

Cluster-green galaxy correlations: Where do these galaxies live?

Selene Levis^{1,2,*}, Facundo Rodriguez^{1,3}, Héctor J. Martínez^{1,3}, Valeria Coenda^{1,3}, and Hernán Muriel^{1,3}

¹ Instituto de Astronomía Teórica y Experimental, CONICET – UNC, Laprida 854, X5000BGR Córdoba, Argentina

² Facultad de Matemática, Astronomía, Física y Computación, Universidad Nacional de Córdoba, Av. Medina Allende s/n, X5000HUA Córdoba, Argentina

³ Observatorio Astronómico, Universidad Nacional de Córdoba, Laprida 854, X5000BGR Córdoba, Argentina

Received 24 July 2025 / Accepted 20 April 2026

ABSTRACT

Context. Green valley (GV) galaxies are thought to represent a transitional population between star-forming and quiescent systems. However, their spatial distribution relative to galaxy systems, such as galaxy groups and clusters, remains unclear, particularly in relation to the large-scale environmental influence on galaxy quenching.

Aims. We aim to determine whether GV galaxies preferentially inhabit specific environments within galaxy systems, and to explore the physical nature of their location on the outskirts of massive haloes.

Methods. We analysed the spatial distribution of GV galaxies using the cluster–galaxy cross-correlation function, based on two datasets: the hydrodynamical simulation Illustris TNG300-1 and observational data from the Sloan Digital Sky Survey (SDSS DR18), cross-matched with the MPA–JHU DR7 catalogue. Galaxy systems with $\log(M_{200}/M_{\odot}) \geq 13.5$ are used as cluster centres, while galaxies classified as blue, green, or red, based on their location in the $(u - r)$ versus stellar mass diagram, serve as tracers for the correlation analysis.

Results. In TNG, GV galaxies show an increasing relative fraction with cluster-centric distance, peaking on the outskirts, particularly for low-mass galaxies and haloes. In some cases, the GV fraction exceeds that of red galaxies. SDSS data reveal qualitatively similar, but weaker trends, with the GV fraction remaining below that of red galaxies at all scales. Most GV galaxies on the outskirts are satellites bound to the central FoF group, consistent with a backplash or infall origin. Mock catalogues built from TNG and matched to SDSS selection functions reproduce the observational signal, indicating that projection effects drive the differences between datasets.

Conclusions. GV galaxies preferentially reside on the outskirts of galaxy systems as satellites bound to the central halo, supporting a scenario in which they are transitioning objects influenced by environmental quenching.

Key words. galaxies: groups: general – galaxies: halos – galaxies: star formation – galaxies: statistics – large-scale structure of Universe

1. Introduction

Galaxies in the Universe are broadly found to occupy two distinct sequences in the colour vs absolute magnitude (or stellar mass) diagram (CMD): the blue cloud (BC), consisting of actively star-forming galaxies, and the red sequence (RS), composed of passive systems (Strateva et al. 2001). This bimodality is well described by fitting the observed colour distribution at fixed absolute magnitude with the sum of two Gaussian functions, representing the blue and red populations, respectively (Baldry et al. 2004). Blue galaxies typically exhibit late-type morphologies (spiral and irregular), whereas red galaxies are generally associated with early-type systems (elliptical and lenticular; Driver et al. 2006). However, incorporating near-ultraviolet (NUV) data from the GALEX satellite Wyder et al. (2007) revealed a more complex structure in the CMD. They identified an excess of galaxies in the region between the BC and the RS, defining what is known as the green valley (GV), a transitional zone occupied by galaxies with declining star formation. These galaxies exhibit lower specific star formation rates (sSFRs) than BC galaxies of a similar stellar mass, suggesting that they are in the process of quenching. Different criteria have been used to define the GV boundaries, including empirical lines in the colour mass diagram (Schawinski et al. 2014;

Coenda et al. 2018), density-based colour ranges (Bremer et al. 2018), the 4000 Å break (Anghopo et al. 2019), and the shape of the sSFR distribution (Estrada-Carpenter et al. 2023). As was noted by Salim (2014), GV galaxies are particularly valuable for studying quenching mechanisms, as their position in the CMD reflects an evolutionary path from the BC towards the RS.

Observational studies based on samples from the Sloan Digital Sky Survey (SDSS; York et al. 2000), namely the Mapping Nearby Galaxies at APO project (MaNGA; Bundy et al. 2015), the SAMI Galaxy Survey (Bryant et al. 2015), and the Galaxy And Mass Assembly survey (GAMA; Baldry et al. 2010), show that GV galaxies exhibit physical properties that are intermediate between those of blue, star-forming galaxies and red, quiescent ones. Specifically, GV galaxies are predominantly early-type spirals and lenticulars (Bait et al. 2017; Mishra et al. 2018). Furthermore, the transition from the BC to the RS appears to correlate with a gradual weakening of the spiral arm structure (Smith et al. 2022), suggesting a continuous morphological evolution. However, Schawinski et al. (2014) showed that the GV does not represent a single evolutionary pathway, but rather comprises two distinct quenching modes: slow quenching in late-type disc galaxies, driven by gradual gas depletion over timescales of several gigayears, and rapid quenching in early-type systems associated with major mergers and morphological transformation. An analysis of radial profiles of the sSFR reveals

