Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **541,** 880–898 (2025) https://doi.org/10.1093/mnras/staf1029
Advance Access publication 2025 July 3

Excursion sets with a 'perfect' collapse model

A. M. Wisłocka,[1]★ J. Stücker [1,2], O. Hahn [1,3] and R. E. Angulo [2]

[1]*Institute for Astronomy, University of Vienna, Türkenschanzstraße 17, A-1180 Vienna, Austria*
[2]*Donostia International Physics Center (DIPC), Manuel Lardizabal Ibilbidea 4, E-20018 Donostia, Guipuzkoa, Spain*
[3]*Department of Mathematics, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria*

Accepted 2025 June 21. Received 2025 June 21; in original form 2025 April 9

ABSTRACT

The Lambda cold dark matter model predicts structure formation across a vast mass range, from massive clusters ($\sim 10^{15}\,\mathrm{M}_\odot$) to Earth-mass micro-haloes ($\sim 10^{-6}\,\mathrm{M}_\odot$), resolving which far exceeds the capabilities of current simulations. Excursion set models are the most efficient theoretical tool to disentangle this hierarchy in mass. We test the excursion set paradigm by combining smoothed initial density fields with a 'perfect' collapse model – N-body simulations. We find that a core excursion set assumption – small-scale perturbations do not impact larger scale collapse – is approximately fulfilled but exhibits small quantitative violations dependent on the smoothing filter. For a sharp $k-$space cut-off ~ 20 per cent of mass elements revert collapse as the smoothing scale decreases, while only 3.5 per cent do for a Gaussian and 5 per cent for a top-hat. Further, we test the simple deterministic mass-mapping $M \propto R^3$ (first-crossing scale to halo mass) relation. We find that particles that are first accreted into haloes at the same smoothing scale may end up in haloes of significantly different masses, with a scatter of 0.4–0.8 dex. We also demonstrate that the proportionality constant of this relation should be considered as a degree of freedom. Finally, we measure the mass fraction in different structure morphologies (voids, pancakes, filaments, and haloes) as a function of filter scale. Typical particles appear to be part of a large-scale pancake, a smaller scale filament, and a notably smaller halo. We conclude that validating predictions of excursion set models on a particle-by-particle basis against simulations may enhance their realism.

Key words: methods: analytical – methods: numerical – galaxies: haloes – dark matter – large-scale structure of Universe – cosmology: theory.

1 INTRODUCTION

A central prediction of the Lambda cold dark matter (ΛCDM) standard model of cosmology is that the density field of the early Universe has significant perturbations on a broad range of scales. These perturbations eventually collapse into dense dark matter haloes, with masses ranging from as large as $10^{15}\,\mathrm{M}_\odot$ (solar masses) to as small as a few $\mathrm{M}_\oplus$ (Earth masses) or even smaller, for Weakly Interacting Massive Particle (WIMP) or Quantum Chromodynamics (QCD) axion dark matter candidates.

The observed distribution of galaxies and clusters is closely linked to the formation of haloes with masses $M \gg 10^9\,\mathrm{M}_\odot$, making precise models of the halo mass function (HMF) essential for reliably interpreting observations. Further, some observational probes may depend on the presence of much smaller mass objects. For example, perturbations in the arcs of strong gravitational lenses may indicate the presence of haloes of masses $M \sim 10^7\,\mathrm{M}_\odot$ and smaller (Koopmans 2005; Vegetti & Koopmans 2009; McKean et al. 2015; Hsueh et al. 2020; Gilman et al. 2024). Further, if dark matter is composed of ~ 100 GeV WIMPs with a significant self-annihilation cross-section, it may even be possible to probe the smallest mass haloes $M \sim \mathrm{M}_\oplus$ (also known as 'prompt cusps', Diemand, Moore & Stadel 2005; Ishiyama, Makino & Ebisuzaki 2010; Bringmann & Weniger 2012; Angulo et al. 2017; Delos & White 2023a; Ondaro-Mallea et al. 2024). Such objects may contribute most of the observable self-annihilation signal and might be an important factor in the cosmic gamma-ray background (Delos & White 2023b; Stücker et al. 2023; Delos et al. 2024).

Therefore, it is important to reliably predict the abundance and properties of such structures. While the formation of high-mass objects ($M \gtrsim 10^8\,\mathrm{M}_\odot$) can be accurately traced with N-body simulations, lower mass structures are generally inaccessible in standard representative volumes due to resolution constraints. A notable exception are the VVV simulations, which resolve 20 orders of magnitude in mass scale through zoom-in simulations of highly underdense regions (Wang et al. 2020; Zheng et al. 2024). They, however, do not represent the general statistics of our Universe. Analytical approaches are ideally suited to bridging this vast range of scales, allowing for a deeper understanding of structure formation in ΛCDM beyond what is currently achievable with N-body simulations.

Arguably, the most successful analytical approaches to predict the abundance of collapsed structures, or the HMF, are the excursion set formalism (Bond et al. 1991), and peak theory (Bardeen et al. 1986). Excursion set theory emerged as an improvement to the seminal work of Press & Schechter (1974) on obtaining an HMF analytically. In this approach, the density field is smoothed with a

★ E-mail: agata.wislocka@univie.ac.at

kernel at a scale R_s, thus allowing one to trace relevant non-local properties of the environment as local properties of the smoothed field. In the smoothed density field, it is then possible to detect collapse that is associated with the scale R_s with a simplified method – for example, a spherical collapse model (Press & Schechter 1974) or an ellipsoidal model (Bond & Myers 1996). Subsequently, decreasing the smoothing scale allows one to identify which volume elements collapse on a given scale and then make a prediction for the HMF (e.g. Press & Schechter 1974; Bond et al. 1991; Sheth & Tormen 1999; Sheth, Mo & Tormen 2001; Tinker et al. 2008; Maggiore & Riotto 2010a; Despali et al. 2016). At its core, this method is conceptually very simple, yet it delivers surprisingly accurate HMFs even when implemented in its most basic form (see e.g. Zentner 2007; Lapi, Salucci & Danese 2013, for a comprehensive review).

Beyond the HMF, the excursion set method has numerous other applications in cosmology, yielding detailed predictions for various statistical quantities of structure formation. These include conditional mass functions (e.g. Shen et al. 2006; Rubiño-Martín, Betancort-Rijo & Patiri 2008; de Simone, Maggiore & Riotto 2011; Lapi et al. 2013; Tramonte et al. 2017), mass accretion, merger histories and formation times (e.g. Bond et al. 1991; Lacey & Cole 1993; Somerville & Kolatt 1999; Somerville et al. 2000; Neistein, Macciò & Dekel 2010; Simone et al. 2011; Lapi et al. 2013), and the mass–concentration relation (e.g. Ludlow et al. 2016; Bohr et al. 2021), among others.

To achieve agreement with these statistical quantities excursion set formalisms tend to employ additional degrees of freedom that often go beyond the original excursion set formulation by Bond et al. (1991) and Press & Schechter (1974). These degrees of freedom include the choice of the smoothing function for mass-mapping (Zentner 2007; Maggiore & Riotto 2010a; Leo et al. 2018; Delos 2024) or the first-crossings distribution function (Bond et al. 1991; Zentner 2007; Maggiore & Riotto 2010b; Farahi & Benson 2013; Schneider, Smith & Reed 2013; Delos et al. 2024), implementations of stochastic barriers (e.g. Maggiore & Riotto 2010b; Corasaniti & Achitouv 2011; Ma et al. 2011; Achitouv et al. 2013, 2014), direct modifications of the spherical collapse threshold (Monaco 1995; Eke, Cole & Frenk 1996; Jenkins et al. 2001; Delos 2024; Verza et al. 2024), accounting for three-dimensionality of collapse (Bond & Myers 1996; Chiueh & Lee 2001; Sheth et al. 2001; Lee 2006; Shen et al. 2006; Sandvik et al. 2007; Stücker, Busch & White 2018; Musso, Despali & Sheth 2024), and other approaches to environmental dependence (Catelan et al. 1998; Taruya & Suto 2000). In particular, ellipsoidal collapse models are especially compelling, as they may additionally capture the anisotropy of collapse of fluid elements. This anisotropy, also observed in simulations, leads to the formation of pancakes, filaments, and haloes (Sheth et al. 2001, 2006; Sandvik et al. 2007; Ludlow, Borzyszkowski & Porciani 2014). However, since an accurate calculation of the triaxial effects on the evolution of fluid elements is rather challenging, Sheth et al. (2001) and Sheth & Tormen (2002) proposed an approximate approach that captures the dynamics statistically, with a set of free parameters and a moving barrier. If the parameters are calibrated against simulations they allow to describe HMF to very high accuracy at the qualitative level. For this reason this approach has achieved significant popularity. None the less, this approach remains an oversimplification of the complex halo formation dynamics, with significant quantitative discrepancies reported between simulations and this model (Reed et al. 2006; Robertson et al. 2009).

While more complex excursion set approaches offer valuable high-accuracy descriptions of statistical quantities, it remains uncertain whether this improved accuracy stems from a more realistic physical model or simply from the increased number of degrees of freedom. For example, the additional parameters in approximate ellipsoidal collapse models might effectively compensate for other shortcomings in the approach. Contrary to the common view, Delos (2024) recently argued that this improvement should not be attributed to the improved physical accuracy in ellipsoidal collapse modelling. In particular, they demonstrated that the spherical collapse HMFs agree nearly exactly with the analytical prediction from ellipsoidal collapse of Sheth et al. (2001), when generated with the top-hat and Gaussian kernels. This exemplifies how the congruence between the simulation and excursion set HMFs is an offshoot of assuming Markovian walks (sharp $k-$space smoothing) in density, not the veracity of the ellipsoidal collapse model. Similarly, in a series of publications, Lucie-Smith et al. (2018), Lucie-Smith, Peiris & Pontzen (2019), and Lucie-Smith et al. (2024) concluded that accounting for the tidal effects in the ellipsoidal collapse models has minimal impact on the HMFs. At the same time, experiments like those of Ludlow & Porciani (2011) and López-Cano et al. (2024) do uphold the claim of the importance of accounting for the environmental tidal effects in developing improved analytical methods for halo formation. Determining the true source of improvement is in particular relevant when making predictions about aspects that have not been rigorously tested. Less explored areas include predictions related to the cosmic web (Shen et al. 2006; Sandvik et al. 2007), extrapolations to scales that remain unresolved in simulations (Angulo & White 2010; Stücker et al. 2018; Liu et al. 2024), or particle-by-particle predictions of the evolution of individual mass elements (Bond et al. 1991; White 1994; Sheth et al. 2001; Monaco, Theuns & Taffoni 2002a; Monaco et al. 2002b).

Here, we propose a novel approach to reliably test the fundamental assumptions that are common in all excursion set formalisms. For this, we evaluate the performance of the excursion set approach if the collapse model – which is typically described by a spherical or ellipsoidal collapse approximation – is instead replaced by a fully non-linear simulation with a structure identification algorithm. We will refer to this as an excursion set with a 'perfect' collapse model. The resulting excursion set draws an upper limit to the degree of realism that may be achieved by excursion set models. Higher accuracy may only be achieved if model parameters compensate for the shortcomings in the excursion set paradigm itself (e.g. Verza et al. 2024).

We would like to emphasize that this type of analysis only became possible due to the phase-space simulations that do not suffer from artificial fragmentation, combined with a tessellation-based structure finder, which allows the identification of the morphology (haloes, sheets, etc.) of fluid elements (Stücker et al. 2020, 2022). These breakthroughs in computational cosmology enable us to directly test certain assumptions of the excursion set theory for the first time. In this regard, our work complements that of Robertson et al. (2009) in the ongoing effort to test the groundwork of the excursion set theory.

This article is organized as follows: In Section 2, we summarize key ingredients that are common among all excursion set formalisms. In Section 3, we describe the details of our simulations, how collapse along different axes is detected, and how we can use simulations as a 'perfect' collapse model. In Section 4, we present our measurements and test how well different assumptions of excursion sets are fulfilled. Finally, in Section 5, we discuss the inherent limitations of the

excursion set formalisms. We also discuss which aspects of the excursion set formalisms can be improved to enhance their realism.

2 EXCURSION SET MODELS

Excursion set formalisms aim to disentangle the hierarchy of structure formation through four core ingredients:

(i) The smoothed linear density field is considered a function of the smoothing scale R – allowing the tracking of relevant non-local properties of an object's environment as local properties of the smoothed environment.
(ii) A collapse criterion distinguishes between fluid elements that are part of a gravitationally collapsed structure and those that are not.
(iii) It is assumed that structure formation can be represented through an excursion set across smoothing scales, i.e. a mass-element is part of a structure given by the largest smoothing scale at which the collapse criterion is fulfilled.
(iv) To define mass functions, it is assumed that the smoothing scale of collapse can be mapped to a halo mass through a simple one-to-one deterministic relation.

In this section, we will briefly summarize these core concepts of excursion set formalisms and point out a few popular examples. Finally, we will motivate that a simulation may be used as a 'perfect' collapse model to test the core ingredients of excursion sets. By 'perfect', we mean that no simplifying assumptions are made – such as approximating collapse as spherical or ellipsoidal with a simple external tide model, etc.

2.1 Scale space

In excursion set formalisms, it is common to define a smoothed linear density field $\delta_R(\boldsymbol{q})$ – where $\boldsymbol{q}$ are Lagrangian coordinates and R is a smoothing scale – by convolving with a filter function

$$\delta_R = W_R \circledast \delta. \tag{1}$$

If one considers the space of Lagrangian coordinates plus the smoothing scale $(\boldsymbol{q}, R)$, one may speak of a 'scale-space'. The smoothed density field can be conveniently evaluated in Fourier space

$$\widetilde{\delta}_R = \widetilde{W}_R \cdot \widetilde{\delta}, \tag{2}$$

where a tilde designates a Fourier transformation.

Here, we will consider a few common choices of the filter function as the sharp k-space

$$\widetilde{W}_{\rm sk}(kR) = \Theta(k - k_R)\,, \tag{3a}$$

$$\text{with } k_R = R^{-1}, \text{ the top} - \text{hat} \tag{3b}$$

$$\widetilde{W}_{\rm th}(kR) = \frac{3\,[\mathrm{Sin}(kR) - kR\,\mathrm{Cos}(kR)]}{(kR)^3} \tag{3c}$$

and the Gaussian (3d)

$$\widetilde{W}_{\rm g}(kR) = \exp\left[-\frac{(kR)^2}{2}\right], \tag{3e}$$

where θ is the Heavyside step-function. Beyond the linear density field at the scale of consideration, the variance is of particular relevance

$$\sigma_R^2 = \frac{1}{2\pi^2}\int_0^\infty k^2\,P(k)\,\widetilde{W}^2(kR)\,\mathrm{d}k\,. \tag{4}$$

We will briefly discuss some of the characteristics of the above window functions. For a more in-depth discussion, we refer the reader to Bond et al. (1991) and Maggiore & Riotto (2010a). For the case of a sharp-k filter, $\widetilde{W}_{\rm sk}$, δ_R performs a Markovian random walk as a function of R. This makes it the only filter for which purely analytic expressions for the HMFs have been obtained. Conversely, any other filter leads to correlated random walks, which makes a fully analytical treatment difficult. The main drawback of the sharp$-k$ function is its awkward shape in configuration space, with wiggly features extending to infinity, resulting in an ill-defined volume and by extension, the mass linked to it. On the other hand, the top-hat function $\widetilde{W}_{\rm th}$ seems to be the most natural choice when working with a spherical collapse model, with the mass associated with a collapsed Lagrangian patch well defined in real space. However, due to its sharp feature in configuration space it mirrors the issue of $\widetilde{W}_{\rm sk}$ in Fourier space, in turn giving rise to wiggly features in smoothed power spectra. The only other commonly used function that avoids the issues of sharp boundaries is the Gaussian window function $\widetilde{W}_{\rm g}$, which has the convenience of being well defined in configuration and Fourier space, although its reach spans all space. In contrast to the other two filters, the Gaussian function is also somewhat intuitive if one considers more realistic density profiles of haloes, with density peaking in the centre, gradually dropping with distance, with no sharp (top-hat) or underdense (sharp$-k$) features weaved into the profile.

