Probabilistic Lagrangian bias estimators and the cumulant bias expansion

Jens Stücker[1], Marcos Pellejero-Ibáñez[2], Raul E. Angulo[1,3], Francisco Maion[1], and Rodrigo Voivodic[1]

[1] Donostia International Physics Center (DIPC), Paseo Manuel de Lardizabal 4, 20018 Donostia-San Sebastian, Spain
e-mail: jstuecker@dipc.org
[2] Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh, EH9 3HJ , UK
e-mail: mpelleje@ed.ac.uk
[3] IKERBASQUE, Basque Foundation for Science, E-48013, Bilbao, Spain

May 2024

ABSTRACT

The spatial distribution of galaxies is a highly complex phenomenon currently impossible to predict deterministically. However, by using a statistical *bias* relation, it becomes possible to robustly model the average abundance of galaxies as a function of the underlying matter density field. Understanding the properties and parametric description of the bias relation is key to extract cosmological information from future galaxy surveys. Here, we contribute to this topic primarily in two ways: (1) We develop a new set of "probabilistic" estimators for bias parameters using the moments of the Lagrangian galaxy environment distribution. These estimators include spatial corrections at different orders to measure bias parameters independently of the damping scale. We report robust measurements of a variety of bias parameters for haloes, including the tidal bias and its dependence with spin at a fixed mass. (2) We propose an alternative formulation of the bias expansion in terms of "cumulant bias parameters" that describe the response of the logarithmic galaxy density to large-scale perturbations. We find that cumulant biases of haloes are consistent with zero at orders $n > 2$. This suggests that: (i) previously reported bias relations at order $n > 2$ are an artefact of the entangled basis of the canonical bias expansion; (ii) the convergence of the bias expansion may be improved by phrasing it in terms of cumulants; (iii) the bias function is very well approximated by a Gaussian – an avenue which we explore in a companion paper.

Key words. Cosmology: theory – large-scale structure of Universe – Methods: analytical

1. Introduction

Observations of the large-scale distribution of galaxies are among the most promising probes to accurately infer the cosmological parameters of our Universe. Past large-scale structure surveys have helped shaping our current model of the Universe by getting precise measurements of the angular diameter distance and growth rate at different epochs (see e.g. Alam et al. 2021). Forthcoming surveys like the Dark Energy Instrument (DESI, Levi et al. 2013), or Euclid (Amendola et al. 2018) will measure sky positions and redshifts (both photometric and spectroscopic) of an unprecedented amount of galaxies. To reliably interpret these datasets it is of crucial importance to accurately model the spatial distributions of galaxies as a function of cosmology.

The evolution of the large-scale distribution of matter is predominantly driven by gravity and can be predicted very reliably through perturbation theory (see Bernardeau et al. 2002, for a review) or by gravity-only N-body simulations (see e.g. Angulo & Hahn 2022, for a review). However, observed galaxies trace the matter distribution only in a *biased* way (Kaiser 1984; Mo & White 1996). Optimally exploiting the information from galaxy surveys therefore requires not only accurately modelling gravity, but also to account for the formation and evolution of galaxies.

There are a variety of methodologies to model the formation of galaxies. The most detailed approach involves employing hydrodynamical simulations to explicitly follow gas dynamics and the formation and evolution of stars, black holes and galaxies

(see e.g. Vogelsberger et al. 2014; Dubois et al. 2014; Schaye et al. 2015; Davé et al. 2019). In principle, such simulations would be the ideal method to account for the formation of galaxies if the underlying physics was modelled reliably and at an affordable cost. However, in practice these simulations are limited in volume and number due to their substantial computational requirements. Additionally, they necessitate the assumption of sub-grid physics e.g. to model the unresolved formation of stars, the growth of active galactic nuclei and their feedback processes. The results of such simulations depend strongly on the associated parameters and are therefore severely limited in their predictive ability (e.g. Genel et al. 2019; Villaescusa-Navarro et al. 2021).

To overcome the limitations of hydrodynamic simulations, a more agnostic approach can be pursued. Semi-analytic models, for example, follow a number of empirical and semi-empirical relations that provide more versatility to the modelling of galaxy formation physics (see e.g. Kauffmann et al. 1999; Henriques et al. 2015; Stevens et al. 2016; Lacey et al. 2016; Croton et al. 2016; Lagos et al. 2018). An even more agnostic strategy involves techniques such as Sub-Halo Abundance Matching (see e.g. Conroy et al. 2006; Reddick et al. 2013; Chaves-Montero et al. 2016; Lehmann et al. 2017; Dragomir et al. 2018; Contreras et al. 2021; Ortega-Martinez et al. 2024) or Halo Occupation Distribution (see e.g.Peacock & Smith 2000; Berlind & Weinberg 2002; Zheng et al. 2005; Cacciato et al. 2012; Salcedo et al. 2022a), which rely solely on gravity simulations and fundamental assumptions about populating collapsed structures

(haloes and sub-haloes) based on their mass. These assumptions must be sufficiently flexible to accommodate various galaxy formation scenarios. Nonetheless, these techniques also have computational constraints, and the principles they rely on may oversimplify reality, necessitating extensions to account for dependencies beyond merely the mass of the collapsed structure (a phenomenon known as assembly bias, Gao et al. 2005; Wechsler et al. 2006; Gao & White 2007; Croton et al. 2007; Dalal et al. 2008; Faltenbacher & White 2010; Montero-Dorta et al. 2017; Zehavi et al. 2018; Ferreras et al. 2019; Sato-Polito et al. 2019; Tucci et al. 2021; Salcedo et al. 2022b; Chaves-Montero et al. 2023, among others).

The most general approach currently known is a perturbative expansion of galaxy bias (see Desjacques et al. 2018 for a review). In this framework, the evolution of matter is assumed to be modelled accurately through purely gravitational effects whereas the galaxy field selectively populates the matter field. The galaxy field is then written in a perturbative manner as a function of the properties of the underlying matter distribution. In Eulerian bias schemes, the galaxy number density is expanded in terms of the final properties of the density field whereas in Lagrangian schemes this is done in terms of the initial properties of the linear density field. This approach offers great flexibility, allowing a single model to describe biased tracers with vastly different properties. Bias approaches based on effective field theory aim to describe the clustering behavior of galaxies down to $k \approx 0.2h/\mathrm{Mpc}$ (Baumann et al. 2012; Baldauf et al. 2016; Vlah et al. 2016). They have been used extensively to extract robust cosmological constraints from surveys (Ivanov et al. 2020; d'Amico et al. 2020; Colas et al. 2020; Nishimichi et al. 2020; Chen et al. 2020; Philcox & Ivanov 2022).

There exist different options for defining the biasing scheme. Traditionally, perturbation theory is used to describe the underlying gravitational evolution of the density field whereas in the recently developed "hybrid approaches" the gravitational evolution is treated exactly through N-body simulations (Modi et al. 2020; Kokron et al. 2021; Zennaro et al. 2023; Pellejero Ibañez et al. 2022, 2023; DeRose et al. 2023). However, in all bias methods the galaxy number density is expanded in a perturbative series with a set of free coefficients known as "bias parameters". These parameters can be interpreted as the response of the galaxy number density to perturbations of the density field. While these parameters exhibit some physical insights (Lazeyras et al. 2016; Lazeyras & Schmidt 2019; Barreira et al. 2021), they are generally treated as nuisance parameters which are marginalized over to extract the cosmological information of interest.

A quantitative understanding of the bias parameters is important for analysing future galaxy surveys. On the one hand, this is necessary to determine how well the bias expansion converges. Whether the truncation at a given order gives accurate results, depends on higher order terms being sufficiently small that they can be neglected. On the other hand, it is of significant interest to limit the prior volume that is used when fitting to cosmological surveys to maximize the extracted cosmological information. In particular, it may significantly benefit an analysis to fix bias parameters to so called "coevolution relations" that relate the parameters to each other. Therefore, significant effort has been put to measuring bias parameters and to constrain coevolution relations in simulations.

One of the most accurate methods for measuring bias parameters is through separate universe simulations (Li et al. 2014; Wagner et al. 2015). In these, it is directly tested how the number of tracers responds to an increase in the large-scale density. This technique has been used to constraint density bias parame-

ters of haloes to very high accuracy (Lazeyras et al. 2016). Further, similar ideas have been used to measure the response to non Gaussian perturbations (Barreira 2020) and to the changes in the Laplacian of the density field (Lazeyras & Schmidt 2018). Other bias parameters have been constrained through different methods. For example, the tidal bias has been measured through Fourier space correlations (Modi et al. 2017; Lazeyras & Schmidt 2018) or all bias parameters can be constrained simultaneously through forward modelling the power spectrum (e.g. Zennaro et al. 2022) or the field level galaxy distribution (e.g. Lazeyras et al. 2021).

In this paper, we propose a novel approach for accurately measuring Lagrangian bias parameters from simulations. In this approach we consider the distribution of the linear density field at the initial (Lagrangian) locations of galaxies – which we call the "galaxy environment distribution". We adopt a probabilistic approach to model this galaxy environment distribution and show that large-scale bias parameters have a simple relation to the moments of this distribution. We use this to derive estimators of the bias parameters that can be extended with simple corrections to practically eliminate their scale dependence. We derive such estimators for a large variety of bias parameters and provide corresponding measurements for the case of haloes. Our method is operationally similar to "Lagrangian cross correlation" methods (Musso et al. 2012; Paranjape et al. 2013b,a; Biagetti et al. 2014). However, while these rely on peak theory to correct scale dependencies, we provide a model independent approach to obtain scale independent bias parameters.

Further, we propose a new set of "cumulant bias parameters" which are defined as the response of the *logarithm* of the galaxy number density to perturbations in the linear field. We show that these parameters have significantly improved properties when compared to their canonical counterparts. For the case of haloes we find that cumulant biases of order three and higher are consistent with zero. Therefore, the canonical coevolution relations for haloes at order order 3 and beyond appear primarily as an effect of a sub-optimal parameterization. We suggest that rephrasing the bias expansion in terms of the cumulant bias parameters can significantly enhance its convergence.

The vanishing of cumulants for haloes at order three and beyond implies that the bias function can be well approximated through a Gaussian. We will explore such a Gaussian bias model in a companion paper Stücker et al. (in prep).

This article is organized as follows: In Section 2 we explain the probabilistic approach to measure Lagrangian bias parameters of scalar variables like the density and the Laplacian and we introduce the concept of the cumulant bias expansion. In Section 3 we present measurements of the corresponding parameters and demonstrate how cumulant bias parameters exhibit several practical advantages. In Section 4 we show how to generalize the concept of the probabilistic estimators for tensorial bias variables, like for example the tidal bias b_{K^2}. In Section 5 we provide measurements of a few select tensorial quantities. In Section 6 we briefly discuss the quantitative importance of different bias terms. Finally, in Section 7 we summarize the benefits of our novel estimators and we discuss under which circumstances it is advantageous to phrase the bias expansion in terms of the cumulant bias parameters.

2. Theory

In this section, we introduce the necessary theory and (1) relate bias parameters to properties of the galaxy environment distribution, (2) express the moment and cumulant generating functions

of the galaxy environment distribution in terms of the large-scale bias function, (3) introduce the concept of a cumulant bias expansion, (4) show how to derive estimators for both canonical and cumulant bias parameters.

2.1. Definitions

The considerations here are based on the idea of the peak-background split (PBS) which states that *a long-wavelength density perturbation acts like a local modification of the background density for the purposes of the formation of halos and galaxies* (Kaiser 1984; Bardeen et al. 1986; Desjacques et al. 2018). Bias parameters describe the response of the galaxy number to such long-wavelength perturbations.

An exact implementation of the PBS is given by the separate universe approach. In this approach one considers some universe in which a measurable number of galaxies $n_{g,0}$ forms. If one were to increase the background density of the universe by a relative amount δ_0 (e.g. in a separate universe simulation, Frenk et al. 1988; Li et al. 2014; Wagner et al. 2015), then in the new universe a different number of galaxies $n_g(\delta_0)$ forms. We call their ratio

$$F(\delta_0) = \frac{n_g(\delta_0)}{n_{g,0}} = 1 + b_1\delta_0 + \frac{1}{2}b_2\delta_0^2 + \cdots + \frac{1}{n!}b_n\delta_0^n + \cdots \quad (1)$$

the "true bias function" or the "separate universe bias function". This function can directly be measured with separate universe simulations (e.g. Lazeyras et al. 2016) and we will refer to the coefficients of the indicated expansion as the "canonical bias parameters" or just "the bias parameters":

$$b_n = \left.\frac{\partial^n F(\delta_0)}{\partial \delta_0^n}\right|_{\delta_0=0} \quad (2)$$

Therefore, the bias parameters physically describe the response of the galaxy density to small perturbations at infinitely large scales. In this article we want to investigate galaxy bias from a probabilistic perspective.

We consider an infinitesimally small Lagrangian volume element which nothing is known about except the linear density contrast δ smoothed at some scale (and possibly other features of the linear field like the Laplacian L or the tidal field). Neglecting primordial non-Gaussianity, the density contrast follows a Gaussian distribution

$$p(\delta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\delta^2}{2\sigma^2}\right) . \quad (3)$$

For simplicity, we will assume throughout this article that the smoothed density contrast is defined with a sharp k-space filter. Most considerations here translate in a simple manner to differently filtered cases, but some additional care must be taken due to the more complicated correlation between large and small scales. This is discussed in more detail in Appendix A.

We call the average probability that a galaxy forms in such a volume element $p(g)$ and we call the conditional probability, given the knowledge of the linear density contrast, $p(g|\delta)$. The excess probability

$$f(\delta) := \frac{p(g|\delta)}{p(g)} \quad (4)$$

is parameterized through a function $f(\delta)$ which we will refer to as the "scale dependent bias function" or just the "bias function" throughout this article. The bias function depends in a predictable manner on the variance of δ at the considered scale, as we will show later.

Since densities at different scales add up linearly, a separate-universe style modification of the large-scale density contrast from 0 to δ_0 will immediately translate to a modification of the linear density in our volume element $\delta \to \delta + \delta_0$. Therefore, F and f should be related through

$$F(\delta_0) = \langle f(\delta + \delta_0) \rangle \quad (5)$$

where the angled brackets indicate an expectation value taken over the Lagrangian volume (see also Desjacques et al. 2018). The relation indicates that in a separate universe experiment the number of galaxies should change according to the average change in probability of forming galaxies when changing the linear density contrast everywhere in space. Therefore, the canonical bias parameters are given in terms of the scale dependent bias function as

$$b_n = \left\langle \frac{\partial^n f}{\partial \delta^n} \right\rangle . \quad (6)$$

Note that the coefficients of an expansion of f around zero are scale dependent, but they can be matched to the canonical bias parameters with equation (6) to parameterize f with scale-independent coefficients – a procedure commonly referred to as "renormalization".

We will see later that equation (5) holds only approximately. This is so, since at any finite smoothing scale, the density contrast δ is correlated with other variables (e.g. the Laplacian) so that a change in the small scale density contrast at the location of all galaxies is not exactly equivalent to the "pure" density change in separate universe experiments. This introduces scale-dependencies that can be accounted for.

Finally, we introduce one further object which we call the "galaxy environment distribution" $p(\delta|g)$ or "halo environment distribution" when we talk specifically about haloes. This quantifies the probability of the linear density contrast in an infinitesimal volume element, given that there is a galaxy at the considered location. This function can easily be measured as a histogram of the linear density at the Lagrangian locations of galaxies as is illustrated in Figure 1 – where we have used a damping scale of $k_d = 0.15h\text{Mpc}^{-1}$ here[1], leading to $\sigma = 0.56$.

Through Bayes' theorem, the relation between the galaxy environment distribution and the bias function is given by

$$p(\delta|g) = \frac{p(\delta \cap g)}{p(g)} \quad (7)$$

$$= \frac{p(g|\delta)p(\delta)}{p(g)} \quad (8)$$

$$= f(\delta)p(\delta) . \quad (9)$$

This means for example, that the bias function $f(\delta)$ can be investigated in a non-parametric way by measuring $p(\delta|g)$ and dividing by the Gaussian background distribution $p(\delta)$ as we will investigate in the companion paper Stücker et al. (in prep).

In this section, we will use probability theory to investigate properties of the galaxy environment distribution. In particular, we will show that bias parameters are simply related to the moments of this distribution and that probability theory motivates the usage of better behaved "cumulant bias parameters".

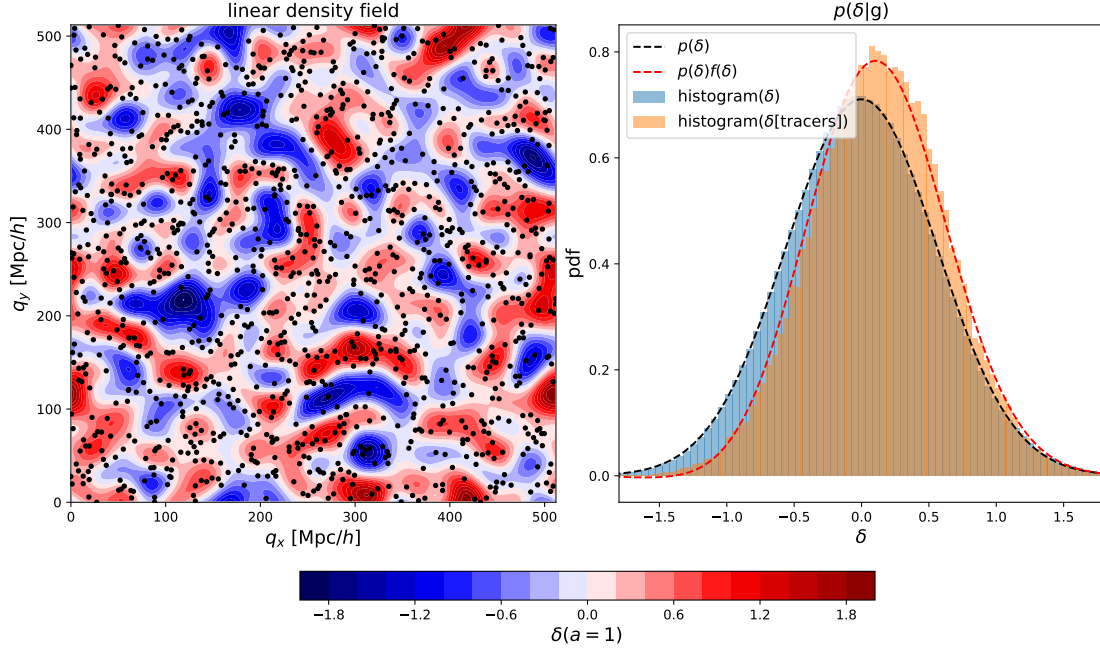[1] Corresponding approximately to a sharp truncation scale in Fourier space, as explained in Section 3.2.

Fig. 1. Illustration of the inference of the halo environment distribution $p(\delta|g)$. Left: we trace back galaxies to their origin in Lagrangian space (marked as black dots) and infer the (smoothed) linear density field δ at their Lagrangian locations. Right: The environment distribution (orange histogram) is given by the distribution of δ at the galaxy locations which is notably biased relative to the matter distribution $p(\delta)$ (blue histogram and a Gaussian represented as dashed line). The galaxy environment distribution is well approximated through $p(\delta)f(\delta)$ where here $f(\delta)$ is a quadratic polynomial bias function.

2.2. Bias estimators

Following up on equation (5) we can write

$$
\begin{aligned}
F(\delta_0) &= \langle f(\delta + \delta_0)\rangle \\
&= \int_{-\infty}^{\infty} p(\delta)f(\delta + \delta_0)\mathrm{d}\delta \\
&= \int_{-\infty}^{\infty} p(\delta - \delta_0)f(\delta)\mathrm{d}\delta \; .
\end{aligned}
\tag{10}
$$

Where in the last line we have made the substitution $\delta \to \delta + \delta_0$, and δ represents the linearly extrapolated overdensity, hence it can assume values between negative and positive infinity. The latter equation can be interpreted as an alternative perspective onto the separate universe experiment: When increasing the background density, the probability of environments is changed by a factor $p(\delta - \delta_0)/p(\delta)$ whereas the likelihood of forming a galaxy when presupposing a given environment stays constant.

Now, combining equations (2) and (10), we can evaluate the bias parameters as

$$
\begin{aligned}
b_n &= \frac{\partial^n}{\partial \delta_0^n} \int_{-\infty}^{\infty} p(\delta - \delta_0)f(\delta)\mathrm{d}\delta \bigg|_{\delta_0=0} \\
&= (-1)^n \int_{-\infty}^{\infty} p^{(n)}(\delta)f(\delta)\mathrm{d}\delta \\
&= (-1)^n \int_{-\infty}^{\infty} \frac{p^{(n)}(\delta)}{p(\delta)} p(\delta|g)\mathrm{d}\delta \\
&= (-1)^n \left\langle \frac{p^{(n)}(\delta)}{p(\delta)} \right\rangle_{\mathrm{g}}
\end{aligned}
\tag{11}
$$

where the angled brackets with a 'g' subscript indicate an expectation value evaluated over the locations of galaxies (rather than

all of Lagrangian space). Further, we can use that p is a Gaussian distribution for which derivatives are given by the (probabilisist's) Hermite polynomials:

$$
\frac{\partial^n}{\partial \delta^n} \exp\left(-\frac{\delta^2}{2\sigma^2}\right) = (-1)^n \sigma^{-n} \exp\left(-\frac{\delta^2}{2\sigma^2}\right) H_n\left(\delta/\sigma\right)
\tag{12}
$$

so that the bias estimators read

$$
b_n = \left\langle \frac{H_n\left(\delta/\sigma\right)}{\sigma^n} \right\rangle_{\mathrm{g}} \; .
\tag{13}
$$

This expression has already been used by Paranjape et al. (2013b,a) to measure bias parameters. It was motivated by its emergence in excursion set frameworks (Musso et al. 2012) and peak statistics (Paranjape & Sheth 2012), but it is clearly also valid outside such frameworks, requiring only to assume the PBS. On a related note, it is worth mentioning that it has already been proposed by Szalay (1988) to expand the field in terms of Hermite polynomials.

