Stochastic modelling and Statistical Analysis of Cosmic Microwave Background Data

Ravindi Nanayakkara

Supervisor: Associate Professor Andriy Olenko

Co-Supervisor: Professor Philip Broadbridge

La Trobe University, Australia

Young Statisticians Showcase
September 24, 2019

Overview

Overview

Overview

Overview

1. Introduction to CMB data
2. Random Fields on a Sphere
3. Multifractality and the Rényi function
4. Estimating Multifractionality
5. Future Work

Overview

Introduction to CMB data

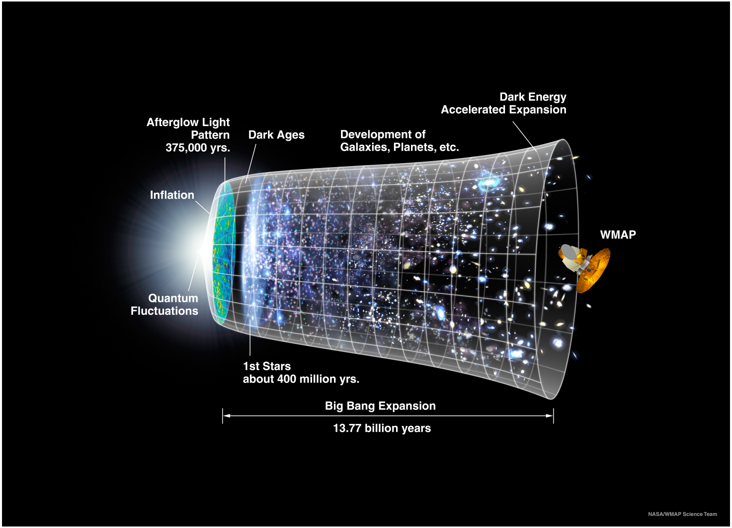


Image credit: NASA / WMAP Science Team

Cosmological Background

- The Cosmic Microwave Background(CMB) is the radiation from the universe since 380,000 years from its birth.
- In Big Bang cosmology, the CMB is an electromagnetic radiation residue from its earliest stage.
- The CMB depicts variations which correspond to different regions and represent the roots for all future formation including the solar system, stars and galaxies in the present world.
- The unforeseen discovery of the CMB was done by Arno Penzias and Robert Wilson who were American radio astronomers.

LA TROBE UNIVERSITY

Cosmological Background

- Earlier, the universe was very hot and dense in nature.
- After the big bang, the universe is expanding and cooling down and had been possible for the atoms to reformulate again after around 400,000 years of its life.
- This phenomenon is known as **Epoch of combination** and since that time photons have been able to move freely escaping from the opaque of the early universe.
- The first light which eliminated from this process is known as the cosmic microwave background.

LA TROBE UNIVERSITY

Missions

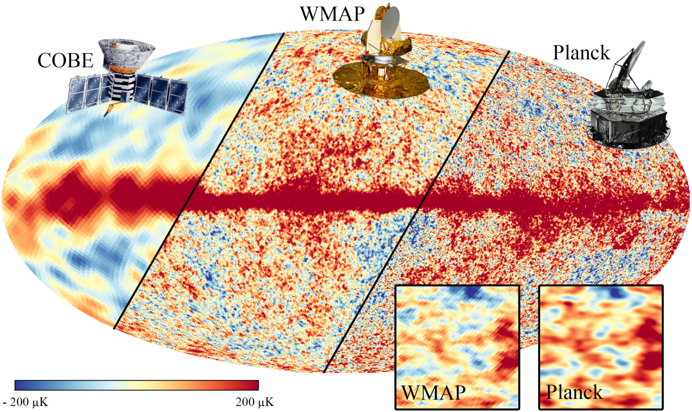


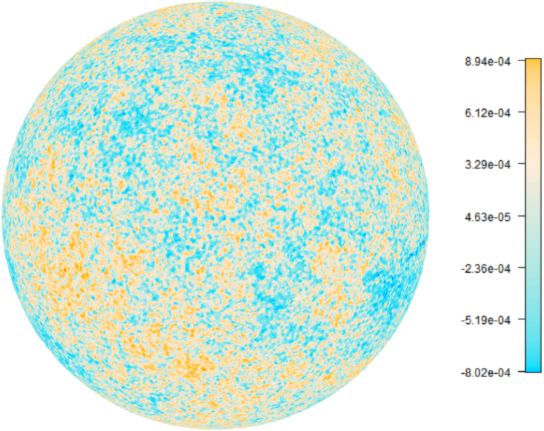
Image credit: https://jgudmunds.wordpress.com

Planck Mission

- In the year 2009, the European Space Agency launched the mission Planck in order to study the CMB thoroughly.
- The frequency range captured by the Planck mission is much wider and its sensitivity is higher than the previous missions of Cosmic Background Explorer(COBE) and Wilkinson Microwave Anisotropy Probe (WMAP).
- Current CMB data are at 5 arc-minutes resolution on the sphere.
- Contains 50,331,648 data collected by Planck mission.