In general, filter functions are chosen based on the goal of recovering some target statistic. For example, Schneider et al. (2013) have demonstrated that, of these three filters, the sharp $k-$space filter is the only one to recover the asymptotic low-mass behaviour of HMFs in warm dark matter cosmologies with a cut-off in the power spectrum, as observed in simulations (Benson et al. 2013; Lapi et al. 2013; Schneider et al. 2013; Schneider 2015; Leo et al. 2018). In contrast, Delos (2024) has shown that for linear power spectra without a cut-off, i.e. CDM cosmologies, the top-hat window function gives very accurate results for many scenarios when no approximations are used in the excursion set formalism.

Clearly, there is no unique optimal window function for all scenarios. Therefore, in this article, we will test how well it is justified to disentangle structure formation in scale space for all three window functions.

2.2 The collapse criterion

Excursion set formalisms define a collapse criterion that distinguishes between fluid elements that are assumed to collapse and those that do not. While the primary focus of most formulations of the excursion set (Bond et al. 1991) is to distinguish between haloes and non-haloes, it is possible to consider more general collapse definitions that classify collapsed structures into voids, pancakes (or walls), filaments, and haloes. To handle these structures systematically, we define a *morphology rank* as a number n that distinguishes structures as follows: 0 = *void*, 1 = *pancake*, 2 = *filament*, and 3 = *halo*. We may understand n as the number of axes of a volume element that have collapsed (see e.g. Falck, Neyrinck & Szalay 2012). This definition implicitly assumes that a volume element within a halo has previously undergone collapse into a filament and a pancake, representing another form of hierarchical structure formation.

We may then define a collapse criterion as a functional that maps a smoothed linear density field on to a morphology rank.

$$c : \delta_R(\boldsymbol{q} - \boldsymbol{q}_0) \to n \in \{0, 1, 2, 3\}, \tag{5}$$

which may be evaluated separately for each location of interest $\boldsymbol{q}_0$. Note that this definition includes translation invariance and, as a functional, allows to incorporate aspects of the linear density field beyond the local density $\delta_R = \delta_R(\boldsymbol{q} - \boldsymbol{q}_0 = 0)$, such as the deformation tensor.

While we do not use analytical collapse models in this study, we still want to mention two important examples for illustration: the most commonly adopted collapse criterion is that of spherical collapse (Press & Schechter 1974), which assumes a volume element to be collapsed into a halo if the smoothed linear density exceeds a value of $\delta_c = 1.686$:

$$c_{\rm sc} = \begin{cases} 3 & \text{if} \quad \delta_R \geq \delta_c\,, \\ 0 & \text{otherwise.} \end{cases} \tag{6}$$

More sophisticated triaxial or ellipsoidal collapse models may distinguish between the collapse along each of the three axes, based on the deformation tensor $d_{R,ij} = (\partial_i \partial_j / \nabla^2)\delta_R$:

$$c_{\rm ec} = \begin{cases} 3 & \text{if} \quad \delta_R \geq f_{\rm ec,h}(e_R, p_R)\,, \\ 2 & \text{else if} \quad \delta_R \geq f_{\rm ec,f}(e_R, p_R)\,, \\ 1 & \text{else if} \quad \delta_R \geq f_{\rm ec,p}(e_R, p_R)\,, \\ 0 & \text{otherwise,} \end{cases} \tag{7}$$

with three separate barriers for pancake, filament, and halo formation that depend on the ellipticity e_R and the prolateness p_R of the deformation tensor $d_{R,ij}$, which are defined through its eigenvalues $\lambda_1 \geq \lambda_2 \geq \lambda_3$ as $e_R = (\lambda_1 - \lambda_3)/2(\lambda_1 + \lambda_2 + \lambda_3)$ and $p_R = (\lambda_1 + \lambda_3 - 2\lambda_2)/2(\lambda_1 + \lambda_2 + \lambda_3)$. Examples of triaxial models are the tide-free ellipsoidal collapse model of White & Silk (1979), the ellipsoidal collapse model with tides by Bond & Myers (1996), the ellipsoidal approximation (only for halo formation) from Sheth et al. (2001), the (shape-free) triaxial collapse model in Stücker et al. (2018), or the Lagrangian perturbation theory expansion of the triple sine-wave considered by Rampf et al. (2023).

In principle, collapse models could incorporate even higher order aspects of the linear field into the collapse criterion. However, it is unclear a priori how fruitful this approach would be, as other aspects of an excursion set model, such as mass assignment (Sheth & Tormen 1999; Verza et al. 2024), might fail independently of the degree of realism employed in the collapse model.

2.3 The excursion set assumption

Once a collapse criterion has been defined, it is possible to test whether this criterion is satisfied when applying filters to the linear field at different smoothing scales R. We will write 'c_R' as a short-hand for the collapse criterion evaluated on the density field at smoothing scale R. Generally, we may encounter different classifications at different length-scales. For example, according to one criterion, it may be that a particle is considered part of a halo $c_{R_1} = 3$ on some scale R_1, but outside of a halo $c_{R_2} < 3$ on a larger scale $R_2 \geq R_1$. It is the core assumption of the excursion set formalisms that a particle has the morphology rank n_R, if it was determined to have been collapsed to a structure of this rank on *any* larger scale:

$$n_R = \sup\{c_{R'} \mid R' \geq R\}. \tag{8}$$

In other words, this would mean that the particle is predicted to be part of a large-scale halo and that the density perturbations on smaller scales are irrelevant to the collapse question.

Note that here we have defined an excursion set on the *morphology rank*, whereas the original spherical collapse formulation by Bond et al. (1991) was defined as an excursion set on the *smoothed linear density field* δ_R, which is then compared with the spherical collapse barrier δ_c (giving a classification n_R). In the case of spherical collapse, these two definitions are equivalent. However, defining the excursion set on the morphology rank provides a clear extension for more general collapse models.

Each collapsing axis has an associated length-scale at which the corresponding collapse criterion is first fulfilled. In particular, we can define the Lagrangian length-scales of pancakes, filaments, and haloes:

$$R_{\rm pancake} = \sup\{R \mid c_R \geq 1\} \tag{9a}$$

$$R_{\rm filament} = \sup\{R \mid c_R \geq 2\} \tag{9b}$$

$$R_{\rm halo} = \sup\{R \mid c_R = 3\} \tag{9c}$$

If the corresponding sets are empty, indicating that a particle has not collapsed at any length-scale, these radii are considered undefined. Note that these are Lagrangian scales and are therefore related to mass scales, but they should not be mistaken for the Eulerian extent of these structures, as the two can differ significantly. Further, note that the definitions imply that strictly $R_{\rm pancake} \geq R_{\rm filament} \geq R_{\rm halo}$. That is, every halo is embedded in a larger scale filament, which in turn is part of an even larger scale pancake. At sufficiently large scales, these structures are not collapsed at all. The latter aspect is, for instance, utilized in the EFTofLSS (effective field theory of large-scale structure) models (cf. Baumann et al. 2012; Carrasco, Hertzberg & Senatore 2012), where all highly non-linear scales are integrated out in order to guarantee the applicability of perturbative techniques on large-enough scales.

2.4 The mass map

The previous definitions already allow for the association of each fluid element with a scale of collapse (for a given collapse model). However, to obtain predictions, e.g. of the HMF, it is additionally necessary to relate the collapse scale to a mass scale. Therefore, all excursion set formalisms (that we know of, with the exception of Hahn & Paranjape 2014) assume that if a fluid element first satisfies the halo collapse criterion at a smoothing scale R, then it is part of a halo with mass

$$M_{\rm halo} = M(R_{\rm halo})\,. \tag{10}$$

It is worth stressing that this relation is assumed to be *deterministic* and that all fluid elements that fulfil the collapse criterion at the same smoothing scale become part of a halo of the same mass. This is a point that we will investigate in more detail later. Generally, all implementations assume that the mass map should scale as $M(R) \propto R^3$. However, there is no universally accepted method for defining the proportionality constant of the relation, except in the case of the top-hat filter. Assuming a unit filter normalization, $\int W_R = 1$ (Maggiore & Riotto 2010a), one has

$$\frac{M(R)}{\rho_{m,0}\, R^3} = \begin{cases} 6\pi^2 & \text{sharp } k\text{-space}, \\ 4\pi/3 & \text{top-hat}, \\ (2\pi)^{3/2} & \text{Gaussian}, \end{cases} \tag{11}$$

where $\rho_{\rm m,0}$ is the mean matter density at $z = 0$. We will also assume (11) as the baseline for the mass-map in our analysis.

2.5 Excursion sets with a perfect collapse model

In general, excursion set formalisms assume simplistic collapse models to describe aspects of the evolution of the fully non-linear density field, such as the HMF. Inaccuracies in the description of the modelled quantities have often been attributed to shortcomings in the realism of collapse models. For instance, it is widely accepted that an ellipsoidal collapse model instead of spherical collapse leads to an improved description of the HMF (Sheth et al. 2001).

In the next section, we will discuss how adopting a simulation as a collapse model is possible. This allows us to test how well excursion sets perform if the degree of realism of the collapse model is maximized. For simplicity, we call this a 'perfect' collapse model. This approach allows us to distinguish between inaccuracies arising from the realism of the collapse model and the limitations arising from the mathematical structure of the formalism outlined above.

3 SIMULATION AS A COLLAPSE MODEL

In this section, we describe how a fully non-linear simulation can serve as a collapse model within an excursion set formalism. In the mathematical language we adopted in Section 2, any functional that predicts morphology rank for a given linear density field qualifies as a collapse model. To use a simulation as a collapse model, we may therefore smooth the linear density field at a scale R, run a simulation with such modified initial conditions, and subsequently determine collapsed regions in the simulated fields to define a morphology rank for each particle. The excursion set may then be obtained by simulations with initial conditions smoothed on many scales.

In Sections 3.1–3.3 we will describe how we infer the morphology rank for particles in a simulation and in Section 3.4, how we set up a large number of simulations to mimic excursion sets.

3.1 Equations of motion

According to the current paradigm, DM is collisionless, cold, and made up of purely gravitationally (or possibly also weakly) interacting particles (Bertone & Tait 2018). The equations of motion (EoMs) for such non-relativistic, collisionless, and self-gravitating matter in terms of a (co-moving) position $\boldsymbol{x}(\boldsymbol{q}, t)$ and canonically conjugate momentum $\boldsymbol{p}(\boldsymbol{q}, t)$ in terms of cosmic time t and indexed by a Lagrangian coordinate $\boldsymbol{q}$ (defined as $\boldsymbol{q} = \boldsymbol{x}(\boldsymbol{q}, t = 0)$) are given by the following ordinary differential equations (e.g. Peebles 1981)

$$\dot{\boldsymbol{x}}(\boldsymbol{q}, t) = a^{-2}(t)\, \boldsymbol{p}(\boldsymbol{q}, t) \tag{12a}$$

$$\dot{\boldsymbol{p}}(\boldsymbol{q}, t) = -a^{-1}(t)\, \boldsymbol{\nabla}_{\boldsymbol{x}} \phi(\boldsymbol{x}, t) \tag{12b}$$

$$\boldsymbol{\nabla}_{\boldsymbol{x}}^2 \phi(\boldsymbol{x}, t) = 4\pi\, G \rho_{\mathrm{m},0}\, \delta(\boldsymbol{x}, t)\,, \tag{12c}$$

where a dot denotes a derivative w.r.t. cosmic time, $a(t)$ is the scale factor, $\phi(\boldsymbol{x}, t)$ is the peculiar potential, $\delta(\boldsymbol{x}, t) = \frac{\rho(\boldsymbol{x}, t) - \rho_{\mathrm{m}}(t)}{\rho_{\mathrm{m}}(t)}$ is the matter density contrast, $\rho_m(t)$ is the time-dependent background matter density. These EoMs are solved in the L-GADGET3 N-body code that we use for our numerical simulations (Springel 2005; Angulo et al. 2012). The perfectly cold limit is already implicit in our notation, in that we think of the fluid as dispersionless in Lagrangian space, i.e. initially, the momentum distribution at location $\boldsymbol{q}$ is single-valued.

In this study, we deliberately exclude any baryonic effects, as the primary objective is to evaluate the assumptions of the excursion set theory only in the context of DM structure formation. Moreover, incorporating baryonic effects is unlikely to significantly influence the simulation outcomes related to the aspects of our investigation. For the smallest DM haloes ($\mathrm{M} \sim \mathrm{M}_{\oplus}$) baryons tend to stream out of the halo potentials, leaving these structures unaffected.

3.2 The distortion of fluid elements

Given its cold nature, i.e. the lack of significant primordial velocity dispersion, the DM distribution is reduced to a three-dimensional sub-manifold (parametrized by the Lagrangian coordinate $\boldsymbol{q}$) embedded in a six-dimensional phase-space; this sub-manifold is often referred to as the DM *phase-space sheet* (Abel, Hahn & Kaehler 2012; Shandarin, Habib & Heitmann 2012).

The EoMs (12) imply a deformation (or distortion) of each fluid element comprising this sheet, which is described by the Jacobian matrix of the Lagrangian map

$$\mathsf{D}_{\boldsymbol{xq}}(\boldsymbol{q}, t) := \tfrac{\mathrm{d}\boldsymbol{x}(\boldsymbol{q}, t)}{\mathrm{d}\boldsymbol{q}}\,. \tag{13}$$

Note that in general, $\mathsf{D}_{\boldsymbol{xq}}$ is not a symmetric tensor. However, the antisymmetric part only represents an axis rotation, hence it only becomes important in a highly non-linear setting, i.e. post shell-crossing.

As the phase-space sheet evolves, it gradually creases under the influence of gravity. In this description, the distortion tensor is directly related to the relative density change of an infinitesimal volume of the DM sheet (i.e. fluid element) as the sheet changes shape. As a direct consequence of mass conservation, the fluid particle density is given by the so-called *stream* density

$$\rho_s(\boldsymbol{q}, t) = \frac{\rho_m(t)}{|\det \mathsf{D}_{\boldsymbol{xq}}(\boldsymbol{q}, t)|}\,. \tag{14}$$

At early times, there is a one-to-one mapping between Lagrangian and Eulerian space. At later times, the mapping will generally become multivalued (i.e. 'multistream'), and multiple Lagrangian fluid elements will map to the same Eulerian coordinate. Specifically, if one axis collapses along a given direction $\boldsymbol{\xi}$, then $\mathsf{D}_{\boldsymbol{xq}} \cdot \boldsymbol{\xi}$ flips its sign at that instant, invalidating the bijective property of the Lagrangian mapping. In other words, the fluid element shell-crosses along $\boldsymbol{\xi}$. This, in turn, results in a singularity in ρ_s, as the determinant will be zero at that moment. The distortion tensor thus entails all (infinitesimal) information about the three-dimensional deformation and possible triaxial collapse of each fluid element.

In the literature, two distinct methods have been discussed that can yield the evolution of $\mathsf{D}_{\boldsymbol{xq}}$ in a fully non-linear cosmological setting, with general random initial conditions. These are the GDE approach (geodesic evolution equation; Vogelsberger et al. 2008) and the Lagrangian sheet tessellation approach (Abel et al. 2012; Shandarin et al. 2012; Hahn & Angulo 2016; Sousbie & Colombi 2016; Stücker et al. 2020). In this paper, we will focus on the former.

Unlike a standard N-body simulation, which follows only particle positions and momenta, the advantage of the simulation we employ for this study is that it follows the per-particle temporal evolution of the distortion tensor with the GDE,

$$\dot{\mathsf{D}}_{\boldsymbol{xq}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\boldsymbol{q}}\right) = a^{-2}\, \mathsf{D}_{\boldsymbol{pq}} \tag{15a}$$

$$\dot{\mathsf{D}}_{\boldsymbol{pq}} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}\boldsymbol{q}}\right) = a^{-1}\, \mathsf{T} \cdot \mathsf{D}_{\boldsymbol{xq}}\,, \tag{15b}$$

where $\mathsf{D}_{\boldsymbol{pq}} := \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}\boldsymbol{q}}$ and $T_{ij} = -\,\partial_i \partial_j \phi$ is the tidal tensor. For an in-depth description of the distortion tensor and the GDE, refer to Vogelsberger et al. (2008) and Vogelsberger & White (2011).