The estimators for the first four bias parameters read:

$$
b_1 = \left\langle \frac{\delta}{\sigma^2} \right\rangle_{\mathrm{g}}
\tag{14}
$$

$$
b_2 = \left\langle \frac{\delta^2 - \sigma^2}{\sigma^4} \right\rangle_{\mathrm{g}}
\tag{15}
$$

$$
b_3 = \left\langle \frac{\delta^3 - 3\delta\sigma^2}{\sigma^6} \right\rangle_{\mathrm{g}}
\tag{16}
$$

$$
b_4 = \left\langle \frac{\delta^4 - 6\delta^2\sigma^2 + 3\sigma^4}{\sigma^8} \right\rangle_{\mathrm{g}} \; .
\tag{17}
$$

2.3. The moment generating function

Equation (13) shows that there exists a simple relation between moments of the galaxy environment distribution and the bias pa-

rameters. We can show this in a very general manner. Expanding the Gaussian background distribution we obtain

$$p(\delta - \delta_0) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{\delta^2 - 2\delta\delta_0 + \delta_0^2}{2\sigma^2}\right)$$

$$= p(\delta) \exp\left(\frac{\delta\delta_0}{\sigma^2}\right) \exp\left(-\frac{\delta_0^2}{2\sigma^2}\right)$$

and inserting this into equation (10) yields

$$F(\delta_0) = \exp\left(-\frac{\delta_0^2}{2\sigma^2}\right) \int_{-\infty}^{\infty} f(\delta) p(\delta) \exp\left(\frac{\delta\delta_0}{\sigma^2}\right) d\delta$$

$$= \exp\left(-\frac{1}{2}t^2\sigma^2\right) \int_{-\infty}^{\infty} p(\delta|g) \exp(t\delta) d\delta$$

$$= \exp\left(-\frac{1}{2}t^2\sigma^2\right) \langle \exp(t\delta) \rangle_g \tag{18}$$

where we have labelled $t = \delta_0/\sigma^2$. The last term can be identified with the moment generating function

$$M_g(t) = \langle \exp(t\delta) \rangle_g \tag{19}$$

of the galaxy environment distribution. Therefore, the moment generating function of the galaxy environment distribution and the separate universe bias function can be directly converted into each other

$$M_g(t) = F(t\sigma^2) \exp\left(\frac{1}{2}t^2\sigma^2\right) \tag{20}$$

and the relation between the moments and the bias parameters as in equation (13) can equivalently be found by taking derivatives of the moment generating function

$$\mu_n := \langle \delta^n \rangle_g = \left.\frac{\partial}{\partial t^n} M_g(t)\right|_{t=0} . \tag{21}$$

It is worth noting that this result may be related to the considerations in White (1979) where for the case of discrete tracers, the moment generating function of the galaxy count frequency distribution is expressed through the void probability function.

2.4. Cumulant bias parameters

We may further ask, how the bias parameters relate to the cumulants of the galaxy environment distribution. In probability theory the value of cumulants are generally thought to characterize a distribution more independently than its moments. For example, if a distribution has a large first moment $\langle x \rangle$, then we should also expect that it has large second and third moments $\langle x^2 \rangle$ and $\langle x^3 \rangle$. However, the first cumulant of a distribution $\langle x \rangle$ says very little about the second and third cumulants, $\langle (x - \langle x \rangle)^2 \rangle$ and $\langle (x - \langle x \rangle)^3 \rangle$ respectively. For example, for a Gaussian distribution the mean is independent of its variance and the third and higher order cumulants are actually zero.

The cumulant generating function is defined as

$$K_g(t) = \log M_g(t)$$

$$= \frac{1}{2}t^2\sigma^2 + \log F(t\sigma^2) . \tag{22}$$

Cumulants of the galaxy environment distribution are simply given by the derivatives of the cumulant generating function

$$\kappa_n = \left.\frac{\partial K_g}{\partial t^n}\right|_{t=0} . \tag{23}$$

Therefore, they evaluate to

$$\kappa_n = \begin{cases} \beta_n \sigma^{2n} & \text{if } n \neq 2 \\ \beta_2 \sigma^4 + \sigma^2 & \text{if } n = 2 \end{cases} \tag{24}$$

where we have defined

$$\beta_n = \left.\frac{\partial^n}{\partial \delta_0^n} \log F(\delta_0)\right|_{\delta_0=0} . \tag{25}$$

Notice that whether β_2 is above or below zero is a direct indication of whether the variance of the halo environment distribution is larger or smaller than that of the background. We will refer to these parameters as "cumulant bias parameters" and they are directly related to the canonical bias parameters b_n. Comparing their definition to the canonical bias parameters in equation (2) shows that they relate to each other exactly in the same way that cumulants relate to moments (compare equations (21) and (23)). For example, at the first four orders we have

$$\beta_1 = b_1 \tag{26}$$
$$\beta_2 = b_2 - b_1^2 \tag{27}$$
$$\beta_3 = b_3 - 3b_1b_2 + 2b_1^3 \tag{28}$$
$$\beta_4 = b_4 - 4b_1b_3 - 3b_2^2 + 12b_1^2b_2 - 6b_1^4 . \tag{29}$$

2.5. Interpretation

Given a set of bias parameters, these relations allow us to directly find the cumulants of the galaxy environment distribution, or alternatively, they allow us to infer canonical bias parameters by measuring cumulants of the galaxy environment distribution.

However, we note that the cumulant bias parameters are, just as the canonical bias parameters, well defined bias parameters and we can consider to phrase the bias expansion directly in terms of the cumulant bias parameters.

$$\log F = \beta_1 \delta_0 + \frac{1}{2}\beta_2 \delta_0^2 + \frac{1}{6}\beta_3 \delta_0^3 + ... \tag{30}$$

There are several reasons to believe that they may form a better set of parameters than the canonical bias parameters:

– The cumulant bias parameters are the derivatives of the logarithm of the galaxy density. Therefore, they presuppose the positivity of the function and may be better behaved, especially in low density regions.
– The canonical bias parameters behave similarly to moments of the galaxy environment distribution. Therefore, one may expect that if b_1 is large automatically b_4 will also be large. On the other hand, we may expect β_4 to be rather independent of the value of β_1 – just as cumulants are relatively independent of each other.
– If the galaxy environment distribution has the form of a Gaussian then $\beta_n = 0$ for all $n > 2$. On the other hand all b_n would be non-zero in this case. Therefore, if all β_n beyond degree 2 are small this motivates the usage of a Gaussian bias model. We will show that this is indeed the case for haloes.

Further, we note that the cumulant bias parameters and the probabilistic considerations in this section motivate novel approaches to parameterize the bias function at finite smoothing scales. There are several possible well motivated approaches and we will just mention them here briefly, but we will leave a more thorough investigation to future studies.

The first and most straight-forward proposition is to assume an expansion of the logarithm of the bias function, leading to an exponential of a polynomial

$$\log f(\delta) = \gamma_0 + \gamma_1 \delta + \gamma_2 \delta^2 + \gamma_3 \delta^3 + ... \tag{31}$$

$$f(\delta) = \exp\left(\gamma_0 + \gamma_1 \delta + \gamma_2 \delta^2 + \gamma_3 \delta^3 + ...\right) \ . \tag{32}$$

Such a bias expansion would have to be truncated at an even order (if $n \geq 2$) and the highest order coefficient must be negative to guarantee a well normalized probability distribution. By definition this guarantees the positivity of the bias function which is arguably a desirable property – especially if one intended to create mocks from a bias model. Further, the resulting distribution would be part of the exponential family which guarantees several desirable properties. The main limitation is that at orders $n > 2$ it might be difficult to properly renormalize the expression analytically[2].

This choice is the only general choice we know of that guarantees $f(\delta) > 0$ for all values of δ even if the expansion is truncated. However, there are a few other noteworthy options that do not fulfill this constraint. One option is motivated by the intriguing simplicity of the relation between cumulant generating function (22) and galaxy number density. This motivates to define the bias function directly through the cumulant generating function

$$K_g(t) = K_m(t) + \sum_n \frac{1}{n!} \beta_n t \sigma^{2n} \tag{33}$$

where $K_m(t)$ is the cumulant generating function of the matter distribution. This general Ansatz might also be feasible in Eulerian space, but for the Lagrangian case it is simply $K_m(t) = \frac{1}{2} t \sigma^2$.

Equation (33) is already in renormalized form (i.e. phrased in terms of scale independent bias parameters). If the expansion is truncated at an order $n = 2$ then this corresponds to a well defined Gaussian bias model, but if it is truncated at a higher order then it is not easy to find an analytical form of $f(\delta)$. However, a numerical expression can be obtained through a Fourier transformation. Unfortunately, such a truncation at higher orders does not lead to a well defined probability density, but instead may have negative function values. In fact the cumulant generating function does not yield a positive pdf for any finite order polynomial of degree $n > 2$ (Lukacs 1970). However, this does not disfavour this approach over the canonical bias expansion which also yields negative pdf.

Further, we note that it is an option to phrase the bias expansion relative to a Gaussian distribution that has the correct cumulants up to order two:

$$p(\delta|g) = \frac{1}{\sqrt{2\pi}\sigma_g} \exp\left(-\frac{(\delta - \mu_g)^2}{2\sigma_g^2}\right)(1 + \gamma_3 \delta^3 + \gamma_4 \delta^4 + ...) \tag{34}$$

where $\mu_g = \kappa_1 = \beta_1 \sigma^2$ and $\sigma_g^2 = \kappa_2 = \beta_2 \sigma^4 + \sigma^2$ and where the γ_n start at order three, since the lowest two orders are already specified through the Gaussian. This type of expansion is similar to a truncated Edgeworth series (which is an expansion to approximate distributions with given cumulants).

Finally, we note that it is possible to use the cumulant bias parameters also in a classical polynomial bias expansion. For example a third order expansion that only uses two parameters, assuming $\beta_3 = 0$ would read

$$f(\delta) = \beta_1 \delta + (\beta_2 + \beta_1^2)(\delta^2 - \sigma^2) + (3\beta_1\beta_2 - \beta_1^3)(\delta^3 - 3\delta\sigma^2) + ... \tag{35}$$

which can be understood as lower order terms predicting the likely behavior of higher order terms.

2.6. Multivariate estimators

We have so far only considered the bias function under the assumption that the only known aspect of the environment is the density δ at the considered location. However, other aspects of the environment may be known, such as the Laplacian

$$L = \nabla^2 \delta \tag{36}$$

or higher order derivatives of the density field

$$P = \nabla^4 \delta \tag{37}$$

or the tidal field. Since the tidal field is inherently a tensorial quantity, its treatment is slightly more complicated and we will therefore discuss it in a dedicated manner in Section 4.

Assuming that we are dealing only with scalar quantities that describe the environment, we may summarize the environment through a single vector $\boldsymbol{x}$ – for example, if we consider the density and Laplacian as variables we have $\boldsymbol{x} = (\delta, L)^T$. The majority of the relations that we have derived in the previous section hold in a similar manner for the multivariate case as well. We will not derive all of these again, since their derivation follows in almost complete analogy, but we list all important relations.

Assuming that all considered environment variables follow from linear operators on the Gaussian random field, their distribution is given by a multivariate Gaussian

$$p(\boldsymbol{x}) = \frac{1}{2\pi \sqrt{\det(\mathsf{C})}} \exp\left(-\frac{1}{2}\boldsymbol{x}^T \mathsf{C}^{-1} \boldsymbol{x}\right) \ . \tag{38}$$

For example, for the case of $\boldsymbol{x} = (\delta, L)^T$ we have the covariance matrix

$$\mathsf{C}_{\delta,L} = \begin{bmatrix} \sigma_0^2 & -\sigma_1^2 \\ -\sigma_1^2 & \sigma_2^2 \end{bmatrix} \tag{39}$$

$$\mathsf{C}_{\delta,L}^{-1} = \begin{bmatrix} \frac{\sigma_2^2}{\sigma_0^2\sigma_2^2 - \sigma_1^4} & \frac{\sigma_1^2}{\sigma_0^2\sigma_2^2 - \sigma_1^4} \\ \frac{\sigma_1^2}{\sigma_0^2\sigma_2^2 - \sigma_1^4} & \frac{\sigma_0^2}{\sigma_0^2\sigma_2^2 - \sigma_1^4} \end{bmatrix} \tag{40}$$

where $\sigma_0^2 = \langle \delta^2 \rangle$, $\sigma_2^2 = \langle L^2 \rangle$ and $\sigma_1^2 = \langle -\delta L \rangle$.

Now, we can conveniently write bias parameters in a matrix form

$$\boldsymbol{b}_1 = \nabla_{\boldsymbol{x}} F |_{\boldsymbol{x}=0} \tag{41}$$

$$\mathsf{b}_2 = (\nabla_{\boldsymbol{x}} \otimes \nabla_{\boldsymbol{x}})F |_{\boldsymbol{x}=0} \tag{42}$$

$$\mathsf{b}_3 = (\nabla_{\boldsymbol{x}} \otimes \nabla_{\boldsymbol{x}} \otimes \nabla_{\boldsymbol{x}})F |_{\boldsymbol{x}=0} \tag{43}$$

$$\mathsf{b}_n = \nabla_{\boldsymbol{x}}^{\otimes n} F |_{\boldsymbol{x}=0} \tag{44}$$

where $\otimes$ designates an outer product, the power notation in the last line designates a repeated outer product, $\boldsymbol{b}_1$ is a vector, b_2 is a symmetric rank two matrix and b_3 a symmetric rank three tensor and so on and $\nabla_{\boldsymbol{x}}$ denotes a gradient with respect to the chosen variables, for example $\nabla_{\boldsymbol{x}} = (\partial/\partial\delta, \partial/\partial L)^T$.

---

[2] That means to write it in terms of scale independent bias parameters.

It is then straightforward to show

$$b_n = (-1)^n \left\langle \frac{\nabla_x^{\otimes n} p}{p} \right\rangle_g \tag{45}$$

which evaluates at the first two orders to

$$\boldsymbol{b}_1 = \left\langle \mathsf{C}^{-1} \boldsymbol{x} \right\rangle_g \tag{46}$$

$$\mathsf{b}_2 = \left\langle \mathsf{C}^{-1} (\boldsymbol{x} \otimes \boldsymbol{x} - \mathsf{C}) \mathsf{C}^{-1} \right\rangle_g \tag{47}$$

Importantly, these expressions do not lead to the same bias estimators for the density if a second variable is present that is correlated with the density. For example, if we use the density and Laplacian as variables, we find

$$b_\delta = \left\langle \frac{\delta \sigma_2^2 + L \sigma_1^2}{\sigma_0^2 \sigma_2^2 - \sigma_1^4} \right\rangle_g \tag{48}$$

$$b_L = \left\langle \frac{L \sigma_0^2 + \delta \sigma_1^2}{\sigma_0^2 \sigma_2^2 - \sigma_1^4} \right\rangle_g \tag{49}$$

for the components of $\boldsymbol{b}_1$ and

$$b_{\delta\delta} = \left\langle \frac{\left(\delta \sigma_2^2 + L \sigma_1^2\right)^2 - \sigma_2^2 \left(\sigma_0^2 \sigma_2^2 - \sigma_1^4\right)}{\left(\sigma_0^2 \sigma_2^2 - \sigma_1^4\right)^2} \right\rangle_g \tag{50}$$

$$b_{LL} = \left\langle \frac{\sigma_0^2 \left(-\sigma_0^2 \sigma_2^2 + \sigma_1^4\right) + \left(L \sigma_0^2 + \delta \sigma_1^2\right)^2}{\left(\sigma_0^2 \sigma_2^2 - \sigma_1^4\right)^2} \right\rangle_g \tag{51}$$

$$b_{\delta L} = \left\langle \frac{\left(L \sigma_0^2 + \delta \sigma_1^2\right)\left(L \sigma_1^2 + \delta \sigma_2^2\right) - \sigma_1^2 \left(\sigma_0^2 \sigma_2^2 - \sigma_1^4\right)}{\left(\sigma_0^2 \sigma_2^2 - \sigma_1^4\right)^2} \right\rangle_g \tag{52}$$

for the components of b_2 where we have written e.g. $b_{\delta\delta}$ as a symbol for what was previously referred to as b_2. Further, we note that a general form for the Nth density bias parameter is given by

$$b_{\delta^N} = \left\langle \frac{H_N((\delta + L \sigma_1^2 / \sigma_2^2)/\sigma_*)}{\sigma_*^N} \right\rangle_g \tag{53}$$

with $\sigma_*^2 = \sigma_0^2 - \sigma_1^4/\sigma_2^2$. The difference between the new estimators and the estimators from equations (14)-(17) is that the new ones correspond to partial derivatives of the density at fixed values of the Laplacian. The old estimators are the derivatives of the projected distribution, which is not the same due to the correlation between density and Laplacian. As we will see in Section 3.3, the new estimators are less scale dependent, since they are closer to pure partial derivatives with respect to the density (in the separate universe sense).

We will refer to the new estimators as the "estimators with spatial corrections of order 2" whereas the previous estimators are without any spatial corrections (i.e. of spatial order 0). It is straightforward to obtain higher order spatial corrections by considering the covariance matrix of the distribution of $\boldsymbol{x} = (\delta, L, P)^T$ with P as in equation (37); these would then be estimators of spatial order 4. Numerically it is very simple to determine the covariance matrix and to evaluate the estimators as in equation (47), so that in principle it is not very difficult to obtain corrections at any spatial order. However, for simplicity we will only evaluate estimators of spatial order 0 and 2 in this article.

We note that our estimators with spatial corrections are quite similar in spirit to the considerations by Musso et al. (2012); Paranjape & Sheth (2012); Paranjape et al. (2013b,a) to map scale-dependent measurements of biases as in equation (13) under the assumption of excursion set (peaks) models to scale-independent large-scale parameters. However, here we can see that such a mapping to the large-scale limit can be done *without* the assumption of any models and we provide therefore a generalization of their approach. Consider for example that the relation between the spatial order 0 and spatial order 2 estimator of b_1 is given by

$$\langle \delta \rangle_g = \mathsf{C} \boldsymbol{b}_1 = \sigma_0^2 b_{1,\text{SO0}}$$
$$= \sigma_0^2 b_{1,\text{SO2}} - \sigma_1^2 b_{L,\text{SO2}} .$$

Therefore, $b_{L,\text{SO2}}$ predicts the scale dependence of $b_{1,\text{SO0}}$ and $b_{1,\text{SO2}}$ posits a more scale-independent estimate:

$$b_{1,\text{SO0}} = b_{1,\text{SO2}} - \frac{\sigma_1^2}{\sigma_0^2} b_{L,\text{SO2}} . \tag{54}$$

The additional cost of our approach (e.g. in comparison to Paranjape et al. 2013a) is that, beneath the density, also the Laplacian has to be evaluated in Lagrangian space, but the benefits are model-independence and a much reduced mathematical complexity.

2.7. Multivariate cumulants

Analogously to the monovariate case, we define multivariate cumulant bias parameters as

$$\boldsymbol{\beta}_N = \nabla_x^N \log(F)\big|_{x=0} \tag{55}$$

which leads immediately to

$$\boldsymbol{\beta}_1 = \boldsymbol{b}_1 \tag{56}$$

$$\boldsymbol{\beta}_2 = \mathsf{b}_2 - \boldsymbol{b}_1 \otimes \boldsymbol{b}_1 \tag{57}$$

$$\beta_{3,ijk} = b_{3,ijk} - (b_{1,i} b_{2,jk} + b_{1,j} b_{2,ki} + b_{1,k} b_{2,ij}) + 2 b_{1,i} b_{1,j} b_{1,k} \tag{58}$$

where we gave $\boldsymbol{\beta}_3$ in index notation, since the central term is difficult to express in vectorial notation. Note that this leads to the same relations between the density cumulant biases and their canonical biases as in equations (26)-(29), but it also includes additional relations like e.g.

$$\beta_{\delta\delta L} = b_{\delta\delta L} - b_{\delta\delta} b_L - 2 b_{\delta L} b_\delta + 2 b_\delta^2 b_L . \tag{59}$$

Finally, the relation between the separate universe bias function and the moment and cumulant generating functions is in analogy to equation to equations (20) and (22) given by

$$M(\boldsymbol{t}) = \exp\left(\frac{1}{2} \boldsymbol{t}^T \mathsf{C} \boldsymbol{t}\right) F(\boldsymbol{x}_0 = \mathsf{C} \boldsymbol{t}) \tag{60}$$

$$K(\boldsymbol{t}) = \frac{1}{2} \boldsymbol{t}^T \mathsf{C} \boldsymbol{t} + \log\left(F(\mathsf{C} \boldsymbol{t})\right) \tag{61}$$

which can be differentiated to show that the cumulants are given in an easy form through the cumulant bias parameters, e.g.

$$\boldsymbol{\kappa}_1 = \mathsf{C} \boldsymbol{\beta}_1 \tag{62}$$

$$\boldsymbol{\kappa}_2 = \mathsf{C} \boldsymbol{\beta}_2 \mathsf{C} + \mathsf{C} \tag{63}$$

$$\kappa_{3,ijk} = \sum_{abc} C_{ai} C_{bj} C_{ck} \beta_{abc} . \tag{64}$$

However, instead of measuring directly the cumulants of the galaxy environment distribution and inverting equations (62) - (64), it is in practice simpler to instead define the variable

$$\boldsymbol{u} = \mathsf{C}^{-1}\boldsymbol{x} \tag{65}$$

and measure its cumulants κ_u which relate to the cumulant biases as

$$\beta_{ijk...} = \begin{cases} \kappa_{u,ijk...} & \text{if} \quad i+j+k+... \neq 2 \\ \kappa_{u,ijk...} - C_{ab}^{-1} & \text{if} \quad i+j+k+... = 2 \end{cases} \tag{66}$$

where a and b indicate the indices of the non-zero variables in the second order case.

