Thermal SZ sky maps and cross-correlation with weak lensing surveys - Implications for cosmology and baryonic physics

Pranav S **Do all the thesis fancy things**
like university logo, filfilment dialogue and supervisor

August 20, 2019

Abstract

The knowledge of the baryonic and non-baryonic distribution of the universe is fundamental in understanding evolution and structure formation in the Universe. However, a large fraction of them cannot be detected directly. One of the major ways of indirectly detecting the baryon distribution, is cross correlation between the thermal Sunyaev Zeldovich effect (tSZ) and Weak Gravitational Lensing. Doing an independent analysis of this cross correlation will also be useful in detecting any systematic errors, if any, in the existing skymaps. We compute the tSZ skymaps using a methodology independent from the Planck collaboration's piplines, and then cross correlate these skymaps with tangential shear to compare with the existing constrains on halo astrophysics and cosmology [1]. This work consists of three parts, a) Independent generation of tSZ skymaps using machine learning (3.3), b) Cross-correlating the tSZ skymaps with weak-lensing maps (4.1), c) Comparison with theory using Halo models(4.4). This work was done partially in collaboration with Prof. Rishi Khatri.

Acknowloedgements

Contents

Chapter 1

Introduction

WRITE

Chapter 2

Relevant Physics

2.1 SZ Effect

2.2 Weak Lensing

Chapter 3

Foreground Component Seperation

One of the ways of removing foregrounds from our data is to have a parametric model for the foregrounds based on known physics and fit the data to these foreground models to get an estimate of the parameters and use them to eliminate the foregrounds in the data. This is the methodology followed by Commander pipeline of the Planck collaboration [2] and LIL Method [3]. These methods suffer from being model dependent and the parameters are insufficient to model the foregrounds because various physical processes contribute to the foreground in a single pixel, making it hard to model these with a single parametric model. Therefore, it is useful to look at algorithms which are model independent and just use the data to estimate the foreground characteristics.
Various such *blind* component seperation algorithms have been proposed where only the spectrum of the signal is needed [4, 5] and have been applied in various variations such as Harmomic space, Needlet frame etc [6, 7]. Since these blind algorithms require the number of foreground components to be lesser than the number of spectral channels available, it is necessary to divide the sky with similar foreground properties into different regions and apply the algorithm seperately on those regions to reduce the number of foreground components. These existing variations try to divide the data into different clusters with similar foregrounds based on heuristic arguments and our current understanding of the properties of the foregrounds. We provide a spectral data based approach extending upon this data driven foreground clustering approach [8], which uses the signature of the foregrounds available in the data. Since the data is clustered only based on the spectra, different regions of the sky can still be in the same cluster based on their foreground properties. To the best of our knowlegde, this is the first attempt at using *unsupervised* machine learning for foreground component seperation.

3.1 Generalized ILC

Also write an appendix about the derivation of the formulas

3.2 Quick Introduction to Machine Learning

Machine learning is a type of an algorithm which, rather than explicitly coding an algorithm for a particular function, gives the computer the ability to *learn* that function. Machine learning usually involves a training set, from which the computer *learns* the function of interest, and a validation set where the computer applies whatever it has learned from the training set and gives us results. Based on the way we use our training set, machine learning can be divided into two categories.

- Supervised Machine Learning
- UnSupervised Machine Learning

3.2.1 Supervised Machine Learning

A machine learning algorithm is said to perform supervised learning, when the training set is differet from the validation set, and we have the expected learning outcomes as part of our training set.

3.2.2 UnSupervised Machine Learning

3.2.3 Neural Networks

3.3 New tSZ maps with unsupervised Machine Learning

In unsupervised machine learning, we take the hard clustering approach, where each data belongs only to a single class. In order to smooth the boundaries of the partitions like the current ILC algorithms, we repeat the hard clustering multiple times with different seeds. Since various seeds produce different partitions, and there is no a-priori reason to consider one partition over the other, we consider all of them with equal probability. We see that this approach essentially smoothens the boundary and no additional smoothing across the cluster boundaries are necessary.

