Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **441,** 1457–1467 (2014) doi:10.1093/mnras/stu680

Measuring non-local Lagrangian peak bias

Matteo Biagetti,[1★] Kwan Chuen Chan,[1] Vincent Desjacques[1] and Aseem Paranjape[2]

[1] *Département de Physique Théorique and Center for Astroparticle Physics (CAP), Université de Genève, 24 quai Ernest Ansermet, CH-1211 Genève, Switzerland*
[2] *Institute for Astronomy, Department of Physics, ETH Zürich, Wolfgang-Pauli-Strasse 27, CH-8093 Zürich, Switzerland*

Accepted 2014 April 3. Received 2014 April 1; in original form 2013 October 15

ABSTRACT

We investigate non-local Lagrangian bias contributions involving gradients of the linear density field, for which we have predictions from the excursion set peak formalism. We begin by writing down a bias expansion which includes all the bias terms, including the non-local ones. Having checked that the model furnishes a reasonable fit to the halo mass function, we develop a one-point cross-correlation technique to measure bias factors associated with χ^2-distributed quantities. We validate the method with numerical realizations of peaks of Gaussian random fields before we apply it to *N*-body simulations. We focus on the lowest (quadratic) order non-local contributions $-2\chi_{10}(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)$ and $\chi_{01}[3(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)^2 - k_1^2 k_2^2]$, where $\boldsymbol{k}_1$, $\boldsymbol{k}_2$ are wave modes. We can reproduce our measurement of χ_{10} if we allow for an offset between the Lagrangian halo centre-of-mass and the peak position. The sign and magnitude of χ_{10} is consistent with Lagrangian haloes sitting near linear density maxima. The resulting contribution to the halo bias can safely be ignored for $M = 10^{13}\,\mathrm{M}_\odot\,h^{-1}$, but could become relevant at larger halo masses. For the second non-local bias χ_{01} however, we measure a much larger magnitude than predicted by our model. We speculate that some of this discrepancy might originate from non-local Lagrangian contributions induced by non-spherical collapse.

Key words: cosmology: theory – dark matter – large-scale structure of Universe.

1 INTRODUCTION

Understanding the clustering of dark matter haloes has been a topic of active research for many years. A number of analytic approaches have been developed to tackle this issue such as the peak model (Bardeen et al. 1986, hereafter BBKS), the excursion set framework (Bond et al. 1991) or perturbation theory (see e.g. Bernardeau et al. 2002, for a review). Heuristic arguments like the peak-background split (Kaiser 1984) and approximations like local bias (Fry & Gaztanaga 1993) have been very helpful for modelling the clustering of dark matter haloes. Nevertheless, improvements in computational power and numerical algorithms as well as the advent of large-scale galaxy surveys have considerably increased the need for an accurate description of halo clustering. Until recently, however, it was unclear how the peak approach, which is thus far the only framework in which biased tracers form a discrete point set, relates to the more widespread excursion set theory, local bias approximation or peak-background split argument.

Working out this connection has been the subject of several recent papers. Desjacques (2013), building on earlier work by Desjacques et al. (2010), showed that correlation functions of discrete density peaks can be computed using an effective (i.e. which does not involve measurable counts-in-cells quantities) generalized bias expansion in which all the bias parameters, including those of the non-local terms,[1] can be computed from a peak-background split. In parallel, Paranjape & Sheth (2012) demonstrated how the peak formalism, which deals with statistics of density maxima at a fixed smoothing scale, can be combined with excursion set theory, whose basic building block is the density contrast at various filtering scales. Similar ideas can already be found in the early work of Bond (1989). Paranjape, Sheth & Desjacques (2013a, hereafter PSD) subsequently computed the mass function and linear bias of haloes within this excursion set peak (ESP) approach and showed that it agrees very well with simulation data.

The focus of this work is on the second-order non-local bias terms predicted by the ESP approach. These generate corrections to the Fourier peak bias of the form $-2\chi_{10}(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)$ and $\chi_{01}[3(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)^2 - k_1^2 k_2^2]$ (Desjacques 2013). What makes them quite interesting is the fact that there are related to χ^2 rather than normally distributed variables. Here, we will show how one can measure their amplitude in the bias of dark matter haloes without computing any correlation function. Of course, this technique can also be applied to measure

---

[1] To facilitate the comparison with other studies, we will call non-local terms all contributions to Lagrangian clustering that are not of the form $\delta^n(\boldsymbol{x})$, where $\delta(\boldsymbol{x})$ is the linear mass density field.

★ E-mail: Matteo.Biagetti@unige.ch

Published by Oxford University Press on behalf of the Royal Astronomical Society

non-local Lagrangian bias contributions induced by e.g. the tidal shear, but this will be the subject of future work.

This paper is organized as follows. In the first part, we will advocate a slight modification of the original ESP formulation of PSD in order to easily write down the corresponding effective bias expansion (Section 2). Next, we will explain how the cross-correlation technique proposed by Musso, Paranjape & Sheth (2012), which has already been successfully applied to the bias factors associated with the density field (PSD; Paranjape et al. 2013b), can be extended to measure the second-order non-local bias factors χ_{10} and χ_{01} that weight the two quadratic, non-local bias contributions (Section 3). Finally, we will validate our method with peaks of Gaussian random fields before measuring χ_{10} and χ_{01} for dark matter haloes (Section 4). We conclude in Section 5.

2 EXCURSION SET PEAKS

In this section, we apply the excursion set approach to the peak model in the case of a moving barrier to get a prediction of the halo mass function which we compare to simulations. We then get expressions for bias parameters, generalizing results in Desjacques (2013) and Desjacques, Gong & Riotto (2013). We also point out a few changes to PSD. We show that, as far as the mass function is concerned, these modifications do not make much difference (only few per cent, in agreement with what PSD found), but they affect first- and second-order bias parameters, as new terms arise.

2.1 Notation

We will adopt the following notation for the variance of the smoothed density field (linearly extrapolated to present day) and its derivatives:

$$\sigma_{j\alpha}^2 = \frac{1}{2\pi^2}\int_0^\infty \mathrm{d}k\, P(k) k^{2(j+1)} W_\alpha^2(kR_\alpha), \tag{1}$$

where $P(k)$ is the power spectrum of the mass density field, $W_\alpha(kR_\alpha)$ and the subscript $\alpha = \mathrm{G}$ or T will denote Gaussian or tophat filtering, respectively. Moreover, R_α is the Lagrangian smoothing scale (which may depend on the choice of kernel). Denoting δ_T and δ_G the linear density field smoothed with a tophat and Gaussian filter, respectively, we introduce the variables

$$\begin{aligned}\nu(\boldsymbol{x}) &= \frac{1}{\sigma_{0\mathrm{T}}}\delta_\mathrm{T}(\boldsymbol{x})\\ u(\boldsymbol{x}) &= -\frac{1}{\sigma_{2\mathrm{G}}}\nabla^2\delta_\mathrm{G}(\boldsymbol{x})\\ \mu(\boldsymbol{x}) &= -\frac{\mathrm{d}\delta_\mathrm{T}}{\mathrm{d}R_\mathrm{T}}(\boldsymbol{x}).\end{aligned} \tag{2}$$

Note that, while ν and u have unit variance, μ is not normalized. We will use the notation $\langle\mu^2\rangle = \Delta_0^2$ in what follows.

Cross-correlations among these three variables are useful and will be denoted as

$$\langle\nu u\rangle = \gamma_1 = \frac{\sigma_{1X}^2}{\sigma_{0T}\sigma_{2\mathrm{G}}} \tag{3}$$

$$\langle\nu\mu\rangle = \gamma_{\nu\mu} = \frac{1}{\sigma_{0T}}\int_0^\infty \frac{\mathrm{d}k}{2\pi^2} P(k)k^2 W_\mathrm{T}(kR_\mathrm{T})\frac{\mathrm{d}W_\mathrm{T}(kR_\mathrm{T})}{\mathrm{d}R_\mathrm{T}} \tag{4}$$

$$\langle u\mu\rangle = \gamma_{u\mu} = \frac{1}{\sigma_{2\mathrm{G}}}\int_0^\infty \frac{\mathrm{d}k}{2\pi^2} P(k)k^4 W_\mathrm{G}(kR_\mathrm{G})\frac{\mathrm{d}W_\mathrm{T}(kR_\mathrm{T})}{\mathrm{d}R_\mathrm{T}}. \tag{5}$$

The first-order, mixed spectral moment σ_{1X} is

$$\sigma_{1X}^2 = \frac{1}{2\pi^2}\int \mathrm{d}k\, P(k)k^4 W_\mathrm{T}(kR_\mathrm{T})W_\mathrm{G}(kR_\mathrm{G}), \tag{6}$$

i.e. one filter is tophat and the other Gaussian.

2.2 First-crossing and moving barrier

2.2.1 *Summary of previous results*

Let us first summarize the basic ideas behind the ESP approach introduced by Paranjape & Sheth (2012) and further developed in PSD and Desjacques et al. (2013).

The excursion set approach states that a region of mass M has virialized when the overdensity $\delta(R)$, where $R \sim M^{1/3}$ is the filtering scale associated with the perturbation, reaches the spherical collapse threshold δ_c provided that, for any $R' > R$, the inequality $\delta(R) < \delta_\mathrm{c}$ holds. This last condition formally implies an infinite set of constraints (one at each smoothing scale). However, as was shown in Musso & Sheth (2012), the requirement $\delta(R + \Delta R) < \delta_\mathrm{c}$ with $\Delta R \ll 1$ furnishes a very good approximation. This follows from the fact that the trajectory described by $\delta(R)$ as a function of R is highly correlated for large radii. As a result, if δ crosses δ_c at R, then it is almost certainly below the threshold at any larger radius.

This first-crossing condition can be combined with the peak constraint, so that peaks on a given smoothing scale are counted only if the inequality above is satisfied. In this case, the effective peak bias expansion introduced in Desjacques (2013) is modified through the presence of a new variable μ (equation 2) which, as was shown in Desjacques et al. (2013), reflects the dependence of bias to the first-crossing condition.

