Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **474,** 5437–5458 (2018)
Advance Access publication 2017 November 25

doi:10.1093/mnras/stx3055

COSMOS2015 photometric redshifts probe the impact of filaments on galaxy properties

C. Laigle,[1]★ C. Pichon,[2,3] S. Arnouts,[4] H. J. McCracken,[2] Y. Dubois,[2] J. Devriendt,[1] A. Slyz,[1] D. Le Borgne,[2] A. Benoit-Lévy,[2] Ho Seong Hwang,[5] O. Ilbert,[4] K. Kraljic,[4] N. Malavasi,[6] Changbom Park[3] and D. Vibert[4]

[1]*Sub-department of Astrophysics, University of Oxford, Keble Road, Oxford OX1 3RH*
[2]*Sorbonne Universités, UPMC Univ Paris 6 et CNRS, UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, F-75014 Paris, France*
[3]*School of Physics, Korea Institute for Advanced Study (KIAS), 85 Hoegiro, Dongdaemun-gu, Seoul 02455, Republic of Korea*
[4]*Aix Marseille Université, CNRS, Laboratoire d'Astrophysique de Marseille, UMR 7326, F-13388 Marseille, France*
[5]*Quantum Universe Center, Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemun-gu, Seoul F-02455, Republic of Korea*
[6]*Department of Physics and Astronomy, Purdue University, 525 Northwestern Avenue, West Lafayette, IN 47907, USA*

Accepted 2017 November 23. Received 2017 November 23; in original form 2017 February 25

ABSTRACT

The variation of galaxy stellar masses and colour types with the distance to projected cosmic filaments are quantified using the precise photometric redshifts of the COSMOS2015 catalogue extracted from Cosmological Evolution Survey (COSMOS) field (2 deg^2). Realistic mock catalogues are also extracted from the lightcone of the cosmological hydrodynamical simulation Horizon-AGN. They show that the photometric redshift accuracy of the observed catalogue ($\sigma_z < 0.015$ at $M_* > 10^{10}\mathrm{M}_\odot$ and $z < 0.9$) is sufficient to provide two-dimensional (2D) filaments that closely match their projected three-dimensional (3D) counterparts. Transverse stellar mass gradients are measured in projected slices of thickness 75 Mpc between $0.5 < z < 0.9$, showing that the most massive galaxies are statistically closer to their neighbouring filament. At fixed stellar mass, passive galaxies are also found closer to their filament, while active star-forming galaxies statistically lie further away. The contributions of nodes and local density are removed from these gradients to highlight the specific role played by the geometry of the filaments. We find that the measured signal does persist after this removal, clearly demonstrating that proximity to a filament is not equivalent to proximity to an overdensity. These findings are in agreement with gradients measured in both 2D and 3D in the Horizon-AGN simulation and those observed in the spectroscopic surveys VIPERS and GAMA (which both rely on the identification of 3D filaments). They are consistent with a picture in which the influence of the geometry of the large-scale environment drives anisotropic tides that impact the assembly history of galaxies, and hence their observed properties.

Key words: methods: numerical – methods: observational – techniques: photometric – galaxies: evolution – galaxies: formation – large-scale structure of Universe.

1 INTRODUCTION

A challenge for current galaxy formation models is to understand the role of the environment, its anisotropy, and its dynamics in shaping the properties of galaxies across cosmic time. Galaxies form in the gravitational potential wells of dark matter haloes. Their mass assembly can thus be approached at a first order by the gravitational history of halo assembly (Blumenthal et al. 1984; Navarro, Frenk & White 1995), while at a second order, the evolution of baryonic matter (gas accretion and cooling, star formation and feedback) induces additional levels of complexity (e.g. Hopkins et al. 2014, and reference therein). Starting from the dark matter only, the theory of cosmological structure formation predicts a strong dependence of halo mass assembly on its environment, driven by gravitational physics.

On the one hand, halo mass assembly history is encoded in the initial conditions of the matter that will end up in the halo (its so-called initial Lagrangian patch). At the simplest level, it is a function of its mass that determines its collapse time encoded in the spherical collapse model (Press & Schechter 1974). An environmental dependence will occur when also considering the impact

★ E-mail: clotilde.laigle@physics.ox.ac.uk

of the long wavelength density mode corresponding to the large-scale cosmic web. As encompassed by the peak-background split theory (Sheth & Tormen 1999a), a large-scale overdensity induces a density boost, allowing the proto-halo to pass earlier the critical threshold of collapse and biasing the mass function in the vicinity of the large-scale structure (shifted Press-Schechter, Kaiser 1984; Efstathiou et al. 1988; Cole & Kaiser 1989; Bond et al. 1991; Mo & White 1996; Sheth & Tormen 1999b): the abundance of massive haloes will thus be enhanced in overdense regions (Alonso, Eardley & Peacock 2015), and these objects will be more strongly clustered. An observational consequence of this prediction is the mass–density relation (Oemler 1974; Davis & Geller 1976; Balogh et al. 1997), where the local density at a given redshift remains a tracer of the initial density boost.

On the other hand, in complement to the local density, the initial geometric distribution of neighbouring structures will also explicitly impacts the accretion and merger history of the proto-halo via the tides it applies on to the patch. It has been found that at a given mass, denser environments are on average populated by older haloes, where the age of a halo is defined as the epoch by which the halo has assembled one half of its current mass. This effect is called 'assembly bias' (e.g. Sheth & Tormen 2004; Gao, Springel & White 2005; Wechsler et al. 2006; Croton, Gao & White 2007) and can be interpreted as follows. At a given mass, if the halo is isolated, it can continue to grow by accreting matter from its neighbourhood, via the accretion of larger shells in the initial density field. If this halo lies in the vicinity of a massive structure, it will stop to grow earlier because the matter will recede towards that massive structure. Hence, the large-scale structure around the proto-halo will affect the halo history via external tides (e.g. Dalal et al. 2008; Hahn et al. 2009; Wang et al. 2011) captured by the ellipsoidal collapse model (Bond & Myers 1996). Those tides span a large range of scales. At small scale they may induce mass losses for haloes in the vicinity of a more massive halo, while on larger scale they imply statistically distinct mass assembly histories in different large-scale environment (see also Castorina et al. 2016), e.g. by delaying or quenching the mass inflow. An important ingredient at large scale is the anisotropic structure of the matter distribution, which is organized as the cosmic web (Bond, Kofman & Pogosyan 1996), made of walls, filaments, and nodes. The shape of the tidal tensor[1] will be precisely determined by this anisotropy. It will affect not only the accretion history of the proto-halo as described above but also its dynamics via tidal torquing.

This dynamical impact of the anisotropic cosmic web has been investigated in dark matter simulations in relation to spin acquisition: Aubert, Pichon & Colombi (2004), Aragón-Calvo et al. (2007), Hahn et al. (2007), Codis et al. (2012), and Aragon-Calvo & Yang (2014) report angular momentum alignment of low-mass haloes with filaments and a perpendicular orientation for more massive haloes. This alignment can be understood when measuring the vorticity content in walls and filaments inherited from flow-crossing (Libeskind et al. 2013; Laigle et al. 2015). Structures that are forming in these large scale flows will end up with a spin preferentially aligned with that vorticity and consequently with the filament. They will accrete matter through secondary infall with a coherent rotational motion, up to a specific transition mass corresponding to the Eulerian size of the quadrant, beyond which inflow will start to advect misaligned angular momentum. This dynamical connection can also be predicted from the initial conditions, and its variation with the geometry of the environment is now well established in the context of the constrained tidal torque theory (Codis, Pichon & Pogosyan 2015b). Interestingly, this dependence still holds for galaxies, despite the scale difference and the more complex baryonic physics in the circumgalactic medium. Tempel & Libeskind (2013) have observed the alignment of galaxy's spin with their closest filament in the SDSS survey (see also e.g. Hirv et al. 2017), and Dubois et al. (2014) have first confirmed this dependence for galaxies in hydrodynamical simulations (see e.g. also Codis et al. 2015a; Chisari et al. 2015, 2016; González et al. 2017). To summarize, it is expected that the large-scale environment impacts both the mass and spin assembly history of the host haloes. Observations and simulations have shown that the dynamical dependence on the large-scale environment remains true also for galaxy spin. The key question now is: does it also affect scalar quantities such as the mass, colour, and specific star formation rate (sSFR) of galaxies?

If the effect of the tidal field on accretion is not completely erased by baryonic physics and feedback, the physical properties that depend on the accretion history should therefore be impacted by the corresponding anisotropic infall, and a specific dependence of galaxy masses, colours, and star-formation rates on the geometry of the environment should be observable. However, very few observational evidences sustain the prediction that the geometry of the environment impacts significantly galaxy accretion history.

The 'galactic conformity' (Weinmann et al. 2006) is a first measured evidence of this effect. It quantifies a correlation between the quenching of the central and the quenching of its satellite galaxies: the fraction of quiescent satellites is higher around quiescent centrals than around star-forming centrals. This effect has been detected up to redshift 2.5 for both low- and high-mass satellite galaxies (Kawinwanichakij et al. 2016). It suggests that galaxy properties depend not only on halo mass but also on the larger scales, and is likely connected to the halo assembly history. Hearin, Behroozi & van den Bosch (2016) observed this galactic conformity up to very large separation (4 Mpc): this effect is difficult to explain in semi-analytical models where galaxy properties depend essentially on halo mass. Attempting to make sense of it, they use the term of 'halo accretion conformity' and connect this effect to the mutual evolution of haloes in the same large-scale tidal field. The effects of these large scale tides are not captured by measurements of the local density.

More recently, results from spectroscopic surveys have reported the dependences of galaxy masses and types on the geometry of the environment using SDSS (Yan, Fan & White 2013; Martínez, Muriel & Coenda 2016; Chen et al. 2017; Poudel et al. 2017), GAMA (Alpaslan et al. 2016; Kraljic et al. 2018) and, at higher redshift, VIPERS (Malavasi et al. 2017), based on the three-dimensional (3D) reconstruction of the cosmic web. In particular, Malavasi et al. (2017) and Kraljic et al. (2018) find a significant trend for galaxies with different stellar masses and type to segregate near the filaments: the most massive and passive galaxies are found closer to the filament centre than the star-forming, low-mass ones. Can photometric redshift surveys unravel such signal at higher redshift?

The recently achieved accuracy of current photometric redshifts opens the prospect of working now with projected two-dimensional (2D) slices that are sufficiently thin to unambiguously study the projected filaments. One of the assets of photometric surveys is their ability to probe different epochs of cosmic evolution to leverage their relative importance in building up galaxies, with complete sample at

[1] Specifically, the traceless part of the tidal tensor, which orientation via its eigenvectors points along the main axes of the embedding cosmic web. Recall that gradients of the potential displace haloes, while the tides distort and rotate them along the directions set by those vectors.

much lower cost than spectroscopic surveys. The potential of photometric surveys may have been initially underestimated because of their insufficient precision along the line of sight, making it seemingly too difficult to reconstruct the 3D field at the required scale to trace the cosmic web. With the large and deep multiwavelength surveys such as the Cosmological Evolution Survey[2] (COSMOS; Scoville et al. 2007), the upcoming Hyper Suprime-Cam survey (HSC; Miyazaki et al. 2012b), and in the long term the Large Synoptic Survey Telescope (LSST; Ivezic et al. 2008), it is timely and interesting to estimate the requested accuracy to achieve this goal via simulations; it is also timely to exploit the latest COSMOS2015 catalogue (Laigle et al. 2016) as a prototype of these upcoming surveys to try and detect the dependences of galactic properties such as stellar mass and galaxy colours with respect to their position within the cosmic web. Previous studies on COSMOS have investigated the effect of local environment on galaxy properties, either relying on isotropic density estimators (e.g. Scoville et al. 2013; Darvish et al. 2017), or dividing the field in morphological components (field, filament or cluster; Darvish et al. 2016). However, we stress that our aim in this work is to investigate the specific impact of the anisotropic tides on galaxy properties, therefore disentangling explicitly the effect of the geometry of the large-scale environment on driving these tides (i.e. the traceless part of the tidal tensor) from effects purely driven by the local density (i.e. the trace of the tidal tensor). In order to achieve this goal, we will focus on mass and colour gradients specifically towards filaments.

The paper is organized as follows. Section 2 describes the observed and simulated data sets, and the tool to extract the density and the skeleton. Section 3 assesses the reliability of the skeleton reconstruction using photometric redshifts, while relying on the realistic virtual photometric mocks built from the HORIZON-AGN hydrodynamical cosmological simulation (Dubois et al. 2014). Section 4 presents results from the data and our interpretation of these results. Section 5 wraps up and outlines future works. Appendices B, C, and D investigate how the reconstruction and the signal vary as a function of the thickness of the slice, the photometric redshift accuracy, and the completeness of the sample, hence providing guidelines for computing a similar reconstruction in surveys with different redshift uncertainties. Appendix E assesses the robustness of the gradients towards filaments when removing the contribution of the nodes. We use a standard Λ cold dark matter (ΛCDM) cosmology with Hubble constant $H_0 = 70.4\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, total matter density $\Omega_\mathrm{m} = 0.272$, and dark energy density $\Omega_\Lambda = 0.728$. All magnitudes are expressed in the AB system. Throughout this work, uncertainties on the distributions are computed from bootstrap resampling. The shaded area around the distributions corresponds to 1σ deviation to the mean for a bootstrap resampling with 100 realizations. All the differential distributions are normalized.

2 DATASETS AND METHODS

Before analysing the relative distribution of galaxies within the projected 2D cosmic web, let us first introduce the COSMOS2015 photometric catalogue, the mock catalogue extracted from the HORIZON-AGN simulation, and the ridge tracing algorithm used to trace the cosmic web.

[2] http://cosmos.astro.caltech.edu/

2.1 Photometric catalogue from COSMOS

Let us first summarize the main properties of the COSMOS2015 photometric catalogue that we built in Laigle et al. (2016).

This catalogue includes apparent magnitudes in 30 bands from ultraviolet (UV) to infra-red (IR): 0.25–8 μm. In optical, it contains the same broad-band images as previous releases (Capak et al. 2007; Ilbert et al. 2009). The COSMOS-20 survey provides most of the optical coverage: 6 broad-bands: B, V, r, i, z^{++}, 12 medium bands, and 2 narrow bands taken with Subaru Suprime-Cam (Taniguchi et al. 2007, 2015). The u^*-band data are obtained from the Canada–Hawaii–France Telescope (CFHT/MegaCam). Intermediate bands are very useful to increase the spectral resolution of the measured spectral energy distribution (SED) in the apparent optical wavelength range, contributing in this way to the accuracy of photometric redshifts especially at low redshift (e.g. Cardamone et al. 2010). In the near-infrared (NIR), we rely on the new Y, J, H, K_s images from the UltraVISTA survey (DR2; McCracken et al. 2012) and the Y band of the Hyper Suprime-Cam at Subaru telescope (Miyazaki et al. 2012a). The photometry is extracted using SExtractor (Bertin & Arnouts 1996) in dual image mode. Following a similar reduction procedure as described in McCracken et al. (2012), the detection image is a χ^2 sum of the four NIR images of UltraVISTA DR2 and the z^{++} band taken with Subaru Suprime-Cam. This combination ensures to increase the completeness of the catalogue. Mid-IR data in the four IRAC channels (i.e. in a wavelength range between ∼3 and 8 μm) come from the SPLASH programme (PI: Capak).

Photometric redshifts (z_phot) are computed using LePhare (Arnouts et al. 2002; Ilbert et al. 2006) with a configuration similar to Ilbert et al. (2013). In particular, the principal nebular emission lines are implemented using an empirical relation between the UV light and the emission line fluxes as described in Ilbert et al. (2009).

From a comparison with the spectroscopic samples zCOSMOS bright (Lilly et al. 2007) and faint (Lilly et al. in preparation), the uncertainty of our photometric redshifts turns out to be $\sigma_{\Delta z}/(1+z) = 0.007$ at $i^+ < 22.5$ and $\sigma_{\Delta z}/(1+z) = 0.03$ at $1.5 < z < 4$. Stellar masses, absolute magnitudes, and star formation rates (SFRs) are also derived with the LePhare code.

We remove from the catalogue all the objects which are flagged as belonging to a bad area or for which the photometry is possibly contaminated by the light of saturated stars, meaning that we keep only objects in $\mathcal{A}^\mathrm{UVISTA}\&\mathcal{A}^{!\mathrm{OPT}}\&\mathcal{A}^\mathrm{COSMOS}$, according to the notations in table 7 of Laigle et al. (2016).

Passive galaxies are identified using their locations in the rest-frame colour–colour plane NUV $-$ r/r $-$ J (Williams et al. 2009; Ilbert et al. 2013). Passive objects are identified as those with $M_\mathrm{NUV} - M_r > 3(M_r - M_J) + 1$ and $M_\mathrm{NUV} - M_r > 3.1$ (fig. 16 in Laigle et al. 2016). In particular, this technique avoids mixing the red dusty galaxies and passive ones. This criterion is expected to be more robust than a segregation based on the SFR derived from the SED fitting, suffering from the poor knowledge of the dust attenuation shape in UV and optical wavelengths. In the following, "colour-type" designs the galaxy type (passive or star-forming) determined based on the colour diagram.