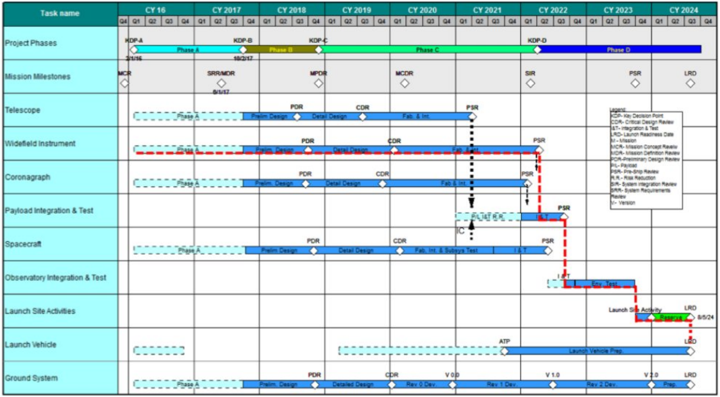
What does CMB data look like?

- Obtained using the newly developed **rcosmo** package.



LA TROBE UNIVERSITY

Next Generation Missions



Next Generation Explorer: CMB-S4

(Simons Foundations, NSF and US Department of Energy)

LA TROBE UNIVERSITY

CMB data are available as FITS(Flexible Image Transport System) files stored in HEALPix (Hierarchical Equal Area Isolatitude Pixelation). Each pixel describes distinct location, intensity, polarisation and other CMB attributes.

The HEALPix format has numerous advantages, compared to other spherical data representations:

- equal area pixels,
- hierarchical tessellations of the sphere,
- iso-latitude rings of pixels.

It is used for an efficient organization of spherical data in a computer memory and providing fast spherical harmonic transforms, search and numerical analysis of spherical data.

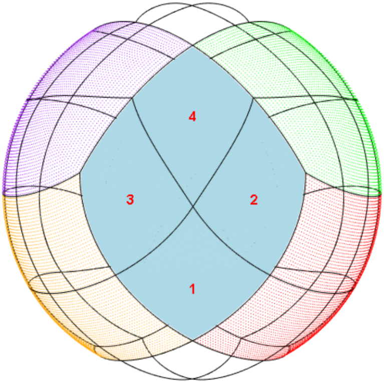
What does CMB Data Frame look like?

```
> df1<-coords(cmbdf, new.coords = "cartesian")
> df1
A CMBDataFrame
# A tibble: 12,582,912 x 4
       x     y        z          I
   <dbl> <dbl>    <dbl>      <dbl>
 1 0.707 0.707 0.000651 -0.0000920
 2 0.707 0.708 0.00130  -0.0000804
 3 0.708 0.707 0.00130  -0.0000899
 4 0.707 0.707 0.00195  -0.0000771
 5 0.706 0.708 0.00195  -0.0000701
 6 0.705 0.709 0.00260  -0.0000606
 7 0.707 0.708 0.00260  -0.0000663
 8 0.706 0.708 0.00326  -0.0000569
 9 0.708 0.706 0.00195  -0.0000872
10 0.708 0.707 0.00260  -0.0000728
# ... with 12,582,902 more rows
```

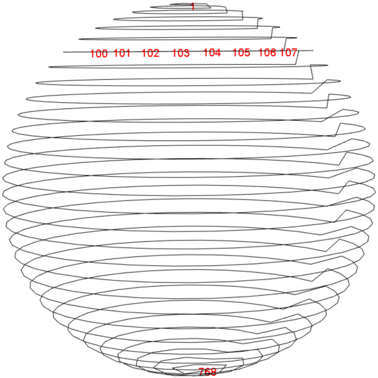
Figure: Cartesian Coordinates View

```
> df2<-coords(cmbdf, new.coords = "spherical")
> df2
A CMBDataFrame
# A tibble: 12,582,912 x 3
   theta   phi          I
   <dbl> <dbl>      <dbl>
 1  1.57 0.785 -0.0000920
 2  1.57 0.786 -0.0000804
 3  1.57 0.785 -0.0000899
 4  1.57 0.785 -0.0000771
 5  1.57 0.787 -0.0000701
 6  1.57 0.788 -0.0000606
 7  1.57 0.786 -0.0000663
 8  1.57 0.787 -0.0000569
 9  1.57 0.784 -0.0000872
10  1.57 0.785 -0.0000728
# ... with 12,582,902 more rows
```

