New Window on Primordial Non-Gaussianity

Enrico Pajer[1] and Matias Zaldarriaga[2]

[1]*Department of Physics, Princeton University, Princeton, New Jersey 08544, USA*
[2]*School of Natural Sciences, Institute for Advanced Study, Princeton, New Jersey 08540, USA*

We know very little about primordial curvature perturbations on scales smaller than about a Mpc. Measurements of the μ distortion of the cosmic microwave background spectrum provide the unique opportunity to probe these scales over the unexplored range from 50 to 10^4 Mpc^{-1}. This is a very clean probe, in that it relies only on well understood linear evolution. Also, just the information about the low multipoles ($l \sim 100$) of μ is necessary. We point out that correlations between μ distortion and temperature anisotropies can be used to test Gaussianity at these very small scales. In particular the μT two-point correlation is proportional to the very squeezed limit of the primordial bispectrum and hence measures f_{NL}^{loc}, while $\mu\mu$ is proportional to the primordial trispectrum and measures τ_{NL}. We present a Fisher matrix forecast of the observational constraints on f_{NL}^{loc} and stress that a cosmic variance limited experiment could in principle reach $\Delta f_{NL}^{loc} \sim \mathcal{O}(10^{-3})$.

PACS numbers: 98.80.Es, 98.70.Vc

The initial conditions of our Universe can be described in the simplest case just by specifying at some early time the probability distribution function of the adiabatic mode on a wide range of scales. Since then most of these scales have evolved in a complicated and often nonlinear way until today. Therefore, it is very hard to find accurate probes of these initial conditions, with two notable exceptions: large scale structures today, which still have not entered a fully nonlinear regime, and the anisotropies of the cosmic microwave background (CMB) radiation. Because of complicated nonlinear dynamics on the one side and Silk damping on the other, both probes are useful only at scales of order a Mpc or larger. Hence, it is very interesting to find ways to investigate and constrain scales outside of the $\{10^{-4} - 1\}$ Mpc^{-1} window, which we have explored so far.

In [1] it was shown that [2] due to the dissipation of acoustic waves, the spectral distortion of the CMB can be used to constrain the (integrated) primordial power spectrum in the approximate range $50 \lesssim k$ Mpc $\lesssim 10^4$. This mechanism provides us with the probably unique chance to probe primordial perturbations that by now have been completely erased by Silk damping and swamped by complicated gravitational dynamics. Although we know that the primordial perturbations in the cosmic microwave background–large-scale structure (CMB-LSS) window are close to Gaussian and, hence, well described by just the power spectrum, we know near to nothing about the statistical properties at much smaller scales.

In this Letter, we show that besides probing the integrated power spectrum, the CMB spectral distortion is a potentially powerful tool to constrain deviations from Gaussianity at scales $50 \lesssim k$ Mpc $\lesssim 10^4$. Correlations between the μ distortion and temperature anisotropies of the CMB provide a direct measurement of the primordial

bispectrum (in the squeezed limit), while self-correlations of μ distortion measure the primordial trispectrum. It is important to stress that this probe is particularly clean in that it only relies on well understood CMB physics and is largely unaffected by gravitational nonlinearities.

The rest of the Letter is organized as follows. We start by reviewing the relevant aspects of CMB physics. Then, following [1,3,4], we derive an analytic estimate for the μ distortion which we use to compute two-point self-correlations and cross correlation with temperature anisotropies. Finally, we present a Fisher matrix forecast together with some consideration of observational prospects and we conclude with a summary.