The stellar mass completeness of the sample is estimated from the K_s magnitude following Pozzetti et al. (2010). The stellar mass limits up to redshift 1.3 are shown in Table 1 for the entire population of galaxies. In the following, our results are based on a galaxy sample with $M_\mathrm{lim} > 10^{10}\,\mathrm{M_\odot}$ and $0.5 < z < 0.9$. This choice is justified in Section 2.4. The final galaxy catalogue contains 11 284 objects.

Table 1. Limiting mass ($M_{\rm lim}$), estimation of the comoving transverse width of the slice ($D_{\rm trans}$), redshift errors ($\Delta z[M_* \sim 10^{10}]$) and corresponding thickness in Mpc ($D_{\Delta z}[M_* \sim 10^{10}]$), and chosen thickness of the slices in redshift ($\Delta z[D = 75\,{\rm Mpc}]$) and in Mpc ($D_{\Delta z}$). The limiting mass is estimated as in Laigle et al. (2016). The comoving transverse distance is estimated at the upper limit of each redshift bin and corresponds to 1.4 deg (side of the COSMOS field). $M_* > 10^{10}{\rm M}_\odot$ is the conservative limiting mass that we choose in this work to build a mass-selected sample at $0.5 < z < 0.9$. $\Delta z[M_* \sim 10^{10}{\rm M}_\odot]$ corresponds to twice the median 1σ redshift error determined by LEPHARE for the galaxies with masses $10^{10} < M_*/M_\odot < 10^{10.5}$ (the faintest galaxies in the bin, for which the redshift uncertainties are the highest). $D_{\Delta z}[M_* \sim 10^{10}{\rm M}_\odot]$ is the corresponding thickness in comoving Mpc at the upper limit of the redshift bin. Finally, $\Delta z[D = 75{\rm Mpc}]$ is the maximum in each redshift bin of the redshift thicknesses chosen in this work corresponding to $D_{\Delta z} = 75$ Mpc. It has been chosen such as to encompass two times the median redshift error in the lowest mass bin at $z \sim 0.9$.

bin	$M_{\rm lim}$	$D_{\rm trans}$ (Mpc)	$\Delta z[M_* \sim 10^{10}]$ ($2 \times 1\sigma$)	$D_{\Delta z}[M_* \sim 10^{10}]$ (Mpc)	$\Delta z[D = 75\,{\rm Mpc}]$	$D_{\Delta z}$ (Mpc)
$0.5 < z < 0.6$	8.72	54	0.019	58.3	0.024	75
$0.6 < z < 0.7$	8.92	62	0.021	61.6	0.025	75
$0.7 < z < 0.8$	8.95	70	0.024	68.0	0.027	75
$0.8 < z < 0.9$	9.17	75	0.029	75.2	0.029	75
$0.9 < z < 1.0$	9.19	82	0.032	80.4	–	–
$1.0 < z < 1.1$	9.23	87	0.045	105.1	–	–
$1.1 < z < 1.2$	9.37	93	0.059	131.4	–	–
$1.2 < z < 1.3$	9.44	98	0.059	128.1	–	–

2.2 Virtual catalogue from HORIZON-AGN

2.2.1 *The HORIZON-AGN lightcone*

The HORIZON-AGN simulation[3] (see Dubois et al. 2014, for more details) is run with a ΛCDM cosmology with total matter density $\Omega_{\rm m} = 0.272$, dark energy density $\Omega_\Lambda = 0.728$, amplitude of the matter power spectrum $\sigma_8 = 0.81$, baryon density $\Omega_{\rm b} = 0.045$, Hubble constant $H_0 = 70.4\,{\rm km\,s^{-1}\,Mpc^{-1}}$, and $n_s = 0.967$ compatible with the WMAP-7 data (Komatsu 2011). The size of the simulation box is $L_{\rm box} = 100\,h^{-1}$ Mpc on a side, and the volume contains 1024^3 dark matter particles, corresponding to a dark matter mass resolution of $M_{\rm DM,res} = 8 \times 10^7\,{\rm M}_\odot$. The simulation is run with the RAMSES code (Teyssier 2002), and the initially coarse 1024^3 grid is adaptively refined down to $\Delta x = 1$ proper kpc, with refinement triggered in a quasi-Lagrangian manner: if the number of dark matter particles becomes greater than 8, or the total baryonic mass reaches 8 times the initial dark matter mass resolution in a cell.

Heating of the gas from a uniform UV background takes place after redshift $z_{\rm reion} = 10$ following Haardt & Madau (1996). Gas can cool down to 10^4 K through H and He collisions with a contribution from metals using rates tabulated by Sutherland & Dopita (1993). Star formation occurs in regions of gas number density above $n_0 = 0.1\,{\rm H\,cm^{-3}}$ following a Schmidt law: $\dot{\rho}_* = \epsilon_* \rho_{\rm g}/t_{\rm ff}$, where $\dot{\rho}_*$ is the star formation rate mass density, $\rho_{\rm g}$ is the gas mass density, $\epsilon_* = 0.02$ is the constant star formation efficiency, and $t_{\rm ff}$ is the local free-fall time of the gas. The stellar mass resolution is $M_{*,\,\rm res} \simeq 2 \times 10^6\,{\rm M}_\odot$. Feedback from stellar winds, supernovae type Ia and type II are included into the simulation with mass, energy, and metal release. The simulation also follows the formation of black holes, which can grow by gas accretion at a Bondi-capped-at-Eddington rate and coalesce when they form a tight enough binary. Black holes release energy in a quasar/radio (heating/jet) mode when the accretion rate is, respectively, above and below one per cent of Eddington, with efficiencies tuned to match the black hole–galaxy scaling relations at $z = 0$ (see Dubois et al. 2012, for details).

To better mimic observed surveys, a lightcone is produced on-the-fly at every coarse time step of the simulation. The lightcone generation is described in Pichon et al. (2010). For the lightcone extraction, we have replaced gas cells by gas particles, and treated them as for the stars and dark matter particles. The area of the lightcone is 5 deg^2 below $z = 1$ and 1 deg^2 above.

2.2.2 *Galaxy catalogue production*

We identify galaxies using the ADAPTAHOP structure finder (Aubert et al. 2004; Tweed et al. 2009), applied to the distribution of star particles in the simulated lightcone. Because ADAPTAHOP works at fixed redshift, the extraction is done in thin slices of the lightcone, each slice being kept at fixed redshift. To avoid cutting galaxies, slices are overlapping, then duplicated galaxies are removed. Structures are selected using a local threshold of 178 times the average matter density, the local density of individual particles being computed from the 20 nearest neighbours. Only structures that have more than 50 particles are considered. (A further mass cut similar to the observed catalogue will be applied at a later stage of the catalogue production.) We compute galaxy fluxes and magnitudes using the single stellar population (SSP) models from Bruzual & Charlot (2003), using a Salpeter (1955) initial mass function. We assume that each star particle behaves as an SSP and compute its contribution to the total SED by logarithmically interpolating the models in metallicity and age and multiplying by the initial mass of the particle. The computation of the photometry includes redshift distortions due to the galaxy peculiar velocities. Galaxy exact redshifts are then perturbed in order to mimic the redshift accuracy of the COSMOS2015 catalogue. More specifically, the simulated galaxy catalogue is divided in bins of both redshift and i-band magnitude, and in each bin galaxy redshifts are perturbed using a random Gaussian error corresponding to the observed one in COSMOS2015, in the same redshift and magnitude bins. A spectroscopic redshift is assigned to the same fraction of bright galaxies as in the observed catalogue. In the following, this mock catalogue is used to test the 2D skeleton reconstruction and our ability to recover mass gradients in two dimensions. Simulated galaxy masses are taken as the intrinsic ones, i.e. the sum of the masses of all stellar particles within the galaxies. Galaxy mass errors are not incorporated in this work, because a comprehensive modelling of uncertainties (including systematics) on masses derived from SED fitting is clearly beyond its scope.

[3] http://www.horizon-simulation.org/

A critical assessment of the HORIZON-AGN simulation in term of the evolution of its galaxy populations is presented in Kaviraj et al. (2017) through comparison of colours, luminosity functions and mass functions. When taking into account uncertainties and systematics, they found that the simulation captures reasonably well the evolutionary trend of galaxy mass assembly as seen through the statistical quantities. The remaining point of tensions between the simulated and observed mass functions concern the high-redshift Universe (underestimation of the mass function at the high-mass end) and the low-mass end of the mass function. The overestimation of the low-mass end in the simulation is not due to an overproduction of the number of low-mass galaxies, but to an overestimation of the mass of these small galaxies. As a matter of fact, the halo mass function is correctly reproduced in the simulation, but the stellar-to-halo mass relation is slightly above the observational data for low-mass galaxies. The reason is likely to be the strength of the stellar feedback, which is not high enough. While the high redshift tension does not impact this study, the low-mass issue can potentially affect the accuracy of the predictions for the skeleton reconstruction: if more density tracers are present in the mocks than in the observations, more structures are likely to be identified. To counterbalance this effect, the simulated catalogue used for tracing the skeleton presented in this work are based on a subset of galaxies extracted from the original catalogue. Only the most massive galaxies are kept from the simulated catalogue, the number of galaxies being determined to match the number density of galaxies in the observed catalogue with a mass $\log M_*/\mathrm{M}_\odot > 10$. The mass limit in this subsample is $10^{10.2}\,\mathrm{M}_\odot$. However, for the sake of simplicity, we keep in the following the same mass bins for the simulated and observed catalogue, but keeping in mind that they may probe slightly different galaxy populations in the lowest mass bins.

Finally, in order to get a sample statistically similar to the COSMOS2015 catalogue, we work only on the central two square degrees of the lightcone and masks similar to the COSMOS field are applied on the simulated catalogue.

The separation between star-forming and passive galaxies based on photometry is also less straightforward in HORIZON-AGN than in data: a large fraction of passive galaxies are indeed moved towards the star-forming region in the colour diagram due to residual star formation, which is not completely extinguished in passive systems. In addition, the low-mass galaxies at low redshift are generally too passive (see e.g. the sSFR–mass relation in Dubois et al. 2016). For this reason relying on the colour diagram itself is not sufficient to properly separate passive and star-forming galaxies. In this work, we do not directly discriminate between passive and star-forming systems. However, we will look for a segregation between the 45 per cent most star-forming and the 45 per cent most passive galaxies in each mass bins, based on the sSFR distribution.

2.3 Density and cosmic network identification

2.3.1 The persistent skeleton algorithm

The filaments of the cosmic web are identified from the density field, which refers either to the number density of dark matter particles (in the case of the 3D reference skeleton in the simulation) or to the number density of galaxies (for the observed and simulated 2D skeletons). When computed from the galaxy distribution, the density is estimated from the Delaunay tessellation of the particles (Schaap & van de Weygaert 2000).

To identify the cosmic network from the density, we use the persistence based filament tracing algorithm (DISPERSE; Sousbie 2011; Sousbie, Pichon & Kawahara 2011). This method identifies ridges from the density field as the special lines connecting topologically robust saddle points to peaks (which are called 'nodes', in the following). The identified filament network is by construction multiscale, while the extraction is robust to noise. The set of all segments defining these ridges is called the skeleton (Pogosyan et al. 2009). DISPERSE has been already successfully used in two dimensions to characterize stellar filaments in the Milky Way (e.g. Arzoumanian et al. 2011; Panopoulou et al. 2017).

Each filament is defined to be a set of connected small segments linking extrema together. Persistence is defined as the ratio of the density value at the two critical points in a topologically significant pair of critical points: maximum-saddle, saddle-saddle (in three dimensions only), and saddle-minimum. This ratio quantifies the robustness of the underlying topological feature characterized by this pair (the fact that the connected critical points are responsible for the appearance of a new topological feature such as connected component, tube, ball, in the corresponding excursion set). Expressed in terms of numbers of σ, persistence quantifies the significance of the critical pairs in the Delaunay tessellation of a random discrete Poisson distribution. In the following, a skeleton with a 2σ persistence level is a skeleton for which all the persistence pairs that have their probability to be found in a random discrete Poisson distribution below 2σ of the mean have been removed. Removing low-persistence pairs is a multiscale non-local method to filter noise/low significance filaments. This method is particularly well adapted to noisy data sets such as redshift catalogues (Sousbie 2011).

Table 2 summarizes the extracted skeletons both on observations and on simulations and their main features, which are also detailed below.

2.3.2 Cosmic web extraction in COSMOS

To identify filaments in the 2D density field from the COSMOS2015 catalogue, we proceed as follows.

We first divide the galaxy sample in slices of constant thickness, the choice of which is justified in Section 2.4. To estimate the density in each slice, we compute the 2D Delaunay tessellation from the galaxy distribution. To take into account the errors on photometric redshifts in this calculation, we associate to each galaxy in each slice a weight $p_{\mathrm{gal},\,i}$ which is the probability for this galaxy to be in the slice i given the probability function distribution of its redshift $P_{\mathrm{gal}}(z)$ (as computed by LEPHARE):

$$p_{\mathrm{gal},i} = \int_{z_{1,i}}^{z_{2,i}} P_{\mathrm{gal}}(z)\,\mathrm{d}z \Big/ \int P_{\mathrm{gal}}(z)\,\mathrm{d}z\,, \tag{1}$$

where $z_{1,\,i}$ and $z_{2,\,i}$ are the lower and upper redshifts of the slice. This probability is used to weight the Delaunay tessellation at each galaxy location in order to sharpen the density estimation. This weighting allows to reduce the pollution effects of foreground and background galaxies for which the probability to be in the slices is low, and in turn, to increase the number of robust detected structures. Based on the simulation, we estimate that about 10 per cent more filaments of the reference skeleton are recovered in the weighted skeleton. When available, the redshift of the galaxy is assigned to its spectroscopic redshift.

To correctly estimate the density and the topology close to boundaries, a surface of 'guard' particles is added outside the bounding box and new particles are added by interpolating the estimated density computed on the boundary of the distribution. Note that the Delaunay tessellation is by construction a robust method to reconstruct

Table 2. The skeletons used in this work to assess the quality of the reconstruction or for gradients measurements. $M_{\rm lim}$ refers to the mass limit of the galaxy sample used for the computation, σ to the persistence threshold and $D_{\Delta z}$ is the thickness of the slice (for 2D reconstructions). For $\mathrm{SKL}_{\rm 3D}^{\rm DM}$, the persistence threshold is not expressed in terms of σ. The quoted values are those used in the main text, but Appendices investigate how the reconstruction changes when these values vary.

Name	Catalogue	Dim		σ [a]	$M_{\rm lim}$ [a]	$D_{\Delta z}$ [a]	Comments
$\mathrm{SKL}_{\rm 2D}^{\rm obs}$	COSMOS2015	2D	Galaxies with photo-z	2	$10^{10}\,\mathrm{M}_\odot$	75 Mpc	Density weighted by the photo-z errors
$\mathrm{SKL}_{\rm 2D}^{\rm phot}$	HORIZON-AGN	2D	Galaxies with photo-z	2	$10^{10.2}\,\mathrm{M}_\odot$	75 Mpc	Density weighted by the photo-z errors
$\mathrm{SKL}_{\rm 2D}^{\rm spec}$	HORIZON-AGN	2D	Galaxies with exact z	2	$10^{10.2}\,\mathrm{M}_\odot$	75 Mpc	–
$\mathrm{SKL}_{\rm 3D}^{\rm DM}$	HORIZON-AGN	3D	Dark matter particles	–	$8\times10^{7}\,\mathrm{M}_\odot$	–	Density computed on a grid

[a] Unless specified otherwise.

reliably the density in a field that contains masks. The triangulation automatically connects galaxies from across the masked regions. The masks are not however filled with a random distribution of galaxies based on the averaged density in the vicinity of the masked regions (e.g. Aragon-Calvo et al. 2015). Appendix A quantifies the effect of masks on the skeleton reconstruction.

Once the density computed, filaments are identified with DISPERSE from the 2D Delaunay tessellation with a 2σ persistence threshold. The properties of the noise are not necessarily the same in two dimensions as in three dimensions. This implies that although a persistence threshold above 3σ would be more conservative in three dimensions (which guarantees that less than 0.3 per cent of the critical pairs are spurious), the 2D skeleton extracted with a 3σ persistence would miss a large number of robust 3D filaments. To compensate this effect, choosing a 2D persistence threshold of 2σ enables to drastically decrease the number of unrecovered 3D filaments (by a factor $\sim$1.7), without sensibly increasing the number of spurious 2D filaments (see also Malavasi et al. 2017, for a motivation behind this mapping; this choice is further justified in Appendix B using the simulation).

Fig. 1 shows the density of the COSMOS field in two slices at redshifts $z\sim0.58$ and $z\sim0.73$ estimated from the Delaunay tessellation and their corresponding skeletons in red.

