Non-Gaussianity from extragalactic point sources

F Lacasa
Institut d’Astrophysique Spatiale (IAS), Orsay, France
E-mail: fabien.lacasa@ias.u-psud.fr

Abstract. We have described briefly the population of compact sources appearing at CMB frequencies, and studied their non-Gaussianity using publicly available full-sky simulations. Introducing a parametrization permitting to visualise efficiently the bispectrum, we have described the configuration and scale dependence of the bispectrum of radio and IR point sources, as well as its frequency dependence, and shown that it is well fitted by an analytical prescription. We have shown further that the clustering of IR sources increases their non-Gaussianity by several orders of magnitude, and that their bispectrum peaks in the squeezed triangles. Examining the impact of these sources on primordial non-Gaussianity estimation, we have found that the radio sources yield an important positive bias to local $f_{NL}$ estimation at low frequencies, but this bias is efficiently reduced by masking detectable sources. On the other hand, IR sources produce a negative bias at high frequencies, which is not dimmed by the masking, as their clustering is dominated by faint sources.

1. Introduction
The Cosmic Microwave Background is a powerful probe of the early and late universe, it is the dominant signal on the sky around 100 GHz, and its observation has contributed to the establishment of the standard model of cosmology, the latest surveys being WMAP ACT and SPT. However, in the 30 – 350 GHz frequency window, other signals are present, and they become important at small angular scales, where the CMB plummets due to Silk damping, and in non-Gaussian part of correlation function, as the CMB is close to Gaussian. This is of special importance, e.g., for Planck, which covers a large frequency range at high angular resolution. There are principally three populations of extragalactic sources:

(i) Galaxy clusters, in which the electrons of the hot ionised gas scatter off the CMB photons leaving a distinct spectral signature (Sunyaev-Zeldovich effect).
(ii) Radio loud galaxies with an Active Galactic Nucleus and emitting through synchrotron and free-free processes.
(iii) Dusty star forming galaxies, where the UV emission from stars heats the dust which consequently re-emits in the infrared domain.

We have concentrated on the spatial distribution of this two former populations. Radio sources can be considered randomly distributed on the sky and can, hence, be modelled as a white-noise, entirely characterized by the number counts of the sources; the harmonic correlation function (power spectrum, bispectrum, etc.) are then flat and related to the corresponding moment of the 1-point distribution. On the other hand, IR sources are highly clustered in dark matter halos and have, hence, non-trivial correlation functions; there are models and observations of their
number counts and power spectrum, but all higher order moments are theoretically necessary to describe the field.

The study of higher order moments also has an interest for the CMB, as it can break the degeneracy between the primordial processes generating the cosmological perturbations, e.g., the inflation models. While most models predict similar power spectra, they may be distinguished at the 3-point or higher level; as an example, standard single-field models predict undetectably small non-Gaussianity, while multifield models, non-canonical kinetic terms or vacuum initial condition, etc. produce potentially detectable non-Gaussianity with distinct shapes. As of today, no definite deviation from Gaussianity has been found in the CMB, even if there is a hint of NG of the ‘local’ shape in WMAP-7 data [1]. The local shape of non-Gaussianity is produced generically by several inflation models and takes the form:

\[ \Phi(x) = \Phi_G(x) + f_{NL} (\Phi_G(x)^2 - \langle \Phi_G(x)^2 \rangle) \]  

(1)

where \( \Phi \) is the Bardeen potential, and the subscript G denotes the Gaussian linear part.

To study the distribution of radio and IR galaxies, we have made use of publicly available full-sky simulated maps of these sources [2] between 30 and 350 GHz at high angular resolution made through N-body dark matter simulations. The results presented in this paper are explained in greater details in [3].

2. Bispectrum

A Gaussian field on the sphere is entirely characterised by its mean and 2-point correlation function (or power spectrum in harmonic space), and the 3-point correlation function is, thus, the lowest order indicator of non-Gaussianity, and possesses a priori more signal to noise ratio than higher order moments. The bispectrum is the 3-point correlation function in harmonic space given as:

\[ \langle a_{\ell_1 m_1} a_{\ell_2 m_2} a_{\ell_3 m_3} \rangle = C_{\ell_1 \ell_2 \ell_3}^{m_1 m_2 m_3} \times b_{\ell_1 \ell_2 \ell_3} \]  

with \( C_{\ell_1 \ell_2 \ell_3}^{m_1 m_2 m_3} = \int d^2 n Y_{\ell_1 m_1}(n) Y_{\ell_2 m_2}(n) Y_{\ell_3 m_3}(n) \).  