* Corresponding author: selenelevis@mi.unc.edu.ar

that GV galaxies show suppressed star formation compared to star-forming galaxies of similar mass, with this suppression evident out to two effective radii (Belfiore et al. 2018). Spatially resolved observations from SAMI reveal that galaxies in denser environments show a decline in sSFR from the outside-in, supporting an environmental cause for quenching such as ram pressure stripping or galaxy interactions (Medling et al. 2018). The dispersion in sSFR along the star-forming sequence exhibits a characteristic U-shaped trend with stellar mass, attributed to stellar feedback at low masses and active galactic nucleus (AGN) feedback at high masses (Davies et al. 2019). This implies that the quenching mechanism affects the galaxy as a whole, rather than being confined to an expanding central region. The relative abundance of GV galaxies suggests typical crossing timescales in the CDM of 1–2 Gyr, and that this transition is independent of the environment (Bremer et al. 2018). These findings support the notion that GV galaxies – mainly early-type disc systems – experience a sustained decline in gas supply, insufficient to maintain previous star formation levels, ultimately leading them to adopt passive colors. The cessation of star formation may be driven by a combination of internal processes and external factors. In this regard, Das et al. (2021) report that the fraction of green galaxies remains consistently around 10–20% regardless of environment, with approximately 10% of these hosting an AGN. The environmental effects appear to further influence the quenching process. Coenda et al. (2018) find that quenching is more efficient in high-density environments, where both the quenching timescale and the star formation rates are lower. Studies of group environments show that the fraction of passive galaxies increases towards halo centres, while star-forming galaxies dominate the infall regions, with a modest (~20%) suppression of the star formation rate already evident in galaxies within low-mass groups (Barsanti et al. 2018). In cluster environments, galaxies with recently quenched star formation, identified through strong Balmer absorption, bear significantly more common than in the field and are preferentially located in cluster outskirts, suggesting outside-in quenching on timescales of ~1.5 Gyr (Owers et al. 2019). Kinematic observations from SAMI show that gas-disturbed galaxies are most frequent in the infall regions of clusters, where ram pressure stripping during pericentric passage drives centrally concentrated star formation and eventual quenching (Bagge et al. 2025). For lenticular galaxies, the dominant formation channel appears to be the fading of spirals in dense environments, though mergers contribute in lower-density settings (Cocato et al. 2022). According to Sampaio et al. (2024), the environment significantly impacts galaxy morphology, particularly in low-mass systems. In summary, transition galaxies carry signatures of both secular evolution and environment-driven quenching.

Regarding the spatial distribution of galaxies, one of the most important findings from large-scale surveys is the strong dependence of galaxy clustering on colour. Early analyses using data from the SDSS revealed that red galaxies exhibit significantly higher clustering amplitudes and steeper correlation functions than blue galaxies (Zehavi et al. 2002). These differences were interpreted as a manifestation of the morphology–density relation (Dressler 1980), with red, early-type galaxies preferentially residing in denser environments and blue, and star-forming galaxies inhabiting less dense regions. Later work confirmed that this colour dependence holds across luminosity bins and is not reducible to luminosity-related effects alone (Zehavi et al. 2005, 2011).

Several studies have extended the clustering analyses to the GV. Coil et al. (2008), analysing galaxies at $z \sim 1$, show that

GV galaxies exhibit large-scale clustering comparable to that of red galaxies, yet their small-scale correlation slope and infall kinematics more closely resemble those of blue galaxies. This implies that a green galaxy is more likely to have a red neighbour nearby than a blue one. This dual behaviour suggests that green galaxies inhabit haloes similar in mass to those of red galaxies, but are more likely to reside on their outskirts rather than near the centres. The one-halo term observed for green galaxies compared to red ones supports this picture, indicating a lower radial concentration within haloes, possibly consistent with a scenario in which green galaxies migrate towards halo centres, becoming redder in the process.

Using SDSS and GALEX data, Heinis et al. (2009) reported that clustering amplitude increases monotonically with redder UV–optical colours, underscoring its sensitivity to recent star formation history. Further support for the transitional nature of GV galaxies comes from halo modelling: Krause et al. (2013) found that green galaxies occupy haloes with intermediate masses and satellite fractions relative to blue and red populations. This intermediate status is also evident at higher redshifts. Lin et al. (2019), studying galaxies in the redshift range $0.5 \lesssim z \lesssim 2.5$, show a redshift-dependent trend: at $z > 1.5$, green galaxies resemble blue galaxies more closely in their clustering, while at $z < 1.5$ they are more similar to red galaxies. Furthermore, these authors report that GV galaxies exhibit a clustering amplitude comparable to that of AGNs at similar redshifts, pointing to a potentially significant role of feedback processes in driving galaxy colour transformations.

The broader picture emerging from multiple studies suggests that blue galaxies are typically central objects in low-mass haloes, while red and green galaxies are more frequently satellites residing in virialised environments (e.g. Loh et al. 2010; Krause et al. 2013). These results support a model in which the blue-to-red transition via the GV reflects not only internal processes – such as gas depletion and stellar ageing – but also environmental evolution, as galaxies migrate from filamentary structures into groups or clusters.

More recently, McNab et al. (2021) investigated the abundance of transition galaxies in clusters, reporting that GV galaxies do not exhibit a significant excess in dense environments at high stellar masses, although a modest enhancement is observed at lower stellar masses. Additionally, Joy et al. (2025) employed marked correlation functions to assess the dependence of galaxy properties on the environment, finding that colour is the strongest tracer of local overdensity among the parameters they probed.

However, these studies have primarily relied on the two-point autocorrelation function (ACF) for colour-selected galaxy samples, or on the cross-correlation function (CCF) between galaxy populations distinguished by colour. To date, only one recent study has explored the dependence of galaxy–cluster clustering on galaxy properties via cluster–galaxy cross-correlations. At low redshift, Comparat et al. (2025) analysed correlations between X-ray selected clusters and colour-defined galaxy samples, finding that red galaxies dominate the cluster–galaxy cross-correlation signal, while blue galaxies show a clear suppression.

Together, these results trace a coherent picture: galaxy colour encodes information about the environment, and the transition from blue to red is accompanied by systematic changes in spatial distribution. The GV galaxies, in particular, exhibit clustering signals that reflect both their intermediate star-forming status and their emerging association with denser structures.