2.3.3 *Cosmic web extraction in the mocks*

We extract the 2D skeletons in the simulation with mock photometric redshifts following the exact method used for the COSMOS field. In addition, a 2D skeleton is also extracted in each slice with true ('spectroscopic' in the following) redshifts (hence without including the weighting with the photometric uncertainties). Finally, a 3D skeleton ($\mathrm{SKL}_{\rm 3D}^{\rm DM}$) is extracted from the dark matter distribution. To compute this skeleton, in order to save computational time, dark matter particles are first projected on a 3D grid with a resolution of 250 comoving kpc using a cloud-in-cell algorithm. We perform a first extraction with a very low persistence threshold. Then the adopted persistence threshold is estimated to be $\sim$5 times the rms level of density fluctuations for the low-density filaments in the previous extraction. Note that our results are robust against the exact value chosen for this persistence threshold. This skeleton becomes the reference skeleton. It is indeed the most suitable, as the galaxy distribution itself is already a biased tracer of the underlying density field. In addition, dark matter particles are more numerous in the simulation, allowing us to extract more accurate filament positions, which will be all the more important when computing distances between 2D and 3D skeletons in Section 3.

2.4 Choice of slices

The thickness of the slices on which the projected density is computed and the skeleton extracted will impact our analysis. A volume-limited sample with a constant stellar mass cut and a constant slice thickness is required, in order to avoid possible systematics in the measurements due to different quality of reconstruction (because of increasing slice thickness with increasing redshifts). In two dimensions, there can be some confusion between nodes (maxima in the density field) and projected 3D filaments, or filaments and projected 3D walls. Furthermore, projected filaments may link some overdensities that are in fact not truly linked in three dimensions. Hence if the slices are too thick, the skeleton derived from the projected density will have little correspondence to its 3D counterpart, with a large number of fake 2D filaments and non-recovered 3D filaments. However, if the chosen thickness of the slices is too small, the 2D skeleton will trace only fragments of the 3D filaments and will be, in turn, a poor representation of the fully connected 3D skeleton.

Following Appendix D, we choose to work up to redshift 0.9 with slices of thickness 75 comoving Mpc, and by considering all galaxies more massive than $10^{10}\,\mathrm{M}_\odot$. This final choice is guided both by the required thickness for an optimal reconstruction when exact redshifts are available and by the current redshift accuracy of our sample. As it is shown in Table 1, with a thickness of 75 Mpc, all galaxies of the final sample have their redshift uncertainties smaller than the thickness of the slice, where uncertainties are estimated from the $2\times1\sigma$ errors[4] given by LEPHARE for each galaxy.

Tests on the virtual photometric catalogue detailed below confirm that this choice is good enough to reconstruct the skeleton. We do not consider galaxies below redshift 0.5 as the comoving transverse width of the field starts to be too small to extract reliably the large-scale structure. The adjacent redshift slices are overlapping and spaced by half of the thickness of the slices. There are 30 slices on the redshift range of $0.5<z<0.9$.

3 EXTRACTION ASSESSMENT IN MOCKS

3.1 2D versus 3D skeletons

The HORIZON-AGN simulated catalogue is now used to estimate the quality of the skeleton reconstructed with photometric redshifts. For this assessment, the 2D photometric skeleton is compared both to the spectroscopic one and to the projected 3D reference one.

[4] The 1σ error is defined as the value enclosing half of 68 per cent of the probability distribution function of the photometric redshift of the galaxy.

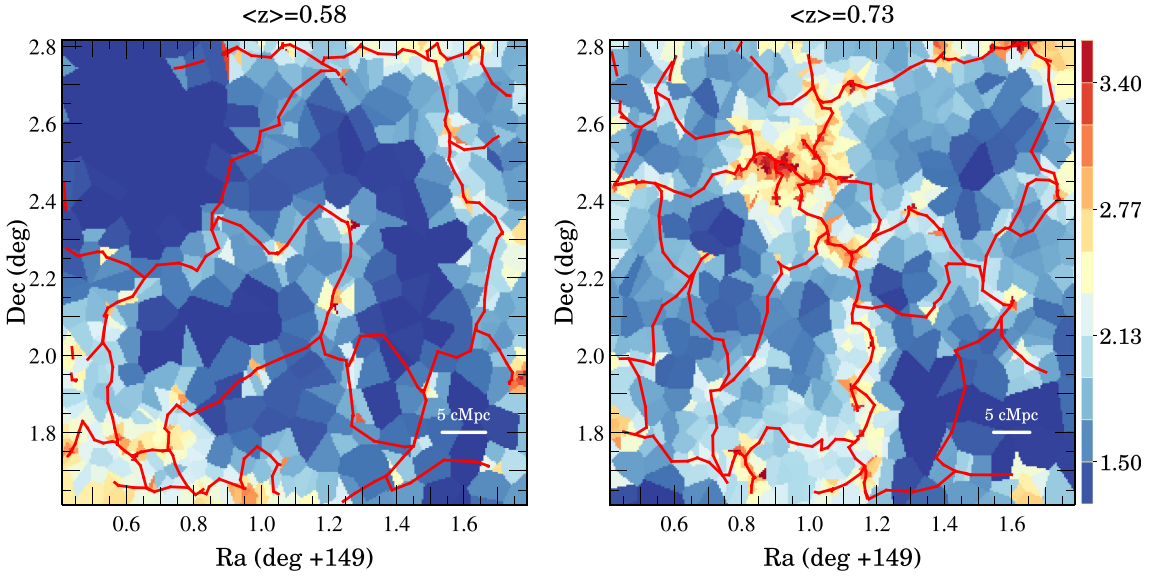

Figure 1. Two examples of the density field estimated from the Delaunay tessellation in slices of thickness 75 Mpc in the COSMOS field. The density is down-weighted by the photometric errors. The red skeleton has been computed with a persistence threshold of 2σ. The horizontal white bar in the lower right indicates the comoving scale length of 5 Mpc.

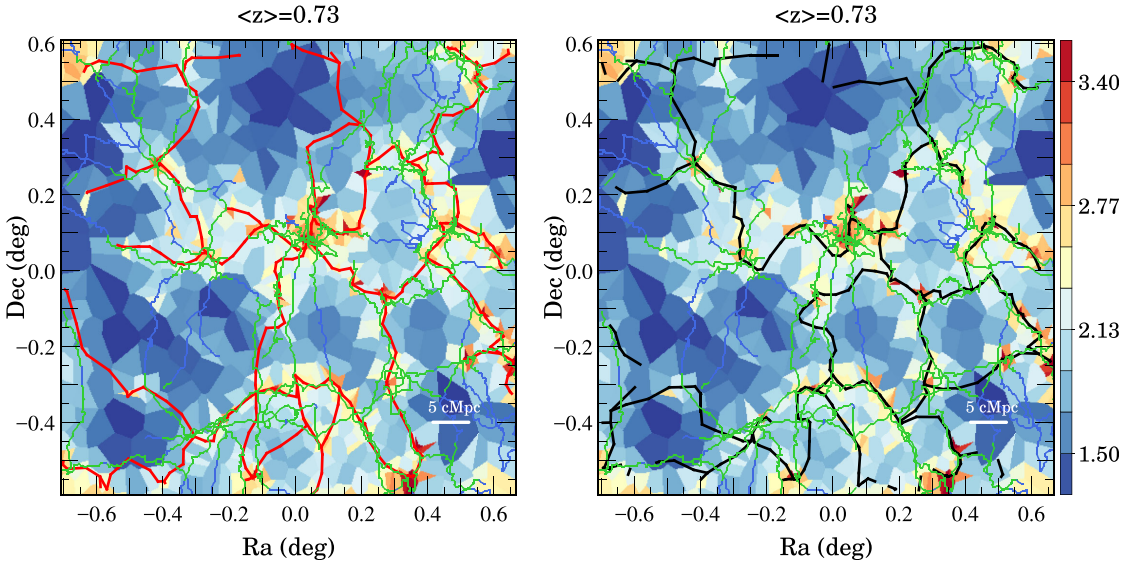

Figure 2. An example of the density field estimated from the Delaunay tessellation at $z = 0.73$ in a slice of thickness 75 Mpc from the HORIZON-AGN lightcone. The left-hand panel displays the photometric SK_{2D}^{phot} (red line) and the projected DM SKL_{3D} (green and blue) skeletons, while the right-hand panel shows the spectroscopic SK_{2D}^{spec} (dark line) and projected DM SKL_{3D} skeletons. SKL_{3D} is computed from the full dark matter particles distribution. Its segments are blue when they do not have a counterpart neither in SK_{2D}^{spec} or SKL_{2D}^{phot} within 0.8 projected Mpc. The photometric and spectroscopic skeletons are computed with a 2σ persistence threshold from the galaxy distribution. The background colour codes, the galaxy number density computed from the weighted Delaunay tessellation with photometric redshifts. Most of the filaments in SKL_{2D}^{spec} have a close counterpart in SKL_{3D}. We note that some 3D filaments are recovered in SKL_{2D}^{phot} but not in SKL_{2D}^{spec} (e.g. the filament in [RA, Dec.] = [−0.4, 0.05]), which is a consequence of the weighting by the photometric redshift uncertainties. The horizontal white bar in the lower right indicates the comoving scale length of 5 Mpc.

Fig. 2 displays a slice of thickness 75 Mpc around redshifts $z \sim 0.73$ in the HORIZON-AGN simulation. The left-hand panel shows the confrontation of the photometric (red) with the reference SKL_{3D}^{DM} (green and blue) skeletons, while the right-hand panel shows the spectroscopic (black) and the reference SKL_{3D}^{DM} skeletons. The segments are blue when they do not have a 2D counterpart (neither in the photometric or spectroscopic skeleton) within 0.8 projected Mpc. The 2D skeletons are visually in relatively good agreement with the projected 3D one. In more details, we see that most filaments of SKL_{2D}^{phot} have a reliable counterpart in SKL_{3D}^{DM}. However, some filaments of the former are not recovered in the latter and some filaments of the latter have no counterpart in the former. Hence, the choice of the persistence threshold is important to mitigate the number of filaments in SKL_{2D}^{phot} which have no counterpart in SKL_{3D}^{DM} and which are thus potentially noisy detection. In addition, filaments detected in SKL_{2D}^{phot} but not in SKL_{3D}^{DM} may potentially be

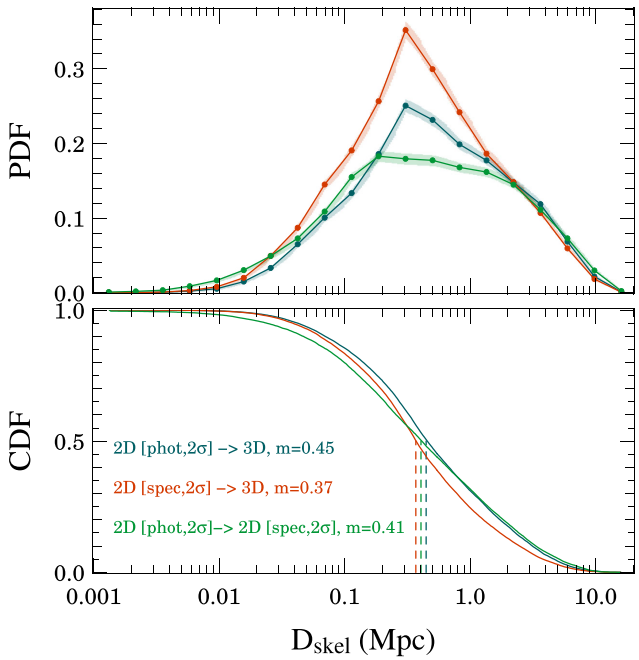

Figure 3. Differential (*top*) and cumulative (*bottom*) distributions of the distances between SKL_{2D}^{phot} and SKL_{3D}^{DM} (blue line), SKL_{2D}^{spec} and SKL_{3D}^{DM} (red line), and SKL_{2D}^{phot} and SKL_{2D}^{spec} (green line) in the HORIZON-AGN lightcone. For each segment in the first skeleton, we measure the distance to its closest segment in the second. The distances are measured in all slices between $z \sim 0.5$ and $z \sim 0.9$. The *m* values specify the medians of the distribution.

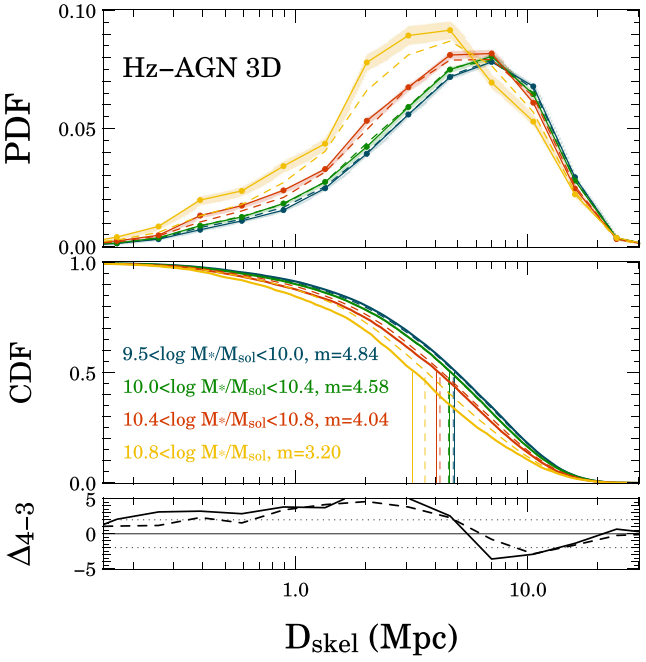

Figure 4. Differential (top) and cumulative (middle) distributions of the distances to the 3D (not projected) dark matter skeleton SKL_{3D}^{DM} for galaxies as a function of their masses in the HORIZON-AGN lightcone for $0.5 < z < 0.9$. The dashed distributions correspond to a random signal preserving the mass–density relation. The *m* values give the medians of each distribution, which are also indicated by vertical lines. The contribution of nodes to the gradient signal has been removed from the analysis to highlight an effect specifically related to the filaments. The bottom row shows the residuals between the two more massive bins ($10.4 < \log M_*/\mathrm{M}_\odot < 10.8$ and $0.8 < \log M_*/\mathrm{M}_\odot < 13.0$) for the original (solid line) and reshuffled (dashed line) samples.

wells seen in projection. Finally, note that some 3D filaments are better recovered in the photometric skeleton than in the spectroscopic skeleton, which is a consequence of the weighting by the photometric redshift uncertainties.

Let us quantify this statistically by measuring the distribution of distances, $d_{2D\to 3D}$, between SKL_{2D}^{phot}, SKL_{2D}^{spec}, and SKL_{3D}^{DM} skeleton. The distance between two skeletons is determined following the method introduced by Sousbie (2011): for every filamentary segments in the first skeleton, the closest neighbours in the second skeleton are identified. The distance between the two skeletons is estimated from the medians of the distances between all individual segments. These distributions give us an estimation of the reliability of the skeleton. They are shown on Fig. 3. Green and blue lines correspond to the distributions of the distances $d_{2D\to 3D}$, using spectroscopic and photometric redshift, respectively, for the computation of SKL_{2D}. Although the photometric skeleton has a slightly larger tail, the medians of the distribution stay relatively close, which suggests that the photometric redshift uncertainties have a small impact on our global ability to recover the 3D structures, at least under the chosen parameters. However, the tail of the distribution implies that a certain number of 2D filaments have no 3D counterpart at all; that is, they correspond to spurious 2D filaments connecting points in two dimensions that are spatially disconnected in three dimensions. We find that, on average, 22.0 per cent and 16.5 per cent of segments have no counterpart closer than 1.5 projected Mpc in the photometric and spectroscopic skeleton, respectively.

Note that SKL_{2D}^{phot} is degraded or biased compared to the projected SKL_{3D}^{DM} skeleton for three reasons. First, the photometric skeleton SKL_{2D}^{phot} will be noisier than the spectroscopic one SKL_{2D}^{spec} due to the galaxy redshift uncertainties. Secondly, even with exact redshifts, SKL_{2D}^{spec} is a biased reconstruction of the projected 3D skeleton computed from galaxies due to the projection effect. Finally, it should be acknowledged that the 3D galaxy skeleton SKL_{3D}^{gal} is biased compared to the 3D dark matter skeleton SKL_{3D}^{DM}. This bias, although relatively negligible, is dependent on the limiting mass of the galaxy catalogue and is investigated in Appendix C.

3.2 3D measurement from 2D at COSMOS2015 accuracy

Although the 3D filaments can be reliably traced from 2D slices at the COSMOS2015 redshift accuracy, it is not obvious that a distance-to-filament trend will be statistically recovered in two dimensions. The projection will inevitably blur the measurement: galaxies far from the filament but on the same line-of-sight will be counted as very close by projection. We will therefore test in the HORIZON-AGN simulation if a signal existing in three dimensions may be recovered in two dimensions.

Following the recent results from Malavasi et al. (2017), we expect to find mass gradients towards 3D filaments, i.e. more massive galaxies closer to the filament centre than less massive ones. We first measure this trend in three dimensions. In a second step, we verify that the signal is also measurable in two dimensions.

