An inventory of galaxies in cosmic filaments feeding galaxy clusters: galaxy groups, backsplash galaxies, and pristine galaxies

Ulrike Kuchner [⬤],[1]★ Roan Haggar [⬤],[1] Alfonso Aragón-Salamanca [⬤],[1] Frazer R. Pearce,[1] Meghan E. Gray,[1] Agustín Rost [⬤],[2] Weiguang Cui [⬤],[3] Alexander Knebe [⬤][4,5,6] and Gustavo Yepes [⬤][4,5]

[1]*School of Physics, Astronomy, University of Nottingham, Nottingham NG7 2RD, UK*
[2]*Instituto de Astronomía Teórica y Experimental (IATE), Laprida 854, Córdoba, Argentina*
[3]*Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, United Kingdom*
[4]*Departamento de Física Teórica, Módulo 15, Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain*
[5]*Centro de Investigación Avanzada en Física Fundamental (CIAFF), Facultad de Ciencias, Universidad Autónoma de Madrid, 28049 Madrid, Spain*
[6]*International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Highway, Crawley, Western Australia 6009, Australia*

ABSTRACT

Galaxy clusters grow by accreting galaxies from the field and along filaments of the cosmic web. As galaxies are accreted they are affected by their local environment before they enter (pre-processing), and traverse the cluster potential. Observations that aim to constrain pre-processing are challenging to interpret because filaments comprise a heterogeneous range of environments including groups of galaxies embedded within them and backsplash galaxies that contain a record of their previous passage through the cluster. This motivates using modern cosmological simulations to dissect the population of galaxies found in filaments that are feeding clusters, to better understand their history, and aid the interpretation of observations. We use zoom-in simulations from THE THREEHUNDRED project to track haloes through time and identify their environment. We establish a benchmark for galaxies in cluster infall regions that supports the reconstruction of the different modes of pre-processing. We find that up to 45 per cent of all galaxies fall into clusters via filaments (closer than $1\,h^{-1}$Mpc from the filament spine). 12 per cent of these filament galaxies are long-established members of groups and between 30 and 60 per cent of filament galaxies at R_{200} are backsplash galaxies. This number depends on the cluster's dynamical state and sharply drops with distance. Backsplash galaxies return to clusters after deflecting widely from their entry trajectory, especially in relaxed clusters. They do not have a preferential location with respect to filaments and cannot collapse to form filaments. The remaining pristine galaxies ($\sim$30–60 per cent) are environmentally affected by cosmic filaments alone.

Key words: large-scale structure of Universe – galaxies: clusters: general – cosmology: observations – methods: data analysis – galaxies: evolution.

1 INTRODUCTION

In a cold dark matter Universe model, low-mass haloes form first. Larger haloes become more common over time, successively building up their mass through merging and accretion of smaller haloes. Galaxy clusters mark the culmination of mass assembly and the peaks of dynamical gravitational structure formation. They are the highest density environments in the large-scale Universe, packed with thousands of galaxies, both in the virialized cluster core and infalling from the highly anisotropic matter distribution surrounding the clusters. Galaxies fall into clusters through a variety of environments: as part of groups, on their own from the general field, or in streams via filaments of the cosmic web (Zel'dovich 1970; van Haarlem & van de Weygaert 1993).

Clusters assemble their *mass* predominantly by massive accretion events, like infalling groups of galaxies (McGee et al. 2009), but build up their *galaxy* population predominantly through the accretion

of lower mass haloes, i.e. isolated galaxies falling in on their own (Berrier et al. 2008). This differentiation could be important for the evolution of galaxies and clusters of galaxies, because different environments evoke different physical processes that depend on the mass of the host (dark matter) halo. Satellite galaxies in high density environments such as clusters and groups differ from isolated galaxies of the same stellar mass in key aspects, such as their colour (e.g. Peng et al. 2010), star formation rate (e.g. Kauffmann et al. 2004; Wetzel et al. 2013), and morphology (e.g. Dressler 1980). Galaxies in denser environments tend to be redder, more elliptical/spheroidal with less gas and ongoing star formation. This well-known finding is grounded on a wealth of observations from galaxies in clusters opposed to galaxies in the general field and tested against a variety of physical processes acting in clusters (Oemler 1974; Dressler 1980; Postman & Geller 1984; Balogh et al. 1997; Poggianti et al. 1999). Galaxies are commonly thought to transform both in terms of star forming activity and morphology as they experience dense environments. Therefore, the environment of galaxies plays a key role in the formation and evolution of galaxies (Blanton & Moustakas 2009). Still, a full description of the

★ E-mail: ulrike.kuchner@nottingham.ac.uk

relationship between galaxies and their environments, including their specific processes (i.e. environmentally driven tidal or hydronamical mechanisms versus internal mass-dependent mechanisms) is still outstanding. A primary complication for an understanding is that we do not know how much of the correlation between galaxy properties and cluster membership is due to a transformation inside the cluster as opposed to in environments prior to entering the cluster, a phenomenon called 'pre-processing' (Zabludoff & Mulchaey 1998; Balogh, Navarro & Morris 2000; Wetzel et al. 2013). While this term is not absolute, it generally refers to any process operating in high density environments that leads to the transformation of galaxies and experienced before the cluster infall. In this paper, 'pre-processing' therefore summarizes *all* environmental effects, including hydrodynamical and gravitaional effects, acting in groups and in large-scale filaments, that affect cluster galaxies before they enter the virialized regions of a cluster. A common indicator for pre-processing is galaxy quenching, as this presents relatively accessible observational evidence of pre-processing, however changes in other galaxy properties like galaxy structures can equally help to constrain pre-processing.

The increasing awareness and current discussion of pre-processing as an important ingredient to galaxy evolution has prompted surveys to focus on cluster outskirts, i.e. observations that go beyond R_{200},[1] in order to identify the sites where galaxies are first affected by their environment before falling into clusters (Fujita & Goto 2004; Porter et al. 2008; Mahajan, Raychaudhury & Pimbblet 2012; Haines et al. 2018b; Sarron et al. 2019; Malavasi et al. 2020). One upcoming dedicated study of cluster infall regions is the WEAVE Wide-Field Cluster Survey (WWFCS) with the multi-object spectrograph WEAVE (WHT Enhanced Area Velocity Explorer) on the William Herschel telescope (Balcells et al. 2010; Dalton et al. 2012). It will systematically observe 20 nearby clusters out to $5R_{200}$ with the goal to determine whether significant pre-processing accelerates the quenching of star formation and/or morphological transformation. Our investigation presented in this and previous papers is motivated by the WWFCS, but the results are universal and equally applicable to a wide range of experiments.

In this paper, we focus on pre-processing in large-scale filaments, which themselves are heterogeneous environments, including galaxy groups embedded within them. Around half of the mass of the Universe is found in cosmic filaments (Cautun et al. 2014; Cui et al. 2018), which, in turn, fundamentally define the spatial organization of galaxies over a vast range of scales from less than one to tens and even hundreds of Megaparsecs (Libeskind et al. 2017; van de Weygaert et al. 2014). A growing body of evidence shows that large-scale filaments play a similar role in shaping the properties of galaxies as clusters do, albeit to a lesser degree. Galaxies close to cosmic web filaments are redder (Kraljic et al. 2018; Laigle et al. 2017), elliptical (Kuutma, Tamm & Tempel 2017), with higher metalliciy (Darvish et al. 2015; Gray et al. 2009), more massive (Malavasi et al. 2016) and more likely to have been quenched (Alpaslan et al. 2016; Winkel et al. 2021) than their counterparts at fixed M∗ at increased distances away from filaments. This can be due to ram pressure that removes the hot haloes especially of lower mass galaxies (Bahé et al. 2013; Benítez-Llambay et al. 2013). While simulations suggest that haloes at the same mass in denser environments form earlier than in less dense environments, owing to the dependence of halo clustering not

only on mass but also on the formation redshift and assembly history (a term coined 'assembly bias', Gao, Springel & White 2005; Jung, Lee & Yi 2014), this may be a simplified view of the problem since mass assembly is driven by different physical processes inside and outside of filaments (Poudel et al. 2017). Mergers, tidal effects and smooth accretion are attributed to different densities and strongly influence the current property of a galaxy beyond its formation time. Differences can also be explained by accretion of pre-enriched filamentary gas (Darvish et al. 2015), which may lead to a star-formation enhancement in filaments (Vulcani et al. 2019) when galaxies are fuelled with gas (Kleiner et al. 2016).

Embedded within the large-scale cosmic web, galaxy groups continue to accrete galaxies and gas (Kauffmann, Li & Heckman 2010). This is especially relevant close to clusters, where infalling groups can easily sweep up field galaxies and grow quickly (Vijayaraghavan & Ricker 2013). Members are likely processed by ram pressure enhanced by feedback within groups prior to their accretion into the clusters themselves (Bahé & McCarthy 2014; Jung et al. 2018), as is evident in observations of galaxy mergers and ram pressure stripping signatures (Jaffé et al. 2016; Bianconi et al. 2017; Haines et al. 2018a; Benavides, Sales & Abadi 2020). Earlier simulations suggest that a significant fraction of all cluster galaxies – some report between a third and half of cluster galaxies at $z = 0$ – could enter clusters as part of groups (McGee et al. 2009; White, Cohn & Smit 2010; De Lucia et al. 2012). However, most galaxies spend relatively little time in groups before falling into the cluster (less than 2.5 Gyrs, Vijayaraghavan & Ricker 2013; Han et al. 2018), so either group environmental mechanisms must act fast to be significant for the cluster population, or only group members that have spent extended periods of time in their host halo are measurably affected and indeed pre-processed. Either way, most groups are part of filaments (e.g. Tempel et al. 2014), and therefore a number of filament galaxies are actually processed by their group environment. To unambiguously identify the effect of filaments on galaxy evolution, it may be necessary to remove the contribution of groups.