In a recent study, Levis et al. (2025) examined the evolution of various properties of galaxies, in and around groups, using the IllustrisTNG simulations (Nelson et al. 2019), classifying them

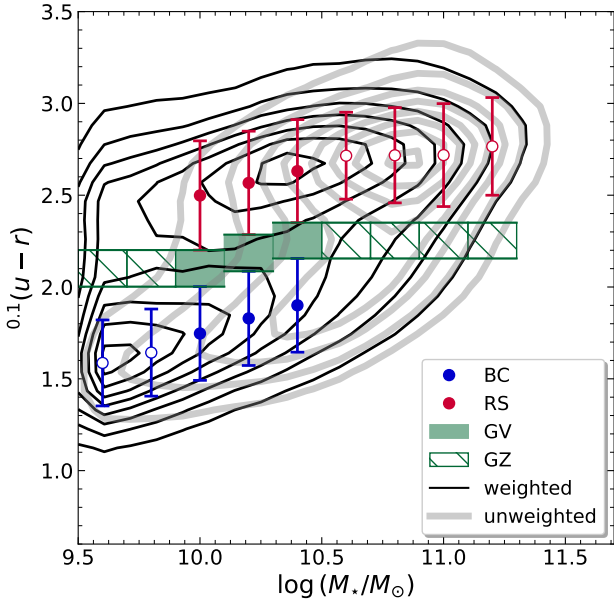


Fig. 1. $0.1(u-r)$ colour-stellar mass diagram (CMD) for MGS DR18 galaxies restricted to $0.025 \leq z \leq 0.15$ and to $\log(M_*/M_\odot) \geq 9.5$. Solid blue and red circles indicate the centres of the Gaussian components that best fit the blue and red populations, respectively. Open symbols of the same colours mark cases where only the BC or RS is present. The shaded green region indicates the GV, while the hatched area marks the GZ. Isocontours of galaxy number density are shown as black lines for the weighted distribution and grey lines for the unweighted case.

according to their trajectories. They found that backplash galaxies exhibit the highest fraction of green galaxies.

The two-point correlation function is a powerful statistical tool for quantifying the spatial distribution of galaxies. In this work, we investigate the colour dependence of galaxy clustering by analysing the cross-correlation between galaxy systems (from groups to clusters) and galaxies, with a particular focus on the GV population, in order to understand the spatial locations that these galaxies occupy in the vicinity of systems.

The article is structured as follows. Section 2 introduces the clusters and galaxies samples used in this study, drawn from the IllustrisTNG simulations and SDSS observations. Section 3 outlines the methodology, focusing on the use of the CCF. In Sect. 4 we present the results from the simulations and compare them with the observational data. Furthermore, it provides a comparison with the mock catalogue to investigate discrepancies between the simulations and observations. Finally, Sect. 5 summarises the conclusions of the work.

2. The samples

2.1. Simulations IllustrisTNG

From a theoretical perspective, we performed our analysis using the TNG300-1 run of the IllustrisTNG project (Nelson et al. 2019), a cosmological gravo-magnetohydrodynamical simulation that evolves the formation of structure from $z = 127$ to the present day under a Λ cold dark matter cosmology consistent with Planck Collaboration XIII (2016). The adopted cosmological parameters are $\Omega_{\Lambda,0} = 0.6911$, $\Omega_{m,0} = 0.3089$, $\Omega_{b,0} = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$, and $h = 0.6774$. This simulation models a 300 Mpc^3 co-moving volume and incorporates a comprehensive treatment of galaxy formation processes,

including gas dynamics, star formation, feedback, and chemical enrichment.

To trace the galaxy population, we selected 175056 subhaloes from the Subfind catalogue (Springel et al. 2001) with stellar masses of $\log(M_*/M_\odot) \geq 9.5$. We assigned each galaxy rest-frame absolute magnitudes in the SDSS u and r bands using synthetic photometry from the IllustrisTNG Supplementary Data Catalogs (Nelson et al. 2018), which include dust attenuation effects. This galaxy sample serves as the tracer population in our cluster-galaxy cross-correlation analysis.

Many authors have used the colour-stellar mass plane (e.g. Wyder et al. 2007; Brammer et al. 2009; Walker et al. 2013; Lee et al. 2015; Trayford et al. 2015; Coenda et al. 2018; Eales et al. 2018; Phillipps et al. 2019; Parente et al. 2025) to define the GV population, adopting different criteria to delineate the boundaries between the BC and the RS. In this work, we adopted the definition proposed by Levis et al. (2025), who identify red, blue, and green galaxy populations in the $(u-r)$ -stellar mass diagram (see their Fig. 1 for further details). Their classification is based on modelling the colour distribution within stellar mass bins using single or double Gaussian components, selected according to the Bayesian information criterion (BIC). From these fits, they define the GV as the region where both red and blue populations co-exist, and extend this definition towards higher stellar masses where either the RS or the BC remains but exhibits a broadened colour distribution. Following their scheme, we classified all galaxies within both the GV and the extended green zone (GZ) as green galaxies. Applying this criterion to our sample yields 92 798 blue galaxies, 29 477 green galaxies, and 52 781 red galaxies.

Furthermore, we extracted 1146 galaxy systems from the friends-of-friends (FoF) group catalogue at $z = 0$. The TNG FoF catalogues were generated from the simulations using the standard FoF algorithm (Springel et al. 2001). This method links together particles separated by less than a fixed fraction of the mean interparticle spacing (typically 0.2), thereby identifying groups that correspond approximately to virialised dark matter haloes. Within each FoF halo, the most bound particle is taken as a reference centre. We selected those with masses $\log(M_{200}/M_\odot) \geq 13.5$. Here, M_{200} refers to the mass enclosed within a radius in which the average density is 200 times the critical density of the Universe at the same redshift. This mass threshold ensures that each system hosts at least five satellite galaxies, providing reliable statistics for clustering measurements.

2.2. Observational data

2.2.1. The galaxy sample

We drew our observational galaxy sample from the Main Galaxy Sample (MGS; Strauss et al. 2002) in SDSS Data Release 18 (DR18; Almeida et al. 2023). We restricted our selection to the contiguous Northern Galactic Cap Legacy footprint ($\sim 7500 \text{ deg}^2$) within $0.025 \leq z \leq 0.15$. To minimise incompleteness at the bright end, we exclude all galaxies with $r \leq 14.5$, resulting in a final apparent-magnitude interval of $14.5 \leq r \leq 17.77$.

