The role of the cosmic web in the scatter of the galaxy stellar mass–gas metallicity relation

Callum T. Donnan1,2, Rita Tojeiro1 and Katarina Kraljic3

The large-scale structure of the Universe can be understood in terms of features such as filaments, nodes and walls, which we collectively term the cosmic web. Galaxies evolve within the cosmic web, naturally raising the question of its impact on that process. There are two main mechanisms by which the cosmic web can influence galaxies: one is by modulating the growth of haloes, and the other is by regulating the gas ecosystem around galaxies. Disentangling the two is difficult, but key to deriving a holistic picture of galaxy formation and observational constraints on the growth of haloes. Here we report a detection of the effect of the cosmic web on the galaxy stellar mass–gas-phase metallicity relation of low-redshift star-forming galaxies using data from the Sloan Digital Sky Survey. The proximity of a galaxy to a node, independently of stellar mass and overdensity, influences its gas-phase metallicity, with galaxies closer to nodes displaying higher chemical enrichment than those farther away. We find a similar, but notably weaker, effect with respect to filaments. We find qualitative agreement in the cosmological hydrodynamical simulation IllustrisTNG (TNG300). Using IllustrisTNG, our results can be explained by both halo assembly bias and gas supply combining in nodes in a way that markedly modulates the metallicity of the gas, contributing to the scatter of this fundamental relation in galaxy evolution.

However, the gas ecosystem of galaxies is complex, and understanding how the cosmic web affects gas in and around galaxies is crucial to understanding its overall impact on the galaxy–halo connection. Some observational studies have focused on correlating the HI gas content in galaxies with some form of cosmic web estimate, with varying success and results1–3. Here we focus on the gas-phase metallicity, which describes the relative mass of elements heavier than helium in the interstellar medium (ISM) of a galaxy. Some insight into the environmental effects on gas-phase metallicity has been gained before4–6, but never within the full context of the cosmic web in a wide-area survey. The shape of the metallicity versus stellar mass relation (MZR) is a key observable in models of galaxy evolution, and offers important insight into baryonic evolutionary mechanisms in galaxies7. The metal content in the ISM is enriched by stellar evolution, and depleted by winds that carry metals away from a galaxy and by dilution through the accretion of lower-metallicity gas from the circumgalactic medium and the intergalactic medium, which in turn flow from the cosmic web. Gas-phase metallicity generally climbs steeply with stellar mass, with a scatter considerably larger than measurement uncertainties8. Understanding what drives this scatter—that is, why do galaxies at fixed stellar mass have different gas-phase metallicities?—is likely to drive our understanding of the physical processes that define the evolutionary paths of galaxies. In simulations, the scatter in the MZR has been linked to outflows and inflows9,10, and in observations it has been shown to correlate with the star-formation rate11, but how much the factors that drive these are purely internal and stochastic, or external and modulated by the cosmic web, remains to be determined.

In this Article, we study the effect of the cosmic web on the gas-phase metallicity of low-redshift galaxies in observations and in simulations. Our goal is to establish whether the cosmic web contributes to the observed scatter in the MZR and to explore possible underlying physical mechanisms to that contribution using...
Large-scale spectroscopic experiments will provide a step change in our understanding of the role of the cosmic web in the evolution of galaxies at low redshift, refining the results presented here.

The MZR observations

Our observational sample consists of emission-line galaxies from the Sloan Digital Sky Survey (SDSS) DR7 Main Galaxy Sample, with a median redshift of 0.071 and median stellar mass of $10^{10}$ solar masses ($M_\star$). Sample selection is detailed in the Methods, which also provides details of the catalogue of galaxy metallicities and estimates of the distance to cosmic web features. In this study, the term cosmic web will be used to refer to regions that are topologically different, and classified as filaments or nodes (see further details in the Methods). There is a sampling effect as a function of redshift, and therefore every distance to filaments or nodes is normalized by the redshift-dependent mean intergalaxy separation ($D_z$).