We show in Appendix A that it is also possible to derive similar estimators if the density field was filtered with a function different than a sharp k-space filter. The only additional complication in that case that is necessary to account for the correlation matrix between the smoothed large scales and unsmoothed small scales. However, for simplicity, we will focus in this article only on measurements with the sharp k-space filter.

2.8. The PBS break-down scale

The core assumption of the bias formalism is that on sufficiently large scales the response to perturbations becomes independent of the length scale – commonly known as the PBS. Under this assumption it is possible to formulate scale-independent bias parameters and to evaluate predictions at different scales. However, the PBS assumption cannot hold to arbitrarily small scales – both for physical reasons (e.g. the response of haloes to perturbations much smaller than their Lagrangian radius should be zero) and for mathematical reasons. Here, we want to briefly highlight the mathematical scale at which the PBS has to break down inevitably.

Recall that we have shown in equation (24), how κ_2, the variance of the galaxy environment distribution, changes with scale. The predicted variance is only well defined, if

$$\sigma^2 \leq \sigma_{\max,\delta}^2 = \frac{1}{-\beta_2} \tag{67}$$

and will become zero if $\sigma = \sigma_{\max,\delta}$. It is important to note that $\beta_2 < 0$ in general – as we will see fr:in Section 3.4. Mathematically, at this scale the environment distribution and the bias function have to become Dirac-delta functions. Beyond this scale, the PBS predicts a galaxy environment distribution with negative variance – making any density-only bias model mathematically inconsistent. It may seem as if it was possible to set up expansion bias models for arbitrary high σ, but this is only so, since the negativity of the bias function makes it formally possible to have a negative variance. Actual galaxies will strictly obey $\kappa_2 > 0$ and therefore, the response of density-only models has to become scale-dependent latest at $\sigma_{\max,\delta}$ – but likely already earlier. We call this the "PBS-breakdown scale".

The break-down scale is different if additional variables are considered. For the multivariate distribution of density and Laplacian (δ, L), we have to require

$$\det(\boldsymbol{\kappa}_2) \geq 0 \tag{68}$$
$$\Leftrightarrow \det(1 + \mathsf{C}\boldsymbol{\beta}_2) \geq 0 \tag{69}$$

which appears slightly more complicated, but it will also typically be violated at some finite damping scale, as we will see in Section 3.7.

3. Density Bias Measurements

In this Section we evaluate the Lagrangian density bias parameters for different sets of haloes. We will verify the consistency of our estimators by comparing them to the vast literature available on the subject; furthermore, we will also compare the canonical bias parameters b_n with the cumulant bias parameters β_n.

3.1. Simulation

For the analysis through this paper, we use a single cosmological box simulation with high resolution. This simulation is part of the "BACCO simulation project" that was first introduced in Angulo et al. (2021). It has a boxsize of $L = 1440h^{-1}\text{Mpc}$ with 4320^3 particles leading to a mass resolution of $m_p = 3.2 \times 10^9 h^{-1} M_\odot$. The cosmological parameters are $\Omega_m = 0.307$, $\Omega_\Lambda = 0.693$, $\Omega_b = 0.048.$, $n_s = 0.9611$, $\sigma_8 = 0.9$, $h = 0.677$ which are similar to the Planck Collaboration et al. (2020) cosmology except for the roughly 10% larger value of σ_8. By default we use this simulation at a scale-factor of $a = 1.08$ corresponding approximately to $a = 1$.

To identify haloes the simulation code uses a modified version of SUBFIND which first identifies haloes through a friends of friends (FoF) algorithm and subsequently calculates for each FoF group the mass M_{200b} in a region that encloses 200 times the mean density of the universe.

3.2. Bias measurements

We define sets of biased tracers by considering haloes selected by their mass. For this we consider 21 equally log-spaced halo masses between $10^{12}h^{-1}M_\odot$ and $10^{16}h^{-1}M_\odot$ and for each halo mass we consider all haloes with masses in the range

$$M_{200b} \in [M_i/1.25, M_i \cdot 1.25] \; . \tag{70}$$

To evaluate our Lagrangian bias estimators we only need to know the linear field evaluated at the Lagrangian locations of the tracers. We approximate the Lagrangian location of each halo through the Lagrangian origin of its most bound particle. Since, the simulation started from a Lagrangian grid, the Lagrangian origin of the most bound particle can easily be inferred from its id i_{mb} as

$$\boldsymbol{q}_{\text{mb}} = \frac{L}{N_{\text{grid}}} \begin{pmatrix} i_x \\ i_y \\ i_z \end{pmatrix} \tag{71}$$

$$i_{\text{mb}} = i_x N_{\text{grid}}^2 + i_y N_{\text{grid}} + i_z \tag{72}$$

where $N_{\text{grid}} = 4320$ is the number of particles per dimension.

We know the linear density field of the simulation through the initial condition generator. To save computation time, we create a low resolution grid representation of the linear density field with N_{lin}^3 grid points. For fields different than the linear density field we additionally multiply by the correct operator (e.g. $-k^2$ for the Laplacian) in Fourier space. We create a smoothed version of this field by multiplying with a filter in Fourier space

$$\delta_k = \delta_{\text{lin},k} \cdot \sqrt{\frac{1}{2}\text{erfc}(100\log_{10}(k/k_\text{d}))}, \tag{73}$$

corresponding to a slightly smoothed version of a sharp k-space filter with damping scale k_d where we test in each measurement different damping scales $k_\text{d} \in [0.1, 0.15, 0.2, 0.25, 0.3]h^{-1}\text{Mpc}$. We then deconvolve this field with a linear interpolation kernel

and interpolate it to the Lagrangian locations of our tracer set. We choose N_{lin} sufficiently large that the resulting interpolated values are virtually independent of this discretization, e.g. $N_{\text{lin}} = 183$ at $k_{\text{d}} = 0.1 h^{-1} \text{Mpc}$ and $N_{\text{lin}} = 549$ at $k_{\text{d}} = 0.3 h^{-1} \text{Mpc}$.

To estimate the covariance of a set of measured bias parameters we use a Jackknife technique. For this we divide the box in Lagrangian space into N_{jk}^3 subboxes with $N_{\text{jk}} = 4$. We perform 64 measurements of the vector of bias parameters $\boldsymbol{b}_i$ by subsequently leaving out all tracers in one of the subboxes. Then we estimate the covariance through

$$\mathbf{C}_{\boldsymbol{b}} = \frac{1}{N_{\text{jk}}^3 - 1} \sum_i (\boldsymbol{b}_i - \boldsymbol{b}_0) \otimes (\boldsymbol{b}_i - \boldsymbol{b}_0) \tag{74}$$

$$\boldsymbol{b}_0 = \frac{1}{N_{\text{jk}}^3} \sum_i \boldsymbol{b}_i . \tag{75}$$

It is worth noting that the Jackknife estimator results in a more reliable estimate of the uncertainty of the measurement than e.g. a simple bootstrap would yield. When comparing these estimators we found that the Jackknife gives larger (more conservative) error estimates, since it also accounts for the uncertainty induced by cosmic variance, which is sampled by leaving out spatially correlated parts of the data sets.

3.3. b_1 and scale dependence of estimators

In Figure 2 we show measurements of b_1 for haloes as a function of M_{200b} using the estimator from equation (14) without spatial corrections (top panel) versus the estimator from equation (48) with spatial corrections of order 2 (bottom panel). The dashed lines indicates the fitting function from Tinker et al. (2010) evaluated for the cosmology of our simulation.

As expected, haloes of large masses $M \sim 10^{15} h^{-1} M_{\odot}$ are highly biased $b_1 \sim 3$ whereas low-mass haloes $M \sim 10^{12} h^{-1} M_{\odot}$ have even a slightly negative (Lagrangian) bias. The spatial order 0 estimators have a very small degree of statistical uncertainty, but exhibit a significant scale dependence and are inconsistent across different damping scales. On the other hand the estimators with spatial order 2 have a larger statistical uncertainty, but seem consistent across different damping scales, except at very large k_{d} and large halo masses where also some scale dependence becomes visible, probably indicating the effect of higher spatial order terms.

Our spatial order two measurements agree at all scales well with the fitting function from Tinker et al. (2010). While they lie systematically about 5% below this fit, this is within the quoted uncertainty of the Tinker et al. (2010) fit and a similar difference can be seen in Lazeyras et al. (2016). In comparison to other studies at $z = 0$ (e.g. Lazeyras et al. 2016), we have lower bias values at the same masses (e.g. $b_1 \sim 0.5$ at $10^{14} h^{-1} M_{\odot}$ instead of $b_1 \sim 1$), because of the later time ($z = -0.08$) and relatively high variance ($\sigma_8 = 0.9$) of our simulation.

To further highlight the difference in the scale dependence of the estimators, we show in Figure 3 the error Δb_1 in b_1 measurements as a function of damping scale for a few different halo masses. Here we phrase the error $\Delta b_1 = b_1 - b_{1,\text{ref}}$ relative to the spatial order 2 estimator at $k_{\text{d}} = 0.2 h \text{Mpc}^{-1}$ as $b_{1,\text{ref}}$. Again the spatial order 0 estimator has a significant scale dependence in all cases whereas the order 2 estimators are almost scale independent. As expected, the order zero estimator seems to approach the selected reference value on large scales.

We can estimate the scale dependence of the spatial order 0 estimator as in equation (54). We mark the estimated scale dependence of the spatial order 0 estimators using the spatial order
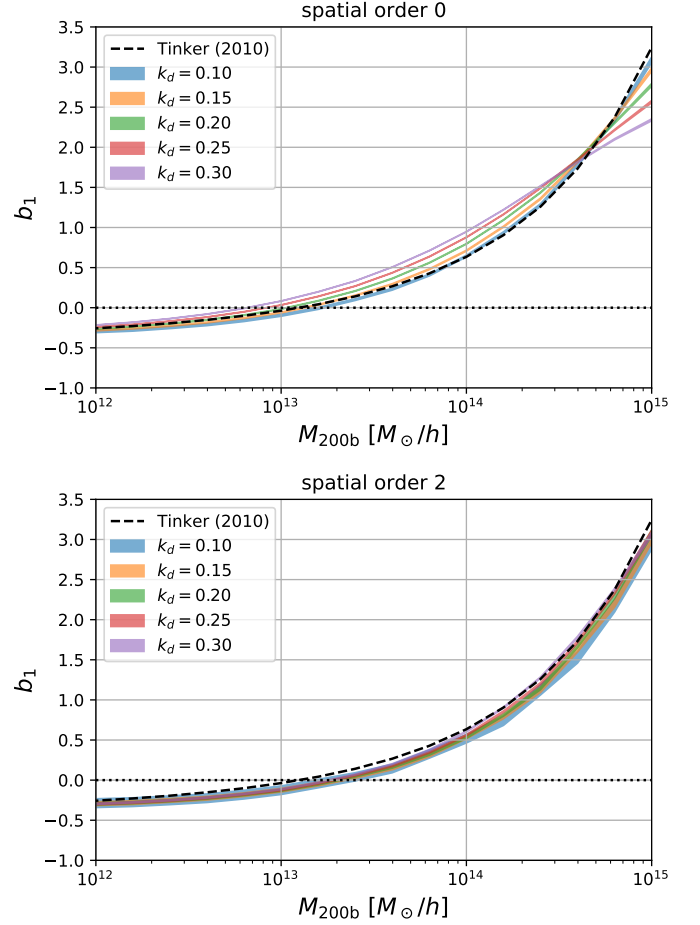

Fig. 2. b_1 as a function of halo mass using the estimators from equation (14) (top) and equation (48) (bottom) measured at different damping scales (different coloured regions). The shaded regions indicate the 1σ certainty region of the estimators. Using the b_1 estimator that includes the Laplacian correction increases the uncertainty of the b_1 estimates, but reduces the dependence on the damping scale, leading to a good agreement across different scales.

2 bias parameters measured at at $k_{\text{d}} = 0.2 h \text{Mpc}^{-1}$ as dashed lines in Figure 3. The scale dependence of the spatial order 0 estimator is predicted quite well, showing that the consideration of the Laplacian introduces a correction of order σ_1^2/σ_0^2 which scales approximately as k_{d}^2.

3.4. b_2 versus β_2

The first cumulant bias parameter that is different from the canonical bias parameter of the same order is $\beta_2 = b_2 - b_1^2$ where $b_1 = \beta_1$. Since there is a one to one relation between (b_1, b_2) and (β_1, β_2) it might seem that there should not be any advantage by using β_2 over b_2 as a parameter when fitting datasets. However, here we will show that β_2 seems to be more independent of β_1 than b_2 of b_1, especially when their covariance matrix is considered.

In Figure 4, we show the measured coevolution relation between b_2 and b_1 (left) versus the coevolution between β_2 and β_1 (right) for the spatial order 2 estimators. In comparison we also show the coevolution relations measured by Lazeyras et al. (2016) as a dashed line, which seems to match our measured relation well up to $k_{\text{d}} \lesssim 0.15 - 0.2 h \text{Mpc}^{-1}$, showing that our method
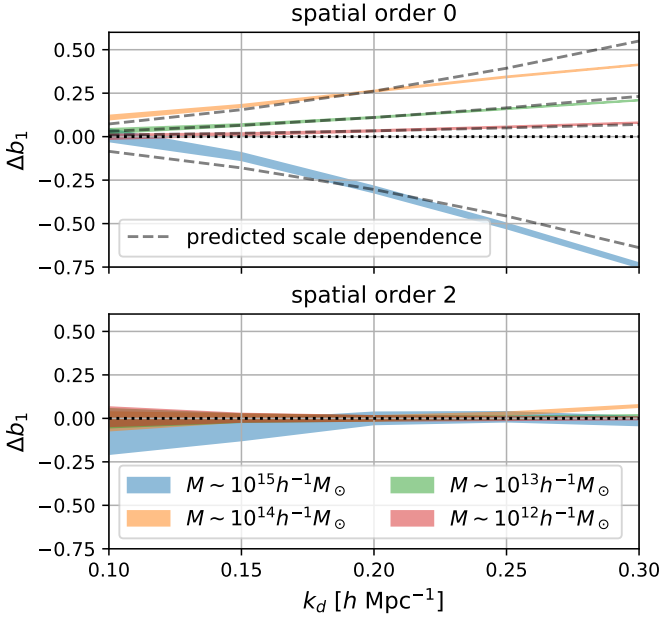
Fig. 3. Scale dependence of the error in b_1 estimates for different halo mass selections. The error is phrased relatively to the measurement with spatial order 2 at $k_d = 0.2h\text{Mpc}^{-1}$. On large scales (small k_d) the spatial order 0 estimators from equation (14) converge well to this estimate, whereas the spatial order 2 estimators from equation (48) agree well with this at all considered damping scales.

for measuring the bias parameters is indeed able to recover the correct large-scale limit of the bias parameters.

We notice that for halos we find $\beta_2 < 0$ across all masses, and with high statistical significance. This means that the width of the halo environment distribution $p(\delta|g)$ is smaller than the width of the background distribution $p(\delta)$ – showing that halo formation is more selective than a random distribution. Possibly, the assumption $\beta_2 < 0$ could be used to limit the considered prior range when fitting certain galaxy surveys. However, it is difficult to anticipate whether this should hold for every possible set of galaxies.

Further, we note that the coevolution relation between β_2 and β_1 is monotonic and roughly linear. For $b_1 \gtrsim 2$ it satisfies $|\beta_2| < |b_2|$. Therefore, β_2 versus β_1 seems slightly simpler than the b_2 to b_1 relation.

The benefit of using β_2 as a parameter is even more evident when considering the covariance matrix of the measurements. In Figure 5 we show the uncertainty of the b_2 measurements (solid lines), versus the uncertainty of the β_2 measurements. Comparing different damping scales shows that the inclusion of smaller scales in the measurements decreases the statistical uncertainty (but increases systematic error). In all cases and especially at high values of b_1 the uncertainty of β_2 is significantly smaller than that of b_2. This is so, since b_2 and b_1 are more correlated than β_2 and β_1.

To highlight this we show in Figure 6 the correlation coefficient

$$\alpha_{12} := \frac{C_{12}}{\sqrt{C_{11}C_{22}}} \qquad (76)$$

where C is the covariance matrix between b_1 and b_2 or β_1 and β_2. For $b_1 \gg 0$ the correlation coefficient of (b_1, b_2) is quite large, even close to 1 in some cases. On the other hand, the correlation coefficient of (β_1, β_2) is quite small and seems on average

consistent with 0. Therefore, measurements of b_1 are quite entangled with b_2, but not so much with β_2. While we do not show it here, we have found that this is also the case for higher order correlations, e.g. of (b_1, b_3) versus (β_1, β_3).

This can be understood when considering the difference between moments and cumulants of a probability distribution. E.g. if we knew about some distribution $p(x)$ that it has a large first moment $\langle x \rangle$, then we might also expect that the second moment $\langle x^2 \rangle$ is large. On the other hand, knowledge about the mean of a distribution, is very uninformative about its width – that is its second cumulant $\sigma_x^2 = \langle (x - \langle x \rangle)^2 \rangle = \langle x^2 \rangle - \langle x \rangle^2$. In the same sense, we expect cumulant bias parameters to be more independent of each other than canonical bias parameters.

3.5. Higher order biases

While the second order cumulant bias β_2 has already slight advantages over b_2 as highlighted in the last section, the benefits are even more significant at higher orders. Here, we will compare the behavior of b_3 and b_4 with the behavior of β_3 and β_4.

In Figure 7, we show b_3, b_4, β_3 and β_4 as a function of b_1 for the spatial order two estimator. The uncertainty of the measurements of b_3 and b_4 are much larger than for b_2 so that we hardly find any signal at $k_d \sim 0.1h\text{Mpc}^{-1}$. However, at $k_d \sim 0.15h\text{Mpc}^{-1}$ we can find a meaningful signal and it is consistent with the b_3 measurements of Lazeyras et al. (2016).

Strikingly, the value of β_3 is extremely small in comparison to b_3 and seems approximately consistent with 0. Noteworthy this is not only the case for scales where b_3 is reasonably scale independent ($k_d \lesssim 0.15h\text{Mpc}^{-1}$), but it is also so for much smaller scales. There seems to be no noteworthy relation between β_3 and β_1. Further, the coevolution relation measured by Lazeyras et al. (2016) seems consistent with $\beta_3 = 0$. Therefore, we could summarize the third order coevolution relation through the relation we get by using $\beta_3 = 0$ in equation (28):

$$b_3 = 3b_1 b_2 - 2b_1^3 \ . \qquad (77)$$

Similarly, we find for β_4 an apparent independence of β_1 and consistency with zero across different scales. Assuming $\beta_4 = 0$ in equation (29) leads to a coevolution relation between b_4, b_3, b_2 and b_1 given by

$$b_4 = 4b_1 b_3 + 3b_2^2 - 12b_1^2 b_2 - 6b_1^4 \qquad (78)$$

We mark this relation with the measurements of b_1, b_2 and b_3 from Lazeyras et al. (2016) as a dashed line in the bottom left panel of Figure 7. This seems in good agreement with the measured b_4.

We conclude that at order $n \geq 3$ the cumulant bias parameters are very close to zero. Coevolution relations at these orders can be summarized simply through $\beta_n = 0$. We might therefore suggest that high order coevolution relations do not represent any physical insights, but rather highlight that the canonical bias parameters form a suboptimal, deeply entangled basis.