We corrected all apparent magnitudes for Galactic extinction using the maps of Schlegel et al. (1998), then computed absolute magnitudes by applying K corrections and shifting to $z = 0.1$ with KCORRECT v4.1 (Blanton et al. 2003). All magnitudes throughout are quoted in the AB system. We further required that $\log(M_*/M_\odot) \geq 9.5$, drawing stellar masses from the MPA-JHU

value-added catalogue based on SDSS-DR7 (Abazajian et al. 2009) through a direct cross-match with our parent sample. We adopted the same cosmological parameters for the observational data as those used in the TNG simulations.

Figure 1 shows the colour-magnitude diagram (CMD) for our sample of 409 348 galaxies. Due to the flux limit of the parent catalogue, each object was assigned a $1/V_{\max}$ weight (Schmidt 1968) in our CDM analysis. Galaxies were then split into blue, green, and red populations, adopting exactly the same classification criteria as those applied to the TNG sample. Figure 1 also shows the difference between the weighted and unweighted CMD. From this classification, we obtained 141 050 blue galaxies, 43 046 green galaxies, and 225 252 red galaxies.

We assessed the impact of stellar mass and photometric uncertainties on the classification of galaxies into blue, green, and red. Regarding the mass uncertainties, according to Kauffmann et al. (2003), the typical 95% confidence interval in stellar mass corresponds to ~ 0.3 dex in $\log(M_{\star}/M_{\odot})$, i.e. $\sigma \sim 0.15$ dex for Gaussian errors. We simulated this by adding a random Gaussian perturbation, ε , to each stellar mass, obtaining $\log(M_{\star})_{\text{pert}} = \log(M_{\star}) + \varepsilon$. After reclassifying galaxies in the CMD, 99% of blue, 95% of green, and 100% of red galaxies retained their original classification. Thus, typical stellar mass uncertainties do not significantly affect our colour-based analysis. We evaluated the impact of photometric errors and k corrections on our classification through a Monte Carlo test. Since k correct does not provide individual k -correction errors, we estimated them using k correct as follows. Input magnitudes were perturbed with Gaussian noise according to their photometric errors, and k corrections were re-computed for 30 realisations per galaxy. From the resulting $(u_{\text{pert}} - r_{\text{pert}}) + (k_{\text{pert}}^u - k_{\text{pert}}^r)$ distributions, we derived standard deviations adopted as individual $\sigma(u-r)$ values. We then generated perturbed colours $(u-r)_{\text{pert}}$ in an analogous way as we did with stellar mass uncertainties and reclassified galaxies in the CMD. The comparison between original and perturbed samples shows that the fractions of blue, green, and red galaxies remain essentially unchanged. After reclassifying galaxies in the CMD, 94% of blue, 82% of green, and 100% of red galaxies retained their original classification. Thus, uncertainties in the photometry should not significantly affect our results.

2.2.2. The sample of galaxy systems

To identify galaxy groups within the SDSS DR18, we employed a methodology outlined by Rodríguez & Merchán (2020). This approach combines the FoF method as in Merchán & Zandivarez (2005) with halo-based techniques (Yang et al. 2005). The process begins by identifying gravitationally bound systems using a percolation method, an adaptation of the technique described by Huchra & Geller (1982). Each identified group is assumed to contain at least one luminous central galaxy, and its initial properties are estimated based on the collective luminosity of its constituent members. Subsequently, an iterative, halo-based refinement step is applied to enhance the reliability of group membership and recalculate halo properties. This iterative procedure, which continues until convergence is achieved, ensures accurate system identification across a wide spectrum of group sizes, from small associations to extensive clusters. The resulting galaxy group catalogue provides essential information, including galaxy membership, spatial co-ordinates, and estimations of halo mass and radius. The halo mass estimation, M_{200} , is derived using an abundance matching technique (Vale & Ostriker 2004; Kravtsov et al. 2004; Conroy et al. 2006; Behroozi et al. 2010),

which posits a direct correlation between a group's characteristic luminosity and its dark matter halo mass. We computed the radius, R_{200} , following Eq. (7) of Rodríguez & Merchán (2020). Furthermore, galaxies within each group were categorised as either central or satellite, with the brightest galaxy consistently designated as the central galaxy (Rodríguez et al. 2021, 2022).

From the identified groups, we selected only those with halo masses $\log(M_{200}/M_{\odot}) \geq 13.5$ to focus on galaxy clusters. We identified the brightest galaxy in each group as the central and adopt its position as the group centre. This choice aligns with the established convention for classifying central and satellite galaxies, whereby the most luminous member traces the gravitational potential of the host halo (Rodríguez et al. 2021, 2022). Our final sample comprises 5612 galaxy systems.

The group catalogue used in our analysis has been validated against weak-lensing masses (Gonzalez et al. 2021) and has been successfully used in a variety of studies of galaxy environments (e.g. Alfaro et al. 2022; Rodríguez-Medrano et al. 2023; Parente et al. 2025). Furthermore, a recent comparative analysis by Kakos et al. (2024) has shown that key conclusions regarding galaxy clustering and the galaxy-halo connection remain robust when using the Yang et al. (2012), Tempel et al. (2017), or Rodríguez & Merchán (2020) catalogues, giving us high confidence in the robustness of our results.

