Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **442,** 69–91 (2014) doi:10.1093/mnras/stu794

Cross-correlating Sunyaev–Zel'dovich and weak lensing maps

Dipak Munshi,[1,2]★ Shahab Joudaki,[3,4] Peter Coles,[1,2] Joseph Smidt[4,5] and Scott T. Kay[6]

[1]*School of Physics and Astronomy, Cardiff University, Queen's Buildings, 5 The Parade, Cardiff CF24 3AA, UK*
[2]*Astronomy Centre, School of Mathematical and Physical Sciences, University of Sussex, Brighton BN1 9QH, UK*
[3]*Swinburne University of Technology, Mail number H30, PO Box 218, Hawthorn, Victoria 3122, Australia*
[4]*Department of Physics and Astronomy, University of California, Irvine, CA 92697, USA*
[5]*Los Alamos National Laboratory, Theoretical Division, PO Box 1663, Mail Stop B283, Los Alamos, NM 87545, USA*
[6]*Jodrell Bank Centre for Astrophysics, Alan Turing Building, The University of Manchester, Manchester M13 9PL, UK*

Accepted 2014 April 21. Received 2014 March 18; in original form 2013 December 24

ABSTRACT

We present novel statistical tools to cross-correlate frequency cleaned thermal Sunyaev–Zel'dovich (tSZ) maps and tomographic weak lensing (wl) convergence maps. Moving beyond the lowest order cross-correlation, we introduce a hierarchy of mixed higher order statistics, the cumulants and cumulant correlators, to analyse non-Gaussianity in real space, as well as corresponding polyspectra in the harmonic domain. Using these moments, we derive analytical expressions for the joint two-point probability distribution function for smoothed tSZ (y) and convergence (κ) maps. The presence of tomographic information allows us to study the evolution of higher order *mixed* tSZ–wl statistics with redshift. We express the *joint* PDFs $p_{\kappa y}(\kappa, y)$ in terms of individual one-point PDFs $[p_\kappa(\kappa), p_y(y)]$ and the relevant bias functions $[b_\kappa(\kappa), b_y(y)]$. Analytical results for two different regimes are presented that correspond to the small and large angular smoothing scales. Results are also obtained for corresponding *hotspots* in the tSZ and convergence maps. In addition to results based on hierarchical techniques and perturbative methods, we present results of calculations based on the lognormal approximation. The analytical expressions derived here are generic and applicable to cross-correlation studies of arbitrary tracers of large-scale structure including, e.g., that of tSZ and soft X-ray background. We provide detailed comparison of our analytical results against state of the art Millennium Gas Simulations with and without non-gravitational effects such as *pre-heating* and *cooling*. Comparison of these results with *gravity only* simulations, shows reasonable agreement and can be used to isolate effect of non-gravitational physics from observational data.

Key words: methods: analytical – methods: numerical – methods: statistical – cosmic background radiation – large-scale structure of Universe.

1 INTRODUCTION

Free electrons in the Universe can be detected through the inverse Compton scattering of cosmic microwave background (CMB) photons (Sunyaev & Zel'dovich 1972, 1980; Birkinshaw 1999; Rephaeli 1995). An inhomogeneous distribution of electrons thus induces a secondary anisotropy in the CMB radiation, which is proportional to the thermal energy of the electrons integrated along the line-of-sight direction. This well-known effect is called the thermal Sunyaev–Zel'dovich (tSZ) effect and it is now routinely being used to image nearby galaxy clusters (Reese et al. 2002; Jones et al. 2005; LaRoque et al. 2006) which the electron temperature can reach the order of 10 keV. The future seems set for a huge increase in Sunyaev–Zel'dovich (SZ) measurements of galaxy clusters; data from Planck alone has generated 1227 candidates with reasonably high signal-to-noise ratios (Planck Collaboration 2013b). In the intergalactic medium (IGM), however, the gas is expected to be in mildly overdense regions and to reach a temperature of only 1 keV or so. While the tSZ effect from clusters can alter the temperature of the CMB by an amount of the order of mK in the Rayleigh–Jeans part of the spectrum, the contribution from the IGM is expected only to reach the μK range. This is below the threshold for detection by *Wilkinson Microwave Anisotropy Probe* (*WMAP*). Nevertheless the Planck

★E-mail: D.Munshi@sussex.ac.uk

Surveyor[1] due to its wider frequency coverage, higher sensitivity, and improved resolution (Hansen et al. 2005; Joudaki et al. 2010), has recently published a near all-sky map of the y-parameter (Planck Collaboration 2013a); also see Hill & Spergel (2014).

The unique spectral dependence of the tSZ effect helps in the task of separating it from other sources of CMB temperature fluctuations so various well-developed component separation schemes are available for construction of frequency-cleaned SZ maps (Bouchet & Gispert 1999; Delabrouillie, Cardoso & Patanchon 2003; Leach et al. 2008). The construction of such maps will provide us with a direct opportunity to probe the thermal history of the Universe, in tandem with other observations: owing to their thermal (peculiar motions) ionized electrons scatter CMB photons, an effect which can be studied using the tSZ effect (Rosati, Borgani & Norman 2002); the neutral component of the IGM can be also be probed via observations of the Lyman α forest (Rauch 1998); X-ray emission due to the thermal bremsstrahlung that originates from the interaction of ionized electrons and nuclei can also provide clues to the nature and evolution of the IGM. These tracers all probe different phases of the IGM and thus play complementary roles. For example, the X-ray emission depends on the square of the density, so is most sensitive to the densest regions in the IGM, primarily in the local Universe, while tSZ studies can probe the more distant Universe because Compton scattering is independent of redshift, as well as being an unbiased tracer of all electrons, as they all participate in the scattering, the tSZ effect is weakly dependent on density so probes the electron density in a wide range of environments.

The modelling of lower order statistics of the tSZ effect can be carried out using various approaches. In the past (Cooray 2000, 2001; Seljak 2000; Zhang & Pen 2001; Komatsu & Seljak 2002; Zhang & Sheth 2007) the modelling has followed the halo model (Cooray & Sheth 2002). The statistical distribution of haloes, specifically their number density as a function of mass (i.e. their mass function), is assumed to be that predicted provided by the Press–Schechter formalism (Press & Schechter 1974) or its generalizations, and the radial profile of such haloes was assumed to be that of Navarro, Frenk & White (1996). The hot gas, assumed to have been heated by shocks, is taken to be in hydrodynamical equilibrium with the dark matter distribution. The typical temperature reached in such systems is sufficient to ionize the hydrogen and helium atoms. These ingredients are sufficient to model the tSZ effect raising from collapsed haloes (Cooray 2000); in addition to this analytical modelling, numerical simulation of the tSZ plays an important role in our understanding of the physics involved (Persi et al. 1995; da Silva et al. 2000; Refregier et al. 2000; Seljak, Burwell & Pen 2001; Springel, White & Hernquist 2001; White, Hernquist & Springel 2002; Lin et al. 2004; Zhang, Pen & Trac 2004; Cao, Liu & Fang 2007; Roncarelli et al. 2007; Hallman et al. 2007, 2009; Shaw et al. 2010; Battaglia et al. 2013; McCarthy et al. 2013). To extend the halo model to larger scales, the extended distribution of free electrons is typically assumed to trace the distribution of dark matter on large scales where the variations in density are in the linear or quasi-linear regime. A perturbative approach along these lines has been developed by several authors over the years (Goldberg & Spergel 1999a,b; Cooray 2000, 2001; Cooray & Hu 2000; Munshi et al. 2013).

Ongoing and proposed ground surveys such as SZA,[2] ACT,[3] APEX,[4] SPT[5] and the recently completed all-sky survey *Planck* has produced a map of the entire y-sky with a great precision (Planck Collaboration 2013a). The SPT collaboration has already reported the measurement of the tSZ power spectrum at $\ell \approx 3000$ (Lueker et al. 2010; High et al. 2012; Hanson et al. 2013; Holder et al. 2013; Saro et al. 2013; Story et al. 2013; Vieira et al. 2013; Hou et al. 2014); the ACT collaboration has also reported analysis on similar scales (Dunkley et al. 2010, 2013; Fowler et al. 2010; Hand et al. 2011; Shegal et al. 2011; Sherwin et al. 2011; Calabrese et al. 2013; Wilson et al. 2012). It is expected that ongoing surveys will improve these measurements due to their improved sky coverage as well as wider frequency range. It is important to realize why the study of secondaries such as tSZ should be an important aspect of any CMB mission. In addition to the important physics the secondaries probe, accurate modelling of the secondary non-Gaussianities is required to avoid 20–30 per cent constraint degradations in future CMB data sets such as *Planck* and CMBPol[6] (Smidt et al. 2010).

While the tSZ surveys described above provide a direct probe of the baryonic Universe, weak lensing (wl) observations (Munshi et al. 2008) on the other hand can map the dark matter distribution in an unbiased way. In recent years there has been tremendous progress on the technical front in terms of specification and control of systematics in wl observables. There are many ongoing and future wl surveys such as Canada–France–Hawaii Telescope (CFHT)[7] legacy survey, Pan-STARRS[8] and the Dark Energy survey (DES)[9] and in future, the Large Synoptic Survey Telescope (LSST),[10] *Joint Dark Energy Mission*[11] will map the dark matter and dark energy distribution of the entire sky in unprecedented details. Among other things, hold great promise in shedding light on the nature of the dark energy and the origin of neutrino masses (Joudaki & Kaplinghat 2011), where the wl signals dominate the others considered by, e.g., the Dark Energy Task Force (Albrecht et al. 2011). However, the optimism associated with lensing is predicated on overcoming the vast systematic uncertainties in both measurement and in theory (Cooray & Hu 2002; Hirata & Seljak 2003, 2004; White 2004; Ma, Hu & Huterer 2005; Huterer et al. 2006; McDonald, Trac & Contaldi 2006). The statistics of the wl convergence has been studied in great detail using an extension of perturbation

[1] http://www.rssd.esa.int/index.php?project=SP
[2] http://astro.uchicago.edu/sza
[3] http://www.physics.princeton.edu/act
[4] http://bolo.berkeley.edu/apexsz
[5] http://pole.uchicago.edu
[6] http://cmbpol.uchicago.edu/
[7] http://www.cfht.hawai.edu/Sciences/CFHLS/
[8] http://pan-starrs.ifa.hawai.edu/
[9] https://www.darkenergysurvey.org/
[10] http://www.lsst.org/llst_home.shtml
[11] http://jdem.gsfc.nasa.gov/

theory (Munshi & Jain 2000, 2001; Munshi, Valageas & Barber 2004; Valageas & Munshi 2004) and methods based on the halo model (Cooray & Hu 2000; Takada & Jain 2002, 2003). These studies developed techniques that can be used to predict the lower order moments (equivalent to the power spectrum and multispectra in the harmonic domain) and the entire probability distribution function (PDF) for a given wl survey. The photometric redshifts of source galaxies are useful for tomographic studies of the dark matter distribution and establish a three-dimensional picture of their distribution (Munshi et al. 2011b).

Due to the line-of-sight integration inherent in the tSZ effect, the redshift evolution is completely lost which degrades the power of tSZ to distinguish different thermal histories. It is however possible to recover some of the information lost by cross-correlating the tSZ with external tracers (Zhang & Pen 2001; Shao et al. 2011b). These tracers could comprise galaxy surveys with spectroscopic redshift information or dark matter distribution from wl surveys in tomographic slices. Such cross-correlations have been studied using two-point statistics. We extend these results to higher order. Higher order statistics such as cumulants and cumulant correlators (Bernardeau et al. 2002) can probe higher order cross-correlations and can in principle provide an independent probe of the bias associated with the baryonic pressure fluctuations. The tomographic cross-correlation statistics that we develop here can be applied to surveys with overlapping sky coverage. Many of the wl surveys and the tSZ surveys will have overlapping sky coverage. For example, DES will overlap with the SPT sky coverage and plans to measure photometric redshifts of roughly 10^8 galaxies up to $z = 1.3$. Tomographic cross-correlation statistics at second or higher order can provide a more detailed picture of the evolution of the thermal history of the baryonic gas by mapping the associated pressure fluctuation. In a recent paper van Waerbeke et al. (2013) have cross-correlated the CFHTLenS data with *Planck* tSZ maps. They measure a non-zero correlation between the two maps out to one degree angular separation on the sky, with an overall significance of six sigma and use the results to conclude a substantial fraction of the 'missing' baryons in the universe may reside in a low-density warm plasma that traces dark matter. An internal detection of tSZ and CMB lensing cross-correlation in *Planck* nominal mission data has also been reported recently at 6.2σ significance (Hill & Spergel 2014).

This paper is organized as follows. In Section 2, we introduce our notations for both tSZ and wl observables. Next, we introduce the mixed cumulant correlators and the cumulants in Section 3. Two different models are used to model the underlying dark matter distribution. We employ the (Smith et al. 2003) prescription, for modelling the evolution of matter power spectrum and use the hierarchical approach for modelling of the higher order correlation hierarchy. This allows us to extend the lower order moment results to arbitrary order and construct the relevant PDF and bias for the smoothed wl κ field and the y-maps in Section 4. In Section 5, we present a short description of the simulations and Section 6 is reserved for the discussion of test of analytical results against numerical simulations. Finally, Section 7 is reserved for concluding remarks. A brief review of the hierarchical ansatz in the quasi-linear and non-linear regimes, as well as the lognormal approach, is provided in the appendix.

The results presented here are generic and can be extended to study other secondaries, such as the cross-correlation involving CMB lensing and tSZ maps.

2 NOTATIONS

In this section, we introduce our notation for the tSZ effect and wl convergence. We will use the following line element:

$$\mathrm{d}s^2 = -c^2\mathrm{d}t^2 + a^2(r)\left[\mathrm{d}r^2 + \mathrm{d}_{\mathrm{A}}^2(r)(\sin^2\theta\,\mathrm{d}\theta^2 + \mathrm{d}\phi^2)\right]. \tag{1}$$

Here $\mathrm{d}_{\mathrm{A}}(r)$ is the comoving angular diameter distance at a (comoving) radial distance r. The angular diameter distance is $\mathrm{d}_{\mathrm{A}}(r) = (-\mathrm{K})^{-1/2}\sinh((-\mathrm{K})^{1/2}r)$ for negative curvature, $\mathrm{d}_{\mathrm{A}}(r) = (\mathrm{K})^{-1/2}\sin((\mathrm{K})^{1/2}r)$ for positive curvature, and for a flat Universe $\mathrm{d}_{\mathrm{A}}(r) = r$. The parameters H_0 and Ω_0 decide the constant K: $\mathrm{K} = (\Omega_0 + \Omega_\Lambda - 1)\mathrm{H}_0^2$. The underlying cosmology that we adopt for numerical study is specified by the following parameter values (to be introduced later): $\Omega_\Lambda = 0.741,\ h = 0.72,\ \Omega_{\mathrm{b}} = 0.044,\ \Omega_{\mathrm{CDM}} = 0.215,\ \Omega_0 = \Omega_{\mathrm{b}} + \Omega_{\mathrm{CDM}},\ n_{\mathrm{s}} = 0.964,\ w_0 = -1,\ w_a = 0,\ \sigma_8 = 0.803,\ \Omega_\nu = 0$. The comoving radial distance at a redshift z is determined by the cosmology (Ω_0, H_0):

$$r(z) = \int_0^z \frac{\mathrm{d}z}{\mathrm{H}_0\sqrt{\Omega_0(1+z)^3 + \Omega_{\mathrm{K}}(1+z)^2 + \Omega_\Lambda}}. \tag{2}$$

Throughout, c will denote the speed of light and will be set to unity.

2.1 Thermal Sunyaev–Zel'dovich effect

The tSZ effect contributes to the CMB temperature fluctuation and is typically expressed as $\delta_T(\nu, \hat{\Omega}) = \Delta T(\hat{\Omega})/\mathrm{T}_0 = g(x_\nu)y(\hat{\Omega})$. In this expression, $g(x_\nu)$ corresponds to the spectral dependence and $y(\hat{\Omega})$ encodes the angular dependence; x_ν represents the dimensionless frequency and $\hat{\Omega} = (\theta, \phi)$ corresponds to a unit vector that signifies pixel positions. A subscript s will be used to denote the smoothed maps, e.g., $y(\hat{\Omega})$. In the non-relativistic limit $g(x_\nu)$ takes the following form:

$$g(x_\nu) = x_\nu\coth\left(\frac{x_\nu}{2}\right) - 4 = \left(x_\nu\frac{e^{x_\nu}+1}{e^{x_\nu}-1} - 4\right); \qquad x_\nu = \frac{h\nu}{k_B\mathrm{T}_0} = \frac{\nu}{56.84\,\mathrm{GHz}} = \frac{5.28\,\mathrm{mm}}{\lambda}. \tag{3}$$

Here k_B and h are the Boltzmann and Planck constants, respectively; ν denotes the frequency of the photon and $\mathrm{T}_0 = 2.726\,\mathrm{K}$ is the temperature of the CMB sky. The tSZ effect is shown as CMB temperature decrement at $\nu \ll 218$ GHz and as an increment at $\nu \gg 218$ GHz, with a null at $\nu = 218$ GHz. In the Rayleigh–Jeans limit, characterized by $x \ll 1$, $g(x) \approx -2$ is roughly independent of frequency and in the

other limiting situation $x_\nu \gg 1$, $g(x_\nu) \approx (x_\nu - 4)$. The key information regarding thermal history of the Universe is encoded in $y(\hat{\Omega})$ maps that are extracted from frequency maps obtained through multifrequency CMB observations. The y-maps are opacity weighted integrated pressure fluctuations along the line of sight,

$$y(\hat{\Omega}) \equiv \int \mathrm{d}s\; n_\mathrm{e}\sigma_\mathrm{T}\frac{k_B T_\mathrm{e}}{m_\mathrm{e}c^2} = \frac{\sigma_\mathrm{T}}{m_\mathrm{e}c^2}\int_0^{r_0}\mathrm{d}r\; a(r) n_\mathrm{e} k_B T_\mathrm{e}(\hat{\Omega}, r) = \frac{\sigma_\mathrm{T}}{m_\mathrm{e}c^2}\int_0^{\eta_\mathrm{H}}\mathrm{d}\eta\; a(\eta)\;\Pi_\mathrm{e}(\eta, \hat{\Omega}) = \int_0^{r_0}\mathrm{d}r\; w_\mathrm{SZ}(r)\pi_\mathrm{e}(r). \tag{4}$$