The MZR for this sample is shown in Fig. 1, as the background histogram. As a first illustration of the effect of the cosmic web on the MZR, we began by plotting the mean MZR in the 10% of galaxies closest to filaments and nodes, such that galaxies closer to filaments and nodes have higher gas-phase metallicities compared with their more distant counterparts.

To further quantify the effect of the cosmic web on MZR, we first computed the gas-phase metallicity residuals with respect to the mean value of the full population, evaluated as a function of stellar mass: that is, for each galaxy we quantified its deviation from the mean metallicity of the galaxies at the same stellar mass. We study these two potential mechanisms in one simulation alone (IllustrisTNG), acknowledging that different cosmological hydrodynamical models may show different results.

Gas-phase metallicity probes ionized, rather than neutral, gas. Ionized gas is less directly linked to gas supply and gas accretion, but it bears the imprint of past stellar evolution, galactic outflows and inflows. Studying it allows one to ask what the effect of the cosmic web on these mechanisms is. Gas-phase metallicity also has the advantage that it is more readily available and will be measurable up to $z < 1.4$ in surveys sufficiently dense and wide to reliably extract the cosmic web, such the forthcoming Dark Energy Spectroscopic Instrument survey. The Dark Energy Spectroscopic Instrument survey and other large-scale spectroscopic experiments will provide a step change in our understanding of the role of the cosmic web in the evolution of galaxies at low redshift, refining the results presented here.

The role of overdensity. Previous studies have shown that galaxies residing in overdensities have higher gas-phase metallicities, and have suggested that in overdense regions the ISM was pre-enriched at higher redshift. A relationship between gas-phase metallicity and overdensity is also seen in our data (Supplementary Fig. 2). As measurements of the cosmic web are typically correlated with overdensity, it was critical to assess how much our results were driven by the distance to the cosmic web, beyond the known effect of overdensity. We made this assessment by shuffling the data such that overdensity information is kept, but cosmic web distances were randomized before recomputing the residuals (see the Methods for full details). In this test, if the observed relationship between metallicity and shuffled distances remained, we could infer that it is overdensity, and not the distance to cosmic web, that drove the signal in the unshuffled data. The results are shown as the dashed
lines in Fig. 2c,d. The gradients are visibly reduced, particularly for nodes. We therefore interpret the relationship between the metallicity residual and distance to cosmic web in our unshuffled data to be primarily driven by the cosmic web, and not by overdensity. Having established this, we turned to IllustrisTNG to help us interpret our results on a qualitative level.

The MZR in simulations

Modern cosmological hydrodynamical simulations can resolve many baryonic processes on galactic scales, within a cosmological volume, and provide good matches to observations [16,26]. Here, we use TNG300 of the IllustrisTNG project [11,31] to: (1) attempt to reproduce the trends seen in our observational work, and (1) investigate the causes. The details of our sample selection and post-processing are given in the Methods. In summary, we selected galaxies in TNG300 by their colour, gas content and stellar mass at $z=0.1$. We did not make mock observations in this study—our goal was not to reproduce the observations exactly, but instead to check whether simulations show the same qualitative trends; and if so, to use simulations to understand what processes may be at play.

The mass–metallicity relation for the TNG300 sample, split by the 10th and 90th percentiles of the distance to the cosmic web, is shown by the solid lines in Fig. 3 (distance to filaments and nodes in Fig. 3a,b, respectively). There is a small difference in gas-phase metallicity with the distance to filaments, with a higher gas-phase metallicity found for the 10% closest galaxies. For nodes, a higher gas-phase metallicity for galaxies in their vicinity is clearly identified at all stellar masses in the range $9 < \log(M_*/M_\odot) < 10.2$ with an average difference of approximately 0.1 dex between the 10% closest and 10% most distant galaxies.