Our modified version of L-GADGET3 (cf. Stücker et al. 2020, for details) uses a standard cosmic-time second-order leapfrog algorithm

to integrate equations (15) in time. This enables us to follow the volume distortion over time and infer from it the cosmic web morphology class of all simulation particles.

3.3 Detecting axis collapse

There is an evident link between axis collapse and the morphological characteristics – voids, pancakes, filaments, and haloes – of the cosmic web, within which simulation particles are embedded (Zel'dovich 1970; Bond, Kofman & Pogosyan 1996; Falck et al. 2012; Ramachandra & Shandarin 2017; Shandarin & Medvedev 2017; Stücker et al. 2020). In essence, by tracing particles from their initial positions, one can tell which structures they end up in, while at the same time recording the correlation between their position and the number of collapsed axes. This tracing of particle morphology/collapse constitutes a substantial aspect of the excursion set theory. Since we will frequently reference these structural components of the cosmic web throughout the text, this subsection details the methodology adopted by Stücker et al. (2020) for classifying fluid elements into the four morphological categories within the simulation code utilized in this study.

Following $\mathsf{D}_{\boldsymbol{xq}}$ over time provides insight into the infinitesimal volume deformation and rotation, equations (15), at discrete locations centred on the N-body particle positions (Stücker et al. 2020). The non-symmetric aspects of the tensor describe rotations, while the symmetric aspects represent the volume deformation along the three dimensions. In the absence of rotations, $\mathsf{D}_{\boldsymbol{xq}}$ is symmetric and therefore can be diagonalized with real eigenvalues, in which case the diagonal elements are simply the eigenvalues that, in turn, describe the stretching of the axes in the principle axis frame (i.e. the initial orientation of the axes). This rotation-less scenario occurs as long as the evolution remains linear. However, the growth of gravitational instability in time leads to non-linear evolution – activating the non-symmetric components of $\mathsf{D}_{\boldsymbol{xq}}$ – and finally collapse along one, two, or all three of the axes. As the evolution of the sheet progresses from linear to non-linear, the orientation of fluid elements inevitably changes, as the DM sheet begins to twist and eventually folds. This folding marks the moment of shell-crossing, and a fluid element at its location is turned inside out.

This sign flip, or rotation, can be traced by comparing the initial (Lagrangian) and final (Eulerian) orientation of a fluid element. The simulations of Stücker et al. (2020), which we employ, use the degree of rotation as a classification scheme for collapse. Specifically, they consider the singular-value decomposition (SVD) of $\mathsf{D}_{\boldsymbol{xq}}(t)$:

$$\mathsf{D}_{\boldsymbol{xq}} = \mathsf{U}\mathsf{S}\mathsf{V}^{\top}\,, \tag{16}$$

where S is a diagonal matrix with singular values s_i along the diagonal, and U and V are orthogonal matrices, where V and U impart the orientation to a volume element in the Lagrangian and the Eulerian configurations, respectively. Simply put, the SVD can be thought of as a generalization of the eigenvalue decomposition to non-symmetric matrices, so that here, the s_i quantifies the degree of stretching of the principal axes. They define the rotation angles

$$\alpha_i = \cos^{-1}(\boldsymbol{v}_i \cdot \boldsymbol{u}_i)\,, \tag{17}$$

where $\boldsymbol{v}_i$ and $\boldsymbol{u}_i$ are the column vectors of V and U. The α_i indicate the degree of rotation of those axes with respect to their initial orientation. Stücker et al. (2020) have found that particles that are part of different types of structures (voids, pancakes, filaments, and haloes) show very different evolution of α_i. Along the axes where a structure is not collapsed α_i is very close to zero, whereas its value along collapsed dimensions tends to vary rapidly between 0 and π.

For this reason, the simulation traces the time-maximum angle

$$\alpha_{\mathrm{max},i}(t) = \sup\{\alpha_i(t') \mid t' \leq t\} \tag{18}$$

to determine axis collapse. If the maximum angle of an axis exceeds a threshold of $\nu = \pi/4$ this axis is classified as collapsed. Following our notation introduced in Section 2.2 we can define the number of collapsed axes identified by the simulation for each fluid element:

$$n = \sum_{i=1}^{3} \Theta(\alpha_{\mathrm{max},i} - \nu)\,, \tag{19}$$

where Θ is the Heaviside step function. Stücker et al. (2020) observed that n is relatively insensitive to the precise value of ν, as $\alpha_{\mathrm{max},i}$ is typically either close to 0 for axes that have not collapsed or close to π for those that have. Further, they have found that this classification is well able to distinguish between different morphologies and we will assume that n corresponds to the morphology rank so that structures can be distinguished based on n as described in Section 2.2 ($0 = \textit{void}$, $1 = \textit{pancake}$, etc.) We refer the reader to Stücker et al. (2020) for further details.

Note that in our set-up, n may refer either to the target morphology or to a collapse classification criterion (denoted as 'c' earlier) depending on the context. We will clarify this in the corresponding places in the text.

3.4 An excursion set of simulations

For the purpose of our investigation we build a simulation set that mimics an excursion set. We do this by evaluating the above-described morphology classification for a variety of linear density fields that are smoothed on different scales by simulating each of those initial conditions.

We initialize each of our simulations at $z = 49$ ($a = 0.02$) and we set the cosmological parameters to $\Omega_{\mathrm{m}} = 0.3051$, $\Omega_{\Lambda} = 0.6948$, $h = 0.676$, $n_s = 0.961$, and $\sigma_8 = 0.8154$. For all simulations, we use a box size of $L = 40\,h^{-1}$Mpc with $N = 256^3$ particles, of mass $M \simeq 3 \times 10^8\,\mathrm{M}_{\odot}h^{-1}$, to trace the dark matter sheet and a larger number of up to $N = 512^3$ released N-body particles. Since, in this work, we are not concerned with the accuracy of the internal structure of haloes and our interest lies simply in identifying the collapse times of various structures, we calculate forces only with a pure particle mesh approach with $N_{\mathrm{pm}} = 512^3$ grid elements leading to a force resolution of $\epsilon = 70$ kpc. The choice of these settings was guided by carrying out a convergence test, which demonstrated that a higher particle mesh resolution does not affect collapse times significantly (see Appendix A for details).

For our investigation, we choose a set of 16 different smoothing scales R_i for each of the three window functions discussed in Section 2.1, so that $\sigma_i^2 := \sigma_{R_i}^2$ takes on a range of pre-specified values. We modified the MUSIC (Hahn & Abel 2011) program to use equation (4) together with a ΛCDM power spectrum $P(k)$ to output fields truncated with the appropriate cut-off, as shown in the top panel of Fig. 1. We select the target variances as $\sigma_i^2 = i$, where $i = [1, ..., 16]$, and throughout this paper, we refer to each simulation by its corresponding target variance. Note that the measured variance of the initial conditions $\langle\delta_R^2\rangle$ often turned out slightly different to the target variance due to cosmic variance. We show the relation between

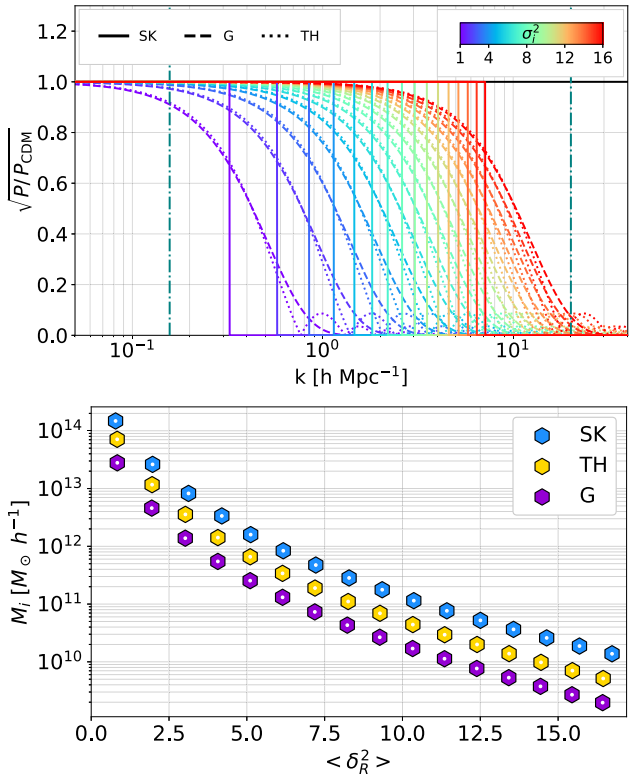

Figure 1. Top: the initial power spectra ($z = 49$) generated with the $\widetilde{W}_{\rm sk}$ (continuous), $\widetilde{W}_{\rm th}$ (dotted), and $\widetilde{W}_{\rm g}$ (dashed) smoothing on 16 different scales, normalized to the CDM spectrum (black). Each spectrum uses a sharp cut in Fourier space to approximately achieve the target variance σ_i^2. The two teal dash–dotted lines indicate the min. (left) and max. (right) k modes of our simulation volume. Bottom: the characteristic mass associated with each k_R scale as a function of the measured variance of the simulation for the sharp k-space (SK), top-hat (TH), and Gaussian (G) filters.

the variance of the density field and the characteristic mass scale as in equation (11) in the bottom panel of Fig. 1.

4 EMULATION EXCURSION SETS WITH SIMULATIONS

Emulating excursion sets with our suits of simulations allows us to test basic assumptions of the excursion set theory independent of the assumed collapse model.

4.1 The excursion set assumption

The assumption that structure formation can be represented as an excursion set that progresses from large to small scales (see Section 2.3) is a powerful tool for understanding the hierarchical emergence of structures. In particular, it simplifies the picture by decoupling the formation of large-scale structures from density perturbations on smaller length-scales. We may phrase the necessary assumption:

If a volume element appears to collapse in the linear field smoothed on some scale, then it will become part of a structure of that length-scale – independently of the perturbations on smaller scales.

This assumption may be stated in condensed form as

$$n_{R_2} \geq n_{R_1} \quad \text{if} \quad R_2 \leq R_1 \tag{20}$$

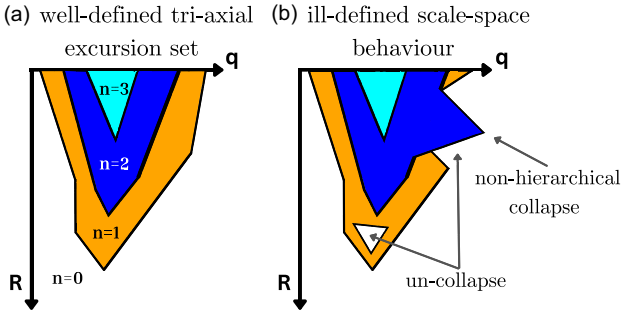

Figure 2. An illustration of an evolving Lagrangian field in the excursion set framework. (a) A well-defined case, where if one was to draw a vertical line starting from any point $\boldsymbol{q}$, the n_R would decrease hierarchically with increasing R. The colour coding associated with the structure rank is white for $n_R = 0$ *(void)*, orange for $n_R = 1$ *(pancake)*, blue for $n_R = 2$ *(filament)*, and cyan for $n_R = 3$ *(halo)*. (b) An ill-defined example, where n_R no longer decreases in a stepwise manner (e.g. a fluid particle to the right of the plot un-collapses from a pancake to a void and then collapses straight into a filament).

and is a direct consequence of equation (8) if $c_R = n_R$ is defined through a perfect collapse model. In particular, it implies that no volume elements *'un-collapse'* when smoothing scales are decreased. For example, let us consider a Lagrangian volume element that collapses in a simulation with some large smoothing scale R_1 into a filament ($n_{R_1} = 2$). In that case, it should also collapse into a filament (or a halo) in any simulation with a smaller smoothing scale $R_2 < R_1$. However, it should not become part of a lower rank structure ($n_{R_2} < 2$) like a pancake or void.

To illustrate this concept, consider Fig. 2, where we plot the evolution of a Lagrangian field $\boldsymbol{q}$ with increasing smoothing scale R. In the left panel we present a well-defined excursion set example employing a perfect tri-axial collapse model, where n_R consistently *increases* with *decreasing* R. The tri-axial nature is represented by the rank of each element, colour-coded to indicate its morphology type: white for $n_R = 0$ *(void)*, orange for $n_R = 1$ *(pancake)*, blue for $n_R = 2$ *(filament)*, and cyan for $n_R = 3$ *(halo)*. We will maintain this colour scheme to represent the morphology classes of fluid elements in all subsequent figures depicting our simulations. In the right panel, we demonstrate a contrasting example where the behaviour in scale-space violates the excursion set assumption, as some regions of $n_R = 1$ revert to $n_R = 0$. This panel also highlights instances of non-hierarchical collapse behaviour, where elements transition directly from $n_R = 0$ to $n_R = 2$ (see Section 2.2) as R decreases.

Further, in Fig. 3 we show the final snapshot ($z = 0$) of the same slice through the evolved density field within our simulation box. Each slice has evolved from different initial conditions, with the sharp k-space cut-off $\widetilde{W}_{\rm sk}$ applied to the $P(k)$ (corresponding to σ_2^2, σ_6^2, σ_{10}^2, and σ_{14}^2 from the top-left to the bottom-right), as described in Section 3.4. These snapshots depict the final distributions of simulation particles at their current coordinates $\boldsymbol{x}$ in the Eulerian frame.

In general, decreasing the smoothing scale primarily introduces additional small-scale structure while preserving structure that has already existed on larger scales. Moreover, structures appear to move up in their morphology rank when the smoothing scale is decreased. For example, filaments may additionally fragment into smaller scale haloes embedded in the larger filament. Pancakes, on the other hand, may fracture into smaller filaments and haloes.

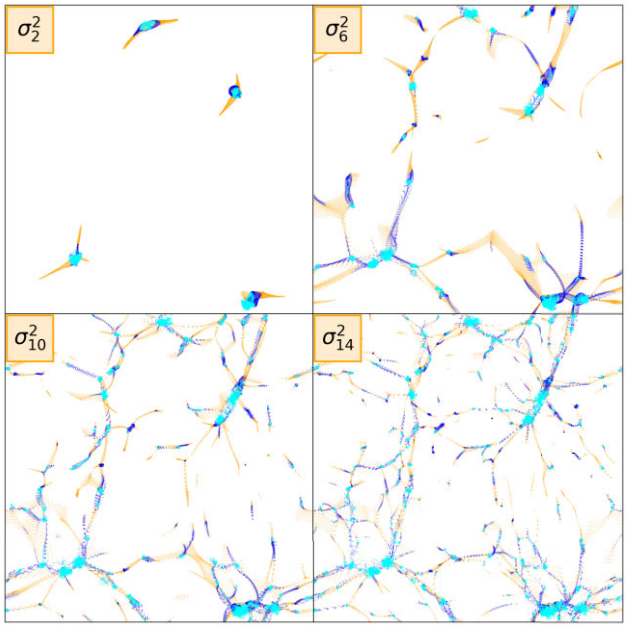

Figure 3. Thin slices through the DM sheet in Eulerian space at $z = 0$ after evolving from four different initial conditions, with σ_2^2, σ_6^2, σ_{10}^2, and σ_{14}^2, smoothed with the $\widetilde{W}_{\rm sk}$ filter. As σ^2 increases (more small-scale k modes included in the initial $P(k)$), more collapsed structures emerge on smaller scales – haloes (cyan), filaments (blue), pancakes (orange), and voids (white).