Further, the fact that all the cumulant biases beyond order two are close to zero indicates that the galaxy environment distribution is very well approximated by a Gaussian distribution. We will investigate the possibility of a Gaussian bias model in Stücker et al. (in prep) where we also explain how this fact may arise naturally from the Gausianity of the background distribution.
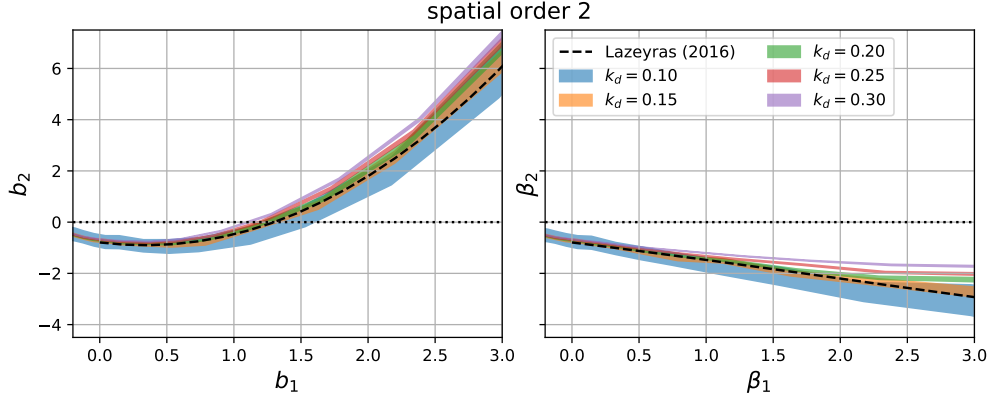
Fig. 4. b_2 and β_2 as a function of $b_1 = \beta_1$ using the spatial order 2 estimators. The black dashed line is the $b_2(b_1)$ relation inferred by Lazeyras et al. (2016) from separate universe simulations. The spatial order 2 estimators seem consistent with the literature coevolution relation down to damping scales of $k_d \sim 0.15 - 0.2h\mathrm{Mpc}^{-1}$. Noteworthy, it seems that $\beta_2 < 0$ holds always, which means that the width of the galaxy environment distribution $p(\delta|g)$ is always smaller than that of the background $p(\delta)$. Therefore, the cumulant bias parameter $\beta_2 - \beta_1$ relation appears slightly simpler than the $b_2 - b_1$ relation.
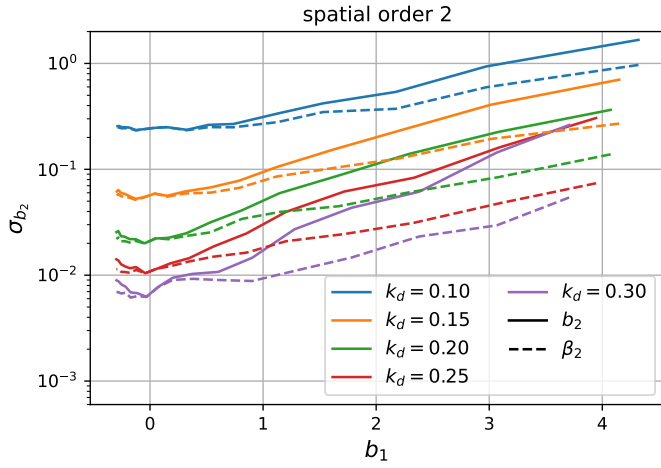

Fig. 5. Uncertainty of b_2 measurement (solid lines) and β_2 measurements (dashed lines) as a function of b_1 for different damping scales. The uncertainty of β_2 is significantly smaller than of b_2 for high values of b_1.

Fig. 6. Correlation coefficient α_{12} between the b_1 and b_2 measurement (solid) and the β_1 and β_2 measurement (dashed). Note that β_2 as a parameter is much more independent of the value of b_1 than b_2.

3.6. Laplacian Bias

In Figure 8, we show the Lagrangian Laplacian bias parameter as a function of halo mass, as inferred from the estimator in equation (49). The Laplacian measurements have a much stronger scale dependence than our previous measurements of density bias parameters. However, it seems that at masses $M \lesssim 3 \times 10^{13}h^{-1}M_\odot$ the scale dependence disappears and a reliable measurement is obtained.

We compare our Lagrangian measurements with the Eulerian measurements of Lazeyras & Schmidt (2019). In general, the relation between the Lagrangian and Eulerian Laplacian bias depends on velocity bias (e.g. Desjacques et al. 2018). The velocity bias b_s quantifies the difference between the displacement field of matter s and galaxies s_g and gets as first order contributions of the form

$$s_g - s \approx b_s \nabla \delta. \tag{79}$$

Therefore, the galaxy density can get an extra contribution proportional to the divergence of this field, that scales as $L = \nabla^2 \delta$. However, for our set of tracers $s_g = s$ by definition, so that $b_s = 0$
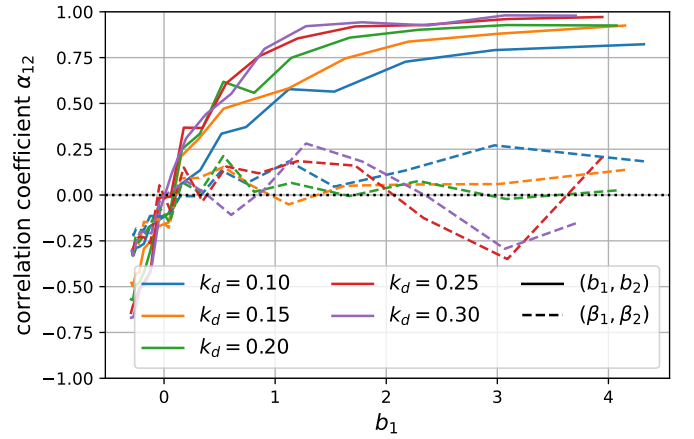
and we simply assume that their Lagrangian and Eulerian Laplacian biases are identical or at least very similar.

In the reliable range $M \lesssim 3 \times 10^{13}h^{-1}M_\odot$ it seems that our measurements agree well with the linear and the quadratic (in halo radius) fit that was inferred by Lazeyras & Schmidt (2019) for the Eulerian Laplacian bias. Therefore, we confirm that the Laplacian bias is negative for haloes and that it can plausibly have amplitudes of order $b_L \sim -20\mathrm{Mpc}^2h^{-2}$. However, we leave more detailed considerations to future studies. In principle, more accurate measurements could be obtained by consideration of higher order corrections and through measurements at higher redshifts.

3.7. The PBS break-down scale

As we have explained in Section 2.8, the PBS has to mathematically break down at some finite scale and bias parameters have to become scale-dependent beyond it. Here we aim to visualize this scale.

The break-down scale depends on the variables that are considered and the second order cumulant bias parameters of the considered variables. For example, it depends through equation
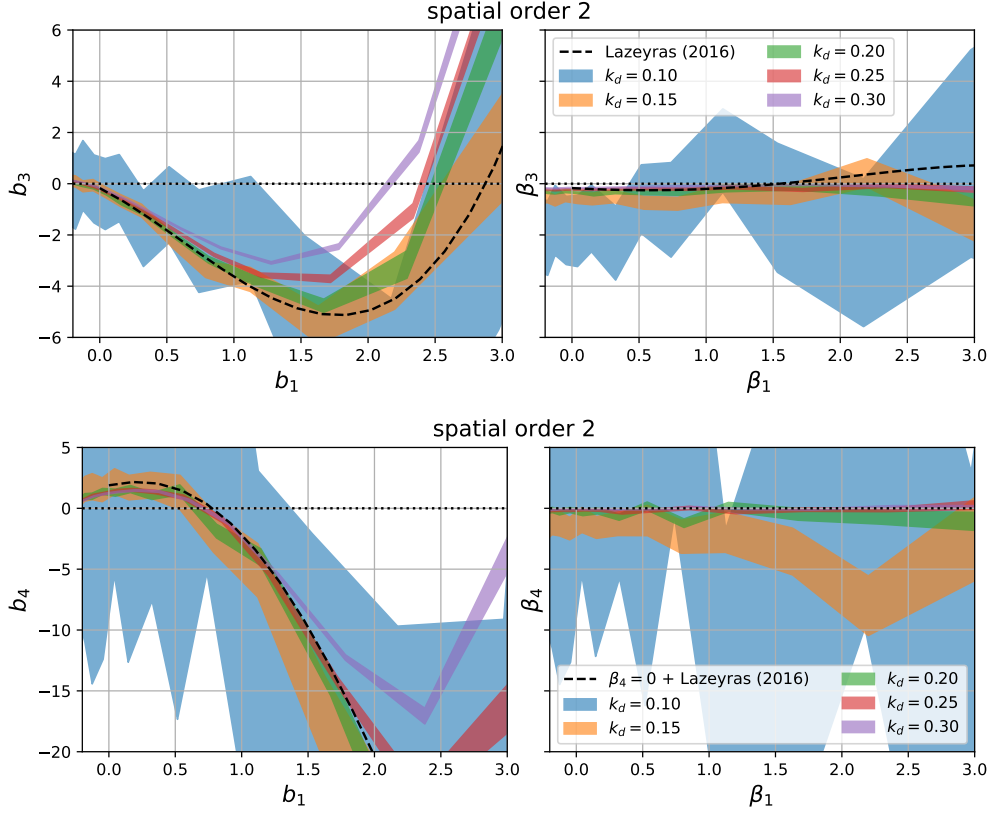
Fig. 7. Coevolution relations of higher order bias parameters b_3 and β_3 (top) and b_4 and β_4 (bottom) for the spatial order 2 bias estimators. For b_4 we indicate as a dashed line a prediction that follows from combining the Lazeyras et al. (2016) measurements of b_1, b_2 and b_3 with equation (29) when using $\beta_4 = 0$. Strikingly, β_3 and β_4 are extremely close to 0 – independently of the value of b_1.
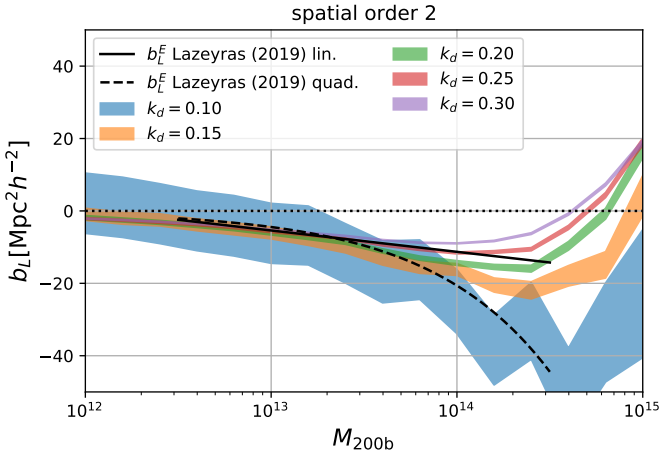

Fig. 8. The Lagrangian Laplacian bias b_L as a function of halo mass. At masses $M \lesssim 3 \times 10^{13} h^{-1} M_\odot$ the scale dependence of the b_L measurements disappears and they agree well with the fits of the Eulerian Laplacian bias from Lazeyras & Schmidt (2019).
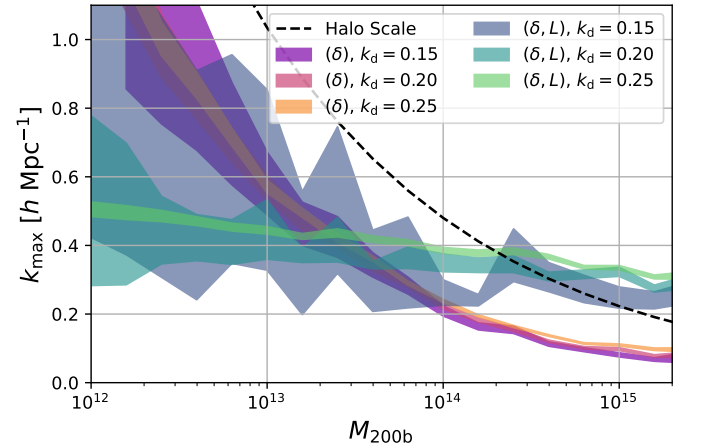
Fig. 9. The maximum damping scale where the PBS can be valid and bias scale-independent. The three reddish contours show the breakdown scale of density-only bias models, estimated with bias parameters measured at different scales, and the three green-blueish contours show it for (δ, L) bias models. The black dashed line shows the wavenumber associated with the Lagrangian radius of haloes. The break-down scale is consistent across different measurements and only seems to scale strongly with halo radius for the (δ) case.

(67) only on $\beta_{\delta\delta}$ for the mono-variate density only case, but for the bi-variate (δ, L) case it depends on $\beta_{\delta\delta}$, $\beta_{\delta L}$ and β_{LL} through equation (69).

For each halo mass, we measure the bias parameters at three different scales $k_d \in (0.15, 0.2, 0.25) h\text{Mpc}^{-1}$. Then, we evaluate the co-variance matrix of the background distribution at 500 different equally log-spaced damping scales between $10^{-3} h\text{Mpc}^{-1}$ and $10^1 h\text{Mpc}^{-1}$ and determine the earliest damping scale $k_{\max}$

where the covariance is large enough to violate equation (67) or equation (69). We show the corresponding results as shaded contours in Figure 9, where we additionally mark the characteristic

wave-number of haloes

$$k_{\text{halo}} = \frac{\pi}{R_{\text{halo}}} \tag{80}$$

where R_{halo} is the Lagrangian radius that encloses the halo mass M_{200b}.

Comparing the measurements at different damping scales, we find that the inferred scale is reasonably converged with the scale that we measured the bias parameters at. The density-only break-down scale is typically a factor two smaller than k_{halo} and seems roughly proportional to it. We therefore conclude that *any* Lagrangian local in matter density (LLIMD) bias model has to break down at a length scale roughly a factor two larger than the consider haloes.

On the other hand, the (δ, L) case shows a notably different break-down scale. It scales only weakly with halo mass and it ranges only between $k_{\text{max}} \sim 0.3 - 0.5h\text{Mpc}^{-1}$. Note that this corresponds roughly to the Lagrangian scale of haloes with $M \sim 2 \times 10^{14} h^{-1} M_\odot$, which have $b_1 \sim 1$. For haloes above $M \gtrsim 3 \times 10^{13} h^{-1} M_\odot$ including the Laplacian increases k_{max} relative to the density-only case, but for lower masses it decreases it.

Comparing with the right panel of Figure 4, we notice that the bias parameters become already scale-dependent at notably smaller damping scales than k_{max}. For example, for $M_{\text{200b}} \sim 10^{15} h^{-1} M_\odot$ with $b_1 \sim 3$, the measurements of β_2 are already scale-dependent beyond $k_{\text{d}} \gtrsim 0.15h\text{Mpc}^{-1}$, whereas the mathematical break-down scale is $k_{\text{max}} \sim 0.3h\text{Mpc}^{-1}$. We therefore suggest that considerable care should be taken when setting up bias models close to the break-down scale. This is particularly relevant for hybrid methods, which might in principle allow to describe galaxy clustering at notably smaller scales than are usually considered in perturbative schemes (e.g. Modi et al. 2020; Zennaro et al. 2022).

4. Estimators for tensorial bias parameters

In Section 2 we have shown how to infer general estimators of bias parameters associated with scalar variables (like the density and Laplacian) with spatial corrections of any order, and in Section 3 we have shown that these can be used to obtain reliable bias measurements from a single simulation.

However, the theory in Section 2 does not explain yet, how to obtain estimators for parameters like the tidal bias b_{K^2} which is defined as the response to

$$K^2 = \text{tr}(\mathsf{K} \cdot \mathsf{K}) \tag{81}$$

$$\mathsf{K} = (\nabla \otimes \nabla)\phi - \frac{\delta}{3}\mathsf{I}_2 \tag{82}$$

where ϕ is the displacement potential, I_2 is the unit matrix and K is the traceless tidal tensor. In this section, we will present a general scheme to measure the related bias parameter b_{K^2} and any other bias parameters that follow from contractions of derivatives of the potential field.

To achieve this, it is *not* optimal to consider directly the distribution of such scalar contracted quantities, since these distributions may get quite complicated. For example, $p(K^2)$ is not a Gaussian distribution, but rather a χ^2 distribution with five degrees of freedom. Further, it is not immediately obvious how to define partial derivatives with respect to such variables, since partial derivatives may depend on what other terms are kept fixed. For example, it is not clear how a term like $K^3 = \text{tr}(\mathsf{K} \cdot \mathsf{K} \cdot \mathsf{K})$ would be derived with respect to K^2. It seems therefore difficult to generalize equation (11) in this manner.

However, a clear and general framework for measuring such "tensorial" bias terms can be developed by instead considering the full (quite high dimensional) distribution of the tidal tensor and its derivatives. Since the potential field of the early universe is a Gaussian random field, these must follow a multivariate Gaussian distribution so that it is easy to compute derivatives of the distribution function in a general manner. The resulting "bias tensors" can be decomposed into isotropic tensors that each have a one to one correspondence with traditionally used bias parameters. In this section, we will introduce the needed mathematical notions step by step and will provide estimators for a few selected bias terms.

4.1. Tensorial Bias Expansion

We write the bias expansion in tensorial form as

$$F = 1 + \mathsf{B_T T} + \frac{1}{2}\mathsf{T B_{TT} T} + ... + \mathsf{B_S S} + ... + \mathsf{B_R R} + ... \tag{83}$$

where again $F = n_\text{g}/n_{\text{g},0}$ (in a "separate universe" sense) and

$$\mathsf{T} = (\nabla \otimes \nabla)\phi \tag{84}$$
$$\mathsf{S} = (\nabla \otimes \nabla \otimes \nabla)\phi \tag{85}$$
$$\mathsf{R} = (\nabla \otimes \nabla \otimes \nabla \otimes \nabla)\phi \tag{86}$$

where $\mathsf{B_T}$ is a canonical bias tensor of rank 2, $\mathsf{B_{TT}}$ of rank 4, $\mathsf{B_S}$ of rank 3 etc. and where an omitted product sign indicates a product over the indices of the last fully symmetric part of the first tensor and the first fully symmetric part of the second tensor. For example:

$$\mathsf{B_T T} = \mathsf{B_T}\overset{(2)}{\cdot}\mathsf{T} = \sum_{ij} B_{\mathsf{T},ij} T_{ij}$$

$$\mathsf{T B_{TT} T} = \mathsf{T}\overset{(2)}{\cdot}\mathsf{B_{TT}}\overset{(2)}{\cdot}\mathsf{T} = \sum_{ijkl} T_{ij} B_{\mathsf{TT},ijkl} T_{kl}$$

$$\mathsf{B_S S} = \mathsf{B_S}\overset{(3)}{\cdot}\mathsf{S} = \sum_{ijk} B_{\mathsf{S},ijk} S_{ijk}$$

where write $\overset{(n)}{\cdot}$ to explicitly denote the number of dimensions that are contracted.

Note that, just as before, the bias tensors correspond to derivatives of the large-scale bias function, e.g.

$$\mathsf{B_T} = \left.\frac{\partial}{\partial \mathsf{T}}F\right|_{\mathsf{T}=0} \tag{87}$$

$$\mathsf{B_{TT}} = \left.\left(\frac{\partial}{\partial \mathsf{T}} \otimes \frac{\partial}{\partial \mathsf{T}}\right)F\right|_{\mathsf{T}=0} \tag{88}$$

4.2. Isotropic Tensors

Due to isotropy of the universe, each of the bias tensors has to be isotropic. That means that a bias tensor should be identical when measued from a rotated frame of reference. For example, it has to be

$$\mathsf{U}^T \mathsf{B_T U} = \mathsf{B_T} \tag{89}$$

for any rotation matrix U. From this it follows immediately that $\mathsf{B_T}$ has to be proportional to the unit matrix I_2 (where the proportionality constant is equal to b_1). In general a rank n tensor A is isotropic if it holds for any rotation matrix U:

$$A_{abc...}U_{ai}U_{bj}U_{ck}... = A_{ijk...} \tag{90}$$

where we have used Einstein's sum convention and where the rotation matrix is applied to each index of A individually. To express whether some tensor is isotropic, we define $\mathbb{U}_n$ as the space of all tensors that are isotropic and of rank n.

In general, all isotropic tensors can be decomposed in index notation through different combinations of the Kronecker-delta symbol δ_{ij} and the Levi-Civita symbol ϵ_{ijk}. For example

$$A \in \mathbb{U}_2 \Rightarrow A_{ij} = a\delta_{ij} \tag{91}$$

$$A \in \mathbb{U}_3 \Rightarrow A_{ijk} = a\epsilon_{ijk} \tag{92}$$

$$A \in \mathbb{U}_4 \Rightarrow A_{ijkl} = a\delta_{ij}\delta_{kl} + b\delta_{ik}\delta_{jl} + c\delta_{il}\delta_{jk} \tag{93}$$

where $a, b, c \in \mathbb{R}$. A compact way of writing the same type of statement is e.g.

$$\mathbb{U}_4 = \text{span}\left(\{\delta_{ij}\delta_{kl}, \delta_{il}\delta_{jk}, \delta_{ik}\delta_{jl}\}\right) \tag{94}$$

which signifies that $\mathbb{U}_4$ is the space of tensors that can be reached through linear combinations of the indicated tensors and we say that these are basis tensors of $\mathbb{U}_4$.

It is easy to see that we should be able to decompose each of the bias tensors into a small number of independent scalars that multiply the basis tensors and that correspond to traditional bias parameters. However, before performing such a decomposition we further need to consider the symmetries of the tensors.

4.3. Symmetric Isotropic Tensors

Since the bias tensors correspond to derivatives of a scalar function with respect to symmetric tensors, they have themselves to obey the same symmetries. For example, the tensor B_{TT} has to be symmetric in the first two and last two indices[3]:

$$B_{TT,ijkl} = B_{TT,jikl} = B_{TT,ijlk} = B_{TT,jilk} . \tag{95}$$

Therefore, B_{TT} cannot be every tensor from the tensorspace $\mathbb{U}_4$, but only from the subspace $\mathbb{V}_{22} \subseteq \mathbb{U}_4$ of isotropic rank four tensors that have the $2, 2$ symmetry.

