Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **505,** 1195–1205 (2021)
Advance Access publication 2021 May 10

https://doi.org/10.1093/mnras/stab1317

A redefinition of the halo boundary leads to a simple yet accurate halo model of large-scale structure

Rafael García,[1]★ Eduardo Rozo,[1] Matthew R. Becker[2] and Surhud More[3,4]

[1]*Department of Physics, University of Arizona, Tucson, AZ 85721, USA*
[2]*High Energy Physics Division, Argonne National Laboratory, Lemont, IL 60439, USA*
[3]*The Inter-University Center for Astronomy and Astrophysics, Post bag 4, Ganeshkind, Pune 411007, India*
[4]*Kavli Institute for the Physics and Mathematics of the Universe (WPI), 5-1-5 Kashiwanoha, Chiba 2778583, Japan*

Accepted 2021 May 6. Received 2021 May 5; in original form 2021 January 14

ABSTRACT

We present a model for the halo–mass correlation function that explicitly incorporates halo exclusion and allows for a redefinition of the halo boundary in a flexible way. We assume that haloes trace mass in a way that can be described using a single scale-independent bias parameter. However, our model exhibits scale-dependent biasing due to the impact of halo-exclusion, the use of a 'soft' (i.e. not infinitely sharp) halo boundary, and differences in the one halo term contributions to $\xi_{\rm hm}$ and $\xi_{\rm mm}$. These features naturally lead us to a redefinition of the halo boundary that lies at the 'by eye' transition radius from the one-halo to the two-halo term in the halo–mass correlation function. When adopting our proposed definition, our model succeeds in describing the halo–mass correlation function with ≈ 2 per cent residuals over the radial range $0.1\,h^{-1}\,{\rm Mpc} < r < 80\,h^{-1}\,{\rm Mpc}$, and for halo masses in the range $10^{13}\,h^{-1}\,{\rm M}_\odot < M < 10^{15}\,h^{-1}\,{\rm M}_\odot$. Our proposed halo boundary is related to the splashback radius by a roughly constant multiplicative factor. Taking the 87 percentile as reference we find $r_{\rm t}/R_{\rm sp} \approx 1.3$. Surprisingly, our proposed definition results in halo abundances that are well described by the Press–Schechter mass function with $\delta_{\rm sc} = 1.449 \pm 0.004$. The clustering bias parameter is offset from the standard background-split prediction by ≈ 10 per cent–15 per cent. This level of agreement is comparable to that achieved with more standard halo definitions.

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

1 INTRODUCTION

The halo model is a powerful formalism for studying the statistical properties of the dark matter and galaxy density fields. In the halo model, the abundance and distribution of galaxies and clusters are linked to the abundance and distribution of dark matter haloes (Cooray & Sheth 2002). The halo model makes several key assumptions. First, it assumes all the matter in the Universe is contained in haloes. This means that the distribution of matter in the Universe can be described by specifying the abundance and distribution of haloes, as well as the mass distribution within these haloes. These statistics are described by the halo mass function dn/dm, the halo bias $b(m)$, and the halo density profile $u(r|m)$. Predicting these halo properties requires large computer simulations that map the matter distribution of the Universe. The output of the simulations is then analysed using a halo finder.

Every halo finding algorithm makes two critical yet relatively arbitrary choices. The first has received plenty of attention, and is the definition of halo mass. Halo mass is typically defined as the mass enclosed within some specific spherical aperture, chosen such that the mean density of the halo within that sphere is equal to some factor of either the critical density or the mass density of the Universe. Spherical overdensity definitions come with a number of issues, such as pseudo-evolution of halo radius and mass (Diemer, More & Kravtsov 2013a; Diemer, Kravtsov & More 2013b). Recent studies have looked into more physically motivated halo boundaries, such as the splashback radius (Diemer & Kravtsov 2014; More, Diemer & Kravtsov 2015). The splashback radius is defined as the radius at which accreted matter reaches its first orbital apocentre after turnaround. This choice of radius solves the issue of pseudo-evolution and cleanly separates infalling material from matter orbiting in the halo. However, other definitions are also commonly used (e.g. friends-of-friends) (see e.g. Knebe et al. 2013). For this reason, one can find calibrations of the halo mass function for multiple definitions (e.g. Tinker et al. 2008; Bhattacharya et al. 2011; McClintock et al. 2019).

The second arbitrary choice is how a halo finding algorithm decides which structures are parent haloes, and which are sub-haloes that 'belong' to a larger halo. We refer to the criteria for categorizing structures as parent haloes versus sub-haloes as percolation or halo exclusion criteria. There is currently no standard percolation scheme, with different halo finders applying different halo exclusion criteria when constructing halo catalogues. The choice of percolation can impact the halo–mass correlation function by up to ≈ 30 per cent (García & Rozo 2019).

The simplest commonly used form of a halo-model description of the halo–mass correlation function ignores both halo boundaries and halo exclusion. One writes $\xi_{\rm hm}(r) = \xi^{\rm 1h}_{\rm hm}(r) + \xi^{\rm 2h}_{\rm hm}(r)$ where the first and second term are referred to as the 1-halo and the 2-halo term,

★ E-mail: rgarciamar@email.arizona.edu

respectively (see e.g. Cooray & Sheth 2002). These two components are usually modelled independently. The one-halo term is described by a halo profile $u(r|m)$, usually an NFW or Einasto profile (Einasto 1965; Navarro, Frenk & White 1997). The two-halo term is modelled by assuming a scale-independent halo bias, where the bias can be defined relative to either the mass correlation function or the linear correlation function. These assumptions result in biases as large as ≈ 20 per cent at translinear scales (Hayashi & White 2008). More recent efforts have introduced scale dependence of the halo bias, allowing for more accurate modelling of the trans-linear regime (van den Bosch et al. 2013).

The fact that parent haloes do not reside inside other parent haloes has important consequences for modelling the trans-linear regime (Sheth & Lemson 1999). Previous studies have incorporated halo exclusion effects into the modelling of the halo distribution in a variety of ways. Smith, Scoccimarro & Sheth (2007), Smith, Desjacques & Marian (2011), and Baldauf et al. (2013) approach halo exclusion by imposing a hard-sphere constraint in the halo–halo correlation function. Philcox, Spergel & Villaescusa-Navarro (2020) addresses halo exclusion in a similar way and incorporates it into the halo covariance. Valageas & Nishimichi (2011a, b) introduced halo exclusion in a Lagrangian framework via a perturbative approach. Unlike the hard-sphere approaches above, our analysis will allow for a 'soft' halo exclusion, in which haloes can partially overlap in a way that enables self-consistent mass definitions.

In this paper, we incorporate both halo edges and halo exclusion into the modelling of the halo–mass correlation function. We demonstrate that by explicitly introducing these two components into the model we achieve much better accuracy from small to large scales for a wide range of halo masses. We emphasize that our model does not require any scale-dependent clustering biases, beyond those brought about because of halo exclusion effects. To achieve good agreement with simulations, our model requires that we redefine halo boundaries. It is this boundary redefinition that is the most important novel aspect of our analysis. Indeed, we show that there is a unique halo radius and mass power-law relation $R(M)$ that ensures consistency between the halo catalogue and our model.

2 A HALO MODEL FOR THE HALO–MASS CORRELATION FUNCTION

2.1 The standard approach

We begin with a brief review of the formalism detailed in Cooray & Sheth (2002), as it forms the basis for our model. Let $\vec{x}_i$ be the position of the ith halo in the Universe. If all mass is contained within haloes, then the mass density of the Universe can be written as

$$\rho_{\rm m}(\vec{x}) = \sum_i m_i u(\vec{x} - \vec{x}_i|m_i), \tag{1}$$

where $u(r|m)$ is the halo profile, and m_i is the mass of the $i^{\rm th}$ halo. Likewise, given a halo selection function $\phi(m)$ (i.e. $\phi(m) = 1$ when $m \in [m - \Delta m, m + \Delta m]$ and 0 otherwise) the corresponding halo density field is

$$n(\vec{x}) = \sum_i \delta(\vec{x} - \vec{x}_i)\phi(m_i). \tag{2}$$

Given these two fields, the halo–mass correlation function is

$$\xi_{\rm hm}(|\vec{x} - \vec{x}'|) = \frac{1}{\bar{n}\bar{\rho}_{\rm m}}\langle n(\vec{x})\rho_{\rm m}(\vec{x}')\rangle - 1, \tag{3}$$

where $\langle\cdots\rangle$ denotes ensemble averaging. We can plug in the expressions for each density field into 3, and predict the two-point correlation function in therms of the halo density profile, the halo mass function, and the clustering of haloes. One has then

$$\begin{aligned}\langle n(\vec{x})\rho_{\rm m}(\vec{x}')\rangle &= \Big\langle \sum_i \sum_j m_j\phi(m_i)\delta(\vec{x} - \vec{x}_i)u(\vec{x}' - \vec{x}_j|m_j)\Big\rangle \\ &= \Big\langle \sum_i m_i\phi(m_i)\delta(\vec{x} - \vec{x}_i)u(\vec{x}' - \vec{x}_i|m_i)\Big\rangle \\ &\quad + \Big\langle \sum_i \sum_{j\neq i} m_j\phi(m_i)\delta(\vec{x} - \vec{x}_i)u(\vec{x}' - \vec{x}_j|m_j)\Big\rangle \\ &\quad .\end{aligned} \tag{4}$$