3.3.1 'k-means clustering' method

EXPLAIN KMEANS k-means clustering partitions n data points into k clusters, by associating each point to the nearest centroid, which serves as a representation of that cluster. By minimising the inertia, We partition the space into k-regions, which serve as the partition for our clusters to be used in ILC.
INSERT IMAGE FOR KMEANS CLUSTERING HERE
The initial positions are initialized using `k-means++` algorithm inorder to avoid the local minimas, which the random initializations can get into. And the data is rescaled with quartile scaling, to prevent the outliers from influenzing the scaling. k-means clustering was implemented using `scikit-learn` [9]

Choice of Variables for k-means clustering

Let the frequency maps in K_{CMB} units be denoted by, T_i, where
$i \in \{30, 44, 70, 100, 143, 217, 353, 545, 857\}$Hz. In order to model the foregrounds, We first subtract the CMB. $A_i = T_i - T_{100Hz}$. Then we substract out the tSZ, by changing to the corresponding y units. The maps we get are mostly dominated by foregrounds. Apart from clustering these foreground maps seperately (labeled as *raw k-means*), we divide out the amplitudes of these *raw maps* to make different measures, which capture how fast the foreground is increasing or decreasing with frequency similar to the measure defined in the FC-ILC (labelled as *kmeans with m*) [8].
Let the frequency maps in K_{CMB} units be denoted by, T_i, where $i \in \{30, 44, 70, 100,$
$143, 217, 353, 545, 857\}$Hz. In order to model the foregrounds, We first subtract the CMB. $A_i = T_i - T_{100Hz}$. Then we substract out the tSZ, by changing to the corresponding y units. The resultant maps we get, are dominated by foregrounds. **EXPLAIN THESE NUMBERS - WRITE APPENDIX ON UNIT CONVERSION**

$$y_1 = A_{857} - 24.371A_{143} \quad y_2 = A_{857} - 7.199A_{217} \quad y_3 = A_{545} - 14.836A_{143} \tag{3.1}$$
$$y_4 = A_{545} - 4.3826A_{217} \quad y_5 = A_{353} - 8.215A_{143} \quad y_6 = A_{353} - 2.4267A_{217} \tag{3.2}$$
$$y_7 = A_{30} - 1.7339A_{70} \quad y_8 = A_{44} - 1.541A_{70} \quad y_9 = A_{217} + 3.671A_{44} \tag{3.3}$$
$$y_{10} = A_{217} + 5.657A_{70} \quad y_{11} = A_{353} + 13.73A_{70} \quad y_{12} = A_{30} - 1.125A_{44} \tag{3.4}$$
$$\tag{3.5}$$

Apart from clustering these y-maps seperately. We divide out the amplitudes to make different measures, which capture how fast the foreground is increasing or decreasing with frequency.

$$m_1 = y_1/y_3 \quad m_2 = y_2/y_4 \quad m_3 = y_3/y_5 \tag{3.6}$$
$$m_4 = y_4/y_6 \quad m_5 = y_7/y_{12} \quad m_6 = y_{12}/y_5 \tag{3.7}$$
$$m_7 = y_7/y_5 \quad m_8 = y_7/y_8 \quad m_9 = y_8/y_5 \tag{3.8}$$

By visually inspecting the maps, We choose the 3 measures. m_1, m_3, m_7, and perform k-means clustering. We also masked the galactic center along with the point sources and clustered the masked regions seperately.

3.3.2 Masks

Write about quartile scaling Masking the ouuliers, and deconvolving the mask using polspice when calculating the power spectrum

3.3.3 Dimensionaliy Reduction

Since, the choice of parameters is upto us in the previous method, we use non-linear dimensionality reduction algorithms in order to reduce the 12 dimensional input data (CMB and tSZ substracted maps) to a lower dimension in order to apply the k-means algorithm on it. Unlike the previous method, we use neural networks to determine the parameters which encode the maximum information about the data. We use two non-linear dimensionality reduction algorithms.
The two algorithms used are,

- Auto Encoders
- Self Organising Maps

3.3.4 Auto Encoders

We use a denoising auto encoder, made from an fully connected neural network, using L^2-norm as the cost. Once the network has been trained we can use the encoder part for dimensionality reduction. And then the k-means algorithm is used in this lower dimensional space to cluster the pixels. Autoencoders were implemented using `tensorflow` [10].