CMB distortion.—At early times ($z \gg 10^6$) the Universe is well described by a hot photon-baryon plasma. The number density of photons $n(\nu)$ per frequency interval is given to very high accuracy by the blackbody spectrum $n(\nu) = (e^x - 1)^{-1}$, where $x \equiv h\nu/(k_B T)$. The equation that describes the subsequent evolution of the photon number density is the full Boltzmann equation (when restricted to Compton scattering this is known as the Kompaneet's equation [5]). This equation has three interesting regimes. Before $z \simeq z_{\mu,i} \equiv 2 \times 10^6$ any energy released into the photon-baryon plasma is quickly thermalized by elastic and double Compton scattering ($e^- \gamma+ \rightarrow e^- + 2\gamma$), which are still very efficient [6]. The end result is again a blackbody spectrum with now a higher T and a larger total number of photons N. After $z_{\mu,i}$ double Compton scattering (as well as bremsstrahlung) becomes less efficient and the total number of photons is approximately frozen. For $z_{\mu,f} \lesssim z \lesssim z_{\mu,i}$, with $z_{\mu,f} \equiv 5 \times 10^4$, equilibrium is still achieved after an energy injection due to elastic Compton scattering, but since this process does not change the number of photons the end result is a Bose-Einstein distribution [7], i.e., $n(\nu) = [e^{x+\mu(x)} - 1]^{-1}$, where μ is a

frequency dependent chemical potential (rescaled by $k_B T$ so that it is dimensionless). The Kompaneet's equation shows that $\mu(x)$ deviates from a constant only at very low frequencies, i.e., $\mu(x) = \mu_0 e^{-x_c/x}$, with $x_c \simeq 5 \times 10^{-3}$. Henceforth, we approximate μ as constant, which is valid everywhere except deep in the Rayleigh-Jeans tail ($\nu \to 0$). Finally, for $z \lesssim z_{\mu,f}$ even Compton scattering is not efficient enough to establish kinetic equilibrium between matter and radiation. The distortion created after this moment is known as y type and is relevant, e.g., for the Sunyaev-Zel'dovich effect [8]. Of course, this is a simplified picture, since there is no sharp transition between one regime and the next. For the purpose of analytical estimates we will take the period responsible for the creation of μ distortion to be $z_{\mu,f} \lesssim z \lesssim z_{\mu,i}$ with the numerical values given above. As we will see, due to a logarithmic dependence on the size of this interval, changing these values by factors of order unity will not alter the main results. It should be clear though that for precise predictions one needs to study the system numerically.

We will be interested in the energy injection coming from the dissipation of acoustic waves of the adiabatic mode (Silk damping) as these reenter the horizon and start oscillating. Other sources of distortion are present (e.g., adiabatic cooling [4]) and the physics of the system is very rich. Our working assumption here is that either all other sources lead to a smaller and therefore negligible distortion, as it is the case if the primordial power spectrum is not too red tilted, or that all other relevant effects are understood with a high enough precision to be subtracted off leaving the μ distortion caused by Silk damping as the only signal.

μ Distortion.—In this section, following [1,3,4], we derive a formula that relates the late time μ distortion to the primordial power spectrum. Using the Bose-Einstein distribution plus the fact that the total number of photons is constant, for an amount of energy (density) released into the plasma δE one finds that $\mu \simeq 1.4 \delta E / E$. Hence, let us estimate the energy injection due to damping of acoustic waves. The energy density of a density wave is given by [9] $Q = \rho \langle \delta_\gamma(x)^2 \rangle_p c_s^2 / (1 + c_s^2)$, with c_s the sound speed, ρ the density and δ the dimensionless amplitude of oscillations averaged over a period (indicated by $\langle \rangle_p$ to differentiate it from the quantum-ensemble average $\langle \rangle$). Since at this time the Universe is dominated by radiation, we take $\rho = \rho_\gamma$ and $c_s^2 \simeq 1/3$. Then, one has

$$\frac{\delta E}{E} \simeq -\int_{z_{\mu,i}}^{z_{\mu,f}} dz \, \frac{d}{dz} \frac{Q}{\rho_\gamma} \simeq \frac{1}{4} \langle \delta_\gamma(x)^2 \rangle_p |_{z_{\mu,f}}^{z_{\mu,i}}, \quad (1)$$