To formalize this, we define $\mathbb{V}_n \subset \mathbb{U}_n$ as the space of all isotropic tensor of rank n that are additionally fully symmetric under any permutation of indices.

$$B_T \in \mathbb{V}_2 \tag{96}$$

$$B_S \in \mathbb{V}_3 \tag{97}$$

$$B_R \in \mathbb{V}_4 . \tag{98}$$

Note that $\mathbb{V}_3 = \{0\}$, i.e. the only symmetric rank three isotropic tensor is 0. Further, we define $\mathbb{V}_{nm}$ as the space of all isotropic tensors of rank $n + m$ that are symmetric in the first n indices and the last m indices. For example,

$$B_{TT} \in \mathbb{V}_{22}. \tag{99}$$

To find a basis for some symmetric isotropic tensor space of rank n, we can consider all basis tensors of $\mathbb{U}_n$, symmetrize these and then discard any tensors that are duplicate or zero. We define the symmetrization operator S_n which symmetrizes a tensor in n indices. E.g.

$$S_2(M_{ij}) = \frac{1}{2}(M_{ij} + M_{ji}) \tag{100}$$

$$S_3(M_{ijk}) = \frac{1}{6}(M_{ijk} + M_{ikj} + M_{jik} + M_{jki} + M_{kij} + M_{kji}) \tag{101}$$

---

[3] And additionally symmetric to exchanges between the first two and last two indices. However, this symmetry follows automatically here, so that we do not consider it explicitly.
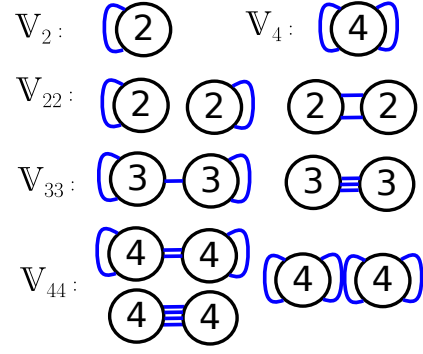
Fig. 10. Graphic representation of the isotropic tensors that form a basis for a few selected isotropic tensor spaces with symmetries. All basis tensors of a space with given symmetry can be constructed by considering the number of different ways that the symmetry groups can be connected. In this figure each circle with number n represents a group of n fully symmetric indices and each connection represents one delta symbol (that can either connect two indices from the same group or from two different groups).

Further, we define a double symmetrization operator S_{nm} that symmetrizes the first n indices and the last m indices. For example, the effect of S_{22} onto a rank four tensor M is given in index notation by

$$S_{22}(M_{ijkl}) = \frac{1}{4}\left(M_{ijkl} + M_{jikl} + M_{ijlk} + M_{jilk}\right) \tag{102}$$

and it acts as follows on the basis tensors of $\mathbb{U}_4$:

$$S_{22}(\delta_{ij}\delta_{kl}) = \delta_{ij}\delta_{kl} =: I_{22,ij} \tag{103}$$

$$S_{22}(\delta_{ik}\delta_{jl}) = \frac{1}{2}(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) =: I_{2=2,ijkl} \tag{104}$$

$$S_{22}(\delta_{il}\delta_{jk}) = \frac{1}{2}(\delta_{il}\delta_{jk} + \delta_{ik}\delta_{jl}) = I_{2=2,ijkl} \tag{105}$$

where we have used that δ_{ij} is itself symmetric. The symmetrized second and third basis tensor are identical. Therefore, $\mathbb{V}_{22}$ has (unlike $\mathbb{U}_4$) only two basis tensors

$$\mathbb{V}_{22} = \text{span}\left(\{I_{22}, I_{2=2}\}\right) . \tag{106}$$

The main difference between $\delta_{ij}\delta_{kl}$ and the two tensors $\delta_{ik}\delta_{jl}$ and $\delta_{il}\delta_{jk}$ is that the delta symbols used to define $\delta_{ij}\delta_{kl}$ each connect internally inside the groups of symmetric indices and have no connections between the groups. (Recall that $i \leftrightarrow j$ and $k \leftrightarrow l$ are to be symmetrized.) However, $\delta_{ik}\delta_{jl}$ and $\delta_{il}\delta_{jk}$ each have zero group internal connections, but two connections between the symmetry groups. In fact, the symmetrization operation identifies all terms that have the same number of intra-/inter- group connections and therefore we only need to consider how many independent possibilities exist to connect the symmetry groups to identify the basis tensors of any space $\mathbb{V}_{mn}$. We can therefore represent the basis tensors of each tensor space through a simple diagram as illustrated in Figure 10. We label these tensors through a symbol that shows the number of inter group connections in the index:

$$\mathbb{V}_{33} = \text{span}\left(\{I_{3-3}, I_{3\equiv3}\}\right) \tag{107}$$

$$\mathbb{V}_{44} = \text{span}\left(\{I_{44}, I_{4=4}, I_{4\equiv4}\}\right) \tag{108}$$

$$\mathbb{V}_{24} = \text{span}\left(\{I_{24}, I_{2=4}\}\right) \tag{109}$$

We explain in Appendix B.1, how to construct tensors with three or more symmetry groups. For these cases the procedure

Tensor	Index Representation	Norm^2	ϕ contr.
J_2	$S_2(\delta_{ij})$	3	δ
J_{22}	$S_{22}(\delta_{ij}\delta_{kl})$	9	δ^2
$J_{2=2}$	$S_{22}(\delta_{ik}\delta_{jl}) - \frac{1}{3}S_{22}(\delta_{ij}\delta_{kl})$	5	K^2
J_4	$S_4(\delta_{ij}\delta_{kl})$	5	L
J_{3-3}	$S_{33}(\delta_{ij}\delta_{kl}\delta_{mn})$	$\frac{25}{3}$	$(\nabla\delta)^2$
$J_{3\equiv3}$	$S_{33}(\delta_{ik}\delta_{mj}\delta_{ln}) - \frac{3}{5}S_{33}(\delta_{ij}\delta_{kl}\delta_{mn})$	7	$S_{3\equiv3}$
J_{24}	$S_{24}(\delta_{ij}\delta_{kl}\delta_{mn})$	15	δL
$J_{2=4}$	$S_{24}(\delta_{ik}\delta_{jl}\delta_{mn}) - \frac{1}{3}S_{24}(\delta_{ij}\delta_{kl}\delta_{mn})$	$\frac{35}{6}$	$\phi_{2=4}$

Table 1. Orthogonal basis tensors that we consider here. The first column indicates the label we use for the tensor, the second column one possible way of defining it in index notation, the third column the square of its norm (as the full contraction with itself) and the last column the term that is obtained when fully contracting it with the corresponding derivatives of the potential as in the bias expansion.

4.4. Orthogonal Basis

The basis tensors highlighted in the last section are sufficient to uniquely decompose the bias tensors. However, it is advantageous to define the basis tensors in such a way that they are orthogonal to each other.

For example, we can decompose an isotropic tensor $\mathsf{M} \in \mathbb{V}_{22}$ as

$$\mathsf{M} = A\mathsf{I}_{22} + B\mathsf{I}_{2=2} . \tag{110}$$

Its full contraction with I_{22} is

$$\mathsf{M} \overset{(4)}{\cdot} \mathsf{I}_{22} = A(\mathsf{I}_{22} \overset{(4)}{\cdot} \mathsf{I}_{22}) + B(\mathsf{I}_{2=2} \overset{(4)}{\cdot} \mathsf{I}_{22}) \tag{111}$$
$$= 9A + 3B \tag{112}$$

which includes contributions of both A and B. We could infer A and B from M by additionally contracting with $\mathsf{I}_{2=2}$ and then solving the resulting system of equations.

However, it would be desirable to have a simple way of finding the coefficients. Therefore, we define an orthogonal basis such that

$$\mathsf{J}_a \overset{(r)}{\cdot} \mathsf{J}_b = \begin{cases} \|\mathsf{J}_a\|^2 & \text{if } a = b \\ 0 & \text{else} \end{cases} \tag{113}$$

where the contraction goes over the full rank r of J_a and J_b (which must have equal rank). Such a basis can be found through Gram-Schmidt orthogonalization and is for example given for $\mathbb{V}_{22}$ by:

$$\mathsf{J}_{22} = \mathsf{I}_{22} \tag{114}$$

$$\mathsf{J}_{2=2} = \mathsf{I}_{2=2} - \frac{1}{3}\mathsf{I}_{22} . \tag{115}$$

Then we can easily find the coefficients as

$$\mathsf{M} = \alpha\mathsf{J}_{22} + \beta\mathsf{J}_{2=2} \tag{116}$$

$$\alpha = \frac{\mathsf{M} \overset{(4)}{\cdot} \mathsf{J}_{22}}{\|\mathsf{J}_{22}\|^2} = \frac{1}{9}\mathsf{M} \overset{(4)}{\cdot} \mathsf{J}_{22} \tag{117}$$

$$\beta = \frac{\mathsf{M} \overset{(4)}{\cdot} \mathsf{J}_{2=2}}{\|\mathsf{J}_{2=2}\|^2} = \frac{1}{5}\mathsf{M} \overset{(4)}{\cdot} \mathsf{J}_{2=2} . \tag{118}$$

Note that the thus inferred coefficients are different than the A and B and the chosen basis will also affect the inferred bias parameters.

It is worth noting that a decomposition of a bias tensor into different bases

$$\begin{aligned}\mathsf{B}_{\mathsf{TT}} &= b_{\mathsf{I}_{22}}\mathsf{I}_{22} + b_{\mathsf{I}_{2=2}}\mathsf{I}_{2=2} \\ &= b_{\mathsf{J}_{22}}\mathsf{J}_{22} + b_{\mathsf{J}_{2=2}}\mathsf{J}_{2=2}\end{aligned} \tag{119}$$

corresponds to different independent scalar variables appearing in the bias expansion. For example, if we track the corresponding term from equation (83), we find

$$\begin{aligned}\mathsf{T}\mathsf{B}_{\mathsf{TT}}\mathsf{T} &= b_{\mathsf{I}_{22}}\operatorname{tr}(\mathsf{T})^2 + b_{\mathsf{I}_{2=2}}\operatorname{tr}(\mathsf{T}^2) \\ &= b_{\mathsf{I}_{22}}\delta^2 + b_{\mathsf{I}_{2=2}}T^2\end{aligned} \tag{120}$$

$$\begin{aligned}\mathsf{T}\mathsf{B}_{\mathsf{TT}}\mathsf{T} &= b_{\mathsf{J}_{22}}\operatorname{tr}(\mathsf{T})^2 + b_{\mathsf{J}_{2=2}}\left(\operatorname{tr}(\mathsf{T}^2) - \frac{\delta^2}{3}\right) \\ &= b_{\mathsf{J}_{22}}\delta^2 + b_{\mathsf{J}_{2=2}}K^2\end{aligned} \tag{121}$$

so that the non-orthogonal basis leads to δ and T^2 as independent terms, whereas the orthogonal basis leads to using δ and K^2 as independent terms. The corresponding bias parameters can be directly identified $b_{\mathsf{J}_{22}} = b_2$ and $\frac{1}{2}b_{\mathsf{J}_{2=2}} = b_{K^2}$.

We list an overview of the orthogonal basis tensors that we will use here in Table 1. While it is possible to derive the full algebra of products between isotropic tensors from the index representations, this is rather cumbersome. Therefore, we have written a short code that creates explicit numerical representations of these tensors and with which it is easy to evaluate (and decompose) different types of products of these tensors. We use this together with the symbolic computer algebra system SYMPY (Meurer et al. 2017) to systematically compute symbolic representations of expressions in the following sections. This code is openly available under `https://github.com/jstuecker/probabilistic-bias`.

4.5. Bias Estimators

Given the orthogonal basis tensors, we define the tensorial bias parameters through

$$b_{\mathsf{J}_X} = \frac{\partial^N F}{\partial \mathsf{T}_0^N} \overset{(2N)}{\cdot} \frac{\mathsf{J}_X}{\|\mathsf{J}_X\|^2}\bigg|_{\mathsf{T}_0=0} \tag{122}$$

where the rank of J_X is $2N$. While for scalar terms, as in equation (2), the bias parameters are simply given by derivatives of the galaxy number with respect to the corresponding scalar, more generally we define bias parameters for any tensorial terms as derivatives of the galaxy number with respect to the tidal tensor (or higher spatial order tensors) that are contracted with the corresponding isotropic tensor.

In complete analogy to equation (45) it follows that these parameters can be estimated as

$$b_{\mathsf{J}_X} = (-1)^N \left\langle \frac{1}{p}\frac{\partial p}{\partial \mathsf{T}^N} \overset{(2N)}{\cdot} \frac{\mathsf{J}_X}{\|\mathsf{J}_X\|^2} \right\rangle_g \tag{123}$$

where p is the full distribution of the tidal tensor T.

Bias terms corresponding to higher spatial derivatives like R and S can be defined analogously.

4.6. Tidal bias

We show in Appendices B.2 and B.3 that the distribution of the tidal tensor is given by

$$p(\mathsf{T}) = N \exp\left(-\frac{1}{2}\mathsf{T}^T \mathsf{C}_\mathsf{T}^+ \mathsf{T}\right) \tag{124}$$

where C_T^+ (the generalized inverse of the covariance matrix, as explained in the Appendix) is given by

$$\mathsf{C}_\mathsf{T}^+ = \frac{1}{\sigma_0^2}\mathsf{J}_{22} + \frac{15}{4\sigma_0^2}\mathsf{J}_{2=2} \ . \tag{125}$$

Taking derivatives yields

$$\frac{1}{p}\frac{\partial p}{\partial \mathsf{T}} = -\mathsf{C}_\mathsf{T}^+\mathsf{T} \tag{126}$$

$$\frac{1}{p}\frac{\partial^2 p}{\partial \mathsf{T}^2} = (\mathsf{T}\mathsf{C}_\mathsf{T}^+) \otimes (\mathsf{C}_\mathsf{T}^+\mathsf{T}) - \mathsf{C}_\mathsf{T}^+ \ . \tag{127}$$

and we find

$$b_{\mathsf{J}_2} = b_1 = \left\langle \frac{\mathsf{J}_2\mathsf{C}^+\mathsf{T}}{\|\mathsf{J}_2\|^2} \right\rangle_\mathrm{g} = \left\langle \frac{\mathsf{J}_2\mathsf{T}}{\sigma_0^2} \right\rangle_\mathrm{g} = \left\langle \frac{\delta}{\sigma_0^2} \right\rangle_\mathrm{g} \ . \tag{128}$$

Here, it appeared a product between isotropic tensors $\mathsf{J}_2 \overset{(2)}{} \mathsf{J}_{22} = 3\mathsf{J}_2$. We evaluate terms of this type systematically with numerical representations as described in Section 4.4. Naturally, this estimator of $b_{\mathsf{J}_2} = b_1$ is consistent with the one that we have inferred previously in equation (14).

At second order we find the estimators

$$b_{\mathsf{J}_{22}} = b_2 = \left\langle \frac{\mathsf{T}^T\mathsf{J}_{22}\mathsf{T} - \sigma_0^2}{\sigma_0^4} \right\rangle_\mathrm{g}$$
$$= \left\langle \frac{\delta^2 - \sigma_0^2}{\sigma_0^4} \right\rangle_\mathrm{g} \tag{129}$$

$$b_{\mathsf{J}_{2=2}} = 2b_{K^2} = \frac{15}{4\sigma_0^4}\left\langle 3\mathsf{T}^T\mathsf{J}_{2=2}\mathsf{T} - 2\sigma_0^2 \right\rangle_\mathrm{g} \tag{130}$$
$$= \frac{15}{4\sigma_0^4}\left\langle 3K^2 - 2\sigma_0^2 \right\rangle_\mathrm{g} \tag{131}$$

We note that conventionally the bias parameter b_{K^2} is defined so that it appears with a pre-factor K^2 in the bias expansion whereas our parameter $b_{\mathsf{J}_{2=2}}$ appears with a factor $\frac{1}{2}K^2$ in the expansion (after contracting the corresponding tensors) so that it is $b_{K^2} = \frac{1}{2}b_{\mathsf{J}_{2=2}}$. Although we think that our fore-factor convention makes in principle more sense, since $b_{\mathsf{J}_{2=2}}$ corresponds to a second derivative term, we will still present results in terms of the conventional notation (e.g. b_{K^2}) throughout this paper.

We note that just as the large-scale value of δ/σ^2 at an object's location is – motivated by equation (14) – sometimes called the "bias-per-object" (e.g. Paranjape et al. 2018; Contreras et al. 2023), one could refer in a similar manner to the value of

$$b_{K^2,1h} = \frac{15}{8\sigma_0^4}(3K^2 - 2\sigma_0^2) \tag{132}$$

as the value of the "tidal bias-per-object".

Just as estimators with spatial corrections could be obtained for the density biases by considering the joint distribution of $p(\delta, L)$, we can obtain higher spatial order estimators for the tidal

bias by considering the joint distribution of second and fourth potential derivatives $p(\mathsf{T}, \mathsf{R})$ and evaluating equation (123) for this distribution. We show how to derive the bias estimator for this case in appendices B.5 and B.6. The resulting estimator is

$$b_{\mathsf{J}_{2=2}} = \frac{15}{4\sigma_*^8}\left\langle 3K^2\sigma_2^4 + 6\phi_{2=4}\sigma_1^2\sigma_2^2 + 3\phi_{4=4}\sigma_1^4 - 2\sigma_2^2\sigma_*^4 \right\rangle_\mathrm{g} \tag{133}$$

where

$$\phi_{2=4} := \mathsf{T}\mathsf{J}_{2=4}\mathsf{R} = \sum_{ij}(\partial_i\partial_j\phi)(\partial_i\partial_j\delta) - \frac{1}{3}\delta\nabla^2\delta \tag{134}$$

$$\phi_{4=4} := \mathsf{T}\mathsf{J}_{4=4}\mathsf{R} = \sum_{ij}(\partial_i\partial_j\delta)^2 - \frac{1}{3}(\nabla^2\delta)^2 \tag{135}$$

and $\sigma_*^2 = \sigma_0^2 - \sigma_1^4/\sigma_2^2$.

4.7. Estimators for third derivative terms

The distribution of third derivatives of the potential S is derived in Appendix B.4 and is given by a multivariate Gaussian with the generalized inverse covariance matrix

$$\mathsf{C}_\mathsf{S}^+ = \frac{3}{\sigma_1^2}\mathsf{J}_{3-3} + \frac{35}{2\sigma_1^2}\mathsf{J}_{3\equiv3} \tag{136}$$

Note that it is fine to consider the third spatial derivatives independently of the second derivatives, since the joint distribution factorizes

$$p(\mathsf{T}, \mathsf{S}) = p(\mathsf{T})p(\mathsf{S}) \tag{137}$$

We find bias estimators by evaluating equation (123) as:

$$b_{\mathsf{J}_{3-3}} = \frac{3}{\sigma_1^4}\left\langle \mathsf{S}\mathsf{J}_{3-3}\mathsf{S} - \sigma_1^2 \right\rangle_\mathrm{g}$$
$$= \frac{3}{\sigma_1^4}\left\langle (\nabla\delta)^2 - \sigma_1^2 \right\rangle_\mathrm{g} \tag{138}$$

$$b_{\mathsf{J}_{3\equiv3}} = \frac{35}{\sigma_1^4}\left\langle 5\mathsf{S}\mathsf{J}_{3\equiv3}\mathsf{S} - 2\sigma_1^2 \right\rangle_\mathrm{g}$$
$$= \frac{35}{\sigma_1^4}\left\langle 5\phi_{3\equiv3} - 2\sigma_1^2 \right\rangle_\mathrm{g} \tag{139}$$

where

$$\phi_{3\equiv3} = \sum_{ijk}(\partial_{ijk}\phi)^2 - \frac{3}{5}\sum_i(\partial_i\delta)^2 \tag{140}$$

and where we note again that conventions may differ by a factor 2 so that e.g. $b_{\mathsf{J}_{3-3}} = 2b_{(\nabla\delta)^2}$ if $b_{(\nabla\delta)^2}$ is the bias parameter that would multiply $(\nabla\delta^2)$ without any forefactor in a bias expansion.

5. Measurements of tensorial bias parameters

In this section we will use the estimators that we have derived in Section 4 to perform measurements of tensorial bias parameters. We will focus here on the tidal bias b_{K^2} and on the biases associated with third potential derivatives $b_{(\nabla\delta)^2} = \frac{1}{2}b_{\mathsf{J}_{3-3}}$ and $b_{\mathsf{J}_{3\equiv3}}$. For the tidal bias we have additionally derived the estimator with spatial corrections of order two, so that we can use it to validate that the presented theory indeed behaves as expected. Further, we will briefly check the effect of assembly bias on b_{K^2} which has previously been reported to be quite significant. For the measurements in this section we use the same simulations and jackknife technique as described in sections 3.1 and 3.2.
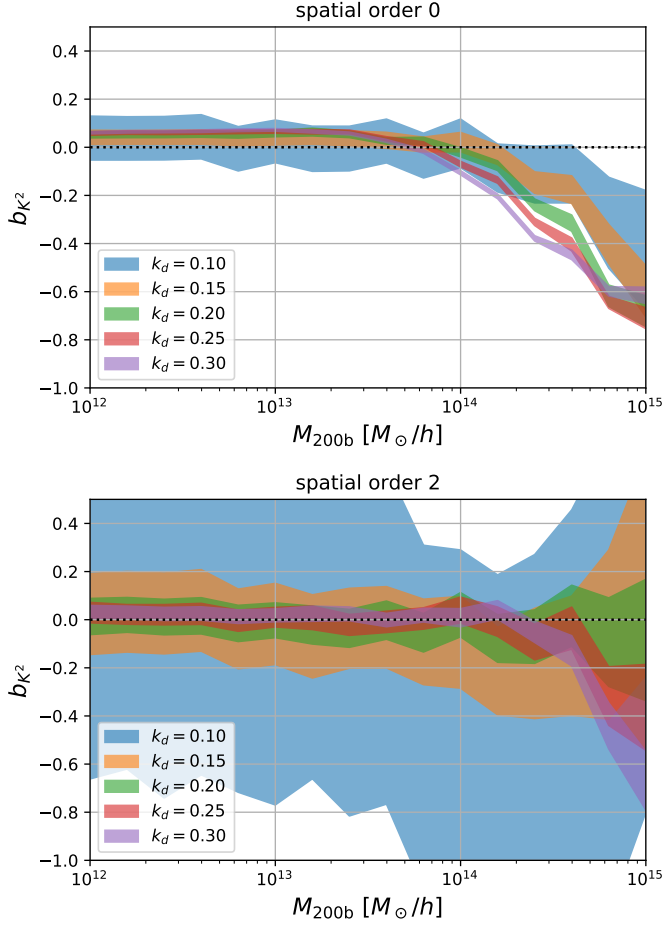
Fig. 11. Tidal bias as a function of mass and for different damping scales measured with the spatial order 0 estimator (top) versus the spatial order 2 estimator (bottom). The spatial order 2 estimator exhibits significantly reduced scale dependence at the price of increased errorbars.