2.3. Mock catalogue

To bridge the gap between theoretical models and observations, we constructed a mock catalogue as an intermediate stage by placing the observer at the origin of the TNG300 simulation box. We replicated the TNG300 volume as needed to match the spatial coverage of the SDSS DR18 spectroscopic survey, leveraging the periodic boundary conditions of the simulation. We calculated the redshift of each galaxy by combining its cosmological distance with the shift caused by peculiar velocities. By combining these redshifts with the intrinsic absolute magnitudes from TNG300, we derived the apparent magnitudes, allowing us to replicate the flux-limited selection of SDSS DR18 through the magnitude cut-off. Also, we constructed an angular selection mask to mimic the DR18 footprint, following methodologies similar to those used by Rodríguez et al. (2015), but adjusting for the updated geometry and completeness of DR18. A crucial consideration is the classification of galaxies as centrals or satellites, which, in observational surveys like SDSS, depends on group-finding algorithms and introduces inherent uncertainties. To ensure consistency, we applied the group-finder algorithm from Rodríguez & Merchán (2020) to our mock catalogue, mirroring the approach taken in observational data.

After applying the redshift range ($0.025 \leq z \leq 0.15$) and the SDSS flux limit ($14.5 \leq r < 17.77$), the final sample contains 128 393 galaxies with stellar masses $\log(M_{\star}/M_{\odot}) \geq 9.5$, and 9745 galaxy systems with $\log(M_{200}/M_{\odot}) \geq 13.5$, used as tracers and cluster centres, respectively, in the projected CCF analysis. Among these galaxies, 26 865 are classified as blue, 19 574 as green, and 81 954 as red.

3. Cluster-galaxy cross-correlation functions

To investigate the spatial distribution of green galaxies, we utilised the CCF, which enables us to quantify differences in the clustering of distinct galaxy samples. For the IllustrisTNG data, we computed the 3D CCF between galaxy system centres (C) and galaxy tracers (G), following the classical estimator of

Davis & Peebles (1983):

$$\xi(r) = \frac{N_R}{N_G} \frac{CG(r)}{CR(r)} - 1. \quad (1)$$

Here, $CG(r)$ and $CR(r)$ are the counts of system–galaxy and system–random pairs separated by a distance, r , and N_G , N_R are the numbers of galaxies and random points in the sample, respectively. The random catalogue (R) was constructed by uniformly distributing points within a box of the same size as the IllustrisTNG300 simulation volume.

To conduct an analysis of observations comparable to that of simulations and to avoid redshift-space distortion effects, we determined the two–co-ordinate CCF. Here, r_l represents the dimension parallel to the line of sight, and r_p represents the dimension perpendicular to it. Subsequently, we integrated over the r_l direction to calculate the projected CCF, $\xi(r_p)$:

$$\xi(r_p) = 2 \int_{r_{\min}}^{r_{\max}} \Xi(r_p, r_l) dr_l, \quad (2)$$

where $r_{\min} = 0.02 h^{-1} \text{ Mpc}$ and $r_{\max} = 30 h^{-1} \text{ Mpc}$ and the estimator for $\Xi(r_p, r_l)$ is given by

$$\Xi(r_p, r_l) = \frac{CG(r_p, r_l)}{CR(r_p, r_l)} \cdot \frac{N_R}{N_G} - 1, \quad (3)$$

where $CG(r_p, r_l)$ and $CR(r_p, r_l)$ represent the counts of cluster–galaxy and cluster–random pairs separated by r_p and r_l , and N_G and N_R are the numbers of galaxies and random points in the sample, respectively. The random catalogue was generated by uniformly distributing points across the SDSS footprint, reproducing the angular distribution, as well as the redshift distribution of the SDSS galaxy sample. For these computations we used the same correlation code as was used in Rodriguez et al. (2022).

4. Results

Figure 2 shows the cluster–galaxy CCFs computed from the IllustrisTNG 300-1 simulation. The upper panels display the CCFs for the full galaxy sample (black curve) and for the colour-selected subsamples: blue, green, and red galaxies (blue, green, and red curves, respectively). The lower panels show the ratio of each subsample’s CCF to that of the full sample, highlighting each subsample’s unique cross-correlation characteristics. Rows correspond to increasing M_{200} terciles (from top to bottom), and columns to increasing stellar mass terciles (from left to right). For each M_{200} tercile, the median value of r_{200} is indicated by a dashed grey line. Shaded regions represent jackknife-estimated uncertainties.

We observe that within the clusters, the CCF of green galaxies exhibits an intermediate behaviour between that of blue and red galaxies. In particular, for galaxies in the highest stellar mass range, the green and blue CCFs become more similar, especially as the stellar mass decreases. When analysing the ratio between the CCF of the colour-defined subsamples and that of the complete sample, we find that the likeness of finding green galaxies increases from the central to the outer regions of galaxy systems and tends to be higher on the outskirts, particularly for galaxies with a lower stellar mass. Specifically, in the lowest and the second M_{200} tercile (for galaxies with low and intermediate stellar masses), green galaxies dominate on the outskirts, even surpassing red galaxies. In the remaining combinations, although the relative fraction of green galaxies also increases with distance from the centre, it does not exceed that of red galaxies.

Table 1. Satellite fractions of GV galaxies on cluster outskirts ($1 \leq R/R_{200} \leq 2$) for TNG and SDSS.

Sample	TNG	SDSS	TNG	SDSS
	[%]		[%] of central FoF	
Whole sample	88.87	78.66	99.17	89.38
1 st tercile	92.73	86.33	100.0	95.83
2 nd tercile	90.85	78.14	99.66	91.61
3 rd tercile	87.00	76.54	98.72	86.25

We performed a cross-correlation analysis using the SDSS sample to examine whether similar trends are observed in real data. Figure 3 presents the observational results in plots analogous to those in Fig. 2. In general, the observational data show trends consistent with the simulation results. The relative fraction of red galaxies increases towards the centres of the systems, whereas the fraction of blue galaxies rises towards the outer regions but never surpasses that of red galaxies within the observed distance range. Green galaxies display an intermediate behaviour: within R_{200} their distribution resembles that of blue galaxies, while in the outer regions it becomes more similar to that of red galaxies. Similarly to in Fig. 2, we note that for green galaxies the relative fraction increases towards the outskirts and the infall region, particularly in low halo-mass systems, where green galaxies tend to match (or even surpass, in the case of the simulations) the relative fraction of red galaxies.