Figure: Spherical Coordinates View



(a) HEALPix nested ordering



(b) HEALPix ring ordering

Random Fields on a Sphere

- The spherical surface in $\mathbb{R}^3$ (as a two-dimensional mainfold) with a given radius $r > 0$ is

$$s(r) = \{x \in \mathbb{R}^3 : \|x\| = r\}$$

- **Statistical model:** CMB can be viewed as a single realization of a random field on a sphere.

- A spherical random field denoted by, $\xi = \{\xi(r, \theta, \varphi) : 0 \leq \theta \leq \pi, 0 \leq \varphi \leq 2\pi, r > 0\}$ is a stochastic function defined on the sphere $s(r)$.

- We consider a real-valued spherical field such that its covariance function depends on the geodesic (or angular) distance between the two points on the sphere.

Random Fields on a Sphere

Definition 1

A random field $\xi(x), x \in T$ where $T \subseteq \mathbb{R}^n$ is called Gaussian if for any $x^{(1)}, ..., x^{(r)} \in T$, the joint distribution of random variables $(\xi(x^{(1)}), ..., \xi(x^{(r)}))'$ are Gaussian.

Definition 2

A real random field $\xi(u), u \in s(1)$ with $E[\xi^2(u)] < \infty$ and $E[\xi(u)] = 0$ is called isotropic on a sphere if,

$$E[\xi(u)\xi(v)] = B(\cos\theta),$$

$u, v \in s(1)$, depends only on the geodesic (angular) distance $\cos\theta$ between u and v.

LA TROBE UNIVERSITY

Random Fields on a Sphere

- We get the following representation of covariance functions of the isotropic random field on the sphere.

$$B(\cos\theta) = \frac{1}{|s(1)|} \sum_{m=0}^{\infty} b_m h(m,n) \frac{C_m^v(\cos\theta)}{C_m^v(1)}, \qquad (1)$$

where $C_m^v, m = 0, 1, \ldots$ are the Gegenbauer polynomials.

- The representation for $B(\cos\theta)$ implies the spectral decomposition of the isotropic field on the sphere $\xi(u), u \in s(1)$, itself. That is, there exists a real-valued sequence of random variables η_m^l such that

$$\xi(u) = \sum_{m=0}^{\infty} \sum_{l=1}^{h(m,n)} S_m^l(u) \eta_m^l, \qquad (2)$$

where S_m^l are spherical harmonics, and

$$E[\eta_m^l] = 0, \qquad E[\eta_m^l \eta_p^q] = b_m \delta_m^p \delta_l^q.$$

LA TROBE UNIVERSITY

Multifractality

- The concept of multifractality initially emerged in the context of measures where Mandelbrot showed the significance of scaling relations in the setting of turbulence modelling.
- Multifractal analysis compromises with the local scaling characteristics of functions distinguished by Hausdorff dimension of classes of points which has the identical Hölder exponent.
- Rényi function which is also known as the deterministic partition function plays a key role in multifractal analysis.
- Leonenko and Shieh (2013) computed the Rényi function for three models of the multifractal random fields and showed some major schemes for the Rényi function that reveal the multifractality of homogeneous and isotropic data.

The Rényi function

- The Rényi function of μ on $s_2(1)$ can be defined as

$$T(q) = \liminf_{m \to \infty} \frac{\log_2 E \sum_l \mu(S_l^{(m)})^q}{log_2 |S_l^{(m)}|}, \quad (3)$$

  where $\{S_l^{(m)}, l, m\},\ l = 0, 1, ..., 2^m - 1$ and $m = 1, 2, ...,$ is the mesh formed by the mth level dyadic decomposition of the spherical surface $s_2(1)$.

- For the CMB data analysis, we use $\mu(S_l^m)$ which equals the cumulative CMB intensity over S_l^m.

- The statistical estimator $\widehat{T(q)}$ is obtained using the equation (3) and for large values of m.