The average over the ensemble has been separated into two parts: one that accounts for the correlation between a halo and the mass contained within it, and one that accounts for the correlation between a halo, and mass that belongs to other haloes. We treat each in turn. We have

$$\begin{aligned}\text{1st term} &= \Big\langle \sum_i m_i\phi(m_i)\delta(\vec{x} - \vec{x}_i)u(\vec{x}' - \vec{x}_i|m_i)\Big\rangle \\ &= \Big\langle \int {\rm d}m \sum_i m_i\phi(m_i)\delta(\vec{x} - \vec{x}_i) \\ &\quad \times u(\vec{x}' - \vec{x}_i|m_i)\delta(m - m_i)\Big\rangle \\ &= \int {\rm d}m\, m\phi(m)u(\vec{x}' - \vec{x}|m)\Big\langle \sum_i \delta(\vec{x} - \vec{x}_i)\delta(m - m_i)\Big\rangle \\ &\quad .\end{aligned} \tag{5}$$

The remaining expectation value corresponds to the mean number of haloes per unit volume per unit mass, that is, the halo mass function,

$$\frac{{\rm d}n}{{\rm d}m} = \Big\langle \sum_i \delta(\vec{x} - \vec{x}_i)\delta(m - m_i)\Big\rangle. \tag{6}$$

Plugging the mass function into the 1st term and integrating over a narrow mass selection function we arrive at

$$\begin{aligned}\text{1st term} &= \int {\rm d}m \frac{{\rm d}n}{{\rm d}m} m\phi(m)u(\vec{x}' - \vec{x}|m) \\ &= \bar{n}mu(\vec{x}' - \vec{x}|m).\end{aligned} \tag{7}$$

This is the so-called one halo term of the halo–mass correlation function.

Now, let's look at the second term,

$$\begin{aligned}\text{2nd term} &= \Big\langle \sum_i \sum_{j\neq i} m_j\phi(m_i)\delta(\vec{x} - \vec{x}_i)u(\vec{x}' - \vec{x}_j|m_j)\Big\rangle \\ &= \int {\rm d}m{\rm d}m'{\rm d}\tilde{\vec{x}}\; m'\phi(m)u(\vec{x}' - \tilde{\vec{x}}|m') \\ &\quad \times \Big\langle \sum_i \sum_{j\neq i} \delta(\vec{x} - \vec{x}_i)\delta(m - m_i)\delta(\tilde{\vec{x}} - \vec{x}_j)\delta(m' - m_j)\Big\rangle \\ &= \int {\rm d}m{\rm d}m'{\rm d}\tilde{\vec{x}}\; m'\phi(m)u(\vec{x}' - \tilde{\vec{x}}|m') \\ &\quad \times \frac{{\rm d}n}{{\rm d}m}\frac{{\rm d}n}{{\rm d}m'}[1 + \xi_{\rm hh}(\vec{x} - \tilde{\vec{x}}|m, m')].\end{aligned} \tag{8}$$

Haloes are biased tracer of the matter density field. At scales much larger than the size of haloes $\xi_{\rm hh}(r|m, m') = b(m)b(m')\xi_L(r)$

$$\xi_{\rm hh}(\vec{x} - \tilde{\vec{x}}|m, m') = b(m)b(m')\xi_L(\vec{x} - \tilde{\vec{x}}) \tag{9}$$

It follows that

$$\text{2nd term} = \bar{n}\bar{\rho}_{\rm m} + \bar{n}b(m)\int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}m'b(m') \times \int {\rm d}\vec{\tilde{x}}\; u(\vec{x}' - \vec{\tilde{x}}|m')\xi_L(\vec{x} - \vec{\tilde{x}}). \quad (10)$$

At large scales, the details of the halo profile become unimportant, and the haloes themselves can be approximated as point masses, so that $u(\vec{x}) \approx \delta(\vec{x})$. With this approximation, and the identity,

$$\int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}m'b(m') = 1\,, \quad (11)$$

the 2nd term becomes

$$\text{2nd term} = \bar{n}\bar{\rho}_{\rm m} + \bar{n}\bar{\rho}_{\rm m}b(m)\xi_L(\vec{x} - \vec{x}')\,. \quad (12)$$

Getting everything together, the product becomes

$$\langle n(\vec{x})\rho_{\rm m}(\vec{x}')\rangle = \bar{n}mu(\vec{x}' - \vec{x}|m) + \bar{n}\bar{\rho}_{\rm m}(1 + b(m)\xi_L(\vec{x} - \vec{x}'))\,. \quad (13)$$

The halo–mass correlation function is

$$\xi_{\rm hm}(r|m) = \tfrac{m}{\bar{\rho}_{\rm m}}u(r|m) + b(m)\xi_L(r), \quad (14)$$

where $r = |\vec{x} - \vec{x}'|$. The first term is known as the one-halo term, $\xi_{\rm hm}^{\rm 1h}$, and accounts for the mass within a single halo. The second term is known as the two-halo term, $\xi_{\rm hm}^{\rm 2h}$, and accounts for the mass across different haloes.

2.2 Incorporating halo exclusion

In the standard approach, we assumed that

$$\left\langle \sum_i \sum_{j\neq i} \delta(\vec{x} - \vec{x}_i)\delta(m - m_i)\delta(\vec{\tilde{x}} - \vec{x}_j)\delta(m' - m_j)\right\rangle = \frac{{\rm d}n}{{\rm d}m}\frac{{\rm d}n}{{\rm d}m'}[1 + b(m)b(m')\xi_L(\vec{x} - \vec{\tilde{x}})]. \quad (15)$$

This is true at large scales, because haloes never overlap. This is not the case at small scales. We introduce a halo exclusion function $E(\vec{x}_i - \vec{x}_j|m_i, m_j)$ which is zero when haloes overlap, and one otherwise. This halo exclusion function multiplies the entire 2nd term, so that equation (8) now becomes

$$\text{2nd} = \int {\rm d}m{\rm d}m'{\rm d}\vec{\tilde{x}}\; m'\phi(m)u(\vec{x}' - \vec{\tilde{x}}|m') \times \frac{{\rm d}n}{{\rm d}m}\frac{{\rm d}n}{{\rm d}m'}[1 + b(m)b(m')\xi_L(\vec{x} - \vec{\tilde{x}})]E(\vec{x} - \vec{\tilde{x}}|m, m'). \quad (16)$$

For a narrow selection function, we get

$$\text{2nd} = \int {\rm d}m'\,\bar{n}\frac{{\rm d}n}{{\rm d}m'}m' \int {\rm d}\vec{\tilde{x}}\; u(\vec{x}' - \vec{\tilde{x}}|m') \times [1 + b(m)b(m')\xi_L(\vec{x} - \vec{\tilde{x}})]E(\vec{x} - \vec{\tilde{x}}|m, m'). \quad (17)$$

The integral over all space is a convolution of the density profile and the 2-halo term with exclusion.

$$\text{2nd} = \int {\rm d}m'\,\bar{n}\frac{{\rm d}n}{{\rm d}m'}m'(u * E)(\vec{x} - \vec{x}'|m, m') + \int {\rm d}m'\,\bar{n}\frac{{\rm d}n}{{\rm d}m'}m'b(m)b(m')(u * E\xi_L)(\vec{x} - \vec{\tilde{x}}). \quad (18)$$

To make further progress, we must specify a halo exclusion function. We assume halo exclusion happens when haloes are separated by a distance $r \leq r_{\rm e}(m, m')$, where $r_{\rm e}$ is the halo exclusion radius. Note that the halo exclusion radius depends on the masses m and m' of the two haloes under consideration. With this definition, the halo exclusion function takes the form

$$E(r|m, m') = 1 - \theta(r_{\rm e}(m, m') - r), \quad (19)$$

where θ is the Heaviside step function. We can set upper and lower bounds for the exclusion radius. For the lower bound, the exclusion radius must be larger than the radius of either of the two haloes. For the upper bound, we use a hard sphere model.

$$\max\{r_{\rm t}(m), r_{\rm t}(m')\} < r_{\rm e}(m, m') < r_{\rm t}(m) + r_{\rm t}(m'). \quad (20)$$

In the above expression, $r_{\rm t}(m)$ is the radius of a halo of mass m.