We investigated the correlation between gas-phase metallicity residuals and distances to the cosmic web in the same way as we did in our observational sample, shown for TNG300 in Fig. 4. There is a strong negative correlation for nodes, with galaxies displaying a higher gas-phase metallicity than average closer to nodes. This is present at both stellar masses with similar strength, although the sample size is notably smaller at $\log(M_*/M_\odot) > 10$ due to the smaller maximum mass in the galaxy selection in Fig. 4. Once again, the hypothesis that the residuals and distance to node are uncorrelated is rejected at greater than 5σ significance using the Spearman’s rank-order coefficient. When looking at distances from filaments, we observe a slight trend in residuals, which increases at very small distances from filaments.

To investigate the role of overdensity in the simulation, we performed the same shuffling technique that we did with the data (shown by the light dashed lines in Fig. 4c,d). As with the data, the correlations were not present in the shuffled simulated data, indicating that overdensity did not dominate the signal we see in the solid lines in Fig. 4c,d. In the case of TNG300, we also took a more straightforward approach, afforded by the larger number of galaxies and cleaner overdensity values (enabled by the higher number density and three-dimensional positions), and simply recomputed the MZR as a function of distance to the cosmic web. However, we used only galaxies within a small bin of overdensity, centred at the

---

Fig. 2 | Gas-phase metallicity residuals as function of the distance to filaments and nodes in SDSS DR7. a,b. One-dimensional histograms of the normalized distance to filaments (a) and nodes (b) for the selected SDSS DR7 galaxy sample. Dashed red lines indicate the 10th and 90th percentiles corresponding to the selection made in Fig. 1. c,d. Binned residuals of the mass–metallicity relation as a function of the normalized distance to filaments (c) and nodes (d) are shown by the solid lines. Binned residuals of the gas-phase metallicity shuffled by overdensity as a function of the normalized distance to filaments (c) and nodes (d) are shown by the dashed lines. Negative correlations between the gas-phase metallicity residuals and the distance to nodes and filaments (at low stellar mass) were removed when shuffling was performed. Error bars represent the s.e.m.
To understand how the cosmic web modulates the gas-phase metallicity, we explored different mechanisms that regulate gas-phase metallicity, including metallicity, we therefore looked at the star-formation histories of galaxies in TNG300 at $z=0.1$ separated into bins by the 10% closest to and 10% most distant from filaments. Galaxies at $d_{\text{node}} < 3.5$ Mpc are not included. The simulation therefore showed a qualitative agreement with observations, in that the effect was larger with nodes than filaments, and in both cases the gas-phase metallicity of galaxies increased with decreasing distance. Quantitatively, we note that the role of filaments was similar in simulations, whereas the role of nodes was amplified. This discrepancy in the role of nodes can be partly explained by the fact that the number density was larger in the TNG300 sample, resulting in typically smaller values of $d_{\text{node}}$. As the strength of the signal increased with decreasing $d_{\text{node}}$, we then expected a stronger signal in our TNG300 sample. The horizontal scales of Figs. 2 and 4 are not directly comparable. However, taking the mean value of $\langle D_{\text{fil}} \rangle = 8$ Mpc, we see that in SDSS DR7 data, we have very few galaxies with $d_{\text{node}} < 0.8$ Mpc, which is where we cut off the plot. The amplitude of the residuals at that distance is consistent between the two figures. The agreement between observations and simulations suggests that there is insight to be gained from further investigating the nature of these trends in the simulation, and we turn to that next.

The impact of the cosmic web on gas-phase metallicity

To understand how the cosmic web modulates the gas-phase metallicity, we explored different mechanisms that regulate gas-phase metallicity in galaxies. The metallicity of the gas in a galaxy is increased by star formation whereby stars produce metals through nuclear fusion and expel metals into the ISM through feedback such as stellar winds and supernovae. It is theorized that the metallicity of the gas could be lowered by the accretion of metal-poor gas from the circumgalactic medium and intergalactic medium. This dilutes the enriched gas in the galaxy, lowering the overall gas-phase metallicity of the ISM.