We can use the transfer function (see, e.g., [11])

$$\Delta_\gamma(k) \simeq 3 \cos(kr) e^{-k^2/k_D^2}, \quad (2)$$

where, using that $R \equiv 3\rho_B / 4\rho_\gamma \ll 1$, the diffusion damping scale is

$$k_D \equiv \left[\int_z^\infty dz \frac{1+z}{6Hn_e \sigma_T (1+R)} \left(\frac{R^2}{1+R} + \frac{16}{15} \right) \right]^{-1/2}$$
$$\simeq (1+z)^{3/2} 4.1 \times 10^{-6} \text{ Mpc}^{-1}, \quad (3)$$

and

$$kr = \int_0^t \frac{k \, dt}{a\sqrt{3(1+R)}} \simeq \frac{2kt}{a\sqrt{3}}. \quad (4)$$

We then have

$$\langle \delta_\gamma(x)^2 \rangle_p = 1.45 \int \frac{d^3 k_1 d^3 k_2}{(2\pi)^6} \mathcal{R}(k_1) \mathcal{R}(k_2)$$
$$\times \langle \Delta_\gamma(k_1) \Delta_\gamma(k_2) \rangle_p e^{i(\vec{k}_1 + \vec{k}_2) \cdot \vec{x}}, \quad (5)$$

where $\mathcal{R}$ describes curvature perturbations. Finally, to account for the fact that μ arises from a thermalization process, we use a top-hat filter in real space $W(x)$, which smears the dissipated energy over a volume of radius $k_s^{-1} \gtrsim k_{D,f}^{-1}$.

Summarizing, the deformation parameter μ is related to primordial perturbations by

$$\mu(x) \simeq 4.6 \int \frac{d^3 k_1 d^3 k_2}{(2\pi)^6} \mathcal{R}(\vec{k}_1) \mathcal{R}(\vec{k}_2) e^{i\vec{k}_+ \cdot \vec{x}} W\left(\frac{k_+}{k_s}\right) \langle \cos(k_1 r)$$
$$\times \cos(k_2 r) \rangle_p [e^{-(k_1^2 + k_2^2)/k_D^2}]_{z_{\mu,f}}^{z_{\mu,i}}, \quad (6)$$

where $W(k) \equiv 3k^{-3}[\sin(k) - k\cos(k)]$ is the Fourier transform of the top-hat filter $W(x)$ and $\vec{k}_\pm \equiv \vec{k}_1 \pm \vec{k}_2$. The quantum-ensemble average of $\mu(x)$ gives the log-integral of the primordial power spectrum from $k_D(z_{\mu,i}) \simeq 1.1 \times 10^4 \text{ Mpc}^{-1}$ to $k_D(z_{\mu,f}) \simeq 46 \text{ Mpc}^{-1}$,

$$\langle \mu(x) \rangle \simeq 2.3 \int d\log k \Delta_{\mathcal{R}}^2(k) [e^{-2k^2/k_D^2}]_f^i, \quad (7)$$

where

$$\langle \mathcal{R}^2 \rangle \equiv (2\pi)^3 \delta^3(\vec{K}_{tot}) P_{\mathcal{R}}(k), \quad P_{\mathcal{R}}(k) \equiv \frac{2\pi^2 \Delta_{\mathcal{R}}^2(k)}{k^3}. \quad (8)$$

To visualize this, in Fig. 1 we plot $\Delta_{\mathcal{R}}^2(k) e^{-2k^2/k_D^2}$ for $k_D(z_{\mu,i})$ (dotted line) and $k_D(z_{\mu,f})$ (dashed line) as well as their difference (red-filled region on the right), which quantifies the amount of dissipated energy and hence μ distortion.