3.3.5 Self Organising Maps

Self organising map reduces the input data to two dimensions by fitting a discrete two dimensional manifold on the data and then use k-means algorithm on this manifold. Self Organising Maps was implemented using the `somoclu` Library [11].

3.3.6 Results

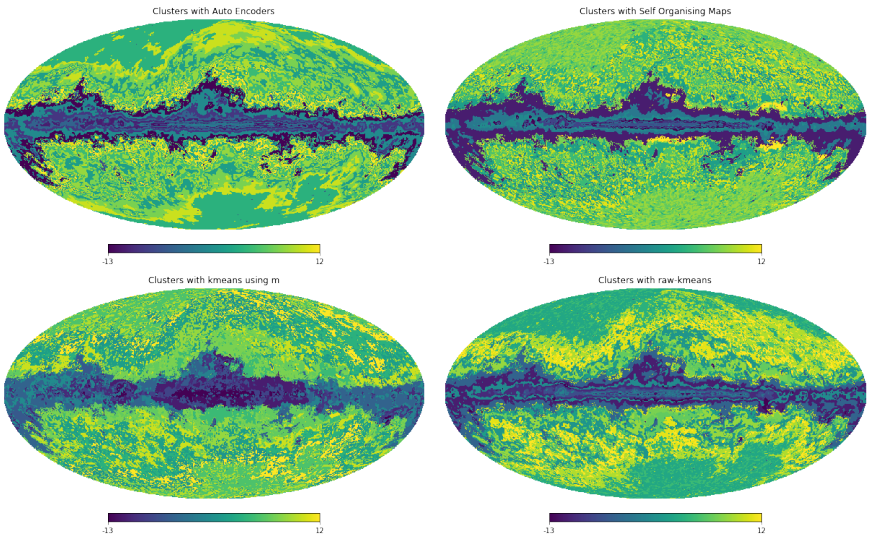

Figure 3.1: A single instance Clustering of the Sky based on foregrounds using various methods (negative values indicate masked regions)

Chapter 4

Thermal-SZ Weak Lensing Cross Correlation

4.1 Calculation from Data

Once the tSZ maps are found using the above method, we cross-correlate the the tangential shear component using the Weak-Lensing datasets available. We compute the cross correlation using the k-D trees algorithm. k-D trees perform efficiently in nearest neighbour searches and help speedup calculating the correlation functions significantly. We use the `treecorr` library for this [12]. We perform the cross correlation using the Red Cluster Sequence Lensing Survey (RCSLens) dataset [13] and the Planck tSZ skymaps [14], initially, to verify our cross correlations by comparing the results with [1]. We plan to do the same with the Kilo Degree Survey (KiDS) dataset [15] and our own tSZ skymaps. Currently this is the state where I am working on. Following this, we will compare with different theoretical predictions will help us constrain the cosmology and baryon models. Especially, We are interested in looking at non-gravitational feedback.

4.2 k-D Trees Algorithm

4.3 Systematic Tests

Explain why B-modes are zero

Explain why Randomization makes the correlation go to zero

4.4 Comparison with Theory

We like to compare our observations with theoretical predicitions based on Halo models, closely following the method developed by [16]. Currently we are concentrating on computing the real space cross correlation $\xi = \langle \gamma_T - y \rangle$, By using

$$\xi^{y-\gamma_T}(\theta) = \int \frac{d^2\vec{l}}{2\pi^2} C_l^{y-k} \cos(2(\phi - \psi)) \exp(i\theta \cos(\phi - \psi)) \tag{4.1}$$

Where, ϕ is the Polar angle with respect to the coordinate system and ψ is the angle between $\vec{l}$ and the coordinate.