Inserting (19) into our previous expressions we find

$$\text{2nd} = \int {\rm d}m'\,\bar{n}\frac{{\rm d}n}{{\rm d}m'}m'[u * (1 - \theta_e)](r) + \int {\rm d}m'\,\bar{n}\frac{{\rm d}n}{{\rm d}m'}m'b(m)b(m')[u * (1 - \theta_e)\xi_L](r). \quad (21)$$

$$= \bar{n}\bar{\rho}_{\rm m}\left[1 - \int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}(u * \theta_e)(r)\right] + \bar{n}\bar{\rho}_{\rm m}b(m)\left[\int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}b(m')[u * (1 - \theta_e)\xi_L](r)\right]. \quad (22)$$

The halo–mass correlation function becomes

$$\xi_{\rm hm}(r|m) = \frac{m}{\bar{\rho}_{\rm m}}u(r|m) - \int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}\theta_e(r|m, m') + b(m)\int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}b(m')[u * (1 - \theta_e)\xi_L](r). \quad (23)$$

Note that this model for the halo–mass correlation function explicitly incorporates halo exclusion in a flexible way, in the sense that the model can be used with any definition for a halo boundary and with any choice of halo percolation.

We can further simplify this expression by using the same approximation as in the standard case, i.e. at large scales the mass profile becomes unimportant, and we can set $u(\vec{x}) \approx \delta(\vec{x})$. With this approximation, the above expression simplifies to

$$\xi_{\rm hm}(r|m) = \frac{m}{\bar{\rho}_{\rm m}}u(r|m) + b(m)\xi_L(r) - \int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}\theta_e(r|m, m') - b(m)\xi_L(r)\int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}b(m')\theta_e(r|m, m'). \quad (24)$$

We wish to incorporate into our model the fact that spherical overdensity halo finders define sharp halo edges such that the mass interior to the halo radius belongs to the halo, while mass exterior to the halo radius does not. This in turn implies that a self-consistent model of the halo–mass correlation function ought to truncate the halo term at the halo boundary. With this truncation in mind, the matter density field can be written as

$$\rho_{\rm m}(\vec{x}) = \sum_i m_i u(\vec{x} - \vec{x}_i|m_i)\theta(r|r_{\rm t}(m)), \quad (25)$$

where $\theta(r|r_{\rm t}(m)) = 1$ when $r < r_{\rm t}(m)$ and 0 otherwise. This imposes a sharp cut in the halo density profile of one halo. In practice, however, we expect that the halo–mass correlation function will exhibit some effective finite width in the radial direction. For instance, we know haloes are triaxial, so even if a halo is defined using a spherical overdensity, we expect 'nature' would prefer a triaxial definition. A triaxial halo definition would 'spread out' the halo boundary across a range of radial scales, naturally leading to a soft truncation of the one-halo term. In short, we expect a soft truncation will produce better results than an infinitely sharp truncation. Of course, this implies that

our model is not entirely consistent with the sharp radial cut imposed by halo finders. We consider this a small price to pay for better precision in our model. Moreover, one could imagine modifying halo finders in order to implement a soft truncation, thereby mimicking our model for the halo–mass correlation function. Indeed, this is how some cluster finders work (e.g. redMaPPer Rykoff et al. 2012). We will leave the task of exploring such modifications of halo finders to future work.

For the above reasons, we choose to model the truncation using the complementary error function centred at the halo edge $r_{\rm t}(m)$ with width $\Delta r_{\rm t}(m)$.

$$\theta_{\rm t}(r|m) = \frac{1}{2}\,{\rm erfc}\left(\frac{r - r_{\rm t}}{\sqrt{2}\Delta r_{\rm t}}\right). \tag{26}$$

Upon including this halo truncation term, the halo–mass correlation function can be written as

$$\begin{aligned}\xi_{\rm hm}(r|m) = &\frac{m}{\bar{\rho}_{\rm m}}u(r|m)\theta_{\rm t}(r|m) + b(m)\xi_L(r)\\ &- \int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}\theta(r|r_{\rm e}(m,m'))\\ &-b(m)\xi_L(r)\int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}b(m')\theta(r|r_{\rm e}(m,m')).\end{aligned} \tag{27}$$

This final expression can still be interpreted in a similar fashion as the standard halo model. It has a one-halo term that accounts for the matter that is contained within the halo boundary, and has a two-halo term that takes into account matter in the rest of the haloes. The difference is that there are correction terms due to halo exclusion. These correction terms can be interpreted as the mass that would have been there in other haloes, were it not for the exclusion volume associated with more massive haloes. Note this 'excluded mass' is comprised of both the excluded mass in the mean, and the 'extra' excluded mass due to halo-mass clustering. The final expressions for the 1-halo and 2-halo terms are

$$\begin{aligned}\xi^{\rm 1h}_{\rm hm}(r|m) = &\frac{m}{\bar{\rho}_{\rm m}}u(r|m)\theta_{\rm t}(r|m) - \int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}\theta(r|r_{\rm e}(m,m'))\\ &-b(m)\xi_L(r)\int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}b(m')\theta(r|r_{\rm e}(m,m')).\end{aligned} \tag{28}$$

$$\xi^{\rm 2h}_{\rm hm}(r|m) = b(m)\xi_L(r). \tag{29}$$

Note we have associated the correction terms with the one-halo term since these represent excluded mass in the vicinity of the halo, i.e. the excluded mass moves in space as one moves haloes in space.

2.3 Refining the two-halo term

In all of the above we have assumed that $\xi_{\rm hh} = b(m)b(m')\xi_L$. This is true at very large scales but not at small scales. If haloes trace matter, then as we move towards non-linear scales, we should expect ξ_L will need to be replaced the matter–matter correlation function $\xi_{\rm mm}$. However, the latter correlation function has a strong 1-halo contribution at small scales. Clearly, linear biasing cannot hold in this regime. The best we could hope for is linear bias relative to the 2-halo term of the matter correlation function, $\xi^{\rm 2h}_{\rm mm} = \xi_{\rm mm} - \xi^{\rm 1h}_{\rm mm}$. This raises the obvious question: how can we remove the 1-halo term of the matter correlation function?

While we cannot give a definitive answer a priori, it is clear what 'removing the 1-halo term' must do to the matter correlation function: it must suppress correlations at small scales. This leads us to adopt a two-halo term for the matter–matter correlation function of the form given by

$$\xi^{\rm 2h}_{\rm mm} = \xi_{\rm mm} \times (1 - \theta_{\rm t}(r_{\rm eff}, \Delta_{\rm eff})). \tag{30}$$

In this expression, $\theta_{\rm t}$ is again a smooth truncation function of the form given by equation (26). The radius $r_{\rm eff}$ sets the scale at which $\xi_{\rm mm}$ transitions from the 1-halo term to the 2-halo term, while $\Delta_{\rm eff}$ determines how quickly this transition occurs.

With these modifications, our final expression for the halo–matter correlation function is

$$\xi_{\rm hm}(r|m) = \xi^{\rm 1h}_{\rm hm}(r|m) + \xi^{\rm 2h}_{\rm hm}(r|m) \tag{31}$$

$$\begin{aligned}\xi^{\rm 1h}_{\rm hm}(r|m) = &\frac{m}{\bar{\rho}_{\rm m}}u(r|m)\theta_{\rm t}(r|m)\\ &- \int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}\theta(r|r_{\rm e}(m,m'))\\ &-b(m)\xi^{\rm 2h}_{\rm mm}(r)\int {\rm d}m'\,\frac{{\rm d}n}{{\rm d}m'}\frac{m'}{\bar{\rho}_{\rm m}}b(m')\theta(r|r_{\rm e}(m,m'))\end{aligned} \tag{32}$$

$$\xi^{\rm 2h}_{\rm hm}(r|m) = b(m)\xi^{\rm 2h}_{\rm mm}(r) \tag{33}$$

We briefly discuss how the expressions written above compare with the expressions for the halo matter correlation function in van den Bosch et al. (2013). In their approach, the first term in equation (32) is $\xi^{\rm 1h}_{\rm hm}$, while the rest of the terms in that equation are accounted for in their two-halo term. The two-halo term that they consider includes radial dependence of the halo bias, halo exclusion, and uses the non-linear matter correlation function. In our case, we consider a simple linear bias relative to $\xi^{\rm 2h}_{\rm mm}$ instead. Thus there is considerable simplicity in the expressions we have derived. In Section 3, we will fit the halo matter correlation function with the results from numerical simulations. For the routinely used halo-mass definition $M_{\rm 200m}$, the model of van den Bosch et al. (2013) performs well while our model performs poorly. As we show below, however, if the halo definitions are made consistent with our formalism – a step that requires fairly simple and straightforward tweaks to the halo finding algorithms – the simpler expressions in our model can describe the halo-mass correlation function with even greater accuracy than that achieved by van den Bosch et al. (2013). In particular, we argue that the complications regarding the radial dependence of the halo bias can be solved by a simple redefinition of the halo boundary, coupled with the use of the 2-halo term of the non-linear matter correlation function.

Similar arguments can be made with regards to perturbative approaches aimed at modelling non-linear biasing in translinear scales (e.g. Smith et al. 2011; Philcox et al. 2020). Our model demonstrates that the introduction of perturbative non-linear bias parameters is unnecessary for describing the clustering of mass around haloes, provided one adopts an appropriate halo boundary, and that one adequately models the impact of halo-exclusion in the data.