The solid lines on Fig. 4 correspond to the differential and cumulative distributions of the 3D distances to SKL_{3D}^{DM} for simulated galaxies in four different bins of masses in the HORIZON-AGN lightcone. The upper mass bin is chosen to contain galaxies above the knee of the observed galaxy mass function, which is found to be $\sim 10^{10.8}\,\mathrm{M}_\odot$ in $0.5 < z < 0.8$ from the fit of the mass

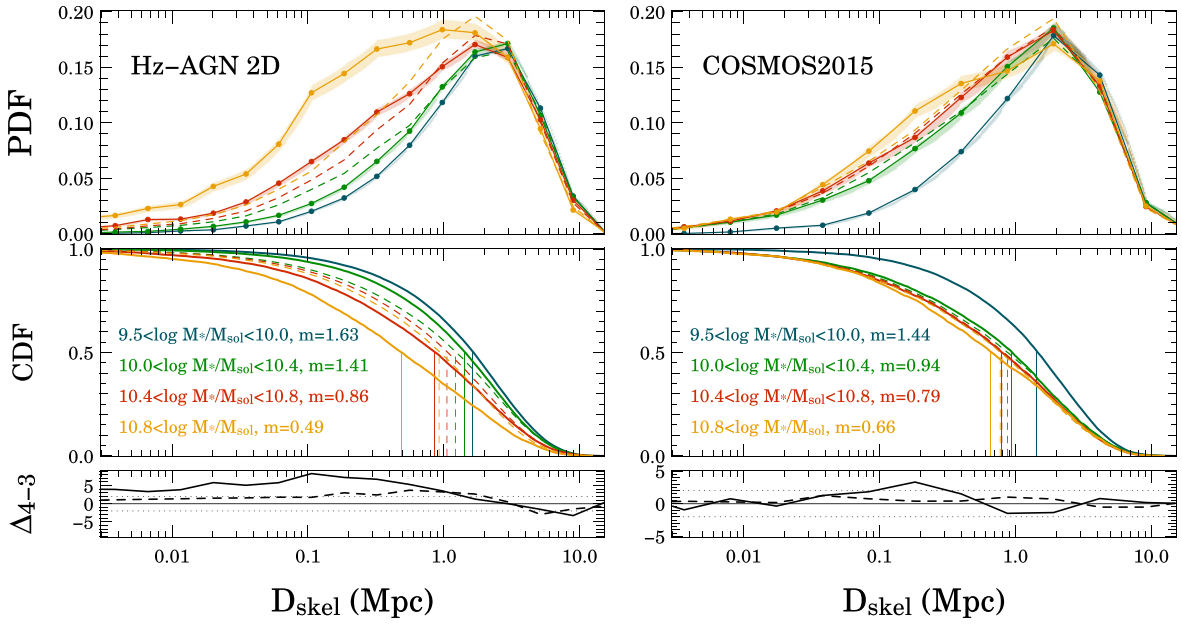

Figure 5. Differential (top) and cumulative (middle) distributions of the distances to the 2D photometric skeleton SKL_{2D}^{phot} for galaxies as a function of their masses in HORIZON-AGN (left) and COSMOS2015 (right). The dashed lines correspond to a random signal preserving the mass–density relation. Note that the lowest mass bin $9.5 < \log M_*/\mathrm{M}_\odot < 10.0$ should be considered with caution because the thickness of the slices has been calibrated for galaxies with a mass $\log M_*/\mathrm{M}_\odot > 10$. Here, the signal measured in all slices between $z \sim 0.5$ and $z \sim 0.9$ is stacked. The *m* values give the medians of each distribution, which are also indicated by vertical lines. The contribution of nodes to the gradient signal has been removed from the analysis to highlight an effect specifically related to the filaments. The bottom row shows the residuals between the two more massive bins ($10.4 < \log M_*/\mathrm{M}_\odot < 10.8$ and $10.8 < \log M_*/\mathrm{M}_\odot < 13.0$) for the original (solid line) and reshuffled (dashed line) samples.

function with a double Schechter profile (Davidzon et al. 2017). To better quantify the significance of the result, the lower panel shows the residual between the distributions of the two intermediate bins expressed in number of σ: $\Delta_{4-3} = \Delta(D_4 - D_3)/\sqrt{(\sigma_3^2 + \sigma_4^2)}$, where the subscripts 3 and 3 refer to the galaxies such that $10.4 < \log M_*/\mathrm{M}_\odot < 10.8$ and $10.8 < \log M_*/\mathrm{M}_\odot < 13.0$, respectively. The deviation exceeds two sigma: the most massive galaxies are found closer to the filament centre than the low mass ones, in agreement with Malavasi et al. (2017).

The same measurements are then carried in two dimensions with SKL_{2D}^{phot} on the HORIZON-AGN lightcone, in all slices between $0.5 < z < 0.9$. The projected transverse distances to the filament, measured in degrees, are then converted in comoving Mpc. The cumulative distributions in all slices are stacked and the mean distribution is presented in the left-hand panel of Fig. 5. Note that the lowest mass bin $9.5 < \log M_*/\mathrm{M}_\odot < 10.0$ should be considered with caution. These galaxies are not used to compute the skeleton, and their redshift uncertainties ($2\times\ 1\sigma$) are generally higher than the thickness of the slices.

3.2.1 *Removing the contribution of the nodes*

Let us now stress that we are looking for an environmental effect driven specifically by filaments. As galaxies are more massive and passive in clusters than in the field, we expect mass gradients towards nodes (defined as the maxima of the density field at the connection of filaments, where clusters are expected to lay), driven by the effect of cluster alone. These gradients could contaminate the mass gradient-trend towards filaments. To minimize node contributions, we remove from the analysis galaxies closer to a node than 3.5 and 0.8 projected Mpc in three and two dimensions, respectively. We show in Appendix E that this operation allows to correctly remove their effect. We still recover mass gradients towards filaments both in three and two dimensions: more massive galaxies are more confined to filaments than less massive ones. These remaining gradients are associated with filaments only.

3.2.2 *Removing the contribution of the background density*

However, it is still unclear what are the respective contributions of the local density (in which case the geometry of the environment does not matter) and the proximity to the filament itself. It is indeed well known that denser environments contain more massive, redder, and much less star-forming galaxies (e.g. Dressler 1980; Balogh et al. 1997; Blanton et al. 2003; Kauffmann et al. 2004; Cucciati et al. 2016; Davidzon et al. 2016), even when excluding cluster regions. Since the density is on average higher closer to the filaments than further away, such local mass–density relation could naturally lead to mass gradients towards filaments.

In order to disentangle the contribution of the proximity to the filament from the effect of the local mass–density relation, a first test is performed, following Malavasi et al. (2017). The galaxy catalogue is split in bins of background density (40 equipopulated density bins), where the density is computed from the galaxy Delaunay tessellation (down-weighted by the photometric uncertainties in the 2D case, as described above) smoothed with a Gaussian kernel of 3 Mpc and 1 *projected* Mpc in three and two dimensions, respectively. In two dimensions, galaxies in the lowest mass bin $\log M_*/\mathrm{M}_\odot < 10.$ are not included in this test, because their redshift uncertainties are fairly large. Then positions are reshuffled between the galaxies in each density bin. This operation preserves the mass–density relation; hence, the role played by the background density. However, within a density bin, it moves galaxy positions with respect to the filaments. If the proximity to the filament does not impact the mass,

Table 3. Summary of the results for the mass gradients towards filaments. The given values are the medians of the distributions and the standard errors from bootstrap resampling with 100 realizations expressed in comoving Mpc. Note that the median values in two and three dimensions are not directly comparable because of the effect of the projection. The bracketed values indicate the percentage of galaxies in each mass bin. The percentages are the same for the original and reshuffled samples. The reshuffled sample corresponds to the random signal preserving the galaxy mass–density relation.

Mass range	Median values of the PDF								
	COSMOS2015			HZ-AGN 2D			HZ-AGN 3D		
	Per cent	Original	Reshuffled	Per cent	Original	Reshuffled	Per cent	Original	Reshuffled
$9.5 < \log M_*/M_\odot < 10$	46	1.44 ± 0.02	–	54	1.63 ± 0.01	–	54	4.84 ± 0.02	4.78 ± 0.03
$10 < \log M_*/M_\odot < 10.4$	23	0.93 ± 0.02	0.87 ± 0.02	27	1.41 ± 0.01	1.22 ± 0.01	27	4.58 ± 0.03	4.55 ± 0.03
$10.4 < \log M_*/M_\odot < 10.8$	18	0.79 ± 0.02	0.80 ± 0.02	13	0.86 ± 0.02	1.06 ± 0.01	13	4.04 ± 0.05	4.19 ± 0.03
$10.8 < \log M_*/M_\odot < 13$	11	0.66 ± 0.02	0.77 ± 0.03	5	0.49 ± 0.02	0.92 ± 0.02	5	3.20 ± 0.08	3.61 ± 0.05

Table 4. Summary of the results from the mass gradients towards filaments after density matching. The density matching algorithm ensures that the selected subset in all mass bins shares the same density distribution. Therefore, the remaining gradient signal is not driven by the background density. The given values are the medians of the distributions and the standard errors from bootstrap resampling with 100 realizations expressed in comoving Mpc.

Mass range	Median values of the PDF after density matching		
	COSMOS2015	HZ-AGN 2D	HZ-AGN 3D
[10, 10.4]	0.78 ± 0.02	1.11 ± 0.02	3.99 ± 0.07
[10.4, 10.8]	0.71 ± 0.01	0.71 ± 0.01	3.82 ± 0.05
[10.8, 13]	0.60 ± 0.02	0.48 ± 0.02	3.46 ± 0.06

the gradients after reshuffling should therefore be identical to the original ones. Conversely, if the background density does not play any role, there should be no gradients after reshuffling. The distributions of the distances to the filament after reshuffling are shown as dashed lines in Figs 4 and 5 and the median values of both original and reshuffled distributions are presented in Table 3. We find that the reshuffling partially preserves the mass gradients towards filaments; however, the reshuffled and original distributions are not identical. The gradient signal is much larger with the original distribution than the reshuffled one both in two and three dimensions. The result of this test suggests that, although the local density plays a role, it is not the only responsible for the mass gradients towards filaments.

To confirm this result, a second and independent test is performed. In each of the 40 equipopulated density bins, the same numbers of galaxies in every ranges of masses are selected. As a result, galaxies in each mass range display the same background density distribution. Therefore, if the density is the only responsible for mass gradients towards filaments, no gradient should be detected in this selected subsample. However, if the proximity to the filament plays a role independent from the density, a gradient signal should be measured. Galaxies in the lowest mass bin are not included in this test. The median values for the density-matched samples are given in Table 4. In the simulation, in three and two dimensions, significant mass gradients are still measured.

Let us finally summarize these results. First, mass gradients are found towards filaments in the HORIZON-AGN simulation. These gradients are specifically induced by filaments and not driven by nodes. In addition, they cannot be entirely explained by the mass–density relation. The particular geometry of the large-scale environment itself plays a role in driving these gradients. Secondly, we demonstrate based on the mocks that at the accuracy of the COSMOS2015 catalogue, mass gradients found in the 3D mocks can also be identified in the 2D ones. In the following, the method can therefore robustly be applied to the observed COSMOS2015 catalogue.

4 GALAXY PROPERTIES IN FILAMENTS

Having confirmed the feasibility to trace the filaments in 2D slices with a thickness of 75 Mpc and to recover the corresponding 3D signal, let us now study in more details the distribution of galaxy masses and colour-types as a function of their distance to filaments using both the observed COSMOS2015 and simulated HORIZON-AGN catalogues.

4.1 Galaxy mass gradients towards filaments in cosmos

The right-hand panel of Fig. 5 presents the distributions of the distance to the filaments for the four bins of mass in COSMOS2015. Similarly to the measurement in the simulation (left-hand panel) presented above, the signal measured in all slices in the redshift range of $0.5 < z < 0.9$ has been stacked and the contribution of the nodes has been removed. In COSMOS, massive galaxies are also found closer to the centre of filaments than the less massive ones, in agreement with the measurement in HORIZON-AGN. However, the median values of the distributions in the simulation span a larger distance range than in the COSMOS2015, indicating a stronger segregation between the mass bins.

One does not expect exactly the same values for the median of the distribution, because galaxy masses are slightly overestimated in the simulation. In addition, galaxy mass uncertainties, which are not modelled in the simulation, are likely to be another reason for this difference. Low-mass galaxies, being more numerous and with larger uncertainties than high-mass galaxies, are probably polluting the higher mass bins. This bias should dilute the signal especially in the high mass bins, as we indeed observe in Fig. 5. Furthermore, galaxy masses derived from SED-fitting are generally underestimated. Such systematics could explain the discrepancy between HORIZON-AGN and COSMOS2015 at a given mass.

The two tests described in Section 3 are also computed here in order to disentangle the contribution of the anisotropic large-scale environment from the effect of the local mass–density relation. The distributions after reshuffling are shown as dashed lines in Fig. 5, and the values after density matching are displayed in Table 4. Consistently to the HORIZON-AGN result, we find that the reshuffling strongly dilutes the gradient signal. However, even after the density matching test, significant gradients are measured. The outcome of these two tests allows us to conclude that, after removing the significant contribution of the nodes, a non-negligible part of the gradient

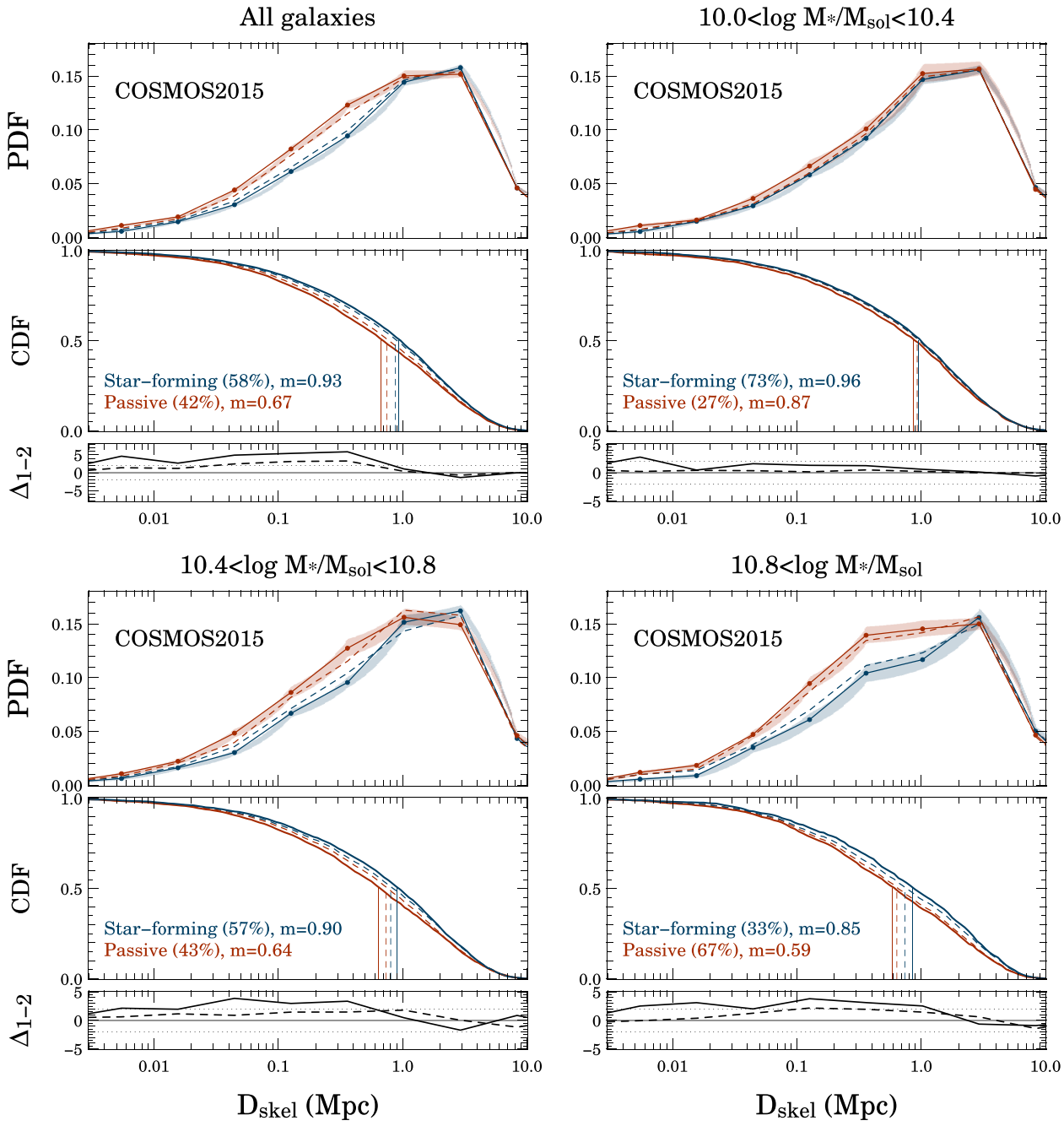

Figure 6. Differential (top) and cumulative (middle) distributions of the distances to the 2D photometric skeleton $\mathrm{SKL}_{\mathrm{2D}}^{\mathrm{phot}}$ for passive (red) and star-forming (blue) galaxies in COSMOS2015 for all galaxies such that 10. $< \log M_*/\mathrm{M}_\odot$ (top left), $10.0 < \log M_*/\mathrm{M}_\odot < 10.4$ (top right), $10.4 < \log M_*/\mathrm{M}_\odot < 10.8$ (bottom left) and $10.8 < \log M_*/\mathrm{M}_\odot$ (bottom right). The value m is the median of each distribution, which is also indicated by vertical lines. N gives the percentage of galaxies of each type in the considered mass bin. The dashed lines correspond to a random signal preserving the galaxy type–density relation in each bin of mass: galaxy types in the catalogue have been randomized with respect to their distances to the filament in given density and mass bins. Here, the signal measured in all slices between $z \sim 0.5$ and $z \sim 0.9$ is stacked. The contribution of nodes to the gradient signal has been removed from the analysis. The bottom rows show the residuals between the passive and star-forming distributions for the original (solid line) and reshuffled (dashed line) samples expressed in number of σ: $\Delta_{1-2} = \Delta(D_1 - D_2)/\sqrt{(\sigma_1^2 + \sigma_2^2)}$, where the subscripts 1 and 2 refer to the passive and star-forming distributions, respectively.

signal towards filament is driven by the geometry of the large-scale environment itself.