Gas mass fractions and the role of inflows. We begin by looking at the gas fraction, $M_{\text{gas}}/M_*$, as a function of stellar mass for galaxies in the 10th and 90th percentiles of the distance to nodes and filaments. The gas mass and stellar mass were computed within the stellar mass half-radius. We thereby considered gas that has reached the ISM, rather than probing gas in the circumgalactic medium. Limiting the gas to the galaxies’ half-light radius is also more comparable to our observational results, due to the SDSS fibre aperture. Figure 3 displays the gas fraction as a function of stellar mass for galaxies in opposing percentiles of the distance to filaments and nodes. Figure 5a shows that the difference between the gas fractions of galaxies close to filaments and those that are distant from filaments is very small; this is true at all stellar masses considered. In the scenario whereby gas-phase metallicity is lowered with increased gas accretion, the similarity in the gas fractions between the two samples is fully consistent with the similarity in their gas-phase metallicities. However, a small difference in the gas fraction of galaxies as a function of the distance to filament is observed if we restricted our analysis to galaxies closer to nodes and more distant from nodes, with the difference becoming more pronounced at low stellar mass. In the scenario above, whereby gas accretion decreases metallicity, once again the results in Figs. 3 and 5 are consistent: in that scenario, galaxies close to nodes have a lower gas fraction than those more distant from nodes, with the difference becoming more pronounced at low stellar mass.

Star-formation histories and the role of halo assembly. Simulations show that the ages of dark-matter haloes are correlated with the age of the stellar populations of the galaxies they host, especially at low halo or stellar mass, and that star-formation histories play a role in regulating the metallicity of the gas through chemical enrichment. To investigate the role of the cosmic web on the gas-phase metallicity, we therefore looked at the star-formation histories of the galaxies in two stellar mass bins (separated at $\log(M_*/M_\odot) = 10$) and in two bins of the distance to the cosmic web (the 10th and 90th percentiles). The resulting sample median $\pm 0.32$—this is shown by the solid lines in Fig. 3. The difference between the MZR computed using all overdensities and the MZR computed using a small bin of overdensity was small, and much smaller than the difference between the MZR computed at different distances from the cosmic web. We again interpret these results as evidence that the effect we observed with respect to the cosmic web was not dominated by overdensity. This conclusion holds when using overdensity computed with Gaussian kernels of 3 Mpc, 6 Mpc or 9 Mpc.

The difference between the MZR computed for galaxies in the full overdensity range. Dashed lines indicate the MZR computed for galaxies fixed to the median overdensity ($\pm 0.32$) in the sample. Coloured numbers indicate the number of galaxies in each MZR, with the smaller number representing the number for fixed overdensity. Grey histograms show two-dimensional histograms of individual galaxies. Error bars indicate the s.e.m. Differences in gas-phase metallicities are present between the closest 10% and most distant 10% of galaxies from filament and nodes in TNG300, independently of overdensity.
in galaxies closer to nodes on the stellar metallicity is evident in respectively. The impact of the earlier star-formation activity gas-phase metallicity of galaxies. Our observational analysis shows This work has shown a connection between the cosmic web and the metallicity in Fig. 3a.

Although not shown here, filaments have a very small impact on the star-formation histories and chemical enrichment, fitting well with the observed negligible effect of filaments on the gas-phase metallicity in Fig. 3a.

### Conclusions

This work has shown a connection between the cosmic web and the gas-phase metallicity of galaxies. Our observational analysis shows that the scatter in the galaxy stellar mass–gas metallicity relation (indicated by the offset in gas-phase metallicity at fixed stellar mass from the median relation) is significantly correlated with the distance to nodes in the cosmic web, but notably less correlated for filaments. Crucially, we have shown that this is an effect that is specific to the anisotropic nature of the cosmic web, and not driven by overdensity. We conclude that nodes and filaments have different roles in the scatter of the mass–metallicity relation—information that is lost if one quantifies the environment by overdensity alone.