2.4 High mass limit

The most massive haloes are much bigger than the rest of their neighbours ($m >> m'$). Consequently, $r_{\rm t}(m) >> r_{\rm t}(m')$. This condition, along with the inequality (20) implies that,

$$r_{\rm e}(m, m') \approx r_{\rm t}(m). \tag{34}$$

Setting the exclusion radius to the halo radius of the more massive halo leads to

$$\xi_{\rm hm}(r|m) = \left[\frac{m}{\bar{\rho}_{\rm m}}u(r|m) - 1\right]\theta_{\rm t}(r|m) + b(m)\xi^{\rm 2h}_{\rm mm}(r)[1-\theta_{\rm t}(r|m)], \quad (35)$$

which is equivalent to

$$\xi_{\rm hm}(r|m) = \begin{cases} \frac{m}{\bar{\rho}_{\rm m}}u(r|m) - 1 & r \le r_{\rm t}(m) \\ b(m)\xi^{\rm 2h}_{\rm mm}(r) & r \ge r_{\rm t}(m) \end{cases}. \quad (36)$$

If instead of the truncated matter–matter correlation function, we use the linear correlation function ξ_L the model turns to

$$\xi_{\rm hm}(r|m) = \begin{cases} \frac{m}{\bar{\rho}_{\rm m}}u(r|m) - 1 & r \le r_{\rm t}(m) \\ b(m)\xi_L(r) & r \ge r_{\rm t}(m) \end{cases}. \quad (37)$$

Note that if $r_{\rm t}(m)$ is not known a priori, one can use the fact that $\xi_{\rm hm}$ is continuous to determine $r_{\rm t}(m)$. That is, our formalism has allowed us to derive from first principles the model proposed by Hayashi & White (2008).

2.5 Halo density profile

We assume that dark matter haloes are spheres whose normalized density distribution is given by the Einasto profile

$$u(r|m) = \frac{\rho_{\rm s}}{m}\exp\left\{-\frac{2}{\alpha}\left[\left(\frac{r}{r_{\rm s}}\right)^{\alpha} - 1\right]\right\}, \quad (38)$$

where $r_{\rm s}$ is the scale radius, $\rho_{\rm s}$ the density at $r_{\rm s}$, and α is the shape parameter. In the following, we use a more convenient parametrization via the mass and concentration and a mass definition. For a particular mass definition, say $M_{\rm 200m}$, the concentration is given by $c = R_{\rm 200m}/r_{\rm s}$, and the density $\rho_{\rm s}$ is obtained by the normalization of $u(r|m)$. This parametrization is

$$u(r|M) = \frac{\alpha c^3(\frac{2}{\alpha})^{3/\alpha}}{3\Gamma(3/\alpha)M}\exp\left\{-\frac{2}{\alpha}\left(\frac{rc}{R(M)}\right)^{\alpha}\right\}, \quad (39)$$

where $R(M)$ is the mass contained within the radius R that defines the boundary of the halo. We use the COLOSSUS (Diemer 2018) python package to compute the Einasto profile.

2.6 Simplifying the halo exclusion terms

The fact that the exclusion terms are integrals over an infinite mass range poses a problem. With the simulation we are using we simply do not know the halo mass function for $M < 10^{12}h^{-1}\,{\rm M}_\odot$. Moreover, lowering this mass limit requires using smaller boxes, which in turn looses larger modes. In short, a brute-force approach to this problem appears unpalatable. While in principle one could attempt performing these integrals over infinite mass range by truncating at the lower mass limit and including an additive term (Schmidt 2016), the redefinition of the halo boundary means that we do not truly know the halo mass function until we find a consistent $R(M)$ relation. Consequently, this 'trick' is not available to us. Despite these limitations, we can take advantage of the fact that the corrections due to halo exclusion are integrals over the halo mass function. Adopting the normalization condition that all mass is contained within haloes, we can think of the mass function $(1/\bar{\rho}_{\rm m})m{\rm d}n/{\rm d}m$ as a probability distribution. Consequently, the mass integrals can all be thought of as expectation values. We assume that the average value of a function f over the probability distribution $(1/\bar{\rho}_{\rm m})m{\rm d}n/{\rm d}m$ can be approximated as the function f evaluated at some input parameter m_0 where we expect $m_0 \approx \langle m\rangle$.

Using this approximation, the corrections due to halo exclusion are simplified and the correlation function can be written as

$$\xi_{\rm hm}(r|m) = \frac{m}{\bar{\rho}_{\rm m}}u(r|m)\theta_{\rm t}(r|m) + b(m)\xi^{\rm 2h}_{\rm mm}(r) - \theta(r|r_{\rm e}(m,m_b)) - b(m)\xi^{\rm 2h}_{\rm mm}(r)\theta(r|r_{\rm e}(m_a,m')), \quad (40)$$

where m_a and m_b are the values of halo mass which approximate the expectation value of the exclusion functions over all halo masses. There are two parameters since the exclusion function is weighted differently in each exclusion term. Note in particular that the parameter m_a arises from the exclusion function integral that includes a bias weighting of the haloes, so we should expect $m_a > m_b$ due to the steep dependence of the halo bias at high masses.

Our approach here again differs from that of van den Bosch et al. (2013), who split the mass integrals into two, and then simplified using the integral conditions. Nevertheless, they must still perform integrals over mass, which our model does away through the introduction of the m_a and m_b parameters.

2.7 Model parameters

Our final model for the halo–mass correlation function depends on several model parameters, namely:

(i) concentration c
(ii) Einasto parameter α
(iii) halo bias b
(iv) truncation parameters of $\xi^{\rm 2h}_{\rm mm}$: $r_{\rm eff}$, $\Delta_{\rm eff}$
(v) effective masses for halo exclusion corrections m_a, m_b

There are additional 'parameters' in our fits, namely

(i) the halo mass m,
(ii) the halo radius $r_{\rm t}$.

The mass m governs the amplitude of the 1-halo term in our fit, while the radius $r_{\rm t}$ sets the boundary of the halo. In principle, these parameters should not be fit parameters. For instance, when using an overdensity criterion Δ when defining halo masses, a self-consistent model should have M_Δ as the mass parameter governing the 1-halo amplitude. Likewise, one should set the radius $r_{\rm t} = R_\Delta$.

As we will see, in practice, using commonly used fixed overdensity criteria results in poor fits to the data. This allows us to ask the question: is the simulation data well fit with some other halo mass m and halo radius $r_{\rm t}$? In this case, we can use our halo model with m and $r_{\rm t}$ as fit parameters to learn about what the mass and radius of the haloes should have been. When doing so, our fits rely on nine parameters for a single mass bin. However, we can vastly reduce this parameter space by enforcing simple power-law scalings of many of our parameters with halo mass. Additionally, the one-halo term that we subtract from the matter–matter correlation function must be independent of mass. Thus, the parameters $r_{\rm eff}$ and $\Delta_{\rm eff}$ have to be shared across all halo mass bins. This forces us to simultaneously fit the model across all available halo masses.

We assume that the halo radius, concentration, the shape parameter, and the effective masses can be parametrized as power laws of halo mass. That is

$$r_{\rm t} = r_p\left(\frac{m}{m_{p1}}\right)^{\beta} \quad (41)$$

$$c = c_p\left(\frac{m}{m_{p2}}\right)^{\gamma} \quad (42)$$

Table 1. Model parameters. The limits in square brackets indicate flat priors.

Parameter	Description	Prior
$\log_{10} m_k$	Halo mass	[11.0, 16.0]
b_k	Halo bias	$[0, \infty]$
$r_{\rm p}$	Halo radius pivot	$[0, \infty]$
β	Halo radius power	$[0, \infty]$
$c_{\rm p}$	Concentration pivot	$[0, \infty]$
γ	Concentration power	$[-\infty, 0]$
$\alpha_{\rm p}$	Shape parameter pivot	$[-\infty, \infty]$
δ	Shape parameter power	$[0, \infty]$
$m_{\rm ap}$	Halo exclusion pivot	$[0, \infty]$
A	Halo exclusion power	$[0, \infty]$
$m_{\rm bp}$	Halo exclusion pivot	$[0, \infty]$
B	Halo exclusion power	$[0, \infty]$
Δ	Width of truncation for one-halo term	$[0, \infty]$
Δ_e	Width of truncation for exclusion terms	$[0, \infty]$
$r_{\rm eff}$	Truncation radius of $\xi_{\rm mm}$	$[0, \infty]$
$\Delta_{\rm eff}$	Truncation width of $\xi_{\rm mm}$	$[0, \infty]$

$$\alpha = \alpha_p \left(\frac{m}{m_{p3}}\right)^{\delta} \tag{43}$$

$$m_a = m_{ap} \left(\frac{m}{m_{p4}}\right)^{A} \tag{44}$$

$$m_b = m_{bp} \left(\frac{m}{m_{p5}}\right)^{B}, \tag{45}$$

where we fit for the amplitudes and exponents in these relations. We select the pivot values $m_{p1} = m_{p3} = 2 \times 10^{14}$, $m_{p2} = 7 \times 10^{14}$, $m_{p4} = m_{p5} = 2 \times 10^{12}$ which are typical values of halo mass in our halo catalogue. The above selection of pivot points roughly decorrelates the slope and amplitude parameters, and was obtained through trial and error.

