Max-Planck-Institut für extraterrestrische Physik

Correlating Cosmological Probes using the Mutual Information

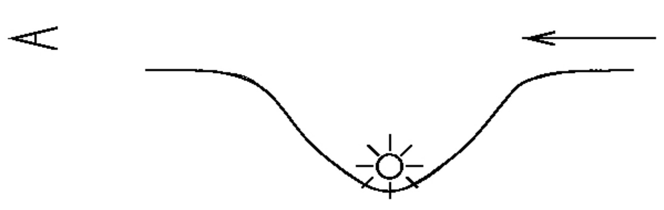
C. Räth[1] , A. Finoguenov[1]
(1) Max-Planck-Institut für extraterrestrische Physik, 85748 Garching, Germany

Abstract

One way to determine parameters describing models for dark energy via the integrated Sachs Wolfe (ISW) effect [1,2] is given by cross-correlating the cosmic microwave background (CMB) with tracers of the large scale structure (LSS) [3].
In several studies it was claimed that there is a significant cross-correlation between the CMB and the densities of galaxies (see e.g. [4-7]). Skeptical reanalyses doubt, however, the significances of those results [8,9].
Up to now the correlations are mostly quantified using linear similarity measures. It is, however, known that in many cases the matching of two quite distinct data sets is more efficiently detected using nonlinear similarity measures like the mutual information (MI) [10] and other entropy-based quantities [11,12].
Here, we propose and discuss the concept of MI and its application as similarity measure for correlating cosmological probes using the WMAP seven year data for the CMB and NVSS and HEAO data as tracers for the LSS. For comparison we also calculate the (linear) cross-correlation (CC) for the data sets.
We find significant correlations between CMB and LSS using MI, where we find that the CC and MI show complementary behaviour. New and significantly improved CMB (PLANCK) and LSS (e.g. eROSITA, EUCLID) data will allow for a full exploitation of the possibilities of MI, which will thus have the potential to become a versatile yet complementary similarity measure for further constraing dark energy parameters via the ISW effect.

Integrated Sachs Wolfe (ISW) - Effect

gravitational potential traced by galaxy or galaxy cluster density



potential depth changes as cmb photons pass through

$$\frac{\delta T}{T} = -2\int \dot{\Phi}(\tau)d\tau$$

The gravitational potential is actually constant in a matter dominated universe on large scales. However, when the equation of state changes, so does the potential, and temperature anisotropies are created.

Linear Correlation Measure: Cross-correlation function CC(ϑ)

Cross-correlation between survey p and q is given by:

$$CC(\vartheta) = \frac{1}{N^{pq}(\vartheta)}\sum_{i,j} f_i^p\left(n_i^p - \langle n^p\rangle\right) f_j^q\left(n_j^q - \langle n^q\rangle\right),$$

$$N^{pq}(\vartheta) = \sum_{i,j} f_i^p f_j^q$$

$N^{pq}(\vartheta)$: Number of pairs of pixels at a given separation

We use 14 angular bins in the range $1° < \vartheta < 14°$.

Nonlinear Correlation Measure: Mutual Information MI(ϑ)

The mutual information, originating from information theory, is a measure of statistical dependencies between two data sets p and q.
It is given by:

$$MI^{pq}(\vartheta) = \sum_{a,b} p(a,b)\log\frac{p(a,b)}{p(a)p(b)}$$

p(a,b) is the joint probability of the values of the data sets p and q found for the separation ϑ.

MI measures the degree of interdependence which is bounded below by complete independence and bounded above by one-to-one mapping.

Due to the limited number of pairs we only found stable results for coarsely binned p(a,b), namely 10 x 10 bins for WMAP vs. HEAO and 150 x 150 bins for WMAP ns. NVSS.