4.2 Colour-type segregation towards filaments in COSMOS

Let us now investigate the effect of the anisotropic environment on galaxy types and star formation.

The COSMOS2015 galaxy sample has been divided in passive and star-forming populations based on the colour diagram, as explained in Section 2.1. Fig. 6 presents the differential and cumulative distributions of the distances to the filaments for star-forming and passive galaxies for different mass bins in COSMOS2015. All galaxies with $M_* > 10^{10}\,\mathrm{M}_\odot$ are considered. The contribution of the nodes has been removed as explained

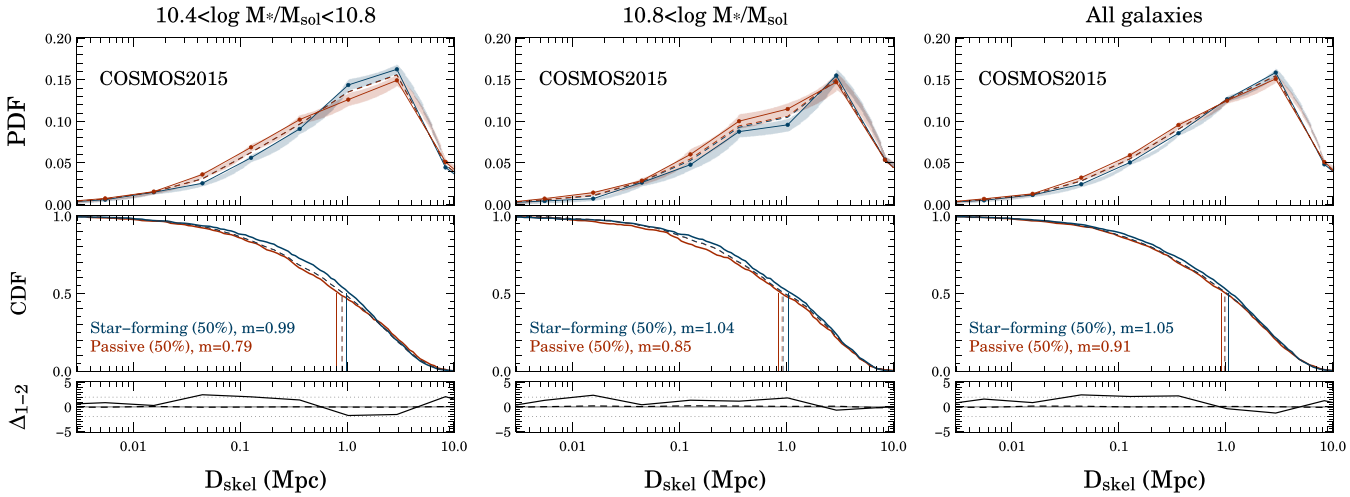

Figure 7. Differential (top) and cumulative (middle) distributions of the distances to the 2D photometric skeleton $\mathrm{SKL}_{\mathrm{2D}}^{\mathrm{phot}}$ for passive (red) and star-forming (blue) galaxies in COSMOS2015 after density matching. See caption Fig. 6 for more details.

Table 5. Summary of the results from the colour-type gradient signal towards filaments in COSMOS2015. The given values are the medians of the distributions and the standard errors from bootstrap resampling with 100 realizations expressed in projected comoving Mpc. The bracketed values indicate the percentage of galaxies within a mass bin in each population. The percentages are the same for the original and reshuffled samples. The reshuffled sample corresponds to a random signal preserving the galaxy type–density relation. The density-matching algorithm ensures to isolate galaxy subsets with the same density distribution in each mass bin.

	Median values of the PDF (COSMOS2015)					
	Original sample		Reshuffled samples		Density-matched samples	
Bin	SF	Passive	SF	Passive	SF	Passive
$10 < \log M_*/\mathrm{M}_\odot < 10.4$	0.96 ± 0.02 [72]	0.87 ± 0.03 [28]	0.95 ± 0.03	0.93 ± 0.02	1.16 ± 0.05 [50]	1.04 ± 0.06 [50]
$10.4 < \log M_*/\mathrm{M}_\odot < 10.8$	0.90 ± 0.03 [57]	0.64 ± 0.02 [43]	0.82 ± 0.04	0.75 ± 0.02	0.99 ± 0.02	0.79 ± 0.03
$10.8 < \log M_*/\mathrm{M}_\odot < 13$	0.85 ± 0.06 [34]	0.59 ± 0.03 [66]	0.74 ± 0.06	0.64 ± 0.03	1.04 ± 0.08	0.85 ± 0.06
All	0.93 ± 0.02 [58]	0.67 ± 0.02 [42]	0.87 ± 0.02	0.75 ± 0.02	1.05 ± 0.03	0.91 ± 0.03

previously. The lower panel displays the residual between the passive and the star-forming distributions, expressed in number of σ: $\Delta_{1-2} = \Delta(D_1 - D_2)/\sqrt{(\sigma_1^2 + \sigma_2^2)}$, where the subscripts 1 and 2 refer to the passive and star-forming distributions, respectively. Passive galaxies are found on average statistically closer to the filament centre than star-forming ones when considering the entire population. This result remains true and significant (above 2σ, except for the lowest mass bin $10^{10} < M_*/\mathrm{M}_\odot < 10^{10.4}$) when measuring the trend in narrow stellar mass bins. As the contribution of galaxies in nodes is minimized, this effect is specifically related to filaments.

However, in accordance with the above-mentioned mass–density relation, it has long been established that low-redshift galaxies are found to be much less star forming in high-density regions relative to low-density regions. As the density is anisotropically distributed in filaments, the density itself could explain these colour-driven gradients towards filaments. To disentangle the effect of the geometry of the large-scale environment from the galaxy type–density relation, we proceed as previously. In each slice and for each mass bin, the galaxy catalogue is split in local density bins (10 density bins equipopulated). Then galaxy types are randomized within these bins (preserving the fraction of passive and star-forming galaxies in each bin). After this reshuffling operation, the gradients are much less significant, whatever the mass bin. We conclude that the colour segregation within the filaments is not driven by the local density only: the geometry of the environment, measured from the proximity to the filament, plays a distinct role.

To confirm this result, we perform as previously a density matching test. In each mass and density bins (4 mass bins and 15 equipopulated density bins in each), the same number of galaxies is selected from both the star-forming and passive populations. This selection ensures that the sub-samples of star-forming and passive galaxies share the same density distribution. The same measurement is then performed on these subsamples, and the result is displayed in Fig. 7 for the two more massive bins and all the galaxies (with $10^{10} < M_*/\mathrm{M}_\odot$) and the right column of Table 5. After density matching, a signal is still measured. The signal is weaker, in particular because the subsamples contain much less galaxies, but still of the order of 2σ. This result confirms the specific role played by the presence of the filament.

4.3 sSFR gradients towards filaments in HORIZON-AGN

Let us now look for the segregation between passive and star-forming galaxies in the simulation. Passive and star-forming galaxies are selected as the 45 per cent least and most star-forming galaxies, respectively, in each mass bin.

We then perform the same measurements as for the observed sample. This measurement is carried here in three dimensions with the reference skeleton, in order to obtain a signal as free of noise and systematics as possible. A signal quantitatively identical to the COSMOS measurement is not expected, given the different ways of selecting the passive and star-forming populations. However, qualitatively, passive galaxies are expected to be closer to the

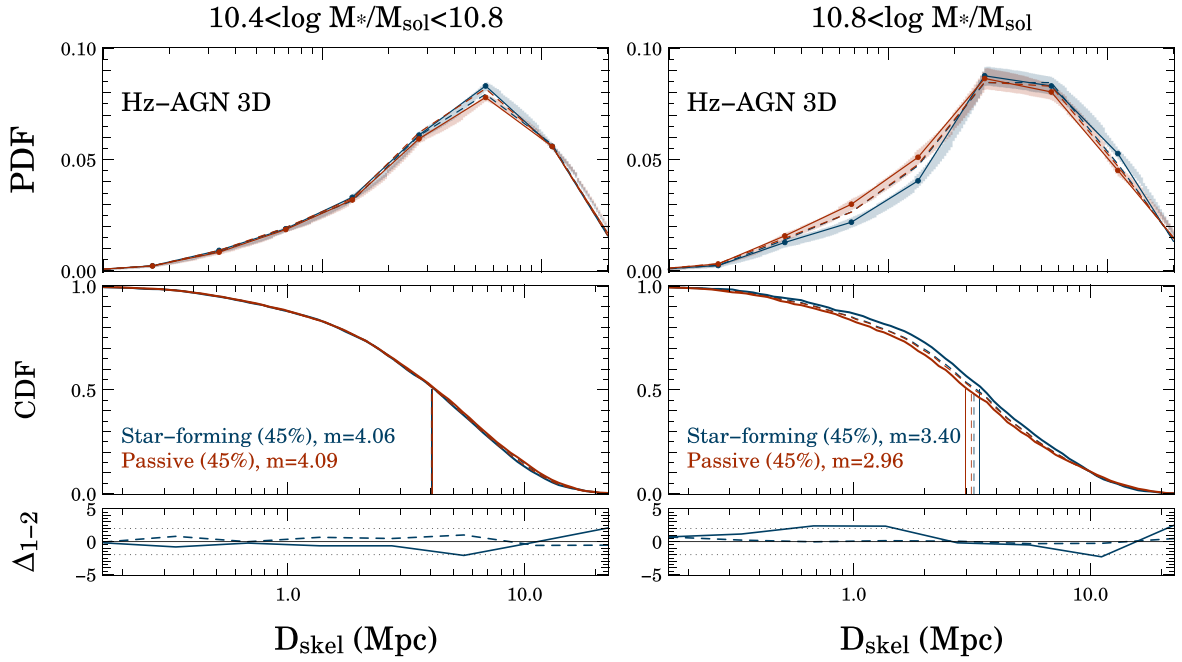

Figure 8. Differential (*top*) and cumulative (*middle*) distributions of the distances to the 3D reference skeleton SKL_{3D}^{DM} for passive (red) and star-forming (blue) galaxies in Horizon-AGN in the two upper mass bins. Passive and star-forming galaxies are selected as the 45 per cent least and most star-forming galaxies respectively in each mass bin. See caption Fig. 6 for more details.

Table 6. Summary of the results from the colour-type gradient signal towards filaments in Horizon-AGN using SKL_{3D}^{DM}. The given values are the medians of the distributions and the standard errors from bootstrap resampling with 100 realizations expressed in comoving Mpc. The bracketed values indicate the percentage of galaxies within a mass bin in each population. The percentages are the same for the original and reshuffled samples. The reshuffled sample corresponds to a random signal preserving the galaxy type–density relation. The density-matching algorithm ensures to isolate galaxy subsets with the same density distribution in each mass bin.

	Median values of the PDF (HORIZON-AGN)					
	Original sample		Reshuffled samples		Density-matched samples	
Bin	SF	Passive	SF	Passive	SF	Passive
$10 < \log M_*/\mathrm{M}_\odot < 10.4$	4.70 ± 0.05 [45]	4.50 ± 0.05 [45]	4.57 ± 0.03	4.53 ± 0.05	4.72 ± 0.06	4.53 ± 0.06
$10.4 < \log M_*/\mathrm{M}_\odot < 10.8$	4.06 ± 0.07 [45]	4.09 ± 0.06 [45]	4.06 ± 0.09	4.03 ± 0.05	4.00 ± 0.08	4.16 ± 0.08
$10.8 < \log M_*/\mathrm{M}_\odot < 13$	3.40 ± 0.09 [45]	2.96 ± 0.10 [45]	3.22 ± 0.10	3.14 ± 0.07	3.33 ± 0.10	2.92 ± 0.08

filament than star-forming ones. The result in different mass bins is presented in Fig. 8 for the two most massive bins and in Table 6. Passive galaxies are found closer to the filament than star-forming ones in the most massive bin, but no signal is found in the intermediate mass bin. From the median of the distributions, there is a slight signal for passive galaxies closer to the filament in the lowest mass bin ($10.0 < \log M_*/\mathrm{M}_\odot < 10.4$). After density matching, the signal is preserved for the most massive bin; however, the trend is reversed for the intermediate mass bin: star-forming galaxies with $10.4 < \log M_*/\mathrm{M}_\odot < 10.8$ are closer to the filament than passive galaxies for the density-matched sample. This seems to indicate a competitive role of the proximity to the filament and the density, and a mass-dependent effect of the filament on galaxies, which can be understood when considering the interplay between accretion and AGN feedback (see Musso et al. 2017; Kraljic et al. 2018, for a discussion).

These measurements are comforting the results from COSMOS2015 and highlight once again the role played by the geometry of the anisotropic environment in driving galaxy properties. In the simulation, the strength of the signal is dependent on the exact sSFR and age cut chosen for the galaxy-type separation (not shown here), as well as the robustness of the filaments included in the analysis (which indicates that the signal is probably multiscale). A more detailed analysis of the galaxy property gradients in the simulation taking into account these variations will be the topic of future work.

4.4 Interpretation and discussion

Our results are quantitatively summarized in Tables 3–6 that report the median values of each distribution. The normalization of these values is quite sensitive to cosmic variance and to the chosen thickness of the slices for the 2D reconstruction. Hence, they have to be taken cautiously when comparing to other data sets. However, the significant difference between the medians of the different populations, either divided in mass bins, in sSFR bins, or in terms of colour types, is a robust result that allows us to report mass gradients towards filaments in COSMOS2015 and Horizon-AGN, and colour-type gradients towards filaments in COSMOS2015. In Horizon-AGN, we also find sSFR gradients in agreement with COSMOS2015 for the most massive bin. They corroborate results found in other fields by independent methods of analysis (Alpaslan

et al. 2016; Chen et al. 2017; Malavasi et al. 2017; Kraljic et al. 2018; Poudel et al. 2017).

Theoretical considerations together with measurements both in observations and in simulations give us clues to how galaxy properties depend on the geometry of their environment. Galaxy harassment (Moore et al. 1996), ram-pressure stripping of the gas (Gunn & Gott 1972; Quilis, Moore & Bower 2000), strangulation of low-mass satellite galaxies (Larson, Tinsley & Caldwell 1980; Balogh, Navarro & Morris 2000), or cumulative galaxy–galaxy hydrodynamic/gravitational interactions (Park & Hwang 2009) may explain the depletion or heating of the cold gas in quenched galaxies in group or cluster environment, i.e. mainly at the node of the cosmic web. At intermediate density ranges, galaxy mergers, while temporarily modestly increasing star formation (e.g. Jogee et al. 2009; Kaviraj et al. 2015), may also halt star-formation and turn star-forming galaxies into passive ones (e.g. Dubois et al. 2016). But none of these processes can be responsible for galaxy segregation depending only on distance to filaments taking place in regions of low-to-intermediate densities, where galaxy interactions are limited.

Since the signal found in this work is sensitive to the large-scale geometry of the matter distribution itself, it must be driven by the large-scale (traceless part of the) tidal tensor, which quantifies this anisotropy. The physical properties of dark matter haloes (Codis et al. 2012) and their host galaxies (Dubois et al. 2014; Welker et al. 2014) indeed correlate with the *dynamics* of their embedding anisotropic large-scale structure. This has identified kinematic signatures: angular momentum advection (Codis et al. 2015b) and internal kinetic anisotropy (Faltenbacher & White 2009) are also set by the larger scale tides of the environment. The induced tides not only impact the merger tree and accretion history of the host but also the filamentary flow of cold gas connecting to the host, hence its coherent gas supply. The past and present efficiency of star formation, as traced by the mass, colours and specific star formation rate, depends critically on the infalling rate and impact parameter of cold gas. It does not only depend on the timing of infall (in connection with the accretion history of the host) but also on its geometry and physical content, in connection with the dynamics of the matter flows (e.g. in plane gas-rich and corotating, or dry mergers with random orbital parameters). As discussed in the introduction, the vorticity-rich large-scale filaments are indeed the locus where low-mass galaxies steadily grow in mass via quasi-polar cold gas accretion (Welker et al. 2015), with their angular momentum aligned with the host filament. Galaxies are expected to accrete more efficiently cold gas when their angular momentum is aligned with the preferential direction of the gas infall, i.e. aligned with the filament (Pichon et al. 2011). Hence, one expects star formation efficiency to be strongest wherever the alignment is tightest. Maximal alignment occurs in the highest vorticity regions (Laigle et al. 2015), i.e. at the edge of filaments (whereas haloes terminate their mass assembly in the filament core while drifting towards nodes following the large-scale flow). The locus of this induced excess of star formation should therefore have measurable signatures in observations when quantified in the metric of the filament, as discussed e.g. in Codis et al. (2015b). This is consistent with what is observed in Figs 6 and 8.