The likelihood of the halo-mass correlation function for haloes in the $k^{\rm th}$ mass bin is

$$\ln \mathcal{L}_k = \ln \mathcal{L}(\xi_{\rm hm}(r|m_{\rm k})|\theta) \propto -\tfrac{1}{2}\mathbf{D}_k^{\top}\mathbf{C}_{\xi_{\rm hm}}^{-1}\mathbf{D}_k \tag{46}$$

where $\boldsymbol{\theta} = (m,\ b,\ r_{\rm p},\ \beta,\ c_{\rm p},\ \gamma,\ \alpha_{\rm p},\ \delta,\ m_{\rm ap},\ A,\ m_{\rm bp},\ B,\ r_{\rm eff},\ \Delta_{\rm eff})$ is the vector of model parameters, $\boldsymbol{D}_k = \boldsymbol{\xi}_{\rm hm}^{\rm data} - \boldsymbol{\xi}_{\rm hm}^{\rm model}$ and $\mathbf{C}_{\xi_{\rm hm}}$ is the covariance matrix of $\boldsymbol{\xi}_{\rm hm}^{\rm data}$. We are looking to fit for all mass bins simultaneously since the parameters $r_{\rm eff}$ and $\Delta_{\rm eff}$ are shared across all halo mass bins. To maintain the jackknife covariance matrix well-conditioned, we ignore the covariance across mass bins. We emphasize that while this assumption will impact the width of the posterior distribution in our analysis, we expect its impact on the precision of the best-fitting model will be minimal. With this assumption, the total likelihood is given by

$$\ln \mathcal{L}(\{\xi_{\rm hm}(r|m_k)\}_{k=1}^{N}|\theta) \propto \textstyle\sum_k \ln \mathcal{L}_k. \tag{47}$$

The priors on the parameters are shown in Table 1. The likelihood is sampled using the python package EMCEE (Foreman-Mackey et al. 2013). The total number of parameters for 12 mass bins is 38 (i.e. just over 3 parameters per correlation function). We use 152 walkers with 50 000 steps each and discard the first 5000 steps of each walker. The chains of each walker become uncorrelated after 400 steps, ensuring a minimum of 17 000 independent samples.

3 RESULTS

3.1 The halo-mass correlation function

We measure the halo-mass correlation function using a cosmological N-body simulation similar to those used in the Aemulus project (DeRose et al. 2019). It is run with the publicly available code GADGET2 (Springel 2005). The simulation is a periodic box of size $1050\,h^{-1}$ Mpc with 1400^3 particles. The cosmology is $h = 0.6704$, $\Omega_m = 0.318$, $\Omega_\Lambda = 0.682$, $\Omega_b = 0.049$, $\sigma_8 = 0.835$, $n_s = 0.962$. The particle mass is $3.7275 \times 10^{10}\,h^{-1}\,{\rm M}_\odot$ and the force softening scale is $20\,h^{-1}$kpc. Haloes were found using the publicly available ROCKSTAR halo finder (Behroozi, Wechsler & Wu 2013) with a spherical overdensity of $\Delta = 200$. ROCKSTAR uses an adaptive friends-of-friends algorithm in 6D phase space to identify dark matter structures. These structures are classified as parent haloes or subhaloes using a soft-sphere halo exclusion scheme: two structures are considered to be in the same parent halo if their separation is less than the radius of the larger structure.

We attempt to fit the halo-mass correlation function data with our halo model. The left-hand panel of Fig. 1 shows the halo-mass correlation function for haloes of mass $M = [1, 2] \times 10^{13}\,h^{-1}\,{\rm M}_\odot$, where mass is defined using an overdensity criterion $\Delta = 200$ relative to the mean density of the Universe. In this fit, we have forced the mass parameter in our halo model to be equal to the mass of the haloes. Likewise, we have forced the truncation parameter $r_{\rm t}$ to coincide with $R_{\rm 200m}$, the radius of the haloes. The latter is shown as a vertical line in the plot, which is left of the 'valley' between the two bumps in the data, which one might expect to correspond to the one and two halo terms. Unsurprisingly, the fit to the data is poor despite the model having seven free parameters.

We now test how well our model works if we let the mass and radius parameters in the halo model be free. This, of course, results in a model that is inconsistent with the halo definitions employed in the creation of the halo catalogue. We will address this point momentarily. For now, let us simply consider how our model fits the data when we let m and $r_{\rm t}$ float.

The right-hand panel in Fig. 1 shows our best-fitting model for the halo-mass correlation function when allowing the mass and truncation radius parameters to float. We see that our halo model now provides an excellent description of the data, and that the best-fitting truncation radius $r_{\rm t}$ (shown as a vertical line) falls close to the 'by-eye' transition between the 1- and 2-halo bumps of the halo-mass correlation function. In other words, the simulation data clearly suggests that the halo boundary should extend further out than $R_{\rm 200m}$, and should be set by $r_{\rm t}$ instead.

These results suggest how to address the lack of consistency between the model parameters m and $r_{\rm t}$, and the mass and halo boundary used to define the haloes in the first place. We consider an iterative approach to halo finding which proceeds as follows. We start by assuming that haloes are defined in iteration i via a radius–mass relation $R_i(M)$. For instance, in iteration $i = 1$, this relation corresponds to the fixed overdensity criterion, $3M/4\pi R_1(M)^3 = \Delta\bar{\rho}_{\rm m}$. Given the relation $R_i(M)$, we perform the following operations:

(i) We generate a halo catalogue using $R_i(M)$ to define the boundaries of haloes and to enforce halo exclusion.

(ii) We measure the halo–matter correlation functions for haloes in fixed mass bins.

(iii) We fit the resulting halo–mass correlation function letting the mass m and radius $r_{\rm t}$ parameters float. These new estimates define the radius–mass relation $R_{i+1}(M)$.

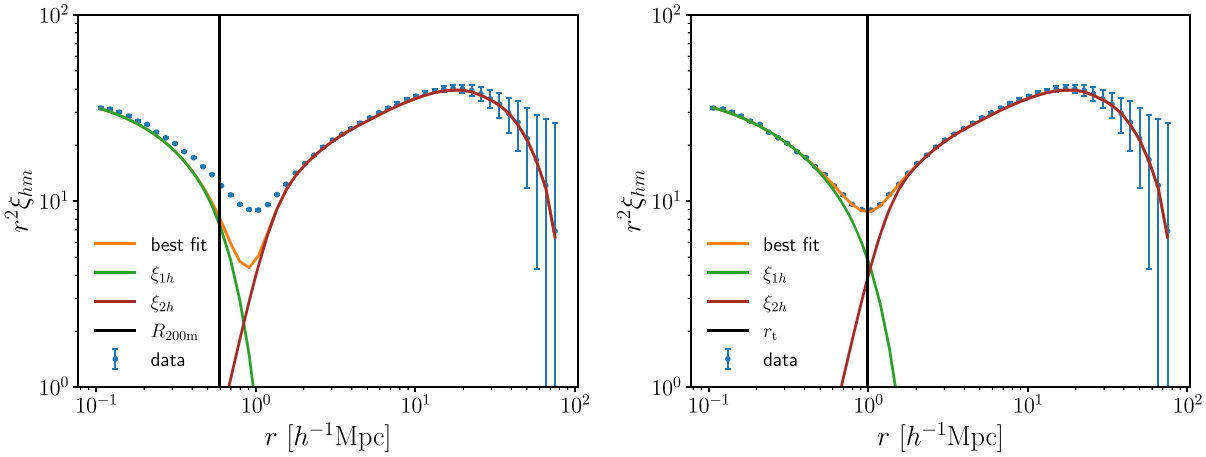

Figure 1. The halo-mass correlation function and model predictions using different choices of halo radius, and letting the amplitude of the one-halo term be given by the mass contained within the aperture used to define haloes. Left-hand panel: Best-fitting model using R_{200m} as the halo boundary and exclusion radius. Right-hand panel: Best-fitting model using our self-consistent halo boundary $(1+\Delta)r_t$ obtained through iteratively applying our model to the simulations. We plot the halo boundary $(1+\Delta)r_t$ instead of r_t because the halo profile does not end at r_t, it is smoothly truncated with a complementary error function of width Δ. Error bars are jackknife. Our halo model provides an accurate description of the simulation data provided the halo boundary is properly defined.