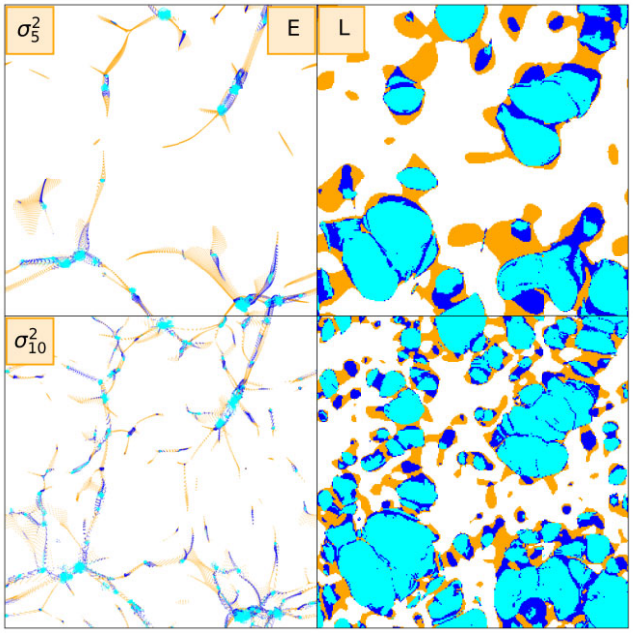

Figure 4. Two slices of $\widetilde{W}_{\rm sk}(kR)$ smoothed simulations, evolved from initial conditions with σ_5^2 (top) and σ_{10}^2 (middle). Left: the fluid element distribution in Eulerian space. Right: the same distribution, but in Lagrangian space. The colour scheme of the top four panels is the same as in Fig. 3.

Fig. 3 suggests that the excursion set assumption holds well on a qualitative level. We will further evaluate this quantitatively in the following subsection.

4.2 Testing the excursion set assumption

Although Fig. 3 demonstrates that structures typically tend to move upwards in their morphology rank when adding small-scale perturbations in Eulerian space, we need to use a Lagrangian perspective to test equation (20) quantitatively.

For illustration, consider Fig. 4, where we present the morphology classification in the final snapshots of two different $\widetilde{W}_{\rm sk}$ smoothed simulations, σ_5^2 (top) and σ_{10}^2 (bottom). The panels on the left show the simulation points plotted in *Eulerian* space (indicated with an 'E'), while in the panels on the right, the points are plotted in *Lagrangian* space (indicated with an 'L'), that is, at their locations in the initial conditions. The colours indicate morphology classification, as outlined in Section 4.1. Note that haloes take up the least volume in Eulerian space and voids the most. However, in Lagrangian space, haloes take up the most volume because they are the densest and most massive structures with the most particles. We again note that most structures that exist in Lagrangian space at σ_5^2 also exist in the σ_{10}^2 simulation.

Fig. 5 shows the σ_{10}^2 distribution of particles in Lagrangian space only, but this time the colouring reflects whether that simulation exhibits additional collapse with respect to the σ_5^2 simulation: grey for *no difference* ($n_{10} = n_5$), black for *collapse* ($n_{10} > n_5$), and red for *un-collapse* ($n_{10} < n_5$). Thus, the grey and black regions are consistent with the excursion set assumption, while the red regions are not expected. In the leftmost panel, corresponding to $\widetilde{W}_{\rm sk}$, we note that most of the volume is black or grey, showing that the assumption is overall quite reasonably satisfied. We observe similar effects in the $\widetilde{W}_{\rm th}$ (middle) and $\widetilde{W}_{\rm g}$ (right) smoothed simulations, but with considerably smaller un-collapsed regions. However, we also note that a small fraction of Lagrangian space seems to un-collapse due to the perturbations from smaller scales – a phenomenon that, by construction, cannot be captured in the excursion set paradigm.

A closer examination of Fig. 5 reveals that the un-collapsing regions are pre-dominantly located near the boundaries of haloes in the σ_5^2 simulation. That these regions are the most affected by smaller scale perturbations is understandable since these particles are just at the threshold of reaching their halo by $z = 0$, and a small difference in the initial conditions may push them over the edge. Further, the coherence scale of the difference field seems more closely related to the smoothing scale of the σ_5^2 than of the σ_{10}^2 simulation. We may therefore speculate that most of the un-collapsing happens due to the perturbations that lie just below the cut-off scale of the σ_5^2 simulation. This can be seen in more detail in Fig. C1 in the appendix, where the same plot is shown for additional examples.

For a quantitative assessment, we analyse the complete set of confusion matrices corresponding to $\widetilde{W}_{\rm sk}$, $\widetilde{W}_{\rm th}$, and $\widetilde{W}_{\rm g}$ from top to bottom in Fig. 6. In each confusion matrix, every row corresponds to all particles associated with a specific morphology in the σ_i^2 simulation (with $n_i = 3$, 2, 1, or 0 for the respective column panels). Each pixel in the same row shows which fraction of these particles is also collapsed to a compatible structure in the σ_j^2 simulation. That means $n_j = 3$ for haloes, $n_j \geq 2$ for filaments, $n_j \geq 1$ for pancakes, and for voids we compare to $n_j = 0$.

Elements left of the diagonal ($\sigma_j^2 < \sigma_i^2$) show the progression of the excursion set: the more small-scale perturbations are added, the more likely it is that a point collapses. The closer σ_j^2 is to σ_i^2, the more likely it is that particles have collapsed to the same type of structure they were selected by at σ_i^2, until the fraction corresponds to 1 at the diagonal by definition.

Elements right of the diagonal ($\sigma_j^2 > \sigma_i^2$), show the fraction of particles that have the same (or a consistent) morphology also at higher variance. For example, in the second column panel, we see

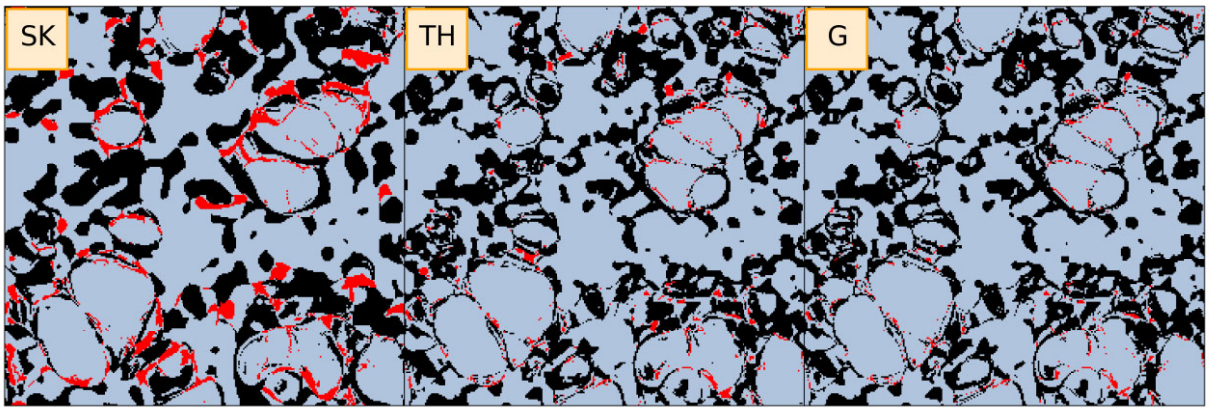

Figure 5. The difference of the two upper simulations $\sigma^2_{10} - \sigma^2_5$, but for all three window functions, $\widetilde{W}_{\rm sk}$, $\widetilde{W}_{\rm th}$, and $\widetilde{W}_{\rm g}$, from left to right. The particle distribution is the same as in σ^2_{10}. Particles belonging to the same morphology class as in σ^2_5 are shaded grey, those which collapsed along more axes $n_{10} > n_5$ are marked in black, and the cases where $n_{10} < n_5$, i.e. un-collapsed particles are coloured red.

particles that are in a filament at $\sigma^2_j = 5$ and are also in a filament (or halo) at higher variances. According to the excursion set assumption, all of the elements to the right of the diagonal should be at 100 per cent for all filters. However, in the simulations smoothed with $\widetilde{W}_{\rm sk}$, we observe that they are approximately 90 per cent instead – indicating that around 10 per cent of the mass un-collapses from filaments to lower rank structures. This deviation from 100 per cent gradually improves as we look to the $\widetilde{W}_{\rm th}$ and $\widetilde{W}_{\rm g}$ panels, where the un-collapsed fraction levels at ~ 4 per cent and ~ 2 per cent for the two filters, respectively.

For the halo, filament, and pancake cases, the deviation of the elements at the top right of the diagonal from unity indicates the degree of the excursion set assumption violation. A key insight from Fig. 6 regarding the $\widetilde{W}_{\rm sk}$ smoothed simulations is that the major un-collapsing events occur when $\Delta\sigma^2 \leq 1$, corresponding to the confusion matrix squares directly adjacent to the diagonal, where the differences range from 5 per cent to 13 per cent.

Beyond this point the fractions remain relatively constant, indicating that perturbations on much smaller length-scales indeed do not impact the collapse of larger scale regions. However, perturbations on slightly smaller length-scales do affect their collapse. For the $\widetilde{W}_{\rm th}$ and $\widetilde{W}_{\rm g}$ smoothed simulations, the un-collapsed fractions with the successive $\Delta\sigma^2 = 1$ steps are visibly more uniform, with no significant un-collapsing events occurring at any particular $\Delta\sigma^2$. In both cases the un-collapsing is rather broadly distributed across the entire $\Delta\sigma^2$ range, with deviations from the diagonal reaching 1 per cent to 4 per cent for the top-hat and 1 per cent to 3 per cent for the Gaussian filters, with no particular pattern.

For voids, the diagram needs to be read in the reverse order. The excursion set assumption implies that everything that is part of a void at the scale σ^2_i must also be void at all larger scales ($\sigma^2_j < \sigma^2_i$). Therefore, all pixels to the left of the diagonal are expected to be at 100 per cent. As with the other column panels, we can see that the fractions are generally of order 90 per cent and that most of the difference comes from scales that are very close to the diagonal for the $\widetilde{W}_{\rm sk}$ smoothed simulations. In contrast, for $\widetilde{W}_{\rm th}$ and $\widetilde{W}_{\rm g}$ cases, this difference is at most 2 per cent and 1 per cent, respectively.

4.3 The simulation excursion set

Since the simulation represents a perfect collapse model, we may ask what happens if we follow through with the excursion set procedure, but use the simulated morphology classification. Using the simulation classification as a collapse model in equation (8), we get the simulation excursion classification:

$$n_{R,\rm ES} = \sup\{n_{R'} \mid R' \geq R\} \tag{21}$$

Under the excursion set assumption, the relationship should hold as $n_{R,\rm ES} = n_R$.

Figs 7–9 demonstrate the result of emulating this excursion set with simulations smoothed with all three windows in the two top panels of each figure, again with the Eulerian and Lagrangian slices left and right, respectively. The middle rows show the simulations with the smallest fixed scale smoothing σ^2_{16} and the bottom panels show the difference between the excursion set and σ^2_{16}, with the following colour assignment: $n_{\rm ES} - n_{16} = 0$ (white), $n_{\rm ES} - n_{16} = 1, 2, 3$ (black, blue, red), with ES denoting excursion set. As anticipated from the previous results, the observation of un-collapsing throughout the 16 runs indicates that the excursion set of the $\widetilde{W}_{\rm sk}$ smoothed simulations contains many more collapsed structures. While voids largely remain intact, haloes, filaments, and pancakes are notably more spread out and smoother in the excursion set plot shown in Fig. 7. Overall, the excursion set generated from these simulations leads to un-collapse of 20 per cent of *all* simulation particles, irrespective of morphology, relative to the σ^2_{16} realization. In this example, we also observe that particles identified as haloes in the excursion set are the least affected by changes to the ICs, with 16 per cent undergoing un-collapse with respect to σ^2_{16}. In contrast, as many as 44 per cent of pancakes and 48 per cent of filaments in the excursion set experience un-collapse compared to σ^2_{16}. Similarly, in line with the findings from Fig. 6, we observe a decrease of the overall un-collapsed regions when comparing the excursion sets generated from the top-hat (Fig. 8) and the Gaussian (Fig. 9) smoothed simulations with their respective σ^2_{16}. While the total fraction of particles undergoing un-collapse between the top-hat excursion set and the fixed scale simulation is 5 per cent, this value is even lower in the Gaussian window example, amounting to just 3.5 per cent. When focusing only on the fraction of particles classified as haloes, filaments, or pancakes in the excursion sets that un-collapse in σ^2_{16}, we observe differences of 5 per cent, 10 per cent, and 6 per cent for the $\widetilde{W}_{\rm th}$ case and 4 per cent, 6 per cent, and 4 per cent for the $\widetilde{W}_{\rm g}$ case, respectively. As in the $\widetilde{W}_{\rm sk}$ example, haloes are the least affected by changes in the ICs, while filaments are the most impacted. In both instances, the smoothing of the collapsed

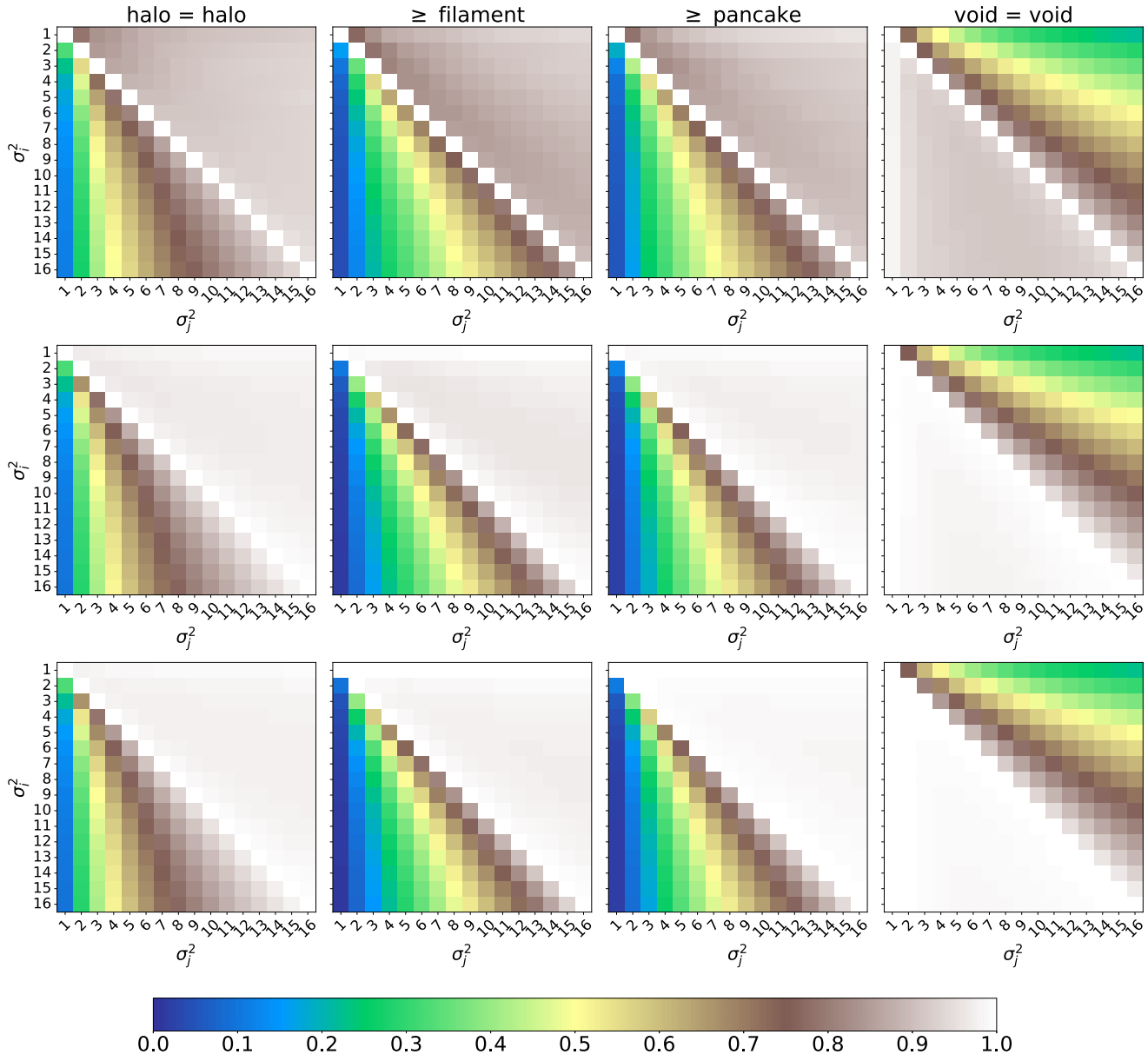

Figure 6. Four confusion matrices per filter: $\widetilde{W}_{\rm sk}$ (top), $\widetilde{W}_{\rm th}$ (middle), and $\widetilde{W}_{\rm g}$ (bottom), each showing the total of particles from σ_i^2 run, which are assigned to the same morphology, or to one of higher rank, in run σ_j^2. I.e. the '≥ pancake' confusion matrix quantifies the fractions of pancake particles from σ_i^2 that are either pancakes, filaments, or haloes in σ_j^2. Moving to the right of the main diagonal, the smoothing scale of the comparison simulation decreases, while moving to the left, the scale increases. The excursion set theory assumption implies that everything to the *right* of the main diagonal, in the three plots from the left, should be white. Conversely, everything to the *left* of the main diagonal in the 'void = void' confusion matrix should also be white.

structures in the excursion sets is hardly visible in the figures and is predominantly observed around haloes, while in the bottom panel of Fig. 7, it is evident that the filament and pancake regions are also significantly affected by un-collapse. These values are summarized in Table 1.