5.1. Tidal bias

We measure the tidal bias $b_{K^2} = \frac{1}{2}b_{J_{2=2}}$ with the estimators from equation (131) at spatial order zero and equation (133) at spatial order two and show the results as a function of mass in Figure 11. The order zero estimator exhibits a strong scale dependence at scales $k_{\mathrm{d}} \gtrsim 0.15h\mathrm{Mpc}^{-1}$. However, overall it seems that the tidal bias is very small in amplitude and that a slightly negative Lagrangian tidal bias is preferred at high halo masses. On the other hand, the spatial order two estimator has a much lower scale dependence and seems mostly consistent across different damping scales, with maybe a slight inconsistency at very high halo masses and large k_{d}. This comes at the price of a significantly increased uncertainty in the measured bias values. However, this demonstrates that the estimators that we have derived in Section 4 behave indeed in the way that we would expect. More precise measurements of the tidal bias could easily be obtained by employing a larger simulation volume or by averaging over several realizations.

In Figure 12 we show the coevolution relation between b_{K^2} and b_1. Here, we show for each estimator only the smallest two scales that appear reasonably unbiased in Figure 11. These are $k_{\mathrm{d}} = 0.1h\mathrm{Mpc}^{-1}$ and $k_{\mathrm{d}} = 0.15h\mathrm{Mpc}^{-1}$ for the spatial order zero estimator and $k_{\mathrm{d}} = 0.25h\mathrm{Mpc}^{-1}$ and $k_{\mathrm{d}} = 0.3h\mathrm{Mpc}^{-1}$ for the spatial order two estimator. Noteworthy, the estimators of different orders agree well with each other. For the larger k_{d} values the
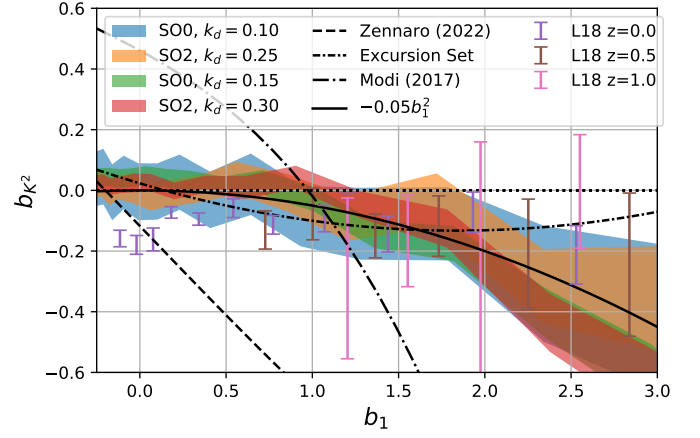

Fig. 12. Coevolution relation b_{K^2} versus b_1. Coloured shaded regions are our measurements and errorbars show the measurements from Lazeyras & Schmidt (2018). Black lines include measurements from Zennaro et al. (2022) and Modi et al. (2017), an excursion set prediction from Sheth et al. (2013) and an approximation that we suggest $b_{K^2} = -0.05b_1^2$.

estimators have smaller errorbars, but appear slightly biased at large values of $b_1 \gtrsim 2$. A rough approximation to the observed functional shape is given by

$$b_{K^2} = -0.05b_1^2 \tag{141}$$

though the uncertainty is of course quite high. However, in comparison to some previous literature estimates, we find very small values of b_{K^2}. The first measurement that we compare with is from Modi et al. (2017). Their relation is stated for Eulerian bias parameters $b_{K^2}^{\mathrm{E}}(b_1^{\mathrm{E}})$ and we convert it to a Lagrangian relation through $b_1 = b_1^{\mathrm{E}} - 1$ and $b_{K^2} = b_{K^2}^{\mathrm{E}} + 2/7b_1$ (Sheth et al. 2013; Desjacques et al. 2018). Similar to other studies (Lazeyras & Schmidt 2018; Zennaro et al. 2022), we find a strong disagreement between our measurements and the Modi et al. (2017) measurements. Further, we show as a dashed line in Figure 12 the coevolution relation of the tidal bias that was presented in Zennaro et al. (2022). This relation has been measured at very small scales $k \sim 0.7h\mathrm{Mpc}^{-1}$ and is not really expected to accurately represent the large-scale tidal bias. However, it is interesting to see that the large-scale tidal response that we measure here is so much weaker than the response that can be measured at small scales. Next, we compare with the excursion set prediction from Sheth et al. (2013) parameterized as $b_1^{\mathrm{E}} = 0.524 - 0.547b_1^{\mathrm{E}} + 0.046(b_1^{\mathrm{E}})^2$. This prediction seems to be in better agreement with our measurements, since it generally predicts smaller amplitudes for the tidal bias and a slight negative tendency. However, the predicted upturn in the function at $b_1 \gtrsim 2$ does not seem to be present in any of our measurements.

Finally, we compare our data with the measurements from Lazeyras & Schmidt (2018), again transformed from Eulerian to Lagrangian parameters, indicated through error-bars in Figure 12. Our measurements agree roughly with that study, in that we favor a very small tidal bias with a preferentially slightly negative amplitude. We suggest therefore, that high Lagrangian tidal biases $|b_{K^2}| > 0.5$ for mass-selected haloes are strongly disfavored for lowly biased objects $b_1 \ll 2$.

However, we also note that there are some statistically significant differences at $b_1 \sim 0$, where we find tidal bias consistent with zero, whereas the data of Lazeyras & Schmidt (2018) clearly favor a negative value of $b_{K^2} \sim -0.2$. Given the fact that
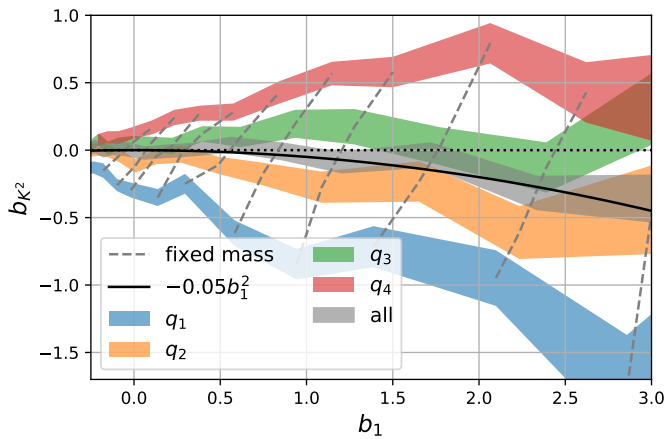
Fig. 13. Assembly bias in the Lagrangian $b_{K^2}(b_1)$ relation for haloes split into four quartiles by spin (at fixed mass). All lines use the spatial order 2 estimator and $k_d = 0.25h\mathrm{Mpc}^{-1}$ as damping scale. Dashed lines show the variation of b_1 and b_{K^2} at fixed mass. The $b_{K^2} - b_1$ relation shows a strong degree of assembly bias – with larger spin selections yielding larger b_{K^2} and larger b_1.

past erroneous measurements have rather predicted too large absolute values of the tidal bias rather than too small ones, and that the three lowest b_1 data points of Lazeyras & Schmidt (2018) do not seem really consistent with the overall behaviour of the curve, we suggest that our measurements are more reliable here and that the Lagrangian tidal bias becomes indeed very small $|b_{K^2}| \ll 0.1$ for $b_1 \lesssim 0.5$.

We conclude that overall the amplitude of the Lagrangian b_{K^2} is quite low for mass selected haloes. Given the challenges of the measurement and the history of errors, it would be desirable to have a clear measurement that does not suffer from the usual scale dependencies and large uncertainties. This could be achieved through anisotropic separate universe simulations as they have been presented in Stücker et al. (2021); Masaki et al. (2020); Akitsu et al. (2021). We checked whether the simulations that were presented in Stücker et al. (2021) are sufficient to measure the tidal bias at the level that we observe here. The simulation parameter choices were not optimal for measuring it, since those simulations do not have sufficient statistics of high mass objects – which are the objects that would have a significant tidal bias according to equation (141). However, a reliable measurement could be obtained by running with larger tidal fields that are tuned to measure the tidal bias at higher redshift, where most objects have a larger b_1 and likely a larger b_{K^2}. However, that said, we find that the simulations have significant enough statistics to constraint an upper bound $|b_{K^2}| < 0.1$ (roughly 1σ level) at low masses $M_{200b} \sim 10^{13}h^{-1}M_\odot$, consistent with our observations here.

5.2. Assembly bias in b_{K^2}

While the tidal bias seems to be quite low for mass selected haloes, it is not at all obvious that this should also be the case for galaxies. The formation and selection of galaxies may depend on other properties that are more sensitive to the tidal field, like the spin, the formation time and the halo concentration. For example, in tidal torque theory the spin of an object should depend on the Lagrangian tidal tensor (White 1984). Therefore, it is important to investigate whether there exists assembly bias with respect to such properties. Lazeyras et al. (2021) have shown

that the tidal bias $b_{K^2}^E$ and the associated $b_{K^2}(b_1)$ coevolution relation depend strongly on secondary halo properties. Here, we will briefly check whether we can reproduce these results.

We measure the spin of each halo in our simulation following Bullock et al. (2001) as

$$\lambda = \frac{|\boldsymbol{L}|}{\sqrt{2}v_{200c}r_{200c}} \qquad (142)$$

where

$$\boldsymbol{L} = \langle \boldsymbol{r} \times \boldsymbol{v} \rangle_{\mathrm{particles}} \qquad (143)$$

is the specific angular momentum averaged over all particles that are within the bound component of the main subhalo according to SUBFIND and r_{200c} is the radius within which the density is 200 times the critical density of the universe and v_{200c} the circular velocity at that radius.

We then group haloes in each mass bin by their spin into four quartiles q_1, q_2, q_3 and q_4, q_1 corresponding to the 25% haloes with lowest spin and q_4 to the 25% with highest spin and measure the bias parameters independently in each quartile q_1. We show the resulting $b_{K^2}(b_1)$ relation in Figure 13. Here, we present only results for the spatial order 2 estimator at a damping scale of $k_d = 0.25h\mathrm{Mpc}^{-1}$, but we have checked that other choices lead to the same results.

The $b_{K^2} - b_1$ relation exhibits a strong degree of assembly bias with halo spin. For example, at $b_1 \sim 1$ the tidal bias of mass selected haloes seems very small $|b_{K^2}| \ll 0.1$, but after the selection on halo spin, values may range between -1 and 0.5. In individual mass bins, the spin selection increases both the value of b_1 and the value of b_{K^2}.

The sign, the b_1 dependence and the overall amplitude of the spin-assembly bias are all in good agreement with the measurements of Lazeyras et al. (2021). For example, at $b_1 = 2$ we find assembly bias variations of $\Delta b_{K^2} \sim 2$ (between low and high spins) consistent with their variations at $b_1^E = 3$ of around $\Delta b_{K^2}^E \sim 2$ and the assembly bias decreases towards lower b_1 and increases towards higher b_1. Therefore, our measurements provide a validation of the results of that study, but it also shows that our method allows to reliably measure the tidal bias.

While we do not present it here, we have also tested whether we can reproduce the concentration assembly bias that was measured by Lazeyras et al. (2021). We have found that if we also define concentration through $v_{\max}/v_{200c}$, then we find an assembly bias of a similar amplitude. However, if concentrations are defined through fits of the halo density profile, the assembly bias is smaller (possibly consistent with 0) and prefers the opposite sign. The concentration assembly bias seems therefore to be a bit more uncertain than that of the spin.

We conclude that even if b_{K^2} seems quite small for mass selected haloes, the tidal bias may not so easily be neglected for tracers that follow more general selection criteria. Values of the Lagrangian b_{K^2} of order unity (and slightly larger) are a plausible possibility for realistic galaxy catalogues.

5.3. Biases of third derivative terms

In Figure 14 we show the measurements that we obtain by evaluating the bias estimators from equations (138) and (139) and where we again divide by a factor 2 to get $b_{(\nabla\delta)^2} = b_{J_{3-3}}/2$ in accordance with the literature convention.

The value of $b_{(\nabla\delta)^2}$ appears to depend strongly on the damping scale and we can therefore not claim a reliable measurement here. However, it seems mostly $b_{(\nabla\delta)^2} < 0$ which seems
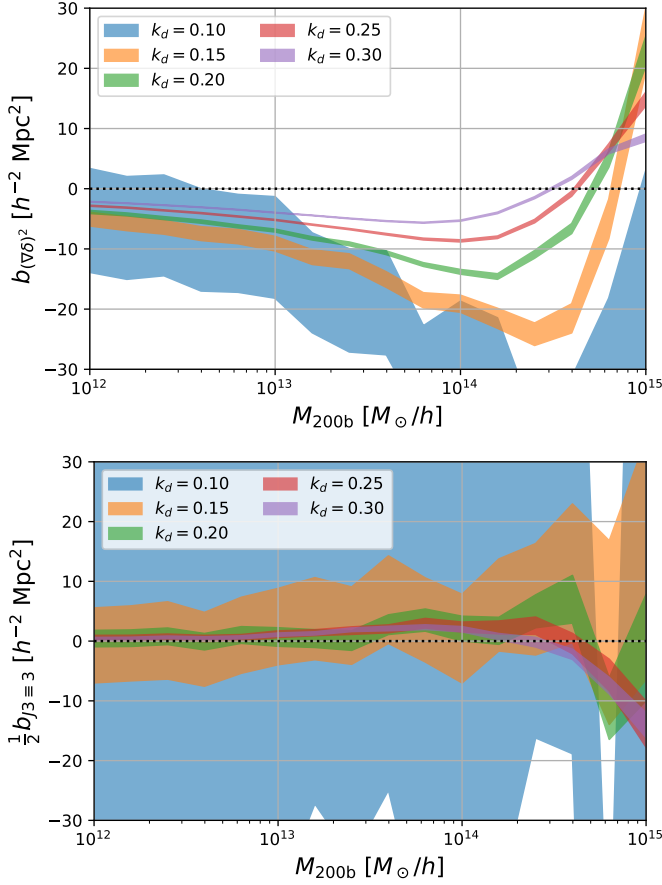
parameter	$10^{14} h^{-1} M_\odot$	$10^{15} h^{-1} M_\odot$	unit
b_1	0.54	3.08	
b_2	-0.89	7.20	
b_3	-1.93	7.81	
b_4	7.81	-32.30	
b_L	-13.92	17.76	$h^{-2}\text{Mpc}^2$
b_{K^2}	0.04	-0.17	
$b_{(\nabla\delta)^2}$	-13.81	22.87	$h^{-2}\text{Mpc}^2$
$\frac{1}{2} b_{3\equiv 3}$	2.10	-1.08	$h^{-2}\text{Mpc}^2$
β_2	-1.13	-2.25	
β_3	-0.33	-0.66	
β_4	0.19	-1.11	

Table 2. Two example sets of bias parameters which we consider in Figure 15. The first set corresponds to $M \sim 10^{14} h^{-1} M_\odot$ and has moderate bias values whereas the second set with $10^{15} h^{-1} M_\odot$ is strongly biased. The last two lines show cumulant bias parameters that can be used instead of b_3 and b_4.

Fig. 14. Measurements of biases terms of third derivatives of the potential. The bias of the density gradient $b_{(\nabla\delta)^2}$ (top) has a very strong scale dependence so that we cannot reliably measure it here. Bottom: the fully contracted third derivatives of the potential seem to have a very small associated bias parameter.

consistent with the idea that haloes should preferentially form at peaks where density gradients should be small. However, at very large masses $b_{(\nabla\delta)^2}$ seems to become positive. We are not quite sure how to interpret this behavior, but it might just be an artefact of the uncorrected scale dependencies. It is worth noting that the negative amplitude of $b_{(\nabla\delta)^2}$ is quite significant $b_{(\nabla\delta)^2} \sim 10 h^{-2}\text{Mpc}^2$ and even grows towards larger scales. We will therefore discuss later whether this (usually neglected) bias term is worth including under some circumstances. We note that $b_{(\nabla\delta)^2}$ has also been measured by Biagetti et al. (2014). However, since we have not properly corrected the scale dependence here, we abstain from a comparison.

On the other hand $b_{J_{3\equiv 3}}$ appears to be quite small $|b_{J_{3\equiv 3}}/2| \lesssim 5 h^{-2}\text{Mpc}^2$ for everything except the very high mass end. Therefore, the contributions of $b_{J_{3\equiv 3}}$ to the bias expansion should be significantly lower than $b_{(\nabla\delta)^2}$.

6. Relevance of bias parameters

We aim to estimate the relevance of different contributions to the bias expansion. If we have a term of the form $\frac{1}{n!} b_X \cdot X$ of the renormalized bias expansion, then we can estimate its relative relevance by comparing the typical difference in the predicted galaxy density field between a bias function f_X that includes the

term versus an expansion f_0 that neglects it

$$\sigma_{\epsilon,X}^2 = \left\langle (f_X - f_0)^2 \right\rangle$$
$$= \left(\frac{b_X}{n!} \right)^2 \left\langle X^2 \right\rangle \tag{144}$$

where the value of $\left\langle X^2 \right\rangle$ depends on the adopted damping scale and where $\sigma_{\epsilon,X}$ is a dimensionless quantity that can be compared across different bias terms. For example, for the contribution of b_3 we have

$$\sigma_{\epsilon,\delta^3}^2 = \left(\frac{b_3}{6} \right)^2 \left\langle (\delta^3 - 3\delta\sigma^2)^2 \right\rangle \tag{145}$$

The value of $\sigma_{\epsilon,X}$ quantifies the typical error that is made at the field level (in Lagrangian space) when neglecting the corresponding term.

We choose two halo sets $M \sim 10^{14} h^{-1} M_\odot$ and $M \sim 10^{15} h^{-1} M_\odot$ as examples and measure all their bias parameters with a damping scale $k_d = 0.2 h\text{Mpc}^{-1}$ and use the spatial order two estimators for all parameters (except $b_{(\nabla\delta)^2}$ and $b_{3\equiv 3}$). We list the corresponding parameters in Table 2. The first set represents a moderately biased population whereas the second set represents an example of strongly biased tracers. We then keep these parameters fixed and evaluate $\left\langle X^2 \right\rangle$ at many different damping scales to obtain $\sigma_{\epsilon,X}$ which we show in Figure 15.

For the moderately biased tracers the relevance of parameters decreases reasonably between different orders. However, this statement depends strongly on the considered damping scale. E.g. at $k_d = 0.4 h\text{Mpc}^{-1}$ many different terms reach similar amplitudes, whereas at $k_d = 0.05 h\text{Mpc}^{-1}$ there is a strong relevance order between different terms. If we aimed to fit a bias expansion at the field level at $k_d = 0.15 h\text{Mpc}^{-1}$ (marked as a vertical dotted line), then we would find that for the moderately biased set $b_1 > b_2 > b_3 > b_4$ all lie within a factor two of each other and for the strongly biased set we even have a changed order $b_2 > b_1 > b_4 > b_3$ and all of these lie within a factor of three with each other. This is of course very problematic, since it does not at all seem clear that higher order terms are smaller than lower order terms here, which casts doubts on the convergence of the expansion. Of course this problem can be controlled by simply going to larger smoothing scales, but at the cost of a reduced constraining power. However, another possibility is to make use of the advantages of the cumulant bias parameters. With these
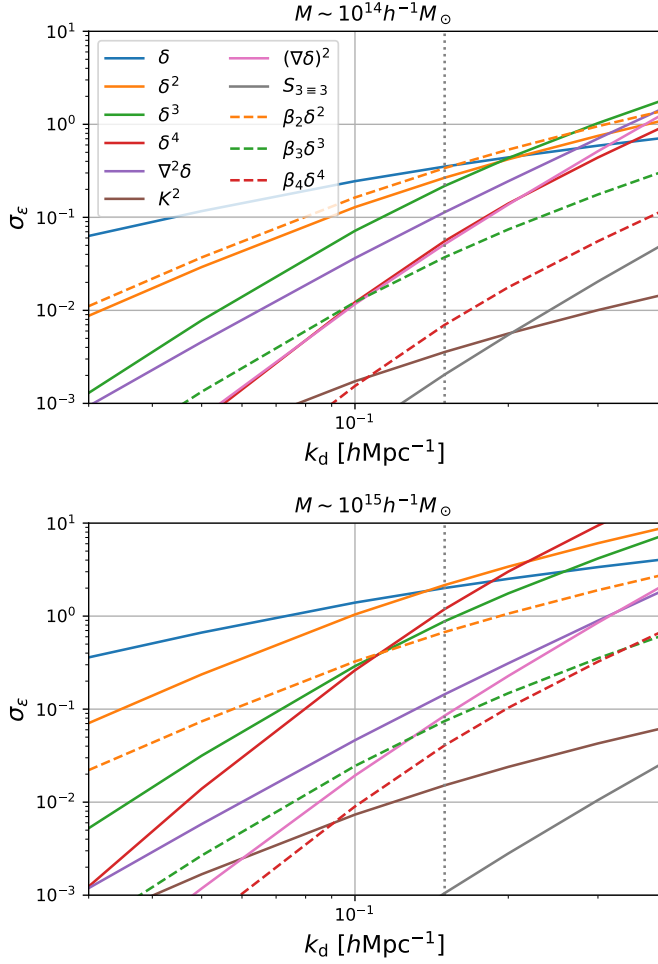
Fig. 15. Quantitative contribution of different bias terms to the bias expansion as a function of damping scale and for two different sets of tracers. To avoid cluttering, we omit the bias fore-factors of some terms in the legend. The vertical dotted line indicates an example scale of interest $k_d = 0.15h\text{Mpc}^{-1}$. The dashed lines show the contributions of cumulant biases at order three and four which can directly be compared to the contributions of b_3 and b_4 (labeled as δ^3 and δ^4).

parameters significant parts of higher order terms get already absorbed by lower order parameters. For example, consider the first three terms of the bias expansion in terms of cumulant parameters:

$$f = 1 + \beta_1\delta + \frac{1}{2}(\beta_2 + \beta_1^2)(\delta^2 - \sigma^2)$$
$$+ \frac{1}{6}(\beta_3 + 3\beta_1\beta_2 + \beta_1^3)(\delta^3 - \delta\sigma^2) + ... \quad (146)$$

Here β_1 and β_2 contribute also to the third term and to infinitely many higher order terms. If we write out the polynomial to high orders, but set $\beta_3 = 0$, then the leading order error we would make is

$$\sigma^2_{\epsilon,\beta_3\delta^3} = \left(\frac{\beta_3}{6}\right)^2 \left\langle (\delta^3 - 3\delta\sigma^2)^2 \right\rangle \quad (147)$$

which is identical to the error term from the canonical bias terms, but with a different fore-factor. Error estimates of these types are indicated as dashed lines in Figure 15. Strikingly these are significantly smaller than the canonical terms and it seems now plausibly to reach good convergence at $k_d = 0.15h\text{Mpc}^{-1}$. E.g.

to reach $\sigma_\epsilon \leq 0.1$ for the strongly biased set we would only need to consider the terms $(\beta_1, \beta_2, \beta_L)$.