Two-point correlation functions.—Let us now compute the two-point correlation between a direction dependent μ distortion $\mu(\hat{n})$ and temperature anisotropy $\Delta T(\hat{n})$. To make contact with the way observations are analyzed, we will decompose both signals in spherical harmonics, which are an orthonormal basis of functions on the sphere. We start with
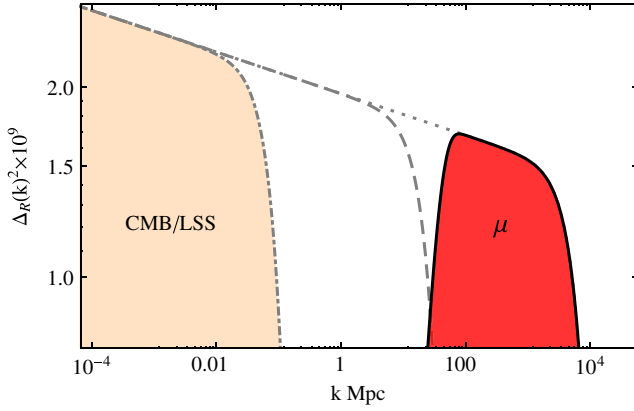
FIG. 1 (color online).   The figure shows the power spectrum with Silk damping as function of log$k$. The dotted, dashed, and dot-dashed lines are $\Delta_R^2 e^{-2k^2/k_D^2}$ at $z_{\mu,i} = 2 \times 10^6$, $z_{\mu,f} = 5 \times 10^4$ and $z_L = 1100$ respectively. The red area on the right indicated by μ is the difference of the power spectrum between $z_{\mu,i}$ and $z_{\mu,f}$.

$$a_{lm}^T \equiv \int d\hat{n} \frac{\Delta T(\hat{n})}{T} Y_{lm}^*(\hat{n})$$
$$= 4\pi \frac{3}{5} (-i)^l \int \frac{d^3k}{(2\pi)^3} \mathcal{R}(\vec{k}) \Delta_l(k) Y_{lm}^*(\hat{k}), \qquad (9)$$

where $\Delta_l(k)$ is the radiation transfer function. Throughout this Letter we will use the Sachs-Wolfe approximation $\Delta_l(k) \simeq j_l(kr_L)/3$ with $r_L \simeq 14$ Gpc the distance from last scattering and j_l a spherical Bessel function. This is in principle valid only in the range $10 \lesssim l \lesssim 50$, but for the purpose of the Fisher forecast it will be a reasonable estimate also for higher l's. Let us define in general

$$\langle (a_{lm}^i)^* a_{l'm'}^j \rangle = \delta_{ll'} \delta_{mm'} C_l^{ij}, \qquad (10)$$

then the TT correlation gives the well-known result

$$C_l^{TT} = \frac{2\pi}{25} \frac{\Delta_{\mathcal{R}}(k_p)^2}{l(l+1)} \simeq \frac{6.0 \times 10^{-10}}{l(l+1)}, \qquad (11)$$

where $k_p \equiv 0.002$ Mpc^{-1} is a common pivot scale at which $\Delta_{\mathcal{R}}^2(k_p) = 2.4 \times 10^{-9}$ [12]. The μ distortion can also be expanded as

$$a_{lm}^\mu = 4.6 \int d\hat{n} Y_{lm}^*(\hat{n}) \frac{d^3k_1 d^3k_2}{(2\pi)^6} \mathcal{R}(\vec{k}_1) \mathcal{R}(\vec{k}_2) e^{i\vec{k}_+ \cdot \vec{x}} W\left(\frac{k_+}{k_s}\right)$$
$$\times \langle \cos(k_1 r) \cos(k_2 r) \rangle_p [e^{-(k_1^2+k_2^2)/k_D^2}]_f^i. \qquad (12)$$