Here $\Pi_\mathrm{e} = n_\mathrm{e}k_\mathrm{B}T_\mathrm{e}$, while m_e corresponds to the electron mass, k_B denotes Boltzmann's constant, $\sigma_\mathrm{T} = 6.6510^{-25}\mathrm{cm}^2$ represents the Thompson cross-section, n_e denotes the number density of electrons, and T_e is the electron temperature. We denote the comoving distance to the surface of last scattering by r_0. Conformal time is denoted by $\mathrm{d}\eta = \mathrm{d}t/a(t)$. The line-of-sight integral depends on the comoving radial coordinate distance r and $a(r)$ is the corresponding scale factor of the Universe. The weight is defined as $w_\mathrm{SZ}(r) = \dot{\tau}(r) = \sigma_T n_\mathrm{e}(r)a(r)$, where the dot defines the derivative with respect to comoving radial distance r, and the dimensionless pressure fluctuation is defined as $\pi_\mathrm{e} = k_\mathrm{B}T_\mathrm{e}/m_\mathrm{e}c^2$. Notice that in accordance with Cooray, Tegmark & Hu (2000), we have defined the dimensionless pressure π_e to be independent of number density of electrons. The $n_\mathrm{e}(r)$ dependence of y-parameter is absorbed in the weight function $w_\mathrm{SZ}(r)$ defined above. We will cross-correlate the Comptonization map $y(\hat{\Omega})$ with tomographic and projected maps from wl surveys to constrain the thermal history of the Universe and its evolution with redshift. Throughout, we will consider the Rayleigh–Jeans part of the spectrum $\delta_T = -2y$; for ACT and SPT operating at $\nu = 150\,\mathrm{GHz}$ from equation (3) we get $g(x) = -0.95$. Detailed modelling of the bias is required only for the computation of variance. The variance $\langle \delta y^2(\hat{\Omega})\rangle$ samples the pressure fluctuation power spectrum $P_{\pi\pi}$ and is expressed as

$$\langle \delta y^2(\hat{\Omega})\rangle_c = \int_0^{r_0}\mathrm{d}r\,\frac{\omega_\mathrm{SZ}^2(r)}{\mathrm{d}_\mathrm{A}^2(r)}\int\frac{\mathrm{d}^2\boldsymbol{l}}{(2\pi)^2}\,\mathrm{P}_{\pi\pi}\left[\frac{\ell}{\mathrm{d}_\mathrm{A}(r)}, r\right] b_\ell^2(\theta_s). \tag{5}$$

The pressure power spectrum $\mathrm{P}_{\pi\pi}(k, z)$ at a redshift z is expressed in terms of the underlying power spectrum $\mathrm{P}_{\delta\delta}(k, z)$ using a bias $b_\pi(k, z)$, i.e. $\mathrm{P}_{\pi\pi}(k, z) = b^2(k, z)\mathrm{P}_{\delta\delta}(k, z)$. The bias $b_\pi(k, z)$ is assumed to be independent of length-scale or equivalently wavenumber k, i.e. $b_\pi(k, z) = b_\pi(z)$. The redshift dependent bias can be expressed as $b_\pi(z) = b_\pi(0)/(1 + z)$. Here $b_\pi(0)$ can be written as $b_\pi(0) = k_B T_\mathrm{e}(0) b_\delta/m_\mathrm{e}c^2$. Different values of b_δ have been reported, e.g., Refregier et al. (2000) found $b_\delta \approx 8-9$ and $T_\mathrm{e}(0) \approx 0.3-0.4$. On the other hand Seljak et al. (2001) found $b_\delta \approx 3-4$. Typical value of $b_\pi(0)$ found by Cen & Ostriker (1999) is $b_\pi(0) = 0.0039$. This is a factor of 2 lower than the value used by Goldberg & Spergel (1999a,b) and Cooray & Hu (2000). A Gaussian beam $b_\ell(\theta_s)$ with full width at half-maximum at θ_s is assumed. We use a redshift dependent generic linear bias model in association with hierarchical ansatz and showed that its predictions are nearly identical to the predictions of lognormal model. The linear biasing model has also been used in generating semi-analytical simulations (White 2003; Schulz & White 2003) and have recently been tested rigorously against numerical simulation (Munshi et al. 2013). Lognormal model has a higher range of validity than perturbative expansion underlying hierarchical ansatz. For numerical calculation we have used the lognormal model as it is much easier to implement. We have found that such an approach works very well for gravity only (GO) simulations.

Some comments are in order at this point. The hierarchical model that we use here was previously used by Zhang, Pen & Wang (2002), Cooray et al. (2000) and Cooray (2000, 2001). It is known that the y-parameter is proportional to the square of the density contrast δ^2, so, e.g., the three-point correlation function of y effectively samples the six-point correlation function of the underlying density contrast δ. However, studies by Cooray (2000, 2001) assumed a linear biasing model for the pressure fluctuations, i.e. in the Fourier domain $\delta_\pi(k, z) = \delta(k, z)b_\pi(k, z)$. The pressure bias model assumes the pressure to be a linear convolution over the density field. Finally Cooray (2000, 2001) further simplified the bias $b_\pi(k, z)$ by factorizing it into a redshift dependent and momentum k-dependent part. The spatial dependence was assumed to be constant. This is the simplification that we will use in our results, which will be useful in computation of correlation hierarchy to an arbitrary order. The hierarchical ansatz and the resulting scaling functions were also used by Valageas, Silk & Schaeffer (2001), Valageas & Silk (1999) and Valageas & Schaeffer (2000). For computation of the number density of collapsed objects these were then used to obtain the tSZ and kinetic Sunyaev–Zel'dovich (kSZ) effect statistics. The results derived here are complementary to calculations based on the halo model and are applicable to tSZ effect originating from the extragalactic ionized medium.

2.2 Weak lensing in projection and tomography

Cosmological wl effects are conveniently encoded in the effective convergence field, which is defined as a weighted projection of the matter density contrast δ. The statistics of the smoothed wl convergence κ are similar to the projected 3D density contrast $\delta(\boldsymbol{x})$ and can be expressed through a line-of-sight integration $\kappa(\hat{\Omega}) = \int_0^\infty \mathrm{d}r\; w_\mathrm{wl}(r)\delta(r, \hat{\Omega})$. Their weight $w_\mathrm{wl}(r)$ is sensitive to the source distribution; for 2D surveys without any source redshift information, it can be expressed as

$$\omega_\mathrm{wl}(r) = \frac{3}{2}\frac{H_0^2}{c^2}\Omega_0 a^{-1}(r)\mathrm{d}_\mathrm{A}(r)\int_r^\infty \mathrm{d}r_s\; p_s(z_s)\,\frac{\mathrm{d}z_s}{\mathrm{d}r_s}\frac{\mathrm{d}_\mathrm{A}(r - r_s)}{\mathrm{d}_\mathrm{A}(r_s)};\quad p_s(z_s) = \bar{n}_g\frac{z_s^2}{2z_0^3}\exp^{-z_s/z_0}. \tag{6}$$

Here $p_s(z_s)$ represents the source redshift distribution and $\bar{n}_g$ denotes surface density of sources. The peak of the distribution is reached at $2z_0$. We will adopt two different survey configurations. We adopt $z_0 = 0.5$ for the LSST survey. For the tomographic calculation we divide the entire source redshift range in five redshift bins with each tomographic bin containing roughly the same number density of source galaxies.

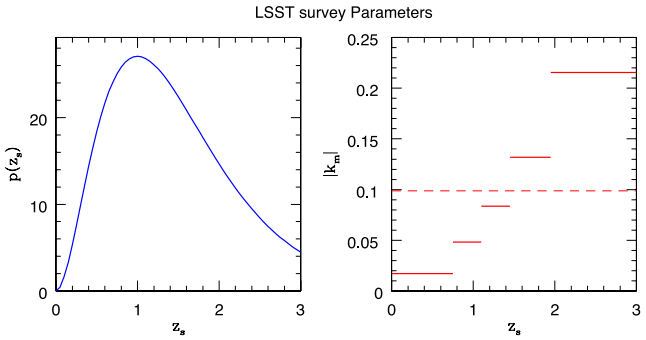

Figure 1. The parameter κ_m for the entire survey and $\kappa_m^{(i)}$ for individual bins (see equation 8 for definitions), that denotes the minimum value of convergence for the survey or respective bins, are displayed as a function of source redshift for the Λ cold dark matter (ΛCDM) cosmology (right-hand panel). These parameters are independent of angular smoothing scale. For individual bins it is shown using solid lines and for the entire survey it is shown using dashed lines. The left-hand panel shows the source distribution as a function of redshift as defined in equation (6). See the text for more details on our choice of tomographic bins.

The redshift bins of the sources are delimited at [0.75, 1.1, 1.45, 1.95, 3.00]. The tomographic convergence maps $\kappa^{(i)}(\hat{\Omega}) = \int_0^\infty \mathrm{d}r\ w_{\rm wl}^{(i)}(r)\delta(r,\hat{\Omega})$ depend on individual weights:

$$\omega_{\rm wl}^{(i)}(r) = \frac{3}{2}\frac{H_0^2}{c^2}\Omega_0\frac{1}{\bar{n}_i}a^{-1}(r)\,\mathrm{d_A}(r)\int_{\max\{r,r_i\}}^{r_{i+1}}\mathrm{d}r\ p_s(r_s)\,\frac{\mathrm{d}z_s}{\mathrm{d}r_s}\frac{\mathrm{d_A}(r-r_s)}{\mathrm{d_A}(r_s)}. \tag{7}$$

We will need the minimum values of the projected convergences for individual tomographic bins $\kappa_m^{(i)}$ as well as the entire survey κ_m which are defined as follows:

$$\kappa_m^{(i)} = -\int_0^\infty \mathrm{d}r\ \omega_{\rm wl}^{(i)}(r); \qquad \kappa_m = -\int_0^\infty \mathrm{d}r\ \omega_{\rm wl}(r). \tag{8}$$

These expressions are derived by setting $\delta = -1$ in the definition of $\kappa^{(i)}$ and κ (Munshi & Jain 2000). We will next use these expressions to derive the cross-correlations at second and higher order both in real space and the harmonic domain. The source redshift distribution for LSST survey and the corresponding $\kappa_m^{(i)}$ defined above are displayed in Fig. 1.

3 MIXED LOWER ORDER MOMENTS: CUMULANTS AND CUMULANT CORRELATORS

In this section, we present results in real space and in the harmonic domain that are completely generic and can provide useful information to study the pressure fluctuations in the baryonic gas. We apply these results to understand the studies involving tomographic bins from wl surveys. In case of CMB lensing one of the source planes is identified with the last scattering surface (LSS). The cross-spectra $\mathrm{P}_{\delta\pi}$ (defined below) can be probed by cross-correlating tomographic wl maps $\kappa_{(i)}$ and the tSZ y-maps of projected pressure fluctuations. These correlations sample the three-dimensional density–pressure cross-spectra $\mathrm{P}_{\delta\pi}$; using the small angle approximation (Kaiser 1998) one can write:

$$\langle\kappa_{(i)}(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c = \int_0^{r_0}\mathrm{d}r\frac{\omega_{\rm wl}^{(i)}(r)\,\omega_{\rm SZ}(r)}{\mathrm{d_A^2}(r)}\int\frac{\mathrm{d}^2\boldsymbol{l}}{(2\pi)^2}\ \exp(\,i\,\boldsymbol{\theta}_{12}\cdot\boldsymbol{l})\,\mathrm{P}_{\delta\pi}\left[\frac{\ell}{\mathrm{d_A}(r)},r\right]b_\ell(\theta_s)W_{\rm TH}(\ell\theta_s); \quad P_{\pi\delta}(k,z) = b_\pi(z)P_\delta(k,z). \tag{9}$$

Here, $b_\ell(\theta_s)$ and $W_{\rm TH}(\ell\theta_s) = (2J_1(\ell\theta_s)/\ell\theta_s)$ are Gaussian and top-hat windows, respectively. The line-of-sight directions $\hat{\Omega}_1$ and $\hat{\Omega}_2$ are separated by an angle θ_{12}. In our notation $|\boldsymbol{l}| = \ell$. The weight function $\omega_{(i)}(r)$ for tomographic convergence and $\omega_{\rm SZ}(r)$ for tSZ surveys are defined in equations (7) and (4), respectively.

To study non-Gaussianity in pressure fluctuation, we need to go beyond the power spectrum analysis; we propose the mixed cumulant correlators for this purpose. Cumulant correlators have the advantage of being very simple to estimate, and are defined in real space so can be useful for smaller surveys. Similar results can be obtained for the kurt-spectra, which we will not consider here. We will consider a top-hat smoothing $W_{\rm TH}(l\theta_s)$ for the convergence maps and a Gaussian beam $b_l(\theta_s)$ for the $y(\hat{\Omega})$ maps

$$\langle\kappa_{(i)}^2(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c = \int_0^{r_0}\mathrm{d}r\frac{[\omega_{\rm wl}^{(i)}]^2(r)\,\omega_{\rm SZ}(r)}{\mathrm{d_A^4}(r)}\int\frac{\mathrm{d}^2\boldsymbol{l}_1}{(2\pi)^2}W_{\rm TH}(\ell_1\theta_s)\int\frac{\mathrm{d}^2\boldsymbol{l}_2}{(2\pi)^2}W_{\rm TH}(\ell_2\theta_s)\int\frac{\mathrm{d}^2\boldsymbol{l}_3}{(2\pi)^2}b_{l_3}(\theta_s)\mathrm{B}_{\delta\delta\pi}\left(\frac{\ell_i}{\mathrm{d_A}(r)},r\right)_{\sum l_i=0} \tag{10}$$

$$\langle\kappa_{(i)}(\hat{\Omega}_1)\,\delta y^2(\hat{\Omega}_2)\rangle_c = \int_0^{r_0}\mathrm{d}r\frac{\omega_{\rm wl}^{(i)}(r)\,\omega_{\rm SZ}^2(r)}{\mathrm{d_A^4}(r)}\int\frac{\mathrm{d}^2\boldsymbol{l}_1}{(2\pi)^2}W_{\rm TH}(\ell_1\theta_s)\int\frac{\mathrm{d}^2\boldsymbol{l}_2}{(2\pi)^2}b_{\ell_2}(\theta_s)\int\frac{\mathrm{d}^2\boldsymbol{l}_3}{(2\pi)^2}b_{\ell_3}(\theta_s)\mathrm{B}_{\delta\pi\pi}\left(\frac{\ell_i}{\mathrm{d_A}(r)},r\right)_{\sum l_i=0}. \tag{11}$$

Here, $\mathrm{B}_{\delta\delta\pi}$ represents the mixed bispectrum involving three dimensional density contrast $\delta(\boldsymbol{x})$ and pressure fluctuation $\pi(\boldsymbol{x})$. These *mixed* cumulant correlators should be considered in addition to the pure ones that probe $\mathrm{B}_{\delta\delta\delta}$ and $\mathrm{B}_{\pi\pi\pi}$. These results can trivially be extended to higher order to compute mixed cumulant correlators. At the level of fourth order they will probe the mixed trispectra of various types; e.g., we can use $\langle\kappa_{(i)}^2(\hat{\Omega}_1)y^2(\hat{\Omega}_2)\rangle_c$ $\langle\kappa_{(i)}^3(\hat{\Omega}_1)y(\hat{\Omega}_2)\rangle_c$ to probe the trispectra $\mathrm{T}_{\delta\delta\pi\pi}$, $\mathrm{T}_{\delta\delta\delta\pi}$ and $\mathrm{T}_{\delta\pi\pi\pi}$. At these level these statistics are completely

general and they can be estimated in real space simply by cross-correlating smoothed $y(\theta_s)$ and $\kappa(\theta_s)$ maps raised to suitable powers. The higher order expressions follow from the same logic.

The derivation so far has been completely generic. It does *not* depend on specific form of the multispectra. Later, we will be using a specific form for the correlation hierarchy which will allow us to include correlation functions to all order. This is done using a generating function formalism.

The results in the real space are suitable for smaller surveys. However, with the increase in survey size it is pertinent to consider a harmonic space approach as real space measurements are often correlated. We will introduce power spectra in the harmonic domain that are Fourier transforms of the corresponding cumulant correlators in real space.

$$\mathcal{C}_\ell^{\kappa\kappa,y} = \sum_{\ell_1\ell_2} \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)}{4\pi(2\ell+1)}} \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ 0 & 0 & 0 \end{pmatrix} \mathcal{B}^{\kappa\kappa y}_{l_1 l_2 l}; \quad \mathcal{C}_\ell^{yy,\kappa} = \sum_{\ell_1\ell_2} \sqrt{\frac{(2\ell_1+1)(2\ell_2+1)}{4\pi(2\ell+1)}} \begin{pmatrix} \ell_1 & \ell_2 & \ell \\ 0 & 0 & 0 \end{pmatrix} \mathcal{B}^{yy\kappa}_{\ell_1\ell_2\ell}. \tag{12}$$

Here, $\mathcal{B}^{\kappa\kappa y}_{\ell_1\ell_2\ell_3}$ and $\mathcal{B}^{\kappa yy}_{\ell_1\ell_2\ell_3}$ denote the angular bispectra. They are related to their 3D counterparts by a line-of-sight projection, e.g.,

$$\mathcal{B}^{\kappa\kappa y}_{\ell_1\ell_2\ell_3} = \int_0^\infty \mathrm{d}r \left[\frac{w^2_{\rm wl}(r) w_{\rm SZ}(r)}{\mathrm{d}^4_{\rm A}(r)}\right] \mathrm{B}_{\delta\delta\pi}\left[\frac{\ell_1}{\mathrm{d}_{\rm A}(r)}, \frac{\ell_2}{\mathrm{d}_{\rm A}(r)}, \frac{\ell_3}{\mathrm{d}_{\rm A}(r)}; r\right]. \tag{13}$$

Similar expressions can be obtained for individual bins by replacing $w(r)$ with $w_{(i)}(r)$. The skew- and kurt-spectra are simple to estimate from real data, and the scatter for such estimates is well understood (Munshi & Heavens 2010; Munshi et al. 2011a). These results can be used to construct a family of skew-spectra for various choices of tomographic slices. We also provide the expressions for the auto- and cross-spectra for tSZ and wl surveys below:

$$\mathcal{C}_\ell^{(i)\kappa,y} = \int_0^\infty \mathrm{d}r \frac{\omega^{(i)}_{\rm wl}(r)\omega_{\rm SZ}(r)}{\mathrm{d}^2_{\rm A}(r)} \mathrm{P}_{\delta\pi}\left(\frac{\ell}{\mathrm{d}_{\rm A}(r)}; r\right). \tag{14}$$

The auto spectra for wl and tSZ can be expressed in terms of $\mathrm{P}_{\delta\delta}$ and $\mathrm{P}_{\pi\pi}$ with suitable changes in the weight functions, i.e. $[\omega^{(i)}_{\rm wl}(r)]^2$ for wl surveys and $\omega^2_{sz}(r)$ for tSZ surveys.

The wl power spectrum (ps) for individual bins and the tSZ (y-parameter) ps are displayed in Fig. 2. The projected or 2D tSZ ps is obtained by replacing both $\omega^{(i)}_{\rm wl}(r)$ in equation (14) by $\omega_{\rm SZ}(r)$. For computing the wl auto-spectra for individual bins, we replace both weight functions by $\omega_{(i)}(r)$. We have displayed the cross-spectra $\mathcal{C}_\ell^{(i)\kappa,y}$ defined in equation (14) in Fig. 3 as a function of the harmonics ℓ. We have shown the results for two different σ_8. Higher redshift bins are more correlated with tSZ. The 2D tSZ cross-spectra is obtained by replacing $\omega^{(i)}_{\rm wl}(r)$ in equation (14) by $\omega_{\rm wl}(r)$. The 2D wl ps (right-hand panel) and the 2D cross-spectra (left-hand panel) are shown in Fig. 4.