It is note-worthy that the contribution of $b_{(\nabla\delta)^2}$ is not much smaller than that of b_L so that it might make sense to include this parameter more commonly in the bias expansion at scales $k_d \gtrsim 0.1h\text{Mpc}^{-1}$. However, we remind the reader that there are some other possibly important parameters here which we have not considered, so that we cannot present a comprehensive picture of the relevance of bias parameters.

We conclude that it may have significant advantages to phrase the bias expansion in terms of the cumulant parameters β_n instead of the canonical bias parameters – especially so when considering highly biased tracers or small smoothing scales where the rate of convergence of the canonical expansion is excruciatingly slow.

7. Conclusions

In this article we have presented a novel method to measure bias parameters through moments of the galaxy environment distribution $p(\delta|g)$. We have shown that such estimators can be derived both for scalar bias parameters like b_1 and b_L and for tensorial ones like b_{K^2} and $b_{(\nabla\delta)^2}$ and that they can be made scale independent by considering spatial corrections at various orders. We have verified the reliability of these estimators by recovering well established relations between known parameters and we have additionally used our estimators to measure so far unknown terms like $b_{3\equiv3}$ and $b_{(\nabla\delta)^2}$ – where the latter appears to be particularly large so that it might be worth to include it more commonly in the bias expansion.

The main benefits of our novel method are its simplicity and its generality. It is only required to evaluate the linear field of a simulation at the Lagrangian locations of a set of tracers – e.g. given by the most bound id in a halo catalogue – and to take simple expectation values. This measurement uses a minimal number of discretization steps – making it numerically robust and easy to understand. In contrast to many other approaches, the measurement of e.g. b_1 is completely independent of the assumed order of the expansion in powers of δ. This makes it generally unlikely that parameters numerically compensate for the absence of other terms. Finally, we note that all bias parameters can be evaluated at a low computational cost with a single simulation and that it is possible to use arbitrary filtering methods with the slight alterations explained in Appendix A.

However, to avoid confusion we want to point out a few limitations of our method for estimating biases. First of all, we note that our method assumes that the background distribution is known analytically. Therefore, it does not immediately translate to situations where this is not the case. For example, at third order in the bias expansion, time-derivatives of the tidal tensor appear which depend on the second Lagrangian perturbation theory potential $\phi^{(2)}$ which is not a Gaussian random field. Therefore, the formalism can only be translated to the associated bias parameters if the distribution of $\phi^{(2)}$ and its derivatives can be written (or at least adequately approximated) analytically. Further, the formalism cannot directly be applied to the Eulerian galaxy environment distribution $p^E(\delta|g)$, since the non-linear Eulerian matter distribution $p^E(\delta)$ does also not follow Gaussian statistics.

Beyond the novel bias estimators, we have also newly introduced the cumulant bias parameters

$$\beta_n = \frac{\partial^n}{\partial\delta_0^n} \log\left(\frac{n_g(\delta_0)}{n_{g,0}}\right)\Bigg|_{\delta_0=0} \quad (148)$$

These cumulant bias parameters are proportional to the cumulants of the galaxy environement distribution[4] and they relate to the canonical bias parameters exactly in the same way as cumulants relate to moments. When measuring these parameters, we have found them to be more independent from each other than canonical bias parameters. Further, we have shown that for haloes cumulant biases of order $n \geq 3$ are very close to zero. This has several intriguing consequences:

(1) The galaxy environment distribution and the bias function of haloes are very well approximated by a Gaussian. This motivates the usage of a Gaussian bias model at order $n = 2$ and it motivates e.g. to expand around a Gaussian bias model for orders $n > 2$. We will discuss the Gaussian bias model in detail in a companion paper and show that it has several desirable properties (Stücker et al., in prep).

(2) Coevolution relations of haloes at order $n \geq 3$ are equivalent to $\beta_n = 0$. Therefore, one may predict such relations between canonical bias parameters at very high orders by combining $\beta_n = 0$ with the mapping between canonical and cumulant bias parameters. However, this also suggests that such coevolution relations may not carry much physical significance, but are rather the artefacts of a suboptimal basis choice. An expansion around a Gaussian model would naturally incorporate all of these relations at any order.

(3) The bias expansion may have significantly improved convergence when phrased in terms of cumulants instead of canonical bias parameters. At the field level the convergence of the canonical expansion seems already questionable at scales $k_{\rm d} \gtrsim 0.15h{\rm Mpc}^{-1}$, since parameters may easily grow as their order grows. However, when phrased in terms of cumulant biases, we expect neglected higher order terms to be significantly less important, as they are already well captured through lower order parameters. Therefore, we expect cumulant biases to be most beneficial wherever the canonical expansion shows poor convergence – e.g. at high masses, at late times or at small smoothing scales.

---

[4] Except for β_2, which has an additional contribution.

Appendix A: Estimators with different filters

In Section 2, we have shown derivations under the assumption of a sharp k-space filter. For this choice all considered bias functions and bias parameters relate trivially to the 'separate universe' parameters that are commonly referred to in the literature. However, in general other filtering kernels could also be chosen. For example if we define the smoothed density as

$$\delta_l(k) = W(k)\delta(k) \tag{A.1}$$

where δ_l is the smoothed density and δ the unsmoothed one. If the filter W is for example chosen to be a Gaussian kernel

$$W_g(k) = \exp\left(-\frac{k^2}{2k_d^2}\right) \tag{A.2}$$

instead of a sharp k filter, then some considerations have to be adapted. In principle, all of the formula from Section 2 could still be applied, but they would refer to a different set of bias parameters that describe the response to perturbations of a different shape. Rather than introducing a new set of bias parameters, it is more convenient to directly write the commonly used parameters in terms of the new basis. Therefore, we will show here how to directly measure the 'separate universe' bias parameters from the moments of the galaxy environment distribution that was obtained with a density field that was filtered with any kernel.

First of all, we consider how a bias function on very large scales $f_l(\delta_l)$ should relate to the bias function on some much smaller scale $f_s(\delta_s)$. If the PBS is still valid on the smaller scale, then the probability of forming a galaxy, given knowledge of δ_s and δ_l depends only on the value of δ_s:

$$p(g|\delta_s, \delta_l) = p(g|\delta_s) . \tag{A.3}$$

Under this assumption, we can re-express the large scale bias function through the small scale function marginalized over the conditional probability of the small scale densities:

$$f_l(\delta_l) = \frac{p(g|\delta_l)}{p(g)} = \int \frac{p(g|\delta_s)}{p(g)} p(\delta_s|\delta_l) d\delta_s$$
$$= \int f_s(\delta_s) p(\delta_s|\delta_l) d\delta_s . \tag{A.4}$$

Since, the joint distribution of δ_s and δ_l is a multivariate Gaussian, the conditional distribution corresponds to a multivariate Gaussian that is conditioned on the value of one variable, which itself is a Gaussian with modified mean and covariance:

$$p(\delta_s|\delta_l) = N(\delta_s, \mu_*, \sigma_*) \tag{A.5}$$
$$\mu_* = \alpha\delta_l \tag{A.6}$$
$$\sigma_*^2 = \sigma_{ss}^2 - \alpha^2\sigma_{ll}^2 \tag{A.7}$$

where $\alpha = \sigma_{ls}^2/\sigma_{ll}^2$ corresponds to the correlation between smoothed large scale and small scale density field, $\sigma_{ll}^2 = \langle\delta_l\delta_l\rangle$, $\sigma_{ls}^2 = \langle\delta_l\delta_s\rangle$ and $\sigma_{ss} = \langle\delta_s\delta_s\rangle$.

If we define the renormalized bias function F_α through the bias function f_l as in Section 2, then we find

$$F_\alpha(\delta_0) := \int f_l(\delta_l) p(\delta_l - \delta_0) d\delta_l$$
$$= \int f_s(\delta_s) p_{ss}(\delta_{ss} - \alpha\delta_0) d\delta_s \tag{A.8}$$

where p_{ss} is simply the background distribution of the small scale densities. This can be shown either by evaluating the integrals (corresponding to convolutions) explicitly or by considering the large scale limit of f_l where $\sigma_* \to \sigma_{ss}$.

We can now relate the renormalized bias functions obtained with different filtering methods, since we may assume that f_s is independent of the filter employed on large scales. Therefore,

$$F_{\alpha_1}\left(\frac{\delta_0}{\alpha_1}\right) = F_{\alpha_2}\left(\frac{\delta_0}{\alpha_2}\right) \tag{A.9}$$

where we have indicated the usage of different filters through different α in the index. Therefore, different filters simply lead to a rescaling of the renormalized bias function. The value of α depends on the filter and the smoothing scale. If we assume that δ_s was obtained with some small scale filter W_s and δ_l with W_l, we have

$$\sigma_{ls}^2 = \left\langle \int \delta(\mathbf{k}_1) W_s(k_1) d^3\mathbf{k}_1 \int \delta(\mathbf{k}_2) W_l(k_2) d^3\mathbf{k}_2 \right\rangle$$
$$= \int\int \langle\delta(\mathbf{k}_1)\delta(\mathbf{k}_2)\rangle W_s(k_1) W_l(k_2) d^3\mathbf{k}_1 d^3\mathbf{k}_2$$
$$= \int\int P(k_1)\delta_D(\mathbf{k}_1 - \mathbf{k}_2) W_s(k_1) W_l(k_2) d^3\mathbf{k}_1 d^3\mathbf{k}_2$$
$$= \int P(k) W_s(k) W_l(k) d^3\mathbf{k} \tag{A.10}$$
$$\approx \int P(k) W_l(k) d^3\mathbf{k} \tag{A.11}$$

where the last approximation assumes that W_s is very close to 1 over the support of W_l, because W_l corresponds to a much larger scale smoothing than W_s. Further, we have

$$\sigma_{ll}^2 = \int P(k) W_l^2(k) d^3\mathbf{k} \tag{A.12}$$
$$\alpha = \frac{\int P(k) W_l(k) d^3\mathbf{k}}{\int P(k) W_l^2(k) d^3\mathbf{k}} . \tag{A.13}$$

It is easy to see that $\alpha = 1$ for any sharp Fourier space filter, independent of the filtering scale. However, for other filters it is important to note that α may be generally larger than 1, for example $\alpha \sim 2$ for a Gaussian filter at scales of $k_d \sim 0.1h\mathrm{Mpc}^{-1}$. Further, α depends generally on the filtering-scale. The renormalized bias function does therefore only correspond to a well-defined limit at arbitrary large scales if α is constant (as for the sharp k filter).

Since the case $\alpha = 1$ seems clearly to be the simplest and leads to a sensibly scale-independent renormalization, we consider it as the reference case and simply label its renormalized bias function F, and we have focused in the main-text on the discussion of this case. However, we can still use the relations derived in the main-text to measure the properties of F directly through measuring aspects of some filter-dependent function F_α.

We will directly show this for the general multivariate case here by adapting the considerations from Section 2.7. For the multivariate case the renormalized bias functions relate as

$$F_A(\mathbf{x}_0) = F(A\mathbf{x}_0) \tag{A.14}$$

where

$$A = C_{ls}C_{ll}^{-1} \tag{A.15}$$
$$C_{ls} = \langle\mathbf{x}_l \otimes \mathbf{x}_s\rangle \tag{A.16}$$
$$C_{ll} = \langle\mathbf{x}_l \otimes \mathbf{x}_l\rangle \tag{A.17}$$
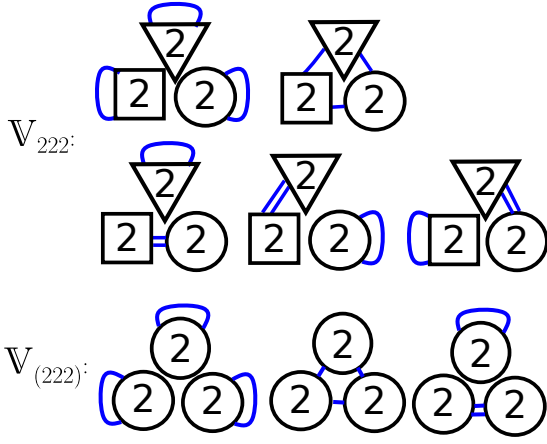
Fig. B.1. Illustration of the basis tensors of $\mathbb{V}_{222}$ versus those of $\mathbb{V}_{(222)}$. Top: without the symmetry requirement between groups (indicated by usage of different symbols for each group) there are five basis tensors. Bottom: With the additional symmetry requirement between groups (indicated by usage of the same symbol for each group), since the latter three tensors of $\mathbb{V}_{222}$ get all united to a single symmetrized term (bottom).

where C_{ls} may more conveniently be measured as $\langle \boldsymbol{x}_m \otimes \boldsymbol{x}_m \rangle$ where $\boldsymbol{x}_m$ is evaluated in the linear fields filtered with $\sqrt{W_l}$.

The cumulant-generating function of the galaxy environment distribution with δ_l is again given through equation (61), just with F_A instead of F:

$$K_\mathsf{A}(t) = \frac{1}{2} t^T \mathsf{C}_{ll} t + \log\left(F_\mathsf{A}(\mathsf{C}_{ll} t)\right)$$
$$= \frac{1}{2} t^T \mathsf{C}_{ll} t + \log\left(F(\mathsf{C}_{ls} t)\right) \tag{A.18}$$

For measuring the cumulant bias parameters, we may then define a more convenient variable

$$\boldsymbol{u}_* = \mathsf{C}_{ls} \boldsymbol{x}_0 \tag{A.19}$$

which has the cumulant generating function

$$K_{\boldsymbol{u}_*}(t) = \frac{1}{2} t^T \mathsf{C}_{ls}^{-1} \mathsf{C}_{ll} \mathsf{C}_{ls}^{-1} t + \log\left(F(t)\right) \tag{A.20}$$

so that the cumulant bias parameters can be inferred by differentiating, leading to:

$$\beta_{ijk\ldots} = \begin{cases} \kappa_{\boldsymbol{u}_*, ijk\ldots} & \text{if} \quad i+j+k+\ldots \neq 2 \\ \kappa_{\boldsymbol{u}_*, ijk\ldots} - C_{*,ab}^{-1} & \text{if} \quad i+j+k+\ldots = 2 \end{cases} \tag{A.21}$$

where $\mathsf{C}_*^{-1} = \mathsf{C}_{ls}^{-1} \mathsf{C}_{ll} \mathsf{C}_{ls}^{-1}$. We can therefore measure the separate universe bias parameters with an arbitrarily filtered distribution in a very similar manner to the one using a sharp k-space filter. It is worth noting that the inferred relations show also an important property: If β_n vanish for $n > 2$, that automatically implies that the corresponding cumulants vanish for any filtered galaxy environment distribution. Therefore, if the bias function is Gaussian for one filter, then it will also appear Gaussian for any other filter.

Appendix B: Isotropic Tensors

Appendix B.1: Tensors with three or more symmetry groups

In the main text we have only explained, how to obtain basis tensors for tensor spaces with at most two symmetry groups.

However, for higher order biases like e.g. the equivalent of b_3 the isotropic tensor decomposition requires consideration of additional symmetries. We define $\mathbb{V}_{222}$ as the space of tensors that are symmetric under permutations of the first two, second two and third two indices. Further, we define $\mathbb{V}_{(222)}$ as the space of tensors that are additionally invariant to exchanges of these symmetry groups. These spaces have different number of basis tensors

$$\mathbb{V}_{222} = \text{span}\left(\{\mathsf{I}_{222}, \mathsf{I}_{2-2-2-}, \mathsf{I}_{2=22}, \mathsf{I}_{22=2}, \mathsf{I}_{222=}\}\right) \tag{B.1}$$
$$\mathbb{V}_{(222)} = \text{span}\left(\{\mathsf{I}_{222}, \mathsf{I}_{2-2-2-}, \mathsf{I}_{(2=22)}\}\right) \tag{B.2}$$

where

$$\mathsf{I}_{(2=22)} = \frac{1}{3}(\mathsf{I}_{2=22} + \mathsf{I}_{22=2} + \mathsf{I}_{222=}) \tag{B.3}$$

is additionally symmetrized between exchanges between the symmetry groups and the other tensors are given by

$$I_{222, ijklmn} = S_{222}(\delta_{ij} \delta_{kl} \delta_{mn}) \tag{B.4}$$
$$I_{2-2-2-, ijklmn} = S_{222}(\delta_{ik} \delta_{lm} \delta_{nj}) \tag{B.5}$$
$$I_{2=22, ijklmn} = S_{222}(\delta_{ik} \delta_{jl} \delta_{mn}) \tag{B.6}$$
$$I_{22=2, ijklmn} = S_{222}(\delta_{ij} \delta_{km} \delta_{ln}) \tag{B.7}$$
$$I_{222=, ijklmn} = S_{222}(\delta_{im} \delta_{kl} \delta_{jn}) \tag{B.8}$$

The basis tensors of $\mathbb{V}_{(222)}$ are visualized through a diagram in Figure B.1. The third derivative tensor of that would appear in the bias expansion can be decomposed in this space

$$\left. \frac{\partial^3 F}{\partial \mathsf{T} \partial \mathsf{T} \partial \mathsf{T}} \right|_{\mathsf{T}=0} \in \mathbb{V}_{(222)} \tag{B.9}$$

Appendix B.2: Covariance of potential derivatives

To find an explicit form for the distribution of the derivatives of the tidal tensor and other higher order derivatives of the potential, we need to find their covariance matrix and invert it.

For this we first consider the covariance between any derivatives of the potential. This is given by

$$C_{ij\ldots,ab\ldots} = \langle \phi_{ij\ldots} \phi_{ab\ldots} \rangle$$
$$= \frac{1}{(2\pi)^6} \left\langle \int \int \frac{(ik_i)(ik_j)\ldots(ik_a)(ik_b)\ldots}{k^4} \right.$$
$$\left. e^{-i(\boldsymbol{k_1}-\boldsymbol{k_2})\boldsymbol{x}} \delta_k^*(\boldsymbol{k_1}) \delta_k(\boldsymbol{k_2}) \mathrm{d}\boldsymbol{k_1} \mathrm{d}\boldsymbol{k_2} \right\rangle \tag{B.10}$$

where δ_k is the Fourier transform of density field δ and we have used that δ is a real field so that $\delta_k(\boldsymbol{k}) = \delta_k^*(-\boldsymbol{k})$. Using the definition of the power spectrum

$$\langle \delta_k^*(\boldsymbol{k_1}) \delta_k(\boldsymbol{k_2}) \rangle = (2\pi^3) P(k) \delta_\mathrm{D}^3(\boldsymbol{k_1} - \boldsymbol{k_2}) \tag{B.11}$$

where δ_D is the Dirac delta function, we can simplify

$$C_{ij\ldots,ab\ldots} = \frac{1}{(2\pi)^3} \int \frac{(ik_i)(ik_j)\ldots(ik_a)(ik_b)\ldots}{k^4} P(k) \mathrm{d}^3 \boldsymbol{k} \tag{B.12}$$

We can see that this expression does not depend on the order of the indices nor whether they stem from the first or the second potential term. Therefore, we just count the total number n_1 of occurrences of k_x, n_2 of k_y and n_3 of k_z. Without loss of generality we can order the coordinates so that $n_1 \geq n_2 \geq n_3$. Further, we

Example Terms	n_1	n_2	n_3	Value
$\langle \phi_{11}^2 \rangle$	4	0	0	$\frac{\sigma_0^2}{5}$
$\langle \phi_{11}\phi_{22} \rangle$, $\langle \phi_{12}^2 \rangle$	2	2	0	$\frac{\sigma_0^2}{15}$
$\langle \phi_{1111}\phi_{11} \rangle$, $\langle \phi_{111}^2 \rangle$	6	0	0	$\frac{\sigma_1^2}{7}$
$\langle \phi_{1111}\phi_{22} \rangle$, $\langle \phi_{1112}\phi_{12} \rangle$, $\langle \phi_{112}^2 \rangle$	4	2	0	$\frac{\sigma_1^2}{35}$
$\langle \phi_{1123}\phi_{23} \rangle$, $\langle \phi_{1122}\phi_{33} \rangle$, $\langle \phi_{123}^2 \rangle$	2	2	2	$\frac{\sigma_1^2}{105}$
$\langle \phi_{11111}\phi_{111} \rangle$, $\langle \phi_{1111}^2 \rangle$	8	0	0	$\frac{\sigma_2^2}{9}$
$\langle \phi_{11111}\phi_{122} \rangle$, $\langle \phi_{1112}^2 \rangle$	6	2	0	$\frac{\sigma_2^2}{63}$
$\langle \phi_{11112}\phi_{222} \rangle$, $\langle \phi_{1111}\phi_{2222} \rangle$, $\langle \phi_{1122}^2 \rangle$	4	4	0	$\frac{\sigma_2^2}{105}$
$\langle \phi_{11112}\phi_{233} \rangle$, $\langle \phi_{1112}\phi_{1233} \rangle$, $\langle \phi_{1123}^2 \rangle$	4	2	2	$\frac{\sigma_2^2}{315}$

Table B.1. Covariances of various derivatives of the potential. The first column is an (incomplete) list of examples to give an idea of what kind of terms fall under the same combination of n_1, n_2 and n_3. All uneven terms are omitted, because they are zero and terms containing 0th and first derivatives are omitted, because the are not Galilei-invariant.

can switch to spherical coordinates and replace the cosine of the azimutal variable:

$$\boldsymbol{k} = \begin{pmatrix} k\mu \\ k(1-\mu)^2 \cos\phi \\ k(1-\mu)^2 \sin\phi)^T \end{pmatrix} \tag{B.13}$$

so that we have

$$\Sigma(ij...ab) = \frac{\sigma_M^2}{4\pi} \int_{-1}^{1} \mu^{n_1} d\mu \int_0^{2\pi} \cos^{n_2}(\phi)\sin^{n_3}(\phi)d\phi \tag{B.14}$$

$$\sigma_M^2 = \frac{1}{2\pi^2} \int P(k)k^{2M}dk \tag{B.15}$$

$$M = \frac{n_1 + n_2 + n_3 - 4}{2} . \tag{B.16}$$

These integrals can easily be evaluated by hand or with a computer algebra system. We list a few examples of the covariances in Table B.1 Note that only terms where each n_1, n_2 and n_3 are even are non-zero. Note that all uneven terms are zero, so that e.g. the 2 second derivatives of the potential have zero covariance with the third derivatives, but a non-zero covariance with the fourth derivatives.