2.2.2 *Modifications to PSD*

We made a couple of modifications to the approach of PSD, which we will now describe in more detail.

First, PSD used the fact that $\mu \equiv u$ when Gaussian filtering is also applied to the density field, so that the first-crossing condition can be accounted for with the variable u only. When δ is smoothed with a tophat filter however, one should in principle deal explicitly with μ and, therefore, consider the trivariate normal distribution $\mathcal{N}(\nu, u, \mu)$. We will proceed this way.

Secondly, Sheth, Mo & Tormen (2001) argued that, owing to the triaxiality of collapse, the critical density for collapse is not constant and equal to $\delta_\mathrm{c} = 1.68$, but rather distributed around a mean value which increases with decreasing halo mass. Analyses of N-body simulations have confirmed this prediction and showed the scatter around the mean barrier is always significant (Dalal et al. 2008; Robertson et al. 2009; Elia, Ludlow & Porciani 2012). Since the stochasticity induced by triaxial collapse is somewhat cumbersome to implement in analytic models of halo collapse (see e.g. Hahn & Paranjape 2014, for a tentative implementation with the peak constraint), we will consider a simple approximation calibrated with numerical simulations (note that it differs from the diffusing barrier approach of Maggiore & Riotto 2010). Namely the square-root stochastic barrier

$$B = \delta_\mathrm{c} + \beta\sigma_0, \tag{7}$$

wherein the stochastic variable β closely follows a log-normal distribution, furnishes a good description of the critical collapse threshold as a function of halo mass (Robertson et al. 2009). In PSD, this result was interpreted as follows: each halo 'sees' a moving barrier

$B = \delta_c + \beta\sigma_0$ with a value of β drawn from a log-normal distribution. Therefore, the first-crossing condition becomes

$$B < \delta < B + \left(B' + \mu\right)\Delta R, \tag{8}$$

where the prime designates a derivative w.r.t. the filtering scale. Here, however, we will assume that each halo 'sees' a constant (flat) barrier, whose height varies from halo to halo. Therefore, we will implement the first-crossing condition simply as

$$B < \delta < B + \mu\,\Delta R. \tag{9}$$

Consequently, the variable μ will satisfy the constraint $\mu > 0$ rather than $\mu > -B'$.

With the aforementioned modifications, the ESP multiplicity function reads

$$\begin{aligned} f_{\rm ESP}(\nu_c) = &\left(\frac{V}{V_*}\right)\frac{1}{\gamma_{\nu\mu}\nu_c}\int_0^\infty {\rm d}\beta\, p(\beta) \\ &\times \int_0^\infty {\rm d}\mu\,\mu\int_0^\infty {\rm d}u\, f(u,\alpha=1)\mathcal{N}(\nu_c+\beta,u,\mu), \end{aligned} \tag{10}$$

where V is the Lagrangian volume associated with the tophat smoothing filter, V_* is the characteristic volume of peaks, $p(\beta)$ is a log-normal distribution for which we take $\langle\beta\rangle = 0.5$ and ${\rm Var}(\beta) = 0.25$ as in PSD and $f(u, \alpha)$ is the slightly modified form (see Desjacques et al. 2010) of the original curvature function of Bardeen et al. (1986); see Appendix A. We can now apply Bayes' theorem and write $\mathcal{N}(\nu,u,\mu) = \mathcal{N}(\nu,u)\mathcal{N}(\mu|\nu,u)$. The integral over μ,

$$\int_0^\infty {\rm d}\mu\,\mu\,\mathcal{N}(\mu|\nu,u), \tag{11}$$

is the same as in Musso & Sheth (2012) and, therefore, is equal to

$$\bar\mu\left[\frac{1+{\rm erf}(\bar\mu/\sqrt{2}\Sigma)}{2} + \frac{\Sigma}{\sqrt{2\pi}\bar\mu}{\rm e}^{-\bar\mu^2/2\Sigma^2}\right], \tag{12}$$

where

$$\bar\mu = u\left(\frac{\gamma_{u\mu}-\gamma_1\gamma_{\nu\mu}}{1-\gamma_1^2}\right) + (\nu+\beta)\left(\frac{\gamma_{\nu\mu}-\gamma_1\gamma_{u\mu}}{1-\gamma_1^2}\right) \tag{13}$$

$$\Sigma^2 = \Delta_0^2 - \frac{\gamma_{\nu\mu}^2 - 2\gamma_1\gamma_{\nu\mu}\gamma_{u\mu} + \gamma_{u\mu}^2}{1-\gamma_1^2}. \tag{14}$$

Substituting this expression into equation (10) and performing numerically the integrals over u and β, we obtain an analytic prediction for the halo mass function without any free parameter. Our ESP mass function differs at most by 2–3 per cent over the mass range 10^{11}–$10^{15}\,{\rm M}_\odot\,h^{-1}$ from that obtained with the prescription of PSD. Likewise, the linear and quadratic local bias parameters are barely affected by our modifications.

2.3 Comparison with numerical simulations

To test the validity of our approach, we compare the ESP mass function with that of haloes extracted from N-body simulations. For this purpose, we ran a series of N-body simulations evolving 1024^3 particles in periodic cubic boxes of size 1500 and 250 h^{-1} Mpc. The particle mass thus is 2.37×10^{11} and $1.10 \times 10^9\,{\rm M}_\odot\,h^{-1}$, respectively. The transfer function was computed with CAMB (Lewis, Challinor & Lasenby 2000) assuming parameter values consistent with those inferred by *WMAP*7 (Komatsu et al. 2011): a flat ΛCDM cosmology with $h = 0.704$, $\Omega_{\rm m} = 0.272$, $\Omega_{\rm b} = 0.0455$, $n_{\rm s} = 0.967$ and a normalization amplitude $\sigma_8 = 0.81$. Initial conditions were laid down at redshift $z = 99$ with an initial particle displacement computed at second order in Lagrangian perturbation theory with 2LPTic (Crocce, Pueblas & Scoccimarro 2006). The simulations were run using the N-body code GADGET-2 (Springel 2005) while the haloes were identified with the spherical overdensity halo finder AHF (Knollmann & Knebe 2009) assuming an overdensity threshold $\Delta_c = 200$ constant throughout redshift.

In Fig. 1, we compare the simulated halo mass function to the ESP prediction at redshift $z = 0$ and 1. The latter can be straightforwardly obtained from the multiplicity function $f_{\rm ESP}(\nu_c)$ as

$$\begin{aligned} \frac{{\rm d}\bar n_{\rm h}}{{\rm d}\ln M} &= \frac{\bar\rho}{M}\nu_c f_{\rm ESP}(\nu_c, R_{\rm s})\frac{{\rm d}\log\nu_c}{{\rm d}\log M} \\ &= -3R_{\rm T}\left(\frac{\gamma_{\nu\mu}\nu_c}{\sigma_{0T}}\right)V^{-1}f_{\rm ESP}(\nu_c), \end{aligned} \tag{15}$$

where we used the fact that $\gamma_{\nu\mu} = \sigma_{0T}'$ to obtain the second equality. The ESP prediction agrees with the simulations at the 10 per cent level or better from $10^{14}\,{\rm M}_\odot\,h^{-1}$ down to a halo mass $10^{11}\,{\rm M}_\odot\,h^{-1}$, where the correspondence between virialized haloes and initial density peaks should be rather vague. The abundance of very rare clusters with $M > 10^{14}\,{\rm M}_\odot\,h^{-1}$ is difficult to predict because of exponential sensitivity to δ_c. In this respect, it might be more appropriate to work with a critical linear density $\delta_c \approx 1.60$ if haloes are defined with a fixed non-linear threshold $\Delta_c = 200$ relative to the mean density (see, e.g. Barkana 2004; Valageas 2009, for a discussion).

2.4 Bias parameters

The bias factors of ESP can be computed using the same formulas as in Desjacques (2013). With the additional variable μ, the 'localized' number density (in the terminology of Matsubara 2012) can be written as (Desjacques et al. 2013)

$$n_{\rm ESP}(\boldsymbol{w}) = -\left(\frac{\mu}{\gamma_{\nu\mu}\nu_c}\right)\theta_H(\mu)\,n_{\rm pk}(\boldsymbol{y}), \tag{16}$$

where $n_{\rm pk}$ is the localized number density of BBKS peaks and $\boldsymbol{w} = (\nu, \eta_i, \zeta_{ij}, \mu) \equiv (\boldsymbol{y}, \mu)$ is an 11-dimensional vector containing all the independent variables of the problem. Therefore,

$$\begin{aligned} \sigma_{0T}^i\sigma_{2G}^j b_{ijk} &= \frac{1}{\bar n_{\rm ESP}}\int {\rm d}^{11}\boldsymbol{w}\, n_{\rm ESP}(\boldsymbol{w})H_{ijk}(\nu,u,\mu)P_1(\boldsymbol{w}) \\ \sigma_{1G}^{2k}\chi_{k0} &= \frac{(-1)^k}{\bar n_{\rm ESP}}\int {\rm d}^{11}\boldsymbol{w}\, n_{\rm ESP}(\boldsymbol{w})L_k^{(1/2)}\!\left(\frac{3\eta^2}{2}\right)P_1(\boldsymbol{w}) \\ \sigma_{2G}^{2k}\chi_{0k} &= \frac{(-1)^k}{\bar n_{\rm ESP}}\int {\rm d}^{11}\boldsymbol{w}\, n_{\rm ESP}(\boldsymbol{w})L_k^{(3/2)}\!\left(\frac{5\zeta^2}{2}\right)P_1(\boldsymbol{w}). \end{aligned} \tag{17}$$

Here, $P_1(\boldsymbol{w})$ is the one-point probability density

$$\begin{aligned} P_1(\boldsymbol{w}){\rm d}^{11}\boldsymbol{w} = &\,\mathcal{N}(\nu,u,\mu)\,{\rm d}\nu\,{\rm d}u\,{\rm d}\mu \times \chi_3^2(3\eta^2)\,{\rm d}(3\eta^2) \\ &\times\chi_5^2(5\zeta^2)\,{\rm d}(5\zeta^2)\times P({\rm angles}), \end{aligned} \tag{18}$$

where $H_{ijk}(\nu, u, \mu)$ are trivariate Hermite polynomials and $\chi_k^2(x)$ is a χ^2-distribution with k degrees of freedom (d.o.f.). The probability density P (angles; which was missing[2] in Desjacques 2013) represents the probability distribution of the five remaining d.o.f. Since they are all angular variables, they do not generate bias factors

---

[2] We thank Marcello Musso for pointing this out to us.