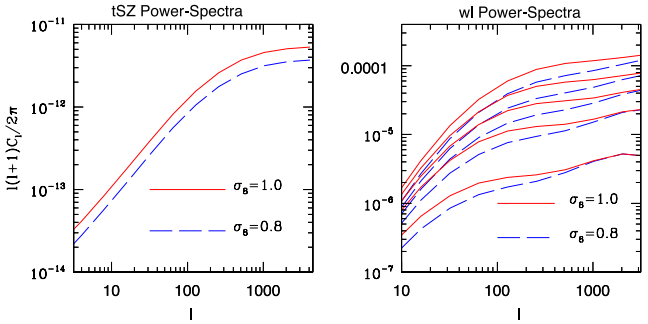

Figure 2. The left-hand panel shows tSZ power spectrum as a function of the harmonics ℓ. The right-hand panel shows redshift-resolved tomographic wl and tSZ cross-spectra, defined in equation (14), as function of ℓ. The background cosmology is assumed ΛCDM. The solid lines in each panel correspond to $\sigma_8 = 1$ and the dashed lines correspond to $\sigma_8 = 0.8$.

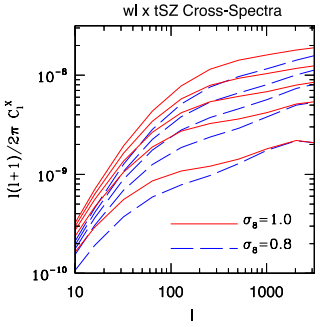

Figure 3. The redshift-resolved cross-spectra $C_\ell^{(i)\kappa,y}$ given in equation (14) for wl tomographic slices and the tSZ survey are plotted as function of harmonics ℓ.

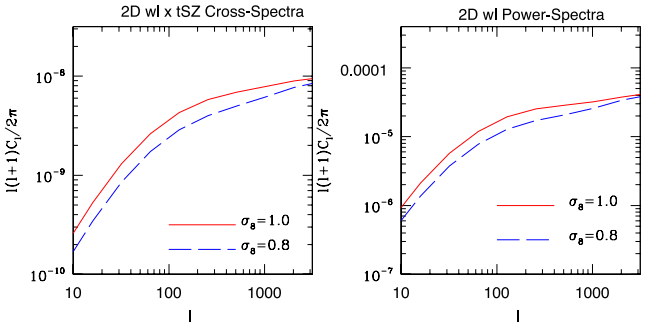

Figure 4. The projected (or 2D) wl and tSZ cross-spectra defined in equation (14) is plotted as a function of harmonics ℓ (left-hand panel). We also show the projected wl convergence power spectra as a function of ℓ (right-hand panel). Two different values of $\sigma_8 = [0.8, 1.0]$ are considered.

3.1 Hierarchical ansatze

In deriving the above expressions, we have not used any specific form for the matter correlation hierarchy, however the length-scales involved in small angles are in the highly non-linear regime. Assuming a tree model for the matter correlation hierarchy in the highly non-linear regime one can write the most general case as (White 1979; Peebles 1980; Fry 1984; Bernardeau & Schaeffer 1992; Szapudi & Szalay 1993; Bernardeau & Schaeffer 1999):

$$\xi_N^{\delta}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \sum_{\alpha,\mathrm{N-trees}} Q_{N,\alpha} \sum_{\mathrm{labellings}} \prod_{\mathrm{edges(i,j)}}^{(N-1)} \xi_2^{\delta}(\boldsymbol{r}_i, \boldsymbol{r}_j). \tag{15}$$

It is interesting to note that an exactly similar hierarchy develops in the quasi-linear regime in the limit of vanishing variance (Bernardeau 1992), however the hierarchical amplitudes $Q_{N,\alpha}$ become shape-dependent in such a case. In the highly non-linear regime there are some indications that these functions become independent of shape parameters as has been suggested by studies of the lowest order parameter $Q_3 = Q$ using high-resolution numerical simulations (Scoccimarro et al. 1998). In the Fourier space such an ansatz will mean that the whole hierarchy of multispectra B_δ, T_δ can be written in terms of sum of products of power-spectra P_δ, e.g., in low orders we can write:

$$B_\delta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3)_{\sum k_i=0} = Q(P_\delta(k_1)P_\delta(k_2) + P_\delta(k_2)P_\delta(k_3) + P_\delta(k_3)P_\delta(k_1))\delta_D(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3), \tag{16}$$

$$\begin{aligned} T_\delta(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3, \boldsymbol{k}_4)_{\sum k_i=0} = {} & [R_a P_\delta(k_1)P_\delta(\boldsymbol{k}_1 + \boldsymbol{k}_2)P_\delta(|\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3|) + \text{cyc.perm.} \\ & + R_b P_\delta(k_1)P_\delta(k_2)P_\delta(k_3) + \text{cyc.perm.}]\delta_D(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 + \boldsymbol{k}_4). \end{aligned} \tag{17}$$

Different hierarchical models differ in the way they predict the amplitudes of different tree topologies. Bernardeau & Schaeffer (1992) considered the case where amplitudes in general are factorizable, at each order one has a new 'star' amplitude and higher order 'snake' and 'hybrid' amplitudes are constructed from lower order 'star' amplitudes (see Munshi, Coles & Melott 1999a,b; Munshi, Melott & Coles 1999c for a detailed description). In models proposed by Szapudi & Szalay (1993), it is assumed that all hierarchical amplitudes of a given order are actually degenerate. We do not use any of these specific models for clustering and only assume the hierarchical nature of the higher order correlation functions. In the past, primarily data from galaxy surveys have been analysed extensively using these ansatze. Our main motivation here is to show that cross-correlation statistics of wl and the tSZ surveys can also be analysed using such techniques. The most general form for the lower order cumulant correlators in the large separation limit can be expressed as

$$\langle\kappa^2(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c = 2Q_3\hat{\mathcal{C}}_3[\mathcal{I}_{\theta_s}\mathcal{I}_{\theta_{12}}] = C_{21}^{\eta\eta'}\hat{\mathcal{C}}_3[\mathcal{I}_{\theta_s}\mathcal{I}_{\theta_{12}}] \equiv C_{21}^{\kappa y}\langle\kappa^2\rangle_c\langle\delta y(\hat{\Omega}_1)\kappa(\hat{\Omega}_2)\rangle_c, \tag{18}$$

$$\langle\kappa^3(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c = (3R_a + 6R_b)\hat{\mathcal{C}}_4[\mathcal{I}_{\theta_s}^2 I_{\theta_{12}}] = C_{31}^{\eta\eta'}\hat{\mathcal{C}}_4[\mathcal{I}_{\theta_s}^2 I_{\theta_{12}}] \equiv C_{31}^{\kappa y}\langle\kappa^2\rangle_c^2\langle\kappa(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c, \tag{19}$$

$$\langle\kappa^2(\hat{\Omega}_1)\delta y^2(\hat{\Omega}_2)\rangle_c = 4R_b\hat{\mathcal{C}}_4[\mathcal{I}_{\theta_s}^2\mathcal{I}_{\theta_{12}}] = C_{22}^{\eta\eta'}\hat{\mathcal{C}}_4[\mathcal{I}_{\theta_s}\mathcal{I}_{\theta_{12}}] \equiv C_{22}^{\kappa y}\langle\kappa^2\rangle_c\langle\delta y^2\rangle_c\langle\kappa(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c, \tag{20}$$

$$\langle\kappa^4(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c = (24S_a + 36S_b + 4S_c)\hat{\mathcal{C}}_5[\mathcal{I}_{\theta_s}^3\mathcal{I}_{\theta_{12}}] = C_{41}^{\eta\eta'}\hat{\mathcal{C}}_5[\mathcal{I}_{\theta_s}^3\mathcal{I}_{\theta_{12}}] \equiv C_{41}^{\kappa y}\langle\kappa^2\rangle_c^3\langle\kappa(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c, \tag{21}$$

$$\langle\kappa^3(\hat{\Omega}_1)\delta y^2(\hat{\Omega}_2)\rangle_c = (12S_a + 6S_b)\hat{\mathcal{C}}_5[I_{\theta_s}^3 I_{\theta_{12}}] = C_{32}^{\eta\eta'}\hat{\mathcal{C}}_5[\mathcal{I}_{\theta_s}^3\mathcal{I}_{\theta_{12}}] \equiv C_{32}^{\kappa y}\langle\kappa^2\rangle_c^2\langle\delta y^2\rangle_c\langle\kappa(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c. \tag{22}$$

Here, $C_{pq}^{\kappa y}$ denotes the cumulant correlators of the convergence field and $C_{pq}^{\eta\eta'}$ denotes the cumulant correlators for the underlying mass distribution. The subscript c denotes the *connected* components of the higher order correlation functions. The amplitudes $R_a = \nu_2^2$, $R_b = \nu_3$ and $S_a = \nu_2^3$, $S_b = \nu_3\nu_2$, $S_c = \nu_4$ are expressed in terms of the vertices which can be evaluated using Hyper Extended Perturbation Theory (HEPT; Scoccimarro et al. 1998). The exact expressions for various $C_{pq}^{\eta\eta'}$ presented in equation (22) are obtained by counting of the relevant

tree diagrams. In the limiting situation when $\hat{\Omega}_1 = \hat{\Omega}_2$ we can recover the corresponding cumulants. Extending the above results to arbitrary order we can write:

$$\langle \kappa^p(\hat{\Omega}_1)\delta y^q(\hat{\Omega}_2)\rangle_c = C^{\eta}_{pq}\hat{\mathcal{C}}_{p+q}[[\mathcal{I}_{\theta_s}]^{(p-1)}_{\rm wl}[\mathcal{I}_{\theta_s}]^{(q-1)}_{\rm sz}[\mathcal{I}_{\theta_{12}}]] = C^{\kappa y}_{pq}[\langle\kappa^2\rangle^{(p-1)}_c][\langle\delta y^2\rangle^{(q-1)}_c]\langle\kappa(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c. \tag{23}$$

This is a generalization of the usual definition of cumulant correlators for the case of two different fields in this particular case convergence κ and δy. The smoothing angular scales and the window function are left completely arbitrary. These definitions can also be used to define mixed normalized one-point $S^{\kappa y}_N$ parameters (with $N = p + q$) involving two different fields by considering the limiting situation $\hat{\Omega}_1 = \hat{\Omega}_2$.

$$\hat{\mathcal{C}}_{p+q}\left[[\mathcal{I}_{\theta_s}]^{p-1}\mathcal{I}_{\theta_{12}}\right] = \int_0^{r_0}\frac{\omega^p_{\rm wl}(r)\omega^q_{\rm sz}(r)b^q_\pi(r)}{{\rm d}^{2({\rm p+q-1})}_{\rm A}(r)}[\mathcal{I}_{\theta_s}]^{(p-1)}_{\rm wl}[\mathcal{I}_{\theta_s}]^{(q-1)}_{\rm sz}[\mathcal{I}_{\theta_{12}}]\,{\rm d}r; \qquad [\mathcal{I}_{\theta_s}]_{\rm wl} = \int\frac{{\rm d}^2\boldsymbol{l}}{(2\pi)^2}P_\delta\left(\frac{l}{{\rm d}_{\rm A}(r)}\right)W^2_{\rm TH}(l\theta_s); \tag{24}$$

$$[\mathcal{I}_{\theta_s}]_{\rm SZ} = \int\frac{{\rm d}^2\boldsymbol{l}}{(2\pi)^2}P_\delta\left(\frac{l}{{\rm d}_{\rm A}(r)}\right)b^2_l(\theta'_s); \qquad [\mathcal{I}_{\theta_{12}}] \equiv \int\frac{{\rm d}^2\boldsymbol{l}}{(2\pi)^2}P_\delta\left(\frac{l}{{\rm d}_{\rm A}(r)}\right)W_{\rm TH}(l\theta_s)b_l(\theta'_s)\exp(\,i\,\boldsymbol{l}\cdot\boldsymbol{\theta}_{12}). \tag{25}$$

Though results can be derived for different smoothing angular scales, for simplicity, we will only consider identical smoothing beam size $\theta_s = \theta'_s$. The hierarchical expression for the lowest order cumulant, i.e. S^κ_3 for convergence was derived by Hui (1999). He showed that his result agrees well with numerical ray tracing experiments. Later studies have shown that higher order cumulants and even the two-point statistics such as cumulant correlators can also be reliably modelled in a similar way (Munshi & Coles 2000; Munshi & Jain 2001). More recently, it was shown (Munshi et al. 2013) that the statistics of tSZ too can be modelled according to the same prescription. In particular, it was shown that the lognormal distribution can be used to predict the PDF and bias associated with the tSZ maps. We extend these results to the case of joint analysis of wl and tSZ maps.

We will develop these results further to construct the full joint 2PDF simply using the individual bias functions for tSZ and wl. The hierarchical ansatz allows us to write the joint 2PDF as

$$p^{(i)}_{\kappa y}(\kappa^{(i)}, y)\,{\rm d}\kappa^{(i)}{\rm d}y = p^{(i)}_\kappa(\kappa^{(i)})p_y(y)(1 + b^{(i)}_\kappa(\kappa)\xi^{\kappa y}_{12}(\theta_{12})b_y(y))\,{\rm d}\kappa^{(i)}\,{\rm d}y, \tag{26}$$

and its relation to the bias associated with collapsed objects in underlying density field $\eta = 1 + \delta$.

As an aside, it is simple to check that if we are dealing with tomographic redshift slices $\kappa^{(i)}$ and $\kappa^{(j)}$ of the same or different wl surveys the corresponding cumulant correlators are defined as

$$\hat{\mathcal{C}}^{ij}_{p+q}\left[[\mathcal{I}_{\theta_s}]^{p-1}_{\rm wl}[\mathcal{I}_{\theta_s}]^{q-1}_{\rm sz}[\mathcal{I}_{\theta_{12}}]\right] = \int_0^{r_0}\frac{\omega^p_{(i)}(r)\omega^q_{(j)}(r)b^q_\pi(r)}{{\rm d}^{2({\rm p+q-1})}_{\rm A}(r)}[\mathcal{I}_{\theta_s}]^{(p-1)}_{\rm wl}[\mathcal{I}_{\theta_s}]^{(q-1)}_{\rm sz}[\mathcal{I}_{\theta_{12}}]\,{\rm d}r. \tag{27}$$

The expressions for the cumulant correlators are constructed by replacing $y(\hat{\Omega})$ in equation (23) with $\kappa^{(j)}(\hat{\Omega})$:

$$\langle\kappa^p_{(i)}(\hat{\Omega}_1)\kappa^q_{(j)}(\hat{\Omega}_2)\rangle_c = C^{\eta}_{pq}\hat{\mathcal{C}}_{p+q}[[\mathcal{I}_{\theta_s}]^{(p+q-2)}[\mathcal{I}_{\theta_{12}}]] = C^{\kappa y}_{pq}\langle\kappa^2_{(i)}\rangle^{(p-1)}_c\langle\kappa^2_{(j)}\rangle^{(q-1)}_c\langle\kappa^{(i)}(\hat{\Omega}_1)\kappa^{(j)}(\hat{\Omega}_2)\rangle_c. \tag{28}$$

These expressions are quoted here for the highly non-linear regime. As is well known, a similar hierarchy develops in the quasi-linear regime so the same analytical tools are applicable in such a situation. We will next use these expressions to compute the lower order moment of one and two-point PDFs.

The corresponding results for the Ostriker–Vishniac effect (and kSZ effect; Castro 2003) will be presented elsewhere.

4 JOINT PROBABILITY DISTRIBUTION FUNCTION κ AND $\mathcal{Y}$

The primary aim in this section is to prove that with a suitable definition of reduced y-parameter, its statistics under certain approximation can be reduced to that of the underlying density contrast δ. The proof depends on a generic hierarchical ansatz and a scale independent biasing model. Analogously, we also define a reduced convergence parameter whose statistics reflects that of the underlying mass distribution. The problem of correlating κ and y then reduces to correlating η and η', and can be modelled using techniques developed for analysing correlation structure of the density distribution.