In order to compute the $y - k$ cross correlation power spectra found in Equation 4.1, we use the 1-halo term as defined in [17].

$$C_l^{y-k,1h} = \int_0^{z_{max}} dz \frac{dV}{dz d\Omega} \int_{M_{min}}^{M_{max}} dM \frac{dn}{dM} y_l(M,z) k_l(M,z) \tag{4.2}$$

For the halo mass function, We use the form suggested by Sheth and Tormen [18] For the convergence profile in fourier space, We use

$$k_l = \frac{W^k(z)}{\xi^2(z)} \frac{1}{\rho_m} 4\pi \int_0^{r_{vir}} dr r^2 \frac{\sin(lr/\xi)}{lr/\xi} \rho(r; M, z) \tag{4.3}$$

And for the fourier transform of the projected gas pressure.

$$y_l = \frac{4\pi r_s}{l_s^2} \frac{\sigma_T}{m_e c^2} \int dx x^2 \frac{\sin(lx/\xi)}{lx/\xi} P_e(x; M, z) \tag{4.4}$$

For electron pressure, We use the *universal pressure profile* and the NFW model. We plan to use the best fit parameters for the Pressure profile from the Planck Collaboration and then compare the halo model predictions from both the Planck and WMAP-7yr cosmologies. To look for any non-gravitational feedback, we plan to use the method developed by [19], applied to tSZ C_ls.

Chapter 5

Main results

The result of using the various algoritms FFP6 simulations is shown in Fig(5). We see that our clustering algorithm is better than the one dimensional version [8]. We hope to compare it with the existing algorithms using FFP6 simulations. We also see that the neural network method like self organising map with a 200x200 grid performs, very close to the k-means algorithm where the measures are chosen by hand. It is interesting to note that *raw-kmeans* performs the same as a self organising map. We plan to extend the work to the grid size of 500x500 and to improve the performance of auto-encoders by doing a hyperparameter optimization. The work regarding the cross-correlation and comparison with the halo model is ongoing.

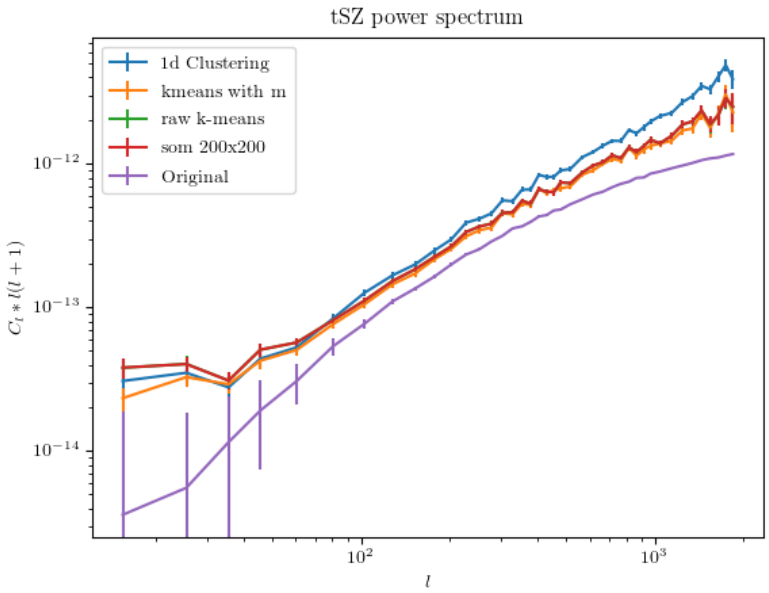

Figure 5.1: Comparison of the tSZ power spectrum for various methods

Chapter 6

Conclusions

To summarize, We saw that doing foreground clustering using a single measure already performed on par with the Planck collaboration's pipelines [8]. We see now that the various machine learning algorithms used for this improves upon the FC-ILC algorithm using a single measure. We hope that using a independent algorithm to produce maps with different residuals, would be useful in testing for the effect of foregrounds and any biases in the existing algorithms for the estimates of the cosmological parameters. We would also like to point out that in future experiments with more channels using machine learning techniques with allow for more accuracy than single parameter clustering. **TEST**

Bibliography

[1] A. Hojjati *et al.*, "Cross-correlating Planck tSZ with RCSLenS weak lensing: Implications for cosmology and AGN feedback," *Mon. Not. Roy. Astron. Soc.* **471** no. 2, (2017) 1565–1580, arXiv:1608.07581 [astro-ph.CO].

[2] H. K. Eriksen *et al.*, "CMB component separation by parameter estimation," *Astrophys. J.* **641** (2006) 665–682, arXiv:astro-ph/0508268 [astro-ph].