In Sect. 2.2 we explore how colour uncertainties could affect our classification of green galaxies in the SDSS sample when colours are perturbed accordingly with expected errors in photometry and k corrections. We computed the global CCF between the sample of colour-perturbed green galaxies and the groups in the SDSS. This is shown in Fig. A.1, where it can be seen that, within the errors, the resulting signal is consistent with that of the original sample. We therefore do not expect colour uncertainties to be an issue in the CCF of green galaxies and systems.

We explored whether the observed CCF of green galaxies is unique to them and not what we would expect from mixing up together blue and red galaxies. We selected a sample that includes all galaxies 0.1 dex above the GV or GZ and 0.1 dex below it, hereafter the (B+R) sample, and repeated the global CCF analysis. These are in Fig. A.2 for TNG galaxies, and in Fig. A.3 for SDSS galaxies. It can be seen that, in both cases, the CCF of the (B+R) sample presents a clearly distinct behaviour compared to green galaxies.

Previous studies suggest that most blue galaxies are typically central and isolated, whereas the fraction of satellites is higher among red and green galaxies, and that they are more commonly found in virialised structures (e.g. Heinis et al. 2009; Loh et al. 2010). We computed the central and satellite fractions of green galaxies located on the outskirts of galaxy clusters, specifically between one and two R_{200} from the group centre, for systems with $\log(M_{200}/M_{\odot}) \geq 13.5$ as well as for the three M_{200} terciles, in both the TNG simulation and SDSS observations, as summarised in Table 1. The table also reports the fraction of those satellites that belong to the central FoF group.

For the full sample, the fraction of centrals and satellites exhibits a qualitatively similar trend in both the simulation and observations: GV galaxies on the outskirts are predominantly satellites. In the SDSS data, only $\sim 21\%$ are classified as centrals, while in the TNG simulation this fraction drops to $\sim 11\%$.

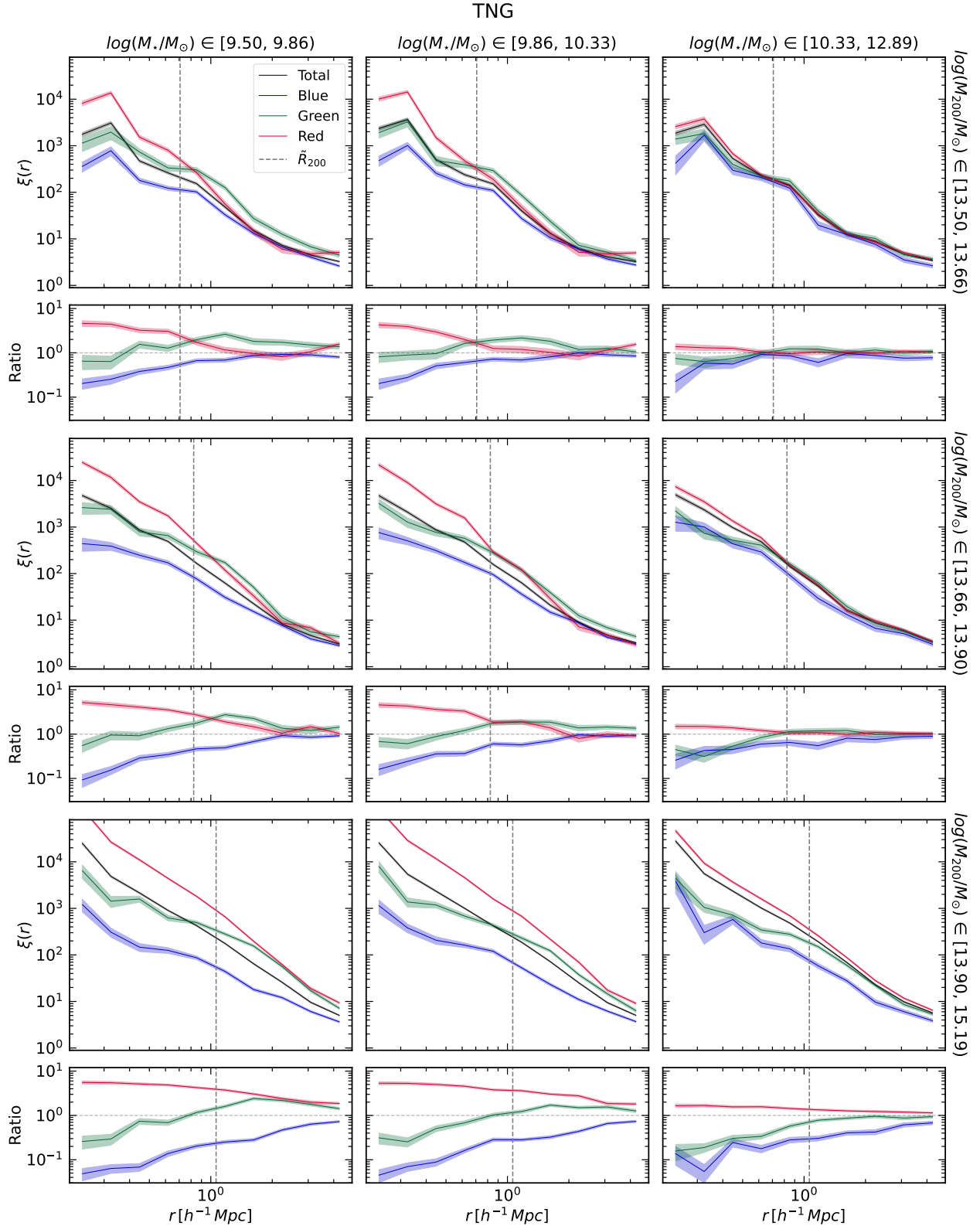


Fig. 2. Upper panels: Cluster-galaxy CCFs derived from the IllustrisTNG300-1 simulation. The black curve shows the full galaxy sample; blue, green, and red curves correspond to the respective colour subsamples. Rows correspond to terciles of galaxy system M_{200} , and columns to terciles of galaxy stellar mass. Shaded regions show uncertainties estimated via jackknife resampling. Lower panels: Ratios between the CCFs of the red, green, and blue subsamples and that of the total sample. Dashed grey lines denote the median R_{200} corresponding to each tercile of M_{200} .