We conclude that the excursion set assumption does draw an appropriate picture on a qualitative level and that it also works reasonably well with the $\widetilde{W}_{\rm sk}$ smoothing and surprisingly well with $\widetilde{W}_{\rm th}$ and $\widetilde{W}_{\rm g}$ smoothing on a quantitative level. In this section, we demonstrated that the core assumptions of the excursion set hold extremely well in practice with a Gaussian window, slightly worse with a top-hat and significantly worse with a sharp $k-$space window functions. That is, if we combine a perfect collapse model (a simulation) with a sharp k-space filtering method, then we find quantitative deviations from the core assumption on the particle-by-particle basis of order 20 per cent. This implies that an excursion set formalism using $\widetilde{W}_{\rm sk}$ for smoothing, along with a maximally physical collapse model, *should* yield a 20 per cent error.

4.4 Morphology mass fractions

So far, we have focused on the particle-by-particle contrast of the excursion sets generated with different window functions and the fixed scale σ_{16}^2 simulations. As already discussed in Section 1, the excursion set formalism has many other useful applications, such as predicting HMFs, which offer valuable insights into the cosmic

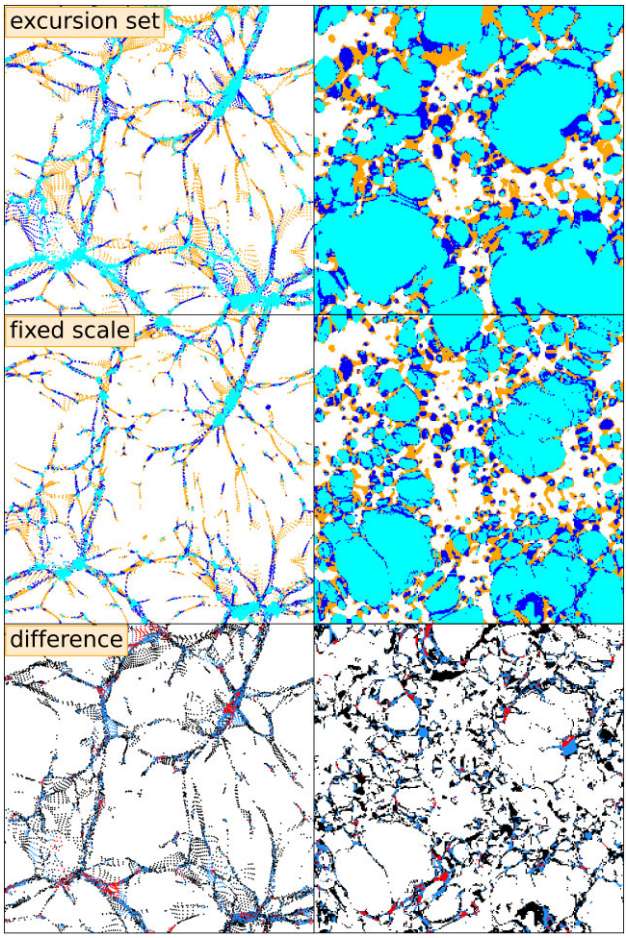

Figure 7. The excursion set constructed from the 16 simulations smoothed with $\widetilde{W}_{\rm sk}$ (top), the σ^2_{16} output (middle), and their difference (bottom). The colour scheme of the difference plot is as follows: white indicates no difference, while black, blue, and red correspond to differences of 1, 2, and 3, respectively, in the morphology rank n_R. The simulation particles are plotted in Eulerian space (left panels) and Lagrangian space (right panels).

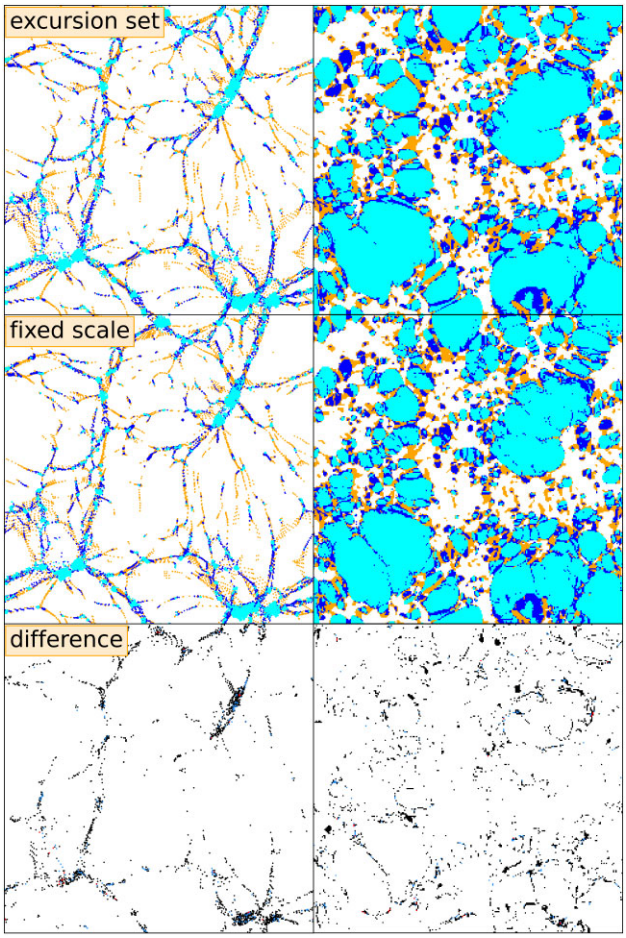

Figure 8. Same as Fig. 7, but for simulation ICs smoothed with $\widetilde{W}_{\rm th}$.

mass distribution within haloes at all scales relevant to the chosen cosmology. This methodology can also be extended to explore the mass distribution in other structures, such as pancakes and filaments. However, gaining a more comprehensive understanding requires data that are not easily accessible through approaches relying on standard structure finders. Using our simulations, we can not only identify the morphologies of the structures into which the simulation particles collapse, but also obtain a general understanding of the scales at which these structures form when combined with the excursion set framework.

We calculate the cumulative mass fraction $F(<\sigma^2)$ or $F(>R)$ for each morphology class. We do this by identifying the scale of the first (and largest) collapse in the excursion set, as defined by equation (9) and determining the fraction of particles with a collapse scale larger than the considered scale. Recount that in our way of counting here, every particle embedded in a filament will also be considered part of a pancake, and so on.

The top panel of Fig. 10 shows the cumulative distributions of particle fractions associated with pancakes (stars), filaments (hexagons), and haloes (diamonds) based on the excursion set of simulations obtained with the $\widetilde{W}_{\rm sk}$ smoothing.[1] In the plot, the upper x-axis indicates the Lagrangian length smoothing scales, while the lower x-axis shows the corresponding values of σ^2. We observe that typical particles appear to be part of a large-scale pancake, a smaller scale filament, and a notably smaller halo. For example, the median particle (where $F = 0.5$) is part of an $R \sim 1.8$ Mpc h^{-1} pancake, an $R \sim 0.8$ Mpc h^{-1} filament, and an $R \sim 0.5$ Mpc h^{-1} halo. In Appendix B we also show the same plot, but for the example of the $\widetilde{W}_{\rm th}$ smoothed simulations.

At the smallest considered smoothing scale ($R \sim 0.14\,{\rm Mpc}\,h^{-1}$) – which is already considerably smaller than what most cosmological simulations resolve – the morphology fractions for pancakes, filaments, and haloes are 88 per cent, 78 per cent, and 68 per cent, respectively. Since the mass fractions are growing moderately in this regime, we would expect that in most cosmological simulations of order 80–90 per cent of particles are embedded in a pancake (or higher), 70–80 per cent in a filament (or halo), and 60–70 per cent in a halo. The exact numbers may naturally vary based on the specifics of the classification (see Knebe et al. 2011, 2013; Angulo & Hahn 2022, for a comprehensive review). Additionally, we anticipate these values to rise considerably if simulations were extended to the physical cut-off in the power spectrum, such as in a neutralino cosmology (Stücker et al. 2018; Wang et al. 2020).

[1] We have checked that other kernels lead to similar results, but for simplicity we focus only on the sharp$-k$ case here.

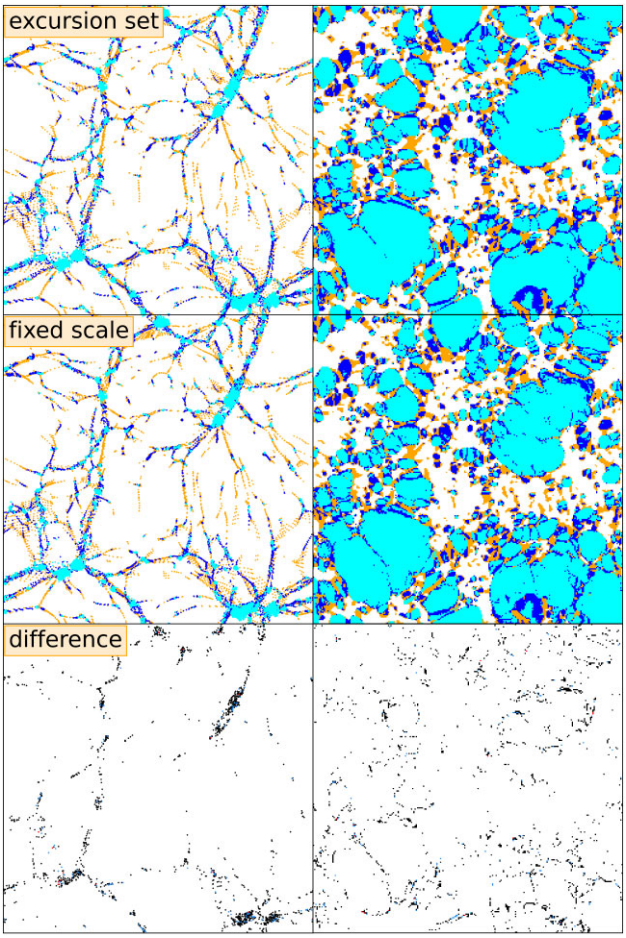

Figure 9. Same as Fig. 7, but for simulation ICs smoothed with $\widetilde{W}_{\rm g}$.

Table 1. The un-collapsed fractions of particles between the three excursion sets and the reference σ^2_{16} of the three suites of smoothed simulations. U indicates the number of un-collapsed particles of the given type X. The fractions are taken with respect to the number N of X particles collapsed in the excursion set and $N_{\rm Total} = 256^3$.

	$\widetilde{W}_{\rm sk}$	$\widetilde{W}_{\rm th}$	$\widetilde{W}_{\rm g}$
X		$U_{\rm X}/N_{\rm X}$	
Halo	0.16	0.05	0.04
Filament	0.48	0.1	0.06
Pancake	0.44	0.06	0.04
Total	0.2	0.05	0.035

We can compare our results to existing predictions from excursion sets. In particular, the mass fractions of pancakes and filaments were previously explored by Shen et al. (2006). In this work, excursion sets in density with three separate barriers for pancakes, filaments, and haloes (based on approximations of the ellipsoidal collapse model) were used to explore the formation of such structures. They found that, in a ΛCDM Universe, at $z = 0$, above 99 per cent of all mass is contained in pancakes, 72 per cent in filaments, and 46 per cent in haloes, with $M > 10^{10}\,{\rm M}_\odot$. Since their resolved mass is exactly in the range of the sharp$-k$ mass at our smallest resolved scale $\sim 2 \times 10^{10}\,{\rm M}_\odot$, we can compare these numbers approximately to ours. Clearly, the model of Shen et al. (2006) strongly overpredicts the mass in pancakes. This is also confirmed by a variety of other

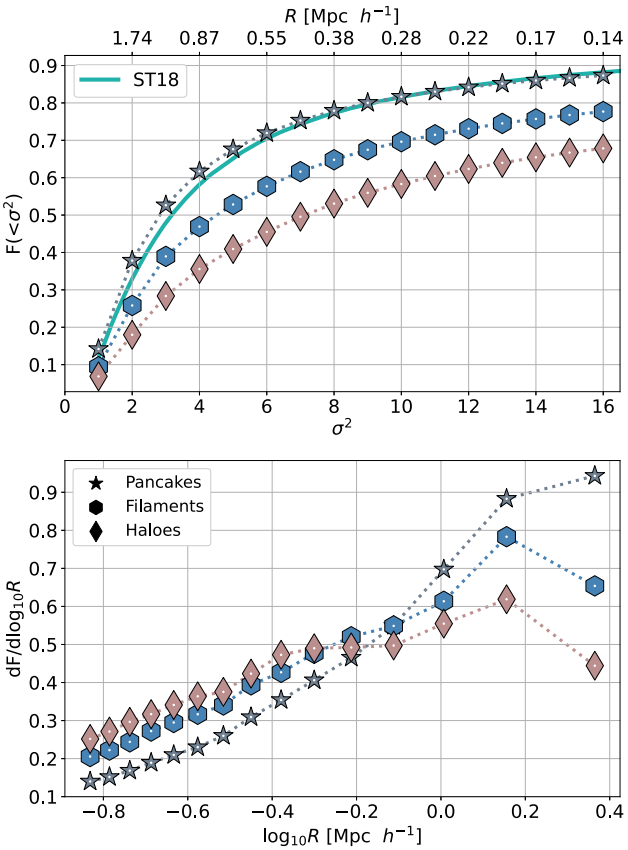

Figure 10. Top: the mass fractions of particles in pancakes (stars), filaments (hexagons), and haloes (diamonds) as a function of the smoothing scale σ^2 and R. The turquoise curve represents $F(<\sigma^2)$ of collapsed particles from Stücker et al. (2018). Bottom: the particle fraction in structures of the excursion set radii. Both panels show the distributions for the $\widetilde{W}_{\rm sk}$ smoothed excursion set of simulations.

studies based on cosmic web classification schemes (Buehlmann & Hahn 2019), which tend to find numbers between 11 per cent and 32.5 per cent for the mass in structures with $n \geq 1$.

In the top panel of Fig. 10, we also plot $F(<\sigma^2)$ adapted from fig. 9 of Stücker et al. (2018) (turquoise). Their triaxial collapse model has been designed to describe the first collapse of voids towards pancakes (or single-stream regions to multistream regions). The authors used random walks of the deformation tensor to determine statistics of single-stream regions ($n = 0$). The fraction of mass in pancakes – given by the remaining collapsed material – appears to be very well predicted by this approach. Consequently, it can be concluded that, at the very least, an accurate description of the full hierarchy of three-dimensional collapse requires accounting for the full deformation tensor and is challenging to achieve using effective density barriers, as is done for haloes (Sheth et al. 2001).