After galaxies are accreted by the cluster, they either remain bound to the gravitational potential well of the cluster, or their trajectories carry them out of the cluster, up to several R_{200}, where they will turn around to fall back in on a subsequent infall. This population of 'backsplash galaxies' is no small fraction: immediately outside of clusters, up to 70 per cent of all galaxies can be backsplash galaxies (Gill, Knebe & Gibson 2005; Haines et al. 2015; Haggar et al. 2020) and have therefore been processed by the cluster itself. By the time they are observed as backsplash galaxies, they may reveal their past environmental history through 'post-processing' signatures that are all but indistinguishable to pre-processing signatures. Beyond this complication, other possible processing mechanisms induced in accretion shocks or when crossing cosmic web walls ('wall stripping', Winkel et al. 2021) can strip halo gas which leads to star formation consumption and quenching, especially in low-mass galaxies.

As a direct consequence of structure formation, galaxies falling into clusters are therefore a combination of 'field galaxies' – both isolated and as pairs and small groups – and galaxies in filaments – again, isolated and as part of groups – as well as backsplash galaxies. Given this diversity, pre-processing studies need to take the entire environmental history of galaxies over a lifetime spent in a hierarchically assembling global environment into account.

This paper sets out to provide a census of the fractions of galaxies that feed clusters from a variety of evolving environments and investigates whether this varying composition depends on the

---

[1]The radius within which the mean density of a cluster is equal to 200 times the critical density of the Universe and used by us as defining the extent of the cluster.

dynamical state or mass of the cluster. This can inform analysis of observational signatures of star formation histories against measured environments which investigate galaxy transformation. Our study involves tracing the filamentary structure beyond the virial radius in large hydrodynamical simulations while also considering the orbital trajectories of infalling galaxies. After detailing the simulations (Section 2.1), we discuss the identification of the main components, filaments (Section 2.2), groups (Section 2.3) and backsplash galaxies (Section 2.4). We then discuss the importance of group galaxies and filaments (Section 3.1) and the contamination of backsplash galaxies in filaments, where we separate galaxies that are leaving the cluster from returning galaxies (Section 3.2). Our final section summarizes the heterogeneous composition of filament galaxies.

2 SIMULATIONS AND METHODS

2.1 THE THREEHUNDRED clusters

To help interpret observations of galaxy properties feeding clusters, we need to know their environmental history during accretion. To know this means to turn to simulations. This paper makes use of hydrodynamical simulations of THE THREEHUNDRED project[2] (Cui et al. 2018). This project selected 324 spherical regions with radius 15 h^{-1}Mpc centered on the most massive clusters ($M_{200} \gtrsim \times 10^{14}\ h^{-1}M_\odot$) in the 1 h^{-1}Gpc volume of the dark-matter-only MultiDark simulation (Klypin et al. 2016) with Planck cosmology (Ade et al. 2016). THE THREEHUNDRED were simulated using a range of different physics models. The suite contains the same clusters simulated with Gadget-Music (Sembolini et al. 2012), Gadget-X (Beck et al. 2015; Rasia et al. 2015) and GIZMO-Simba (Davé et al. 2019) amongst others, as well as several semi-analytic models, producing 129 snapshots from redshift $z \sim 17$ to 0. For a comprehensive description and discussion of the full-physics treatment, comparison and limitations of codes and the AHF-halo finding of THE THREEHUNDRED, we refer to the survey description paper by Cui et al. (2018) and references therein. For the work presented in this paper, we only use the mass distribution of the full physics simulations performed with Gadget-X to generate our filamentary network, because the goal of this investigation does not require further information. With the exception of tracing the infall of galaxies to identify backsplash galaxies, we restrict our current investigation to redshift z = 0, both motivated by the wish to minimize evolutionary effects, and preparing for upcoming observations with WEAVE. We will expand on this in a future publication (Cornwell et al. in prep.).

In summary, these hydrodynamic simulations of galaxy clusters return information in six dimensional phase space over numerous time steps in a volume of several virial radii of the clusters, i.e. large enough to include many additional groups and filaments, which may or may not be physically associated with the central cluster and useful to track infall. The sample also includes volumes that host pairs of clusters. We assess the dynamical state of the cluster, 'relaxedness', based on a combination of three characteristic parameters that capture signatures of activity. These are 1) the virial ratio (a measure of virialization of the cluster), 2) the centre-of-mass offset from the maximum density point, and 3) the fraction of mass in subhalos (see Cui et al. 2017, 2018; Haggar et al. 2020, where this has been discussed in detail). A cluster is considered relaxed if it has a low fraction of mass in subhalos, low centre-of-mass offset and virial
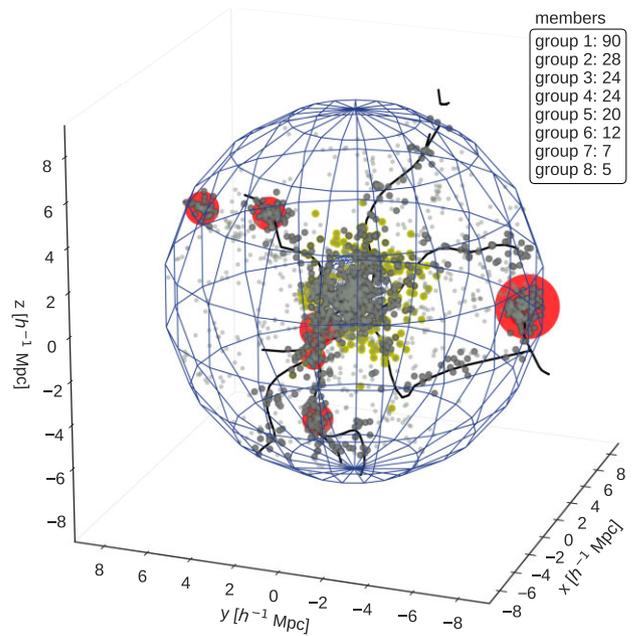
Figure 1. The example cluster (cluster 0066) of THE THREEHUNDRED project illustrates the variety of environments and processing histories galaxies around clusters can have. Shown are galaxies within 1 h^{-1}Mpc of galaxy-detected filaments (dark grey), groups (highlighted by red disks), the general 'field' (light grey) and backsplash galaxies (yellow). Each environment is related to mechanisms that may pre-process the galaxies. Backsplash galaxies have been environmentally affected by the cluster itself on their previous pass through the cluster. The large mesh sphere indicates $5R_{200}$, the small sphere $1R_{200}$. The insert lists the number of group members for this example.

ratio approaching 1. Clusters with a higher 'relaxedness' parameter R and specifically with $R > 1$ are considered more relaxed, clusters with $R < 1$ as unrelaxed or dynamically active.

How reliably we can separate galaxies in groups, filaments or cluster outskirts is fundamental for studying the effects of galaxy evolution and pre-processing. Systematics in classifications can bias our view of pre-processing and hamper the compatibility of simulations and observations. Simulations can help to quantify the effect of every environment a galaxy experiences during its lifetime but some care needs to be taken to bridge simulations to observations. In previous publications, we have detailed how transforming the simulations into realistic mock observations allow to forecast the impact of projection effects and the reliability of filament finding for upcoming wide-field spectroscopic surveys (Kuchner et al. 2020, 2021). In a next step, we will investigate the effects of further observational constraints such as fibre collisions during the production of observing blocks on finding filaments in the crowded regions of galaxy clusters (Cornwell et al in prep.). While we especially focus on mimicking observations that will be obtained with the WEAVE Wide-Field Cluster survey as part of the community-led surveys with the new spectroscopic facility WEAVE at the WHT (see Introduction Section 1, as well as Kuchner et al. (2020) and Jin et al in prep.), we emphasize that the results are more general, and valid for a number of observational applications.

2.2 Filament identification

The paper considers major filaments around clusters that can be thought of as *highways* or *transport channels* of the Universe (Fig. 1), along which mass and galaxies get funneled into clusters (e.g. van Haarlem & van de Weygaert 1993; Knebe et al. 2004). To extract

cosmic web filaments in each volume of our sample at $z = 0$, we used the robust filament finding algorithm DISPERSE (Sousbie 2011). We have applied the software on a discrete point distribution of mock galaxies in 3D and 2D – a useful and well established approach in both simulated and observed datasets across scales from sub-galactic to cosmological interests (e.g. Malavasi et al. 2016, 2020; Kraljic et al. 2018; Hess et al. 2018; Arzoumanian et al. 2019; Winkel et al. 2021). For our purpose, we define mock galaxies as all haloes with masses $M_{\text{halo}} > 3 \times 10^{10} \, h^{-1}\text{M}_\odot$ (comparable to $M_* > 3 \times 10^9 \, h^{-1}\text{M}_\odot{}^3$) and use them as input to DISPERSE. The software processes the data in two steps. For filaments used in this paper, the software first computes the density from the Delaunay Tesselation on the 3D halo distribution, which we post-process by weighting by halo mass of each mock galaxy. Then, DISPERSE identifies the critical points in the density field; in 3D, these are minima, two kind of saddle points, and maxima. The filament extraction is theoretically motivated: filaments are defined as the spatial lines following the gradients of the density field that connect saddle points to peaks. Not all critical points that DISPERSE extracts have the same significance with respect to noise. The significance of a pair of critical points (e.g. a saddle-to-peak pair) with respect to the noise is quantified by the persistence parameter σ, which is a user-controlled input parameter. This way, low persistence features can be filtered out, which in turn allows to work with noisy data sets and to remove features that are not physically meaningful. In Kuchner et al. (2020) we compared filament networks based on mock galaxies in true 3D coordinates to the networks based on the underlying gas distribution, which we considered as our reference network. The result of this assessment was a persistence threshold of $\sigma = 6.5$ appropriate for finding filaments based on mock galaxies around massive clusters.