To compute the bias associated with peaks in the convergence field, we have to first develop an analytic expression for the generating function $\beta_{\kappa y}(z_1, z_2)$ for the convergence field κ and the tSZ field $\delta y = y - \langle y\rangle$. For that, we will use the usual definition for the mixed two-point cumulant correlator $C^{\kappa y}_{pq}$:

$$C^{\kappa y}_{pq} = \frac{\langle\kappa^p(\hat{\Omega}_1)\delta y^q(\hat{\Omega}_2)\rangle_c}{\langle\kappa^2\rangle^{p-1}_c\langle\delta y^2\rangle^{q-1}_c\langle\kappa(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c}. \tag{29}$$

(for a more detailed description of cumulant correlators see. This is a natural generalization of cumulant correlators found generally in the literature for individual fields (Munshi, Coles & Viel 2012). In the limiting case of $\hat{\Omega}_1 = \hat{\Omega}_2$, we recover the limiting case of one point cumulants: $S^{\kappa y}_{p+q} = C^{\kappa y}_{pq}$. We will show that, like its density field counterpart, the two-point generating function for the convergence field κ can also be expressed (under certain simplifying assumptions) as a product of two generating functions $\beta(z)$ which can then be directly related to the bias associated with 'hotspots' in the convergence field. It is clear that the factorization of generating function actually depends on the

factorization property of the cumulant correlators, i.e. $C^{\eta\eta'}_{pq} = C^{\eta}_{p1}C^{\eta'}_{q1}$. Note that, such a factorization is possible when the correlation of two patches in the directions $\hat{\Omega}_1$ and $\hat{\Omega}_2$ $\langle\kappa(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c$ is smaller compared to the variance $\langle y_s^2\rangle$ for the smoothed patches:

$$\beta_{\kappa y}(z_1, z_2) = \sum_{p,q}^{\infty} \frac{C^{\kappa y}_{pq}}{p!q!} z_1^p z_2^q = \sum_{p,q}^{\infty} \frac{1}{p!q!} \frac{z_1^p z_2^q}{\langle\kappa^2\rangle_c^{p-1}\langle\delta y^2\rangle_c^{q-1}} \frac{\langle\kappa^p(\hat{\Omega}_1)\delta y^q(\hat{\Omega}_2)\rangle_c}{\langle\kappa(\hat{\Omega}_1)\delta y(\hat{\Omega}_2)\rangle_c}. \tag{30}$$

Here, z_1 and z_2 are dummy variables. We will now use the integral expression for cumulant correlators (Munshi & Coles 2002) to express the generating function which in turn uses the hierarchical ansatz and the far field approximation as explained above:

$$\beta_{\kappa y}(z_1, z_2) = \sum_{p,q}^{\infty} \frac{C^{\kappa y}_{pq}}{p!q!} \frac{z_1^p}{\langle\kappa^2\rangle_c^{p-1}} \frac{z_2^q}{\langle\delta y^2\rangle_c^{q-1}} \frac{1}{\xi^{12}_{y\kappa}} \int_0^{r_0} \mathrm{d}r\ \mathrm{d}_{\mathrm{A}}^2(r) \frac{\omega^{\mathrm{p}}_{\mathrm{wl}}(r)\omega^{\mathrm{q}}_{\kappa}(r)}{[\mathrm{d}_{\mathrm{A}}(r)]^{2\mathrm{p}}[\mathrm{d}_{\mathrm{A}}(r)]^{2\mathrm{q}}} [\mathcal{I}_{\theta_s}]^{\mathrm{p}-1}_{\mathrm{wl}} [\mathcal{I}_{\theta_s}]^{\mathrm{q}-1}_{\mathrm{sz}} \mathcal{I}_{\theta_{12}}. \tag{31}$$

It is possible to further simplify the above expression by separating the summation over dummy variables z_1 and z_2, which will be useful to establish the factorization property of two-point generating function for bias $\beta_{\kappa y}(z_1, z_2)$. We can now decompose the double sum over the two indices into two separate sums over individual indices. We do not use any of these specific models for clustering and only assume the hierarchical nature of the higher order correlation functions. In the past, primarily data from galaxy surveys have been analysed extensively using these ansatze. The motivation here is to show that cross-correlation statistics of wl surveys against the tSZ surveys can also be analysed using such techniques. The most general result for the lower order cumulant correlators in the large separation limit can be expressed as

$$\beta_{\kappa y}(z_1, z_2) = \int_0^{r_0} \mathrm{d}r\ \mathrm{d}_{\mathrm{A}}^2(r) \frac{\mathcal{I}_{\theta_{12}}}{\xi^{\kappa y}_{12}} \frac{\langle\kappa^2\rangle_c}{[\mathcal{I}_{\theta_s}]_{\mathrm{wl}}} \beta_{\eta}\left(\frac{z_1}{\langle\kappa^2\rangle_c} \frac{\omega_{\mathrm{wl}}(r)}{\mathrm{d}_{\mathrm{A}}^2(r)} [\mathcal{I}_{\theta_s}]_{\mathrm{wl}}\right) \frac{\langle\delta y^2\rangle_c}{[\mathcal{I}_{\theta_s}]_{\mathrm{sz}}} \beta_{\eta'}\left(\frac{z_2}{\langle\delta y^2\rangle_c} \frac{\omega_{\mathrm{SZ}}(r) b_{\pi}(r)}{\mathrm{d}_{\mathrm{A}}^2(r)} [\mathcal{I}_{\theta_s}]_{\mathrm{sz}}\right). \tag{32}$$

The above expression is quite general and depends only on the small angle approximation (Limber 1954) and the large separation approximation, and is valid for any given specific model for the generating function $\mathcal{G}_{\delta}(\tau)$. However, it is easy to notice that the projection effects as encoded in the line-of-sight integration do not allow us to write down the two-point generating function $\beta_{\kappa y}(z_1, z_2)$ simply as a product of two one-point generating functions $\beta_{\kappa}(z_1)$ and $\beta_y(z_2)$, as generally is the case for the density field $\eta = 1 + \delta$.

As in the case of the derivation of the PDF for the smoothed convergence field κ, it will be much easier if we define a reduced smoothed convergence field η_s. The statistical properties of η_s are very similar to that of the underlying 3D density field (under certain simplifying approximation), and are roughly independent of the background geometry and dynamics of the universe. In a similar manner, we also define the reduced y field η_s'. We define the reduced convergence η_s and the reduced tSZ y-map η_s', respectively, by the following expressions: $\eta_s = (\kappa - \kappa_m)/(-\kappa_m) = (1 + \kappa/|\kappa_m|)$ and $\eta_s' = \delta y/\langle y\rangle$, where the minimum value of κ_m is defined as $\kappa_m = -\int_0^{r_s} \mathrm{d}r\ \omega_{\mathrm{wl}}(r)$ and $\langle y\rangle$ is the average value of the y-parameter. It is easy to notice that the minimum value of the convergence field will occur in those lines of sight which are completely devoid of any matter, i.e. $\delta = -1$ all along the line of sight. We will also find out later that the cosmological dependence of the statistics of the κ field is encoded in k_m, and this choice of the new variable η_s makes its related statistics almost independent of the background cosmology. Repeating the above analysis again for the η_s field, we can express the cumulant correlator generating function for the reduced convergence field η_s as

$$\beta_{\eta\eta'}(z_1, z_2) = \int_0^{r_0} \mathrm{d}r \frac{1}{|\kappa_m|} \frac{1}{\langle y\rangle} \mathrm{d}_{\mathrm{A}}^2(r) \frac{\mathcal{I}_{\theta_{12}}}{\xi^{\kappa y}_{12}} \frac{\langle\kappa^2\rangle}{[\mathcal{I}_{\theta_s}]_{\mathrm{wl}}} \beta_{\eta}\left(|\kappa_m| \frac{z_1}{\langle\kappa^2\rangle_c} \frac{\omega_{\mathrm{wl}}(r)}{\mathrm{d}_{\mathrm{A}}^2(r)} [\mathcal{I}_{\theta_s}]_{\mathrm{wl}}\right) \frac{\langle\delta y^2\rangle}{[\mathcal{I}_{\theta_s}]_{\mathrm{sz}}} \beta_{\eta'}\left(\langle y\rangle \frac{z_2}{\langle\delta y^2\rangle_c} \frac{\omega_{\mathrm{SZ}}(r) b_{\pi}(r)}{\mathrm{d}_{\mathrm{A}}^2(r)} [\mathcal{I}_{\theta_s}]_{\mathrm{sz}}\right). \tag{33}$$

While the above expression is indeed very accurate and relates the generating function of the density field with that of the convergence field, it is difficult to handle for any realistic practical applications.

In a scaling analysis, the scaling function $h(x)$ can be expressed as an inverse Laplace transform of the generating function $\phi(z)$ which is a generating function of the normalized one-point cumulants S_p (a detailed relevant discussion can be found in Munshi et al. 2013):

$$h(x) = -\frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \exp(x\, z)\phi(z); \qquad \phi(z) \equiv \sum_p S_p \frac{z^p}{p!}. \tag{34}$$

Similarly, the bias function $b(x)$ in scaling analysis is related to the function $\tau(z)$. The generating function defined as $\tau(z) = \sum_p z^p C_{p1}/p!$ acts as a generator for the C_{p1} parameters. The entire hierarchy of the C_{pq} can be constructed from the C_{p1} parameters, i.e. $C_{pq} = C_{p1}C_{q1}$ at large separation limit.

$$h(x)b(x) = \frac{1}{2\pi i} \int_{-i\infty}^{i\infty} \exp(x\, z)\tau(z); \qquad \tau(z) \equiv \sum_p C_{p1} \frac{z^p}{p!}. \tag{35}$$

It is important to notice that the scaling functions such as $h(x)$ for the density PDF and $b(x)$ for the bias associated with overdense objects (with $x = (1 + \delta)/\bar{\xi}_2^{\delta}$) are typically estimated from numerical simulations especially in the highly non-linear regime ($\bar{\xi}_2^{\delta}$ is the volume average of two-point correlation function). Such estimations are plagued by several uncertainties, such as the finite size of the simulation box. It was noted in earlier studies that such uncertainties lead to only a rather approximate estimation of $h(x)$. The estimation of the scaling function associated with the bias, i.e. $b(x)$ is even more complicated due to the fact that the two-point quantities such as the cumulant correlators and the bias are more affected by the finite size of the catalogues. So it is not fruitful to actually integrate the exact integral expression we have derived above and we will replace all line-of-sight integrals with its approximate values. The previous study by Munshi & Jain (2000) have used an exactly similar approximation to simplify the one-point PDF for κ and found good agreement with ray tracing simulations. We will

show that our approximation reproduces the numerical results quite accurately for a wide range of smoothing angle,

$$|\kappa_m| \approx \frac{1}{2} r_s \omega_{\rm wl}(r_c); \qquad \langle y \rangle \approx \frac{1}{2} r_s \omega_{\rm SZ}(r_c) b_\pi(r_c); \qquad 0 < r_c < r_s; \tag{36}$$

$$\langle \kappa^2 \rangle_c \approx \frac{1}{2} r_s \frac{\omega^2_{\rm wl}(r_c)}{{\rm d}^2_{\rm A}(r_c)} \left[\int \frac{{\rm d}^2 \boldsymbol{l}}{(2\pi)^2} P_\delta \left(\frac{l}{{\rm d}_{\rm A}(r_c)} \right) W^2_{\rm TH}(l\theta_s) \right]; \qquad \langle \delta y^2 \rangle_c \approx \frac{1}{2} r_s \frac{b^2_\pi(r_c) \omega^2_{\rm SZ}(r_c)}{{\rm d}^2_{\rm A}(r_c)} \left[\int \frac{{\rm d}^2 \boldsymbol{l}}{(2\pi)^2} P_\delta \left(\frac{l}{{\rm d}_{\rm A}(r_c)} \right) b^2_l(\theta_s) \right]; \tag{37}$$

$$\langle \kappa(\hat{\Omega}_1) \delta y(\hat{\Omega}_2) \rangle_c \approx \frac{1}{2} r_s \frac{\omega_{\rm SZ}(r_c) \omega_{\rm wl}(r_c)}{{\rm d}^2_{\rm A}(r_c)} \left[\int \frac{{\rm d}^2 \boldsymbol{l}}{(2\pi)^2} P_\delta \left(\frac{l}{{\rm d}_{\rm A}(r_c)} \right) W_{\rm TH}(l\theta_s) b_l(\theta_s) \exp[i \boldsymbol{l} \cdot \theta_{12}] \right]. \tag{38}$$

Use of these approximations gives us the leading-order contributions to these integrals, and we can check that to this order we recover the factorization property of the generating function, i.e. $\beta_{\eta\eta'}(z_1, z_2) = \beta_\eta(z_1)\beta_{\eta'}(z_2)$,

$$\beta_{\eta\eta'}(z_1, z_2) = \beta_\eta(z_1)\beta_{\eta'}(z_2) = \beta_{1+\delta}(z_1)\beta_{1+\delta'}(z_2) \equiv \tau(z_1)\tau(z_2). \tag{39}$$

So it is clear that at this level of approximation, due to the factorization property of the cumulant correlators, the bias function $b_\eta(x)$ associated with the peaks in the convergence field κ, beyond certain threshold, obeys a similar factorization property too, which is exactly the same as its density field counterpart. Earlier studies have established such a correspondence between the convergence field and the density field in case of the one-point PDF $p(\delta)$ (Munshi & Jain 2000),

$$b_\eta(x_1)h_\eta(x_1)b_{\eta'}(x_2)h_{\eta'}(x_2) = b_{1+\delta}(x_1)h_{1+\delta}(x_1)b_{1+\delta'}(x_2)h_{1+\delta'}(x_2). \tag{40}$$

The following relation between $\beta_\eta(z)$ and $b_\eta(x)$ is useful for modelling. However, for all practical purpose, we find that the differential bias $\beta_{>\eta}(y)$ as defined is simpler to measure from numerical simulations due to its cumulative nature:

$$b_\eta(x)h_\eta(x) = -\frac{1}{2\pi i} \int_{-i\infty}^{i\infty} {\rm d}z \tau(z) \exp(xz); \quad b_\eta(>x)h_\eta(>x) = -\frac{1}{2\pi i} \int_{-i\infty}^{i\infty} {\rm d}z \frac{\tau(z)}{z} \exp(xz). \tag{41}$$

Using these results we can finally write down:

$$p_{\eta\eta'}(\eta, \eta')\,{\rm d}\eta\,{\rm d}\eta' = p_\eta(\eta)p_{\eta'}(\eta')(1 + b_\eta(\eta)\xi_{12}^{\eta\eta'} b_{\eta'}(\eta'))\,{\rm d}\eta\,{\rm d}\eta'; \quad \eta = \kappa_s/|\kappa_m|; \quad \eta' = \delta y/\langle y \rangle; \qquad \xi_{12}^{\eta\eta'} = \langle \eta(\hat{\Omega}_1)\eta'(\hat{\Omega}_2) \rangle. \tag{42}$$

It is important to notice that although the bias $b_\eta(x)$ associated with the reduced convergence field and the underlying density field are exactly same, the variance associated with the density field is very high but the projection effects in the convergence field bring down the variance in the convergence field to less than unity which indicates that we have to use the integral definition of bias to recover it from its generating function (see Appendix A for more details). Finally, writing down the joint PDF $p_{\kappa y}(\kappa, y)$ for the smoothed projected convergence field κ and the y in terms of their reduced versions η and η':

$$p_{\kappa y}(\kappa, y)\,{\rm d}\kappa\,{\rm d}y = p_\kappa(\kappa)p_y(y)(1 + b_\kappa(\kappa)\,\xi_{12}^{\kappa y}\,b_y(y))\,{\rm d}\kappa\,{\rm d}yn;$$

$$b_y(y) = b_\eta(y/\langle y \rangle)/\langle y \rangle; \quad b_\kappa(\kappa) = b_\eta(\kappa/|\kappa_m|)/|\kappa_m|; \qquad \xi_{12}^{\kappa y} \equiv \langle \kappa(\hat{\Omega}_1) y(\hat{\Omega}_2) \rangle = \xi_{12}^{\eta\eta'}/(|\kappa_m|\langle y \rangle). \tag{43}$$

In an earlier study, Munshi et al. (1999b) used similar arguments for the convergence maps to show that $p_\kappa(\kappa) = p_\eta(\kappa/|\kappa_m|)/|\kappa_m|$, and recently it was generalized by (Munshi et al. 2011) for the y-maps to $p_y(y) = p_\eta(y/\langle y \rangle)/\langle y \rangle$. If we notice that $\xi_{12}^{\kappa y} = \xi_{12}^{\eta\eta'}/(|\kappa_m|\langle y \rangle)$; then the above expressions helps us to write $b_\kappa(\kappa) = b_\eta(\kappa/|\kappa_m|)/|\kappa_m|$ and $b_y(y) = b_\eta(y/\langle y \rangle)/\langle y \rangle$.

Equation (43) is the main result of our analysis – it provides an expression for the joint PDF of κ and y which encapsulates knowledge of cumulant correlators of all orders. We have shown that we can define two variables η (reduced convergence) and η' (reduced y-parameter) that can simplify the derivation considerably. Assuming a scale independent biasing model, we can directly link the statistics of y-parameter to that of the underlying density contrast δ. Using these variables, we have shown that the PDF and bias of y can be directly linked to that of δ. The mixed lower order cumulant correlators involving κ and y-maps can then be expressed in terms of the bias functions of κ and y. We also stress here that the presence of significant non-gravitational effect will violate the assumptions that we have used in our analytical derivation.

A similar result can of course be derived for cross-correlation of tSZ y-map and wl convergence maps $\kappa^{(i)}$ from individual tomographic bins. The joint PDF $p^{(i)}(\kappa^{(i)}, y)$ for tomographic maps $\kappa^{(i)}$ and projected y-maps can be expressed in terms of the individual PDF $p^{(i)}(\kappa)$, $p(y)$ maps and the bias $b^{(i)}_\kappa(\kappa)$ and $b_y(y)$:

$$p^{(i)}(\kappa^{(i)}, y)\,{\rm d}\kappa^{(i)}\,{\rm d}y = p^{(i)}(\kappa)p(y)(1 + b^{(i)}_\kappa(\kappa^{(i)})\,\xi^{(i)}_{12}\,b_y(y))\,{\rm d}\kappa^{(i)}{\rm d}y;$$
$$b^{(i)}_\kappa(\kappa) = b_\eta(\eta)/|\kappa^{(i)}_m|; \quad b_y(y) = b_\eta(\eta')/\langle y \rangle; \quad \xi^{(i)}_{12} \equiv \xi_{12}^{\eta\eta'}/(|\kappa^{(i)}_m|\langle y \rangle). \tag{44}$$

Notice that, the bias and one point PDF for both the projected convergence maps κ and tomographic maps $\kappa^{(i)}$ as well as for the tSZ map y are all constructed from the same underlying PDF p_η and the bias b_η associated with the 3D density distribution. The individual PDF $p_\kappa(\kappa)$ and $p_y(y)$ and the bias $b_\kappa(\kappa)$ and $b_y(y)$ have already been tested against numerical simulations and were found to be remarkably successful.

This technique will also allow for the computing of joint PDF of convergence maps for CMB lensing and the tSZ y-maps. This can be achieved by replacing the κ_m in equation (43) with corresponding κ_m for the LSS, i.e. $\kappa_m = -\int_0^{r_{\rm LSS}} {\rm d}r\, w_{\rm wl}(r)$, where the source distance r_s

is now replaced by the comoving distance to the LSS $r_{\rm LSS}$. This is especially relevant given the recent evidence for a CMB lensing signal with *WMAP* and ACT (Das et al. 2011; Smidt et al. 2011). The results for the cumulant correlators can be modified analogously. Though, the results in equations (43) and (44) are derived using the hierarchical ansatz, the final results are remarkably independent of details of any of the assumptions that were used in deriving them. This is an indication of their more generic validity. Indeed, it has been shown that the PDF and the bias for the tSZ and wl fields can be constructed from equivalent but other valid descriptions of PDF and bias of underlying 3D density contrast. The lognormal distribution for the underlying PDF and bias provide one such description. In Appendix B, we have provided a short description of the lognormal distribution. The lognormal model based PDF and bias for κ and y are given by

$$p_\kappa(\kappa) = p_{\rm ln}(\kappa/|\kappa_m|)/|\kappa_m|; \quad p_y(\delta y) = p_{\rm ln}(\delta y/\langle y\rangle)/\langle y\rangle; \quad b_\kappa(\kappa) = b_{\rm ln}(\kappa/|\kappa_m|)/|\kappa_m|; \quad b_y(\delta y) = b_{\rm ln}(\delta y/\langle y\rangle)/\langle y\rangle. \tag{45}$$

The resulting PDF and bias matches with the ones from perturbative calculations for higher smoothing angular scales.