The bottom panel of Fig. 10 shows the differential mass fractions, represented by the first crossing distribution $\mathrm{d}F/\mathrm{d}\log_{10} R$, as a function of scale, while Fig. 11 (and the supplementary Fig. B2) shows a $\sigma-$scaled first crossing distribution $\mathrm{d}F/\mathrm{dln}\sigma = f(\sigma)$, as a function of σ for all smoothing kernels: $\widetilde{W}_{\rm sk}$ (blue), $\widetilde{W}_{\rm th}$ (yellow), and $\widetilde{W}_{\rm g}$ (purple). The plots reveal that large-scale pancakes appear more significant compared to haloes, whereas small-scale pancakes appear less dominant. We may conclude that most of the smaller scale web patterns of the Universe will likely be accreted into larger

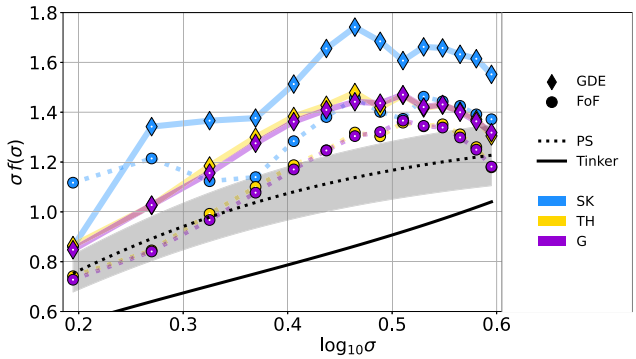

Figure 11. The multiplicity functions $f(\sigma)$ obtained from the sharp$-k$ (SK), top hat (TH), and Gaussian (G) smoothed simulations, accompanied by functions of Press & Schechter (1974) (dotted) and Tinker et al. (2008) (continuous) in black, where the region within 10 per cent of the Press & Schechter curve is shaded grey. The diamonds represent the GDE halo particles (see Sections 3.2 and 3.3), whereas the circles represent the FoF haloes.

scale pancakes. Further, in Fig. 11 we also plot the theoretical first crossing distribution $f(\sigma)$ of Press & Schechter (1974) (black dotted) and the fitting function of Tinker et al. (2008) (black continuous). The region within 10 per cent of the Press & Schechter prediction is shaded grey. We also plot two sets of halo $f(\sigma)$s obtained by applying the excursion set formalism to our simulations. In method (1) halo particles were identified within the simulation via the GDE scheme (see Sections 3.2 and 3.3); we plot these as diamonds. In method (2) particles comprising haloes were identified via the friends-of-friends (FoF) algorithm, with linking length 0.16, of GADGET4 (Springel et al. 2021); these we plot as circles. For (2) the same excursion set assumption was applied – a particle is assigned the largest halo mass of all σ_i^2. Notice how the (1) and (2) functions differ significantly form one another, with FoF haloes always containing smaller particle fractions per σ, as compared to the results obtained with the GDE method. The contrast between the FoF and GDE $f(\sigma)$s of the top-hat and Gaussian smoothed simulations is notably smaller than that of the sharp$-k$ example. In general, more particles fulfil the condition to count as belonging to a halo, according to the halo finder in the sharp$-k$ smoothed simulations – a consequence of the extent of the $\widetilde{W}_{\rm sk}$ in real space. However, the effect of such smoothing appears to impact the GDE morphology assignment algorithm more than $\widetilde{W}_{\rm th}$ and $\widetilde{W}_{\rm g}$ do. Fig. 11 demonstrates that the GDE method is more permissive in classifying particles as collapsed into structures, leading to higher collapsed fractions. In line with our previous results, the Gaussian and top-hat points are in a very tight agreement, while the sharp$-k$ simulations result in a higher fraction of mass collapsed into all structures on all scales, which is reflected in Fig. 7. Note how the $f(\sigma)$ obtained with a 'perfect' collapse model, which is also clearly of triaxial nature, lie significantly closer to the simple analytical result of spherical collapse model, than it does to the empirically obtained Tinker et al. (2008) function. It is a surprising result, given that the latter is motivated by triaxial (ellipsoidal) collapse, even if not directly derived from it. At the same time, all three excursion sets of simulations result in higher numbers of collapsed haloes than the two predictions suggest. For completeness we also plot the filament and pancake multiplicity functions in Appendix B, Fig. B2.

4.5 Collapse scale to mass relation

So far, we have tested the general assumptions of excursion set formalisms, particularly in terms of how the hierarchy of structure formation progresses across scale-space. However, excursion sets are very commonly used to predict the mass functions of haloes, which requires an additional assumption. It is typically assumed that when a particle first satisfies a collapse criterion at a smoothing scale R_s, it becomes part of a halo with mass $M \propto R_s^3$, as outlined in equation (11). This mapping between collapse scales and the assumed halo mass is a crucial ingredient that we can examine with our excursion set of simulations.

Consider Fig. 12, where we label each particle in Lagrangian space by the predicted halo mass, based on the default mass mapping (11), for the excursion sets obtained from the simulations that have been smoothed with the different filter functions. Additionally, in the rightmost panel, we also show each particle labelled by the mass $\rm M_{FoF}$ of the halo it has been assigned to in a CDM (no smoothing) simulation[2] with an FoF halo finder with linking length 0.16.

We note that our collapse criterion is more permissive than the traditional FoF definition. For FoF haloes, it is possible for particles that have entered the group to leave it again, e.g. when they reach their first orbital apocentre at the splash-back radius. In contrast, in the GDE framework we consider any particles that have ever gone through any 3D crossing, as permanently associated with a halo – naturally also including particles at their splash-back radii. The effect of this difference can be seen in Fig. 11, whereby the multiplicity functions of FoF haloes are systematically below $f(\sigma)$ of the GDE halo particles, which implies more mass collapsed into haloes per smoothing scale in the latter case. Further, we point out that the excursion sets may label particles that are part of the same structure with different masses, because some particles have only become part of the same structure at smaller smoothing scales. It is clear that the true collapse scale-to-mass relation is not a simple one-to-one relation, and is likely to be significantly more complex. In Fig. 12, the excursion set masses correlate reasonably well with the actual halo masses, but deviations are quite large in amplitude.

To evaluate this more quantitatively, in Fig. 13 we show the relation between the excursion set collapse scale (indicated through mass, as in equation 11) versus the measured masses of haloes of the same particles in a CDM simulation. We have to restrict this comparison to particles that are both part of an FoF halo in the simulation and have collapsed according to the simulation excursion set. We show this in two ways: (1) the greyscale, background hex bin distribution of particle counts, with finer two-dimensional binning and (2) blue violin distributions of six coarsely binned counts along the x-axis. The orange circles indicate the medians of particle counts in the x-axis bins, and the diagonal white-dashed lines show the benchmark $\rm M_{FoF}$ masses. Notice that the excursion set mass distribution is not continuous (we have exactly 16 masses corresponding to our sets of smoothed simulations), and due to our choice of y-axis binning, there are horizontal gaps in the hexbin distribution. Likewise, due to our small simulation volume ($L = 40$ Mpc), there is a lack of particle counts in high-mass haloes ($> 10^{13}\,\rm M_\odot h^{-1}$) in the simulations, giving rise to horizontal data gaps.

We find two key insights from this figure: (1) the median relation tends to have an offset from the naive expectation from equation (11). For the considered window functions the offset is

[2] All the parameters in this simulation were set to the same values as described in Section 3.4.

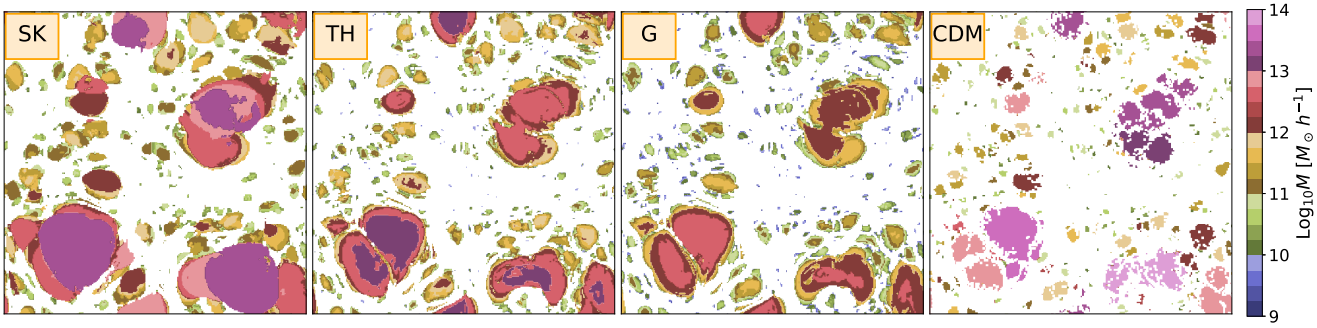

Figure 12. Lagrangian plots of excursion set mass assignment to simulation particles collapsed into haloes, based on the scale of collapse σ_i^2 according to the excursion set assumption. The first three panels show the results from the suite of simulations ran with initial $P(k)$ smoothed with the $\widetilde{W}_{\rm sk}$, $\widetilde{W}_{\rm th}$, and $\widetilde{W}_{\rm g}$, from left to right. The rightmost panel shows the FoF haloes from a CDM (no smoothing) simulation.

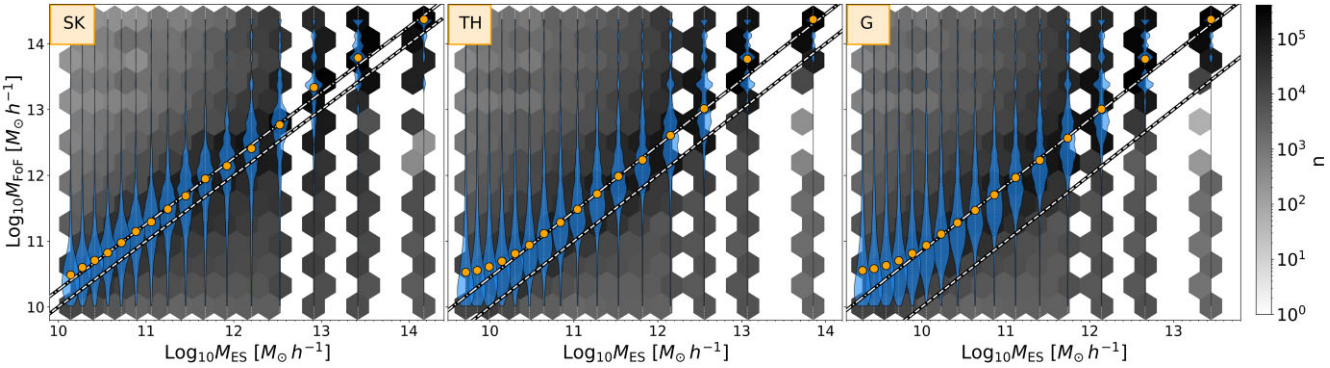

Figure 13. The background hexagonal plot shows the particle count n in the 2D bins, while the violin plots (blue) show the count distribution along the y-axis in the 16 bins corresponding to each $M_{\rm ES}$. Medians of the $M_{\rm FoF}$ distributions w.r.t. to the $M_{\rm ES}$ (orange points). The benchmark masses (white dashed) and the diagonal shifted along the y-axis by a factor α to match the medians (white dash–dotted).

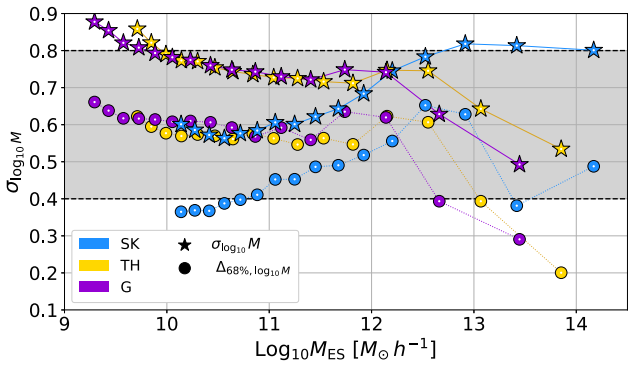

Figure 14. Two estimators of the logarithmic scatter in the mass-map. The stars indicate the standard deviation, and the circles are a percentile-based estimate. While the precise amplitude depends on the details of the filter and the estimator, the overall amplitude ranges between 0.4 and 0.8 dex (shaded area).

$\alpha = 0.26, 0.45, 0.85$ for $\widetilde{W}_{\rm sk}$, $\widetilde{W}_{\rm th}$, and $\widetilde{W}_{\rm g}$ respectively. We suggest that the pre-factor of the mass map should generally be considered as a degree of freedom that may be fitted in excursion set models. This may account for the uncertainty associated with the halo definition and the effective scale of the smoothing function. (2) The collapse scale-to-mass relation is much more complicated than commonly assumed. Particles that collapse at a singular smoothing scale often end up in haloes of different masses (see also White 1994; Sheth et al. 2001; Monaco et al. 2002a; Achitouv et al. 2013; Hahn & Paranjape 2014). This will make it challenging to reliably predict masses, even when adopting a 'perfect' collapse model.

We evaluate this scatter quantitatively in Fig. 14. Here, we show two estimates of the scatter amplitude of $M_{\rm FoF}$ in fixed bins of $M_{\rm ES}$. The first estimate is the standard deviation in log-space

$$\sigma_{\log_{10} M} = \sqrt{\langle (\log_{10} M - \langle \log_{10} M \rangle)^2 \rangle} \tag{22}$$

and the second is half the extent of the 68 per cent interval

$$\Delta_{68\ \text{per cent},\log_{10} M} = \frac{1}{2}(\log_{10} M_{84\ \text{per cent}} - \log_{10} M_{16\ \text{per cent}}), \tag{23}$$

where $M_{84\ \text{per cent}}$ and $M_{16\ \text{per cent}}$ are the 84th and 16th percentiles, respectively. If $M_{\rm FoF}$ would follow a lognormal distribution at fixed $M_{\rm ES}$, then the two estimators should lead to identical results.

We find that the estimator $\Delta_{68\ \text{per cent},\log_{10} M}$ generally gives smaller values than $\sigma_{\log_{10} M}$ by about 0.2 dex, showing that the error distribution is not perfectly lognormal. Overall, amplitudes range between 0.4 and 0.8 dex, and the scatter seems to be relatively constant with mass, generally not changing by more than 0.2 dex in the range below $M \lesssim 10^{13}{\rm M}_\odot h^{-1}$. The scatter can be even a bit smaller at higher masses, but the measurement is likely unreliable here due to finite-size effects and a very limited sample. Interestingly, the $\widetilde{W}_{\rm sk}$ leads to a slightly smaller scatter in the mass map than the other filters at masses $M \sim 10^{11}{\rm M}_\odot h^{-1}$.

Independently of these details, we may summarize that excursion sets with a 'perfect' collapse model still have at least an uncertainty of order 0.4–0.8 dex in the mass map. Lucie-Smith et al. (2024) reported

a similar mass scatter within this range (for 10^{11}–$10^{13.4}\,\mathrm{M}_\odot\,h^{-1}$ haloes), in their halo collapse model based on convolutional neural networks (see also López-Cano et al. 2024). Improvements in the collapse model cannot reduce this uncertainty; instead, it is rather due to the limitations imposed by the spherical kernel that is used to traverse scale-space (see also Hahn & Paranjape 2014; Lucie-Smith et al. 2024; Verza et al. 2024).

5 DISCUSSION AND CONCLUSION

Excursion sets are an invaluable tool to predict and understand the formation of structures. While they are primarily utilized to predict statistical quantities, such as HMFs, they may also be used to infer detailed particle-by-particle predictions, such as those related to accretion and merger histories, as well as the formation of all types of structures like voids, pancakes, filaments, and haloes. Such predictions are crucial for understanding structure formation beyond scales that can be resolved reliably in simulations. However, to make such predictions, excursion sets must be tested thoroughly at scales that *can* be resolved in simulations.