The output of the algorithm is a set of critical points and spatial lines presented as small segments of the filament axes (or skeleton). We can therefore compute the distance of each mock galaxy to the filament axes, a useful parameter to investigate gradients of galaxy properties (e.g. Laigle et al. 2017; Kraljic et al. 2018). DISPERSE does not give information whether a galaxy is 'inside' or 'outside' a filament. In order to compute the filament diameter or width, an additional parametrization is required. In Kuchner et al. (2020), we have defined the filament width based on density profiles of gas particles as a way to provide a convenient 'inside/outside' definition for observational applications. We have taken care to choose an appropriate fixed width, trying to optimize completeness without increasing the contamination. Depending on the science goal, we defined mock galaxies with distances to filament axes (skeleton) of $D_{\text{skel}} < 0.7 \, h^{-1}\text{Mpc}$ (for maximum purity) or $D_{\text{skel}} < 1 \, h^{-1}\text{Mpc}$ (for maximum completeness) to be 'inside' filaments. Note that a constant thickness and basic segregation is a simplification that does not properly reflect the diffuse characteristic of filament gas and galaxies collapsing towards filament spines, nor does it properly capture the variation of filament thickness closer to haloes including at locations of massive groups (Dolag et al. 2006; Rost et al. 2020).

In this paper, we define filaments with a constant thickness of $1 \, h^{-1}\text{Mpc}$, i.e. all mock galaxies with a distance of less than $1 \, h^{-1}\text{Mpc}$ to the skeleton ($D_{\text{skel}} < 1 \, h^{-1}\text{Mpc}$) are considered filament galaxies (Fig. 1). This is similar to choices made in previous publications (e.g. Colberg, Krughoff & Connolly 2005; Tempel et al. 2014; Kooistra et al. 2019). We also note values for a more restricted

filament thickness of $0.7 \, h^{-1}\text{Mpc}$ in the text. All numbers thus depend on the choice of filament thickness, which in turn depends on the science case and emphasis on e.g. purity versus completeness. The density profile discussed in Fig. 6 in Kuchner et al. (2020) shows that the profile drops steeply beyond $1 \, h^{-1}\text{Mpc}$. Increasing the filament thickness by a factor of two ($D_{\text{skel}} < 2 \, h^{-1}\text{Mpc}$) therefore leads to a large increase of contamination while overall only adding ~4 per cent of galaxies that are located in the true periphery of filaments. Importantly, it is not clear whether these galaxies will experience any environmental effect in filament peripheries, since at $2 \, h^{-1}\text{Mpc}$ from the filament spine, the density has dropped by a factor of ~12 (depending on proximity to the node), which will be difficult to verify observationally. In summary, the choice of a constant and unique thickness for all filaments remains a simplification and does not fully capture the variation in filaments, but it considers the majority of true filament galaxies that experience a significant increase of gas density while keeping the contamination at bay.

2.3 Group identification

Our group identification is motivated by observations and the overall objective to identify group members that can experience pre-processing. Finding groups in observations is a challenging problem, because, while groups comprise all gravitationally bound galaxies residing in a dark matter halo, often only the brightest (usually central) galaxy or galaxies can be detected due to the survey's magnitude limit. Background and foreground objects and redshift space distortions lead to high false positive rates. In that case, one might choose to first identify bright group galaxies based on their spectroscopic or line-of-sight velocity data. Then, an excess of fainter galaxies in comparison to a field sample can be assigned to the group. Alternatively, a number of automated ways (geometrical, colour and model-based methods as well as probabilistic techniques) to identify galaxy agglomerations in large-scale survey observations exist, including methods like the Dressler-Shectman tests (DS; Dressler & Shectman 1988), halo-based group finders (e.g. Yang et al. 2005, using halo occupation statistics), Voronoi-Delaunay Method (Marinoni et al. 2002) and Friends-of-Friends algorithm (Geller & Huchra 1983) or through X-ray observations that bypass the uncertainty from small numbers of luminous galaxies in groups. Each recipe to find group members comes with benefits and drawbacks, and fair comparisons are understandably challenging. If spectroscopic data is available, groups in and around clusters specifically have often been identified using positions and velocities (e.g. Eke et al. 2004; Lisker et al. 2018; Iodice et al. 2019). The aim is to select galaxies that most likely represent the true bound structures, however, science-specific considerations (e.g. completeness versus purity) will control choices.

Similar to this idea, we define group galaxies in THE THREEHUNDRED simulations by first locating group centre haloes outside of $1R_{200}$ and within $5R_{200}$ of the central cluster. These are haloes with velocity dispersion $\sigma_v > 300 \, h^{-1}\text{km/s}$ and mimic the most luminous central galaxy of the group. For reference, this is slightly higher than the median velocity dispersion of groups in the Two-degree Field Galaxy Redshift Survey (2dFGRS, Eke et al. 2004). Then, we identify group members as all mock galaxies (i.e. haloes above $\sim3 \times 10^9 \, h^{-1}\text{M}_\odot$) within $1R_{200}$ of this central halo. An additional criterion based on the distance to the group centre assures that each halo can only be a member of one group. Note that by lowering the velocity dispersion threshold to $150 \, h^{-1}\text{km/s}$, many more galaxies in groups can be identified, and the fraction of galaxies that are

---

[3]In Kuchner et al. (2020), we discussed how halo mass limits compare to stellar masses expected for upcoming WWFCS observations that motivate this choice. In the present paper, we continue to use halo masses.

members of groups almost doubles. However, we prefer to select group members with a higher probability to have been affected by the high density environment, i.e. galaxies that have experienced pre-processing. This is because close to clusters, infalling groups can easily sweep up field galaxies and grow quickly (Vijayaraghavan & Ricker 2013). Therefore, while larger fractions of galaxies may enter the cluster through lose groups or pairs, most have only had a brief pre-processing period (Han et al. 2018). Our approach does not exclude very rich sub-structures that could be considered as discrete clusters (see Section 3.1.1). We do not impose a sharp divide between a group and a cluster since observationally, numbers of members depend on further quantities like magnitude- or volume limits. Furthermore, group definitions span a variety of properties like size and richness, with a wide range of velocities and morphologies of its members.

Fig. 1 shows an example of groups highlighted in red identified in one cluster volume. The insert prints the number of group members for each group: the largest group found in this volume has 90 galaxies, the smallest 5. The figure also highlights filaments in black (Section 2.2) with associated filament galaxies in dark grey and backsplash galaxies (see next Section 2.4) in yellow. Because filaments can be understood as ridges that connect maxima (nodes) in the density field of the galaxy distribution, we see filaments linking groups and clusters. Therefore, most group members will be part of the filament network, located in the cores of filaments. Likewise, filament galaxies, as defined by their distance to the skeleton ($D_{\rm skel} < 1\ h^{-1}$Mpc), can be group members. The simplification of a fixed filament width also means that some group members of massive groups will be located further than $1\ h^{-1}$Mpc from the filament spine.

2.4 Backsplash galaxy identification

In the most general terms, backsplash galaxies are galaxies that are observed outside R_{200} of the cluster, but have been inside of the cluster previously (Gill et al. 2005; Bahé et al. 2013). As a result, these galaxies have likely undergone significant disruption. They can be either departing (leaving) the cluster after its passage through, or they area on a subsequent infall (returning). This definition does not assume that the galaxy is bound to the cluster halo and does not include the location of the galaxy outside the cluster. Note that this definition is not unique.[4] While they are found in the same location as infalling galaxies, and are thus only distinguishable through kinematics for the observer (Gill et al. 2005; Pimbblet 2010), backsplash galaxies have been affected by the cluster environment itself.

In THE THREEHUNDRED simulations, we identify backsplash galaxies based on the orbital history of each galaxy relative to R_{200}. The backsplash galaxy population consists of all galaxies with a distance to the cluster center at $z = 0$ of $D_{z=0} > R_{200}$ and a minimum distance to the cluster centre at any time in their history $D_{min} < R_{200}$. For an analysis that includes backsplash galaxies – which require knowledge of previous snapshots $z > z_0$ – we use a subsample of 257 clusters. Briefly, clusters and their backsplash population are excluded from the sample in cases when the main branch cannot be tracked back to before $z = 0.5$ and when large apparent jumps in the position of the cluster merit a judgement on $D_{z=0} > R_{200}$ unreliable (see Haggar et al. 2020, for details). Fig. 1 shows backsplash galaxies in yellow: their distribution forms a cloud around the clusters' R_{200}. Note that affiliation to the backsplash

---

[4]E.g., Haines et al. (2015) consider all galaxies on their outward radial velocity past pericenter as backsplash galaxies.

population and group membership are not exclusive. Backsplash galaxies can be part of groups, however with only 9 per cent of backsplash galaxies in groups, this is relatively rare.

3 RESULTS AND DISCUSSION

3.1 The importance of galaxies in groups and filaments surrounding clusters

Galaxies that are part of groups and filaments prior to the final cluster environment may have been environmentally affected, i.e. pre-processed. Because of their sufficiently high densities, but lower velocity dispersions (and therefore higher dynamical friction force), transformation and merging occur more frequently in groups than in clusters. Therefore, in order to understand the role of pre-processing, it is important to know how many infalling galaxies are part of groups. Furthermore, groups are usually part of the wider filamentary network, as they represent maxima in the density field. Thus, most group galaxies are also filament galaxies.

We therefore ask: are the filaments feeding galaxy clusters dominated by distinct infalling groups, or are they largely coherent streams of individual galaxies? And does the fraction of group galaxies and filament galaxies vary with the mass or dynamical state of the cluster? Fig. 2 shows the fractions of mock galaxies outside the cluster's R_{200} and inside $5 \times R_{200}$ in groups and in filaments (i.e. galaxies with a distance to the skeleton of $D_{\rm skel} < 1\ h^{-1}$Mpc) as a function of cluster mass (Fig. 2A) and relaxedness (Fig. 2B). Each point represents the fraction in one cluster, the bands indicate the means of the point distributions and corresponding 1σ errors.