(2)

A full calculation of the bispectrum at WMAP or Planck resolution is computationally very intensive, and so one usually resort to binning multipoles. The bispectrum is invariant under permutation of \((\ell_1, \ell_2, \ell_3)\), and it is a function of the triangle shape only. Three triangle shapes are of particular importance: squeezed (one side is much smaller than the other two sides), equilateral, and folded (flat isosceles).

To plot a bispectrum in an efficient way, e.g., without redundancy of information, we have introduced a parametrization accounting for this permutation invariance, and defined three parameters \((P, F, S)\) based on the elementary symmetric polynomials \(\sigma_1 = \ell_1 + \ell_2 + \ell_3\), \(\sigma_2 = \ell_1 \ell_2 + \ell_1 \ell_3 + \ell_2 \ell_3\), and \(\sigma_3 = \ell_1 \ell_2 \ell_3\). The locations of the different triangles of constant perimeter \(P\) in the \((F, S)\) plan is shown in Figure 1. One can then plot a bispectrum by making slices of perimeters and colour-coding the value of the bispectrum.

The local type primordial non-Gaussianity defined in Eq. (1) leaves a characteristic imprint on the CMB, which can be detected at the 3-point level. The bispectrum, thus, generated peaks in the squeezed triangles, and because it has a separable form, it has been shown [4] that a fast estimator for its amplitude factor \(f_{NL}\) can be built, which does not need the prohibitive computation of the full bispectrum. This, commonly called \textit{KSW estimator} is optimal in the
sense that it minimizes the $\chi^2$ of the fit of the observed bispectrum to the local bispectrum by:

$$\hat{f}_{NL} = \sigma^2(f_{NL}) \sum_{\ell_1 \leq \ell_2 \leq \ell_3} N_{\ell_1,\ell_2,\ell_3} \frac{b_{\text{obs}}^{\ell_1,\ell_2,\ell_3}}{C_{\ell_1} C_{\ell_2} C_{\ell_3}},$$

(3)

with

$$\sigma^2(f_{NL}) = \sum_{\ell_1 \leq \ell_2 \leq \ell_3} N_{\ell_1,\ell_2,\ell_3} \left( \frac{b_{\text{loc}}^{\ell_1,\ell_2,\ell_3}}{C_{\ell_1} C_{\ell_2} C_{\ell_3}} \right)^2,$$

(4)

where $N_{\ell_1,\ell_2,\ell_3} = \frac{(2\ell_1+1)(2\ell_2+1)(2\ell_3+1)}{4\pi} \left( \begin{array}{ccc} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{array} \right)^2$ is the number of such configurations on the sky.

3. Point sources

As stated in the introduction, radio sources are randomly distributed on the sky, and so they produce a constant bispectrum $b_{ps} \propto S^3 dN/dS$, but as higher frequencies are surveyed, e.g., with Planck, the IR population becomes of growing importance, and one must account for the superposition of these signals. We have used a prescription, originally proposed by [5], and developed it to a full sky multi-population and analytical prescription; its assumptions are that radio and IR populations are independent, and that IR distribution can be modelled as a white-noise convolved by the known power spectrum. Analytically, at the bispectrum level, it yields:

$$b_{\ell_1,\ell_2,\ell_3}^{\text{IR}} = \alpha \sqrt{C_{\ell_1} C_{\ell_2} C_{\ell_3}} \left( \begin{array}{ccc} \ell_1 & \ell_2 & \ell_3 \\ 0 & 0 & 0 \end{array} \right)^2,$$