Our analysis of IllustrisTNG demonstrates that the relationship between gas-phase metallicity and the distance to nodes can be explained by galaxies closer to nodes having limited access to low-metallicity gas and, at the same time, having experienced a more active star-formation history in the past, meaning that they have enriched their interstellar medium with more metals than galaxies of the same mass in different cosmic web environments. The qualitative agreement between TNG300 and SDSS DR7—despite quantitative differences—is encouraging, and suggests that comparing TNG300 and SDSS DR7 was a worthwhile exercise. However, current limitations of cosmological hydrodynamical simulations, particularly with respect to baryonic physics, mean that quantitative agreement with observations is not yet always possible. Our current work, therefore, offers a valuable opportunity in terms of model constraints: it will be informative to contrast other current and future models with our observations.

Disentangling the contributions from different physical mechanisms, and unfolding the different roles of filaments and nodes, will

---

**Fig. 4 | Gas-phase metallicity residuals as function of the distance to filaments and nodes in IllustrisTNG.**

- **a,b,** One-dimensional histograms of the distance to filaments (a) and nodes (b) for the selected TNG300 galaxy sample. Dashed red lines indicate the 10th and 90th percentiles corresponding to the selection made in Fig. 3.
- **c,d,** Binned residuals of the mass–metallicity relation as a function of the distance to filaments (c) and nodes (d) are shown by the solid lines. Binned residuals of the gas-phase metallicity shuffled by overdensity as a function of the distance to filaments (c) and nodes (d) are shown by the dashed lines. Negative correlations between the gas-phase metallicity residuals and distances to filaments (up to $d_{\text{fil}} \approx 10^{2.7}$ Mpc) and nodes were removed when shuffling was performed, pointing towards the role of the cosmic web, beyond overdensity. Error bars represent the s.e.m.
soon be possible by combining forthcoming statistically powerful surveys such as the Dark Energy Spectroscopic Instrument survey with detailed analysis of various simulations and forward-modelled mock observations. Our work here is an important first step towards that goal.

**Methods**

SDSS. The SDSS\(^5\) is a large galaxy imaging\(^5\) and spectroscopic\(^6\) survey covering over 14,000 square degrees of sky at present. In this study, we used data from the DR7\(^6\) Main Galaxy Sample\(^6\), which consists of a magnitude-limited sample to a petrosian r-band magnitude of \(r_p = 17.77\). This sample is well suited to our work, as the survey provides a complete and dense sample of galaxies over a
large contiguous area for the determination of cosmic web features, as well as high-quality spectra enabling the determination of detailed galaxy properties (such as the stellar mass and gas-phase metallicity).

**IllustrisTNG.** IllustrisTNG is a cosmological magneto-hydrodynamical suite of simulations that models galaxy formation and evolution on cosmological scales\(^{1-7}\) by solving the coupled evolution of dark matter and baryons. The IllustrisTNG collaboration provides simulation outputs from runs performed in different volumes and at different resolutions. We mainly utilized the TNG300-1 run, corresponding to a volume of 302.6 cMpc\(^3\) (co-moving Mpc) with a gas particle resolution of \(m_{\text{gas}} = 7.6 \times 10^4 M_{\odot} h^{-1}\). This is the largest of the TNG300 simulations, providing a wider variety of environments and better statistics. The smaller, higher-resolution boxes were used only to check the robustness of the gas-phase metallicity relation in TNG300.