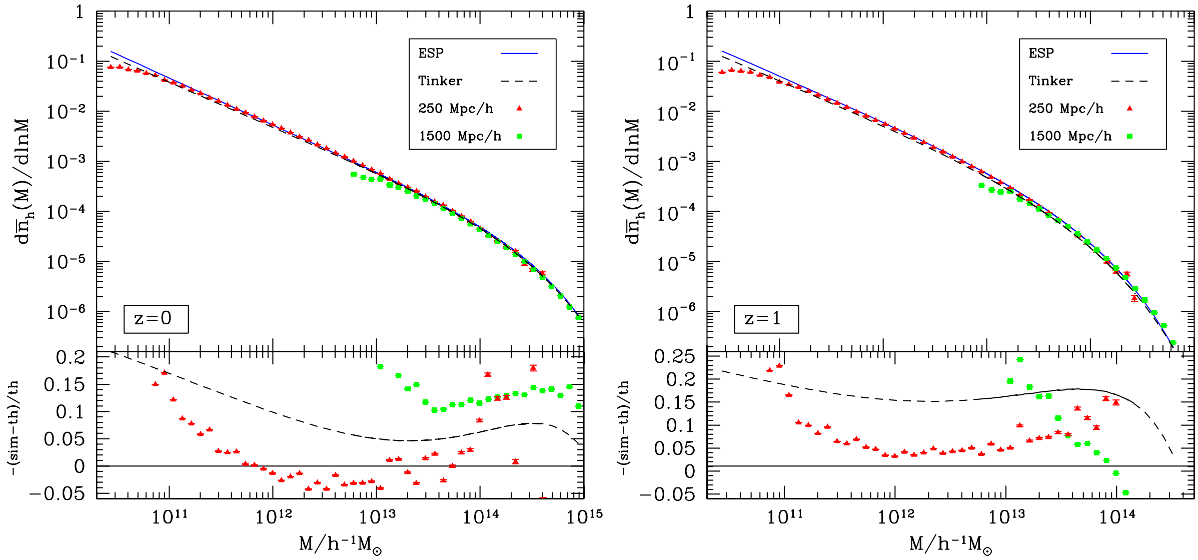

Figure 1. Halo mass function measured from *N*-body simulation at redshift $z = 0$ (left-hand panel) and $z = 1$ (right-hand panel) with different box sizes as indicated in the figures. The error bars are Poisson. The data are compared to the theoretical prediction (equation 15) based on the ESP formalism and the fitting formula of Tinker et al. (2008). We also show the fractional deviation of the Tinker et al. (2008) and the measured halo mass function relative to our theoretical prediction.

because the peak (and halo) overabundance can only depend on scalar quantities (e.g. Catelan, Matarrese & Porciani 1998; McDonald & Roy 2009).

The behaviour of the bias factors b_{ij0} and χ_{kl} as a function of halo mass is similar to that seen in fig. 1 of Desjacques (2013). The bias factors b_{ijk} with $k \geq 1$ weight the contributions of μ^k terms to the clustering of ESP that are proportional to derivatives of the tophat filter w.r.t. the filtering scale $R_{\rm T}$. Similar contributions appear in the clustering of thresholded regions (Matsubara 2012; Ferraro et al. 2013) since their definition also involve a first-crossing condition.

The effective bias expansion takes the form (Desjacques 2013; Desjacques et al. 2013)

$$\begin{aligned}\delta_{\rm pk}(\boldsymbol{x}) = {} & \sigma_{0T} b_{100}\nu(\boldsymbol{x}) + \sigma_{2\rm G} b_{010} u(\boldsymbol{x}) + b_{001}\mu(\boldsymbol{x}) \\ & + \frac{1}{2}\sigma_{0T}^2 b_{200}\nu^2(\boldsymbol{x}) + \sigma_{0T}\sigma_{2\rm G} b_{110}\nu(\boldsymbol{x})u(\boldsymbol{x}) \\ & + \frac{1}{2}\sigma_{2\rm G}^2 b_{020}u^2(\boldsymbol{x}) + \frac{1}{2} b_{002}\mu^2(\boldsymbol{x}) \\ & + \sigma_{1\rm G}^2\chi_{10}\eta^2(\boldsymbol{x}) + \sigma_{2\rm G}^2\chi_{01}\zeta^2(\boldsymbol{x}) \\ & + \sigma_{0T} b_{101}\nu(\boldsymbol{x})\mu(\boldsymbol{x}) + \sigma_{2\rm G} b_{011}u(\boldsymbol{x})\mu(\boldsymbol{x}) + \cdots . \end{aligned} \tag{19}$$

Here, the rule of thumb is that one should ignore all the terms involving zero-lag moments in the computation of $\langle\delta_{\rm pk}(\boldsymbol{x}_1), \ldots, \delta_{\rm pk}(\boldsymbol{x}_N)\rangle$ in order to correctly predict the *N*-point correlation function, as demonstrated explicitly in Desjacques (2013). The appearance of rotationally invariant quantities is, again, only dictated by the scalar nature of the peak overabundance. The variables of interest here are

$$\begin{aligned}\eta^2(\boldsymbol{x}) &= \frac{1}{\sigma_{1\rm G}^2}(\nabla\delta)^2(\boldsymbol{x}) \\ \zeta^2(\boldsymbol{x}) &= \frac{3}{2\sigma_{2\rm G}^2}\mathrm{tr}\left[\left(\partial_i\partial_j\delta - \frac{1}{3}\delta_{ij}\nabla^2\delta\right)^2\right](\boldsymbol{x}),\end{aligned} \tag{20}$$

so that $3\eta^2(\boldsymbol{x})$ and $5\zeta^2(\boldsymbol{x})$ are χ^2-distributed with 3 and 5 d.o.f., respectively.

3 BIASES FROM CROSS-CORRELATION: EXTENSION TO χ^2 VARIABLES

In this section, we will demonstrate that the bias factors χ_{ij} can be measured with a one-point statistics. We will test our method on density peaks of a Gaussian random field before applying it to dark matter haloes.

3.1 Bias factors b_{ijk}: Hermite polynomials

Musso et al. (2012) showed that the bias factors of discrete tracers (relative to the mass density δ) can be computed from one-point measurements rather than computationally more expensive *n*-point correlations. Their idea was implemented by PSD and Paranjape et al. (2013b) to haloes extracted from *N*-body simulations in order to test the predictions of the ESP formalism. Namely haloes were traced back to their 'protohalo' patch (since one is interested in measuring Lagrangian biases) in the initial conditions, the linear density field was smoothed on some 'large-scale' $R_{\rm l}$ and the quantity $H_n(\nu_{\rm l} = \delta_{\rm l}/\sigma_{0l})$ was computed (for $n = 1, 2$ only) at the location of each protohalo. The average of $H_n(\nu_{\rm l})$ over all protohaloes reads

$$\frac{1}{N}\sum_{i=1}^{N} H_n(\nu_{\rm l}) = \int_{-\infty}^{+\infty} \mathrm{d}\nu_{\rm l}\,\mathcal{N}(\nu_{\rm l})\langle 1 + \delta_{\rm h}|\nu_{\rm l}\rangle H_n(\nu_{\rm l}), \tag{21}$$

where $\delta_{\rm h}$ is the overdensity of protohaloes. This expression assumes that the first-crossing condition can be implemented through a constraint of the form equation (9), so that $P(\nu_{\rm l})$ is well approximated by a Gaussian (Musso et al. 2012). For the ESP considered here, this ensemble average reads

$$\begin{aligned}&\frac{1}{\bar{n}_{\rm ESP}}\int_{-\infty}^{+\infty}\mathrm{d}\nu_{\rm l}\,\mathcal{N}(\nu_{\rm l})\langle n_{\rm ESP}|\nu_{\rm l}\rangle H_n(\nu_{\rm l}) \\ &\quad = \frac{1}{\bar{n}_{\rm ESP}}\int \mathrm{d}^{11}\boldsymbol{w}\, n_{\rm ESP}(\boldsymbol{w})\,(-\epsilon_\nu)^n \\ &\qquad \times \left(\frac{\partial}{\partial\nu} + \frac{\epsilon_u}{\epsilon_\nu}\frac{\partial}{\partial u} + \frac{\epsilon_\mu}{\epsilon_\nu}\frac{\partial}{\partial\mu}\right)^n P_{\rm l}(\boldsymbol{w}).\end{aligned} \tag{22}$$

Here, ϵ_X denotes the cross-correlation between ν_l and the variables $X = (\nu, u, \mu)$ defined at the halo smoothing scale. The right-hand side reduces to a sum of nth-order bias factors b_{ijk} weighted by products of ϵ_ν, ϵ_u and ϵ_μ. Relations between bias factors of a given order (which arise owing to their close connection with Hermite polynomials, see e.g. Musso et al. 2012) can then be used to extract a measurement of each b_{ijk}.

Before we generalize this approach to the chi-squared bias factors χ_{ij}, we emphasize that, in this cross-correlation approach, the smoothing scale R_l can take any value as long as it is distinct from the halo smoothing scale. Paranjape et al. (2013b) chose $R_\mathrm{l} \gg R_\mathrm{s}$ in the spirit of the peak-background split but this requirement is, in fact, not necessary as long as the correlation between the two scales is taken into account. In any case, we will stick with the notation R_l for convenience.