Using the identities

$$e^{i\vec{k}\cdot\vec{x}} = \sum_l (2l+1) i^l P_l(\hat{k}\cdot\hat{x}) j_l(kx), \qquad (13)$$

$$P_l(\hat{k}\cdot\hat{x}) = \frac{4\pi}{2l+1} \sum_{m=-l}^l Y_{lm}(\hat{k}) Y_{lm}^*(\hat{x}), \qquad (14)$$

one can conveniently rewrite (12) as

$$18.4\pi(-i)^l \int \frac{d^3k_1 d^3k_3}{(2\pi)^6} Y_{lm}^*(\hat{k}_+) \mathcal{R}(\vec{k}_1) \mathcal{R}(\vec{k}_2) W\left(\frac{k_+}{k_s}\right)$$
$$\times j_l(k_+ r_L) \langle \cos(k_1 r) \cos(k_2 r) \rangle_p [e^{-(k_1^2+k_2^2)/k_D^2}]_f^i. \qquad (15)$$

The $T\mu$ correlation is found to be

$$C_l^{\mu T} = 6.1\pi \frac{9}{25} f_{\mathrm{NL}} b \frac{\Delta_{\mathcal{R}}^4(k_p)}{l(l+1)} \ln\left(\frac{k_{D,i}}{k_{D,f}}\right)$$
$$\simeq \frac{2.2 \times 10^{-16}}{l(l+1)} f_{\mathrm{NL}} b, \qquad (16)$$

where we used the local bispectrum [13]

$$\langle \mathcal{R}^3 \rangle = (2\pi)^3 \delta^3(\vec{K}_{\mathrm{tot}}) \left(-\frac{6}{5} f_{\mathrm{NL}}\right) [P_{\mathcal{R}}(k_1) P_{\mathcal{R}}(k_2) + 2\mathrm{perm's}] \qquad (17)$$

and defined

$$\frac{b}{l(l+1)} \equiv \frac{2}{\log\left(\frac{k_{D,i}}{k_{D,f}}\right)} \int d\log k_+ j_l(k_+ r_L)^2$$
$$\times \int d\log k_- \frac{\Delta_{\mathcal{R}}^2(k_-/2)}{\Delta_{\mathcal{R}}^2(k_p)} \frac{\Delta_{\mathcal{R}}^2(k_+)}{\Delta_{\mathcal{R}}^2(k_p)}$$
$$\times \left[e^{-(k_+^2+k_-^2)/(2k_D^2)}\right]_f^i W\left(\frac{k_+}{k_s}\right). \qquad (18)$$

This formula can be simplified if one assumes a weak scale dependence. Taking $\Delta_{\mathcal{R}}^2(k) = \Delta_{\mathcal{R}}^2(k_p)(k/k_p)^{n_s-1}$, one finds

$$b \simeq \frac{\left[\left(\frac{k_D}{2k_p}\right)^{n_s-1}\right]_f^i}{(n_s-1)\log\left(\frac{k_{D,i}}{k_{D,i}}\right)} \simeq 1 + \frac{n_s-1}{2} \log\left(\frac{k_{D,i} k_{D,f}}{4k_p^2}\right),$$

so that $b \simeq 1 + 12(n_s - 1)$ close to the scale invariant limit $n_s \to 1$.

The $\mu\mu$ self-correlation has a Gaussian and a non-Gaussian contribution. For the former one finds

$$C_{l,\mathrm{Gauss}}^{\mu\mu} \sim 3.5 \times 10^{-17} \frac{\Delta_{\mathcal{R}}^4(k_{D,f})}{\Delta_{\mathcal{R}}^4(k_p)} \frac{k_s r_L^{-2}}{k_{D,f}^3} \lesssim 9 \times 10^{-29}, \qquad (19)$$

where in the last line we took $\Delta_R(k_{D,f}) = \Delta_R(k_p)$ and in order to get an upper bound $k_s = k_{D,f}$, which is the shortest scale that contributes sizably to dissipation. This scaling can be understood as follows. For a Gaussian field, consider the power spectrum (i.e., μ distortion) at small scales $k_{D,f}$. Its fluctuations in regions much more distant than $k_{D,f}^{-1}$ are completely independent. Hence, the expansion of these power fluctuations in multiples is independent of l, i.e., just white noise, for low multiples $l/r_L \ll k_{D,f}$. Now let us consider the effect of smearing. Measuring the low μ-distortion multiples, we are effectively averaging over

many small-scale (again of order $k_{D,f}$) independent real-izations, and hence we are suppressing the individual variance by $N^{-1/2}$. The number of $k_{D,f}$ modes in a spherical shell around the last scattering surface of thickness k_s is given by $N \sim k_{D,f}^3/(k_s r_L^{-2})$.