While unrelaxed clusters have accreted large amounts of material (including through groups) in their recent history, they have also rapidly grown their R_{200} as a consequence. The fraction of cluster mass in subhalos inside R_{200} at present day is high, but we do not see evidence that the fraction of galaxies in filaments (closer than $1\ h^{-1}$Mpc) and groups outside R_{200} and within $5R_{200}$ of the cluster is higher in unrelaxed clusters (Fig. 2B). Independent of masses and dynamical status, approximately 10 per cent of all mock galaxies outside $1R_{200}$ can be found in groups and roughly 45 per cent (30 per cent) of all mock galaxies outside $1R_{200}$ are in filaments where filament thickness is defined as $D_{\rm skel} < 1\ h^{-1}$Mpc ($D_{\rm skel} < 0.7\ h^{-1}$Mpc, see Section 2.2). Most group host haloes are located in filaments: 93 per cent of all group centres are located within 1Mpc/h of filament spines. Note that this reduces to 77 per cent for a more restricted filament thickness definition of 0.7Mpc/h. It is not surprising that most groups are part of filaments given DISPERSE identifies filaments by connecting maxima in the density field – an *a-posteriori* confirmation of the filament extraction. Nevertheless, this could be an important consideration for pre-processing studies since group galaxies in filaments have been shown to experience increased pre-processing compared to group galaxies outside filaments (Poudel et al. 2017).

The low fraction of galaxies in groups may at first appear in tension with recent observational studies that typically report higher fractions (e.g. McGee et al. 2009; Dressler et al. 2013; Cybulski et al. 2014). We caution that a comparison is not straightforward given the differences in defining groups and mass thresholds. As discussed earlier, our cautiously identified group members represent galaxies that have spent a significant amount of time as part of groups and are thus likely to be environmentally effected by the group. Han et al. (2018) found that only ∼12 per cent of cluster members have spent more than 4 Gyr in a group and have therefore had enough time to quench (satellite star formation rates evolve unaffected for

(a) dependence on cluster mass

(b) dependence on dynamical state of the cluster
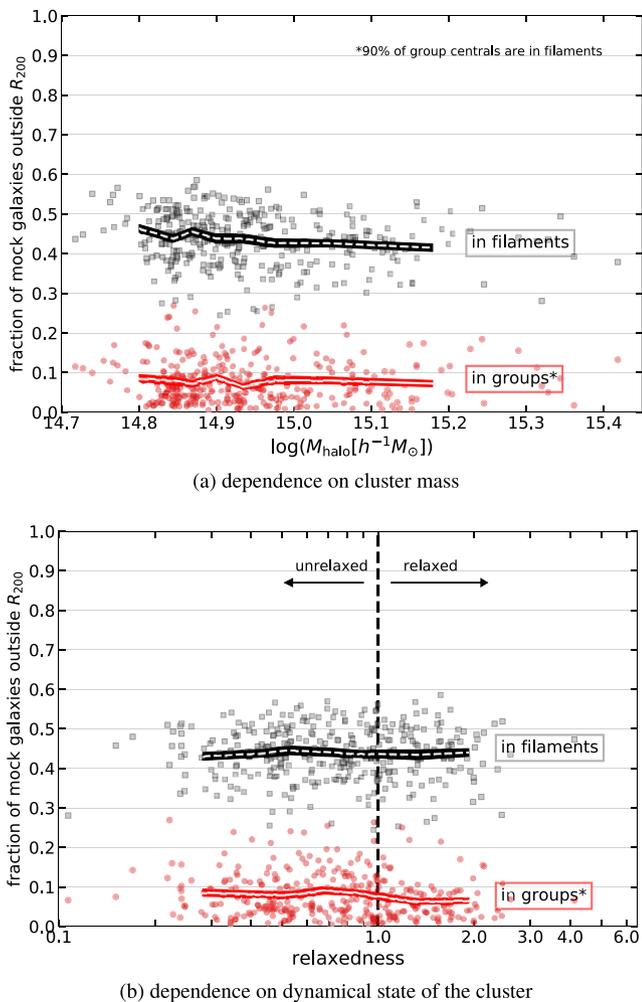
Figure 2. The fraction of all galaxies in filaments (black band, dashed line, defined as haloes with distance to filament axes $D_{\rm skel} < 1\ h^{-1}{\rm Mpc}$) and in groups (red band, solid line) outside R_{200} does not depend on mass (top) or dynamical state (bottom) of the cluster. About 45 per cent of all galaxies down to $M_* > 10^9 M_\odot$ are in filaments, and ~10 per cent of all galaxies are in the groups. The number of filament galaxies depends on the choice of filament thickness. Here we consider filament cores with a constant radius of $1\ h^{-1}{\rm Mpc}$. The fractions are not exclusive: 90 per cent of central group haloes are part of filaments and therefore a large fraction of group galaxies are also in filaments. Unrelaxed clusters are roughly defined as cluster with relaxedness $R < 1$. Coloured bands are 1σ errors on the mean.

2–4 Gyr after infall, Wetzel et al. 2013). Many more galaxies spend only a limited amount of time (half of the galaxies spent less than 2.5 Gyr) in the host before joining the cluster population. Note also that observational analyses are complicated by high contamination rates, especially in the most typical groups that only host a few galaxies. It is important to keep in mind that observationally defined groups may include an additional 40 per cent of interlopers as group members (Eke et al. 2004).

Just like clusters, groups grow over time by merging and accreting members from their surroundings – in most cases, this will be from the filament environment. We therefore investigate whether the fraction of group galaxies changes as a function of distance to the cluster centre. An increase could imply that even our cautious selection overestimates the fraction of group galaxies that have had enough time to be efficiently pre-processed, e.g. quenched as group satellites during infall. Fig. 3 shows the fraction of galaxies

in filaments (solid line, black error band), in groups (dashed line, red error band) and backsplash galaxies (dotted line, yellow error band) as a function of distance to the cluster centre. Fractions are calculated in 30 shells out to $5R_{200}$ surrounding the cluster. We do not show fractions inside R_{200}, because at very small distances the volume of filaments quickly encompasses the entire volume, and fractions become meaningless. The red dot-dashed line in Fig. 3 shows that the fraction of galaxies in groups remains constant with distance. Similarly, we found that the average richness of groups stays constant as a function of distance. Richness is defined as the number of cluster members, i.e. all galaxies within R_{200} of the group host. This is independent of whether they are located within filaments or outside of filaments. While this may seem in contrast to observations that report that groups in filaments have more satellites than outside of filaments (Guo, Tempel & Libeskind 2015), we again point towards differences (and difficulties) in defining groups consistently in simulations and observations and refer to our reasoning and choices (Section 2.3). The constant fraction of galaxies in groups suggests that we indeed capture galaxies that have had a chance to pre-process.

The black band (solid line) in Fig. 3 shows the *fraction of all galaxies* in filaments as a function of distance, the grey line is the volume-corrected fraction. Because the galaxy density and the relative volume of filaments increases towards clusters,[5] the measured fraction of filament galaxies naturally increases. This can be seen by the upturn of the black solid line at smaller distances to the cluster. We reproduce and correct for this by calculating and subtracting the fraction of galaxies in randomly placed filament networks, which is shown by the dot-dashed line and light grey error band, i.e. for each cluster we calculate the fraction of galaxies in a network from another random cluster. The resulting volume-corrected fraction is shown in the solid dark grey line. For completion, we note that we have tested randomized orientations of the same cluster network as well as networks of a random different cluster for this correction. While results are not identical, both are valid ways to demonstrate the volume correction and differences are at the level of 10 per cent at small distances to the cluster centre. The correction removes the increase of galaxies towards the cluster centre and flattens the curve – a slight divergence from our results based on reference networks extracted from the underlying gas distribution discussed in Kuchner et al. (2020). In this 'best case scenario' of gas filaments, we had found a small increase of galaxies in filaments closer to clusters (by about 8 per cent). However, given our choices for filament extraction, we found that it was most challenging to correctly identify filaments very close to clusters. As a consequence, such small effects may not have been picked up.

3.1.1 Scatter on the extremes

THE THREEHUNDRED simulations include 56 cluster volumes with very rich infalling groups of more than 150 members. These large groups can be treated as cluster-like systems with their own filament networks which will eventually merge with the more massive cluster. In Kuchner et al. (2020) we have shown that these second most massive haloes (SMH) are connected to the central clusters with thick bridges, as has also been described in numerous observations (e.g. Durret et al. 2008; Tanimura et al. 2019; Umehata et al. 2019; Reiprich et al. 2021).