3.2 Bias factors χ_{ij}: Laguerre polynomials

The approach presented above can be generalized to χ^2 distributions. The main difference is the appearance of Laguerre polynomials $L_n^{(\alpha)}$. Consider for instance the χ^2-quantity $3\eta^2$ smoothed at the scale R_l, i.e. $3\eta_\mathrm{l}^2$. In analogy with equation (21), the ensemble average of $L_n^{(1/2)}(3\eta_\mathrm{l}^2)$ at the peak positions is

$$\frac{1}{N}\sum_{i=1}^{N} L_n^{(1/2)}\left(\frac{3\eta_\mathrm{l}^2}{2}\right) = \int_0^\infty \mathrm{d}(3\eta_\mathrm{l}^2)\,\chi_3^2(3\eta_\mathrm{l}^2) \times \left\langle 1+\delta_\mathrm{h}|3\eta_\mathrm{l}^2\right\rangle L_n^{(1/2)}\left(\frac{3\eta_\mathrm{l}^2}{2}\right). \tag{23}$$

The conditional average $\langle 1+\delta_\mathrm{h}|3\eta_\mathrm{l}^2\rangle$ reads

$$\begin{aligned}\left\langle 1+\delta_\mathrm{h}|3\eta_\mathrm{l}^2\right\rangle &= \frac{1}{\bar{n}_\mathrm{ESP}}\int \mathrm{d}^{11}\boldsymbol{w}\, n_\mathrm{ESP}(\boldsymbol{w}) P_1\left(\boldsymbol{w}|3\eta_\mathrm{l}^2\right)\\ &= \frac{1}{\bar{n}_\mathrm{ESP}}\int \mathrm{d}u\,\mathrm{d}\nu\,\mathrm{d}\mu\,\mathcal{N}(\nu,u,\mu)\\ &\quad\times\int \mathrm{d}(3\eta^2)\,\chi_3^2(3\eta^2|3\eta_\mathrm{l}^2)\int \mathrm{d}(5\zeta^2)\,\chi_5^2(5\zeta^2)\\ &\quad\times\int \mathrm{d(angles)}\,P(\mathrm{angles})\,n_\mathrm{ESP}(\boldsymbol{w}).\end{aligned} \tag{24}$$

We substitute this relation into equation (23) and begin with the integration over the variable $3\eta_\mathrm{l}^2$.

We use the following relation (which can be inferred from equation 7.414 of Gradshteyn & Ryzhik 1994):

$$\int_0^\infty \mathrm{d}x\,\mathrm{e}^{-x} x^{j+\alpha} L_n^{(\alpha)}(x) = \frac{(-1)^n}{n!}\frac{j!\,\Gamma(j+\alpha+1)}{(j-n)!}. \tag{25}$$

With the aid of this result and on expanding the conditional χ^2-distribution $\chi_3^2(3\eta^2|3\eta_\mathrm{l}^2)$ in Laguerre polynomials (see Appendix B for details), we obtain

$$\begin{aligned}&\int_0^\infty \mathrm{d}(3\eta_\mathrm{l}^2)\,\chi_3^2(3\eta_\mathrm{l}^2)\,L_n^{(1/2)}\left(\frac{3\eta_\mathrm{l}^2}{2}\right)\,\chi_3^2(3\eta^2|3\eta_\mathrm{l}^2)\\ &= \frac{(-1)^n}{n!}\frac{1}{\Gamma(3/2)}\left(\frac{3\eta^2}{2}\right)^\alpha \frac{\mathrm{e}^{-3\eta^2/2(1-\epsilon^2)}}{2\left(1-\epsilon^2\right)^{\alpha+1}}\\ &\quad\times\sum_{j=0}^{\infty}\frac{j!}{(j-n)!}\left(\frac{-\epsilon^2}{1-\epsilon^2}\right)^j L_j^{(1/2)}\left[\frac{3\eta^2}{2(1-\epsilon^2)}\right].\end{aligned} \tag{26}$$

For simplicity, let us consider the cases $n = 0$, 1 solely. For $n = 0$, the sum simplifies to

$$\begin{aligned}&\sum_{j=0}^{\infty}\left(\frac{-\epsilon^2}{1-\epsilon^2}\right)^j L_j^{(1/2)}\left[\frac{3\eta^2}{2(1-\epsilon^2)}\right]\\ &\quad= \left(1-\epsilon^2\right)^{3/2}\exp\left[\left(\frac{\epsilon^2}{1-\epsilon^2}\right)\frac{3\eta^2}{2}\right],\end{aligned} \tag{27}$$

and the integral equation (26) ($L_0^{(1/2)}(3\eta_\mathrm{l}^2/2) \equiv 1$) is trivially equal to $\chi_3^2(3\eta^2)$ (as it should be, since we are essentially marginalizing over $3\eta_\mathrm{l}^2$).

For $n \geq 1$, the sum can be evaluated upon taking suitable derivatives of the right-hand side of equation (27), which indeed is a generating function for the Laguerre polynomials $L_n^{(1/2)}$. For $n = 1$, a little algebra leads to

$$\begin{aligned}&\sum_{j=0}^{\infty} j\left(\frac{-\epsilon^2}{1-\epsilon^2}\right)^{j-1} L_j^{(1/2)}\left[\frac{3\eta^2}{2(1-\epsilon^2)}\right]\\ &\quad= \left(1-\epsilon^2\right)^{5/2} L_1^{(1/2)}\left(\frac{3\eta^2}{2}\right)\exp\left[\left(\frac{\epsilon^2}{1-\epsilon^2}\right)\frac{3\eta^2}{2}\right].\end{aligned} \tag{28}$$

Hence, equation (26) with $n = 1$ equals $\epsilon^2 L_1^{(1/2)}(3\eta^2/2)\chi_3^2(3\eta^2)$. Performing the remaining integrals over ν, u, μ and $5\zeta^2$ (the integral over the angles is trivially unity) and taking into account the ESP constraint through the multiplicative factor $n_\mathrm{ESP}(\boldsymbol{w})$, equation (23) simplifies to

$$\int_0^\infty \mathrm{d}(3\eta_\mathrm{l}^2)\,\chi_3^2(3\eta_\mathrm{l}^2)\,\left\langle 1+\delta_\mathrm{h}|3\eta_\mathrm{l}^2\right\rangle\,L_1^{(1/2)}\left(\frac{3\eta_\mathrm{l}^2}{2}\right) = -\epsilon^2\sigma_1^2\chi_{10}. \tag{29}$$

For the variable $3\eta^2$, the cross-correlation coefficient ϵ is

$$\epsilon^2 \equiv \frac{\langle\eta^2\eta_\mathrm{l}^2\rangle - \langle\eta^2\rangle\langle\eta_\mathrm{l}^2\rangle}{\sqrt{\left(\langle\eta^4\rangle-\langle\eta^2\rangle^2\right)\left(\langle\eta_\mathrm{l}^4\rangle-\langle\eta_\mathrm{l}^2\rangle^2\right)}} = \left(\frac{\sigma_{1X}^2}{\sigma_{1\mathrm{s}}\sigma_{1\mathrm{l}}}\right)^2, \tag{30}$$

which we shall denote as ϵ_1 in what follows. Furthermore,

$$\sigma_{n\times}^2 = \frac{1}{2\pi^2}\int_0^\infty \mathrm{d}k\,k^{2(n+1)}\,P(k)W_\mathrm{G}(kR_\mathrm{s})W_\mathrm{G}(kR_\mathrm{l}) \tag{31}$$

designates the splitting of filtering scales, i.e. one filter is on scale R_s while the second is on scale R_l. It should be noted that, unlike σ_{1X} defined in equation (6), both filtering kernels are Gaussian.

The derivation of the bias factors χ_{0k} associated with the quadratic variable ζ^2 proceeds analogously. In particular,

$$\int_0^\infty \mathrm{d}(5\zeta_\mathrm{l}^2)\,\chi_5^2(5\zeta_\mathrm{l}^2)\,\left\langle 1+\delta_\mathrm{h}|5\zeta_\mathrm{l}^2\right\rangle\,L_1^{(3/2)}\left(\frac{5\zeta_\mathrm{l}^2}{2}\right) = -\epsilon^2\sigma_2^2\chi_{01}. \tag{32}$$

Here, the cross-correlation coefficient is $\epsilon = \sigma_{2X}^2/(\sigma_{2\mathrm{s}}\sigma_{2\mathrm{l}}) \equiv \epsilon_2$. Note that, in both cases, the cross-correlation coefficient drops very rapidly as R_l moves away from R_s for realistic CDM power spectra. In addition, one could in principle choose $R_\mathrm{l} < R_\mathrm{s}$ (if there is enough numerical resolution) to measure χ_{ij}.

4 TEST WITH NUMERICAL SIMULATIONS

In this section, we first validate our predictions based on peaks of Gaussian random fields with measurements extracted from random realizations of the Gaussian linear density field, and then move on to calculate χ_{10} and χ_{01} for $M \gtrsim M_\star$ haloes, where $M_\star$ is the characteristic mass of the haloes.