As mentioned earlier, the joint probability of $p_{\kappa y}(\kappa, y)$ is a noisy statistics. The integrated measure or cumulative PDF of the convergence crossing a threshold κ and the tSZ field y crossing a threshold κ_0 and y_0, respectively, is given by

$$p_{\kappa y}(>\kappa_0, > y_0) = \int_{\kappa_0}^{\infty} {\rm d}\kappa \int_{y_0}^{\infty} {\rm d}y\ p_{\kappa y}(\kappa, y); \quad p_\kappa(>\kappa_0) = \int_{\kappa_0}^{\infty} {\rm d}\kappa\, p_\kappa(\kappa); \quad p_y(>y_0) = \int_{y_0}^{\infty} {\rm d}y p(y); \tag{46}$$

$$b_\kappa(>\kappa_0) = \int_{\kappa_0}^{\infty} {\rm d}\kappa\, p(\kappa) b_\kappa(\kappa) \Big/ \int_{\kappa_0}^{\infty} {\rm d}\kappa\, p(\kappa); \qquad b_y(>y_0) = \int_{\kappa_0}^{\infty} {\rm d}\kappa\, p_y(y) b_y(y) \Big/ \int_{y_0}^{\infty} {\rm d}y p_y(y). \tag{47}$$

The following statistics can be constructed from the cumulative bias to probe the joint SZ and wl cross-correlation to all orders:

$$\mathcal{B}(>\kappa_0, > y_0) = [b_\kappa(>\kappa_0) b_y(>y_0)]^{1/2} = \frac{1}{\sqrt{\xi_{12}^{\kappa y}}} \left[\frac{p_{\kappa y}(>\kappa_0, > y_0)}{p_\kappa(>\kappa_0) p_y(>y_0)} - 1\right]^{1/2}. \tag{48}$$

It is important to notice that our modelling of the PDF p_κ and p_y is independent of the modelling of respective variances. The variance calculations for y depend on inputs such as the detailed modelling of bias and its redshift evolution that can be independently checked by comparing with simulations. However, given a correct variance the lognormal model or the hierarchical ansatz can be used to model the entire PDF as we have shown. For construction of the PDF of the scaled or reduced tSZ, i.e. η we also need the average $\langle y\rangle$. The modelling of average $\langle y\rangle$ is independent of the construction of PDF. However, what we have shown here is that given these two inputs we can reliably predict the PDF of y. While the PDF of reduced tSZ field η is independent of cosmology, the dependence of y-parameter is encoded in the definition of reduced y field $\eta' = \delta y/\langle y\rangle$. The convergence PDF is defined in terms of the PDF for the reduced convergence $\eta = \kappa/|\kappa_m|$. The PDF for the reduced convergence κ as well as the scaled tSZ map y are both shown to be the same as the PDF of underlying ark matter distribution δ. The dependence of p_κ on cosmological parameters is encoded in the definition of κ_m. The cosmological parameter dependence of bias $b_\kappa(\kappa)$ and $b_y(y)$ is encoded in $\langle y\rangle$ and κ_m.

It is also interesting that our calculations show that the PDF of *projected* or 2D convergence as well as the tSZ maps are described by the underlying 3D density contrast δ which is in agreement with (Munshi & Jain 2000, 2001), and provides robust mathematical justification for the use of lognormal model and its link to linear biasing prescription.

A simple order of magnitude estimate for the wl cumulants is $S_p^\kappa = \langle\kappa_s^p\rangle/\langle\kappa_s^2\rangle^{p-1}$ which is given by $S_p^\kappa = S_p/|\kappa_m|^{p-2}$. The cumulants of the field $\delta y/\langle y\rangle$ simply follow that of the underlying density contrast. The normalized cumulant correlators of the smoothed convergence field κ and the field $\delta y/\langle y\rangle$ is given by $C_{pq}^{\kappa y} = C_{pq}^{\eta\eta'}/|\kappa_m|^{p-1}$.

The numerical results for the cumulative bias for reduced convergence $b(>\eta)$ are shown in Fig. 5. In the left-hand panel, we show results for smoothing angular scale $\theta_s = 10$ arcmin and in the right-hand panel, we show results for $\theta_s = 5$ arcmin The results are displayed

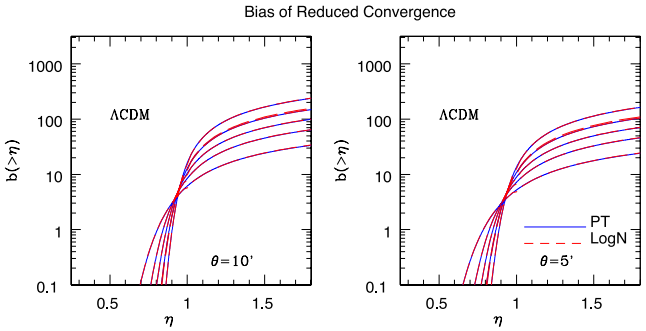

Figure 5. The bias function $b(\eta)$ defined in equation (45) of the reduced convergence is depicted for two different smoothing angular scales. The dashed lines show the analytical predictions from the hierarchical ansatz and the solid lines show predictions from the lognormal distribution. The left-hand panel corresponds to $\theta_s = 10$ arcmin and the right-hand panel corresponds to $\theta_s = 5$ arcmin. The different curves in each panel correspond to different redshift bins. The lower curves correspond to shallower (lower redshift) bins and the higher curves correspond to deeper (higher redshift) bins. The top-most curve corresponds to the projected survey without any tomographic information.

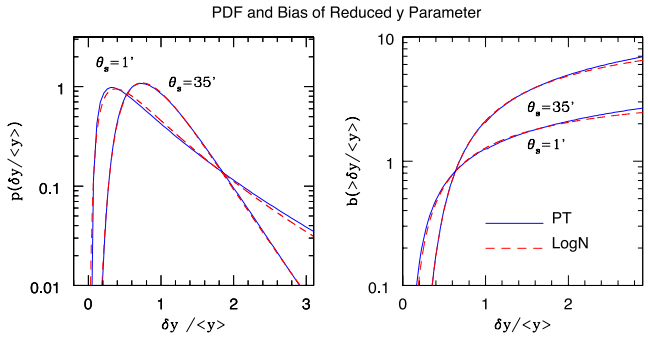

Figure 6. The PDF defined in equation (46) (left-hand panel) and bias function defined in equation (47) (right-hand panel) $b(>\delta y/\langle y\rangle)$ are plotted as a function of $\eta' = \delta y/\langle y\rangle$. They are computed using two different approaches. The dashed lines show the analytical predictions from the hierarchical ansatz and the solid lines show the predictions from the lognormal distribution. The lines correspond to different angular scales as depicted.

for different redshift bins. Lower curves correspond to lower redshift bins. We have presented results for both lognormal distribution as well as perturbative calculations.

The corresponding results for the reduced y-parameter are shown in Fig. 6. The left-hand panel shows the PDF for δy and the right-hand panel shows the cumulative bias. The angular scales are $\theta_s = 1$ and $\theta_s = 35$ arcmin. Both results of perturbative calculations and log-normal approximations are shown in each panel.

5 HYDRODYNAMICAL (SPH) SIMULATIONS

The simulated y-maps that we have used were generated by Kay et al. (2012) using *millennium gas simulations* (Hartley et al. 2008; Starek, Rudd & Evrard 2009; Short et al. 2010; Young et al. 2011). Which in turn were generated to provide hydrodynamic versions of the Virgo consortium's[12,13] dark matter Millennium Simulations and were performed using publicly available GADGET2 N-body/hydrodynamics code (Springel 2005). Two different versions of the simulations use same initial conditions and box size. In the first run, the gas was modelled as an ideal non-radiative fluid and was allowed to go adiabatic changes in regions of non-zero pressure gradient. The evolution was modelled using smooth particle hydrodynamics (sph). An artificial viscosity too was used to convert bulk kinetic energy of the gas into its internal energy. This is essential to capture the physics of shock, and thus generate quasi-hydrostatic equilibrium. These process ensures quasi-hydrostatic equilibrium inside vitalized haloes. See text for more details of the hydrodynamic simulations used to generate these maps. These set of simulations will be referred as GO simulations. Non-radiative descriptions of inter-cluster gas do not reproduce the observed X-ray properties of the clusters (Voit 2005). So the next set of simulations that we use pre-heated gas at high redshift that can generate the required core entropy and capable of producing a steeper X-ray luminosity-temperature in agreement with observations. The entropy level of these second set of simulations were chosen to match the mean X-ray luminosity temperature relation at $z = 0$ (Edward & Henry 1991; Kaiser 1991). These simulations also include radiative cooling and an entropy sink. We will refer to these simulations as PC. Cooling in these simulations does not play an important role as the cooling time for the preheated gas is long compared to the Hubble time. The cosmological parameters of these simulations are $\Omega_{\rm CDM} = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_{\rm b} = 0.045$, $h = 0.73$ and $\sigma_8 = 0.9$. The projected mass distribution for both GO and PC simulations are presented in Fig. 7. The scaled $\log_{10}[y/\langle y\rangle]$ parameter distribution of a realization is shown in Fig. 8 for GO simulations and Fig. 9 for simulations with pre-heating and cooling (PC). The left-hand panels show contribution from all individual components. The middle panels represents contribution from overdense regions that satisfy the constraint $1 + \delta < 100$. Finally, the right-hand panels correspond to the contribution to the y-map from gas which satisfy the constraint $\mathrm{T} < 10^5$ K.

There is a very clear and obvious difference between the two sets of maps in that the GO maps have more substructure. The smoothness of the PC maps is due to the external thermal energy added to the gas by the pre-heating process. The mean y-parameter in the GO simulations is $\langle y\rangle = 2.3 \times 10^{-6}$ and in the PC simulations it is nearly four times higher $\langle y\rangle = 9.9 \times 10^{-6}$. These values are consistent with *COBE*/Far Infrared Absolute Spectrophotometer constraint $\langle y\rangle \leq 1.5 \times 10^{-5}$. However, it is believed such a high level of pre-heating would definitely remove some of the absorption features seen in the Lyman α spectrum observed towards quasars (Theuns, Mo & Schaye 2001; Shang, Crotts & Haiman 2007; Borgani & Viel 2009). Indeed, the PC model studied here should be treated as an *extreme test* of the effect of a high pre-heating scenario.

In terms of source contributions, the bulk of the y-signal comes from low redshift, i.e. $z < 2$. However, in case of PC simulations the opposite is true, where 80 per cent of the signal originates from $z < 3.5$. The overdense regions such as the group or clusters are the sources of y-signal in the GO simulations which are primarily embedded in structures that collapsed at relatively lower redshift. In case of the PC simulations most of the signals comes from mildly over-dense gas at high redshift. It is interesting to notice that the GO simulations do get contributions from the gas at high redshift $z > 4$. However, such contributions are completely erased in case of the PC simulations. This

[12] http://www.virgo.dur.ac.uk/
[13] http://www.mpa-garching.mpg.de/galform/millennium/

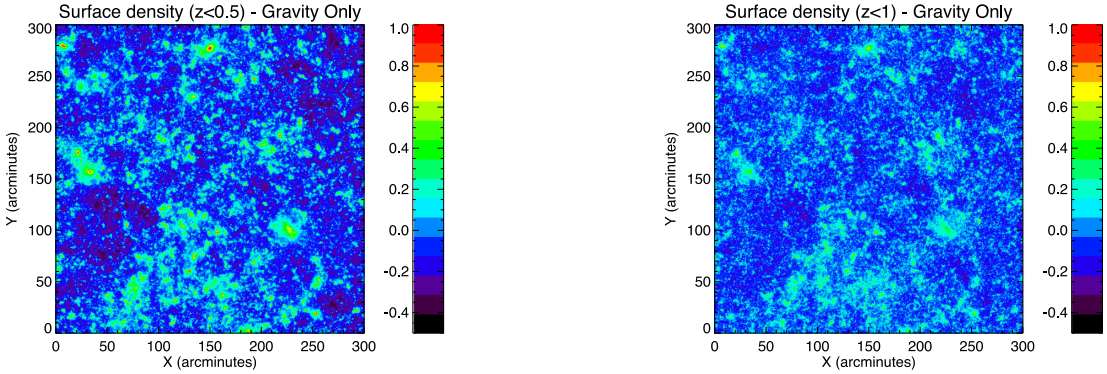

Figure 7. Simulated $5^\circ \times 5^\circ$ projected density contrast maps $\log_{10}[1+\delta]$ are depicted. The maps were generated using Virgo consortium's Millennium Gas Simulations. Maps were constructed using line-of-sight integration out to redshift $z = 0.5$ (left-hand panel) and $z = 1$ (right-hand panel). We will be using these projected mass maps as a proxy for wl convergence to test analytical predictions.

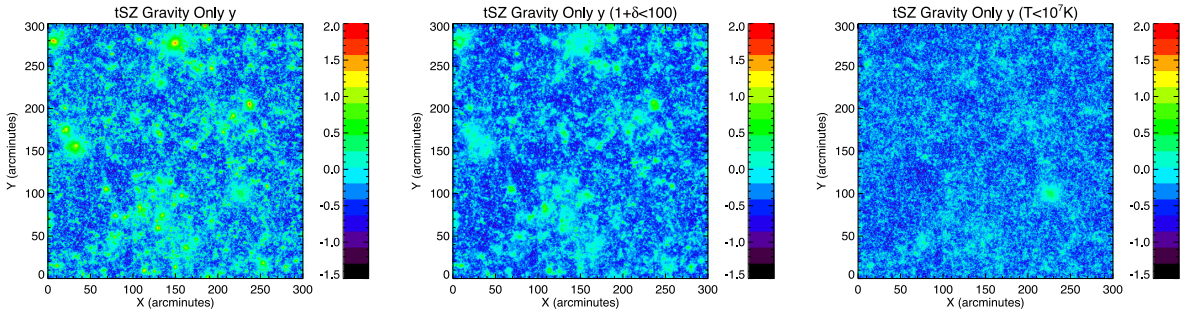

Figure 8. Simulated $5^\circ \times 5^\circ$ dimensionless scaled tSZ maps $\log_{10}[y/\langle y\rangle]$ are depicted. The maps were generated using Virgo consortium's Millennium Gas Simulation. The left-hand panel shows the resulting y-map. The middle panel corresponds to map generated using low-density regions and will be referred to as a y_ρ map. Only overdense regions with density $1+\delta < 100$ were considered. The right-hand panel correspond to low-temperature regions $T < 10^7$K and will be referred to as a y_T map. These set of hydrodynamic simulations ignore pre-heating but take into account adiabatic cooling. We will refer to them as GO simulations (see the text for more details). We use $\langle y\rangle$ of the total signal when showing the results for low density and temperatures.

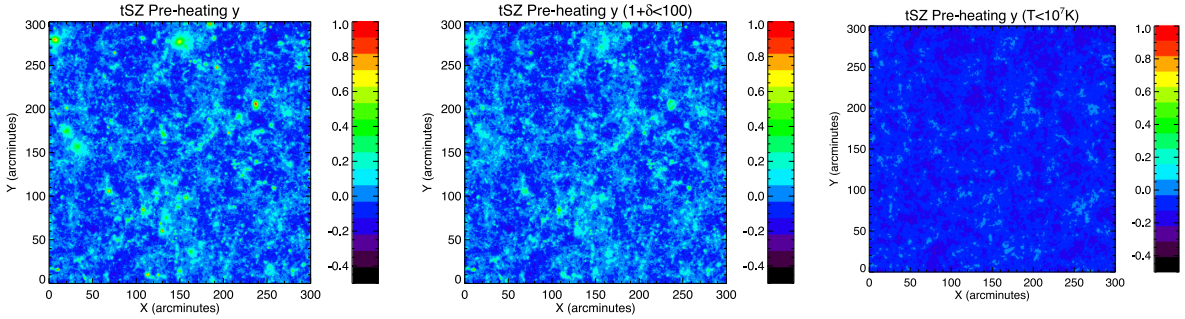

Figure 9. Same as previous figure, but for simulations with PC. These simulations will be referred to as PC simulations.

is primarily due to the fact that radiative cooling erases most of the ionized gas at these redshifts. Due to unavailability of joint tSZ and ray-tracing simulations for wl, we have used the projected density as a proxy for wl convergence. Indeed, wl convergence is identical to projected density contrast with a redshift dependent weight factor (Munshi & Jain 2000, 2001).

6 TESTS AGAINST NUMERICAL SIMULATION

In this section, we provide a detailed discussion of the statistics we employed to analyse the simulation.

Cumulants and cumulant correlators. We have computed the lower order moments (cumulants) namely the skewness parameter S_3 and the kurtosis parameter S_4 from both the GO and PC simulations. These results are presented in Figs. 10 (skewness) and 11 (kurtosis), respectively. The range of smoothing angular scale θ_s probed corresponds roughly to 15 arcsec$-$2 arcmin. We find that at angular scales $\theta_s > 1$ arcmin the skewness computed from GO simulation matches very closely with analytical predictions. In addition to the skewness parameter, the kurtosis parameter too matches with theoretical predictions at smoothing angular scales $\theta_s > 1$arcmin. For both the redshift bins, we have tested that the numerical results are close to lognormal predictions $S_3 = 3$ and $S_4 = 16$. While the higher order moments are reproduced well

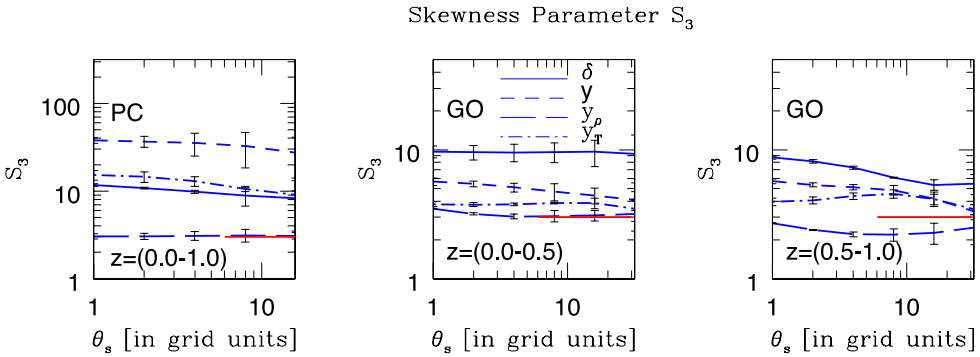

Figure 10. The skewness parameter $S_3 = \langle\delta^3\rangle_c/\langle\delta^2\rangle^2$, signifying the lowest order departure from Gaussianity is plotted as a function of the smoothing angular scale θ_s for various simulations. The size of the grid interval is 15 arcsec. The left-hand panel shows the results from PC simulations. The middle and right-hand panels correspond to GO simulations. The redshift range considered is $z = (0.0-1.0)$ for the PC simulations. Two different redshift ranges are considered for the GO simulations. The middle panel corresponds to $z = (0.0-0.5)$ and the right-hand panel corresponds to $z = (0.5-1.0)$. The scatter or error bars are computed using three realizations for each simulation type. Various lines correspond to density contrast δ, y-parameter (short-dashed lines), y-parameter contributed only by the low-temperature phase (long-dashed lines) and the low-density phase (dot–dashed lines), respectively. The smooth solid lines in middle and right-hand panels correspond to the lognormal prediction for S_3, i.e. $S_3 = 3$. The finite survey size not only introduces scatter or variance but it also introduces a bias in the estimator (see the text for more details.)