---

[5]Depending on the mass of the cluster, between 20 and 30 per cent of the volume immediately outside R_{200} is taken up by filaments (calculated in a shell of 100 kpc thickness).
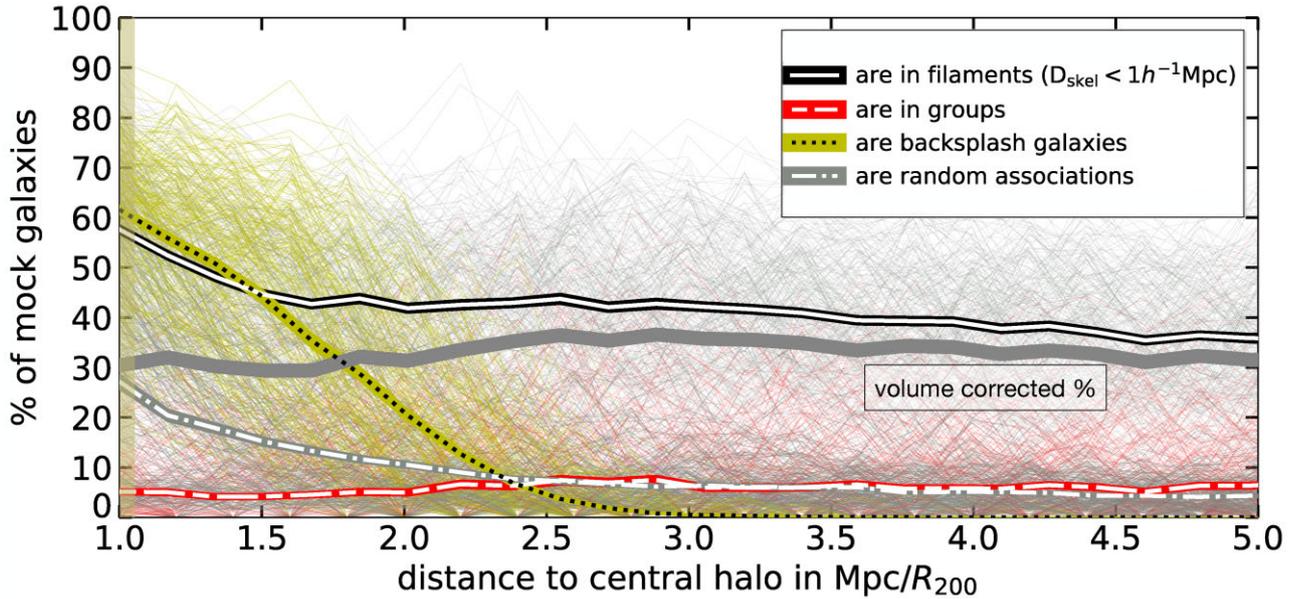
Figure 3. The fraction of galaxies in groups and filaments does not depend on the cluster distance, backsplash galaxies increase steeply closer to the cluster. Shown are percentages of mock galaxies in filaments (black band, solid line), groups (red band, dashed line) and that are backsplash (yellow band, dotted line) as a function of distance to the cluster centre (normalized by R_{200}). Note that the number of filament galaxies is for a characteristic filament thickness of 1Mpc ($D_{skel} < 1\,h^{-1}$Mpc) and bands are the 1σ error on the mean. The fractions are not exclusive. The thick solid dark grey line is the volume-corrected number of galaxies in filaments based on a randomly rotated networks (light grey dot-dashed line).

As a consequence, the 'contamination' of filaments with group galaxies varies strongly across the sample (Fig. 4) and we do not find a correlation of contamination with cluster properties. Overall, the contamination, i.e, the number of groups in filaments in cluster outskirts or their richness does not depend on the mass or dynamical state of the central cluster itself. The example figure shows fractions of galaxies in filaments and in groups in three clusters of the sample: the left panel shows a system where almost all filament galaxies are pristine filament galaxies. The contamination of filament galaxies that are in groups are shown in grey in the lower panel. We can identify two areas (at distance ~ 2 and $\sim 5R_{200}$) with groups. At these distances, ~ 20 per cent of the filament galaxies are in groups. The example in the middle shows one larger group embedded in a rich filament. The example on the right highlights a complex system with two large groups (akin to lower-mass clusters) that will merge with the cluster in the future: at two separate distances from the cluster centre, groups dominate the filaments and therefore around half of all filament galaxies in this system are found in groups.

3.2 Backsplash galaxies in filaments

It is challenging to unambiguously identify individual backsplash galaxies in observations, i.e. galaxies whose orbital trajectories have taken them through and out of the cluster after first or second infall. Usually, a variety of signatures need to add up: their gas morphologies could be altered due to ram pressure stripping (Haynes, Giovanelli & Chincarini 1984; Abramson et al. 2011; Jaffé et al. 2015) and their stellar masses are lower due to tidal stripping (Poggianti et al. 2017; Ramatsoku et al. 2019). Further, stellar spectra might indicate post-starburst signatures (Paccagnella et al. 2017; Kelkar et al. 2019). In addition, backsplash galaxies show a stronger radial alignment than infalling galaxies (Knebe et al. 2020). But more commonly, backsplash galaxies are identified in phase-space diagrams through their positions and velocities. Statistically, they have recession

velocities comparable to that of the cluster and are found in its immediate vicinity. However, we do not yet know how backsplash galaxies relate to filaments feeding clusters, i.e. if they have a preferential location with respect to filaments. Knowing whether they preferentially lie inside or outside of filaments could help to identify them. In addition, knowing how many backsplash galaxies are in filaments creates awareness that some observational signatures of galaxies in filaments (that possibly look like evidence of group- or filamentary pre-processing) may in fact be due to the galaxy's past environmental history of having gone through the cluster.

Close to the cluster, backsplash galaxies become an increasingly important ingredient of the galaxy population mix, which can be appreciated by looking back to Fig. 3 where the backsplash population is denoted by the dotted line and yellow error band. The average fraction of backsplash galaxies rises to $\sim 65\%$ close to the cluster and is virtually absent outside of $\sim 2.5\,R_{200}$. Note, however, that backsplash galaxies extend far beyond the typical virial radius of a cluster ($\sim 1.5R_{200}$) and extend to the splashback radius,[6] beyond which material is not expected to be virialized. We further investigate the positions and paths of backsplash galaxies in Fig. 5. It shows the distribution of a representative sample of backsplash galaxies at redshift $z = 0$ around clusters (indicated by the black circle), relative to the position at which they first entered the cluster, and their trajectories, for relaxed (top) and unrelaxed (bottom) clusters separately. We produced this plot by rotating the path taken through the cluster by each backsplash galaxy, so that each galaxy is on the x-axis (y=z = 0) in their last snapshot before entering the cluster. We also rotated the paths such that the motion in the z-direction is minimized, and hence the galaxy paths are (approximately) in the

---

[6]R_{sp} is a physically motivated definition of the halo boundary where particles reach the apocenter of their first orbit; typically in the range [1,2.5]R_{vir} (Diemer & Kravtsov 2014).

(a) low contamination   (b) medium contamination   (c) high contamination
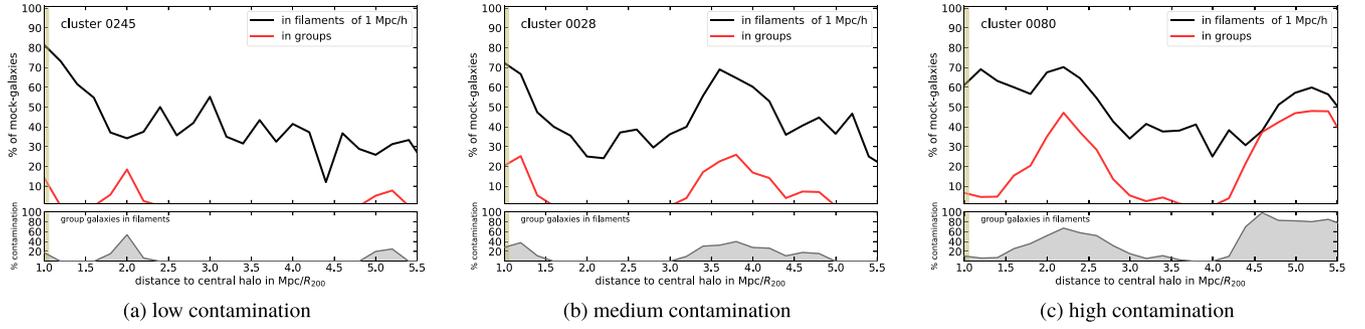
Figure 4. Three examples of varying contamination of group galaxies in filaments of $1\ h^{-1}$Mpc thickness feeding galaxy clusters. It shows distances to the cluster centre versus the fraction of galaxies in filaments (in black) and galaxies in groups (in red). The peaks in the percentage of mock galaxies indicate the positions of the groups and also shows that the groups are part of filaments. The extreme example on the right is a complex system including a cluster-sized group of 180 members.

(a) trajectory of backsplash galaxies in relaxed clusters

(b) trajectory of backsplash galaxies in unrelaxed clusters
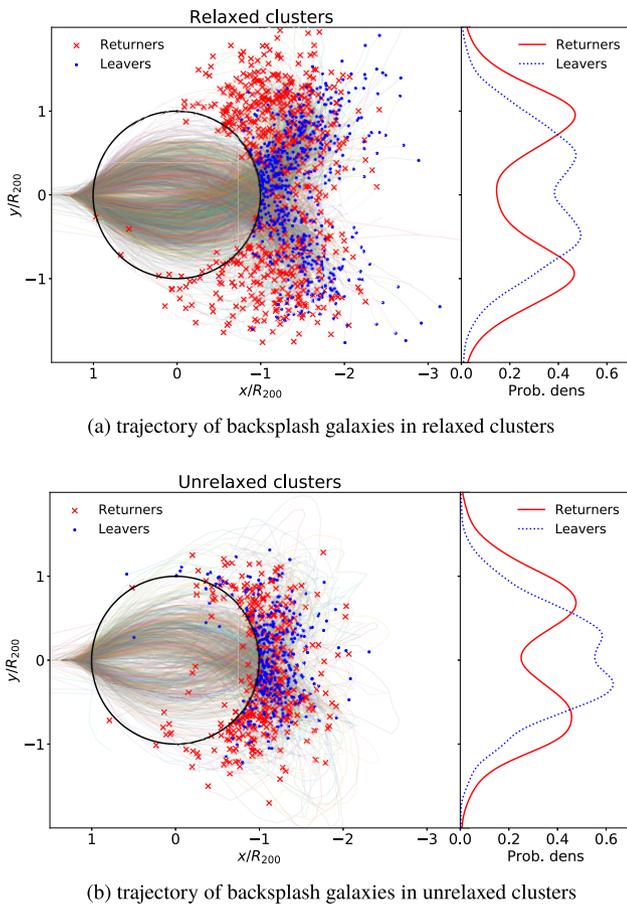
Figure 5. Galaxies deflect from a straight line on their way through and out of the cluster. They leave and return to the cluster in a wide cone. Shown are traces of backsplash galaxies in relaxed (top) and unrelaxed (bottom) clusters at $z = 0$. 'Leavers' are marked with blue dots and dotted line; they have not yet reached apocenter. 'Returners', in red crosses and solid line, are on their next infall to the cluster. For clarity, points and tracks are shown for three clusters each, representative of the whole sample. The kernel density estimation on the right hand panel shows the degree of deviation from a straight line through the cluster for returners and leavers.

plane of the page. Backsplash galaxies leave a cluster typically after ~ 2 Gyr opposite the location where they entered and build a 'cone' of opening angle 23^{+14}_{-12} degrees in relaxed clusters and s21^{+17}_{-11} degrees in unrelaxed clusters. They return in a wider cone of 41^{+20}_{-16} degrees in re-

laxed clusters and 35^{+22}_{-16} degrees in unrelaxed clusters (see also Knebe et al. 2004). The angles are the median values and 1-sigma spread, where 0 degrees corresponds to a galaxy that has passed straight through a cluster with no deviation ($y = 0$ in Fig. 5). For both relaxed and unrelaxed clusters, the returners are slightly more deflected.