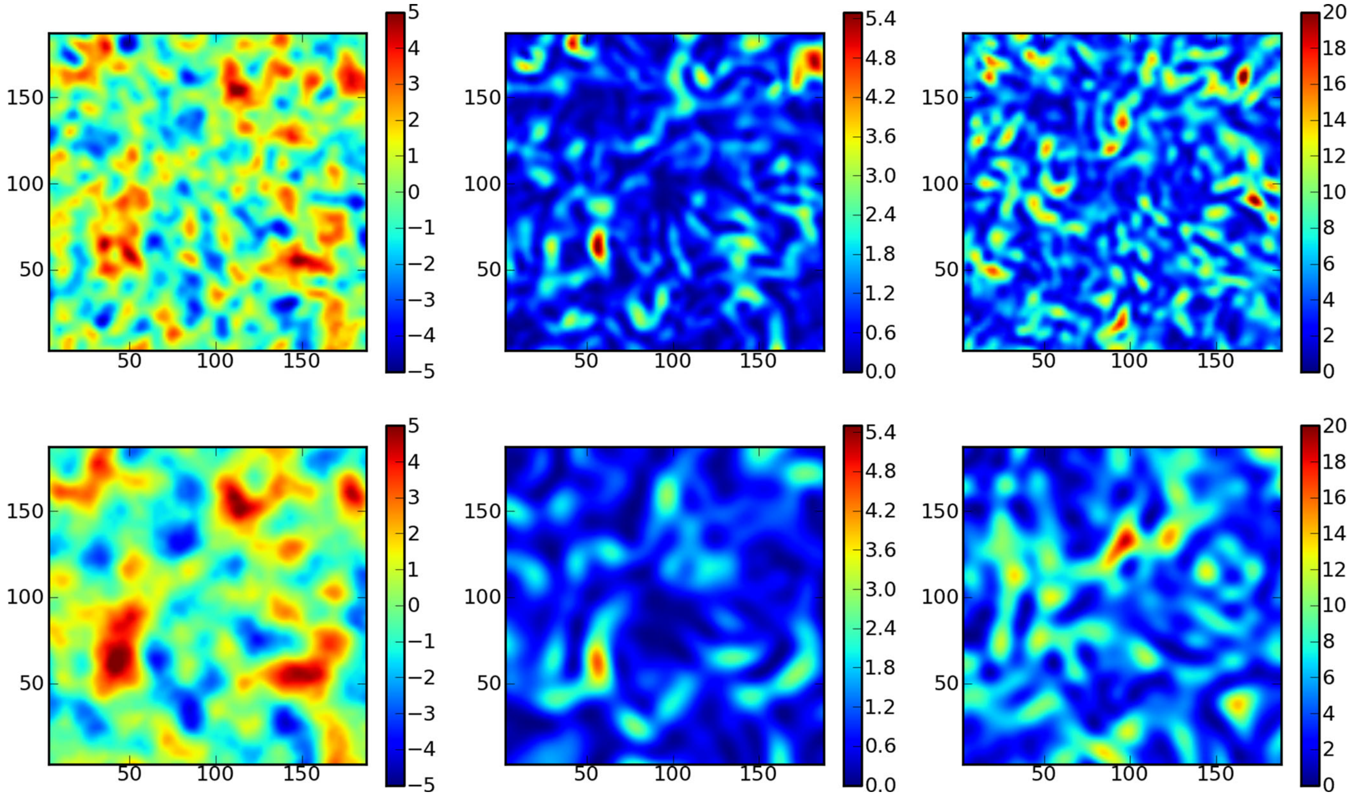

Figure 2. Sections for ν_l, $3\eta_\mathrm{l}^2$ and $5\zeta_\mathrm{l}^2$ (from left to right). A filtering scale of $R_\mathrm{l} = 5$ and $10\,h^{-1}$ Mpc is used for the first and second row, respectively. Note that a tophat kernel is applied for ν_l, while a Gaussian window is used for η_l^2 and ζ_l^2. In each panel, the dimension of the section is $200 \times 200\,h^{-2}\,\mathrm{Mpc}^2$.

4.1 Peaks of Gaussian random fields

We generate random realizations of the Gaussian, linear density field with a power spectrum equal to that used to seed the *N*-body simulations described above. To take advantage of Fast Fourier Transforms (FFTs), we simulate the linear density field in periodic, cubic boxes of side $1000\,h^{-1}$ Mpc. The size of the mesh along each dimension is 1536. We smooth the density field on scale $R_\mathrm{s} = 5\,h^{-1}$ Mpc with a tophat filter and find the local maxima by comparing the density at each grid point with its 26 neighbouring values.

We then smooth the density field on the larger scales $R_\mathrm{l} = 10$, 15 and $20\,h^{-1}$ Mpc with a Gaussian filter and compute

$$\eta_\mathrm{l}^2 = \frac{1}{\sigma_{1\mathrm{l}}^2}(\nabla\delta_\mathrm{l})^2 \tag{33}$$

$$\zeta_\mathrm{l}^2 = \frac{3}{2\sigma_{2\mathrm{l}}^2}\mathrm{tr}\left[\left(\partial_i\partial_j\delta_\mathrm{l} - \frac{1}{3}\delta_{ij}\nabla^2\delta_\mathrm{l}\right)^2\right]. \tag{34}$$

These density fields with derivatives sensitively depend on the smoothing scales used. To illustrate this, we show in Fig. 2 sections of ν_l, $3\eta_\mathrm{l}^2$ and $5\zeta_\mathrm{l}^2$. The sections, each of which of dimensions $200 \times 200\,h^{-2}\,\mathrm{Mpc}^2$, were generated at $z = 99$ with the same random seed. The first row corresponds to $R_\mathrm{s} = 5\,h^{-1}$ Mpc, whereas the second row displays results on the filtering scale $R_\mathrm{l} = 10\,h^{-1}$ Mpc. We note that, for the normalized density field ν_l, an increase in the smoothing scale washes out the small-scale features, but the large-scale pattern remains. For the quadratic variable η_l^2 however, the resemblance between the features at the small and large filtering scale is tenuous. This is even worse for ζ_l^2.

Compared to ν_l, the fields η_l^2 and ζ_l^2 have one and two additional derivatives which give rise to an effective window function whose isotropic part is given by

$$W_\mathrm{eff}(k, R) = k^n \mathrm{e}^{-(kR)^2/2}, \tag{35}$$

where $n = 0$, 1 and 2 for ν_l, η_l^2 and ζ_l^2, respectively. For $n = 0$, the window becomes narrower as R_l increases, yet remains unity for wavenumbers $k \lesssim 1/R_\mathrm{l}$. W_eff reaches a maximum at $\sqrt{n}/R$. Hence, for $n = 1$ and 2, W_eff selects predominantly wavemodes with $k \sim 1/R$. Consequently, since in a Gaussian random field the wavemodes at different scales are uncorrelated, patterns in the fields η_l^2 and ζ_l^2 can change drastically as R_l varies. This effect is expected to be most significant for $n = 2$, i.e. ζ_l^2.

For each local density maxima, we store the peak height ν as well as the value of η_l^2 and ζ_l^2 at the peak position. The left-hand panel of Fig. 3 displays as histograms the resulting probability distribution $P(3\eta_\mathrm{l}^2|\mathrm{pk})$ for three different values of $R_\mathrm{l} = 10$, 15 and $20\,h^{-1}$ Mpc. The solid curves represent the theoretical prediction equation (B4) with $x = \langle 3\eta^2|\mathrm{pk}\rangle = 0$ and $\epsilon_1 = 0.71$, 0.44 and 0.29 (from the smallest to largest R_l) as was measured from the random realizations. The dashed curve is the unconditional χ^2-distribution with 3 d.o.f. The theory gives excellent agreement with the simulations. Note also that we did not find any evidence for a dependence on the peak height, as expected from the absence of a correlation between ν and η_l^2. The right-hand panel of Fig. 3 shows results for ζ_l^2. Here, however, since the cross-correlation coefficient diminishes very quickly when R_l even slightly departs from R_s, we show result for $R_\mathrm{l} = 10\,h^{-1}$ Mpc only, which corresponds to $\epsilon_2 = 0.57$. In addition, because one should expect a dependence of the shape of the density profile around peaks to the peak height, we

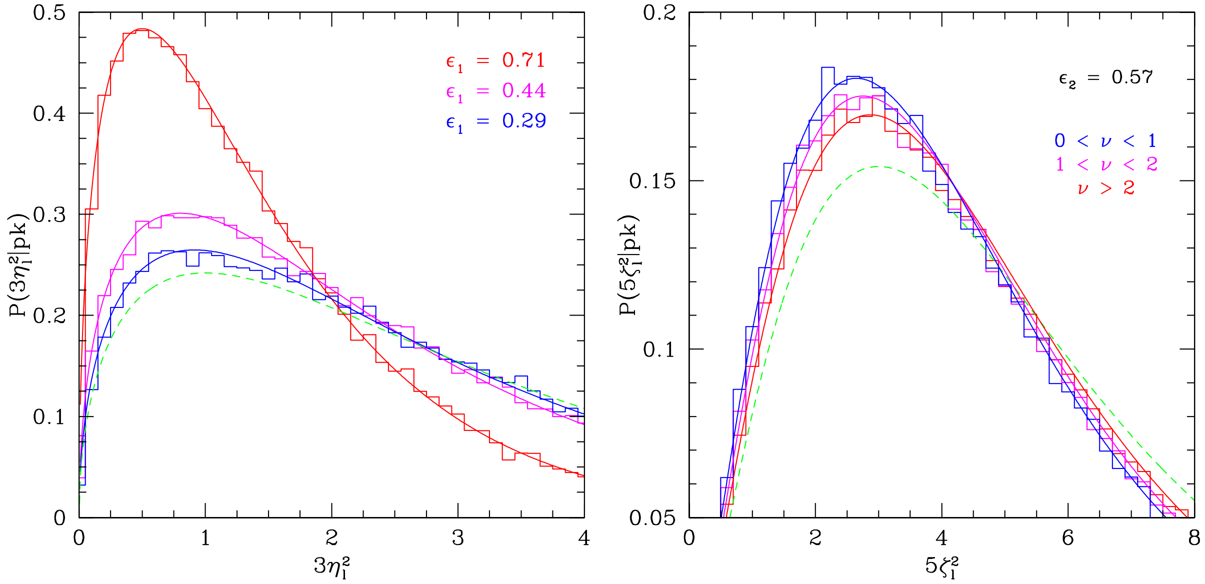

Figure 3. Conditional probability distribution for the variables $3\eta_l^2$ (left-hand panel) and $5\zeta_l^2$ (right-hand panel) measured at the position of maxima of the linear density field smoothed with a Gaussian filter on scale $R = 5\,h^{-1}$ Mpc. Left-hand panel: histograms indicate the results for $R_l = 10$, 15 and $20\,h^{-1}$ Mpc, which leads to $\epsilon_1 = 0.71$, 0.44 and 0.29 as quoted in the figure. Right-hand panel: histograms show the results for a fixed $R_l = 10\,h^{-1}$ Mpc (which implies $\epsilon_2 = 0.57$) but several peak height intervals. In all cases, the solid curves are the theoretical prediction (see text) whereas the dashed (green) curves represents the unconditional distribution $\chi_k^2(y)$.

consider three different ranges of ν as indicated in the figure. The solid curves indicate the theoretical prediction equation (B4) with $\epsilon_2 = 0.57$ and $x = \langle 5\zeta^2|\mathrm{pk}\rangle$, where

$$\langle 5\zeta^2|\mathrm{pk}\rangle = -2\partial_\alpha \ln \int_{\nu_{\min}}^{\nu_{\max}} \mathrm{d}\nu\, G_0^{(\alpha)}(\gamma_1, \gamma_1\nu). \tag{36}$$

Here, $G_0^{(\alpha)}$ is the integral of $f(u, \alpha)$ over all the allowed peak curvatures. The average $\langle 5\zeta^2|\mathrm{pk}\rangle$ increases with the peak height to reach 5 in the limit $\nu \to \infty$. The figure shows a clear deviation from the unconditional distribution $\chi_5^2(5\zeta_l^2)$ (shown as the dashed curve) and a dependence on ν consistent with theoretical predictions.