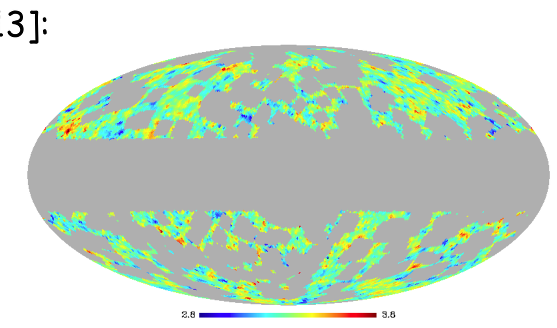
Error Estimations Monte Carlo Error (MCE) and Resampling Error (RSE)

MCE: We generate 500 Monte Carlo simulations of CMB anisotropy maps. Specifically, we simulate maps of the co-added V and W bands including noise and beam-effects of the of WMAP satellite. Further, the WMAP 7yr best fit power spectrum was used.

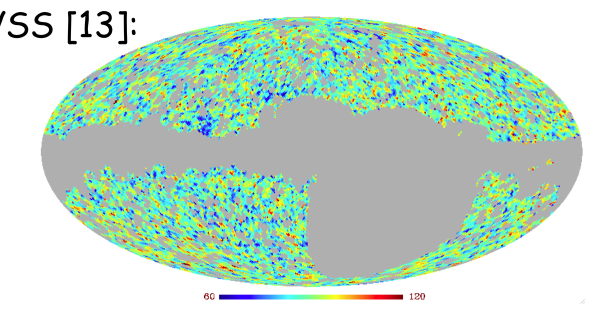
RSE: In addition, we calculated errors estimated from a resampling of LSS maps. These resampled maps are obtained by a random shuffle of the density values. 500 realisations of resampled LSS maps are generated.

The Data Sets

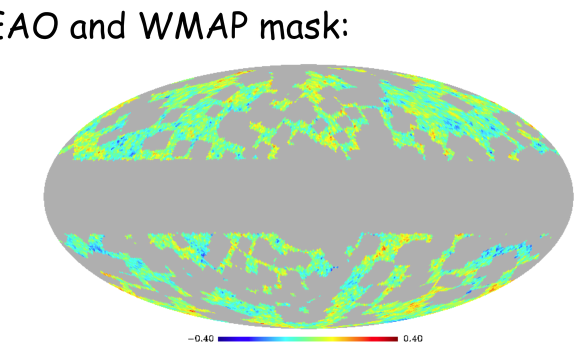
HEAO [13]:



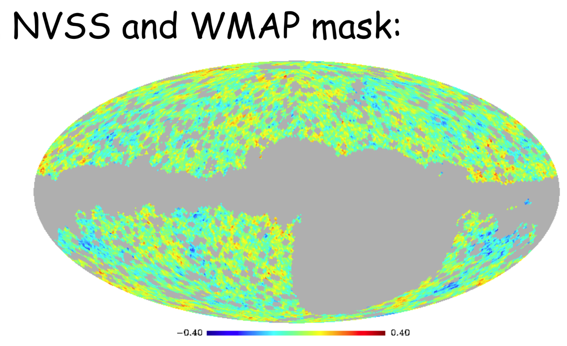
NVSS [13]:



CMB with HEAO and WMAP mask:

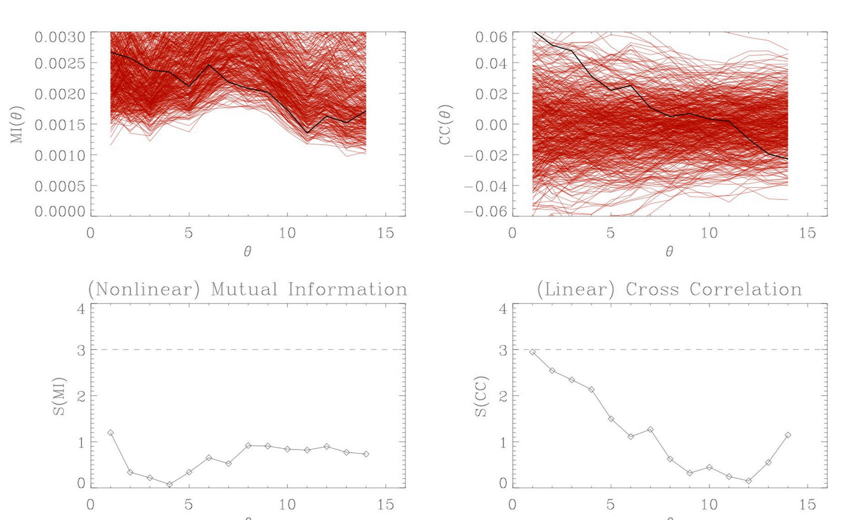


CMB with NVSS and WMAP mask:

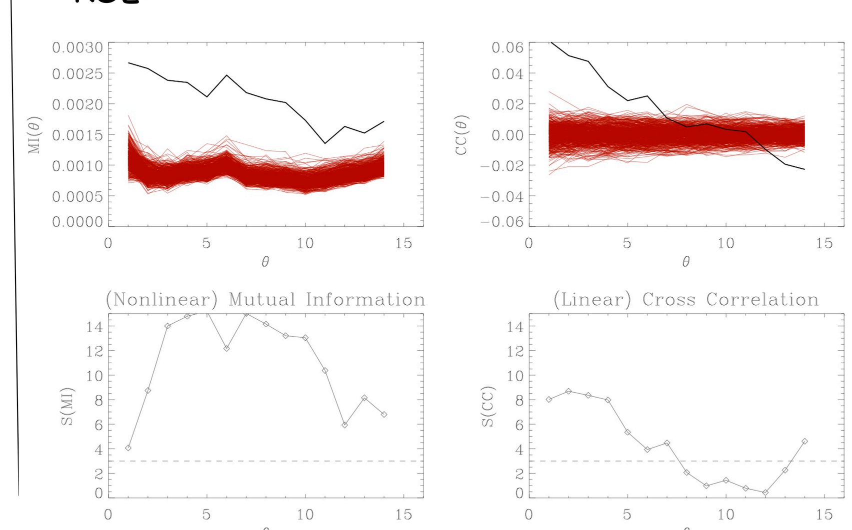


Results: WMAP - HEAO

MCE:



RSE:



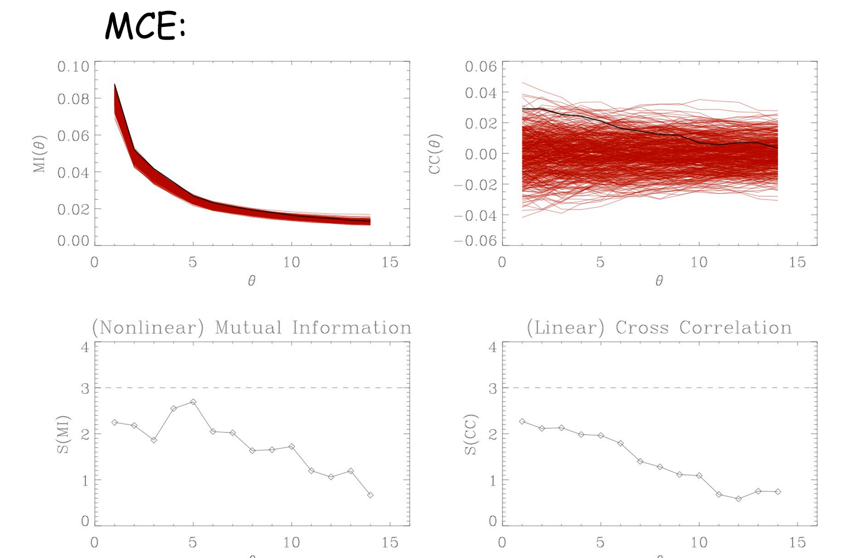
$$S(X) = \frac{X(real) - \langle X(simulation/resample)\rangle}{\sigma(X(simulation/resample))},$$

