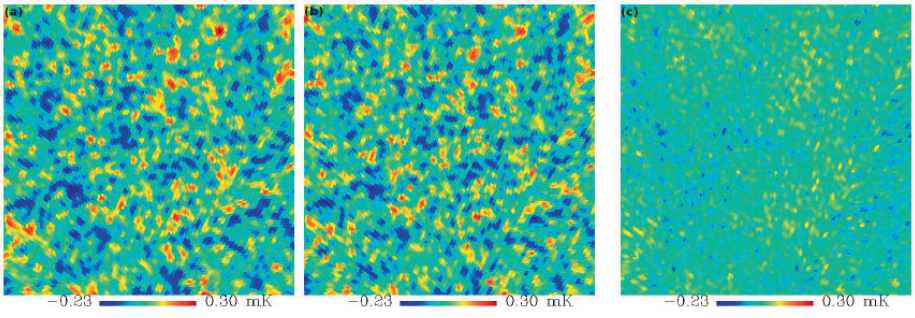

Figure 3 – Maps covering a $50^\circ \times 50^\circ$ patch of the sky with three different signals: (a) input HI, (b) GNILC reconstructed HI, and (c) residuals. The maps are centred at Galactic coordinates $(120^\circ, -60^\circ)$ and observed at 1117.5 MHz (redshift $z \sim 0.3$).

spectrum, C_ℓ^R, of the HI signal that is around 6.1% for multipoles in the range between 30 and 150. Figure 2 shows the power spectra of the recovered HI signal, the input HI signal, and the residual map (recovered HI signal minus input HI signal) for our simulation with 20 frequency channels. Figure 3 shows the input HI signal, the reconstructed HI signal and the residual contamination at 1117.5 MHz. These maps are shown with an angular resolution of 40 arcmin. We see that, after cleaning foregrounds and thermal noise of the observed signal, the GNILC method is able to reconstruct the temperature fluctuations of the HI signal with a r.m.s residual equals to 0.04 mK at this $50^\circ \times 50^\circ$ patch of the sky.

4 Cosmology and Component Separation

Once we have a set of reconstructed HI maps, we can use these maps to calculate the angular power spectra that are needed to study the large-scale properties of the Universe and thus constraint the cosmological parameters of the particular cosmology being studied. As any component separation method always loses some HI power, we may have as a consequence a bias in the power spectra. This bias in the power spectra may leads to a bias in some cosmological parameters (e.g. optical depth and Hubble parameter). A careful analysis of the effect of component separation in the cosmological parameters is then necessary; this analysis is being done by the present author and soon will be available in the literature [7].

Acknowledgements

LCO acknowledges funding from CNPq, Conselho Nacional de Desenvolvimento Cientifico e Tecnologico - Brazil.

References

1. J. B. Peterson *et al* in Astronomy, Vol. 2010, astro2010: The Astronomy and Astrophysics Decadal Survey, p. 234 (2009).
2. R. Ansari *et al*, A&A **540**, A129 (2012).
3. M. Remazeilles *et al*, MNRAS **418**, 467 (2011).
4. L. C. Olivari *et al*, MNRAS **456**, 27490 (2016).
5. H. Akaike in IEEE Transactions on Automatic Control, 19, 716 (1974).
6. R. A. Battye. *et al*, MNRAS **434**, 1239 (2013).
7. L. C. Olivari *et al* in preparation.