4.2 Dark matter haloes

Having successfully tested the theory against numerical simulations of Gaussian peaks, we will now attempt to estimate the bias factors χ_{10} and χ_{01} associated with dark matter haloes. For this purpose, we first trace back all dark matter particles belonging to virialized haloes at redshift $z = 0$ to their initial position at $z = 99$. We then compute the centre-of-mass positions of these Lagrangian regions and assume that they define the locations of protohaloes. We can now proceed as for the Gaussian peaks and compute ν, η_l^2 and ζ_l^2 at the position of protohaloes.

The quadratic bias factors χ_{10} and χ_{01} could be in principle computed analogously to Paranjape et al. (2013b), i.e. by stacking measurements of η_l^2 and ζ_l^2 at the locations of protohaloes as

$$\sigma_{1\mathrm{s}}^2 \hat{\chi}_{10} = -\frac{1}{N\epsilon_1^2} \sum_{i=1}^{N} L_1^{(1/2)}\left(\frac{3\eta_l^2}{2}\right) \tag{37}$$

and

$$\sigma_{2\mathrm{s}}^2 \hat{\chi}_{01} = -\frac{1}{N\epsilon_2^2} \sum_{i=1}^{N} L_1^{(3/2)}\left(\frac{5\zeta_l^2}{2}\right). \tag{38}$$

Here, N is the number of haloes, s designates smoothing at the halo mass scale with a Gaussian filter W_G on scale $R_\mathrm{G}(R_\mathrm{T})$, whereas l designates Gaussian smoothing at the large-scale R_l. However, because the cross-correlation coefficient is fairly small unless R_l is very close to R_G, we decided to compute χ_{10} and χ_{01} by fitting the probability distribution $P(3\eta_l^2|\mathrm{halo})$ and $P(5\zeta_l^2|\mathrm{halo})$ with the conditional χ^2-distribution $\chi_k^2(y|x)$. Namely

$$\sigma_{1\mathrm{s}}^2 \hat{\chi}_{10} = \frac{1}{2}\left(\langle 3\eta^2|\mathrm{halo}\rangle - 3\right)$$

$$\sigma_{2\mathrm{s}}^2 \hat{\chi}_{01} = \frac{1}{2}\left(\langle 5\zeta^2|\mathrm{halo}\rangle - 5\right), \tag{39}$$

where $\langle 3\eta^2|\mathrm{halo}\rangle$ and $\langle 5\zeta^2|\mathrm{halo}\rangle$ are the best-fitting values obtained for x. We used measurements obtained at the smoothing scale $R_l = 10\,h^{-1}$ Mpc only to maximize the signal.

To predict the value of R_G given R_T, we followed PSD and assumed that $R_\mathrm{G}(R_\mathrm{T})$ can be computed through the requirement that $\langle \delta_\mathrm{G}|\delta_\mathrm{T}\rangle = \delta_\mathrm{T}$. This yields a prediction for the value of the cross-correlation coefficients ϵ_1 and ϵ_2 as a function of halo mass, which we can use as an input to $\chi_k^2(y|x)$ and only fit for x. However, we found that using the predicted ϵ_1 leads to unphysical (negative) values for x when one attempts to fit $P(3\eta_l^2|\mathrm{halo})$. Therefore, we decided to proceed as follows.

(i) Estimate both ϵ_1 and $x = \langle 3\eta^2|\mathrm{halo}\rangle$ by fitting the model $\chi_3^2(y|x;\epsilon_1)$ to the measured $P(3\eta_l^2|\mathrm{halo})$.

(ii) Compute ϵ_2 assuming that the same R_G enters the spectral moments.

(iii) Estimate $x = \langle 5\zeta_l^2|\mathrm{halo}\rangle$ by fitting the theoretical model $\chi_5^2(y|x;\epsilon_2)$ to the simulated $P(5\zeta_l^2|\mathrm{halo})$.

We considered data in the range $0 < 3\eta_l^2 < 8$ and $0 < 5\zeta_l^2 < 12$ and gave equal weight to all the measurements (assuming Poisson errors does not affect our results significantly). Table 1 summarizes the best-fitting values obtained for four different halo bins

Table 1. Best-fitting parameter values as a function of halo mass. The latter is in unit of $10^{13}\,\mathrm{M}_\odot\,h^{-1}$. Note that we also list the values of ϵ_2 even though it is not directly fitted to the data (see text for details).

Halo mass	$\langle 3\eta^2\vert\mathrm{halo}\rangle$	ϵ_1	$\langle 5\zeta^2\vert\mathrm{halo}\rangle$	(ϵ_2)
$M > 30$	0.71	0.80	2.98	(0.70)
$10 < M < 30$	1.24	0.66	4.49	(0.52)
$3 < M < 10$	1.62	0.54	5.82	(0.37)
$1 < M < 3$	1.94	0.49	6.12	(0.31)

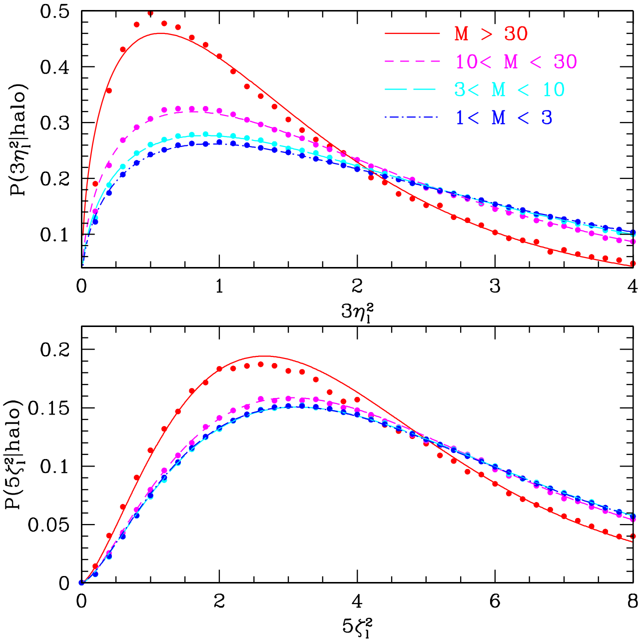

Figure 4. Conditional probability distribution for $3\eta_1^2$ (top panel) and $5\zeta_1^2$ (bottom panel) measured at the centre-of-mass position of protohaloes. The filter is Gaussian with $R_\mathrm{l} = 10\,h^{-1}$ Mpc. The various curves show the best-fitting theoretical predictions for the halo mass bins considered here. Halo mass range is in unit of $10^{13}\,\mathrm{M}_\odot\,h^{-1}$. Poisson errors are much smaller than the size of the data points and, therefore, do not show up in the figure.

spanning the mass range 10^{13}–$10^{15}\,\mathrm{M}_\odot\,h^{-1}$, whereas the measured probability distributions together with the best-fitting models are shown in Fig. 4. The data are reasonably well described by a conditional χ^2-distribution, but the fit is somewhat poorer when the cross-correlation coefficient is close to unity.

The second-order bias factors χ_{10} and χ_{01} of the dark matter haloes at $z = 0$ can be readily computed from equation (39) using the best-fitting values of $\langle 3\eta^2|\mathrm{halo}\rangle$ and $\langle 5\zeta^2|\mathrm{halo}\rangle$. The results are shown in Fig. 5 as the data points. Error bars indicate the scatter among the various realizations and, therefore, likely strongly underestimate the true uncertainty. The dashed curves indicate the predictions of the ESP formalism. The measurements, albeit of the same magnitude as the theoretical predictions, quite disagree with expectations based on our ESP approach, especially χ_{01} which reverses sign as the halo mass drops below $10^{14}\,\mathrm{M}_\odot\,h^{-1}$.

4.3 Interpretation of the measurements

To begin with, we note that, if haloes were forming out of randomly distributed patches in the initial conditions, then both χ_{10} and χ_{01} would be zero since $\langle 3\eta^2\rangle = 3$ and $\langle 5\zeta^2\rangle = 5$ for random field points.

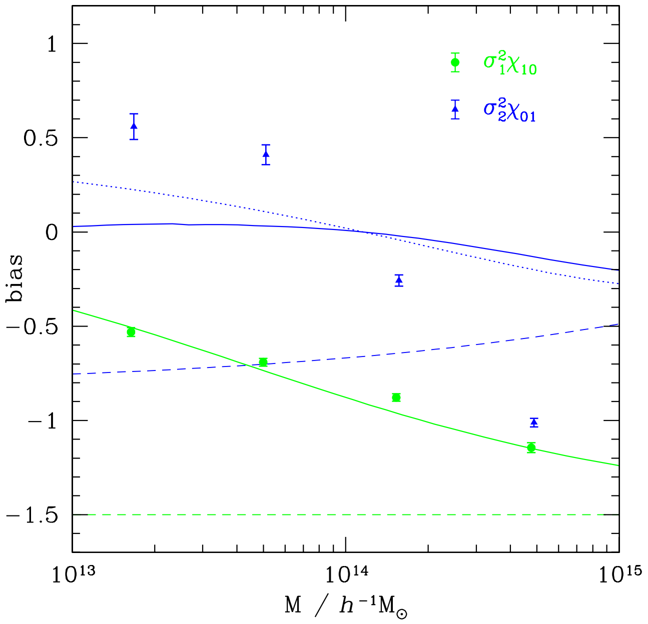

Figure 5. The bias factors $\sigma_1^2\chi_{10}$ and $\sigma_2^2\chi_{01}$ of dark matter haloes identified in the N-body simulations at $z = 0$ are shown as filled (green) circle and (blue) triangle, respectively. Error bars indicate the scatter among six realizations. The horizontal dashed (green) line at $-3/2$ and the dashed (blue) curve are the corresponding ESP predictions. The dotted (blue) curve is $\sigma_2^2\chi_{01}$ in a model where haloes are allowed to collapse in filamentary-like structures. The solid curves are our final predictions, which take into account the offset between peak position and protohalo centre-of-mass (see text for details).

