NonGaussian Detection:
Comparing Performances of Different Statistical Tests

Jiashun Jin

Statistics Department
Purdue University
Collaborators

Alphabetically:

<table>
<thead>
<tr>
<th>Name</th>
<th>Institution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nabila Aghanim</td>
<td>Université Paris Sud</td>
</tr>
<tr>
<td>Laura Cayon</td>
<td>Purdue University</td>
</tr>
<tr>
<td>David Donoho</td>
<td>Stanford University</td>
</tr>
<tr>
<td>Olivier Forni</td>
<td>Université Paris Sud</td>
</tr>
<tr>
<td>Jeau-Luc Starck</td>
<td>Service d’Astrophysique</td>
</tr>
<tr>
<td>Anna Treaster</td>
<td>Purdue University</td>
</tr>
</tbody>
</table>
Cosmic Microwave Background (CMB)

CMB:

- Oldest light in the universe, a direct link to early universe
- A relic of radiation when the universe \( \approx 380,000 \) years old
- An almost perfect black body at a temperature \( \approx 2.725 \) Kelvin
Figure 1: Small angular fluctuations in CMB are predicted as the imprints of initial densities perturbation which gave rise to large scale structures today. Red color: strong emission from the Milky way.
Why study CMB:

- Discriminate different models for early universe
- Disentangle nonGaussianity of one source from another
- Understand early universe: how does it evolve into the large scale galaxies today
Wavelet Approach of nonGaussian Detection

- Standard inflation model predicts that the CMB is Gaussian
- Other models or secondary effects have nonGaussian signatures
- Wavelet transform is a powerful tool for detect nonGaussian signature
  - isotropic à trous algorithm (Starck et al. 1998)
  - bi-orthogonal wavelet transform
For Today

- Consider $n$ transform coefficients of CMB: $X_i$
- Test the hypothesis:
  \[ H_0 : X_i \overset{iid}{\sim} N(0, 1), \quad 1 \leq i \leq n \]

Goal. By comparing different statistics:

- learn the strength and weakness of different tests
- look for the optimal tests in idealized cases
Wavelet Based nonGaussian Tests

1. Excess kurtosis ($\kappa$):

$$\kappa(X_1, \ldots, X_n) = \sum_i [X_i^4 - 3]$$

2. Maximum (Max):

$$\text{Max}(X_1, \ldots, X_n) = \max\{|X_1|, |X_2|, \ldots, |X_n|\}$$

3. Higher Criticism (HC)
   - mentioned in passing by Tukey (1976)
   - showed to be optimal in resolving a very subtle testing problem
Only need $p$-values to implement HC

Obtain individual p-values by:

$$p_i = P\{|N(0, 1)| \geq |X_i|\}$$

• sort p-values:

$$p(1) < p(2) \ldots < p(n)$$

• calculate $i^{th}$ z-score:

$$HC_{n,i} = \sqrt{n}\left[\frac{i/n - p(i)}{\sqrt{p(i)(1 - p(i))}}\right]$$

• take maximum:

$$HC_n^* = \max\{1 \leq i \leq n\} \; HC_{n,i}$$
Heuristic Comparison

A test only sensitive to certain type of nonGaussianity:

- $\kappa$: deviation of 4-th moment from Gaussian
- Max: unusual behavior of very large observations
- HC:
  - unusual behavior of extreme values
  - unusual behavior of moderately large values
Example I: Detecting Cosmic String (CS)

- given superposed image:
  \[ \sqrt{1 - \lambda} \cdot CMB + \sqrt{\lambda} \cdot CS, \quad \lambda \approx 0 \]

- test: \( \lambda = 0 \) vs. \( \lambda > 0 \)
• equivalent to test:

\[ H_0 : X_i = z_i, \quad 1 \leq i \leq n, \]

\[ H_1^{(n)} : X_i = \sqrt{1 - \lambda} \cdot z_i + \sqrt{\lambda} \cdot w_i, \quad 1 \leq i \leq n. \]

- \( z_i \) \( \overset{i.i.d.}{\sim} N(0, 1) \): wavelet coefficients of CMB
- \( w_i \) \( \overset{i.i.d.}{\sim} W \): wavelet coefficients of CS
- \( W \) unknown, but symmetric and heavy tail
Calibrations

Need careful calibrations for subtle analysis:

- increasing amount of data are offset by increasingly challenges:
  \[ \lambda = \lambda_n = n^{-r}, \quad 0 < r < 1 \]