Appendix B.3: The distribution of the tidal tensor

From equation (B.12) it can be seen that the covariance tensor of any derivatives of the potential has to be an isotropic tensor with full symmetry in all indices. Therefore, the covariance matrix of the tidal tensor is $C_T \in \mathbb{V}_4$ so that it has to be proportional to I_4 with a coefficient that we can read off from table B.1:

$$C_T = \langle T \otimes T \rangle$$

$$= \frac{\sigma^2}{5} J_4 . \tag{B.17}$$

This tensor is not invertible, since e.g. the T_{12} component is perfectly degenerate with the T_{21} component. However, in such cases the distribution of T can still be inferred by constructing a generalized inverse C_T^+ which has the property

$$CC_T^+C = C_T \tag{B.18}$$

It then holds

$$p(T) = N \exp\left(-\frac{1}{2}T^T C_T^+ T\right) \tag{B.19}$$

To find the generalized inverse we may assume that $C_T^+ \in \mathbb{V}_{22}$ so that we can make the Ansatz

$$C_T^+ = A_{22}J_{22} + A_{2=2}J_{2=2} \tag{B.20}$$

We can use the decomposition $J_4 = \frac{5}{9}J_{22} + \frac{2}{3}J_{2=2}$ (determined computationally) and evaluate equation (B.18) term by term:

$$C_T C_T^+ = A_{22}\frac{\sigma^2}{9}(J_{22} {}^{(2)} J_{22}) + A_{2=2}\frac{2\sigma^2}{15}(J_{2=2} {}^{(2)} J_{2=2}) \tag{B.21}$$

$$= 3A_{22}\frac{\sigma^2}{9}J_{22} + A_{2=2}\frac{2\sigma^2}{15}J_{2=2} \tag{B.22}$$

$$C_T C_T^+ C_T = 9A_{22}\left(\frac{\sigma^2}{9}\right)^2 J_{22} + A_{2=2}\left(\frac{2\sigma^2}{15}\right)^2 J_{2=2} \tag{B.23}$$

$$\equiv C_T = \frac{\sigma^2}{9}J_{22} + \frac{2\sigma^2}{15}J_{2=2} \tag{B.24}$$

where we have evaluated the appearing tensor products computationally. By identifying coefficients, we find $A_{22} = \frac{1}{\sigma^2}$ and $A_{2=2} = \frac{15}{2\sigma^2}$ and therefore have the generalized inverse of the covariance matrix:

$$C_T^+ = \frac{1}{\sigma^2}J_{22} + \frac{15}{2\sigma^2}J_{2=2} \tag{B.25}$$

and the distribution of the tidal tensor is given by

$$p(T) = N \exp\left(-\frac{2T^T J_{22}T + 15T^T J_{2=2}T}{4\sigma^2}\right) \tag{B.26}$$

$$= N \exp\left(-\frac{2\delta^2 + 15K^2}{4\sigma^2}\right) . \tag{B.27}$$

Appendix B.4: The distribution of third derivatives

The third derivative tensor of the potential S has the covariance matrix

$$C_S = \langle S \otimes S \rangle \tag{B.28}$$

$$= \frac{\sigma_1^2}{7}J_6 \tag{B.29}$$

$$= \frac{3\sigma_1^2}{25}J_{3-3} + \frac{2\sigma_1^2}{35}J_{3\equiv3} . \tag{B.30}$$

Analogously to the calculation in Section B.3, we find the generalized inverse

$$C_S^+ = \frac{3}{\sigma_1^2}J_{3-3} + \frac{35}{2\sigma_1^2}J_{3\equiv3} \tag{B.31}$$

where we have again used numerical representations to evaluate products between isotropic tensors when evaluating equation (B.18). Therefore, the distribution of S is given by

$$p(S) = N \exp\left(-\frac{35S J_{3\equiv3}S - 15S J_{3-3}S}{4\sigma_1^2}\right)$$

$$=: N \exp\left(-\frac{35S_{3\equiv3} - 15S_{3-3}}{4\sigma_1^2}\right) . \tag{B.32}$$

where in the last line we have defined some new scalar variables through the given contractions, where $S_{3-3} = (\nabla\delta)^2$.

Appendix B.5: Joint Distribution of Second and Fourth Derivatives

We can write the covariance matrix of fourth and second derivatives as a block matrix

$$\mathsf{C}_{\mathsf{TR}} = \begin{pmatrix} \langle \mathsf{T} \otimes \mathsf{T} \rangle & \langle \mathsf{T} \otimes \mathsf{R} \rangle \\ \langle \mathsf{R} \otimes \mathsf{T} \rangle & \langle \mathsf{R} \otimes \mathsf{R} \rangle \end{pmatrix} \tag{B.33}$$

$$= \begin{pmatrix} \frac{\sigma_2^2}{5}\mathsf{J}_4 & \frac{\sigma_1^2}{7}\mathsf{J}_6 \\ \frac{\sigma_1^2}{7}\mathsf{J}_6 & \frac{\sigma_2^2}{9}\mathsf{J}_8 \end{pmatrix} \tag{B.34}$$

$$= \begin{pmatrix} \frac{\sigma_0^2}{9}\mathsf{J}_{22} + \frac{2\sigma_0^2}{15}\mathsf{J}_{2=2} & \frac{\sigma_1^2}{15}\mathsf{J}_{24} + \frac{4\sigma_1^2}{35}\mathsf{J}_{2=4} \\ \frac{\sigma_1^2}{15}\mathsf{J}_{42} + \frac{4\sigma_1^2}{35}\mathsf{J}_{4=2} & \frac{\sigma_2^2}{25}\mathsf{J}_{44} + \frac{24\sigma_2^2}{245}\mathsf{J}_{4=4} + \frac{8\sigma_2^2}{315}\mathsf{J}_{4==4} \end{pmatrix} . \tag{B.35}$$

For its inverse we make the Ansatz

$$\mathsf{C}_{\mathsf{TR}}^+ = \begin{pmatrix} A\mathsf{J}_{22} + B\mathsf{J}_{2=2} & C\mathsf{J}_{24} + D\mathsf{J}_{2=4} \\ C\mathsf{J}_{42} + D\mathsf{J}_{4=2} & E\mathsf{J}_{44} + F\mathsf{J}_{4=4} + G\mathsf{J}_{4==4} \end{pmatrix} \tag{B.36}$$

and we find the solution to its generalized inverse by solving

$$\mathsf{C}_{\mathsf{TR}} = \mathsf{C}_{\mathsf{TR}}\mathsf{C}_{\mathsf{TR}}^+\mathsf{C}_{\mathsf{TR}} \tag{B.37}$$

with a computer algebra system and find

$$\mathsf{C}_{\mathsf{TR}}^+ = \begin{bmatrix} \frac{15\sigma_2^2\mathsf{J}_{2=2}}{2\sigma_0^2\sigma_2^2 - 2\sigma_1^4} + \frac{\sigma_2^2\mathsf{J}_{22}}{\sigma_0^2\sigma_2^2 - \sigma_1^4} & -\frac{15\sigma_1^2\mathsf{J}_{2=4}}{2\sigma_0^2\sigma_2^2 - 2\sigma_1^4} - \frac{\sigma_1^2\mathsf{J}_{24}}{\sigma_0^2\sigma_2^2 - \sigma_1^4} \\ -\frac{15\sigma_1^2\mathsf{J}_{4=2}}{2\sigma_0^2\sigma_2^2 - 2\sigma_1^4} - \frac{\sigma_1^2\mathsf{J}_{42}}{\sigma_0^2\sigma_2^2 - \sigma_1^4} & \frac{15\sigma_0^2\mathsf{J}_{4=4}}{2\sigma_0^2\sigma_2^2 - 2\sigma_1^4} + \frac{\sigma_0^2\mathsf{J}_{44}}{\sigma_0^2\sigma_2^2 - \sigma_1^4} + \frac{315\mathsf{J}_{4==4}}{8\sigma_2^2} \end{bmatrix} \tag{B.38}$$

The joint distribution of second and fourth derivatives of the potential can be written as

$$p(\mathsf{T}, \mathsf{R}) = N \exp\left(-\frac{1}{2}\begin{pmatrix} \mathsf{T} & \mathsf{R} \end{pmatrix}\mathsf{C}_{\mathsf{TR}}^+\begin{pmatrix} \mathsf{T} \\ \mathsf{R} \end{pmatrix}\right) . \tag{B.39}$$

Appendix B.6: Tidal estimator with spatial correction

We find the tidal estimator with spatial corrections of order two as

$$b_{\mathsf{J}_{2=2}} = \left\langle \frac{1}{p(\mathsf{T}, \mathsf{R})} \frac{\partial^2 p(\mathsf{T}, \mathsf{R})}{\partial \mathsf{T} \partial \mathsf{T}} \overset{(4)}{\cdot} \frac{\mathsf{J}_{2=2}}{\|\mathsf{J}_{2=2}\|^2} \right\rangle_{\mathsf{g}} \tag{B.40}$$

$$= \frac{15}{4\sigma_*^8} \left\langle 3K^2\sigma_2^4 + 6\phi_{2=4}\sigma_1^2\sigma_2^2 + 3\phi_{4=4}\sigma_1^4 - 2\sigma_2^2\sigma_*^4 \right\rangle_{\mathsf{g}} \tag{B.41}$$

where

$$\phi_{2=4} := \mathsf{T}\mathsf{J}_{2=4}\mathsf{R} = \sum_{ij}\partial_i\partial_j\phi\partial_i\partial_j\delta - \frac{1}{3}\delta\nabla^2\delta \tag{B.42}$$

$$\phi_{4=4} := \mathsf{T}\mathsf{J}_{4=4}\mathsf{R} = \sum_{ij}(\partial_i\partial_j\delta)^2 - \frac{1}{3}(\nabla^2\delta)^2 \tag{B.43}$$

While we do not list them here, we have verified that if we analogously evaluate the expressions for b_{J_2}, b_{J_4}, $b_{\mathsf{J}_{22}}$ and $b_{\mathsf{J}_{24}}$ that they are identical to the estimators for b_1, b_L, b_2 and $b_{\delta,L}$ respectively from equations (48) - (52).

References

Akitsu, K., Li, Y., & Okumura, T. 2021, J. Cosmology Astropart. Phys., 2021, 041

Alam, S., Aubert, M., Avila, S., et al. 2021, Phys. Rev. D, 103, 083533

Amendola, L., Appleby, S., Avgoustidis, A., et al. 2018, Living Reviews in Relativity, 21, 2

Angulo, R. E. & Hahn, O. 2022, Living Reviews in Computational Astrophysics, 8, 1

Angulo, R. E., Zennaro, M., Contreras, S., et al. 2021, MNRAS, 507, 5869

Baldauf, T., Mirbabayi, M., Simonović, M., & Zaldarriaga, M. 2016, arXiv e-prints, arXiv:1602.00674

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

Barreira, A. 2020, J. Cosmology Astropart. Phys., 2020, 031

Barreira, A., Lazeyras, T., & Schmidt, F. 2021, J. Cosmology Astropart. Phys., 2021, 029

Baumann, D., Nicolis, A., Senatore, L., & Zaldarriaga, M. 2012, J. Cosmology Astropart. Phys., 2012, 051

Berlind, A. A. & Weinberg, D. H. 2002, ApJ, 575, 587

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

Biagetti, M., Chan, K. C., Desjacques, V., & Paranjape, A. 2014, MNRAS, 441, 1457

Bullock, J. S., Dekel, A., Kolatt, T. S., et al. 2001, ApJ, 555, 240

Cacciato, M., Lahav, O., van den Bosch, F. C., Hoekstra, H., & Dekel, A. 2012, MNRAS, 426, 566

Chaves-Montero, J., Angulo, R. E., & Contreras, S. 2023, MNRAS, 521, 937

Chaves-Montero, J., Angulo, R. E., Schaye, J., et al. 2016, MNRAS, 460, 3100

Chen, S.-F., Vlah, Z., & White, M. 2020, J. Cosmology Astropart. Phys., 2020, 062

Colas, T., d'Amico, G., Senatore, L., Zhang, P., & Beutler, F. 2020, J. Cosmology Astropart. Phys., 2020, 001

Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, ApJ, 647, 201

Contreras, S., Angulo, R. E., & Zennaro, M. 2021, MNRAS, 508, 175

Contreras, S., Chaves-Montero, J., & Angulo, R. E. 2023, MNRAS, 525, 3149

Croton, D. J., Gao, L., & White, S. D. M. 2007, MNRAS, 374, 1303

Croton, D. J., Stevens, A. R. H., Tonini, C., et al. 2016, ApJS, 222, 22

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

d'Amico, G., Gleyzes, J., Kokron, N., et al. 2020, J. Cosmology Astropart. Phys., 2020, 005

Davé, R., Anglés-Alcázar, D., Narayanan, D., et al. 2019, MNRAS, 486, 2827

DeRose, J., Kokron, N., Banerjee, A., et al. 2023, J. Cosmology Astropart. Phys., 2023, 054

Desjacques, V., Jeong, D., & Schmidt, F. 2018, Phys. Rep., 733, 1

Dragomir, R., Rodríguez-Puebla, A., Primack, J. R., & Lee, C. T. 2018, MNRAS, 476, 741

Dubois, Y., Pichon, C., Welker, C., et al. 2014, MNRAS, 444, 1453

Faltenbacher, A. & White, S. D. M. 2010, ApJ, 708, 469

Ferreras, I., Hopkins, A. M., Lagos, C., et al. 2019, MNRAS, 487, 435

Frenk, C. S., White, S. D. M., Davis, M., & Efstathiou, G. 1988, ApJ, 327, 507

Gao, L., Springel, V., & White, S. D. M. 2005, MNRAS, 363, L66

Gao, L. & White, S. D. M. 2007, MNRAS, 377, L5

Genel, S., Bryan, G. L., Springel, V., et al. 2019, ApJ, 871, 21

Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, MNRAS, 451, 2663

Ivanov, M. M., Simonović, M., & Zaldarriaga, M. 2020, J. Cosmology Astropart. Phys., 2020, 042

Kaiser, N. 1984, ApJ, 284, L9

Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, MNRAS, 303, 188

Kokron, N., DeRose, J., Chen, S.-F., White, M., & Wechsler, R. H. 2021, MNRAS, 505, 1422

Lacey, C. G., Baugh, C. M., Frenk, C. S., et al. 2016, MNRAS, 462, 3854

Lagos, C. d. P., Tobar, R. J., Robotham, A. S. G., et al. 2018, MNRAS, 481, 3573

Lazeyras, T., Barreira, A., & Schmidt, F. 2021, J. Cosmology Astropart. Phys., 2021, 063

Lazeyras, T. & Schmidt, F. 2018, J. Cosmology Astropart. Phys., 2018, 008

Lazeyras, T. & Schmidt, F. 2019, J. Cosmology Astropart. Phys., 2019, 041

Lazeyras, T., Wagner, C., Baldauf, T., & Schmidt, F. 2016, J. Cosmology Astropart. Phys., 2016, 018

Lehmann, B. V., Mao, Y.-Y., Becker, M. R., Skillman, S. W., & Wechsler, R. H. 2017, ApJ, 834, 37

Levi, M., Bebek, C., Beers, T., et al. 2013, ArXiv e-prints [arXiv:1308.0847]

Li, Y., Hu, W., & Takada, M. 2014, Phys. Rev. D, 89, 083519

Lukacs, E. 1970, Characteristic Functions (2nd Edition) (Griffin)

Masaki, S., Nishimichi, T., & Takada, M. 2020, MNRAS, 496, 483

Meurer, A., Smith, C. P., Paprocki, M., et al. 2017, PeerJ Computer Science, 3, e103

Mo, H. J. & White, S. D. M. 1996, MNRAS, 282, 347

Modi, C., Castorina, E., & Seljak, U. 2017, MNRAS, 472, 3959

Modi, C., Chen, S.-F., & White, M. 2020, MNRAS, 492, 5754

Montero-Dorta, A. D., Pérez, E., Prada, F., et al. 2017, ApJ, 848, L2

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

Nishimichi, T., D'Amico, G., Ivanov, M. M., et al. 2020, Phys. Rev. D, 102, 123541

Ortega-Martinez, S., Contreras, S., & Angulo, R. 2024, arXiv e-prints, arXiv:2401.17374

Paranjape, A., Hahn, O., & Sheth, R. K. 2018, MNRAS, 476, 3631

Paranjape, A., Sefusatti, E., Chan, K. C., et al. 2013a, MNRAS, 436, 449

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

Paranjape, A., Sheth, R. K., & Desjacques, V. 2013b, MNRAS, 431, 1503

Peacock, J. A. & Smith, R. E. 2000, MNRAS, 318, 1144

Pellejero Ibañez, M., Angulo, R. E., Zennaro, M., et al. 2023, MNRAS, 520, 3725

Pellejero Ibañez, M., Stücker, J., Angulo, R. E., et al. 2022, MNRAS, 514, 3993

Philcox, O. H. E. & Ivanov, M. M. 2022, Phys. Rev. D, 105, 043517

Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6

Reddick, R. M., Wechsler, R. H., Tinker, J. L., & Behroozi, P. S. 2013, ApJ, 771, 30

Salcedo, A. N., Weinberg, D. H., Wu, H.-Y., & Wibking, B. D. 2022a, MNRAS, 510, 5376

Salcedo, A. N., Zu, Y., Zhang, Y., et al. 2022b, Science China Physics, Mechanics, and Astronomy, 65, 109811

Sato-Polito, G., Montero-Dorta, A. D., Abramo, L. R., Prada, F., & Klypin, A. 2019, MNRAS, 487, 1570

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

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

Stevens, A. R. H., Croton, D. J., & Mutch, S. J. 2016, MNRAS, 461, 859

Stücker, J., Schmidt, A. S., White, S. D. M., Schmidt, F., & Hahn, O. 2021, MNRAS, 503, 1473

Szalay, A. S. 1988, ApJ, 333, 21

Tinker, J. L., Robertson, B. E., Kravtsov, A. V., et al. 2010, ApJ, 724, 878

Tucci, B., Montero-Dorta, A. D., Abramo, L. R., Sato-Polito, G., & Artale, M. C. 2021, MNRAS, 500, 2777

Villaescusa-Navarro, F., Anglés-Alcázar, D., Genel, S., et al. 2021, ApJ, 915, 71

Vlah, Z., Castorina, E., & White, M. 2016, J. Cosmology Astropart. Phys., 2016, 007

Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518

Wagner, C., Schmidt, F., Chiang, C. T., & Komatsu, E. 2015, MNRAS, 448, L11

Wechsler, R. H., Zentner, A. R., Bullock, J. S., Kravtsov, A. V., & Allgood, B. 2006, ApJ, 652, 71

White, S. D. M. 1979, MNRAS, 186, 145

White, S. D. M. 1984, ApJ, 286, 38

Zehavi, I., Contreras, S., Padilla, N., et al. 2018, ApJ, 853, 84

Zennaro, M., Angulo, R. E., Contreras, S., Pellejero-Ibáñez, M., & Maion, F. 2022, MNRAS, 514, 5443

Zennaro, M., Angulo, R. E., Pellejero-Ibáñez, M., et al. 2023, MNRAS, 524, 2407

Zheng, Z., Berlind, A. A., Weinberg, D. H., et al. 2005, ApJ, 633, 791