The procedure is then iterated until convergence is achieved, that is, we iterate until $R_{i+1}(M) = R_i(M)$. In practice, we find that $R(M)$ converges to within ≈ 1 per cent by the end of the second iteration, and it is converged to 0.01 per cent in $\approx$5 iterations. Our fully converged radius–mass relation takes the form

$$r_t(m) = r_p \left(\frac{m}{m_p}\right)^{\beta}, \tag{48}$$

where we select the pivot value $m_p = 2 \times 10^{14} h^{-1}\,\mathrm{M}_\odot$, and the converged parameters are $r_p = 1.558 \pm 0.001 h^{-1}\,\mathrm{Mpc}$, $\beta = 0.200 \pm 0.001$. This relation can be recast as a mass-dependent overdensity criterion,

$$\Delta(M) = \frac{3m_p^{3\beta}}{4\pi\rho r_p^3} M^{1-3\beta}. \tag{49}$$

We emphasize that once convergence is achieved, the halo mass and radius should no longer be considered fit parameters. That is, when adopting the halo mass definition described in equation (48), the parameters m and r_t in the halo model are given precisely by the mass and radius used to define the haloes.

Fig. 2 shows a comparison between the halo–matter correlations measured in the simulation to our best-fitting model after convergence is achieved. The model performs well in a wide range of halo masses and scales. It achieves 2 per cent accuracy for haloes of mass $10^{13}\,h^{-1}\,\mathrm{M}_\odot$ from 0.2 to $60\,h^{-1}\,\mathrm{Mpc}$. Larger halo masses exhibit larger ($\sim$ 10 per cent) deviations, though these are consistent with noise as estimated using jackknife resampling. In other words, our simulation box is not sufficiently large for us to give a robust estimate of the precision of our model at high halo masses. Likewise, the Press–Schechter fit presented here has only been validated for haloes with mass $M \geq 10^{13}\,M_\odot$. We will provide improved calibrations of the precision of our model in future work.

3.2 Is r_t related to the splashback radius?

We have seen that our analysis naturally leads us to redefine halo boundaries. Recently, the so-called splashback radius has been proposed as a physical halo boundary (Diemer & Kravtsov 2014; More et al. 2015). We compare the halo radius we derive to the splashback radius as defined using the SPARTA algorithm (Diemer 2017; Diemer et al. 2017). SPARTA tracks the orbits of all particles in a halo and measures the location of the first apocentre of all particles. The splashback radius of a halo is defined as the smoothed average of the apocentre radii of a fraction of the particles. Common choices are the 75th and 87th percentiles, which roughly match the splashback radius defined as the steepest point of the logarithmic slope of the spherically averaged density profile (More et al. 2015), and as the radius of the sphere with volume equal to the splashback shell of a halo, as first introduced in the code SHELLFISH (Mansfield, Kravtsov & Diemer 2017), respectively. When computing the splashback radius of a halo in the simulation, we rely on the M_{200m} mass of the halo as measured in the simulation.

Fig. 3 shows the ratio r_t/R_{sp} for several splashback definitions, specifically the median, 75th and 87th percentiles. For each mass bin, the splashback radius is the average R_{sp} of all haloes in that bin, as estimated from the M_{200m} masses of the haloes using the SPARTA code (Diemer 2017). We see that these ratios are roughly constant throughout the mass range $[10^{13}, 10^{15}]\,h^{-1}\,\mathrm{M}_\odot$. Taking the 87 percentile splashback radius as our reference, we find that $r_t/R_{sp} \approx 1.3$. This value is close to but somewhat smaller than the edge radius $R_{edge}/R_{sp} \approx 1.55$ identified in Aung et al. (2021). It is interesting that both the edge radius and the radius r_t defined here are roughly constant factors of the splashback radius, and that they are both somewhat larger than the splashback radius. We leave a detailed analysis of how these two different radial scales are related to future work.

As we move to smaller masses, the ratio r_t/R_{sp} grows. We caution, however, the splashback radii measured at low masses as estimated from SPARTA are likely biased for our halo population, with the bias almost certainly increasing with decreasing mass. To see this, recall that SPARTA was calibrated using parent haloes identified with the ROCKSTAR halo finder using R_{200m} as the halo radius. Since our halo boundary is significantly larger than R_{200m}, a low-mass halo neighbouring a high mass haloes will become a substructure of the high mass halo upon applying our new halo definition. These 'haloes' currently contribute to the estimates in SPARTA, but are not included

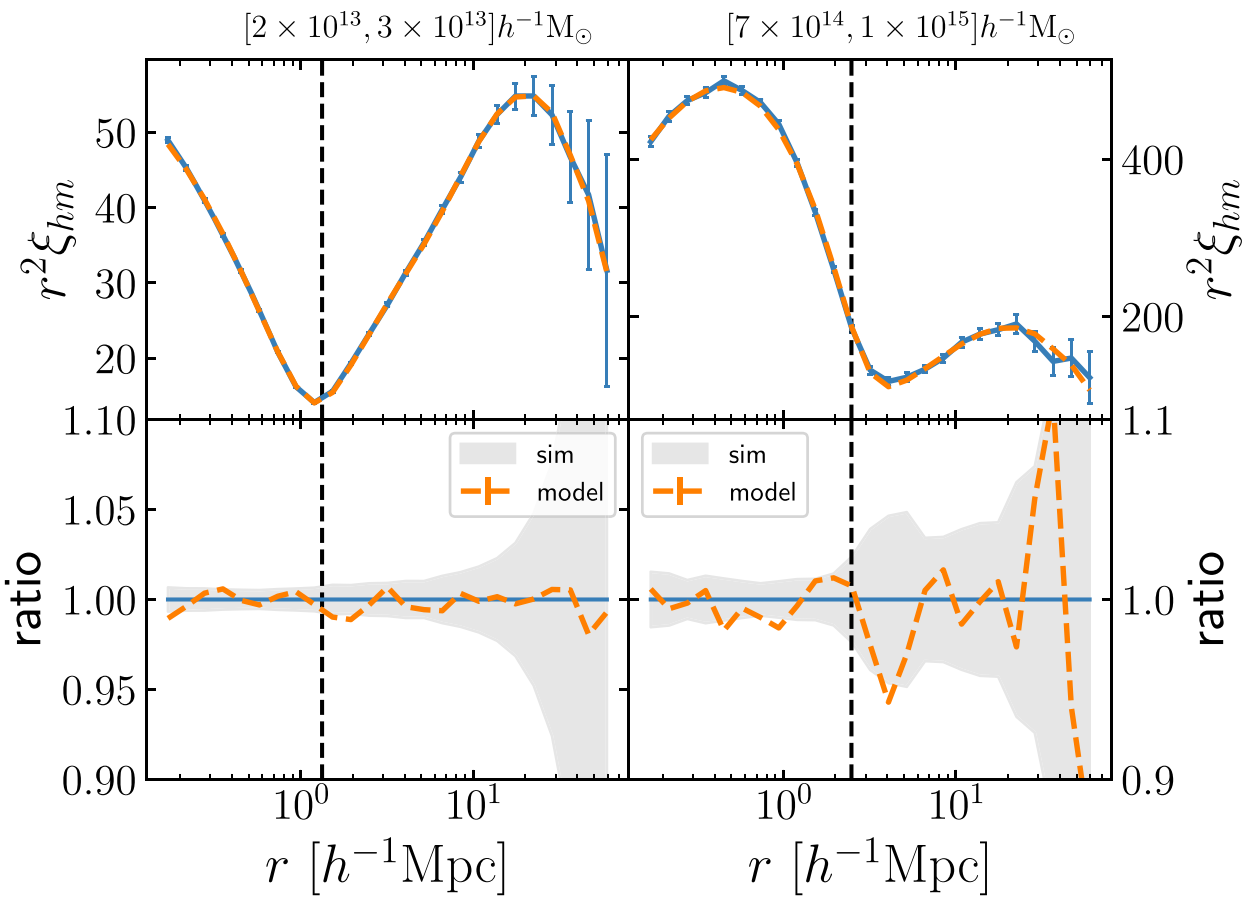

Figure 2. The halo–mass correlation function for haloes of different masses, as labelled. Top row: Halo–matter correlation functions. Bottom row: Fractional difference between the model and the measurement. Error bars are jackknife. The vertical dashed lines corresponds to $r_{\rm t}(1 + \Delta)$.