Following this picture, it is easy to imagine that if a galaxy falls in through a filament, chances are high that that the backsplash galaxy will collapse on to a filament on the other side of the cluster and thus return as part of filaments. This is because filaments are not randomly positioned either: they preferentially follow the semi-major axis of the main halo or connect to their second most massive halo as a bridge (Kuchner et al. 2020). In addition, so far we cannot rule out that backsplash galaxies (help to) *form* a filament when they return to the cluster. To investigate this, we divide backsplash galaxies by their velocities into leavers and returners, i.e, galaxies that have gone through the cluster and are moving away from it in the final snapshot (either for the first or second time) are labeled as 'leavers' (blue dots and dotted line in Fig. 5) and galaxies that have gone out, turned around and are approaching the cluster (either for the first or second time) are labeled 'returners' (red 'x' and solid line). In other words, 'leavers' have left the cluster but have not yet reached the apocentre of their orbit, while 'returners' have passed apocentre, and are now on a second or further infall towards the cluster. Note that all 257 clusters were used in this analysis (see Section 2.4), however for clarity we only show the paths and final positions of galaxies in three clusters, which are representative of the larger sample.

Galaxies that have gone through the cluster are likely to have been deflected from the central axis they each start from, rather than pass straight through. This can be inferred from the double-peaked probability distributions of the y-coordinate of galaxy positions at redshift $z = 0$ in the right-hand panels of Fig. 5. It represents the distance each galaxy has deviated from a straight path through the cluster and can be seen in both relaxed and unrelaxed clusters.[7] Following the expectation that clusters are embedded in a large cosmic filament extended along the major axis, we can now anticipate that backsplash galaxies deviate from a major filament. This is supported by our finding in Rost et al. (2020) where we found that gas preferentially falls into nodes inside filaments, but preferentially leaves the cluster outside filaments. Returners are more heavily deflected sideways, in both relaxed and unrelaxed clusters.

---

[7]The kernel density estimation (KDE) is made using data from all clusters, not just those shown in the left panel of the plot.

(a) fraction of galaxies and backsplash galaxies that are in filaments   (b) fraction of galaxies and filament galaxies that are backsplash
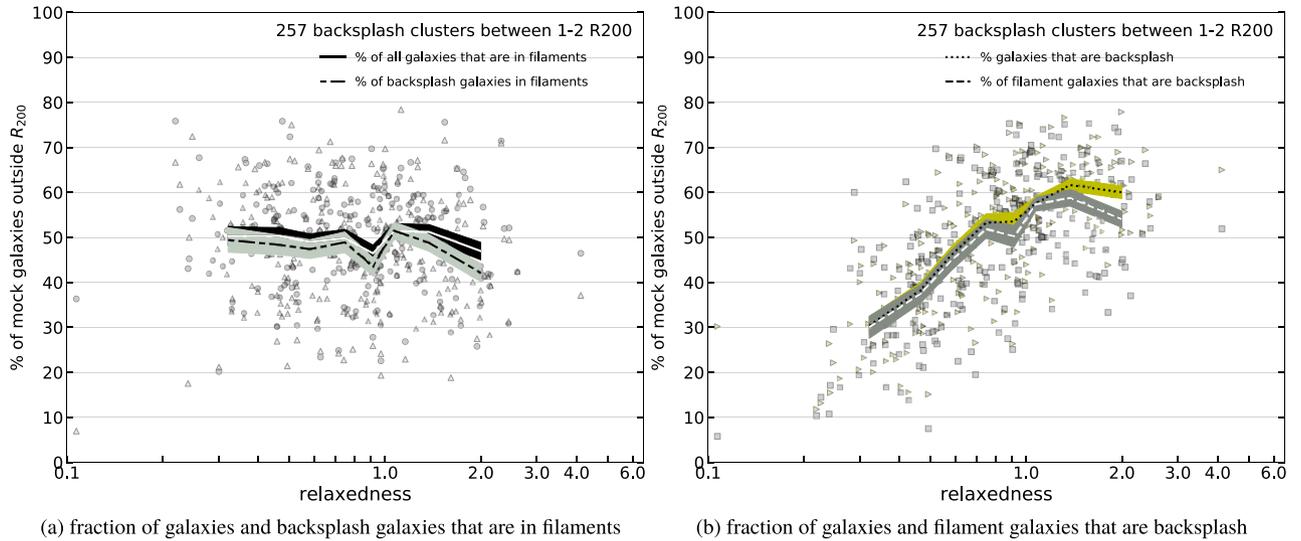
Figure 6. The fraction of backsplash galaxies in filaments of 1 h^{-1}Mpc thickness is nearly identical to the fraction of all galaxies in filaments. This is independent of the dynamical state of the cluster (left, dot dashed line, grey error band). However, the overall fraction of backsplash galaxies increases from 30 per cent in unrelaxed (R<1) to 60 per cent in relaxed (R>1) clusters (right, dotted line, yellow error band). Consequently, the fraction of backsplash galaxies in filaments increases in nearly the same way. Bands show 1 σ errors on the mean. Shown are fractions within 1 h^{-1}Mpc of filament spines. Note that this figure only reports galaxies within 1 and 2 R_{200} and a reduced sample of 257 clusters due to the requirement of continuous snapshot tracking to before $z = 0.5$.

In this analysis, every backsplash galaxy is constructed to start at the same point. In reality, backsplash galaxies enter the cluster from a number of positions around the cluster – through filaments, as groups and as isolated galaxies – and each one deflects and scatters dynamically. The many infall and therefore scatter directions add up to create a cloud of backsplash galaxies, which can be appreciated as yellow points around the cluster in Fig. 1. Importantly, this cloud of additional galaxies close to R_{200} does not influence the filament finding process. In practice, the homogeneous cloud of backsplash galaxies close to the cluster is not an important feature for DISPERSE, provided enough volume or area is available. Even in a (hypothetical) extreme case where all galaxies come in through filaments (and we know from Fig. 2 that statistically this is not the case), they leave the cluster scattered in wide cones that overlap, again smearing out to a cloud of backsplash galaxies that is very similar to the overall distribution of infalling galaxies. We thus see very little evidence that backsplash galaxies are distributed differently to infalling galaxies with respect to filaments. This is evident in Fig. 6A, which shows the percentage of all galaxies in filaments (black points, solid line, black error band) and that of backsplash galaxies in filaments (grey triangles for individual points and dot-dashed line, grey error band for the 1σ error on the mean). The two curves are nearly identical, signifying that backsplash galaxies are neither more nor less likely to re-enter the cluster through filaments than a galaxy on its first infall. We therefore see no evidence that backsplash galaxies collapse to form a filament. Filaments are stable geometrical features that do not quickly change or form. However, Fig. 5 suggests that there may be a difference between galaxies leaving and returning in relaxed and unrelaxed clusters.

3.2.1 Dependence on dynamical state of the cluster

The fraction and extent of backsplash galaxies around clusters not only varies strongly with distance to the cluster but also with dynamical state of the cluster. Fig. 5 shows that backsplash galaxies around relaxed clusters spray further than in unrelaxed clusters,

where the entire backsplash population is typically contained within $2R_{200}$. To investigate whether this resulted from the fact that our relaxed clusters have a lower average radius, we reproduced these plots, normalizing by 2 Mpc (which is approximately the average cluster radius) instead of R_{200}. These plots are not shown, but changing this normalization had very little effect on the results. Rather, this difference is due to the rapid increase of the cluster's radius following mergers that lead to unrelaxed dynamical states – faster than backsplash galaxies replenish (Haggar et al. 2020). The difference is significant: The fraction of backsplash galaxies increases from 30 per cent in unrelaxed (R<1) to 60 per cent in relaxed (R>1) clusters (dashed line and yellow error band in Fig. 6B). As a direct consequence of Fig. 6A, the fraction of backsplash galaxies in filaments rises at the same rate (dashed line, grey error band). We see some hints of a deviation in the most relaxed clusters of the sample, in the sense of a lower fraction of backsplash compared to infalling galaxies in filaments.