[3] R. Khatri, "Linearized iterative least-squares (LIL): a parameter-fitting algorithm for component separation in multifrequency cosmic microwave background experiments such as Planck," *Mon. Not. Roy. Astron. Soc.* **451** no. 3, (2015) 3321–3339, arXiv:1410.7396 [astro-ph.CO].

[4] H. K. Eriksen, A. J. Banday, K. M. Gorski, and P. B. Lilje, "Foreground removal by an internal linear combination method: Limitations and implications," *Astrophys. J.* **612** (2004) 633–646, arXiv:astro-ph/0403098 [astro-ph].

[5] M. Remazeilles, J. Delabrouille, and J.-F. Cardoso, "Foreground component separation with generalised ILC," *Mon. Not. Roy. Astron. Soc.* **418** (2011) 467, arXiv:1103.1166 [astro-ph.CO].

[6] J.-F. Cardoso, M. Martin, J. Delabrouille, M. Betoule, and G. Patanchon, "Component separation with flexible models. Application to the separation of astrophysical emissions," arXiv:0803.1814 [astro-ph].

[7] D. Marinucci, D. Pietrobon, A. Balbi, *et al.*, "Spherical Needlets for CMB Data Analysis," *Mon. Not. Roy. Astron. Soc.* **383** (2008) 539, arXiv:0707.0844 [astro-ph].

[8] R. Khatri, "Data driven foreground clustering approach to component separation in multifrequency CMB experiments: A new Planck CMB map," *JCAP* **1902** (2019) 039, arXiv:1808.05224 [astro-ph.CO].

[9] F. Pedregosa, G. Varoquaux, A. Gramfort, *et al.*, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research* **12** (2011) 2825–2830.

[10] M. Abadi, A. Agarwal, *et al.*, "Tensorflow: Large-scale machine learning on heterogeneous systems," 2015. https://www.tensorflow.org/. Software available from tensorflow.org.

[11] P. Wittek, S. Gao, I. Lim, and L. Zhao, "somoclu: An efficient parallel library for self-organizing maps," *Journal of Statistical Software, Articles* **78** no. 9, (2017) 1–21. https://www.jstatsoft.org/v078/i09.

[12] M. Jarvis, G. Bernstein, and B. Jain, "The skewness of the aperture mass statistic," *Mon. Not. Roy. Astron. Soc.* **352** (2004) 338–352, arXiv:astro-ph/0307393 [astro-ph].

[13] H. Hildebrandt *et al.*, "RCSLenS: The Red Cluster Sequence Lensing Survey," *Mon. Not. Roy. Astron. Soc.* **463** no. 1, (2016) 635–654, arXiv:1603.07722 [astro-ph.CO].

[14] **Planck** Collaboration, N. Aghanim *et al.*, "Planck 2015 results. XXII. A map of the thermal Sunyaev-Zeldovich effect," *Astron. Astrophys.* **594** (2016) A22, arXiv:1502.01596 [astro-ph.CO].

[15] K. Kuijken *et al.*, "The fourth data release of the Kilo-Degree Survey: ugri imaging and nine-band optical-IR photometry over 1000 square degrees," arXiv:1902.11265 [astro-ph.GA].

[16] Y.-Z. Ma, L. Van Waerbeke, G. Hinshaw, A. Hojjati, D. Scott, and J. Zuntz, "Probing the diffuse baryon distribution with the lensing-tSZ cross-correlation," *JCAP* **1509** no. 09, (2015) 046, arXiv:1404.4808 [astro-ph.CO].

[17] A. Cooray and R. K. Sheth, "Halo Models of Large Scale Structure," *Phys. Rept.* **372** (2002) 1–129, arXiv:astro-ph/0206508 [astro-ph].

[18] R. K. Sheth and G. Tormen, "An Excursion Set Model of Hierarchical Clustering : Ellipsoidal Collapse and the Moving Barrier," *Mon. Not. Roy. Astron. Soc.* **329** (2002) 61, arXiv:astro-ph/0105113 [astro-ph].

[19] A. Iqbal, S. Majumdar, B. B. Nath, S. Ettori, D. Eckert, and M. A. Malik, "Excess entropy and energy feedback from within cluster cores up to r_{200}," *Mon. Not. Roy. Astron. Soc.* **472** no. 1, (2017) 713–726, arXiv:1703.00028 [astro-ph.CO].