(5)

with

$$\alpha = \frac{S^2 dN/dS}{\left( \int S^2 dN/dS dS \right)^{3/2}}.$$

We have used publicly available simulations of these two populations between 30 and 350 GHz by [2] to compare with this prescription. Figure 2 shows the observed and predicted bispectra at 350 GHz in the case of superposition of IR and radio sources. We have found a good agreement, within cosmic variance error bars, between the prescription and the observed bispectrum at all frequencies and in all cases: radio sources alone, IR alone, and radio and IR together. It is also remarkable that the clustering of IR sources increases its non-Gaussianity, compared to un-clustered case, by several order of magnitudes at large angular scales. While the radio bispectrum is flat, the IR one exhibits a configuration dependence due to clustering, which can be visualized with the parametrization described previously (see Figure 3).
Figure 2. IR+radio bispectrum at 350 GHz in some configurations. Black line is the observed bispectrum of Sehgal et al simulations, red line is the prescription, blue dashed line is the case, where radio and IR populations are assumed totally correlated.

Figure 3. IR bispectrum at 148 GHz in the parametrization

We have found that the IR bispectrum peaks in the squeezed triangles, as predicted by the prescription, with a slight flattening at high multipoles due to shot-noise, especially at the higher frequencies.

4. Contamination of primordial non-Gaussianity estimate
We have used the KSW estimator [4] described previously by Eq. (3) on the simulated maps, to estimate how much foreground signals would bias the estimation of the primordial local NG parameter \( f_{\text{NL}} \), and computed this bias for several maximum angular resolution (WMAP-like, Planck-like), with the raw maps and after masking sources with flux above the flux cut of the Planck ERCSC catalogue [6]. Figure 1 summarizes these results respectively for radio and IR sources.
Table 1. Bias from radio (left panel) and IR (right panel) to local NG parameter $f_{\text{NL}}$, depending on frequency, maximum angular resolution, and masking or not detectable sources

| $\nu$ (GHz) | without flux cut | with flux cut | without flux cut | with flux cut |
|-------------|------------------|--------------|------------------|--------------|
| $\ell_{\text{MAX}} = 50$ | -4.2, -0.025 to -0.00037, -0.00021 to -0.00027 to -0.00098 | | | |
| $\ell_{\text{MAX}} = 700$ | 3850, 2.5, 0.38, 0.21, 0.27, 0.65 | 1770, 117, 15, 9.7, 12, 30 | 1770, 117, 15, 9.7, 12, 30 |
| $\ell_{\text{MAX}} = 2048$ | 4030, 7.5, 0.31, 0.14, 0.16, 0.29 | | | |

The relative error bars of these biases are of 1.5 – 2.5% for radio sources and 3 to 7% for IR sources. We have found that for both population, the bias increases at higher maximum multipoles, as the CMB signal plummets due to Silk damping. The radio bias is positive and maximal at low frequencies, while the IR bias is negative and peaks at the higher frequencies. At a WMAP-like resolution, except at 30 GHz, these biases can be neglected compared to the expected error bars on $f_{\text{NL}}$ due to cosmic variance. However, at Planck-like resolution, these biases become non-negligible; while the radio bias is very efficiently reduced by masking detectable sources, it is not the case for IR sources, where the bias is unaffected. Indeed, the IR population, and especially the clustered galaxies, are dominated by faint sources much below the detection limit.

References
[1] Komatsu E, Smith K M, Dunkley J, Bennett C L, Gold B, Hinshaw G, Jarosik N, Larson D et al 2011 Ap. J. Supp. Ser. 192 18 [ISSN 0067-0049]
[2] Sehgal N, Bode P, Das S, Hernandez-Monteagudo C, Huffenberger K, Lin Y T, Ostriker J P and Trac H 2010 Ap. J. 709 920 [ISSN 0004-637X] [arXiv:0908.0540]
[3] Lacasa F, Aghanim N, Kunz M and Frommert M 2011 To be published in MNRAS [arXiv:1107.2251]
[4] Komatsu E, Spergel D N and Wandelt B D 2005 Ap. J. 634 14 [ISSN 0004-637X]
[5] Argueso F, GonzalezNuevo J and Toffolatti L 2003 Ap. J. 598 86 [ISSN 0004-637X]
[6] Planck Collaboration 2011 A. & A. 536 A7 [arXiv:1101.2041]