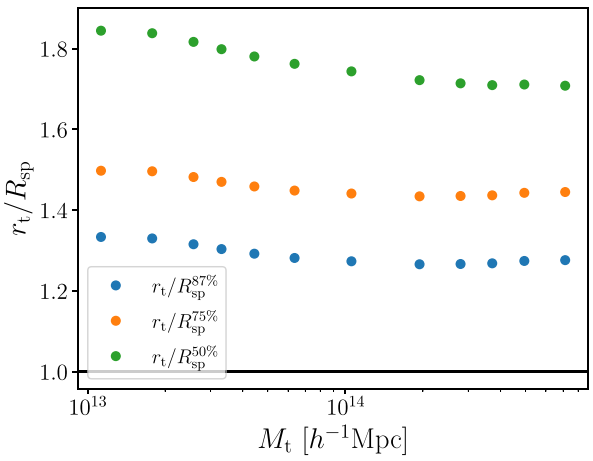

Figure 3. Ratio between our proposed halo boundary $r_{\rm t}$ and the splashback radii from the SPARTA algorithm for a variety of percentiles of the apocentre distribution of dark matter particles. The ratio seems to be roughly constant. We believe that for the halo population we identified the SPARTA estimates of the splashback radius becomes increasingly biased as we move to lower masses. See the text for further details.

in our analysis as parent haloes due to the change in percolation in our halo catalogue. In other words, the halo population in which SPARTA was calibrated does not match our halo population, except at the very highest masses. This implies that a proper comparison of the splashback radius to our proposed halo boundary $r_{\rm t}$ requires recalibration of the particle orbits based on the haloes identified by our algorithm only. We defer this recalibration to future work.

In short, we believe that splashback radii, the halo edge proposed in Aung et al. (2021), and the truncation radius we identified as naturally arising from the halo–mass correlation function are all related, though exactly what this relation is remains unclear. Clarifying the relation between these radii is an ongoing work.

3.3 The halo mass function

The change in halo definition we suggest directly impacts the halo–mass function. We measured the halo–mass function of the final halo catalogue produced by our iterative algorithm. The extended Press–Schechter formalism (Press & Schechter 1974) leads to a theoretical prediction of the halo–mass function of the form

$$\frac{\mathrm{d}n}{\mathrm{d}m} = f(\sigma)\frac{\bar{\rho}_{\rm m}}{m}\frac{\mathrm{d}\ln\sigma^{-1}}{\mathrm{d}m}, \tag{50}$$

where $f(\sigma)$ is some function, and $\sigma(M)$ is the variance of the linear density field over an aperture $r_{\rm t}(M)$. Press and Schechter (Press &

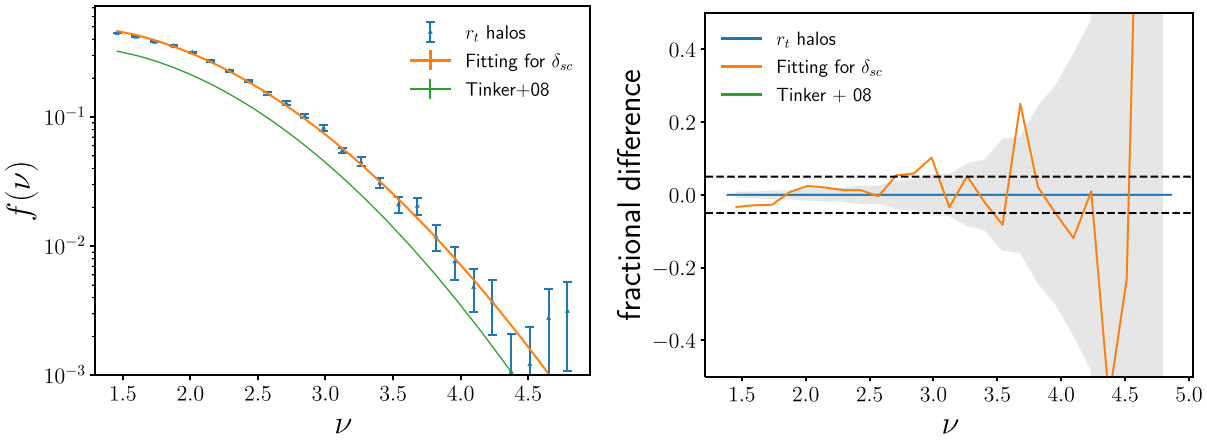

Figure 4. Left-hand panel: The halo mass function in the simulation using our proposed halo definition (blue points with error bars) and the best-fitting Press–Schechter mass function (orange line). The Tinker et al. (2008) mass function (green line) is shown for reference. Halo mass functions plotted as a function of peak height ν. Error bars are jackknife. Right-hand panel: Residuals between the simulation data and the best-fitting Press–Schechter model. The horizontal dashed lines correspond to 5 per cent deviations.

Schechter 1974) derived a first expression for $f(\sigma)$ on the basis of the spherical collapse model. The Press–Schechter multiplicity function $f(\sigma)$ is given by

$$f(\sigma) = \sqrt{\frac{2}{\pi}}\frac{\delta_{\rm sc}}{\sigma}\exp\left[-\frac{\delta_{\rm sc}^2}{2\sigma^2}\right], \tag{51}$$

where $\delta_{\rm sc}$ is the critical density required for spherical collapse. At $z = 0$, and assuming a matter density $\Omega_{\rm m} = 1$, one finds $\delta_{\rm sc} = 1.686$. The quantity $\nu \equiv \delta_{\rm sc}/\sigma(M)$ is typically referred to as the peak height.

As shown in Fig. 4, we find that the Press–Schechter mass function gives an excellent fit (~ 5 per cent precision) to the mass function of our final halo catalogue, provided we fit for the value of $\delta_{\rm sc}$. The posterior on the critical density for collapse $\delta_{\rm sc}$ from our best-fitting Press–Schechter model is $\delta_{\rm sc} = 1.449 \pm 0.004$. We note that the fidelity of our fit is not simply a consequence of the mass range sampled by the simulation: Press–Schecther fails to provide a reasonable fit to the halo–mass function in the simulation when using $M_{\rm 200m}$ as the halo mass. In this light, the excellent agreement between the simulation and the Press–Schechter mass function is surprising, as our analysis did not make any assumptions about halo abundances. Rather, it relied exclusively on features in the halo–mass correlation function to motivate the redefinition of halo boundaries. However, the best-fitting critical threshold for collapse $\delta_{\rm sc}$ is significantly smaller than expected (e.g. Pace, Meyer & Bartelmann 2017). Whether this specific value can be predicted theoretically remains to be seen.

3.4 Halo bias and the peak-background split

We have shown that the theory of spherical collapse can accurately predict the halo–mass function in a simulation, provided we use the correct halo definition and fit for the value of the critical overdensity. In this section, we test whether the peak-background split model of halo bias provides an equally accurate description of our data. The peak-background split predicts the bias as a function of peak height is given by (Cole & Kaiser 1989; Mo & White 1996)

$$b_{\rm PB}(\nu) = 1 + \frac{\nu^2 - 1}{\delta_{\rm sc}}. \tag{52}$$

We calculated the halo bias using the previous equation, where ν is the peak height as defined in the previous section. Fig. 5 shows a comparison between the halo bias as measured using the halo–mass correlation function, and the halo bias derived from the peak background split. The orange band shows the prediction based on our Press–Schechter fit to the halo–mass function. We see the peak-background split model is roughly ~ 10 per cent $-$ 15 per cent) consistent with the data, a level of accuracy comparable to the performance of the peak background split for other halo–mass definitions (e.g. Tinker et al. 2010; Hoffmann, Bel & Gaztañaga 2015; Desjacques, Jeong & Schmidt 2018). However, the predicted bias is clearly too high. We fit our data with a bias of the form derived from the peak-background split, but allow $\delta_{\rm sc}$ to vary independently, finding $\delta_{\rm sc} = 1.375 \pm 0.012$. This model can describe our data with ≈ 5 per cent accuracy, though the residuals clearly exhibit structure as a function of peak height. Note than when evaluating the bias model in equation (52), we vary $\delta_{\rm sc}$ both in the denominator and in the definition of the peak height ν. This is obviously inconsistent with the fit from the halo–mass function, but can be thought of simply as a useful empirical fitting function.

The excursion set peaks formalism of Paranjape, Sheth & Desjacques (2013) has succeeded in presenting a framework capable of jointly fitting the abundance and bias functions of dark matter haloes within the context of the peak-background split hypothesis. Critical to this success is the adoption of a mass-dependent stochastic barrier for collapse. Testing whether or not this formalism can successfully account for the discrepancy between our bias measurements and the peak-background split prediction is beyond the scope of this work.