In the present study, we have attempted to disentangle the specific signature of this anisotropy while relying on the physical properties of galaxies at fixed mass. The various tests presented in this paper confirm that the anisotropy of the environment traced by transverse gradients with respect to filaments is detected in COSMOS2015 catalogue, as anticipated and quantified from mocks derived from the light-cone of Horizon-AGN. We provide observational evidence that what has been a pillar of galaxy formation, i.e. the assumption that the average effects of the environment can be approximated by the effect of the (in particular azimuthally) average environment is only valid at first order. This assumption is obviously strictly true for linear dynamic and linear observables, which does not capture the full complexity of what builds up the properties of galaxies. This past simplification was partially driven by the community's strong focus on two-point functions, which cannot capture the anisotropy of the cosmic web. Yet it had long been known that for instance angular momentum – which is undoubtedly built from anisotropic tides – is a key underlying property driving morphology, which correlates strongly with colour and star formation efficiency. Hence, one could attempt to extend, e.g. assembly bias theory (Sheth & Tormen 2004) to try and explain the observed diversity at a given mass, so long as it is understood that mass assembly is in fact also driven by anisotropic large-scale tides, which will impact gas inflow towards galaxies, hence their properties (Yan et al. 2013; Tramonte et al. 2017).

More elaborate – but still restricted in scope – bias theories such as excursion set peak theories involving a moving barrier account for the other invariants of the tidal tensor (see Desjacques, Jeong & Schmidt 2016, and reference therein), but fail to incorporate the effect of the orientation of its eigen vectors. Semi-analytical models (e.g. Lacey & Silk 1991) do account for the timings encoded in merging trees measured in simulations– which reflect in part the cosmic web – but are restricted to following dark halo only and typically leave out the configuration of their mergers.

An improved model for galaxy properties should therefore explicitly integrate the diversity of the topology of the large environment on multiple scales (following, e.g. Hanami 2001) and quantify the impact of its anisotropy on galactic mass assembly history, and more generally on the kinematic history of galaxies (Musso et al. 2017). What also remains to be understood is what process dominates at what redshift: does the offset of merger and accretion rate imposed by the large-scale turbulent flow explain most of the environment dependence in observed physical properties (Aragon-Calvo, Neyrinck & Silk 2016), or is the geometry of gas inflow within filaments prevalent in feeding galactic discs coherently (Pichon et al. 2011; Stewart et al. 2011)? The kinematic of the large-scale flow is neither strictly coherent nor fully turbulent. Both processes will, in turn, modify the efficiency of star formation and feedback, which is critically non-linear, currently poorly understood and fine tuned. In order to disentangle these effects, the next step will involve studying the connection between galaxy properties, angular momentum, and distance to filaments in the Horizon-AGN simulation and in connection with the COSMOS2015 catalogue.

5 CONCLUSION

Using the photometric catalogue COSMOS2015, we have investigated how galaxy properties depend on their anisotropic environment. A realistic photometric catalogue extracted from the Horizon-AGN hydrodynamical simulation was used to quantify our ability to recover the filamentary structures from 2D slices at the same photometric redshift precision and to predict the expected gradients. The main findings of this work are the following:

(i) Photometric surveys with a photometric precision similar to COSMOS2015 offer the prospect to reliably study the 3D properties of cosmic web from projected 2D slices.

(ii) We observe both in HORIZON-AGN and in COSMOS2015 a significant tendency for massive galaxies to be closer to the filament centre than for less massive galaxies.

(iii) We also observe segregation of passive and star-forming galaxies near the filaments both in COSMOS2015 and in HORIZON-AGN for galaxies with a mass $\log M_*/\mathrm{M}_\odot > 10.4$. At fixed mass, passive galaxies are more confined in the core of the filament than star-forming ones. These two signals persist when minimizing the contribution of nodes and cannot be explained by the local mass–density relation alone.

These findings are qualitatively consistent between simulation and data, and in agreement with studies relying on spectroscopic redshifts (e.g. Malavasi et al. 2017). They underline the specific role played by the anisotropy of the large-scale cosmic web in shaping galaxies properties. The geometry of the large-scale environment drives tides, which impact the merger and accretion history of galaxies, their spin, and in turn their observed properties. Mass and density alone are not sufficient to characterize the observed properties of galaxies away from the cosmic web.

More observations are needed to disentangle the effects of all these competing processes, while relying on a wider fields and (photometric) redshift baseline, towards the peak of the cosmic star formation, where cold streams are expected to be more efficient. This will be challenging as one will need to carefully disentangle the evolutionary effects and selection biases. An alternative observational strategy is to rely on ongoing (SAMI and MANGA Croom et al. 2012; Bundy et al. 2015) or upcoming (Hector, Bland-Hawthorn 2015) integral field spectroscopy surveys and explore spin properties of galaxies (e.g. Harrison et al. 2017) such as their physical and kinematic axis orientation as a function of the large-scale environment. Simulated end-to-end catalogues such as those produced here will remain key to qualify these investigations.

ACKNOWLEDGEMENTS

CL was supported by the ILP LABEX (under reference ANR-10-LABX-63 and ANR-11-IDEX-0004-02) during the first part of this work and is now supported by a Beecroft Fellowship. JD and AS acknowledges funding support from Adrian Beecroft, the Oxford Martin School, and the Science and Technology Facilities Council (STFC). OI acknowledges the funding of the French Agence Nationale de la Recherche for the project 'SAGACE'. This work relied on the HPC resources of CINES (Jade) under the allocation 2013047012 and c2014047012 made by GENCI and on the Horizon Cluster hosted by Institut d'Astrophysique de Paris. We warmly thank S. Rouberol for running the cluster on which the simulation was post-processed. This research is part of Spin(e) (ANR-13-BS05-0005, http://cosmicorigin.org) and horizon-UK. This research is also partly supported by the Centre National d'Etudes Spatiales (CNES). This work is based on data products from observations made with ESO Telescopes at the La Silla Paranal Observatory under ESO programme ID 179.A-2005 and on data products produced by TERAPIX and the Cambridge Astronomy Survey Unit on behalf of the UltraVISTA consortium. We thank D. Munro for freely distributing his YORICK programming language and OpenGL interface (available at http://yorick.sourceforge.net/).

REFERENCES

Alonso D., Eardley E., Peacock J. A., 2015, MNRAS, 447, 2683
Alpaslan M. et al., 2016, MNRAS, 457, 2287
Aragon-Calvo M. A., Yang L. F., 2014, MNRAS, 440, L46
Aragón-Calvo M. A., Jones B. J. T., van de Weygaert R., van der Hulst J. M., 2007, A&A, 474, 315
Aragon-Calvo M. A., Weygaert R. v. d., Jones B. J. T., Mobasher B., 2015, MNRAS, 454, 463
Aragon-Calvo M. A., Neyrinck M. C., Silk J., 2016, MNRAS, preprint (arXiv:1607.07881)
Arnouts S. et al., 2002, MNRAS, 329, 355
Arzoumanian D. et al., 2011, A&A, 529, L6
Aubert D., Pichon C., Colombi S., 2004a, MNRAS, 352, 376
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
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Bland-Hawthorn J., 2015, in Ziegler B. L., Combes F., Dannerbauer H., Verdugo M., eds, Proc. IAU Symp. 309, Galaxies in 3D across the Universe. p. 21
Blanton M. R. et al., 2003, ApJ, 594, 186
Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517
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
Bruzual G., Charlot S., 2003, MNRAS, 344, 1000
Bundy K. et al., 2015, ApJ, 798, 7
Capak P. et al., 2007, ApJS, 172, 99
Cardamone C. N. et al., 2010, ApJS, 189, 270
Castorina E., Paranjape A., Hahn O., Sheth R. K., 2016, preprint (arXiv:1611.03619)
Chen Y.-C. et al., 2017, MNRAS, 466, 1880
Chisari N. et al., 2015, MNRAS, 454, 2736
Chisari N. et al., 2016, MNRAS, 461, 2702
Codis S., Pichon C., Devriendt J., Slyz A., Pogosyan D., Dubois Y., Sousbie T., 2012, MNRAS, 427, 3320
Codis S. et al., 2015a, MNRAS, 448, 3391
Codis S., Pichon C., Pogosyan D., 2015b, MNRAS, 452, 3369
Cole S., Kaiser N., 1989, MNRAS, 237, 1127
Croom S. M. et al., 2012, MNRAS, 421, 872
Croton D. J., Gao L., White S. D. M., 2007, MNRAS, 374, 1303
Cucciati O. et al., 2016, A&A, 602, A15
Dalal N., White M., Bond J. R., Shirokov A., 2008, ApJ, 687, 12
Darvish B., Mobasher B., Sobral D., Rettura A., Scoville N., Faisst A., Capak P., 2016, ApJ, 825, 113
Darvish B., Mobasher B., Martin D. C., Sobral D., Scoville N. Z., Stroe A., Hemmati S., Kartaltepe J., 2017, ApJ, 837, 16
Davidzon I. et al., 2016, A&A, 586, A23
Davidzon I. et al., 2017, A&A, 605, A70
Davis M., Geller M. J., 1976, ApJ, 208, 13
Desjacques V., Jeong D., Schmidt F., 2016, preprint (arXiv:1611.09787)
Dressler A., 1980, ApJ, 236, 351
Dubois Y., Devriendt J., Slyz A., Teyssier R., 2012, MNRAS, 420, 2662
Dubois Y., Pichon C., Welker C. et al., 2014, MNRAS, 444, 1453
Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, MNRAS, 463, 3948
Efstathiou G., Frenk C. S., White S. D. M., Davis M., 1988, MNRAS, 235, 715
Faltenbacher A., White S. D. M., 2009, ApJ, 708, 469
Gao L., Springel V., White S. D. M., 2005, MNRAS, 363, L66
González R. E., Prieto J., Padilla N., Jimenez R., 2017, MNRAS, 464, 4666
Gunn J. E., Gott III J. R., 1972, ApJ, 176, 1
Haardt F., Madau P., 1996, ApJ, 461, 20
Hahn O., Carollo C. M., Porciani C., Dekel A., 2007, MNRAS, 381, 41
Hahn O., Porciani C., Dekel A., Carollo C. M., 2009, MNRAS, 398, 1742
Hanami H., 2001, MNRAS, 327, 721
Harrison C. M. et al., 2017, MNRAS, 467, 1965
Hearin A. P., Behroozi P. S., van den Bosch F. C., 2016, MNRAS, 461, 2135
Hirv A., Pelt J., Saar E., Tago E., Tamm A., Tempel E., Einasto M., 2017, A&A, 599, A31

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581
Ilbert O. et al., 2006, A&A, 457, 841
Ilbert O. et al., 2009, ApJ, 690, 1236
Ilbert O. et al., 2013, A&A, 556, A55
Ivezic Z. et al., 2008, preprint (arXiv:0805.2366)
Jogee S. et al., 2009, ApJ, 697, 1971
Kaiser N., 1984, ApJ, 284, L9
Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, MNRAS, 353, 713
Kaviraj S., Devriendt J., Dubois Y., Slyz A., Welker C., Pichon C., Peirani S., Le Borgne D., 2015, MNRAS, 452, 2845
Kaviraj S. et al., 2017, MNRAS, 467, 4739
Kawinwanichakij L. et al., 2016, ApJ, 817, 9
Komatsu E. et al., 2011, ApJS, 192, 18
Kraljic K. et al., 2018, MNRAS, 474, 547
Lacey C., Silk J., 1991, ApJ, 381, 14
Laigle C. et al., 2015, MNRAS, 446, 2744
Laigle C. et al., 2016, ApJS, 224, 24
Larson R. B., Tinsley B. M., Caldwell C. N., 1980, ApJ, 237, 692
Libeskind N. I., Hoffman Y., Forero-Romero J., Gottlöber S., Knebe A., Steinmetz M., Klypin A., 2013, MNRAS, 428, 2489
Lilly S. J. et al., 2007, ApJS, 172, 70
Malavasi N. et al., 2017, MNRAS, 465, 3817
Martínez H. J., Muriel H., Coenda V., 2016, MNRAS, 455, 127
McCracken H. J. et al., 2012, A&A, 544, A156
Miyazaki S. et al., 2012a, in Proc. SPIE Conf. Ser. Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460Z
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, Nature, 379, 613
Musso M., Cadiou C., Pichon C., Codis S., Kraljic K., Dubois Y., 2017, MNRAS, preprint (arXiv:1709.00834)
Navarro J. F., Frenk C. S., White S. D. M., 1995, MNRAS, 275, 56
Oemler A., Jr, 1974, ApJ, 194, 1
Panopoulou G. V., Psaradaki I., Skalidis R., Tassis K., Andrews J. J., 2017, MNRAS, 466, 2529
Park C., Hwang H. S., 2009, ApJ, 699, 1595
Pichon C., Thiébaut E., Prunet S., Benabed K., Colombi S., Sousbie T., Teyssier R., 2010, MNRAS, 401, 705
Pichon C., Pogosyan D., Kimm T., Slyz A., Devriendt J., Dubois Y., 2011, MNRAS, 418, 2493
Pogosyan D., Pichon C., Gay C., Prunet S., Cardoso J. F., Sousbie T., Colombi S., 2009, MNRAS, 396, 635
Poudel A., Heinämäki P., Tempel E., Einasto M., Lietzen H., Nurmi P., 2017, A&A, 597, A86
Pozzetti L. et al., 2010, A&A, 523, A13
Press W. H., Schechter P., 1974, ApJ, 187, 425
Quilis V., Moore B., Bower R., 2000, Science, 288, 1617
Salpeter E. E., 1955, ApJ, 121, 161
Schaap W. E., van de Weygaert R., 2000, A&A, 363, L29
Scoville N. et al., 2007, ApJS, 172, 1
Scoville N. et al., 2013, ApJS, 206, 3
Sheth R. K., Tormen G., 1999a, MNRAS, 308, 119
Sheth R. K., Tormen G., 1999b, MNRAS, 308, 119
Sheth R. K., Tormen G., 2004, MNRAS, 350, 1385
Sousbie T., 2011, MNRAS, 414, 350
Sousbie T., Pichon C., Kawahara H., 2011, MNRAS, 414, 384
Stewart K. R., Kaufmann T., Bullock J. S., Barton E. J., Maller A. H., Diemand J., Wadsley J., 2011, ApJ, 738, 39
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
Taniguchi Y. et al., 2007, ApJS, 172, 9
Taniguchi Y. et al., 2015, PASJ, 67, 104
Tempel E., Libeskind N. I., 2013, ApJ, 775
Teyssier R., 2002, A&A, 385, 337
Tramonte D., Rubino-Martin J. A., Betancort-Rijo J., Dalla Vecchia C., 2017, MNRAS, 467, 3424
Tweed D., Devriendt J., Blaizot J., Colombi S., Slyz A., 2009, A&A, 506, 647
Wang J. et al., 2011, MNRAS, 413, 1373
Wechsler R. H., Zentner A. R., Bullock J. S., Kravtsov A. V., Allgood B., 2006, ApJ, 652, 71
Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., 2006, MNRAS, 366, 2
Welker C., Devriendt J., Dubois Y., Pichon C., Peirani S., 2014, MNRAS, 445, L46
Welker C., Dubois Y., Pichon C., Devriendt J., Chisari E. N., 2015, A&A, preprint (arXiv:1512.00400)
Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, ApJ, 691, 1879
Yan H., Fan Z., White S. D. M., 2013, MNRAS, 430, 3432

APPENDIX A: IMPACT OF MASKS

The observed catalogue used in this work contains masks over the field. The presence of these masks biases the density distribution and the subsequent skeleton extraction. In the work presented in this paper, masks similar to the COSMOS field have been applied to the simulated catalogue. In the following, the impact of these masks is quantified. Fig. A1 presents the skeleton extraction with and without the masks at $z \sim 0.73$. The red skeleton is the one using the masked catalogue, while the black skeleton is the one using the entire skeleton, for the same mass cut. The green and blue skeleton is the reference skeleton SKL_{3D}^{DM}. As it can been seen in this figure, the masked skeleton misses some 3D filaments. However, it allows to recover some less prominent filaments that were not detected at a given persistent threshold in the non-masked skeleton (e.g. the filament in [RA, Dec.] = [0.45,

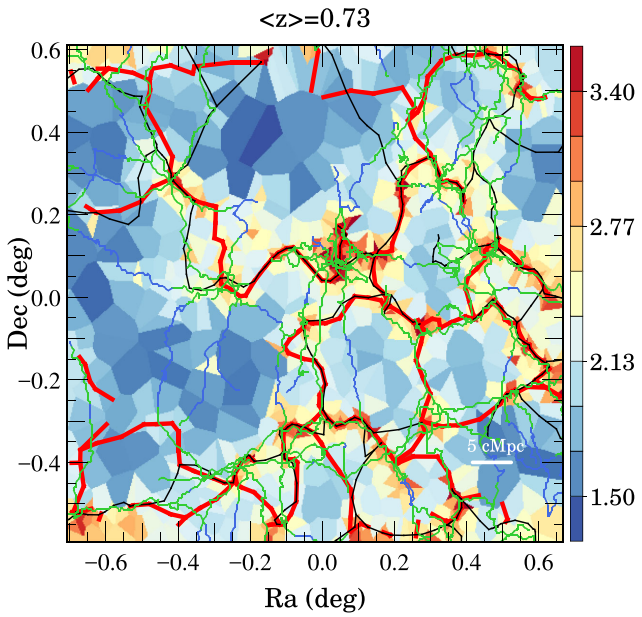

Figure A1. The comparison between the spectroscopic skeleton SK_{2D}^{spec} using the masked (red line) or non-masked (black line) catalogues and the projected DM SKL_{3D} (green and blue) skeletons. Most of the filaments in the masked and non-masked skeletons are similar. The masked skeleton misses some filaments that were correctly recovered in the non-masked skeleton. However, the new filaments that are present in the masked skeleton and not in the non-masked one have generally 3D counterparts, except in the underdense regions. The horizontal white bar in the lower right indicates the comoving scale length of 5 Mpc.