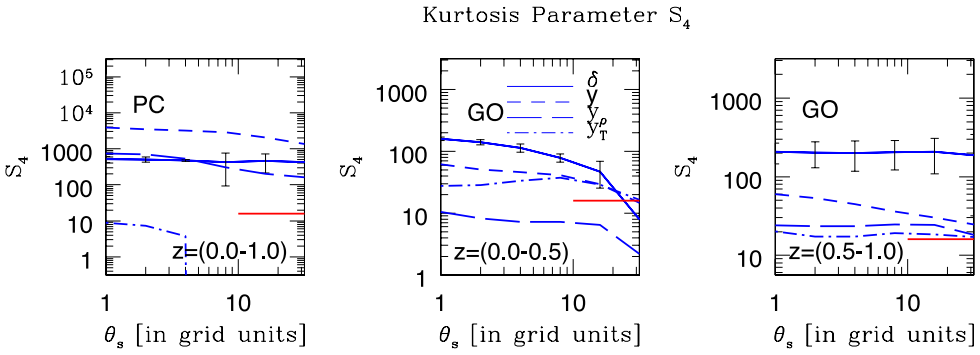

Figure 11. The kurtosis parameter $S_4 = \langle\delta^4\rangle_c/\langle\delta^2\rangle^3$ and similarly for other maps, signifying the next to leading-order departure from Gaussianity is plotted as a function of the smoothing angular scale θ_s for various simulations (in grid units). The size of the grid is 15 arcsec. The left-hand panel shows the results from PC simulations. The middle and right-hand panel correspond to GO simulations. The redshift range considered is $z = (0.0-1.0)$ for the PC simulations. Two different redshift ranges are considered for the GO simulations. The middle panel corresponds to the redshift $z = (0.0-0.5)$ and the right-hand panel corresponds to $z = (0.5-1.0)$. The scatter (or error bars) are computed using three realizations for each simulation type. Various lines correspond to different maps we have studied, i.e. density contrast δ (solid lines), y-parameter (short-dashed lines), y-parameter contributed only by the low-temperature phase (long-dashed lines) and the low-density phase (dot–dashed lines), respectively. The smooth solid lines in middle and right-hand panels correspond to the lognormal prediction for S_4 at large smoothing angular scale, i.e. $S_4 = 16$ (see the text for more details).

in GO simulations, the results from PC simulations where non-Gravitational effects are dominant, do not match with theoretical predictions. This is consistent with our previous studies, where we found that only in case of the GO simulations, where gravitational dynamics and adiabatic cooling are the main factors influencing structure formation, the numerical PDF is reasonably reproduced by theoretical predictions. The lognormal prediction starts to break down at smaller scale and the linear biasing scheme too is less accurate at scales smaller than few arcminutes. We have used three realisations to compute the scatter in S_3 and S_4. Notice that, small survey size can introduce not only a scatter but a significant bias in estimation of S_N parameters in general at large θ_s. We have used a 1024×1024 grid to compute the higher order moments. The lowest probability that we can resolve using this grid is roughly 10^{-6}. For larger θ_s, the number is higher.

In addition to the cumulants for the y-parameter, we have also studied the cumulant correlators C_{21}^{yy} defined in equation (29) for y-parameter, and presented in Fig. 12. The mixed cumulant correlator $C^{y\kappa}$ is presented in Fig. 13. The results for cumulants and their correlators follow a very similar pattern. The relevant scales for cumulant correlators are the smoothing angular scale θ_0 and the separation angular scale θ_{12}. The perturbative results are typically obtained in the limiting case of $\theta_{12} \gg \theta_s$. For numerical implementation, we have chosen $\theta_{12} \approx 4\theta_0$. It is important to realize that being two-point statistics the cumulant correlators are affected a lot more by the finite volume effect than their one-point counterparts. The theoretical predictions for the lowest order predictions for cumulant correlator for lognormal distribution is $C_{21} = 2$. The theoretical results match reasonably well with numerical results for GO simulations. Both cumulants and cumulant correlators can be used to differentiate the effect of non-gravitational processes as well as for sanity checks.

Contributions from individual components. In addition to studying the $\hat{y}$-maps, we have divided the entire contribution from various baryonic components, to check how our theoretical prescriptions compare with that from simulations for individual components. In this context, we notice that, thermodynamic states of baryons, as well as their clustering, at low to medium redshift $z < 5$, has been studied,

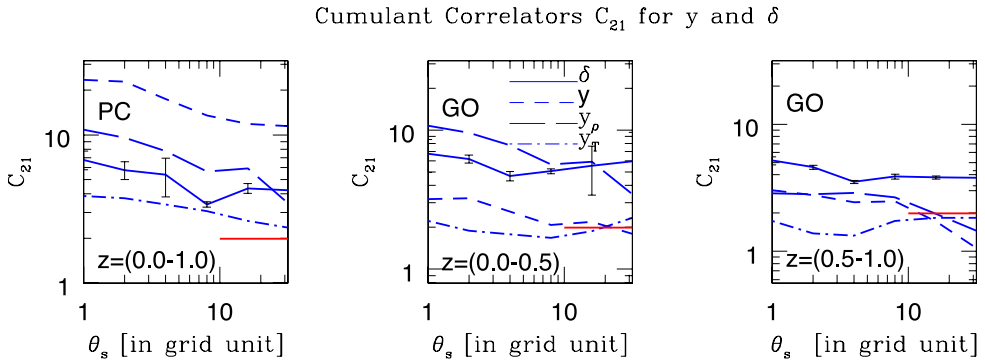

Figure 12. The normalized cumulant correlators $C_{21}^{\delta\delta} \equiv \langle\delta_1^2\delta_2\rangle_c/\langle\delta_1^2\rangle\langle\delta_1\delta_2\rangle$ defined in equation (29) where $(\delta_1 \equiv \delta(\hat{\Omega}_1))$ and equivalently for other maps are plotted as a function of smoothing angular scale θ_0. The left-hand panel shows results from PC simulations. The other two panels correspond to GO simulations. The solid line corresponds to $C_{21}^{\delta\delta}$. The small dashed lines correspond to C_{21}^{yy}. The long-dashed lines and dot–dashed lines are results computed from low-density (ρ) and low-temperature (T) maps, respectively. The middle and right-hand panels correspond to GO simulation with tomographic bins $z = (0.0-0.5)$ and $z = (0.5-1.0)$, respectively. The error bars are computed using three realizations. The angular separation θ_{12} is fixed at $\theta_{12} = 4\theta_s$. The cumulant correlators are two-point statistics and are more susceptible to finite volume effects. The solid line at lower-right-hand corner represents the analytical predictions from lognormal approximation $C_{21} = 2.0$.

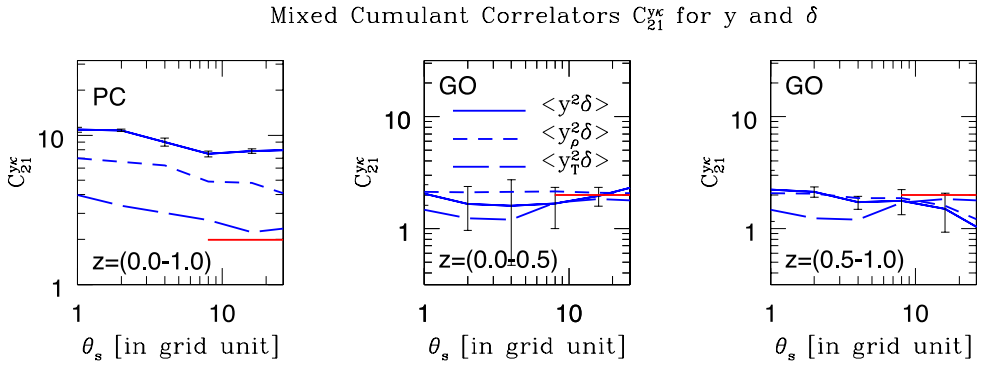

Figure 13. The normalized *mixed* cumulant correlators $C_{21}^{y\kappa} = \langle y^2\kappa\rangle_c/\langle y^2\rangle\langle y_1\kappa_2\rangle$ (defined in equation (29) are plotted as a function of smoothing angular scale θ_0. The left-hand panel shows results from PC simulations. The other two panels correspond to GO simulations at different redshift bins. The solid line corresponds to $C_{21}^{y\kappa}$. The short- and long-dashed lines are results computed from y_ρ and $y_{\rm T}$ maps, respectively. The error bars are computed using three realizations. The angular separation θ_{12} is fixed at $\theta_{12} = 4\theta_0$. The cumulant correlators are two-point statistics and are more susceptible to finite volume effects. The solid line at lower-right-hand corner represents the analytical predictions from lognormal approximation $C_{21} = 2.0$. Notice that we use the projected density δ as a proxy for wl convergence κ.

using both numerical as well as analytical techniques. In their studies, Valageas, Schaeffer & Silk (2002) has used the hierarchical ansatz, to study the *phase-diagrams* of cosmological baryons as function of redshift. The low-temperature 'cool' component of the IGM represented by Lyman α forest typically satisfies the constrain $10^3\ {\rm K} < {\rm T} < 10^4\ {\rm K}$. The exact values of the lower and upper-limit depends somewhat on the redshift. The 'cold' component of the IGM is very well characterized by a well-defined equation of state. The 'warm' component of the IGM on the other hand is shock-heated to a temperature range of $10^4\ {\rm K} < {\rm T} < 10^7\ {\rm K}$ due to the collapse of non-linear structure and cannot be defined by a well-defined equation of state. Though, the 'warm' component does follow a mean temperature–density relation, the scatter around this relation however is more significant than for the 'cool' component. Both the 'cool' and 'warm' components originate outside the collapsed haloes and typically reside in moderate overdensities $1 + \delta < 100$ (Munshi et al. 2013). Finally, the remaining contribution comes from the hot baryonic component of the virialized high-density haloes with temperatures $\rm T > 10^7\ K$.

In Figs 10 and 11, we have plotted the skewness and kurtosis parameters from y_ρ maps made using only the medium overdensity regions $1 + \delta < 100$ which is caused by both 'warm' and the 'cool' component of the IGM, as well as the map made from low-temperature component $y_{\rm T}$. The contribution from high-density virialized haloes are shown too. The skewness and kurtosis corresponding to these individual components are computed individually. The results are presented for the same range of smoothing scales considered before. It is interesting to note that, for GO simulations the skewness computed using component maps y_ρ and $y_{\rm T}$ too are very close to predictions from lognormal model though clearly due to truncation of high ρ tail of the underlying density PDF responsible for construction of y_ρ maps the S_N parameters computed from y_ρ maps are lower than their counterparts constructed from y-maps. In PC simulations, only $y_{\rm T}$ maps agree with lognormal model as explained above the low-temperature regions are less likely to be disturbed by non-gravitational processes. This is in agreement with our previous finding for the entire PDF (Munshi et al. 2013). The cumulant correlators for y-maps, i.e. C_{21} and mixed cumulant correlators $C_{21}^{y\kappa}$ follow a similar patterns.

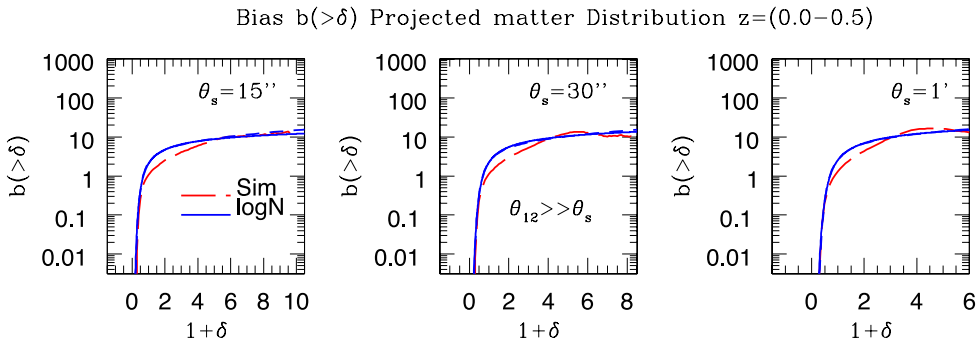

Figure 14. The cumulative bias $b(>\delta)$ defined in equation (42) as a function of δ. Three different smoothing angular scales are considered $\theta_s = 15$ arcsec (left-hand panel), $\theta_s = 30$ arcsec (middle panel) and $\theta_s = 1$ arcmin (right-hand panel), respectively. The results from the simulations are plotted in long-dashed lines. The lognormal results are shown using solid lines. Different levels of dilution of the data were considered.

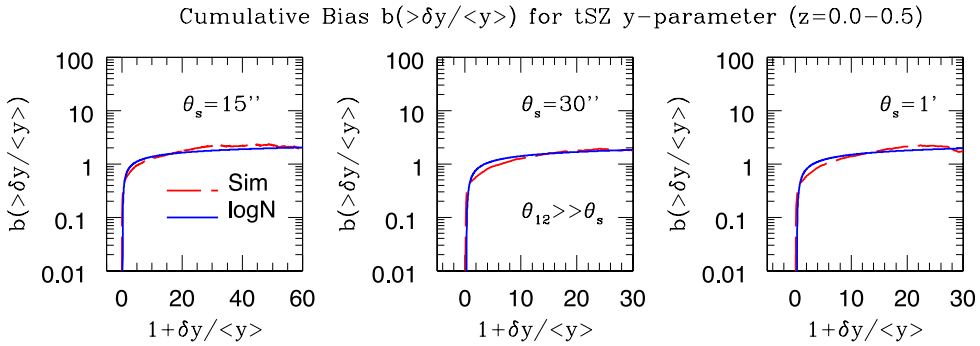

Figure 15. The cumulative bias $b(>\delta y/\langle y\rangle)$ defined in equation (44) is plotted as a function of $\delta y/\langle y\rangle$ for the tSZ y-parameter (GO simulations). Three different smoothing angular scales are considered $\theta_s = 15$ arcsec (left-hand panel), $\theta_s = 30$ arcsec (middle panel) and $\theta_s = 1$ arcmin (right-hand panel), respectively. The simulation results are plotted with long-dashed lines. The lognormal results are shown with solid lines. Different levels of dilution of the data were considered. The variance associated with y-maps is higher compared to the δ maps which explains larger dynamic range for δy.

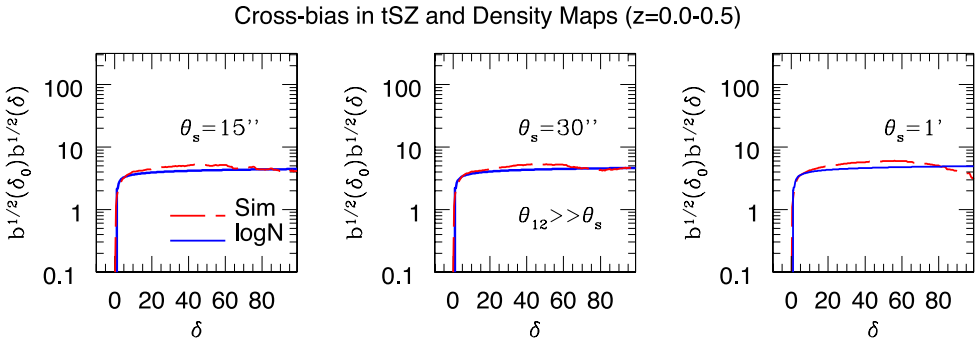

Figure 16. The cross-bias function $\mathcal{B}(>\delta y, >\kappa)$ defined in equation (48) is plotted as a function of $\delta y/\langle y\rangle$. The solid lines show the predictions from the lognormal distribution. The simulation (GO) results are plotted with long-dashed lines. Three different smoothing angular scales are considered $\theta_s = 15$ arcsec (left-hand panel), $\theta_s = 30$ arcsec (middle panel) and $\theta_s = 1$ arcmin (right-hand panel), respectively. The density threshold was fixed at its saturation value and the cross-bias was computed as a function of y-threshold. The redshift bin considered is $z = 0.0-0.5$.

Integrated bias. The computation of integrated bias $b(>y)$ was carried out using techniques discussed in (Munshi et al. 2013) which relies on using a grid to compute the joint PDF using cells separated by a distance. Indeed, the computation of joint PDF is equivalent to a computation of the bias function $b(\delta)$. The differential bias defined this way is noisy so we use the integrated bias $b(>\delta_0)$ beyond a fixed threshold δ_0 as a stable indicator (see equation 44). The results for our computation for bias are presented in Figs 14 and 15. The bias function carries information regarding cumulant correlators to all orders. We find that the analytical predictions for bias to match accurately with the one recovered from numerical simulations. The bias associated with joint PDF of κ and y-parameters are defined in equation (43). These bias functions are related to mixed cumulant correlators defined in equation (48). These results are presented in Figs 16 and 17. The recovery of mixed bias function follows exactly the similar procedure and involve cross-correlating cells from two different maps at a fixed distance. We used the analytical predictions for the individual bias function for κ and y-maps to compute the *mixed* bias function, which is essentially related to the joint PDF of both κ and y-maps to cross specific thresholds (not necessarily the same). The results again match with lognormal

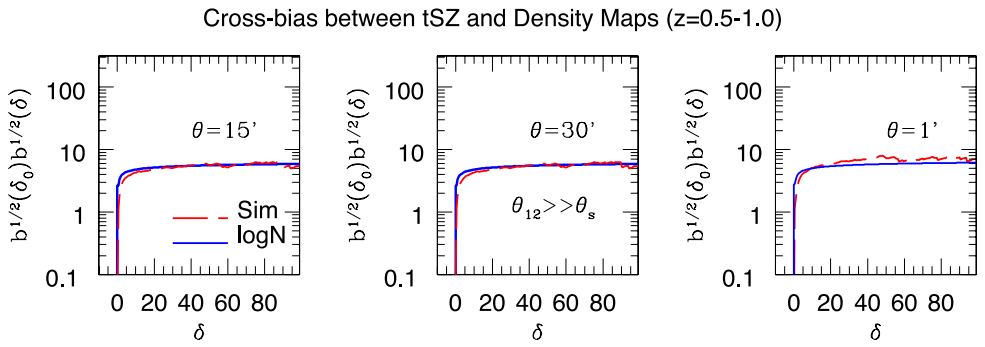

Figure 17. The mixed bias $\mathcal{B}(>\delta y, >\kappa)$ defined in equation (48) is plotted as a function of δy for projected mass distribution. Three different smoothing angular scales are considered $\theta_s = 15$ arcsec (left-hand panel), $\theta_s = 30$ arcsec (middle panel) and $\theta_s = 1$ arcmin (right-hand panel), respectively. The simulation results are plotted with long-dashed lines. The lognormal results are shown with solid lines. The redshift bin considered is $z = 0.5-1.0$. The variance associated with y-maps is higher compared to the δ maps which explains larger dynamic range for δy.

predictions. The mixed bias function contains the knowledge of mixed cumulants of all order. Three different smoothing angular scales were considered for computation of the integrated bias $\theta_s = 15$, 30 arcsec and 1 arcmin. The bias is weakly dependent on θ_s.

To summarize our results about PC simulations, the smoothness of these maps are reflected in their low variance. This is primarily due to the high level of the pre-heating that erases many substructures resulting in maps with less features. We include these simulations in our studies mainly to test the limitations of analytical predictions. The fundamental assumption in our analytical modelling is of gravity-induced structure formation, where baryons are considered as the biased tracers of underlying dark matter clustering. We find significant deviation of the numerical results from theoretical predictions in the presence of high-level pre-heating at small angular scales (see Fig. 8). These deviations are more pronounced at smaller angular scales. The PDFs become Gaussian at scales $\theta_s \sim 10$ arcmin or larger. The deviation at all scales is less pronounced if we remove the collapsed objects and focus on the maps with overdensities $1 + \delta < 100$. However, the PDF of the y-distribution from 'cold' intergalactic gas even in PC simulations is represented very accurately by our analytical results at all angular scales in the presence of pre-heating.