Using the local trispectrum

$$\langle \mathcal{R}^4 \rangle = (2\pi)^3 \delta^3(\vec{K}_{\text{tot}})\tau_{\text{NL}} \times [P_\mathcal{R}(k_1)P_\mathcal{R}(k_2)P_\mathcal{R}(|\vec{k}_1 + \vec{k}_3|)$$
$$+ 11\text{perm's}]$$

for the non-Gaussian contribution one finds

$$C_{l,NG}^{\mu\mu} \sim 42\pi \frac{\Delta_\mathcal{R}(k_p)^6}{l(l+1)} \tau_{\text{NL}} b' \ln^2\left(\frac{k_{D,f}}{k_{D,i}}\right)$$
$$\simeq 5.3 \times 10^{-23} \tau_{\text{NL}} \frac{b'}{l(l+1)}, \quad (20)$$

where b' is defined by

$$\frac{b'}{l(l+1)} = \frac{2}{\ln^2\left(\frac{k_{D,f}}{k_{D,i}}\right)} \int d\ln k_+ \, d\ln k_- \, d\ln k_3$$
$$\times \frac{\Delta_\mathcal{R}(k_-/2)^2 \Delta_\mathcal{R}(k_+)^2 \Delta_\mathcal{R}(k_3)^2}{\Delta_\mathcal{R}(k_p)^6}$$
$$\times [e^{-k_-^2/(2k_D^2)}]_f^i [e^{-2k_3^2/k_D^2}]_f^i, \quad (21)$$

such that $b' \simeq 1$ for a scale invariant power spectrum. Notice that, for a given amount of primordial non-Gaussianity, the ratio of the non-Gaussian to Gaussian contributions to the two-point self correlation is much larger for μ than it is for T. In other words, $C_l^{\mu\mu}$ is more sensitive in relative terms to non-Gaussianity than C_l^{TT}.

To discuss the practical observability of the above signals we need to consider experimental noise. For T this is negligible since we will only consider low multiples ($l \lesssim \mathcal{O}(100)$). To model the noise for μ we assume a Gaussian beam and use [14]

$$C_l^{\mu\mu,N} \simeq w_\mu^{-1} e^{l^2/l_{\max}^2}, \quad (22)$$

where w_μ is the sensitivity to μ and $l_{\max}$ is related to the beam size of the experiment. For example, for an experiment like PIXIE [15] the beam size is $\theta_{FWHM} = 1°.6$ which leads to $l_{\max} \simeq 84$ and the sensitivity is $w_\mu^{-1/2} \simeq \sqrt{4\pi} \times 10^{-8}$. So, a figure of merit to keep in mind is

$$C_l^{\mu\mu,N} \simeq 1.3 \times 10^{-15} e^{l^2/l_{\max}^2}. \quad (23)$$

Fisher matrix analysis.—We are now going to use the above results to perform a Fisher forecast for the bounds that μ distortion can put on f_{NL} local. The fact that we invoked the Sachs-Wolfe approximation instead of using numerical transfer functions should not alter much the estimate of the signal to noise ratio because the acoustic peaks in T mostly scale away. The 1×1 Fisher matrix,

$$F \equiv -\frac{\partial^2}{\partial f_{\text{NL}}^2} \ln L = \frac{\partial^2}{\partial f_{\text{NL}}^2} \frac{\chi^2}{2} = \sum_l \frac{C_l^{\mu T} C_l^{\mu T}}{\sigma_l^2},$$