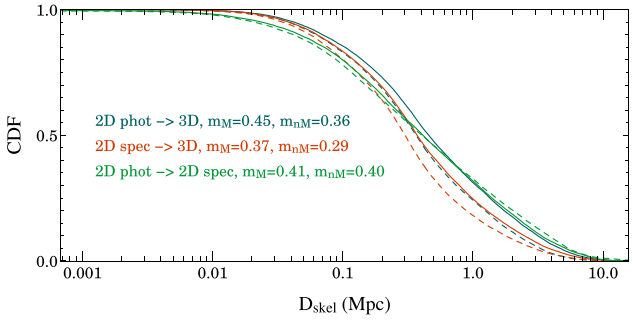

Figure A2. Cumulative distributions of the distances between SKL_{2D}^{phot} and SKL_{3D}^{DM} (blue curve), SKL_{2D}^{spec} and SKL_{3D}^{DM} (red curve) and SKL_{2D}^{phot} and SKL_{2D}^{spec} (green curve) using the masked (solid line) and unmasked (dashed line) simulated catalogues. The m_M and m_{nM} values specify the medians of the distribution for the masked and non-masked skeletons, respectively.

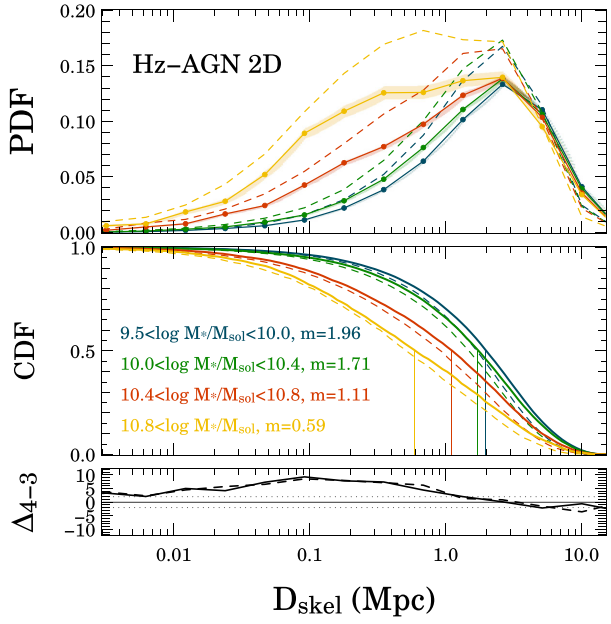

Figure A3. Differential (top) and cumulative (middle) distributions of the distances to the 2D photometric skeleton SKL_{2D}^{phot} for galaxies as a function of their masses in the Horizon-AGN lightcone for $0.5 < z < 0.9$. The dashed distributions correspond to the distance to the masked skeleton, while the solid distributions correspond to the distances to the non-masked skeleton. The m values give the medians of distribution for the non-masked skeleton, which are also indicated by vertical lines. The contribution of nodes to the gradient signal has been removed from the analysis to highlight an effect specifically related to the filaments. The bottom row shows the residuals between the two more massive bins ($10.4 < \log M_*/\mathrm{M}_\odot < 10.8$ and $0.8 < \log M_*/\mathrm{M}_\odot < 13.0$) for the non-masked (solid line) and masked (dashed line) samples.

-0.27]). Hence, there are only few new fake filaments in the masked skeleton, and these new filaments generally lay in the underdense regions.

Fig. A2 shows the comparison of the distances between the 2D skeletons and SKL_{3D}^{DM} when applying these masks. When applying the masks, the typical distance between SKL_{2D}^{phot} and SKL_{3D}^{DM}increases only by ∼0.09 projected Mpc, which is essentially due the fact that the number of unmatched 2D segments increases by a factor of ∼1.3. In overall, the agreement between SKL_{2D}^{phot} and SKL_{3D}^{DM} remains good.

Finally, Fig. A3 presents the mass gradient signal when using the masked or un-masked catalogue for the skeleton reconstruction. The distributions with and without the masks are not identical, but the cumulative distributions and the strength of the signal remain similar.

APPENDIX B: CHOICE OF PERSISTENCE IN TWO DIMENSION

As stated in Section 2.3, the extraction of skeletons with the same persistence thresholds σ in two and three dimensions will typically not lead to the detection of the same filaments. Indeed, one must significantly decrease σ with respect to its 3D value to recover in two dimensions most of the 3D filaments. Let us estimate the optimal 2D σ that recovers most of the filaments without introducing spurious detections. This optimal σ is estimated from the simulated catalogue of Horizon-AGN, by comparing the number of unmatched segments between SKL_{2D}^{spec} and SKL_{3D}^{DM}. We compute first the distances between the two skeletons. In order to compare equitably the percentage of unmatched segments for different slice thicknesses, the distances between the two skeletons are divided by the mean 2D inter-galaxy distance D_{int} (which increases with the slice thickness), taken as the square root of the area divided by the number of galaxies in the slice. Then we define the percentage of unmatched segments as the fraction of segments such as their closest counterpart is further than D_{int} (changing this threshold would move Fig. B1 vertically). The percentage of unmatched segments is shown in Fig. B1. In order to reduce as much as possible the number of SKL_{3D}^{DM} segments that do not have a counterpart in SKL_{2D}^{spec}, without sensitively increasing spurious filament extraction in SKL_{2D}^{spec}, we find that a 2σ persistent threshold for SKL_{2D}^{spec} offers a reasonable balance between the percentage of unmatched segments in SKL_{2D}^{spec} and SKL_{3D}^{DM}, whatever the slice thickness. Although SKL_{2D}^{spec} is computed for Fig. B1 with a mass limit of $10^{10}\,\mathrm{M}_\odot$, this result does not significantly change with a different mass limit.

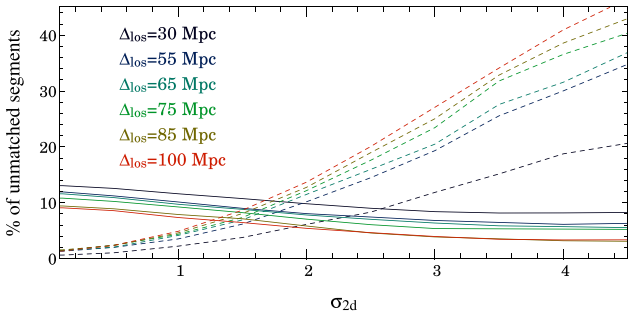

Figure B1. Percentage of unmatched segments between SKL_{3D}^{DM} and SKL_{2D}^{spec} for different slice thicknesses as a function of the persistence threshold used for the computation of SKL_{2D}^{spec}. The solid lines refer to the percentage of SKL_{2D}^{spec} segments that do not have a counterpart in SKL_{3D}^{DM}, and the dashed lines refer to the percentage of SKL_{3D}^{DM} segments that do not have a counterpart in SKL_{2D}^{spec}. The mass limit used for the computation of SKL_{2D}^{spec} is $10^{10}\mathrm{M}_\odot$, but the trend remains in overall the same when this limit varies.

APPENDIX C: GALAXY VERSUS DM SKELETONS

The skeleton computed from dark matter particles is the most suitable to stand as the reference skeleton, since the galaxy distribution is a biased tracer of the underlying density field. In this appendix, we compare the 3D galaxy and dark matter skeletons and we highlight how their relative difference depends on the mass limit of the galaxy sample. To get values directly comparable to Fig. 3, we measure the distances between the galaxy and dark matter skeletons in projection, although both skeletons are computed in three dimensions.

The distribution of the distance between the dark matter and galaxy skeletons SKL_{3D}^{gal} and SKL_{3D}^{DM} are plotted in Fig. C1 for different

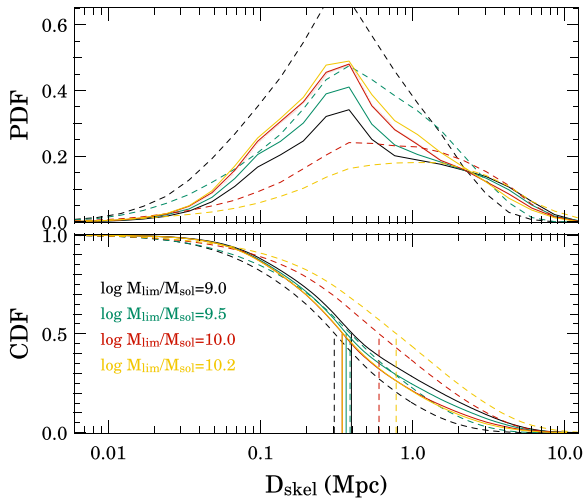

Figure C1. Differential (top) and cumulative (bottom) distributions of the distances from SKL_{3D}^{gal} towards SKL_{3D}^{DM} (solid lines) and conversely (dashed lines) calculated with different mass limits ranging from $10^9\,\mathrm{M}_\odot$ to $10^{10.2}\,\mathrm{M}_\odot$.

mass limits. The solid lines refer to the distance from the galaxy skeleton towards the dark matter one, while the dashed lines refer to converse. All solid lines are almost overlapping. These distribution are peaked at very close separation, but still display a long tail. This suggests that whatever the chosen mass limits for the galaxy skeleton reconstruction, the galaxy filaments have generally a very close counterpart in the dark matter skeleton, although a small fraction has no counterpart at all. We note a slight improvement in recovering the DM skeleton when working with increasing masses. However, when decreasing the mass limit of the galaxy sample, the number of unmatched filaments in the dark matter skeleton decreases strongly (the tail of the dashed distribution decreases). To summarize, the lower limiting mass, the better, but the final choice should be driven by the photometric redshift uncertainties.

APPENDIX D: CHOICE OF SLICE THICKNESS

We aim now to determine if there is an optimal slice thickness to extract the cosmic web and estimate transverse gradients. We proceed in two steps. First, given an ideal survey (without photometric redshift uncertainties), we estimate how the accuracy of the reconstruction varies with the thickness of the slices. Secondly, we estimate how the quality of the reconstruction varies as a function

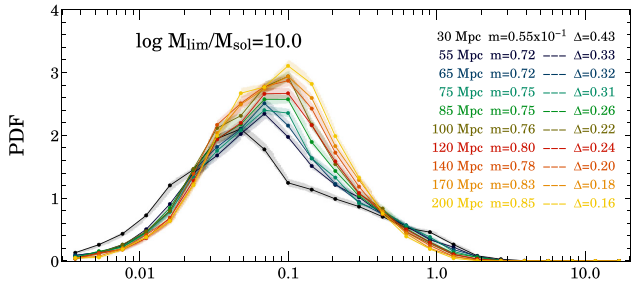

Figure D1. Distribution of the distance between SKL_{2D}^{spec} and SKL_{3D}^{DM}, for slices of different thicknesses ranging from 30 to 200 Mpc, with a limiting mass of $10^{10}\,\mathrm{M}_\odot$. Redshift uncertainties are not introduced here. The values m and Δ indicate the median and the width (68 per cent around the median) of the distributions. The distances are measured in all slices between $z \sim 0.5$ and $z \sim 0.9$.

of mass limits given a photometric redshift sample. Given the redshift accuracy of the COSMOS2015 catalogue, we choose then the optimal slice thickness and mass limit.

D1 Optimization of slice thickness with exact redshifts

The optimal slice thickness should lead to the best match between the 2D skeleton and its projected 3D counterpart. The agreement is quantified from the distribution of the distances between the projected SKL_{3D}^{DM} skeleton and SKL_{2D}. From the distribution, we measure the median and the width that encompasses 68 per cent of the distribution around the median. Good agreement between the skeletons requires both the median and the width to be small. Let us first test the single effect of the projection for different thicknesses, independently of the redshift accuracy.

The distances are normalized by the mean inter-galaxy distance D_{int}. Fig. D3 shows three slices of SKL_{2D}^{spec} and the projected SKL_{3D}^{DM}, corresponding to three slice thickness 30, 75, and 170 Mpc around redshift $z \sim 0.59$. In overall, the more robust 2D filaments trace existing 3D filaments. For the lowest thickness slice, most of the 3D filaments appear to have a counterpart in the 2D skeleton, but there are a lot of fake filaments, generally with a low robustness because the skeleton traces only fragments of filaments. For the highest thickness slice, this fraction of fake 2D filaments decreases strongly, but much more 3D filaments have no counterpart. We statistically quantified it in Fig. D1, which shows the distribution of the distances between SKL_{2D}^{spec} ($M_{\mathrm{lim}} = 10^{10}\,\mathrm{M}_\odot$) and SKL_{3D}^{DM}, for slices of different thicknesses ranging from 30 to 200 Mpc. It therefore appears that for exact redshifts, there is no optimal slice thickness within the considered range. Working with thin slices ensures accurate filaments identification, but a non-negligible fraction of 2D filaments have no counterpart at all in the 3D skeleton. This fraction diminish significantly when increasing the slice thickness, but at the price of less accurate filament identifications. In addition, as displayed by the black line in Fig. D2, the fraction of unmatched 3D segments increases significantly with the thickness. Hence, a reasonable choice could lie between 60 and 100 Mpc. But our final choice of the slice thickness will be in any case limited by the redshift accuracy of the sample. Finally, Fig. D3 indicates that filtering low robustness segments would be a way to remove the filaments that have probably no counterpart in three dimensions. This could be of interest when working with thicker slices. However, as the robustness is an estimation of the density of the filament with respect to the local background, keeping only the more robust filament would bias the measure towards the denser filaments. In addition, keeping

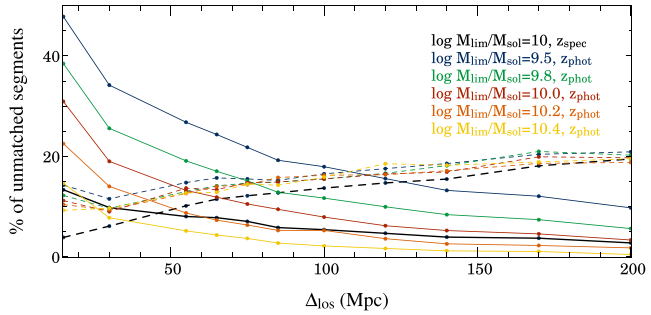

Figure D2. Percentage of unmatched segments within $D_{\rm int}$ between $\mathrm{SKL}_{\rm 3D}^{\rm DM}$ and $\mathrm{SKL}_{\rm 2D}^{\rm phot}$ as a function of the slice thicknesses used for the computation of $\mathrm{SKL}_{\rm 2D}^{\rm phot}$, for different mass limits (galaxies with lower mass have higher redshift uncertainties) for $0.5 < z < 0.9$. The solid lines refer to the percentage of $\mathrm{SKL}_{\rm 2D}^{\rm phot}$ segments that do not have a counterpart in $\mathrm{SKL}_{\rm 3D}^{\rm DM}$, and the dashed lines refer to the opposite. For comparison, the black line displays the distribution for $\mathrm{SKL}_{\rm 2D}^{\rm spec}$ computed with a mass limit of $10^{10}\,\mathrm{M}_{\odot}$.

only segments above a certain robustness threshold would destroy the connectivity of the cosmic web by removing entire sections of filaments. It would bias statistical measurements on the cosmic web, hence prevent accurate cosmological measurements based on such statistics. Note that although only the results for a mass limit of $10^{10}\,\mathrm{M}_{\odot}$ are shown in this Section, the trend is preserved when the mass limit of the sample varies.