7 CONCLUSION

Extending our previous work (Munshi et al. 2013), we have presented a detailed analysis of the higher order (non-Gaussian) cross-correlation of tSZ maps and the projected or redshift-resolved tomographic maps from wl surveys. We use state-of-the-art *Millennium Gas Simulations* that includes realistic description of gas physics test various ingredients in our theoretical calculations and find reasonable agreement. Below we list our main results.

Tomographic resolution and evolution with redshift. What distinguishes our work in this paper is that it extends earlier studies which have focused mainly on cross-correlation analysis at the level of power spectrum. The tSZ maps as well as the maps derived from wl surveys are intrinsically non-Gaussian even at large angular scales. In this study, we extend the cross-correlation statistics to higher order. We employ the mixed cumulants and their correlators that can extract higher order statistics that are useful in probing the thermal history of the Universe when studied in association with redshift-resolved tomographic wl data. The statistics that we introduce can be used in projection (2D) or with tomographic information, independent of any analytical modelling to analyse simulations or observational data. The estimators presented here are generic and can be used for other cross-correlation studies involving two different *arbitrary* fields, e.g., kSZ and lensing or soft X-ray and tSZ. The mixed cumulants and their correlators defined in this paper are the lower order moments of the corresponding one-point or two-point PDFs and associated bias. These statistics generalize cumulant correlators defined in (Munshi 2000) for wl surveys and in (Munshi et al. 2013) for tSZ data sets, and are useful in cross-correlating these two data sets.

Beyond order-by-order calculation. Beyond the order-by-order calculation, a generic prediction for the cumulants or their correlators to an arbitrary order is required. This was achieved by adopting the generating function formalism inherent in models based on the *hierarchical ansatz*. We have employed a generic version of the hierarchical ansatz as well as the lognormal distribution to model the underlying mass distribution. Both of these have been studied extensively in the literature. These particular models go beyond the lowest order in correlation hierarchy. For the hierarchical model, the generating function is parametrized in terms of a single parameter ω which is fixed using inputs from numerical simulations, and is the only freedom allowed in this model. In the perturbative regime this parameter can be calculated analytically. Indeed, modelling of pressure fluctuations in terms of the underlying density distribution is more complicated, and we rely on a redshift-dependent linear biasing model that is expected to be valid at large angular scales (Munshi et al. 2013). More complicated models can be considered in this framework using additional inputs from numerical simulations; here, we have tried to keep the modelling as simple as possible. The halo model can be used for the construction of the lowest order cumulant correlators. However, in such an approach, results can be derived only in an order-by-order manner, and the relevant PDF and bias needs to be constructed using suitably truncated Edgeworth expansions. As mentioned previously, the PDF and bias for the wl data have been studied by many authors using both the hierarchical ansatz and the lognormal distribution. The studies have shown how accurately these distributions can model numerical simulations. In more recent

studies, the hierarchical ansatz has also been used successfully in tSZ modelling. Based on this success, our primary goal in this study has been to extend such analysis to compute the joint PDF of κ and y along same or different lines of sight. In lieu of individual haloes, we consider here the diffuse tSZ component that correlates with the large-scale wl shear and convergence.

New variables to isolate background cosmology. We have shown that the two different variables, i.e. the scaled tSZ parameter $\eta' = \delta y/\langle y\rangle$ and the reduced convergence field $\eta = \kappa/\kappa_m$ can be introduced to simplify the analysis. Employing the hierarchical ansatz, we next show that the statistics of these reduced variables are the same as the underlying density $\rho/\rho_b = 1 + \delta$ under certain simplifying approximations, which is true both at the level of one- and two-point statistics. Thus, the final results do not depend on the specific details of the input parameters of hierarchical ansatz, and may point to a wider applicability. The joint distribution involving κ and y can thereby be written in terms of $p(\eta)$ and $b(\eta)$ with suitable scaling involving $\langle y\rangle$ and κ_m. The detailed construction of the PDF is insensitive to modelling of the bias of the hot ionized gas with respect to the dark matter distribution. The detailed prescription for such bias enters the results through the computation of the variance and the two-point correlation function $\xi_{y\kappa}(\theta_{12})$.

Linear biasing. Specific models of hierarchical clustering have previously been employed to understand the tSZ effect (see e.g. Valageas & Silk 1999; Valageas & Schaeffer 2000; Valageas et al. 2001). However, they have been employed to model contribution from the collapsed objects (e.g. massive clusters). In this case, the tail of the scaling functions $h(x)$ and $b(x)$ are relevant as they describe the collapsed objects. In addition to these inputs, the modelling would involve a prescription for the hydrodynamic equilibrium of the gas residing in the halo. These can be used to model the PDF of $y(\hat{\Omega})$, i.e. $p_y(y)$ or $b_y(y)$. The modelling considered here is complementary to our analysis. We have considered redshift dependent linear biasing to model the shock heated gas in the IGM that produces a diffuse tSZ effect. The linear biasing model has been studied at the level of the power spectrum by many authors (Goldberg & Spergel 1999a,b; Peiris & Spergel 2000). The hierarchical model that we use here was also previously used by Zhang et al. (2002), Cooray et al. (2000) and (Cooray 2000, 2001) for modelling lower order moments where the linear biasing model for the pressure fluctuations was assumed. Our results are natural generalizations of such studies to higher order multispectra to study the non-Gaussian aspects of tSZ maps.

Effect of pre-heating. It is known that the IGM is most likely have been preheated by non-gravitational sources. The feedback from SN or AGN can play an important role. The analytical modelling of such non-gravitational processes is rather difficult. Numerical simulations (da Silva et al. 2000, 2004; Springel et al. 2001; White et al. 2002; Lin et al. 2004) have shown that the amplitude of the tSZ signal is sensitive to the non-gravitational processes, e.g., the amount of radiative cooling and energy feedback. It is also not straightforward to disentangle contributions from competing processes. The inputs from simulations are vital for any progress. Our analytical results should be treated as a first step in this direction. We have focused mainly on large angular scales where we expect the gravitational process to dominate and such effects to be minimal. We find that the simulations without pre-heating (GO) can be understood using analytical arguments. Thus, the effect of additional baryonic physics can be separated using the formalism developed here. To understand the effect of baryonic physics we have developed the simulated maps in different components and studied them individually. We also find that even in the presence of preheating the *cool* component $y_{\rm T}$ of the IGM remains relatively undisturbed and can be modelled using a linear biasing scheme.

From tomography to 3D. Our results are derived using cross-correlating tomographic slices from wl surveys and the projected tSZ surveys. With suitable modification, they can be equally applicable to cross-correlation studies using tomographic maps from the same or different wl surveys. The results can also be used to study the wl of CMB by cross-correlating it against tomographic wl maps or tSZ maps. To move beyond a tomographic or projected survey, a complete 3D analysis can be invoked. An analysis similar to what has been presented here for the kSZ effect (e.g. Shao et al. 2011a) can provide valuable information about the reionization history of the Universe (Munshi et al. 2013; in preparation).

ACKNOWLEDGEMENTS

DM and PC acknowledge support from STFC standard grant ST/G002231/1 at School of Physics and Astronomy at Cardiff University where this work was completed. SJ and JS acknowledge support from the US Department of Education through GAANN at UCI. We thank Alan Heavens for useful discussions. It is also a pleasure for us to thank Martin Kilbinger for related collaboration. We thank Francis Bernardeau for supplying a copy of his code which was used to compute the bias and the PDF from hierarchical ansatz.

REFERENCES

Albrecht A. et al., 2011, arXiv:0901.0721
Balian R., Schaeffer R., 1989, A&A, 220, 1
Battaglia N., Bond J. R., Pfrommer C., Sievers J. L., 2013, ApJ, 777, 123 [Erratum; 780, 189 (2014)]
Bernardeau F., 1992, ApJ, 392, 1
Bernardeau F., 1994, A&A, 291, 697
Bernardeau F., Kofman L., 1995, ApJ, 443, 479
Bernardeau F., Schaeffer R., 1992, A&A, 255, 1
Bernardeau F., Schaeffer R., 1999, A&A, 349, 697
Bernardeau F., Colombi S., Gaztanaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1
Bi H. G., Davidson A. F., 1997, ApJ, 479, 523
Birkinshaw M., 1999, Phys. Rep., 310, 98
Borgani S., Viel M., 2009, MNRAS, 392, L26
Bouchet F. R., Gispert R., 1999, New Astron., 4, 443

Bouchet F., Strauss M. A., Davis M., Fisher K. B., Yahil A., Huchra J. P., 1993, ApJ, 417, 36
Calabrese E. et al., 2013, Phys. Rev. D, 87, 103012
Cao L., Liu J., Fang L.-Z., 2007, ApJ, 661, 641
Castro P. G., 2003, Phys. Rev. D, 67, 123001
Cen R., Ostriker J. P., 1999, ApJ, 514, 1
Coles P., Jones B., 1991, MNRAS, 248, 1
Colombi S., 1994, ApJ, 435, L536
Colombi S., Bouchet F. R., Schaeffer R., 1996a, ApJS, 96, 401
Colombi S., Bouchet F. R., Hernquist L., 1996b, ApJ, 465, 14
Colombi S., Bernardeau F., Bouchet F. R., Hernquist L., 1997, MNRAS, 287, 241
Cooray A., 2000, Phys. Rev. D, 62, 103506
Cooray A., 2001, Phys. Rev. D, 64, 063514
Cooray A., Hu W., 2000, ApJ, 535, L9
Cooray A., Hu W., 2002, ApJ, 574, 19
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
Cooray A., Tegmark M., Hu W., 2000, ApJ, 540, 1
da Silva A. C., Barbosa A. C., Liddle A. R., Thomas P. A., 2000, MNRAS, 317, 37
da Silva A. C., Kay S. T., Liddle A. R., Thomas P. A., 2004, MNRAS, 348, 1401
Das S. et al., 2011, Phys. Rev. Lett., 107, 021301
Dunkley J. et al., 2011, ApJ, 739, 52
Dunkley J. et al., 2013, J. Cosmol. Astropart. Phys., 07, 025
Edward A. E., Henry J. P., 1991, ApJ, 383, 95
Fowler J. W. et al., 2010, ApJ, 722, 1148
Fry J. N., 1984, ApJ, 279, 499
Goldberg D. M., Spergel D. N., 1999a, Phys. Rev. D, 59, 103001
Goldberg D. M., Spergel D. N., 1999b, Phys. Rev. D, 59, 103002
Hallman E. J., O'Shea B. W., Smith B. D., Norman M. L., Harkness R., Wagner R., 2007, ApJ, 671, 27
Hallman E. J., O'Shea B. W., Smith B. D., Burns J. O., Norman M. L., 2009, ApJ, 698, 1759
Hamilton A. J. S., 1985, ApJ, 292, L35
Hand N. et al., 2011, ApJ, 736, 39
Hansen F. K., Branchini E., Mazzotta P., Cabella P., Dolag K., 2005, MNRAS, 361, 753
Hanson D. et al., 2013, Phys. Rev. Lett., 111, 141301
Hartley W. G., Gazzola L., Pearce F. R., Kay S. T., Thomas P. A., 2008, MNRAS, 386, 2015
High F. W. et al., 2012, ApJ, 758, 68
Hill J. C., Spergel D. N., 2014, J. Cosmol. Astropart. Phys., 02, 030
Hirata C. M., Seljak U., 2003, Phys. Rev. D, 68, 083002
Hirata C. M., Seljak U., 2004, Phys. Rev. D, 70, 063526
Holder G. et al., 2013, ApJ, 771, 16
Hou Z. et al., 2014, ApJ, 782, 74
Hui L., 1999, ApJ, 519, L9
Huterer D., Takada M., Bernstein G., Jain B., 2006, MNRAS, 366, 101
Jones M. E. et al., 2005, MNRAS, 357, 518
Joudaki S., Kaplinghat M., 2011, Phys. Rev. D, 86, 023526
Joudaki S., Smidt J., Amblard A., Cooray A., 2010, J. Cosmol. Astropart. Phys., 1008, 027
Kaiser N., 1991, ApJ, 383, 104
Kaiser N., 1998, ApJ, 498, 26
Kay S. T., Peel M. W., Short C. J., Thomas P. A., Young O. E., Battye R. A., Liddle A. R., Pearce F. R., 2012, MNRAS, 422, 1999
Kayo I., Taruya A., Suto Y., 2001, ApJ, 561, 22
Kofman L., Bertschinger E., Gelb J. M., Nusser A., Dekel A., 1994, ApJ, 420, 44
Komatsu E., Seljak U., 2002, MNRAS, 336, 1256
LaRoque S. J., Bonamente M., Carlstorm J. E., Joy M. K., Nagai D., Reese E. D., Dawson K. S., 2006, ApJ, 652, 917
Leach S. M. et al., 2008, A&A, 491, 597
Limber D. N., 1954, ApJ, 119, 665
Lin K.-Y., Woo T.-P., Tseng Y.-H., Lin L., Chiueh T., 2004, ApJ, 608, L1
Lueker M. et al., 2010, ApJ, 719, 1045
Ma Z., Hu W., Huterer D., 2005, ApJ, 636, 21
McCarthy I. G., Le B., Amandine M. C., Schaye J., Holder G. P., 2013, MNRAS, 440, 3645
McDonald P., Trac H., Contaldi C., 2006, MNRAS, 366, 547
Matarrese S., Lucchin F., Moscardini L., Saez D., 1992, MNRAS, 259, 437
Munshi D., 2000, MNRAS, 318, 145
Munshi D., Coles P., 2000, MNRAS, 313, 148
Munshi D., Coles P., 2002, MNRAS, 329, 797
Munshi D., Heavens A., 2010, MNRAS, 401, 2406
Munshi D., Jain B., 2000, MNRAS, 318, 109
Munshi D., Jain B., 2001, MNRAS, 322, 107
Munshi D., Sahni V., Starobinsky A., 1994, ApJ, 436, 517
Munshi D., Coles P., Melott A. L., 1999a, MNRAS, 307, 387
Munshi D., Coles P., Melott A. L., 1999b, MNRAS, 310, 892

Munshi D., Melott A. L., Coles P., 1999c, MNRAS, 311, 149
Munshi D., Valageas P., Barber A. J., 2004, MNRAS, 350, 77
Munshi D., Valageas P., van Waerbeke L., Heavens A., 2008, Phys. Rep., 462, 67
Munshi D., Heavens A., Cooray A., Valageas P., 2011a, MNRAS, 414, 3173
Munshi D., Kitching T., Heavens A., Coles P., 2011b, MNRAS, 416, 629
Munshi D., Coles P., Viel M., 2012, MNRAS, 427, 2359
Munshi D., Joudaki S., Smidt J., Coles P., Kay S. T., 2013, MNRAS, 429, 1564
Navarro J., Frenk C., White S. D. M., 1996, ApJ, 462, 563
Peebles P. J. E., 1971, Princeton Series in Physics. Princeton Univ. Press, Princeton, NJ
Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ
Peiris H., Spergel D., 2000, ApJ, 540, 605
Persi F., Spergel D., Cen R., Ostriker J., 1995, ApJ, 442, 1
Planck Collaboration, 2013a, arXiv:1303.5081
Planck Collaboration, 2013b, arXiv:1303.5089
Press W., Schechter P., 1974, ApJ, 187, 425
Rauch M., 1998, ARA&A, 1, 267
Reese E. D., Carlstorm J. E., Joy M., Mohr J. J., Grego L., Holzapfel W. L., 2002, ApJ, 581, 53
Refregier A., Komatsu E., Spergel D. N., Pen U.-L., 2000, Phys. Rev. D, 61, 123001
Rephaeli Y., 1995, ARA&A, 33, 541
Roncarelli M., Moscardini L., Borgano S., Dolag K., 2007, MNRAS, 378, 1259
Rosati P., Borgani S., Norman C., 2002, ARA&A, 40, 539
Saro A. et al., 2013, MNRAS, 440, 2610
Schulz A. E., White M., 2003, ApJ, 586, 723
Scoccimarro R., Colombi S., Fry J. N., Frieman J. A., Hivon E., Melott A., 1998, ApJ, 496, 586
Sehgal et al., 2011, ApJ, 732, 44
Seljak U., 2000, MNRAS, 318, 203
Seljak U., Burwell J., Pen U.-L., 2001, Phys. Rev. D, 63, 063001
Shang C., Crotts A., Haiman Z., 2007, ApJ, 671, 136
Shao J., Zhang P., Lin W., Jing Y., Pan J., 2011a, MNRAS, 413, 628
Shao J., Zhang P., Lin W., Jing Y., 2011b, ApJ, 730, 127
Shaw L. D., Nagai D., Bhattacharya S., Lau E. T., 2010, ApJ, 725, 1452
Sherwin et al., 2011, Phys. Rev. Lett., 107, 02132
Short C. J., Thomas P. A., Young O. E., Pearce F. R., Jenkins A., Muanwong O., 2010, MNRAS, 408, 2213
Smidt J., Joudaki S., Serra P., Amblard A., Cooray A., 2010, Phys. Rev. D, 81, 123528
Smidt J., Cooray A., Amblard A., Joudaki S., Munshi D., Santos M. G., Serra P., 2011, ApJ, 728, 1, L1
Smith R. E. et al., 2003, MNRAS, 341, 1311
Springel V., 2005, MNRAS, 364, 1105
Springel V., White M., Hernquist L., 2001, ApJ, 549, 681
Starek R., Rudd D., Evrard A. E., 2009, MNRAS, 394, L1
Story K. T. et al., 2013, ApJ, 779, 86
Sunyaev R. A., Zel'dovich Ya. B., 1972, Comments Astrophys. Space Phys., 4, 173
Sunyaev R. A., Zel'dovich Ya. B., 1980, ARA&A, 18, 537
Szapudi I., Szalay A. S., 1993, ApJ, 408, 43
Takada M., Jain B., 2002, MNRAS, 337, 875
Takada M., Jain B., 2003, MNRAS, 340, 580
Taruya A., Takada M., Hamana T., Futamese T., 2002, ApJ, 571, 638
Theuns T., Mo H. J., Schaye J., 2001, MNRAS, 321, 450
Valageas P., 2000a, A&A, 354, 767
Valageas P., 2000b, A&A, 356, 771
Valageas P., 2002, A&A, 382, 412
Valageas P., Schaeffer R., 2000, A&A, 359, 821
Valageas P., Munshi D., 2004, MNRAS, 354, 1146
Valageas P., Silk J., 1999, A&A, 350, 725
Valageas P., Silk J., Schaeffer R., 2001, A&A, 366, 363
Valageas P., Schaeffer R., Silk J., 2002, A&A, 388, 741
van Waerbeke L., Hinshaw G., Murray N., 2014, Phys. Rev. D, 89, 023508
Vieira J. D. et al., 2013, Nature, 495, 344
Voit G. M., 2005, Rev. Mod. Phys., 77, 207
White S. D. M., 1979, MNRAS, 186, 145
White M., 2003, ApJ, 597, 650
White M., 2004, Astropart. Phys., 22, 211
White M., Hernquist V., Springel V., 2002, ApJ, 579, 16
Wilson et al., 2012, Phys. Rev. D, 86, 122005
Young O. E., Thomas P. A., Short C. J., Pearce F., 2011, MNRAS, 413, 691
Zhang P., Pen U-Li., 2001, ApJ, 549, 18
Zhang P., Sheth R. K., 2007, ApJ, 579, 16
Zhang P., Pen U-Li., Wang B., 2002, ApJ, 577, 555
Zhang P., Pen U-L., Trac H., 2004, MNRAS, 355, 451

APPENDIX A: GRAVITATIONAL CLUSTERING AND THE HIERARCHICAL ANSATZ

The lower order cumulants and cumulant correlators as well as the one- and two-point PDFs are commonly used statistics in cosmology to quantify clustering (Bernardeau et al. 2002). Use of generating function techniques to go beyond order by order calculations is an important milestone in this direction (Peebles 1971; Balian & Schaeffer 1989; Bernardeau 1992, 1994; Bernardeau & Schaeffer 1992). This approach is complimentary to the halo model based approach that relies on Press–Sechter mass formalism (Press & Schechter 1974). We provide a very brief review of this formalism in this appendix.