Known results about the Rényi function

Model 1: Log-Normal Scenario

$$T(q) = q\left(1 + \frac{\sigma_Y^2}{4\log b}\right) - q^2\left(\frac{\sigma_Y^2}{4\log b}\right) - 1, \quad q \in [1,2].$$

Model 2: Log-Gamma Scenario

$$T(q) = q\left(1 + \frac{1}{n\log b}\log\frac{1}{(1-\frac{1}{\lambda})^{\beta}}\right) + \left(\frac{\beta}{n\log b}\right)\log\left(1-\frac{q}{\lambda}\right) - 1,$$

where $q \in Q = \{0 < q < \lambda, \lambda > 2\} \cap [1,2] \cap L_{\beta,\lambda}$.

Model 3: Log-Negative-Inverted-Gamma Scenario

$$T(q) = q\left(1 + \frac{c_U \log\frac{2\lambda^{\beta}}{\Gamma(\beta)}}{n\log b}\right) - \frac{1}{n\log b}\log\{q^{\beta/2}K_{\beta}(2\sqrt{q\lambda})\} - \left(1 + \frac{\log\frac{2\lambda^{\beta/2}}{\Gamma(\beta)}}{n\log b}\right),$$

where $q \in Q = [1,2] \cap L_{\beta,\lambda}$.

Computing $\widehat{T(q)}$ for real data

- The estimator of the Rényi function was calculated for real cosmological data.
- First the estimator was computed for the whole sky and then for different window sizes located at different places of the sphere.

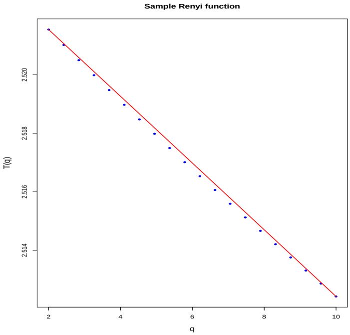


Figure: For the whole sky

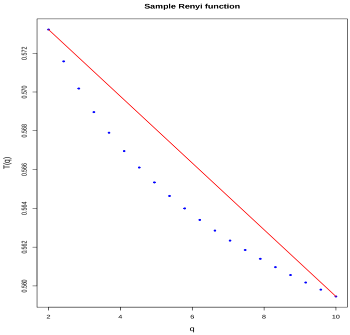


Figure: For a very small window

LA TROBE UNIVERSITY

Computing $\widehat{T(q)}$ for real data

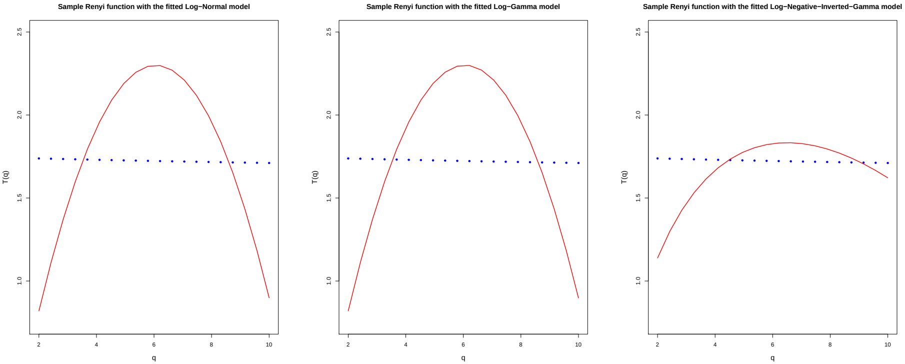


Figure: Plots of Sample Rényi function with the fitted models

Simulation of Random Fields

- There are numerous cases with no explicit expression form for the Rényi function.
- Random field simulations can be used to obtain and study random fields.
- One can simulate random fields for different theoretical models and obtain empirical Rényi function.

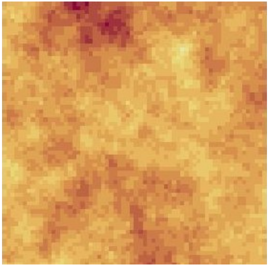


Figure: Realization of a random field with exponential covariance model

Computing $\widehat{T(q)}$ for real and simulated data

- $\widehat{T(q)}$ was computed for a large window area in both real CMB data and simulated data from the exponential covariance model.