To determine whether these satellites are part of the central halo or belong to infalling structures, we analysed their group membership. In SDSS, approximately 89% of the satellites in this

radial range are gravitationally bound to the central FoF group whose outskirts they inhabit; in TNG, this fraction increases to 99%.

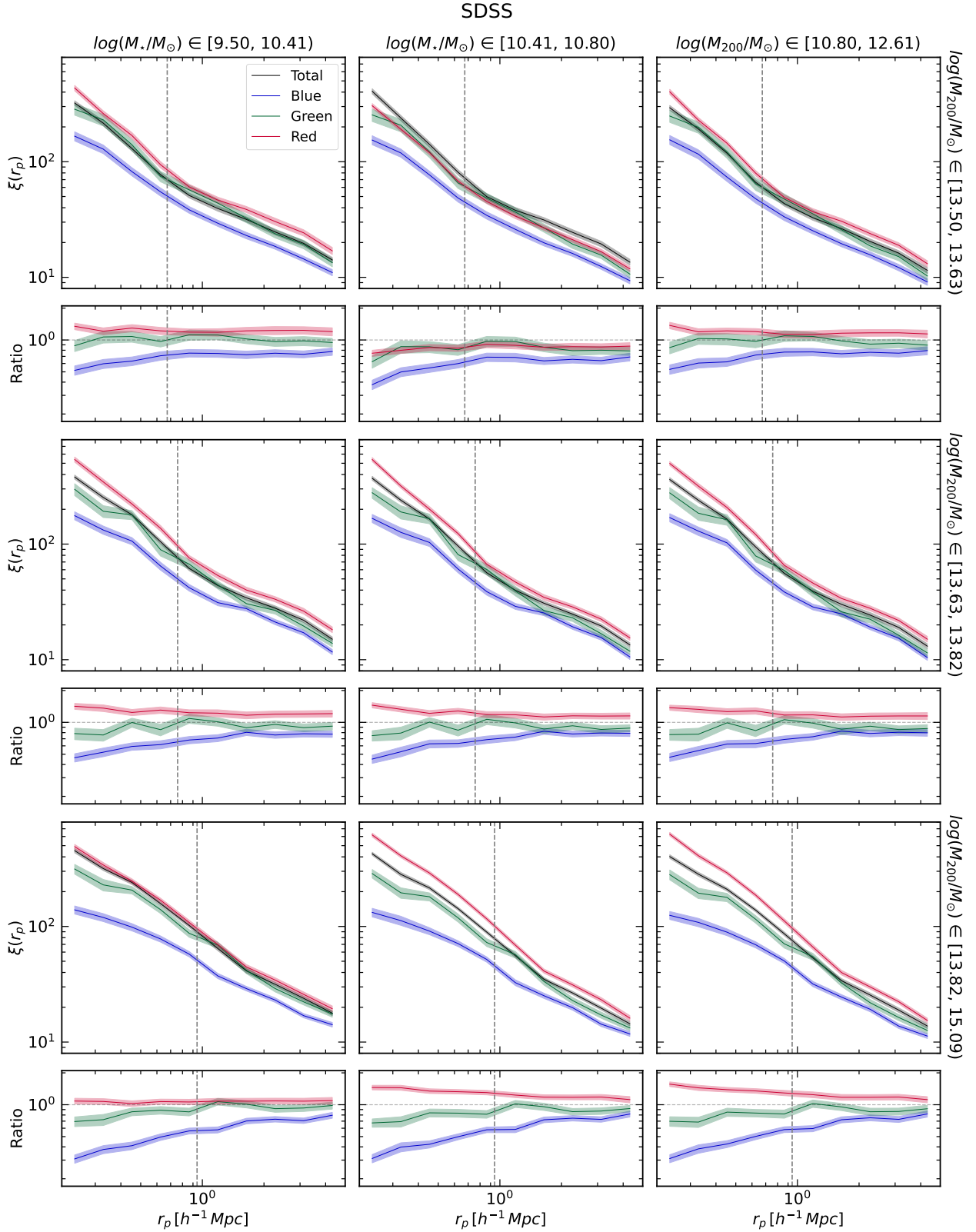


Fig. 3. CCFs and their ratios for the SDSS DR18 sample, analogous to Fig. 2. Rows correspond to terciles of galaxy system M_{200} , and columns to terciles of galaxy stellar mass. Shaded regions indicate uncertainties estimated via jackknife resampling. The dashed grey lines indicate the median R_{200} for each of the respective terciles of M_{200} .

Additionally, we find that the proportion of GV galaxies on the outskirts that are satellites is higher in lower-mass clusters, in both the simulations and the observations. The fraction of

these satellites that are gravitationally bound to the central FoF group also increases with decreasing cluster mass. In all cases, the corresponding fractions are slightly higher in the simulations