Historically, excursion sets have been primarily evaluated based on their ability to reproduce the simulated HMF. However, in the era of precision cosmology, they have taken a backseat role to simulation-based fits (e.g. Jenkins et al. 2001; Sheth et al. 2001; Reed et al. 2003, 2006; Warren et al. 2006; Tinker et al. 2008; Despali et al. 2016; Verza et al. 2024) and HMF emulators (e.g. Monaco et al. 2002a, b; Bocquet et al. 2020). These simulation-derived HMFs typically achieve an accuracy better than 10 per cent for all common halo mass definitions and can include possible non-universal behaviour of the HMF. It must be stressed that, by definition, these mass functions are limited only to the regime where simulations can be performed and, as such, have little to no predictive capability beyond their calibrated range, consequently yielding no analytical insights. Predictions of unresolved quantities – such as the abundance of very low-mass haloes – thus still rests on excursion set models. If their accuracy were improved, these models could yield very economic alternatives to the large simulations required to calibrate high-accuracy mass functions. Such development would facilitate quick exploration of alternative cosmological models.

The aspect of modelling, which has typically been optimized and improved upon, is the model describing the collapse of Lagrangian volume elements. Examples include spherical collapse (Press & Schechter 1974; White & Silk 1979; Bond et al. 1991), 3D ellipsoidal collapse (Bond & Myers 1996; Sandvik et al. 2007), effective density-based ellipsoidal collapse (Sheth et al. 2001, 2006), triaxial collapse (Monaco 1995; Monaco et al. 2002a, b; Stücker et al. 2018), energy-based spherical collapse (Musso & Sheth 2021), and virial-equation-based collapse (Musso et al. 2024).

In this paper, we have shown that excursion sets can also be created using a 'perfect' collapse model, namely, a simulation. This requires running several simulations with initial conditions smoothed on successively reduced scales. Collapse may be detected using conventional methods like an FoF classification or more advanced methods that distinguish between different structure morphologies, as we have adopted here.

This 'excursion set of simulations' has allowed us to (1) test assumptions of the excursion set models independently of the adopted collapse model, (2) define an upper limit to the possible realism of excursion set models, and (3) identify bottlenecks in the current paradigm.

In this paper, we have tested two core assumptions of excursion sets. The first assumption states that collapse does not revert when decreasing Lagrangian smoothing scales. We find that this assumption is overall satisfied surprisingly well and the quantitative accuracy depends significantly on the applied window function. For the sharp $k-$space filter, we find that up to 20 per cent of fluid elements uncollapse across the resolved scales, whereas for the Gaussian and top-hat windows, this percentage is only of order 3.5 per cent and 5 per cent. This degree of accuracy may be a reassurance for any models attempting to disentangle structure formation in scale-space. However, we also note that the best possible excursion set formalism (or Peak model) may not surpass this level of accuracy unless a more optimal window function or other degrees of freedom are chosen to compensate for this effect. For instance, it would be interesting to carry out a particle-by-particle analysis with the smooth$-k$ space filter proposed by Leo et al. (2018). They smooth the sharp transition from kR to 0 of the $\widetilde{W}_{\rm sk}$ by introducing a free parameter β. Notably, the HMF prediction with this filter only improves that of the sharp $k-$space filter for WDM with free-streaming mass in the order of $10^{10}\,M_\odot\,h^{-1}$, corresponding to ~ 1 keV DM and the smallest smoothing mass we calculate for the $\widetilde{W}_{\rm sk}$[3] (Fig. 1). While their analysis focuses only on improving the HMF predictions, we anticipate it might, at best, increase particle-by-particle accuracy to a level comparable to that of the top-hat and Gaussian filters.

The second assumption we tested concerns the *simple deterministic mass-mapping relation* (equation 11), which accounts for the amount of mass enclosed within a (spherically symmetric) filter scale. This process requires mapping the Lagrangian smoothing scale – where the collapse criterion is first fulfilled – to an estimated halo mass and is essential for inferring HMFs, merger histories, etc. By comparing the collapse smoothing scales with actual halo masses from a CDM simulation, we found two key insights:

(i) *Deviations from the standard mass relation*: The proportionality constant in the conventional $M \propto R^3$ relation often deviates from naive expectations and should be treated as a degree of freedom in excursion set formalisms (see also Maggiore & Riotto 2010a). Reports of excursion set models performing better with collapse barriers above (e.g. Verza et al. 2024) the standard spherical collapse threshold $\delta_c = 1.686$ may stem from misattribution of this degree of freedom. Alternative parametrizations of this relation have also been proposed – such as those for sharp and smooth $k-$space filters (e.g. Schneider 2015; Leo et al. 2018) – but still require calibration against simulations. Additionally, Hahn & Paranjape (2014) found a strong correlation between the scatter in excursion set-predicted masses (at fixed halo masses) and the shape of Lagrangian patches. Particularly, they found them to be mostly aspherical for low-mass, and spherical for high-mass haloes, which raises questions about focusing on spherically symmetric smoothing functions in excursion set analysis (see also Achitouv et al. 2014).

(ii) *Significant uncertainty in the mass relation*: The mass-mapping relation exhibits substantial uncertainty of order 0.4–0.8 dex. This stochasticity effectively introduces additional smoothing in the mass-weighted HMF. A similar observation was made by Hahn & Paranjape (2014), who incorporated mass scatter into a modified HMF expression, achieving better agreement with simulations. In future work we could address these insights by running simulations in a larger volume to increase the sample of high-mass haloes ($\geq 10^{13}\,M_\odot\,h^{-1}$), confirming the extent of mass scatter at these

[3]They use the top-hat mass-mapping relation with an added free parameter c (also calibrated with simulations), yielding slightly different M values.

scales. While this effect is unlikely to significantly impact HMF predictions in CDM universes, it could be crucial in cases where the mass function has sharp features, such as those in universes with a power spectrum cut-off. Whether the excursion set theory can be improved in such cosmologies by including this effect will be interesting to investigate in future studies.

These insights highlight the necessity for further research in order to understand how the halo mass of a collapsed fluid element can be efficiently determined from additional simple criteria.

Further, we showed that multiplicity functions $f(\sigma)$ obtained with a 'perfect' collapse model, which is by nature triaxial, lie significantly closer to the analytical result of Press & Schechter (1974) than they do to the Tinker et al. (2008) function, even though the latter is motivated by the ellipsoidal collapse model. All three excursion sets of simulations result in higher numbers of collapsed haloes than the two predictions suggest.

Lastly, we have also used our excursion sets to measure the Lagrangian scales of pancakes, filaments, and haloes that typical fluid elements may be part of. Interestingly, the typical Lagrangian scales of pancakes tend to be significantly larger than those of haloes and filaments. Approximately 50 per cent of particles will be part of a halo with $R \gtrsim 0.5$ Mpc h^{-1}, a filament with $R \gtrsim 0.8$ Mpc h^{-1}, and a pancake with $R \gtrsim 1.8$ Mpc h^{-1}. Our measurements can serve as a benchmark for anisotropic collapse models,[4] and they suggest that an appropriate treatment of the cosmic web requires collapse models based on the full deformation tensor, rather than effective density-based descriptions (Shen et al. 2006).

Based on the findings of our investigation, we conclude that testing the detailed predictions of excursion set models against simulations may guide enhanced models, as well as help to build better intuition about hierarchical clustering. We will continue to explore these avenues in future work, where we will investigate the extent to which the approximate collapse models can accurately represent the triaxial evolution of a cosmic fluid element. Finally, we also note that due attention should be given to exploring the excursion set theory for merger tree statistics and conditional mass functions in vein of the approach adapted by us in this work.

ACKNOWLEDGEMENTS

The authors thank the reviewer, Sten Delos, for his insightful report with valuable suggestions for improvement of this manuscript. The authors also thank Simon White for his useful comments on the draft. JS acknowledges support from the Austrian Science Fund (FWF) under the ESPRIT project number ESP 705-N. REA acknowledges support from project PID2021-128338NB-I00 from the Spanish Ministry of Science and support from the European Research Executive Agency HORIZON-MSCA-2021-SE-01 Research and Innovation program under the Marie Skłodowska-Curie grant agreement number 101086388 (LACEGAL). The computational results presented have been achieved using the Vienna Scientific Cluster (VSC).

DATA AVAILABILITY

All data used in this article are available upon reasonable request from the corresponding author.

[4]The values quoted here are for the sharp$-k$ window, however, other kernels lead to similar results.

REFERENCES

Abel T., Hahn O., Kaehler R., 2012, MNRAS, 427, 61
Achitouv I., Rasera Y., Sheth R. K., Corasaniti P. S., 2013, Phys. Rev. Lett., 111, 231303
Achitouv I., Wagner C., Weller J., Rasera Y., 2014, J. Cosmol. Astropart. Phys., 2014, 077
Angulo R. E., Hahn O., 2022, Living Rev. Comput. Astrophys., 8, 1
Angulo R. E., White S. D. M., 2010, MNRAS, 401, 1796
Angulo R. E., Springel V., White S. D. M., Jenkins A., Baugh C. M., Frenk C. S., 2012, MNRAS, 426, 2046
Angulo R. E., Hahn O., Ludlow A. D., Bonoli S., 2017, MNRAS, 471, 4687
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Baumann D., Nicolis A., Senatore L., Zaldarriaga M., 2012, J. Cosmol. Astropart. Phys., 2012, 051
Benson A. J. et al., 2013, MNRAS, 428, 1774
Bertone G., Tait T. M. P., 2018, Nature, 562, 51
Bocquet S., Heitmann K., Habib S., Lawrence E., Uram T., Frontiere N., Pope A., Finkel H., 2020, ApJ, 901, 5
Bohr S., Zavala J., Cyr-Racine F.-Y., Vogelsberger M., 2021, MNRAS, 506, 128
Bond J. R., Myers S. T., 1996, ApJS, 103, 1
Bond J. R., Cole S., Efstathiou G., Kaiser N., 1991, ApJ, 379, 440
Bond J. R., Kofman L., Pogosyan D., 1996, Nature, 380, 603
Bringmann T., Weniger C., 2012, Phys. Dark Universe, 1, 194
Buehlmann M., Hahn O., 2019, MNRAS, 487, 228
Carrasco J. J. M., Hertzberg M. P., Senatore L., 2012, J. High Energy Phys., 2012, 82
Catelan P., Lucchin F., Matarrese S., Porciani C., 1998, MNRAS, 297, 692
Chiueh T., Lee J., 2001, ApJ, 555, 83
Corasaniti P. S., Achitouv I., 2011, Phys. Rev. D, 84, 023009
Delos M. S., 2024, MNRAS, 528, 1372
Delos M. S., White S. D. M., 2023a, MNRAS, 518, 3509
Delos M. S., White S. D. M., 2023b, J. Cosmol. Astropart. Phys., 2023, 008
Delos M. S., Korsmeier M., Widmark A., Blanco C., Linden T., White S. D. M., 2024, Phys. Rev. D, 109, 083512
Despali G., Giocoli C., Angulo R. E., Tormen G., Sheth R. K., Baso G., Moscardini L., 2016, MNRAS, 456, 2486
Diemand J., Moore B., Stadel J., 2005, Nature, 433, 389
Eke V. R., Cole S., Frenk C. S., 1996, MNRAS, 282, 263
Falck B. L., Neyrinck M. C., Szalay A. S., 2012, ApJ, 754, 126
Farahi A., Benson A. J., 2013, MNRAS, 433, 3428
Gilman D., Birrer S., Nierenberg A., Oh M. S. H., 2024, MNRAS, 533, 1687
Hahn O., Abel T., 2011, MNRAS, 415, 2101
Hahn O., Angulo R. E., 2016, MNRAS, 455, 1115
Hahn O., Paranjape A., 2014, MNRAS, 438, 878
Hsueh J. W., Enzi W., Vegetti S., Auger M. W., Fassnacht C. D., Despali G., Koopmans L. V. E., McKean J. P., 2020, MNRAS, 492, 3047
Ishiyama T., Makino J., Ebisuzaki T., 2010, ApJ, 723, L195
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Knebe A. et al., 2011, MNRAS, 415, 2293
Knebe A. et al., 2013, MNRAS, 435, 1618
Koopmans L. V. E., 2005, MNRAS, 363, 1136
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lapi A., Salucci P., Danese L., 2013, ApJ, 772, 85
Lee J., 2006, preprint(astro–ph/0605697)
Leo M., Baugh C. M., Li B., Pascoli S., 2018, J. Cosmol. Astropart. Phys., 2018, 010
Liu Y. et al., 2024, MNRAS, 527, 11740
López-Cano D., Stücker J., Pellejero Ibañez M., Angulo R. E., Franco-Barranco D., 2024, A&A, 685, A37
Lucie-Smith L., Peiris H. V., Pontzen A., Lochner M., 2018, MNRAS, 479, 3405
Lucie-Smith L., Peiris H. V., Pontzen A., 2019, MNRAS, 490, 331
Lucie-Smith L., Peiris H. V., Pontzen A., Nord B., Thiyagalingam J., 2024, Phys. Rev. D, 109, 063524
Ludlow A. D., Porciani C., 2011, MNRAS, 413, 1961

Ludlow A. D., Borzyszkowski M., Porciani C., 2014, MNRAS, 445, 4110
Ludlow A. D., Bose S., Angulo R. E., Wang L., Hellwing W. A., Navarro J. F., Cole S., Frenk C. S., 2016, MNRAS, 460, 1214
Ma C.-P., Maggiore M., Riotto A., Zhang J., 2011, MNRAS, 411, 2644
Maggiore M., Riotto A., 2010a, ApJ, 711, 907
Maggiore M., Riotto A., 2010b, ApJ, 717, 515
McKean J. et al., 2015, Proc. Sci., Strong Gravitational Lensing with the SKA. SISSA, Trieste, PoS 84
Monaco P., 1995, ApJ, 447, 23
Monaco P., Theuns T., Taffoni G., 2002a, MNRAS, 331, 587
Monaco P., Theuns T., Taffoni G., Governato F., Quinn T., Stadel J., 2002b, ApJ, 564, 8
Musso M., Sheth R. K., 2021, MNRAS, 508, 3634
Musso M., Despali G., Sheth R. K., 2024, A&A, 690, A214
Neistein E., Macciò A. V., Dekel A., 2010, MNRAS, 403, 984
Ondaro-Mallea L., Angulo R. E., Stücker J., Hahn O., White S. D. M., 2024, MNRAS, 527, 10802
Peebles P. J. E., 1981, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, https://doi.org/10.1515/9780691206714
Press W. H., Schechter P., 1974, ApJ, 187, 425
Ramachandra N. S., Shandarin S. F., 2017, MNRAS, 467, 1748
Rampf C., Saga S., Taruya A., Colombi S., 2023, Phys. Rev. D, 108, 103513
Reed D., Gardner J., Quinn T., Stadel J., Fardal M., Lake G., Governato F., 2003, MNRAS, 346, 565
Reed D. S., Bower R., Frenk C. S., Jenkins A., Theuns T., 2006, MNRAS, 374, 2
Robertson B. E., Kravtsov A. V., Tinker J., Zentner A. R., 2009, ApJ, 696, 636
Rubiño-Martín J. A., Betancort-Rijo J., Patiri S. G., 2008, MNRAS, 386, 2181
Sandvik H. B., Möller O., Lee J., White S. D. M., 2007, MNRAS, 377, 234
Schneider A., 2015, MNRAS, 451, 3117
Schneider A., Smith R. E., Reed D., 2013, MNRAS, 433, 1573
Shandarin S. F., Medvedev M. V., 2017, MNRAS, 468, 4056
Shandarin S., Habib S., Heitmann K., 2012, Phys. Rev. D, 85, 083005
Shen J., Abel T., Mo H. J., Sheth R. K., 2006, ApJ, 645, 783
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Sheth R. K., Tormen G., 2002, MNRAS, 329, 61
Sheth R. K., Mo H. J., Tormen G., 2001, MNRAS, 323, 1
de Simone A., Maggiore M., Riotto A., 2011, MNRAS, 418, 2403
Somerville R. S., Kolatt T. S., 1999, MNRAS, 305, 1
Somerville R. S., Lemson G., Kolatt T. S., Dekel A., 2000, MNRAS, 316, 479
Sousbie T., Colombi S., 2016, J. Comput. Phys., 321, 644
Springel V., 2005, MNRAS, 364, 1105
Springel V., Pakmor R., Zier O., Reinecke M., 2021, MNRAS, 506, 2871
Stücker J., Busch P., White S. D. M., 2018, MNRAS, 477, 3230
Stücker J., Hahn O., Angulo R. E., White S. D. M., 2020, MNRAS, 495, 4943
Stücker J., Angulo R. E., Hahn O., White S. D. M., 2022, MNRAS, 509, 1703
Stücker J., Ogiya G., White S. D. M., Angulo R. E., 2023, MNRAS, 523, 1067
Taruya A., Suto Y., 2000, ApJ, 542, 559
Tinker J., Kravtsov A. V., Klypin A., Abazajian K., Warren M., Yepes G., Gottlöber S., Holz D. E., 2008, ApJ, 688, 709
Tramonte D., Rubiño-Martín J. A., Betancort-Rijo J., Dalla Vecchia C., 2017, MNRAS, 467, 3424
Vegetti S., Koopmans L. V. E., 2009, MNRAS, 392, 945
Verza G., Carbone C., Pisani A., Porciani C., Matarrese S., 2024, J. Cosmol. Astropart. Phys., 2024, 079
Vogelsberger M., White S. D. M., 2011, MNRAS, 413, 1419
Vogelsberger M., White S. D. M., Helmi A., Springel V., 2008, MNRAS, 385, 236
Wang J., Bose S., Frenk C. S., Gao L., Jenkins A., Springel V., White S. D. M., 2020, Nature, 585, 39
Warren M. S., Abazajian K., Holz D. E., Teodoro L., 2006, ApJ, 646, 881
White S. D. M., 1994, preprint(astro–ph/9410043)
White S. D. M., Silk J., 1979, ApJ, 231, 1
Zel'dovich Y. B., 1970, A&A, 5, 84
Zentner A. R., 2007, Int. J. Mod. Phys. D, 16, 763
Zheng H., Bose S., Frenk C. S., Gao L., Jenkins A., Liao S., Liu Y., Wang J., 2024, MNRAS, 528, 7300