4 SUMMARY AND CONCLUSIONS

We presented a model for the halo–mass correlation function that assumes a scale-independent bias and explicitly incorporates halo edges and halo exclusion. We emphasize that all the qualitative features in our model are well motivated a priori. The specific parametrizations used to implement these features are arbitrary (e.g. one could replace complementary error functions by Fermi–Dirac functions), but their qualitative form are not. Importantly, our model contains a single scale-independent bias parameter. The ratio of $\xi_{\rm hm}$ to $\xi_{\rm mm}$ does have a scale dependence, but this scale dependence is entirely accounted for by the modifications to the naive halo model

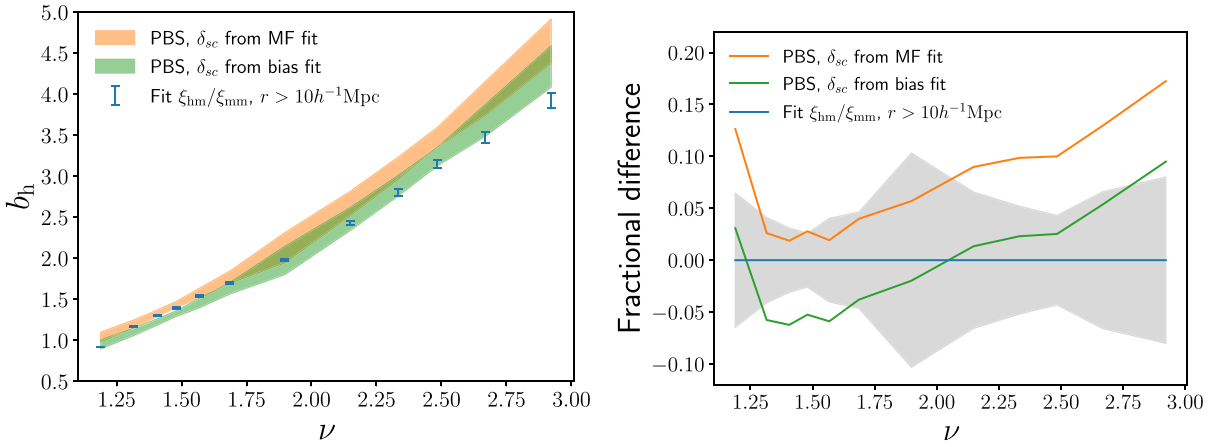

Figure 5. The halo bias measured from the ratio of the halo–mass and mass–mass correlation functions and the prediction from the peak-background split approach. Left-hand panel: The halo bias measured from fitting the halo–mass correlation function is shown in blue. The halo bias calculated using peak-background split prediction with $\delta_{\rm sc}$ the best-fitting value from the halo mass function fit is shown in orange. The green curve shows the peak-background split fit where we vary $\delta_{\rm sc}$. Right-hand panel: Fractional difference with respect to the halo bias $b = \xi_{\rm hm}/\xi_{\rm mm}$.

due to softly truncated halo profiles, halo exclusion, and the different one-halo terms of $\xi_{\rm hm}$ and $\xi_{\rm mm}$.

Our main findings can be summarized as follows:

(i) We derived the model proposed by Hayashi & White (2008) from first principles, and showed that this model is the high mass limit of a halo model that self-consistently incorporates halo edges and halo exclusion.

(ii) Halo exclusion introduces corrections in the halo–mass correlation function at the translinear regime.

(iii) In our model, there is a unique radius–mass power-law relation that can be used to define haloes for which our model provides an accurate (≈ 2 per cent) description of the halo–mass correlation function across a wide range of scales.

(iv) The halo radius identified in our analysis is located at the 'by eye' transition from the one-halo term to the two-halo term.

(v) The halo radius identified in this paper and the splashback radius (calibrated with $R_{\rm 200m}$ haloes) are related by a roughly constant multiplicative factor. However, the exact relation between these two scales, and the edge radius advocated by Aung et al. (2020), remains unclear, and is the focus of ongoing work.

(vi) The mass function of haloes defined using the halo radius identified in this work is well described by the Press–Schechter formula, though the best-fitting value for the critical density for spherical collapse $\delta_{\rm sc}$ ($\delta_{\rm sc} = 1.449 \pm 0.004$) is below its expected value $\delta_{\rm sc} \approx 1.686$.

(vii) The halo bias predictions from the peak-background split approach are not consistent with the halo bias measured from the simulation, exhibiting 10 per cent to 15 per cent offsets depend on halo mass. These differences are comparable to the deviations from the peak-background split prediction for more traditional fixed-overdensity halo definitions. Remarkably, however, if we independently fit for $\delta_{\rm sc}$ in the halo bias expression derived from the peak-background split, we find that a model with $\delta_{\rm sc} = 1.375 \pm 0.012$ can describe our data with ≈ 5 per cent accuracy.

It is very encouraging that multiple lines of evidence are now pointing towards the existence of a true halo boundary that extends well beyond $R_{\rm 200m}$, even if the precise relation between these works is still unclear. Encouragingly, we have shown that defining haloes using our proposed halo boundary significantly simplifies the halo model while improving accuracy. When coupled with new insights into the halo model, we may soon arrive at a complete theory of large-scale structure capable of describing observations at all scales, with the necessary precision required to make full use of upcoming photometric and spectroscopic surveys.

Our long-term goal is to develop this theoretical framework to the point which we can use it to perform cosmological analyses of galaxy clustering down to small scales. Doing so will require calibrating the model as a function of redshift and cosmology. These calibrations necessitate intensive simulation efforts, and building emulators that interpolate our model parameters across cosmology and redshift. Further, we must adequately characterize the precision of the resulting emulators. We will pursue this calibration, along with further theoretical improvements to our model, in future work.

This type of analytic approach may appear quaint given the existence of emulators and simulation-rescaling techniques capable of making high accuracy predictions (Nishimichi et al. 2019; Angulo et al. 2020). Moreover, given that our model needs to be calibrated in simulations, one may wonder why should not we skip the modelling part and let machine learning methods characterize the theory directly. In this context, we believe there remains significant value to the insights gained from our analytic treatment. To paraphrase Eugene Wigner, it is nice that computers can understand the problem, but we would like to understand it too.

ACKNOWLEDGEMENTS

ER and RF were supported by the DOE grant DE-SC0015975. RG is also supported by CONACyT scholarship 710106. ER also acknowledges funding from the Cottrell Scholar program of the Research Corporation for Science Advancement. This research was supported by the Munich Institute for Astro- and Particle Physics (MIAPP) which is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2094 - 390783311. The authors would like to thank Han Aung for useful comments on an early version of this manuscript. ER would like to thank Ravi Sheth, and Bhuvnesh Jain for useful discussions and suggestions regarding the content of this work.

DATA AVAILABILITY

The data underlying this article will be shared on request to the corresponding author.

REFERENCES

Angulo R. E., Zennaro M., Contreras S., Aricò G., Pellejero-Ibañez M., Stücker J., 2020, preprint (arXiv:2004.06245)
Aung H., Nagai D., Rozo E., Garcia R., 2021, MNRAS, 502, 1041
Baldauf T., Seljak U., Smith R. E., Hamaus N., Desjacques V., 2013, Phys. Rev. D, 88, 083507
Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, ApJ, 762, 109
Bhattacharya S., Heitmann K., White M., Lukić Z., Wagner C., Habib S., 2011, ApJ, 732, 122
Cole S., Kaiser N., 1989, MNRAS, 237, 1127
Cooray A., Sheth R., 2002, Phys. Rep., 372, 1
DeRose J. et al., 2019, ApJ, 875, 69
Desjacques V., Jeong D., Schmidt F., 2018, Phys. Rep., 733, 1
Diemer B., 2017, ApJS, 231, 5
Diemer B., 2018, ApJS, 239, 35
Diemer B., Kravtsov A. V., 2014, ApJ, 789, 1
Diemer B., More S., Kravtsov A. V., 2013a, ApJ, 766, 25
Diemer B., Kravtsov A. V., More S., 2013b, ApJ, 779, 159
Diemer B., Mansfield P., Kravtsov A. V., More S., 2017, ApJ, 843, 140
Einasto J., 1965, TrAlm, 5, 87
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
García R., Rozo E., 2019, MNRAS, 489, 4170
Hayashi E., White S. D. M., 2008, MNRAS, 388, 2
Hoffmann K., Bel J., Gaztañaga E., 2015, MNRAS, 450, 1674
Knebe A. et al., 2013, MNRAS, 435, 1618
Mansfield P., Kravtsov A. V., Diemer B., 2017, ApJ, 841, 34
McClintock T. et al., 2019, ApJ, 872, 53
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
More S., Diemer B., Kravtsov A. V., 2015, ApJ, 810, 36
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493.
Nishimichi T. et al., 2019, ApJ, 884, 29
Pace F., Meyer S., Bartelmann M., 2017, J. Cosmol. Astropart. Phys., 2017, 040
Paranjape A., Sheth R. K., Desjacques V., 2013, MNRAS, 431, 1503
Philcox O. H. E., Spergel D. N., Villaescusa-Navarro F., 2020, Phys. Rev. D, 101, 123520
Press W. H., Schechter P., 1974, ApJ, 187, 425
Rykoff E. S. et al., 2012, ApJ, 746, 178
Schmidt F., 2016, Phys. Rev. D, 93, 063512
Sheth R. K., Lemson G., 1999, MNRAS, 304, 767
Smith R. E., Scoccimarro R., Sheth R. K., 2007, Phys. Rev. D, 75, 063512
Smith R. E., Desjacques V., Marian L., 2011, Phys. Rev. D, 83, 043526
Springel V., 2005, MNRAS, 364, 1105
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878
Valageas P., Nishimichi T., 2011a, A&A, 527, A87
Valageas P., Nishimichi T., 2011b, A&A, 532, A4
van den Bosch F. C., More S., Cacciato M., Mo H., Yang X., 2013, MNRAS, 430, 725

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