D2 Robustness w.r.t. photometric redshift accuracy

We found (Fig. C1) that decreasing the mass limit of the sample is desirable to bring a better agreement between the galaxy and dark matter skeletons. At fixed mass limits, the optimal choice of the slice thickness broadly stands between 60 and 100 Mpc (Fig. D2), although increasing the slice thickness is possible at price of missing a larger number of 3D filaments. Let us now consider how the accuracy of the reconstruction is affected by the redshift uncertainties of the sample. This will drive our final choice both for the mass limit and the slice thickness, knowing that lower mass galaxies have higher redshift uncertainties. Fig. D2 shows the percentage of unmatched segments within $D_{\rm int}$ between $\mathrm{SKL}_{\rm 3D}^{\rm DM}$ and $\mathrm{SKL}_{\rm 2D}^{\rm phot}$ as a function of the slice thicknesses for different mass limits. For comparison, the black line displays the distribution for $\mathrm{SKL}_{\rm 2D}^{\rm spec}$ computed with a mass limit of $10^{10}\mathrm{M}_{\odot}$. When the mass limit decreases, more galaxies with high redshift uncertainties are included in the computation. It appears therefore that it is not reasonable to use a mass limit of $10^{9.5}\,\mathrm{M}_{\odot}$, as the fraction of unmatched 2D filaments is very consequent (always greater than 20 per cent for thickness lower than 100 Mpc). Increasing the mass limit to $10^{10}\,\mathrm{M}_{\odot}$ brings this fraction below 10 per cent, while the fraction of unmatched 3D segments do not increase significantly. Given the redshift uncertainties of COSMOS2015, we finally choose to work with galaxies more massive than $10^{10}\,\mathrm{M}_{\odot}$, which guarantees both that the galaxy skeleton is not strongly biased compared to the dark matter one and that the fraction of unmatched 2D segments is not too high. Within the allowing thickness range, we pick the value of 75 Mpc: it ensures that all the galaxies in the sample have their redshift uncertainties ($2\times 1\sigma$) lower than the thickness of the slice.

APPENDIX E: FILAMENTS OR NODES GRADIENTS?

Let us finally assess the robustness of the gradients towards filaments after removing the nodes. Nodes do indeed also drive mass gradients. These gradients could be either isotropic around the overdensity, or anisotropic along the filament, or a combination of both. In the first case, gradients towards nodes could mimic gradients towards filaments, even if they play no role in driving them. To remove the effect of nodes in this work, regions around nodes are excluded from the analysis, with an exclusion radius of 3.5 Mpc in three dimensions and 0.8 Mpc in two dimensions. These cuts are justified by Figs E1 and E2. The figures present gradients towards nodes with and without taking into account the galaxies in these exclusion regions in the simulation Horizon-AGN. First of all, let us notice that they are many more nodes in three dimensions than in two dimensions, which explain why the distributions look so different. When excluding galaxies in the vicinity of the nodes, some remaining gradients are measured. However, it is not possible to get totally rid off them, because he medians of the distribution are

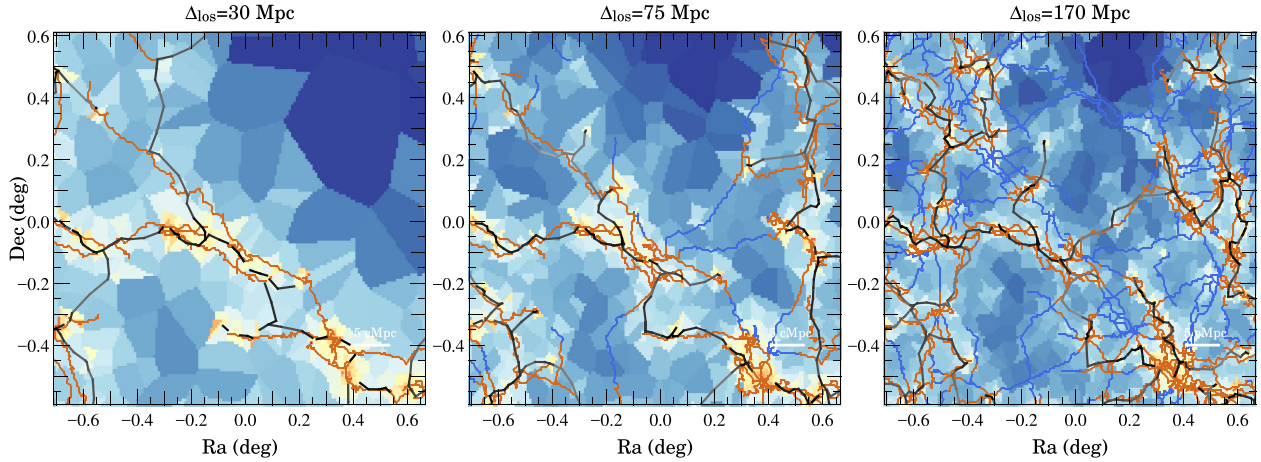

Figure D3. Skeleton reconstruction for different thicknesses ($\Delta_{\rm los}$) in the Horizon-AGN simulation around $z \sim 0.61$. The background density is computed from the Delaunay tessellation in two dimensions. The orange/blue skeleton is $\mathrm{SKL}_{\rm 3D}^{\rm DM}$, while the grey/black skeleton is $\mathrm{SKL}_{\rm 2D}^{\rm spec}$. The blue segments in $\mathrm{SKL}_{\rm 3D}^{\rm DM}$ are those which do not have a counterpart in $\mathrm{SKL}_{\rm 2D}^{\rm spec}$ within $0.5\times D_{\rm int}$, where $D_{\rm int}$ is the mean inter-galaxy distance. The grey-scale for $\mathrm{SKL}_{\rm 2D}^{\rm spec}$ encodes the robustness of the segments. More robust filaments are darker. In overall, the more robust segments in $\mathrm{SKL}_{\rm 2D}^{\rm spec}$ have a counterpart within $0.5\times D_{\rm int}$. Even with a large thickness, the robust filaments in $\mathrm{SKL}_{\rm 2D}^{\rm spec}$ represent existing 3D filaments.

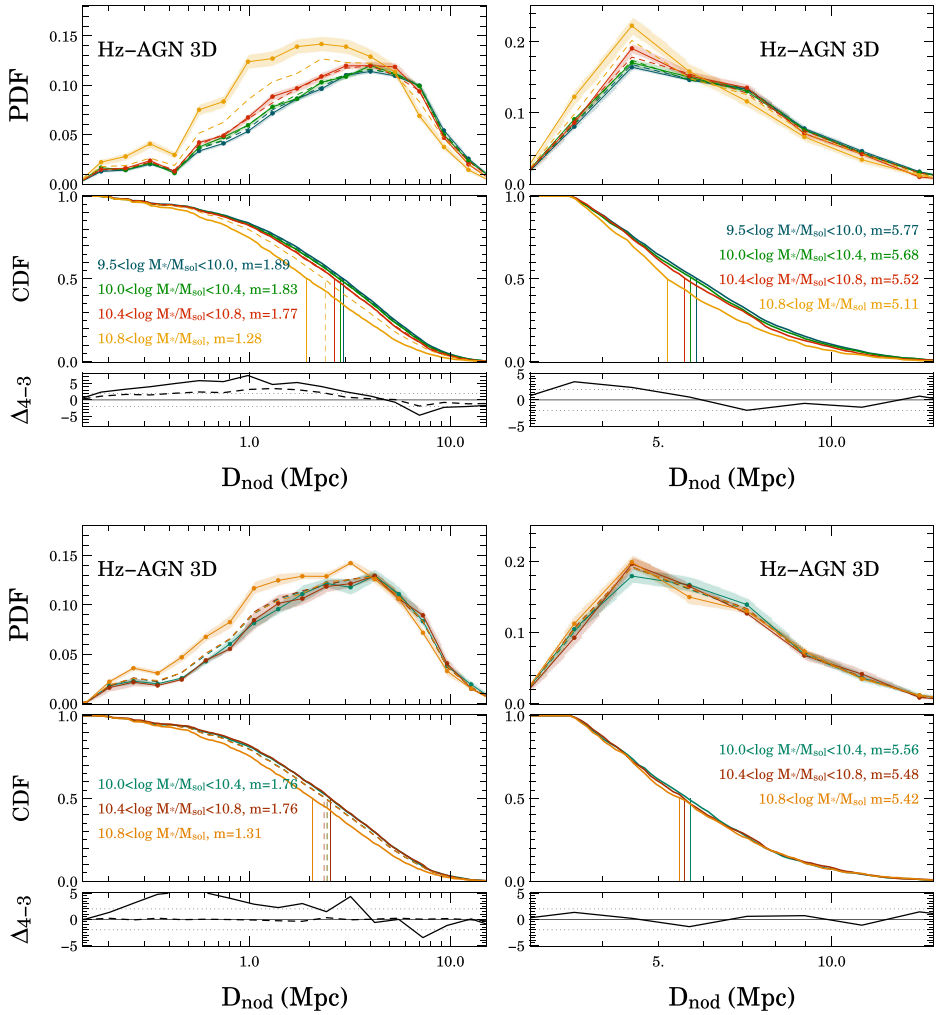

Figure E1. Differential and cumulative distributions displaying mass gradients towards nodes with (left) and without (right) taking into account the galaxies closer than 3.5 Mpc of the nodes. The two bottom panels show the same measurement but using density-matched galaxy populations in the three more mass bins. The signal is computed for $0.5 < z < 0.9$ in the Horizon-AGN simulated lightcone in three dimensions using $\mathrm{SKL}_{3\mathrm{D}}^{\mathrm{DM}}$. Vertical lines indicate the median of each distribution. The bottom row shows the residual between the two upper mass bins ($10.8 < \log M_*/M_\odot < 13$ and $10.4 < \log M_*/M_\odot < 10.8$). When removing galaxies closer than 3.5 Mpc to the nodes, all the gradient signal towards nodes is driven by the density (no signal after density matching).

quite large: removing totally the gradients towards nodes would imply removing a too large fraction of galaxies. Nevertheless, when doing the reshuffling test (see Section 3.2.2), we find that the reshuffled distributions are very close to the original ones, as shown by the dashed line in the top right-hand panels of Figs E2 and E1 (see also Table E1). In addition, no remaining gradients are found after the density matching algorithm (see Section 3.2.2), as displayed by the bottom right-hand panels of Figs E2 and E1 (see also Table E2). We conclude that, after excluding galaxies closer than 3.5 and 0.8 Mpc in three and two dimensions, respectively, the remaining gradients towards nodes are entirely driven by the density. Therefore, they should not impact the gradients towards filaments when these latter are themselves disentangled from the effect of the density.

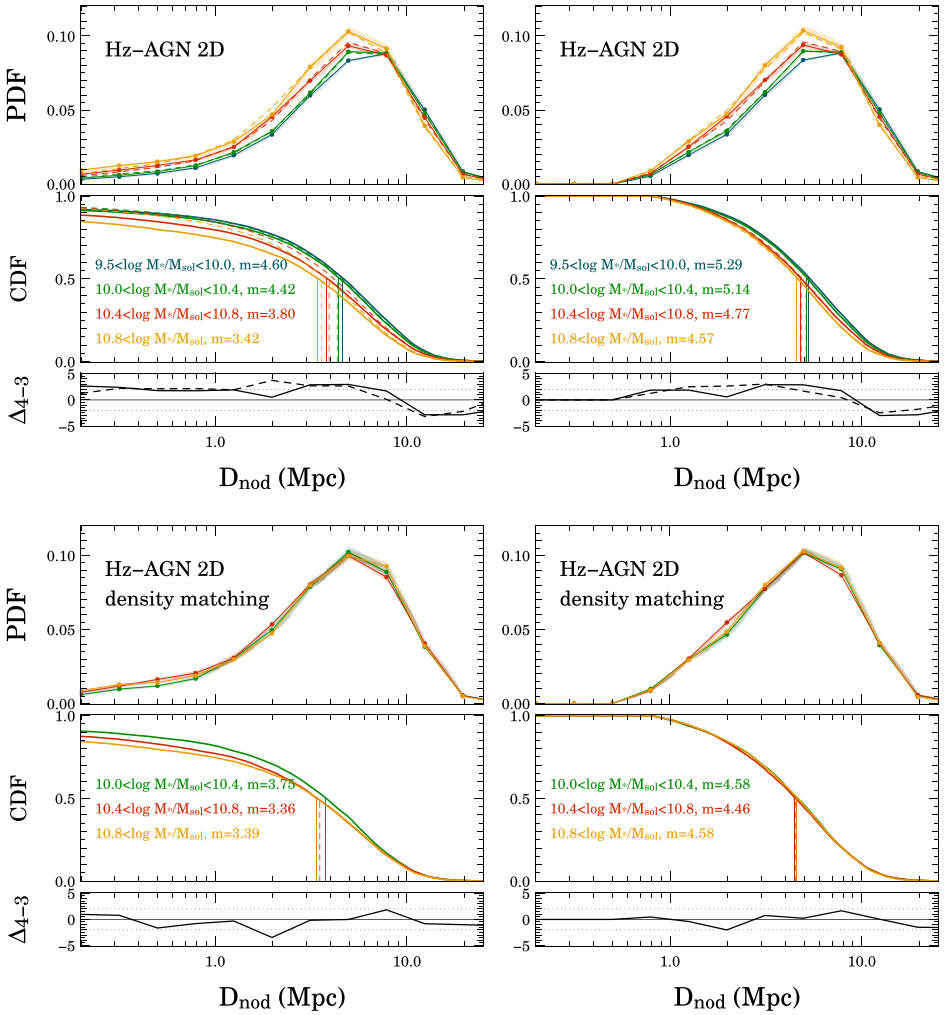

Figure E2. Differential and cumulative distributions displaying mass gradients towards nodes with (left) and without (right) taking into account the galaxies closer than 0.8 projected Mpc of the nodes. The two bottom panels show the same measurement but using density-matched galaxy populations in the three more mass bins. The signal is computed for $0.5 < z < 0.9$ in the Horizon-AGN simulated lightcone in two dimensions using $\mathrm{SKL}_{\mathrm{2D}}^{\mathrm{phot}}$. Vertical lines indicate the median of each distribution. The bottom row in each plot shows the residual between the two upper mass bins ($10.8 < \log M_*/M_\odot < 13$ and $10.4 < \log M_*/M_\odot < 10.8$). When removing galaxies closer than 0.8 projected Mpc to the nodes, all the gradient signal towards nodes is driven by the density (no signal after density matching).

Table E1. Summary of the results from the mass gradients towards nodes, in two and three dimensions, without (*all*) and with ($d_{\mathrm{nod}} > 0.8$ Mpc and $d_{\mathrm{nod}} > 3.5$ Mpc in two and three dimensions, respectively) excluding galaxies in the vicinity of the nodes. The given values are the medians of the distributions and the standard errors from bootstrap resampling with 100 realizations expressed in comoving Mpc. The reshuffled sample corresponds to the random signal preserving the galaxy mass–density relation.

	Median values of the PDF							
	HZ-AGN 2D				HZ-AGN 3D			
Mass range	All		$d_{\mathrm{nod}} > 0.8$		All		$d_{\mathrm{nod}} > 3.5$	
	Original	Reshuffled	Original	Reshuffled	Original	Reshuffled	Original	Reshuffled
[9.5, 10]	4.60 ± 0.02	–	5.29 ± 0.02	–	1.89 ± 0.02	–	5.77 ± 0.03	–
[10, 10.4]	4.42 ± 0.03	4.32 ± 0.03	5.14 ± 0.03	5.11 ± 0.03	1.83 ± 0.02	1.83 ± 0.02	5.68 ± 0.03	5.70 ± 0.05
[10.4, 10.8]	3.80 ± 0.05	3.94 ± 0.04	4.77 ± 0.04	4.82 ± 0.04	1.76 ± 0.03	1.79 ± 0.02	5.52 ± 0.05	5.61 ± 0.05
[10.8, 13]	4.32 ± 0.06	3.56 ± 0.06	4.57 ± 0.07	4.51 ± 0.05	1.28 ± 0.03	1.64 ± 0.04	5.11 ± 0.08	5.41 ± 0.09

Table E2. Summary of the results from the mass gradients towards nodes after density matching, without (*all*) and with ($d_{\rm nod} > 0.8$Mpc and $d_{\rm nod} > 3.5$ Mpc in two and three dimensions, respectively) excluding galaxies in the vicinity of the nodes. The density matching algorithm ensures that the selected subset in all mass bins shares the same density distribution. Therefore, the remaining gradient signal is not driven by the background density. The given values are the medians of the distributions and the standard errors from bootstrap resampling with 100 realizations expressed in comoving Mpc. After density matching, there is no signal driven by the nodes themselves when excluding galaxies closer than 0.8 projected Mpc and 3.5 Mpc in two and three dimensions, respectively.

Mass range	Medians of $\mathrm{PDF}_{d_{\rm nodes}}$ after density matching (Mpc)			
	HZ-AGN 2D		HZ-AGN 3D	
	All	$d_{\rm nod} > 0.8$	All	$d_{\rm nod} > 3.5$
[10, 10.4]	3.71 ± 0.03	4.59 ± 0.07	1.76 ± 0.06	5.56 ± 0.10
[10.4, 10.8]	3.38 ± 0.03	4.49 ± 0.02	1.76 ± 0.04	5.48 ± 0.07
[10.8, 13]	3.36 ± 0.05	4.58 ± 0.05	1.31 ± 0.03	5.42 ± 0.07

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