Leavers and returners may be more clearly separable in relaxed clusters than in unrelaxed clusters (compare the two peaks in the KDE of the right-hand panel Fig. 5), but we do not see a dependence on dynamical state of the cluster in relation to filaments: backsplash galaxies are distributed in the same way with respect to filaments, whether they are in relaxed or unrelaxed clusters and whether they are leaving or returning to the cluster (explained by the picture of a homogenous cloud of backsplash galaxies due to the scatter dynamics of galaxies passing through the cluster, as discussed in the previous section). Fig. 7 underpins this uniformity. It shows PDFs of measured distances from leavers and returners to filament spines in relaxed (top) and unrelaxed (bottom) clusters. Clearly, there is no difference between returners and leavers and also no difference in relaxed and unrelaxed clusters, with the small exception of very unrelaxed clusters (insert in Fig. 7). In a sample of the most unrelaxed clusters with relaxedness parameters R<0.3, returners came back to the clusters significantly closer to filament spines than they had left the cluster. This may indicate some memory of a merger, where in unrelaxed clusters the preferential direction of velocities before the merger is retained, and in relaxed clusters this axis was lost when
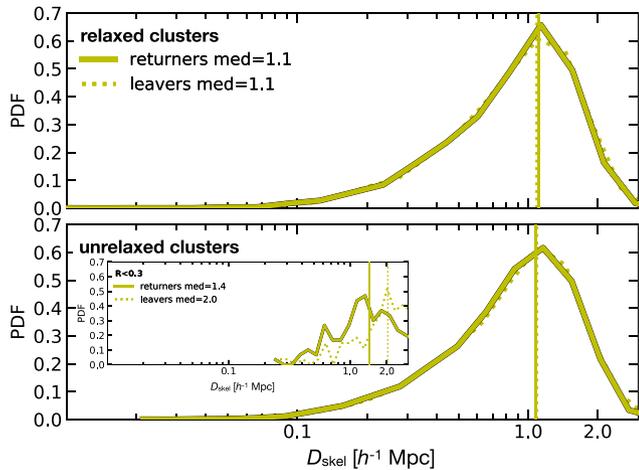
Figure 7. Despite the different distributions of backsplash galaxies in relaxed and unrelaxed clusters, they are distributed in the same way in filaments around relaxed and unrelaxed clusters. Because they are distributed close to the R200, a lot of the volume is actually made up of filaments. The yellow curve and the dot-dashed curve rise in similar ways at least inside 1.5R200, where backsplash galaxies are dominant.

velocities randomized. For the majority of clusters, however, we see no dependence of the location of backsplash galaxies in relaxed and unrelaxed and in leavers and returners with respect to filaments. Note that the area where backsplash galaxies prevail is a turbulent region characterized by accretion shocks where the infalling gas is significantly slowed down and heated while becoming part of the intracluster medium. The turbulence close to the cluster induced by the mixing of material that collapses towards filaments, as well as into the cluster, and gas shocks triggered by substructures is further complicated by signatures of backsplash galaxies. We described these complex gas velocity fields close to THE THREEHUNDRED clusters in Rost et al. (2020).

Finally, some backsplash galaxies may be in groups if they have fallen in as part of groups and leave R_{200} still maintaining group membership. Following the trajectories of groups in THE THREEHUNDRED, we found that backsplash groups are relatively rare since infalling groups lose the majority of their members inside R_{200} of the cluster. With our group definition, only 9 per cent of backsplash galaxies are members of a group at cluster infall.

4 CONCLUSIONS: HETEROGENEOUS FILAMENT ENVIRONMENTS

Cosmic filaments that feed clusters host galaxies with diverse star formation histories. The galaxies may get affected by their current environment, either denser large-scale filaments and/or groups that are part of the filament network ('pre-processed'). This could lead to measurable changes, e.g. of the gas content, star-formation activity and galaxy morphology. Others may have been processed in the past by the galaxy cluster during their first infall. The galaxy mixture strongly depends on the distance from the cluster core and dynamical state of the cluster. Understanding the constituents of galaxy cluster outskirts as a combination of different environments, where the important environment of filaments themselves are heterogeneous, helps to better understand the nature and relative importance of environmental processes on galaxy mass assembly and quenching. Fig. 8 summarizes this non-uniform environment and shows an inventory of galaxies in filaments around simulated THE THREE-
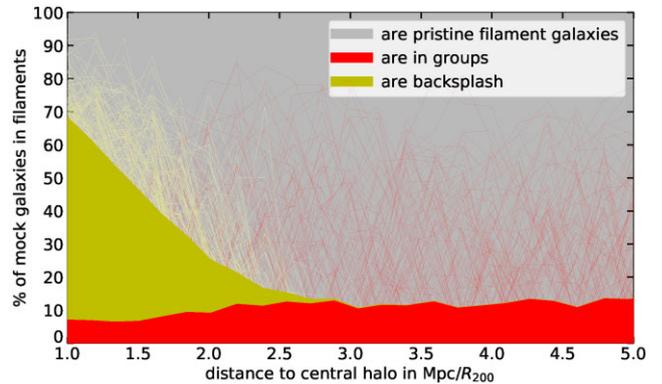
Figure 8. As a way to compare total values of filament members across distances from the cluster ($R_{200} = 1.0$), we plot a stacked histogram. It shows the expected fractions of galaxies in each environment within filaments of constant thickness as a function of distance to the cluster centre normalized by R_{200}: group galaxies as defined in Section 2.3 (red area), backsplash galaxies (yellow area) and 'pristine' filament galaxies (grey area). By inference, only about 30 per cent of all galaxies that fall into the cluster through filaments are 'pristine'. The thin lines highlight the large cluster-by-cluster variations.

HUNDRED clusters, a benchmark to compare observational signatures with. These numbers are based on a characteristic filament core thickness of $1\ h^{-1}$Mpc and halo masses of $M_{halo} > 3 \times 10^{10}\ h^{-1}M_\odot$ (comparable to $M_* > 3 \times 10^9\ h^{-1}M_\odot$). They change according to choices that will depend on the individual science case and emphasis on e.g. purity, completeness, accuracy or precision.[8] Importantly, fractions do not depend on cluster halo mass. The figure summarizes the composition of filaments feeding clusters as a function of distance to the cluster centre. From it we conclude the following:

(i) **Group galaxies:** 12 per cent of all filament galaxies in cluster outskirts (between 1 and 5 R_{200} of the cluster) are located in groups in filaments, where we expect pre-processing by group environments. This number is highly dependent on the exact definition of group membership (Section 2.3). In the context of THE THREEHUNDRED and keeping observational challenges and goals of pre-processing studies in mind, we define groups as galaxies within $1R_{200}$ of a halo with $\sigma_v > 300\ h^{-1}$km/s. This likely captures the correct number of galaxies that has spent a significant time (longer than 4 Gyrs) in groups (Han et al. 2018). The fraction of galaxies in groups doubles when this criterion is lowered to $150h^{-1}$km/s. 90 per cent of group hosts are located in filaments, owing to a large extend to the fact that they mark maxima in the galaxy distribution which are used to construct the filament network (see Section 2.2). While there is considerable cluster-to-cluster variation (Fig. 4), on average the fraction of group galaxies in filaments remains constant with distance from the cluster.

(ii) **Backsplash galaxies:** close to the cluster centre, between 30 per cent (in unrelaxed) and 60 per cent (in relaxed clusters) of all galaxies are members of the backsplash populations (Fig. 6B), i.e. they have been processed by the cluster. This number is highly dependent on the dynamical state of the cluster and distance to the cluster centre: we find more backsplash galaxies in relaxed clusters and close to R_{200}. The number drops sharply with increasing distance and we find no backsplash galaxies beyond 2.5R_{200}. The increasing prevalence of backsplash galaxies around clusters make it

---

[8]We refer the reader to Kuchner et al. (2020) for a detailed overview of how these choices may bias expectations.

% of galaxies in cluster outskirts

% of galaxies entering a cluster via filaments

non-filaments 55%

filaments (1Mpc) 45%

groups 12%

pristine 28 — 58%

backsplash 30 — 60% *

* depends on dynamical state of the cluster; some backsplash galaxies are in groups
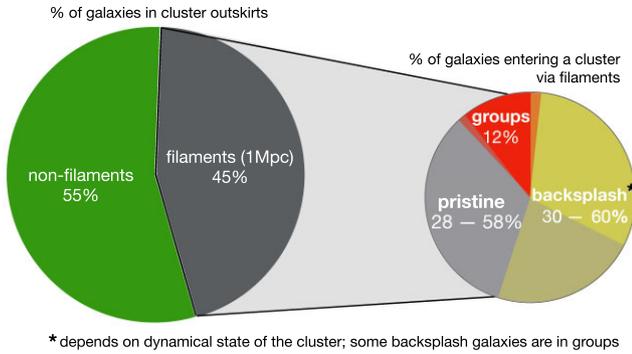
Figure 9. Up to 45 per cent of galaxies accreted by clusters are closer than $1\,h^{-1}$Mpc to a filament spine, which we define as being 'inside filaments' (see text for a discussion on choosing an optimal filament thickness). Filaments themselves are heterogeneous environments that host groups, and backsplash galaxies alongside galaxies that have been environmentally effected by the cosmic filament alone. The pie chart on the right details the breakdown of galaxies in different environments inside filaments at R_{200}. The number of backsplash galaxies is highly dependant on the cluster's dynamical state.

challenging to disentangle the post-processing effects of clusters and the pre-processing of cosmic web environments. Backsplash galaxies are deflected on their pass through the cluster and scatter on their way out which produces a cloud of backsplash galaxies around the cluster. Therefore, they have no preferred location with respect to filaments, i.e. they are not more likely to fall back on to clusters through filaments (Fig. 6A) – neither in relaxed nor unrelaxed clusters (Fig. 7).

(iii) **Pristine filament galaxies**: The remaining $\sim 33\%$ of galaxies in filaments at cluster R_{200} are 'pristine filament galaxies'. These are galaxies entering a cluster via coherent streams of individual galaxies. Importantly, this scenario assumes filaments of constant thickness – a simplification, since filaments are likely growing thicker closer to massive nodes with an increase of galaxies in filaments (see Kuchner et al. (2020) for a discussion).

The complex cluster outskirt physics make the reconstruction of environmental histories of galaxies falling into clusters not only challenging but dependent on factors such as the dynamical state of the cluster or the distance to the cluster centre. Measurements are challenging since this is a regime where the infall, merging and virialization of matter intertwine. Near clusters, accretion shocks and backsplash galaxies dominate and complicate the velocities of galaxies and measurements of their host environments. Further out, galaxy groups and large-scale filaments of the cosmic web may take over. Each relate to specific environmental mechanisms and thus influence expectations for observational evidence of pre-processing (observed effects due to increased densities) in galaxies around clusters. The results presented in this paper demonstrate a statistical breakdown of galaxies in cluster outskirt environments, emphasizing the variety of environments and environmental histories galaxies in filaments can have and typical journeys of galaxies before falling into clusters. Groups and filaments are the instantaneous environment we find galaxies in, backsplash galaxies contain a record of where the galaxies have been in the past. In addition, these are not absolutes: some backsplash galaxies are in filaments, some are in groups, some are in the remaining 'field' around the cluster. In summary, while up to 45 per cent of all galaxies fall into clusters via filaments (closer than $1\,h^{-1}$Mpc from the extracted filament spine), filaments themselves are heterogeneous environments that host groups and backsplash galaxies, alongside a minority of galaxies that have been environmentally effected by the cosmic filament alone (Fig. 9).