The measured dimensionless bias factor $\sigma_1^2\chi_{10}$ is always negative, which indicates that haloes collapse out of regions which have values of η^2 smaller than average. In our ESP approach, we assume that the centre-of-mass position of protohaloes exactly coincides with that of a local density peak, so that $\sigma_1^2\chi_{10} \equiv -3/2$. However, simulations indicate that, while there is a good correspondence between protohaloes and linear density peaks, the centre-of-mass of the former is somewhat offset relative to the peak position (see e.g. Porciani, Dekel & Hoffman 2002; Ludlow & Porciani 2011). To model this effect, we note that, if the protohalo is at a distance R from a peak, then the average value of $3\eta^2$ is $\langle 3\eta^2\rangle(R) = \epsilon_1^2(R)(\langle 3\eta^2|\mathrm{pk}\rangle - 3)$ (in analogy with the fact that the average density at a distance R from a position where $\delta \equiv \delta_\mathrm{c}$ is $\langle\delta\rangle(R) = \xi_\delta(R)\,\delta_\mathrm{c}$). Assuming that the offset R follows a Gaussian distribution, the halo bias factor is

$$\sigma_{1\mathrm{s}}^2\chi_{10} = -\frac{3}{2}\sqrt{\frac{2}{\pi}}\int_0^\infty \frac{\mathrm{d}R}{\sigma}\left(\frac{R}{\sigma}\right)^2 \mathrm{e}^{-R^2/2\sigma^2}\epsilon_1^2(R). \tag{40}$$

The rms variance $\sigma(M)$ of the offset distribution, which generally depends on the halo mass, can be constrained from our measurements of χ_{10} for dark matter haloes. The best-fitting power-law function,

$$\sigma(M) = 2.50\left(\frac{M}{10^{13}\,\mathrm{M}_\odot\,h^{-1}}\right)^{0.063} h^{-1}\,\mathrm{Mpc}, \tag{41}$$

turns out to be a weak function of halo mass. In unit of the (tophat) Lagrangian halo radius, this translates into $\sigma/R_\mathrm{T} \approx 0.79$ and $\approx$0.36

for a halo mass $M = 10^{13}$ and $10^{14}\,\mathrm{M}_{\odot}\,h^{-1}$, respectively. The resulting theoretical prediction is shown as the solid curve in Fig. 5 and agrees reasonably well with our data. This crude approximation demonstrates that an offset between the protohalo centre-of-mass and the peak position can have a large impact on the inferred value of χ_{10}, since the latter is very sensitive to small-scale mass distribution.

Likewise, an offset between the protohalo centre-of-mass and the position of the linear density peak will also impact the measurement of χ_{01}, yet cannot explain the observed sign reversal. In this regard, one should first remember that density peaks become increasingly spherical as $\nu \to \infty$. Nevertheless, while their mean ellipticity $\langle e\rangle$ and prolateness $\langle p\rangle$ converge towards zero in this limit, $\langle v\rangle = \langle ue\rangle$ approaches 1/5 at fixed u (see equation 7.7 of Bardeen et al. 1986). Hence, $\langle\zeta^2\rangle = \langle 3v^2 + w^2\rangle$ does not tend towards zero but rather unity, like for random field points. Consequently, $\sigma_2^2\chi_{01} \to 0$ in the limit $\nu \to \infty$. Secondly, at any finite ν, our ESP approach predicts that χ_{01} be negative because we have assumed that protohaloes only form near a density peak ($\lambda_3 > 0$, where $\lambda_1 \geq \lambda_2 \geq \lambda_3$ are the eigenvalues of $-\partial_i\partial_j\delta$). However, *N*-body simulations strongly suggest that a fraction of the protohaloes collapse along the ridges or filaments connecting two density maxima, and that this fraction increases with decreasing halo mass (Ludlow & Porciani 2011). To qualitatively assess the impact of such primeval configurations on χ_{01}, we extend the integration domain in the plane (v, w) to include all the points with $\lambda_2 > 0$ and $\lambda_3 < 0$ (but still require that the curvature u be positive). This way we not only consider density peaks, but also extrema that correspond to filamentary configurations. The resulting curvature function $f(u, \alpha)$ can be cast into the compact form

$$f(u,\alpha) = \frac{1}{\alpha^4}\left\{\frac{\mathrm{e}^{-\frac{5\alpha u^2}{2}}}{\sqrt{10\pi}}\left(\alpha u^2 - \frac{16}{5}\right) + \frac{\mathrm{e}^{-\frac{5\alpha u^2}{8}}}{\sqrt{10\pi}}\left(31\alpha u^2 + \frac{32}{5}\right) + \frac{\sqrt{\alpha}}{2}\left(\alpha u^3 - 3u\right) \times \left[\mathrm{erf}\left(\sqrt{\frac{5\alpha}{2}}\,\frac{u}{2}\right) + \mathrm{erf}\left(\sqrt{\frac{5\alpha}{2}}\,u\right) - 1\right]\right\}. \quad (42)$$

The dotted curve in Fig. 5 shows $\sigma_2^2\chi_{01}$ when the filamentary configurations are included. While it agrees with the original ESP prediction at large halo mass, it reverses sign around $10^{14}\,\mathrm{M}_{\odot}\,h^{-1}$ because, as the peak height decreases, configurations with $\lambda_3 < 0$ or, equivalently, large values of ζ^2 become more probable. The solid curve takes into account, in addition to filamentary configurations, an offset between the protohalo and the peak position according to the simple prescription discussed above. This is our final prediction for $\sigma_2^2\chi_{01}$. It is clearly at odds with the measurements, which strongly suggest that $\sigma_2^2\chi_{01}$ can be very different from zero for $M \gtrsim 10^{13}\,\mathrm{M}_{\odot}\,h^{-1}$.

It is beyond the scope of this paper to work out a detailed description of the measurements. Using a value of R_{G} different than that obtained through the condition $\langle\delta_{\mathrm{G}}|\delta_{\mathrm{T}}\rangle = \delta_{\mathrm{T}}$ has a large impact on the mass function, suggesting that it will be difficult to get a good fit of both the mass function and the bias factors χ_{10} and χ_{01}. Before concluding however, we note that, if the Lagrangian clustering of haloes also depends on $s_2(\boldsymbol{x}) = s_{ij}(\boldsymbol{x})s^{ij}(\boldsymbol{x})$, where (in suitable units)

$$s_{ij}(\boldsymbol{x}) = \partial_i\partial_j\phi(\boldsymbol{x}) - \frac{1}{3}\delta_{ij}\delta(\boldsymbol{x}), \quad (43)$$

then we are not measuring χ_{01} but some weighted and scale-dependent combination of both χ_{01} and the Lagrangian bias γ_2 associated with $s_2(\boldsymbol{x})$. Recent numerical work indeed suggests that γ_2 might be non-zero for massive haloes (Baldauf et al. 2012; Chan, Scoccimarro & Sheth 2012; Sheth, Chan & Scoccimarro 2013). In this regards, our approach can furnish a useful cross-check of these results since it can provide a measurement of γ_2 which is independent of the bispectrum.

5 CONCLUSION

Dark matter haloes and galaxies are inherently biased relative to the mass density field, and this bias can manifest itself not only in *n*-point statistics such as the power spectrum or bispectrum, but also in simpler one-point statistics. In this work, we took advantage of this to ascertain the importance of certain non-local Lagrangian bias factors independently of a two-point measurement. We extended the cross-correlation technique of Musso et al. (2012) to χ^2-distributed variables, focusing on the quadratic terms $\eta^2(\boldsymbol{x})$ and $\zeta^2(\boldsymbol{x})$ (see equation 20) which arise from the peak constraint and for which we have theoretical predictions. In principle, however, our approach could be applied to measure the Lagrangian bias factor associated with any χ^2-distributed variable such as the tidal shear for instance. We validated our method with peaks of Gaussian random field before applying it to a catalogue of dark matter haloes with mass $M > 10^{13}\,\mathrm{M}_{\odot}\,h^{-1}$. Including an offset between the protohalo centre-of-mass and the peak position in the modelling (motivated by the analysis of Ludlow & Porciani 2011), we were able to reproduce our measurements of the non-local bias $\sigma_1^2\chi_{10}$. Our result $\chi_{10} < 0$ is consistent with the findings of Ludlow & Porciani (2011), who demonstrated that protohaloes with $M > 10^{13}\,\mathrm{M}_{\odot}\,h^{-1}$ preferentially form near initial density peaks ($\chi_{10} \equiv 0$ for a random distribution). However, we were unable to explain the measurements of $\sigma_2^2\chi_{01}$, even with the additional assumption that a fraction of the haloes collapse from filamentary-like structures rather than density peaks. We speculate that a dependence of the halo Lagrangian bias on $s_2(\boldsymbol{x})$ might be needed to explain this discrepancy.

The dependence on $\eta^2(\boldsymbol{x})$ induces a correction $-2\chi_{10}(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)$ to the halo bias which, for collinear wavevectors $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ of wavenumber $0.1\,h^{-1}$ Mpc, is $\Delta b \approx 0.02$ (0.05) and $\approx$0.30 (0.88) for haloes of mass $M = 10^{13}$ and $10^{14}\,\mathrm{M}_{\odot}\,h^{-1}$ at redshift $z = 0$ (1), respectively. Relative to the evolved, linear halo bias $b_1^{\mathrm{E}} \equiv 1 + b_{100}$, the fractional correction is $\Delta b/b_1^{\mathrm{E}} \sim$2 per cent and $\sim$15 per cent for the same low and high halo mass in the redshift range $0 < z < 1$. Hence, this correction can safely be ignored for $M = 10^{13}\,h^{-1}$ Mpc, but it could become relevant at larger halo masses.