$$X = MI(\vartheta), CC(\vartheta)$$

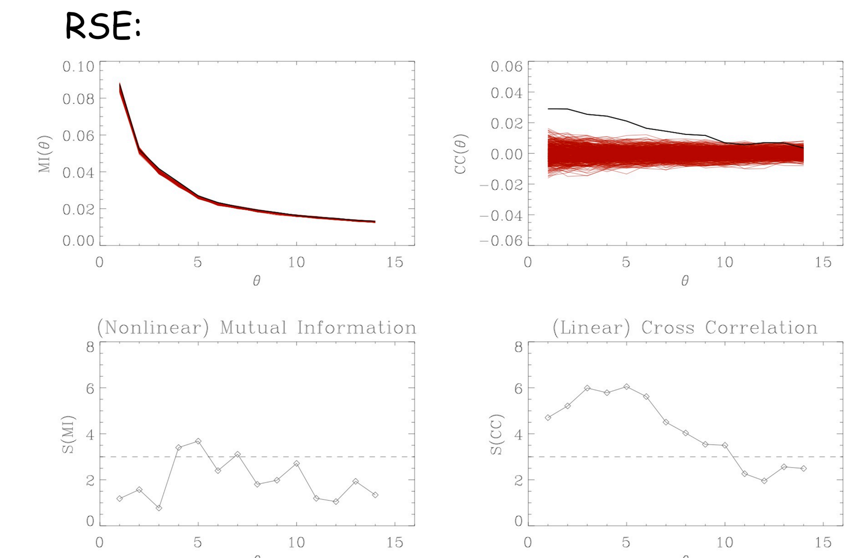
CC shows larger effect than MI for MCE, while a large deviation for MI is found for the RSE case.

Results: WMAP - NVSS

MCE:



RSE:



MI shows larger effect than CC for MCE, while a large deviation is only found for CC for the RSE case.

Conclusions - Outlook

- We realised the full simulation and analysis pipeline for cross-correlating CMB data with traces of the LSS using different similarity measures.
- Besides the linear cross-correlation we calculateded the mutual information as a novel nonlinear similarity measure.
- Although the calculation of the mutual information is still far from being optimised, we achieve similar and also sometimes significantly better results as with the linear cross-correlation.
- In the cases studied so far, CCF and MI show somewhat complementary behaviour.
- MI thus represents a versatile similarity measure for further constraing dark energy parameters via the ISW effect.
- New and significantly improved CMB (PLANCK) and LSS (e.g. eROSITA, EUCLID) data will allow for a full exploitation of the possibilities of MI.

References

[1] R. K. Sachs, A. M. Wolfe, ApJ, 147, 73 (1967).
[2] M. J. Rees, D. W. Sciama, Nature, 217, 511 (1968).
[3] R. G. Crittenden, N. Turok, PRL, 79, 575 (1996).
[4] M. R. Nolta et al., ApJ, 608, 10 (2004).
[5] S. Boughn, R. G. Crittenden, Nature, 472, 45 (2004).
[6] P. Fosalba, E. Gaztanaga, F. J. Castander, ApJ, 597, L89 (2003).
[7] T. Giannantonio et al., PRD, 77, 123520 (2008).
[8] M. Lopez-Corredoira, F. Sylos Labini, J. Betancourt-Rijo, A & A, 513, A3 (2010).
[9] C. Hernandez-Monteagudo, A & A, 520, A101 (2010).
[10] T. M. Cover, J. A. Thomas, *Elements of Information Theory*, Wiley, New York (1991).
[11] B. Zitova, J. Flusser, Image and Vision Computing, 21, 977 (2003).
[12] J. P. W. Pluim, J. B. A. Maintz, M. A. Viergever, IEEE Trans. Med. Img., 22, 986 (2003).
[13] Thanks to T. Giannantonio for providing us with the LSS data.