DATA AVAILABILITY

Data available on request to THE THREEHUNDRED collaboration: https://www.the300-project.org.

REFERENCES

Abramson A., Kenney J. D. P., Crowl H. H., Chung A., van Gorkom J. H., Vollmer B., Schiminovich D., 2011, AJ, 141, 164
Ade P. A. R. et al., 2016, A&A, 594, A13
Alpaslan M. et al., 2016, MNRAS, 457, 2287
Arzoumanian D. et al., 2019, A&A, 621, A42
Bahé Y. M., McCarthy I. G., 2014, MNRAS, 447, 969
Bahé Y. M., McCarthy I. G., Balogh M. L., Font A. S., 2013, MNRAS, 430, 3017
Balcells M. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, Proc.eedings of the SPIE, Ground-based and Airborne Instrumentation for Astronomy III, Vol. 7735. SPIE, p. 77357G
Balogh M. L., Morris S. L., Yee H. K. C., Carlberg R. G., Ellingson E., 1997, ApJ, 488, L75
Balogh M. L., Navarro J. F., Morris S. L., 2000, ApJ, 540, 113
Beck A. M. et al., 2015, MNRAS, 455, 2110
Benavides J. A., Sales L. V., Abadi M. G., 2020, MNRAS, 498, 3852
Benítez-Llambay A., Navarro J. F., Abadi M. G., Gottlöber S., Yepes G., Hoffman Y., Steinmetz M., 2013, ApJ, 763, L41
Berrier J. C., Stewart K. R., Bullock J. S., Purcell C. W., Barton E. J., Wechsler R. H., 2008, ApJ, 690, 1292
Bianconi M., Smith G. P., Haines C. P., McGee S. L., Finoguenov A., Egami E., 2017, MNRAS, 473, L79
Blanton M. R., Moustakas J., 2009, ARA&A, 47, 159
Cautun M., van de Weygaert R., Jones B. J. T., Frenk C. S., 2014, Proceedings of the International Astronomical Union, 11, 47
Colberg J. M., Krughoff K. S., Connolly A. J., 2005, MNRAS, 359, 272
Cui W. et al., 2018, MNRAS, 480, 2898
Cui W., Knebe A., Yepes G., Yang X., Borgani S., Kang X., Power C., Staveley-Smith L., 2017, MNRAS, 473, 68

Cybulski R., Yun M. S., Fazio G. G., Gutermuth R. A., 2014, MNRAS, 439, 3564

Dalton G. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proceedings of the SPIE, Ground-based and Airborne Instrumentation for Astronomy IV, Vol. 8446. SPIE, p. 84460P

Darvish B., Mobasher B., Sobral D., Hemmati S., Nayyeri H., Shivaei I., 2015, ApJ, 814, 84

Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827

De Lucia G., Weinmann S., Poggianti B. M., Aragón-Salamanca A., Zaritsky D., 2012, MNRAS, 423, 1277

Diemer B., Kravtsov A. V., 2014, ApJ, 789, 1

Dolag K., Meneghetti M., Moscardini L., Rasia E., Bonaldi A., 2006, MNRAS, 370, 656

Dressler A., 1980, ApJ, 236, 351

Dressler A., Shectman S. A., 1988, AJ, 95, 985

Dressler A., Oemler A., Jr, Poggianti B. M., Gladders M. D., Abramson L., Vulcani B., 2013, ApJ, 770, 62

Durret F., Kaastra J. S., Nevalainen J., Ohashi T., Werner N., 2008, Space Sci Rev, 134, 51

Eke V. R. et al., 2004, MNRAS, 348, 866

Fujita Y., Goto T., 2004, PASJ, 56, 621

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

Geller M. J., Huchra J. P., 1983, ApJS, 52, 61

Gill S. P. D., Knebe A., Gibson B. K., 2005, MNRAS, 356, 1327

Gray M. E. et al., 2009, MNRAS, 393, 1275

Guo Q., Tempel E., Libeskind N. I., 2015, ApJ, 800, 112

Haggar R., Gray M. E., Pearce F. R., Knebe A., Cui W., Mostoghiu R., Yepes G., 2020, MNRAS, 492, 6074

Haines C. P. et al., 2015, ApJ, 806, 101

Haines C. P. et al., 2018a, MNRAS, 477, 4931

Haines C. P. et al., 2018b, MNRAS, 481, 1055

Han S., Smith R., Choi H., Cortese L., Catinella B., Contini E., Yi S. K., 2018, ApJ, 866, 78

Haynes M. P., Giovanelli R., Chincarini G. L., 1984, ARA&A, 22, 445

Hess K. M. et al., 2018, MNRAS, 484, 2234

Iodice E. et al., 2019, A&A, 627, A136

Jaffé Y. L. et al., 2016, MNRAS, 461, 1202

Jaffé Y. L., Smith R., Candlish G. N., Poggianti B. M., Sheen Y.-K., Verheijen M. A. W., 2015, MNRAS, 448, 1715

Jung I., Lee J., Yi S. K., 2014, ApJ, 794, 74

Jung S. L., Choi H., Wong O. I., Kimm T., Chung A., Yi S. K., 2018, ApJ, 865, 156

Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, MNRAS, 353, 713

Kauffmann G., Li C., Heckman T. M., 2010, MNRAS, 409, 491

Kelkar K., Gray M. E., Aragón-Salamanca A., Rudnick G., Jaffé Y. L., Jablonka P., Moustakas J., Milvang-Jensen B., 2019, MNRAS, 486, 868

Kleiner D., Pimbblet K. A., Jones D. H., Koribalski B. S., Serra P., 2017, MNRAS, 466, 4692

Klypin A., Yepes G., Gottlöber S., Prada F., Heß S., 2016, MNRAS, 457, 4340

Knebe A. et al., 2020, MNRAS, 495, 3002

Knebe A., Gill S. P. D., Gibson B. K., Lewis G. F., Ibata R. A., Dopita M. A., 2004, ApJ, 603, 7

Kooistra R., Silva M. B., Zaroubi S., Verheijen M. A. W., Tempel E., Hess K. M., 2019, MNRAS, 490, 1415

Kraljic K. et al., 2018, MNRAS, 483, 3227

Kuchner U. et al., 2020, MNRAS, 494, 5473

Kuchner U. et al., 2021, MNRAS, 503, 2065

Kuutma T., Tamm A., Tempel E., 2017, A&A, 600, L6

Laigle C. et al., 2017, MNRAS, 474, 5437

Libeskind N. I. et al., 2017, MNRAS, 473, 1195

Lisker T., Vijayaraghavan R., Janz J., Gallagher J. S., Engler C., Urich L., 2018, ApJ, 865, 40

Mahajan S., Raychaudhury S., Pimbblet K. A., 2012, MNRAS, 427, 1252

Malavasi N. et al., 2016, MNRAS, 465, 3817

Malavasi N., Aghanim N., Tanimura H., Bonjean V., Douspis M., 2020, A&A, 634, A30

Marinoni C., Davis M., Newman J. A., Coil A. L., 2002, ApJ, 580, 122

McGee S. L., Balogh M. L., Bower R. G., Font A. S., McCarthy I. G., 2009, MNRAS, 400, 937

Oemler A., 1974, ApJ, 194, 1

Paccagnella A. et al., 2017, ApJ, 838, 148

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2010, AJ, 139, 2097

Pimbblet K. A., 2010, MNRAS, 411, 2637

Poggianti B. M. et al., 2017, ApJ, 844, 48

Poggianti B. M., Smail I., Dressler A., Couch W. J., Barger A. J., Butcher H., Ellis R. S., Augustus Oemler J., 1999, ApJ, 518, 576

Porter S. C., Raychaudhury S., Pimbblet K. A., Drinkwater M. J., MNRAS, 2008, 388, 1152

Postman M., Geller M. J., 1984, ApJ, 281, 95

Poudel A., Heinämäki P., Tempel E., Einasto M., Lietzen H., Nurmi P., 2017, A&A, 597, A86

Ramatsoku M. et al., 2019, MNRAS, 487, 4580

Rasia E. et al., 2015, ApJ, 813, L17

Reiprich T. H. et al., 2021, A&A, 647, A2

Rost A. et al., 2020, MNRAS, 502, 714

Sarron F., Adami C., Durret F., Laigle C., 2019, A&A, 632, A49

Sembolini F., Yepes G., Petris M. D., Gottlöber S., Lamagna L., Comis B., 2012, MNRAS, 429, 323

Sousbie T., 2011, MNRAS, 414, 350

Tanimura H., Aghanim N., Douspis M., Beelen A., Bonjean V., 2019, A&A, 625, A67

Tempel E., Stoica R. S., Martínez V. J., Liivamägi L. J., Castellan G., Saar E., 2014, MNRAS, 438, 3465

Umehata H. et al., 2019, Science, 366, 97

van Haarlem M., van de Weygaert R., 1993, ApJ, 418, 544

van de Weygaert R., 2014, in Shandarin S., Saar E., Einasto J., eds, IAU S308 The Zeldovich Universe: Genesis and Growth of the Cosmic Web IAU symposium proceedings series, vol. 11. Cambridge University Press (CUP)

Vijayaraghavan R., Ricker P. M., 2013, MNRAS, 435, 2713

Vulcani B. et al., 2019, MNRAS, 487, 2278

Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013, MNRAS, 432, 336

White M., Cohn J. D., Smit R., 2010, MNRAS, 408, 1818

Winkel N., Pasquali A., Kraljic K., Smith R., Gallazzi A., Jackson T. M., 2021, MNRAS, 505, 4920

Yang X., Mo H. J., Jing Y. P., van den Bosch F. C., 2005, MNRAS, 358, 217

Zabludoff A. I., Mulchaey J. S., 1998, ApJ, 496, 39

Zel'dovich Y. B., 1970, A&A, 5 84

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