**Catalogues and sample selections.** We use the MPA/JHU catalogue from SDSS DR7\(^{11}\), which contains stellar masses\(^{12}\) and gas-phase metallicities\(^{10}\). The stellar masses used are total stellar masses from photometry fits. The gas-phase metallicities in the catalogue were determined using a photoionization model with the measurements of multiple emission lines. The metallicities have a median 1σ error of 0.03 dex. In IllustrisTNG, we used the group catalogues provided by the simulation, which contain stellar masses, gas-phase metallicities and gas masses for each galaxy in snapshot 91. To better match the metallicities in SDSS DR7 data, we restricted our gas-phase metallicity measurements in TNG300 to gas that is actively forming stars. Galaxies for IllustrisTNG were provided by a supplementary catalogue\(^{11}\), from which we extracted colour-corrected colours and magnitudes.

To facilitate the comparison between observations and simulations, we performed colour and magnitude cuts in TNG300. The goal here was to approximate the two samples, acknowledging that matching samples exactly is not possible nor desirable, given that simulations do not reproduce observed galaxies precisely. We broadly addressed sample differences due to light-cone effects and voids in the snapshot, and the need to select galaxies in TNG that would have a measurable gas-phase metallicity—in the SDSS, these are galaxies with emission lines strong enough to be measured in individual spectra with the required signal-to-noise ratio. Effectively, this means removing lower-mass and redder galaxies in TNG, as well as those without gas. We selected galaxies with \(21.4<r<18.4\) and \(0.3<r<0.7\). The minimum stellar mass taken was \(M_{\text{stellar}} / M_{\odot} = 9.25\), although the magnitude and colour cuts above effectively remove most galaxies with \(M_{\text{stellar}} / M_{\odot} < 9.5\). Finally, a maximum stellar mass of \(M_{\text{stellar}} / M_{\odot} = 10.2\) was chosen as we found discrepancies in the mass–metallicity relation between TNG300, TNG100 and TNG50 that seemed to be associated with TNG300 failing to resolve relevant physical processes at high stellar mass. This is a similar range to that used in another study of the mass–metallicity relation in Illustris\(^{11}\), which showed that the MZR in IllustrisTNG is a good match to observations below that stellar mass cutoff. Our selection yielded a total sample of 97,679 galaxies in TNG300 out of a total of 225,105 galaxies above the minimum stellar mass. No cuts were made to SDSS DR7, other than the natural selection effect of requiring gas-phase metallicity measurements. This selected 65,984 galaxies from the 302,312 in the full sample. Further information and plots on the resulting samples are given in the Supplementary Information, with Supplementary Fig. 1 showing the colour selection on a colour–magnitude diagram and Supplementary Fig. 2 showing the resulting mass distribution.

This 302,312 galaxy sample was initially selected from the full catalogue of 927,552 galaxies in SDSS DR7 where the sample was selected to galaxies that were within a right ascension of 120° and 220° as this provided a contiguous area allowing DisPerSE\(^{30}\) catalogues to measure the distances to the cosmic web. The cuts described here generated a sample of 65,984 SDSS galaxies over approximately 6,700 deg\(^2\).

**Classification of the cosmic web and distances to nodes and filaments.** The cosmic web was extracted using DisPerSE\(^{30}\), which measures persistent, filamentary structures in three dimensions. This traces out the large-scale structure using the density fluctuations, categorizing the structure into nodes, filaments, walls and voids. This allowed us to generate a catalogue of \(d_{\text{node}}\) and \(d_{\text{fil}}\) for every galaxy in SDSS DR7 using the procedure from a recent paper\(^{41}\). DisPerSE was also used to generate a catalogue of the same type for every galaxy in snapshot 91 of TNG300. DisPerSE measures the significance of the attachment between two points in the density field in terms of the number of standard deviations, \(\sigma\). This is called the persistence level. A persistence level of 3\(\sigma\) was used to generate the catalogue for SDSS DR7 and a level of 4\(\sigma\) was used for TNG300.