gives us the signal-to-noise ratio via $S/N = f_{\text{NL}}\sqrt{F}$. For mild non-Gaussianity, $f_{\text{NL}} \lesssim 10^5$, we can estimate the noise in the measurement of each l as

$$\sigma_l^2 = \langle (C_l^{\mu T})^2 \rangle - \langle C_l^{\mu T} \rangle^2 \simeq \frac{1}{2l+1} C_l^{TT} C_l^{\mu\mu,N}, \quad (24)$$

where we used that μ and T instrumental noises are uncorrelated, $C_l^{TT} \gg C_l^{TT,N}$ and $C_l^{\mu\mu} \ll C_l^{\mu\mu,N}$ as seen by comparing (19) with (23). Then, the signal-to-noise ratio can be written as

$$\frac{S}{N} \simeq \frac{12}{5} f_{\text{NL}} \Delta_R \sqrt{\log\left(\frac{l_{\max}}{2}\right)} \left(\frac{S}{N}\right)_\mu, \quad (25)$$

where $(S/N)_\mu$ is the signal-to-noise ratio for the μ distortion averaged over the whole sky. Apart from the numerical coefficient, this relation is to be expected. The long-scale fluctuations of μ around its average that correlate with temperature fluctuations are generated by the non-Gaussian part of $\mathcal{R}$ and hence are suppressed with respect to the μ monopole by $f_{\text{NL}}\mathcal{R}$. Notice that reducing the beam size of the experiment improves the signal-to-noise ratio only logarithmically. Plugging in numbers one finds the figure of merit

$$\frac{S}{N} \simeq 0.7 \times 10^{-3} b f_{\text{NL}} \left(\frac{\sqrt{4\pi} \times 10^{-8}}{w_\mu^{-1/2}}\right), \quad (26)$$

where $\Delta\mu = 10^{-8}$ is the estimated one sigma error on the μ-distortion monopole of an experiment like PIXIE and b, defined in (18), is of order one if the primordial power spectrum is approximately scale invariant.

There are at least two classes of models in which the μT correlation could provide the strongest constraints on non-Gaussianity already with PIXIE's sensitivity. First, models in which the power spectrum grows at small scales [16], since then $b \gg 1$. In most of these models b can be as large as $\mathcal{O}(10^2)$ leading to $\Delta f_{\text{NL}}^{\text{loc}} \sim \mathcal{O}(10)$ for PIXIE. Second, models in which the bispectrum diverges faster than the local template in the squeezed limit, i.e. $\langle \mathcal{R}^3 \rangle \propto k_3^{-\alpha}$ for $k_3 \to 0$ with $\alpha > 3$ (see, e.g., [17] for a phenomenological model).

Before concluding, let us remark that while instrumental noise can be improved, there is a lower bound for the noise imposed by nature, often referred to as cosmic variance. In this respect the μT cross correlations possess an important advantage as compared with the temperature bispectrum, which is another probe of the primordial local bispectrum. Cosmic variance, and more generally sampling variance, scales with the number of modes N as $N^{-1/2}$. For a primordial bispectrum of the local type it is useful to distinguish between long and short modes. The latter are the same for both observables, provided one measures the

low multiples of μ ($l \sim 100$). On the other hand, the number of short modes in the temperature bispectrum is about l_{max}^2 and due to Silk damping l_{max} can never be larger than a few thousand. In μT cross correlations the number of short modes is much larger and, as we said, can be estimated as $k_{D,f}^3/(k_s r_L^{-2}) \gtrsim 10^{12}$. This means that, for comic variance limited experiments, the bispectrum's sensitivity is $\Delta f_{NL} \sim \mathcal{O}(5)$, while using μT correlation one can in principle reach $\Delta f_{NL} \sim \mathcal{O}(10^{-3})$.