We also refined the ESP approach of PSD so that clustering statistics can be straightforwardly computed from the (effective) bias expansion equation (19) (following the prescription detailed in Desjacques 2013). We checked that the predicted halo mass function, from which all the bias factors can be derived, agrees well with the numerical data. However, some of the model ingredients, especially the filtering of the density field, will have to be better understood if one wants to make predictions that are also accurate at small scales.

ACKNOWLEDGEMENTS

VD would like to thank the Perimeter Institute for Theoretical Physics and CCPP at New York University for their hospitality while some of this work was being completed there. MB, KCC and VD acknowledge support by the Swiss National Science Foundation.

REFERENCES

Baldauf T., Seljak U., Desjacques V., McDonald P., 2012, Phys. Rev. D, 86, 083540

Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15

Barkana R., 2004, MNRAS, 347, 59

Bernardeau F., Colombi S., Gaztañaga E., Scoccimarro R., 2002, Phys. Rep., 367, 1

Bond J. R., 1989, in Astbury A., Campbell B., Israel W., Kamal A., Khanna F., eds, Frontiers in Physics: From Colliders to Cosmology. World Scientific Press, Singapore, p. 182

Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440

Catelan P., Matarrese S., Porciani C., 1998, ApJ, 502, L1

Chan K. C., Scoccimarro R., Sheth R. K., 2012, Phys. Rev. D, 85, 083509

Crocce M., Pueblas S., Scoccimarro R., 2006, MNRAS, 373, 369

Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12

Desjacques V., 2013, Phys. Rev. D, 87, 043505

Desjacques V., Crocce M., Scoccimarro R., Sheth R. K., 2010, Phys. Rev. D, 82, 103529

Desjacques V., Gong J.-O., Riotto A., 2013, J. Cosmol. Astropart. Phys., 9, 6

Elia A., Ludlow A. D., Porciani C., 2012, MNRAS, 421, 3472

Ferraro S., Smith K. M., Green D., Baumann D., 2013, MNRAS, 435, 934

Fry J. N., Gaztanaga E., 1993, ApJ, 413, 447

Gradshteyn I. S., Ryzhik I. M., 1994, in Jeffrey A., ed., Table of Integrals, Series and Products, 5th edn. Academic Press, New York, p. 849

Gunst R. F., Webster J. T., 1973, J. Stat. Comput. Simul., 2, 275

Hahn O., Paranjape A., 2014, MNRAS, 438, 878

Kaiser N., 1984, ApJ, 284, L9

Knollmann S. R., Knebe A., 2009, ApJS, 182, 608

Komatsu E. et al., 2011, ApJS, 192, 18

Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473

Ludlow A. D., Porciani C., 2011, MNRAS, 413, 1961

McDonald P., Roy A., 2009, J. Cosmol. Astropart. Phys., 8, 20

Maggiore M., Riotto A., 2010, ApJ, 717, 515

Matsubara T., 2012, Phys. Rev. D, 86, 063518

Musso M., Sheth R. K., 2012, MNRAS, 423, L102

Musso M., Paranjape A., Sheth R. K., 2012, MNRAS, 427, 3145

Paranjape A., Sheth R. K., 2012, MNRAS, 426, 2789

Paranjape A., Sheth R. K., Desjacques V., 2013a, MNRAS, 431, 1503 (PSD)

Paranjape A., Sefusatti E., Chan K. C., Desjacques V., Monaco P., Sheth R. K., 2013b, MNRAS, 436, 449

Porciani C., Dekel A., Hoffman Y., 2002, MNRAS, 332, 339

Robertson B. E., Kravtsov A. V., Tinker J., Zentner A. R., 2009, ApJ, 696, 636

Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1

Sheth R. K., Chan K. C., Scoccimarro R., 2013, Phys. Rev. D, 87, 083002

Springel V., 2005, MNRAS, 364, 1105

Tiku M. L., 1965, Biometrika, 52, 415

Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709

Valageas P., 2009, A&A, 508, 93

APPENDIX A: THE CURVATURE FUNCTION OF DENSITY PEAKS

The curvature function of density peaks is (Bardeen et al. 1986)

$$\begin{aligned} f(u,\alpha) = \frac{1}{\alpha^4}\Bigg\{ & \frac{\mathrm{e}^{-\frac{5\alpha u^2}{2}}}{\sqrt{10\pi}}\left(\alpha u^2 - \frac{16}{5}\right) \\ & + \frac{\mathrm{e}^{-\frac{5\alpha u^2}{8}}}{\sqrt{10\pi}}\left(\frac{31}{2}\alpha u^2 + \frac{16}{5}\right) + \frac{\sqrt{\alpha}}{2}\left(\alpha u^3 - 3u\right) \\ & \times \left[\mathrm{erf}\left(\sqrt{\frac{5\alpha}{2}}\frac{u}{2}\right) + \mathrm{erf}\left(\sqrt{\frac{5\alpha}{2}}u\right)\right]\Bigg\}. \end{aligned} \tag{A1}$$

Note that Desjacques et al. (2010) introduced the extra variable α in order to get a closed form expression for their two-point peak correlation, while Desjacques (2013) showed that $\alpha \neq 1$ can be interpreted as a long-wavelength perturbation in $\zeta^2(\boldsymbol{x})$.

APPENDIX B: BIVARIATE χ^2 DISTRIBUTIONS

We take the following expression for the bivariate χ^2-distribution (Gunst & Webster 1973)

$$\begin{aligned} \chi^2_k(x,y;\epsilon) = & \frac{(xy)^{k/2-1}}{2^k\Gamma^2(k/2)}\left(1-\epsilon^2\right)^{-k/2}\mathrm{e}^{-\frac{x+y}{2(1-\epsilon^2)}} \\ & \times {}_0F_1\left(\frac{k}{2};\frac{\epsilon^2 xy}{4(1-\epsilon^2)^2}\right), \end{aligned} \tag{B1}$$

where x and y are distributed as χ^2-variables with k d.o.f., $\epsilon^2 \leq 1$ is their correlation and ${}_0F_1$ is a confluent hypergeometric function.

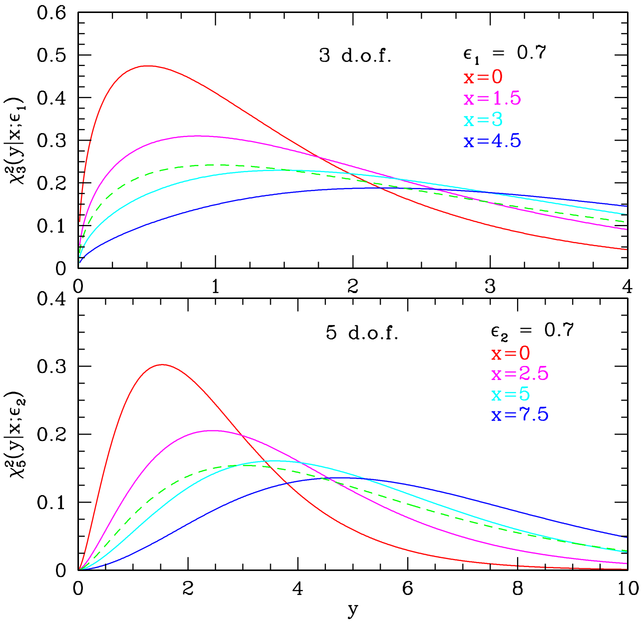

Figure A1. Conditional chi-squared distribution $\chi^2_k(y|x;\epsilon)$ for 3 and 5 d.o.f. Results are shown for several values of x and a fixed cross-correlation coefficient $\epsilon = 0.7$. The dashed (green) curve represents the unconditional distribution $\chi^2_k(y)$.

On using the fact that modified Bessel functions of the first kind can be written as $I_\alpha(x) = \mathrm{i}^{-\alpha} J_\alpha(\mathrm{i}x)$, where

$$J_\alpha(x) = \frac{(x/2)^\alpha}{\Gamma(\alpha+1)}\,{}_0F_1\left(\alpha+1; -\frac{x^2}{4}\right), \tag{B2}$$

the bivariate χ^2-distribution can be reorganized into the product

$$\chi^2_k(x, y; \epsilon) = \chi^2_k(x)\chi^2_k(y|x; \epsilon), \tag{B3}$$

where

$$\chi^2_k(y|x;\epsilon) = \frac{\mathrm{e}^{-\frac{y+\epsilon^2 x}{2(1-\epsilon^2)}}}{2(1-\epsilon^2)}\left(\frac{y}{\epsilon^2 x}\right)^{\alpha/2} I_\alpha\left(\frac{\epsilon\sqrt{xy}}{1-\epsilon^2}\right), \tag{B4}$$

and $\alpha = k/2 - 1$. This conditional distribution takes a form similar to that of a non-central χ^2-distribution $\chi^{2'}_k(x;\lambda)$, where λ is the non-centrality parameter. Fig. A1 displays $\chi^2_k(y|x;\epsilon)$ for several values of x, assuming $k = 3$ and 5. Note that $\chi^2_k(y|x = k;\epsilon)$ is different from $\chi^2_k(y)$.

Using the series expansion of $\chi^{2'}_k(x;\lambda)$ in terms of Laguerre polynomials (Tiku 1965), we arrive at

$$\chi^2_k(y|x;\epsilon) = \frac{\mathrm{e}^{-\frac{y}{2(1-\epsilon^2)}}}{2(1-\epsilon^2)^{\alpha+1}}\left(\frac{y}{2}\right)^\alpha \times \sum_{j=0}^{\infty} \frac{\left(\frac{-\epsilon^2}{1-\epsilon^2}\right)^j}{\Gamma\left(\frac{1}{2}k+j\right)}\left(\frac{x}{2}\right)^j L^{(\alpha)}_j\left[\frac{y}{2(1-\epsilon^2)}\right]. \tag{B5}$$

This series expansion is used to obtain the right-hand side of equation (26).