A1 The generating function

In scaling analysis of the PDF, the void probability distribution function (VPF), denoted as $P_v(0)$, plays most fundamental role, which can be related to the generating function of the cumulants or S_N parameters, $\phi(z)$ (White 1979; Balian & Schaeffer 1989):

$$P_v(0) = \exp(-\bar{N}\sigma(N_c)) = \exp\left(-\phi(N_c)/\bar{\xi}_2^{\delta}\right), \tag{A1}$$

where $P_v(0)$ is the probability of having no 'particles' in a cell of volume v, $\bar{N}$ is the average occupancy of these 'particles' and $N_c = \bar{N}\bar{\xi}_2^{\delta}$ and $\sigma(z) = -\phi(z)/z$. Here $\phi(z) = \sum_{p=1}^{\infty} S_p/p!\, z^p$ is a generating function for S_N parameters. The generating function $\mathcal{G}(\tau)$ for the vertex amplitudes ν_p is related to $\phi(z)$. A more specific model for $\mathcal{G}(\tau)$ is generally used to make more specific predictions (Balian & Schaeffer 1989):

$$\mathcal{G}_{\delta}(\tau) = (1 + \tau/\kappa_a)^{-\kappa_a}. \tag{A2}$$

The parameter κ_a is related to the parameter ω_a to be introduced later. The two generating functions are related by the following expression (Balian & Schaeffer 1989; Bernardeau & Schaeffer 1992, 1999):

$$\phi(z) = z\mathcal{G}_{\delta}(\tau) - \frac{1}{2}z\tau\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{G}_{\delta}(\tau); \quad \tau = -z\frac{\mathrm{d}}{\mathrm{d}\tau}\mathcal{G}_{\delta}(\tau). \tag{A3}$$

However, a more detailed analysis is needed to include the effect of correlation between two or more correlated volume elements which will provide information about bias, cumulants and cumulant correlators (Bernardeau & Schaeffer 1992; Munshi et al. 1999a,b,c) Notice that $\tau(z)$ (a bias also denoted by $\beta(z)$ in the literature) plays the role of generating function for factorized cumulant correlators $C_{p1}^{\eta\eta'}$ ($C_{pq} = C_{p1}^{\eta\eta}C_{q1}^{\eta'\eta'}$): $\tau(z) = \sum_{p=1}^{\infty} C_{p1}^{\eta\eta}/p!\, z^p$.

A2 Generating functions and the construction of the PDF and bias

The hierarchical form of higher order correlation functions appear in two completely different regimes in gravitational clustering. The generating function approach is a convenient technique to sum up to arbitrary information that leads to construction of entire PDFs.

A2.1 The highly non-linear regime

The PDF $p(\delta)$ and bias $b(\delta)$ can be related to their generating functions VPF $\phi(z)$ and $\tau(z)$, respectively, by following equations (Balian & Schaeffer 1989; Bernardeau & Schaeffer 1992, 1999)

$$p(\delta) = \int_{-i\infty}^{i\infty}\frac{\mathrm{d}z}{2\pi i}\exp\left[\frac{(1+\delta)z - \phi(z)}{\bar{\xi}_2}\right]; \quad b(\delta)p(\delta) = \int_{-i\infty}^{i\infty}\frac{\mathrm{d}z}{2\pi i}\tau(z)\exp\left[\frac{(1+\delta)z - \phi(z)}{\bar{\xi}_2}\right]. \tag{A4}$$

It is clear that the function $\phi(z)$ completely determines the behaviour of the PDF $p(\delta)$ for all values of δ. However different asymptotic expressions of $\phi(z)$ govern the behaviour of $p(\delta)$ for different intervals of δ. For large y we can express $\phi(z)$ as $\phi(z) = az^{1-\omega}$. Here we have introduced a new parameter ω for the description of VPF. Typically initial power spectrum with spectral index $n = -2$ (which should model cold dark matter (CDM) like spectra we considered in our simulations at small length-scales) produces a value of 0.3 which we will be using in our analysis (Colombi, Bouchet & Schaeffer 1996a; Colombi, Bouchet & Hernquist 1996b; Colombi et al. 1997). The VPF $\phi(z)$ and its two-point analogue $\tau(z)$ both exhibit singularity for small but negative value of z_*,

$$\phi(z) = \phi_* - a_*\Gamma(\omega_*)(z - z_*)^{-\omega_*}; \quad \tau(z) = \tau_* - b_*(z - z_*)^{-\omega_*-1}. \tag{A5}$$

For the factorizable model of hierarchical clustering the parameter ω_* takes the value $-3/2$ and a_* and b_* can be expressed in terms of the nature of the generating function $\mathcal{G}_{\delta}(\tau)$ and its derivatives near the singularity τ_* (Bernardeau & Schaeffer 1992):

$$a_* = \frac{1}{\Gamma(-1/2)}\mathcal{G}'_{\delta}(\tau_*)\mathcal{G}''_{\delta}(\tau_*)\left[\frac{2\mathcal{G}'_{\delta}(\tau_*)\mathcal{G}''_{\delta}(\tau_*)}{\mathcal{G}'''_{\delta}(\tau_*)}\right]^{3/2}; \quad b_* = \left[\frac{2\mathcal{G}'_{\delta}(\tau_s)\mathcal{G}''_{\delta}(\tau_*)}{\mathcal{G}'''_{\delta}(\tau_*)}\right]^{1/2}. \tag{A6}$$

As mentioned before the parameter k_a which we have introduced in the definition of $\mathcal{G}_{\delta}(\tau)$ can be related to the parameters a and ω appearing in the asymptotic expressions of $\phi(z)$ (Balian & Schaeffer 1989; Bernardeau & Schaeffer 1992)

$$\omega = k_a/(k_a + 2); \quad a = \frac{k_a + 2}{2}k_a^{k_a/k_a+2}. \tag{A7}$$

Similarly, the parameter z_s which describe as the behaviour of the function $\phi(z)$ near its singularity can be related to the behaviour of $\mathcal{G}(\tau)$ near τ_s which is the solution of the equation (Balian & Schaeffer 1989; Bernardeau & Schaeffer 1992)

$$\tau_* = \mathcal{G}'_\delta(\tau_*)/\mathcal{G}''_\delta(\tau_*), \tag{A8}$$

finally we can relate k_a to y_* by following expression (see equation A7):

$$z_* = -\frac{\tau_*}{\mathcal{G}'(\tau_*)}; \quad -\frac{1}{z_*} = x_\star = \frac{1}{k_a}\frac{(k_a+2)^{k_a+2}}{(k_a+1)^{k_a+1}}. \tag{A9}$$

The newly introduced variable $x_\star$ will be useful to define the large-δ tail of the PDF $p(\delta)$ and the bias $b(\delta)$. Different asymptotes in $\phi(z)$ are linked with behaviour of $p(\delta)$ for various regimes of δ. For very large values of variance, i.e. ξ_2 it is possible to define a scaling function $p(\delta) = h(x)/\bar{\xi}_2^2$ which will encode the scaling behaviour of PDF, where x plays the role of the scaling variable and is defined as $x = (1+\delta)/\bar{\xi}_2$. We list different ranges of δ and specify the behaviour of $p(\delta)$ and $b(\delta)$ in these regimes (Balian & Schaeffer 1989)

$$(\bar{\xi}_2^\delta)^{-\omega/(1-\omega)} \ll 1+\delta \ll \bar{\xi}_2^\delta; \quad p(\delta) = \frac{a}{(\bar{\xi}_2^\delta)^2}\frac{1-\omega}{\Gamma(\omega)}\left(\frac{1+\delta}{\bar{\xi}_2^\delta}\right)^{\omega-2}; \quad b(\delta) = \left(\frac{\omega}{2a}\right)^{1/2}\frac{\Gamma(\omega)}{\Gamma\left[\frac{1}{2}(1+\omega)\right]}\left(\frac{1+\delta}{\bar{\xi}_2^\delta}\right)^{(1-\omega)/2} \tag{A10}$$

$$1+\delta \gg \bar{\xi}_2^\delta; \quad p(\delta) = \frac{a_s}{(\bar{\xi}_2^\delta)^2}\left(\frac{1+\delta}{\bar{\xi}_2^\delta}\right)\exp\left(-\frac{1+\delta}{x_\star\bar{\xi}_2^\delta}\right); \quad b(\delta) = -\frac{1}{\mathcal{G}'(\tau_s)}\frac{(1+\delta)}{\bar{\xi}_2^\delta}. \tag{A11}$$

The integral constraints satisfied by scaling function are $S_1 = \int_0^\infty x h(x)\,\mathrm{d}x = 1$ and $S_2 = \int_0^\infty x^2\,h(x)\,\mathrm{d}x = 1$. These take care of normalization of the function $p(\delta)$. Similarly the normalization constraint over $b(\delta)$ can be expressed as $C_{11} = \int_0^\infty x b(x)h(x)\,\mathrm{d}x = 1$, which translates into $\int_{-1}^\infty \mathrm{d}\delta b(\delta)p(\delta) = 0$ and $\int_{-1}^\infty \mathrm{d}\delta\delta b(\delta)p(\delta) = 1$. Several numerical studies have been conducted to study the behaviour of $h(x)$ and $b(x)$, for different initial conditions see, e.g. Colombi et al. (1996a,b, 1997). For very small values of δ the behaviour of $p(\delta)$ is determined by the asymptotic behaviour of $\phi(z)$ for large values of y, and it is possible to define another scaling function $g(z)$ which is completely determined by ω, the scaling parameter can be expressed as $z' = (1+\delta)a^{-1/(1-\omega)}(\bar{\xi}_2^\delta)^{\omega/(1-\omega)}$. However, numerically it is much simpler to determine ω from the study of $\sigma(y)$ compared to the study of $g(z)$:

$$1+\delta \ll \bar{\xi}_2^\delta; \quad p(\delta) = a^{-1/(1-\omega)}(\bar{\xi}_2^\delta)^{\omega/(1-\omega)}\sqrt{\frac{(1-\omega)^{1/\omega}}{2\pi\omega z'^{(1+\omega)/\omega}}}\exp\left[-\omega\left(\frac{z'}{1-\omega}\right)^{-(1-\omega)/\omega}\right]; \quad b(\delta) = -\left(\frac{2\omega}{\bar{\xi}_2^\delta}\right)^{1/2}\left(\frac{1-\omega}{z'}\right)^{(1-\omega)/2\omega} \tag{A12}$$

To summarize, we can say that the entire behaviour of the PDF $p(\delta)$ is encoded in two different scaling functions, $h(x)$ and $g(z')$ and one can also study the scaling properties of $b(\delta)$ in terms of the scaling variables x and z in a very similar way.

A2.2 The quasi-linear regime

The first departure from Gaussianity can be studied analytically using perturbative techniques (Bernardeau 1992, 1994) as well as using steepest descent methods (Valageas 2000a,b, 2002). These techniques are also invaluable in modelling the statistics of extreme underdensities. The PDF and bias now can be expressed in terms of $G_\delta(\tau)$ (Bernardeau 1992, 1994):

$$p(\delta)\,\mathrm{d}\delta = \frac{1}{-\mathcal{G}'_\delta(\tau)}\left[\frac{1-\tau\mathcal{G}''_\delta(\tau)/\mathcal{G}'_\delta(\tau)}{2\pi\bar{\xi}_2^\delta}\right]^{1/2}\exp\left(-\frac{\tau^2}{2\bar{\xi}_2^\delta}\right)\mathrm{d}\tau; \quad b(\delta) = -\left(\frac{k_a}{\bar{\xi}_2^\delta}\right)\left[(1+\mathcal{G}_\delta(\tau))^{1/k_a}-1\right], \tag{A13}$$

$$\mathcal{G}_\delta(\tau) = \mathcal{G}(\tau) - 1 = \delta. \tag{A14}$$

The above expression is valid for $\delta < \delta_c$ where the δ_c is the value of δ which cancels the numerator of the prefactor of the exponential function appearing in the above expression. For $\delta > \delta_c$ the PDF develops an exponential tail which is related to the presence of singularity in $\phi(z)$ in a very similar way as in the case of its highly non-linear counterpart (Bernardeau 1992, 1994):

$$p(\delta)\mathrm{d}\delta = 3a_s\sqrt{\bar{\xi}_2^\delta}4\sqrt{\pi}\delta^{-5/2}\exp\left[-|z_s|\frac{\delta}{\bar{\xi}_2^\delta} + \frac{|\phi_s|}{\bar{\xi}_2^\delta}\right]\mathrm{d}\delta; \quad b(\delta) = -\frac{1}{\mathcal{G}'(\tau_s)}\frac{(1+\delta)}{\bar{\xi}_2^\delta}. \tag{A15}$$

The quasi-linear regime remains relatively better understood than the highly non-linear regime, as the connection to dynamics can be made using analytical schemes. Attempts have been made to extend the perturbative results to the highly non-linear regime (Colombi et al. 1997). The lower order cumulant correlators are related to the moments of the bias function. For the case of top-hat smoothing and spectral index n_e the lowest order of the cumulant correlator is given by (Bernardeau 1994) $C_{21}^{\eta\eta'} = 68/21 - (n_\mathrm{e}+3)/3$.

APPENDIX B: THE LOGNORMAL DISTRIBUTION

In addition to the hierarchical ansatz *lognormal* distribution too is a popular analytical model commonly used in cosmology (Hamilton 1985; Coles & Jones 1991; Bouchet et al. 1993; Colombi 1994; Kofman et al. 1994; Bernardeau & Kofman 1995). It appears in the quasi-linear regime (Munshi, Sahni & Starobinsky 1994) as a natural outcome of gravitational dynamics, under certain simplifying assumptions (Matarrese

et al. 1992). Lognormal distribution is routinely used to model the statistics of wl observables (Munshi 2000; Taruya et al. 2002), clustering of Lyman α absorption systems (e.g. Bi & Davidson 1997) and more recently by Munshi et al. (2013) for modelling of the tSZ statistics.

To understand the construction of lognormal distribution, we introduce a Gaussian PDF in variable s: $p(s) = (2\pi\bar{\Xi})^{-1/2}\exp[-(s-\mu)^2/2\bar{\Xi}]$. With a change of variable $s = \ln(r)$ we can write down the PDF of y which is a lognormal distribution $p(r) = (2\pi\bar{\Xi})^{-1/2}\exp[-(\ln(r)-\mu)^2/2\bar{\Xi}]/r$. The extra factor of $(1/r)$ stems from the fact: $\mathrm{d}r/r = \mathrm{d}s$. Note that s is positive definite and is often associated with $\rho/\langle\rho\rangle = 1+\delta$ which means $\langle r\rangle = 1$. The moment generating function for the lognormal in terms of the mean μ and the variance $\bar{\Xi}$ has the following form: $\langle r^n\rangle = \exp(n\mu + n^2\bar{\Xi}/2)$. This however leads to the fact that if the underlying distribution of r or the density is Gaussian we will have to impose the condition: $\mu = -\bar{\Xi}/2$. In our notation above Ξ is the distribution of the underlying Gaussian field. The variance of r defined as $\langle r^2\rangle - \langle r\rangle^2 = \exp(\Xi) - 1 = \xi_2^\delta$. So we can write $\bar{\Xi} = \ln(1+\bar{\xi}_2^\delta)$. This is the result that was used above. The generalization to two-point or bivariate PDF can be done following the same arguments and can be found in Kayo, Taruya & Suto (2001).

In the limit of large separation $\Xi_{12} \to 0$ we can write down the two-point PDF as

$$p_{\ln}(\delta_1,\delta_2) = p_{\ln}(\delta_1)p_{\ln}(\delta_2)[1 + b_{\ln}(\delta_1)\xi_{12}^\delta b_{\ln}(\delta_2)]; \qquad b_{\ln}(\delta_i) = \Lambda_i/\bar{\Xi}. \tag{B1}$$

However, it is simpler to estimate the cumulative or integrated bias associated with objects beyond a certain density threshold δ_0. This is defined as $b(\delta > \delta_0) = \int_{\delta_0}^\infty p(\delta)b(\delta)\,\mathrm{d}\delta / \int_{\delta_0}^\infty p(\delta)\,\mathrm{d}\delta$. In the low variance limit $\bar{\xi}_2^\delta \to 0$ the usual Gaussian result is restored $b(\delta) = \delta/\bar{\xi}_2^\delta$. The parameters Λ, Λ_i, Ξ_{12}, $\bar{\Xi}$ that we have introduced above can be expressed in terms of the two-point (non-linear) correlation function $\xi_{12}^\delta = \langle\delta_1\delta_2\rangle$ and $\bar{\xi}_2^\delta$ is the volume average of the non-linear two-point correlation function ξ_{12} of the smoothed density field.

The validity and limitations of various aspects of the one-point and two-point PDFs have been studied extensively in the literature against N-body simulations Bernardeau (1992, 1994). It is known that in the weakly non-linear regime the lognormal distribution is equivalent to the hierarchical model with a generating function $\mathcal{G}(\tau) = \exp(-\tau)$. This leads to $S_p = p^{p-2}$. The loop level corrections can be computed exactly for the lognormal distribution, which gives $S_3 = 3 + \bar{\xi}_2^\delta$ and $S_4 = 16 + 15\bar{\xi}_2^\delta + 6(\bar{\xi}_2^\delta)^2 + (\bar{\xi}_2^\delta)^3$. It has been shown (see e.g. Kayo et al. 2001) that the lognormal distribution very accurately describes the cosmological distribution functions even in the non-linear regime $\bar{\xi}_2^\delta \leq 4$ for relatively high values of density contrast $\delta < 100$.

This paper has been typeset from a TEX/LATEX file prepared by the author.