Conclusions.—Measurements of the μ distortion of the CMB spectrum offer the unique opportunity to probe primordial perturbations in the otherwise inaccessible range $50 \lesssim k$ Mpc $\lesssim 10^4$. We have shown that the two-point self correlations of μ-distortion anisotropies and cross correlations with temperature anisotropies provide a direct measurement of the primordial tri- and bispectrum in the squeezed limit, respectively. We have performed a Fisher matrix forecast for the bounds on f_{NL} local and found that a bound of $|\Delta f_{NL}| \lesssim 10^3$ is achievable already with current technology [15]. It is worth stressing again that in the case of mild non-Gaussianity the monopole μ distortion is expected to be larger than the anisotropic part [see (25) for a quantitative statement].

It would be interesting to further study the performances of a dedicated experiment. We should stress that one should be careful in comparing this bound with those from CMB-LSS, since it applies to very small scales, around a kpc, which are completely unconstrained so far.

We have also noticed that the μT cross correlation is contaminated by a much smaller cosmic variance than the temperature bispectrum and for an ideal experiment it could probe an f_{NL} much smaller than unity. In this respect it is important to stress that for all practical purposes single-field inflation predicts a primordial bispectrum that vanishes in the squeezed limit. The well known consistency relation [18] $f_{NL} = n_s - 1$ is derived using comoving momenta. When comparing with observations one should use physical momenta instead. Doing this leads to an additional term that exactly cancels the $n_s - 1$. As long as one considers other sources of primordial local non-Gaussianity so that $f_{NL} \gg n_s - 1$, this slow-roll correction can be safely neglected as we have done in the present derivation.

[1] W. Hu, D. Scott, and J. Silk, Astrophys. J. **430,** L5 (1994).
[2] For a recent review and a comprehensive bibliography, see [3]
[3] J. Chluba and R. A. Sunyaev, arXiv:1109.6552 [Mon. Not. R. Astron. Soc. (to be published)].
[4] R. Khatri, R. A. Sunyaev, and J. Chluba, Astron. Astrophys. **540,** A124 (2012).
[5] A. S. Kompaneets, Zh. Eksp. Teor. Fiz. **31,** 876 (1957) [Sov. Phys. JETP **4,** 730 (1957)].
[6] L. Danese, G. de Zotti, Astron. Astrophys. **107,** 39 (1982); W. Hu and J. Silk, Phys. Rev. D **48,** 485 (1993); L. Danese and G. de Zotti, Nuovo Cimento **7,** 277 (1971); J. D. Barrow and P. Coles, Mon. Not. R. Astron. Soc. **248,** 52 (1991).
[7] R. A. Sunyaev and Y. B. Zeldovich, Astrophys. Space Sci. **7,** 20 (1970); A. F. Illarionov and R. A. Sunyaev, Sov. Astron. **18,** 413 (1975).
[8] Y. B. Zeldovich and R. A. Sunyaev, Astrophys. Space Sci. **4,** 301 (1969).
[9] This formula contains the correct relativistic factor $c_s^2/(1 + w)$ and for small c_s^2 reduces to the non-relativistic result used e.g. in [1]. For $c_s^2 = 1/3$ The difference is exactly the factor 3/4 found in [10].
[10] J. Chluba, R. Khatri, and R. A. Sunyaev, arXiv:1202.0057.
[11] S. Weinberg, *Cosmology* (Oxford University New York, 2008), p. 593.
[12] E. Komatsu *et al.* (WMAP Collaboration), Astrophys. J. Suppl. Ser. **192,** 18 (2011).
[13] To simplify our notation, we omit the label loc in f_{NL}^{loc}.
[14] S. Dodelson, *Modern Cosmology* (Academic New York, 2003), p. 440.
[15] A. Kogut *et al.*, J. Cosmol. Astropart. Phys. 07 (2011) 025.
[16] See, e.g., J. Chluba, A. L. Erickcek, and I. Ben-Dayan, arXiv:1203.2681, and references therein.
[17] J. Ganc, Phys. Rev. D **84,** 063514 (2011).
[18] J. M. Maldacena, J. High Energy Phys. 05 (2003) 013.