- \( W \): symmetric and has a \textit{power-law} tail with index \( \alpha \):
  \[ \lim_{x \to \infty} x^\alpha p\{|W| > x\} = C_\alpha, \quad C_\alpha: \text{ constant} \]

**Question**: Fixed \((r, \alpha)\) and let \( n \to \infty \), what is the optimal test?
Interpretation

\( \alpha = 8 \) is the separating line:

- \( E[W^8] < \infty \): Kurtosis is better
  - \( W \) has a relatively thin tail, nonGaussianity affects the bulk of the data
  - best tests: tests based on moments

- \( E[W^8] = \infty \): HC/Max is better
  - \( W \) has relatively heavy tail
  - nonGaussianity has little effect on the bulk of data, but large effect on extreme values and moderately large values
  - best tests: tests based on data tails


Estimating $\alpha$

- Analysis supports the power-law tail assumption of $W$
- A classical estimator for $\alpha$ is the Hill’s estimator \((Ann. Statist. 1975)\)
- Implementation of Hill’s estimator:

  \[ \hat{\alpha} \approx 6.1, \quad \text{std}(\hat{\alpha}) \approx 0.9, \]
The finer resolution of the image, the larger the $n$

Need large $n$
- to see the real advantage of $HC$
- better answer whether $\alpha < 8$ or $\alpha > 8$
- better answer which of HC and Kurtosis is better
Example II: WMAP First Year Data

http://map.gsfc.nasa.gov/

- WMAP radiometers observe at 5 frequency bands with one or more receivers: K (1), Ka (1), Q (2), V (2), W (4).
- WMAP team suggested use the weighted average of Q-V-W bands (8 receivers)
- Foreground cleaned
- Mask added: strong emission of Milky way etc..
- Downgraded from $n_{side} = 512$ to $n_{side} = 256$: measurement noise dominant in the smallest scale
1. Generate 5,000 simulated Gaussian maps of CMB.

2. For WMAP and each simulated map:
   - Use Spherical Mexican Hat Wavelets (SMHW):
     2-D-spherical wavelets
   - Normalize the wavelet coefficients
   - Apply kurtosis, Max, and HC to wavelet coefficients
Figure 2: Test scores on WMAP and 67%, 95%, and 99% confidence regions on 5,000 simulated CMB maps.
Comparisons of Different Statistics

1. Almost equally powerful for detection, kurtosis is slightly better
   - Define empirical confidence of detection:
     \[
     \frac{\#\{\text{test scores based on simulations} \leq \text{score on WMAP}\}}{5000}
     \]
     - Kurtosis: 99.7%
     - HC: 99.46%
     - Max: 99.44%

2. Higher Criticism: automatically identify a tiny portion data as the source of nonGaussianity
A Second Look of HC

- $H C_\alpha^* = \max_{0 < \alpha < 1} \{ H C_{n, \alpha} \}$,

  $$H C_{n, \alpha} = \sqrt{n} \cdot \left[ \text{Fraction Significant at } \alpha \right] - \alpha / \sqrt{\alpha(1 - \alpha)}$$

- $H C_{n, \alpha} \gg 1$ implies nonGaussianity

---

Figure 3: Plot of $H C_{n, \alpha}$ versus $(1 - \alpha)$ for wavelet coefficients of WMAP at Scale 9
Source for nonGaussianity: the Ring?

Figure 4: Bottom: Ring centered at $(209^\circ, -57^\circ)$
Figure 5: Kurtosis, Max, and HC after the ring removed from the WMAP. No detection of non-Gaussianity at the level $\geq 90\%$. 

Values of $K$, $\text{Max}$, and $\text{HC}$, and the 68%, 95%, and 99% confidence regions. 

Figure 7: The dashed, dotted-dashed and solid lines correspond to the 68%, 95%, and 99% confidence regions.
Comparison to Other Works

• Some work has $\geq 99\%$ confidence of nonGaussian detection, and some work identify the cold spot centered at $(209^\circ, -57^\circ)$.

• Our contribution:
  – Add new statistics to nonGaussian detection: HC and Max
  – Almost equally powerful as kurtosis
  – HC offers automatical identification of a tiny portion of data as the source of nonGaussianity
  – Suggests the ring, not the whole spot, is the source for nonGaussianity (Viela et al. 2004, Cruz et al. 2005)
Take Home Messages

- nonGaussian detection in CMB is an exciting field
- Higher Criticism is a promising new detection tool, adds more discussion to nonGaussian detection
- better answer is expected in future study with a larger $n$