APPENDIX A: CONVERGENCE TEST

We conducted a convergence test on the σ^2_{16} simulation, with the sharp$-k$ smoothing. Given that this simulation has the smallest-scale smoothing, it is the most susceptible to undesirable effects of all the simulations in our sample. Namely, the particle-mesh spacing of 78 kpc h^{-1} in an $N_{\rm pm} = 512^3$ simulation is only moderately smaller compared to the $\sim$ 140 kpc h^{-1} smoothing radius of σ^2_{16}. Such proximity in scales raises a concern that the size of a collapsing region may approach or fall below the force resolution limit, potentially reducing the number of particles assigned to high morphology rank due to delayed collapse in our simulations.

As a benchmark for convergence testing we ran a simulation with the particle mesh resolution of $N_{\rm pm} = 1024^3$ (particle-mesh spacing of 39 kpc h^{-1}) in contrast to $N_{\rm pm} = 512^3$, which we chose for the simulations used in our analysis. We refer to the former as N_{1024} and N_{512} to the latter in what follows. In addition, we increased the particle deposit level parameter eightfold in the N_{1024} simulation. This parameter controls the number of additional 'fake' particles created in the process of defining the density field in high-density regions.

In Fig. A1 we plot the relative error

$$\varepsilon = \frac{a_{512} - a_{1024}}{a_{1024}} \tag{A1}$$

of collapse times $a_{\rm x}$ into pancakes (rosy brown), filaments (grey), and haloes (teal), between the two simulations, where the benchmark value is 0 (light-grey dashed). The boundaries enclosing the shaded regions represent the 90th (top) and the 10th (bottom) percentiles of the N_{512} distribution relative to N_{1024} (light-grey dashed), while the medians of the distributions are shown as coloured dashed lines. Negative values of ε indicate that those particles experience delayed collapse in the higher resolution simulation relative to N_{512}, that is, $a_{512} < a_{1024}$, while positive ε implies delayed collapse $a_{512} > a_{1024}$ in the lower resolution simulation.

For all three morphology classes $a_{512} > 0$ for the overwhelming majority of particles. This means that in the lower resolution simulation most of the particles indeed collapse later relative to the benchmark simulation. However, the relative error for the collapse of the first axis (pancake) remains within $\sim$ 5 per cent for 80 per cent of particles collapsed along at least one axis, through nearly the entire evolution of the simulation volume. For the second axis collapse (filament) ε is also $\sim$ 5 per cent at earlier times, increasing to $\sim$ 7 per cent at later times for 80 per cent of the particles. Unsurprisingly, the biggest deviation from the N_{1024} benchmark occurs in the collapse time of the third axis (halo), where 80 per cent of particles remain within $\varepsilon = 10$ per cent until $a \gtrsim 0.95$. At $a = 1$ the relative error reaches 25 per cent (not visible in the plot). The complexity of chaotic conditions in the environment of particles accreting on to haloes acts to diminish the predictive abilities of simulations. Despite these variations, all of the median error curves stay well within the $\varepsilon \sim$ 2–3 per cent of the benchmark for $\sim$ 70 per cent of the simulation's evolution, with slightly higher error ($\varepsilon \sim 5$ per cent) at earlier times.

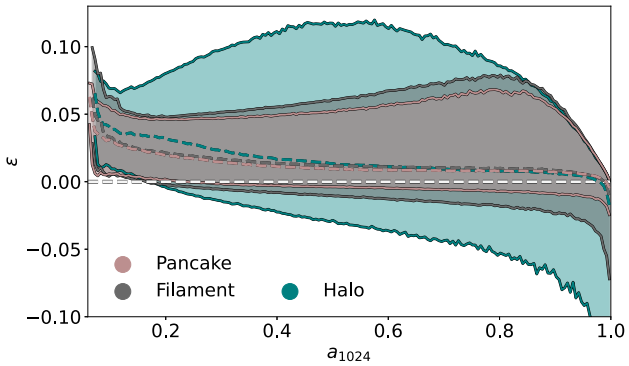

Figure A1. The relative error ε of collapse times into pancakes (rosy brown), filaments (grey), and haloes (teal), between the N_{512} (ours) and N_{1024} (benchmark) simulations as a function of the scale factor a_{1024} of the benchmark simulation. The boundaries around the shaded regions are the 90th (top) and the 10th (bottom) percentiles of the N_{512} distribution about N_{1024} (light-grey dashed). Medians of the distributions are plotted as coloured dashed lines.

Given the minor level of difference in collapse times between these two simulations, we conclude that using $N_{\rm pm} = 512^3$ does not have a significant impact on the conclusions reached in this work. Therefore, we chose the lower particle-mesh resolution for this work as it saves on simulation run time – a N_{512} simulation requires significantly less time, ~ 1 d, as compared to 4 d for a N_{1024} simulation. We expect similar behaviour in ΛCDM simulations, hence we are confident that the force resolution of 78 kpc h^{-1} is also an appropriate choice in this case and will not significantly impact our conclusions.

APPENDIX B: MASS FRACTIONS

Here, we also present the mass fraction $F(<\sigma^2)$ plot for the example of the top-hat smoothing function. Notice how the turquoise curve, adapted from Stücker et al. (2018), is not as good a match for the the mass fraction of $n \geq 1$ regions, as it was when the smoothing was carried out with the sharp$-k$ space filter. The reason for the discrepancy observed in Fig. B1 is the fact that Stücker et al. (2018) also utilized the sharp$-k$ space filter for generating their random walks. Therefore, their results are more relevant for simulations smoothed out with this particular filter.

Out of curiosity, in Fig. B2 we show the multiplicity functions $f(\sigma)$ for pancakes (stars) and filaments (hexagons), resulting from excursion sets of our top-hat (yellow), Gaussian (violet), and sharp$-k$ (blue) smoothed simulations. For reference, we also plot the $f(\sigma)$s of haloes from Fig. 11 faintly in the background. An interesting feature, which stands out in this figure, is the behaviour of the small-scale end of the pancake $f(\sigma)$s. Notice how, contrary to what is observed in the filament and the halo examples, the top-hat and Gaussian excursion sets result in *higher* mass fractions winding up in pancakes of scales with $\sigma > 0.5$, compared to the sharp $-k$ function.

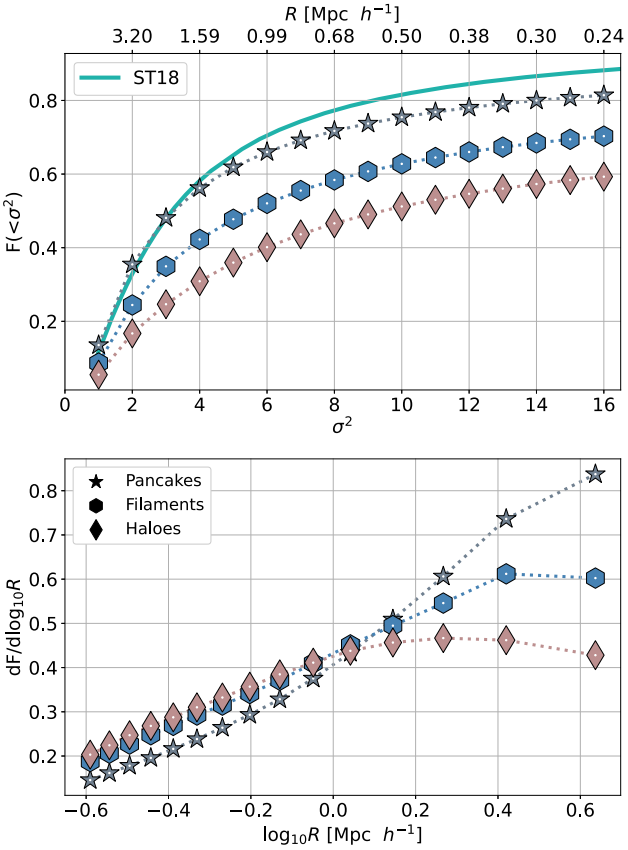

Figure B1. Top: the mass fractions of particles in pancakes (stars), filaments (hexagons), and haloes (diamonds) as a function of the smoothing scale σ^2 and R. The turquoise curve represents $F(<\sigma^2)$ of collapsed particles from Stücker et al. (2018). Bottom: the particle fraction in structures of the ES radii. Both panels show the distributions for the $\widetilde{W}_{\rm th}$ smoothed excursion set of simulations.

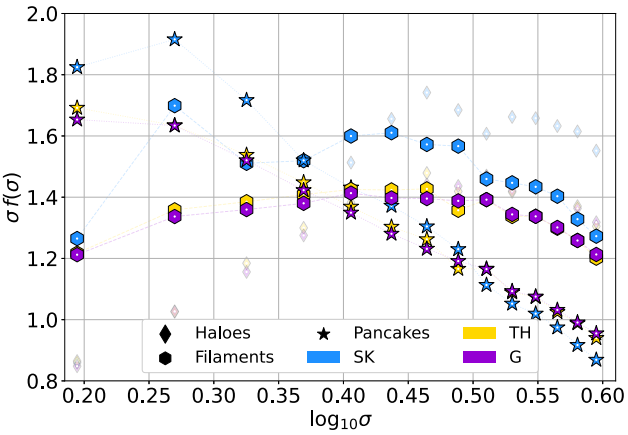

Figure B2. The multiplicity functions $f(\sigma)$ for pancakes (stars) and filaments (hexagons). The halo (diamonds) $f(\sigma)$s from Fig. 11 are plotted faintly in the background for reference.

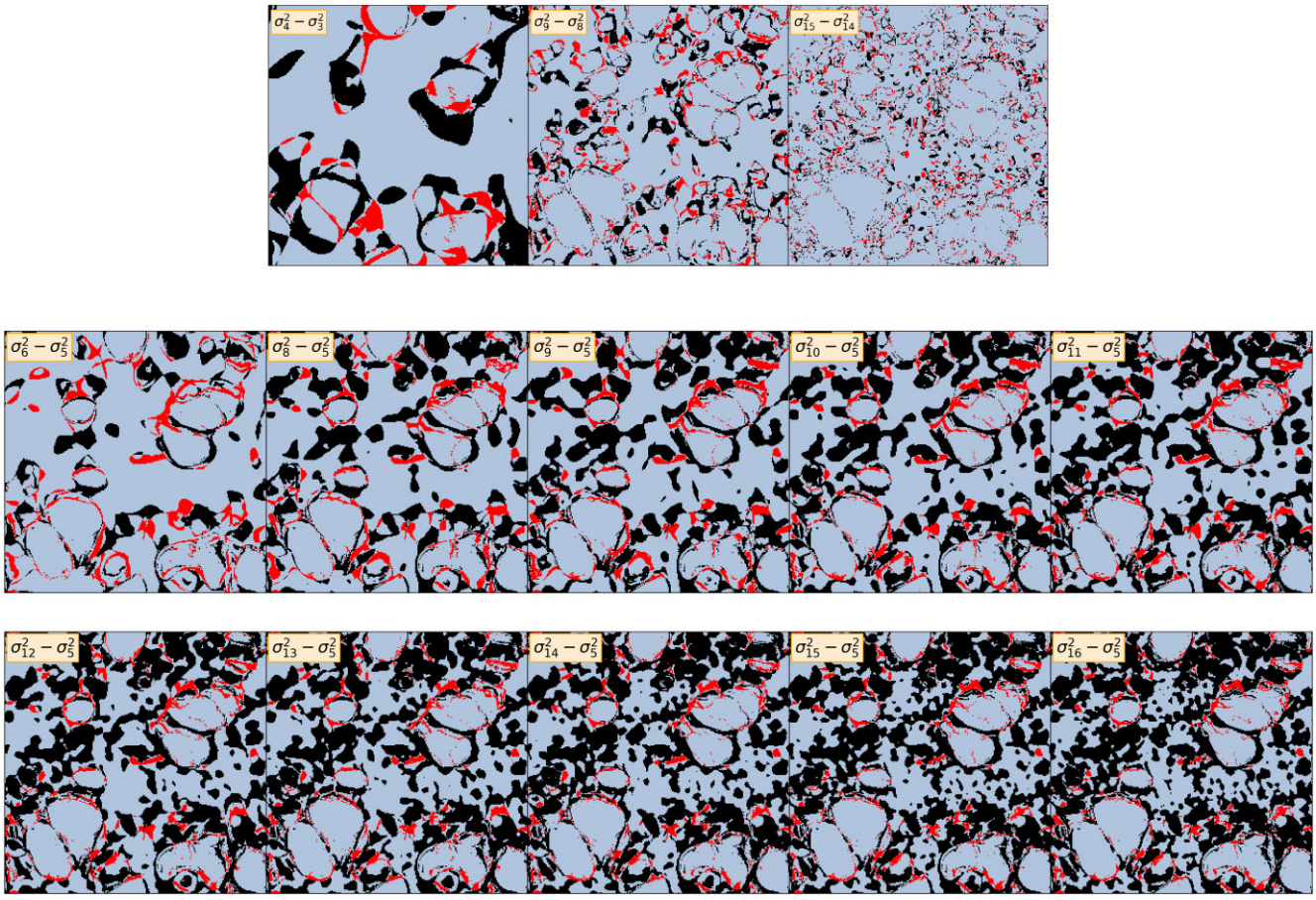

Figure C1. Difference plots of $\Delta\sigma^2 = \sigma_i^2 - \sigma_j^2$, where $i > j$. Top: three examples of $\Delta\sigma^2 = 1$, where the scale of un-collapsing is most prominent. Bottom: evolution of the difference field with increasing $\Delta\sigma^2$ (rightwards direction) for σ_5^2.

APPENDIX C: UN-COLLAPSE

Fig. C1 presents several examples in Lagrangian coordinates, where un-collapse (red regions) is observed for different values of $\Delta\sigma^2$ in simulations smoothed with the sharp$-k$ space window function. This plot complements Fig. 6, where it can be seen that as the scale of smoothing decreases, so does the scale of un-collapsed objects with an increasing $\Delta\sigma^2$. The grey regions represent coordinates where there is no difference between σ_i^2 and σ_j^2, while the black patches indicate a higher rank of collapse in σ_i^2 with respect to σ_j^2.

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