In observations, there is a redshift dependence on the sampling of galaxies. If distances to the cosmic web were computed without accounting for this issue, redshifts with a higher number density \(n(z)\) would have artificially lower distances to nodes and filaments. This issue is typically resolved by binning every distance by the mean intergalaxy separation, \(D(z)\), as a function of \(z\):

\[
D(z) = m(z)^{-1}
\]

In SDSS DR7, \(D(z)\) varies between approximately 4 Mpc and 12 Mpc. When selecting the distance bins by the distance to filament in SDSS DR7, we excluded any galaxies at \(d_{\text{fil}} < 1\) Mpc to remove any influence from nodes and therefore isolate the relationship between metallicity and distance to filaments. This distance has been shown to be a robust cutoff for removing the influence of nodes in SDSS DR7\(^{11}\). In TNG300, we excluded galaxies at \(d_{\text{fil}} < 2\) Mpc, which was determined using the method described in a previous paper\(^{42}\).

**Residuals as a function of the distance to filaments and nodes.** The mean mass–metallicity relations (Figs. 1 and 3) were determined using stellar mass bins of width 0.1 dex by taking the mean gas-phase metallicity in each stellar mass bin. The uncertainty of the mean metallicity in each bin was given as the standard error on the mean, \(\sigma_{\text{mean}} = \sqrt{\frac{\sigma}{n}}\), where \(\sigma\) is the standard deviation of each bin and \(n\) is the number of galaxies in each bin. This procedure was performed for the SDSS DR7 and IllustrisTNG samples.

The residuals of the mass–metallicity relation (Figs. 2 and 4) were determined by taking the difference between the gas-phase metallicity of each galaxy and the mean metallicity for that stellar mass. The metallicity residuals were then binned and plotted as a function of distance to filament and node in bins of width 0.2 dex between a log(d/\(D(z)\)) of \(-1\) and 1 for SDSS DR7 and between a log(d/\(D(z)\)) of \(-1\) and 1.2 in TNG300 (where \(D(z)\) refers to filaments or nodes). The uncertainty is given as the error on the mean, \(\sigma_{\text{residual}}\).

Our interpretation of the results shown in Fig. 2 rely on quantifying how significant the correlation between gas metallicity residual and distance to the cosmic web is. We used the Spearman's rank correlation coefficient, \(r_s\), to quantify the level of correlation between two variables in a non-parametric way, without assuming a linear relationship between the two (only that the relationship is monotonic). Associated with each \(r_s\) we computed a P value that indicated the probability that a given value of \(r_s\) would occur from two uncorrelated variables. \(r_s\) and P values were computed using unbinned data.

We accounted for the effects of overdensity by shuffling the data with respect to overdensity. In this procedure, we constructed a new binned sample such that the distribution of overdensity of each original bin of \(\log(d/\(D(z)\)) was reproduced, but we drew galaxies from the full galaxy population. The result is that each bin of our shuffled sample had an overdensity distribution that matched, by construction, a very large contiguous area for the determination of cosmic web features, as well as high-quality spectra enabling the determination of detailed galaxy properties (such as the stellar mass and gas-phase metallicity).

**Data availability**

Data for SDSS DR7 and IllustrisTNG are publicly available at https://www.wmaпа. mpa-garching.mpg.de/SDSS/DR7/ and https://www.tng-project.org/data/. DisPerSE catalogues are available from the corresponding author upon reasonable request.

**Code availability**

DisPerSE is publicly available at: http://www2.lap.fr/users/sousbie/web/html/indexba87.html?category/Install. All other code used in this project is available from the corresponding author upon reasonable request.

Received: 26 August 2021; Accepted: 31 January 2022; Published online: 14 March 2022.

**References**

1. Wechsler, R. H. & Tinker, J. L. The connection between galaxies and their dark matter halos. *Annu. Rev. Astron. Astrophys.* 56, 435–487 (2018).
2. Blanton, M. R. & Moustakas, J. Physical properties and environments of nearby galaxies. *Annu. Rev. Astron. Astrophys.* 47, 159–210 (2009).
3. Bond, J. R., Kofman, L. & Pogosyan, D. How filaments of galaxies are woven into the cosmic web. *Nature* 380, 603–606 (1996).