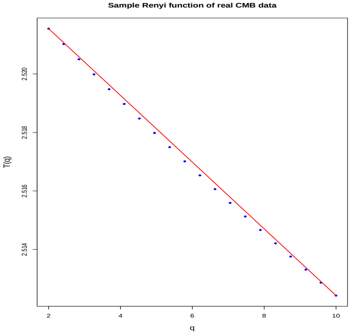


Figure: For real data

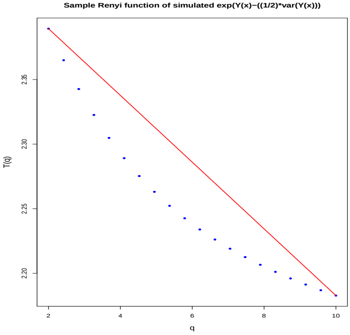


Figure: For simulated data

Hölder Exponent

- Used to measure roughness in a rigorous mathematical way.
- If the data are multifractional, H(t) $\neq$ constant, t is a location on S^2.

Definition 3

Let $Y(\cdot,\cdot)$'s be observations collected on a Healpix grid on the sphere. We define the Hölder exponent as,

$$H(t) = \lim_{N\to\infty} \frac{1}{2}\left(d(1-\gamma) - \frac{log V_N(t)}{log N}\right), \tag{4}$$

where

$$V_N(t) = \sum_{p \in V_N(t)} \left(\sum_{k\in F} d_k Y\left(\frac{p+k}{N}, \frac{p+k}{N}\right)\right)^2.$$

- The estimator $\widehat{H(t)}$ is obtained using the equation (4) for large N.

Computing Hölder Exponent for real CMB data

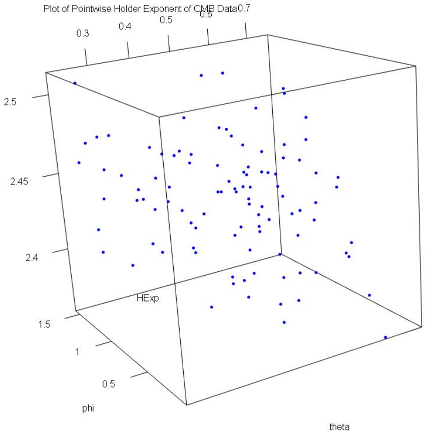


Figure: Plot of $\widehat{H(t)}$ for a sample of cold and warm areas

Conclusions

- According to the computed $\widehat{T(q)}$, we don't have multifractality when considered the real CMB data.
- Although we considered different window sizes of the sphere, we don't have evidence to suggest that we have multifractality since the deviation is not substantial.
- When considered a very small window of the sphere, the sample Rényi function gives a parabolic shape depicting that we have a very minor multifractality in very small scales of the sphere.
- It seems that the CMB data are unifractal.
- According to the computed pointwise Hölder Exponent values, we have a positive sign indicating multifractionality in CMB data.

Future Work

- To investigate multifractality:
  - To obtain explicit Rényi functions for different statistical models.
  - To develop R package for cases where theoretical models and Rényi functions are unknown.

- To study multifractionality of cosmic microwave background data.

- To study statistical properties of spherical random fields(test for non-Gaussianity).

- To investigate high frequency asymptotics for angular spectrum.

References

- Anh, V. Broadbridge, P. Olenko, A. & Wang, Y. (2018). On approximation for fractional stochastic partial differential equations on the sphere. Stochastic Environmental Research and Risk Assessment. 32:2585-2603.
- Ayache, A. & Véhel, L. (2004). On the identification of the pointwise Hölder exponent of the generalized multifractional Brownian motion. Stochastic Processes and their Applications. 0304-4149
- Broadbridge, P. Kolesnik, A. Leonenko, N. & Olenko, A. (2019). Random spherical hyperbolic diffusion. will appear in Journal of Statistical Physics, arXiv:1904.12243.
- Fryer, D. Olenko, A. Li, M. & Wang, Y. (2019). rcosmo: R Cosmic Microwave Background Data Analysis. R package version: 1.1.1.
- Fryer, D. Li, M. & Olenko, A. (2019). rcosmo: R Package for Analysis of Spherical, HEALPix and Cosmological Data, arXiv preprint arXiv:1907.05648.
- Leonenko, N. & Shieh, N.-R. (2013). Rényi function for multifractal random fields. Fractals, 21(02), 1350009.
- NASA/IPAC INFRARED SCIENCE ARCHIVE. (2019). Retrieved September 11, 2019, from https://irsa.ipac.caltech.edu/data/Planck/release_2.

Thank you

Statistical Society of Australia

LA TROBE UNIVERSITY