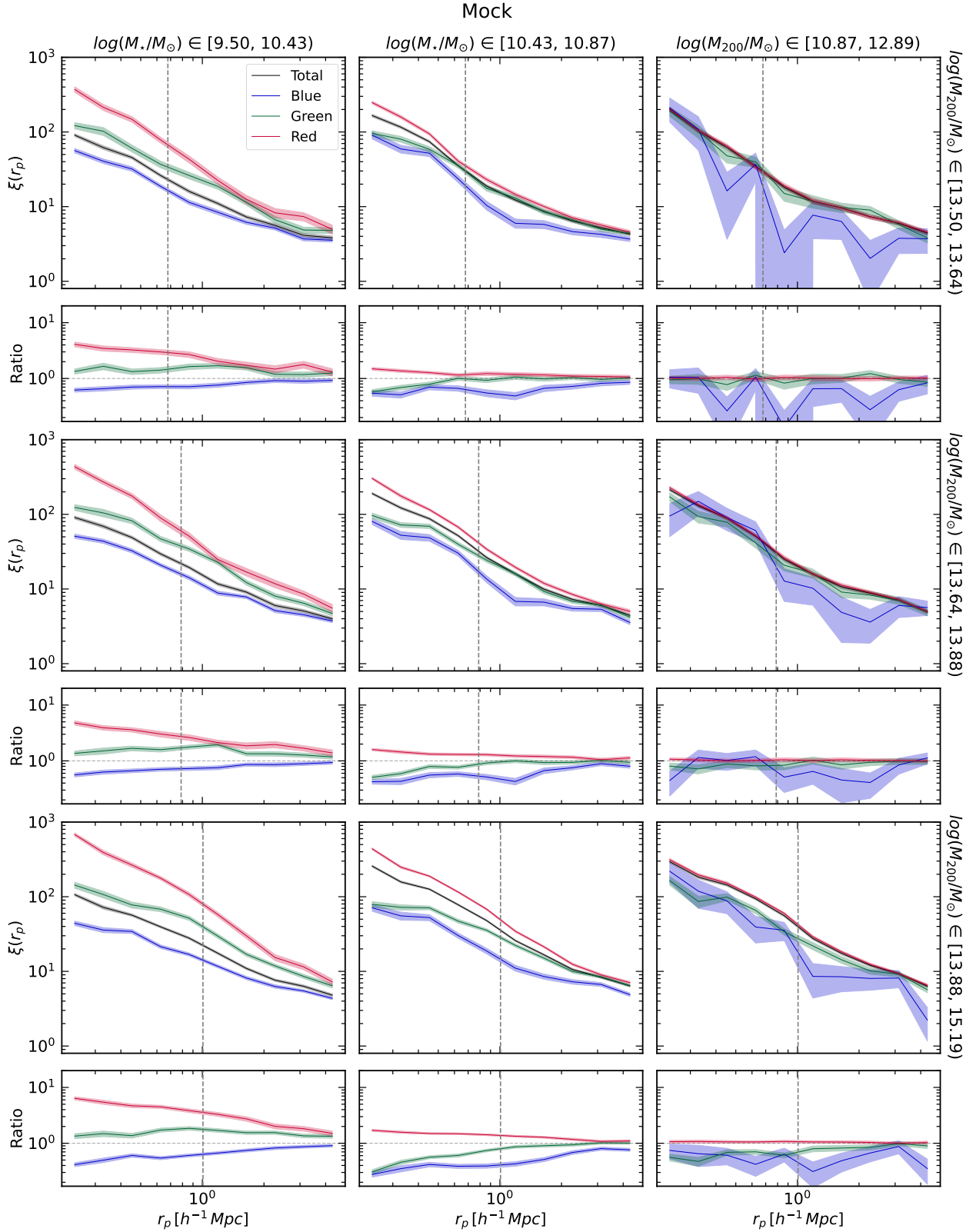


Fig. 4. CCFs and their ratios for the mock sample, analogous to Fig. 2. Rows correspond to terciles of galaxy system M_{200} , and columns to terciles of galaxy stellar mass. Shaded regions indicate uncertainties estimated via jackknife resampling. The dashed grey lines indicate the median R_{200} for each of the respective terciles of M_{200} .

than in the observations. The predominance of GV galaxies as satellites bound to the central FoF is consistent with the results of Levis et al. (2025), who, using a different methodology and

a sample of group galaxies from the TNG simulations, report that most GV galaxies are backsplash galaxies, i.e. galaxies that have experienced a single pericentric passage and are therefore

likely already bound to the system and undergoing quenching. An alternative scenario is that these galaxies are on their first infall into the central group.

To understand the origin of the discrepancies between the results from simulations and those derived from observational data, we performed an analogous analysis over the mock catalogue described in Sect. 2.3, which is shown in Fig. 4. Overall, the trends for green galaxies in the mock are more similar to SDSS. The relative rise of the correlation signal of green galaxies on the outskirts of systems seen in TNG vanishes in the mock. This behaviour indicates that the better part of the simulation–observation tension seen when comparing TNG and SDSS is naturally explained by projection effects, and the associated mixing of galaxies at different distances along the line of sight.

However, even when mock results are more similar to SDSS, the differences between green galaxies and the others are stronger in the mock. These differences between the mock and the observations may be due to a number of factors that prevents a one-to-one comparison. For instance the modelling of stellar populations in the simulation does not fully reproduce the observations (e.g. Nelson et al. 2019). On the other hand, the biases introduced by the SDSS photometry measurement procedure are not fully accounted for in the simulation modelling. Furthermore, in the observations, the best way to determine the GV in a colour-mass diagram requires the use of colours not available in the TNG data.

5. Conclusions

In this work, we investigate the spatial distribution of GV galaxies in relation to galaxy systems using the cluster-galaxy CCF. Our main objective is to determine whether GV galaxies preferentially inhabit specific environments.

We employed two complementary datasets: the magneto-hydrodynamical simulation IllustrisTNG 300-1, and observational data from the Sloan Digital Sky Survey (SDSS-DR18), cross-matched with the DR7 MPA–JHU catalogue. In both samples, we used galaxies with $\log(M_*/M_\odot) \geq 9.5$ at $z = 0$ as tracers, and galaxy systems with $\log(M_{200}/M_\odot) \geq 13.5$ at $z = 0$ as centres for the CCF. Galaxies were classified by colour (blue, green, and red) based on their location in the $(u-r)$ colour–stellar mass diagram. We applied the appropriate correlation function for each dataset – 3D for TNG and projected for SDSS – to perform the analysis. We recall that this comparison is qualitative in nature, aimed at identifying general trends rather than establishing a one-to-one correspondence between the simulations and observations.

The results from the IllustrisTNG simulation show a clear trend: the relative fraction of GV galaxies tends to increase towards the outer regions of galaxy systems and reaches a maximum on their outskirts, particularly for galaxies of lower stellar mass and in less massive systems. In these cases, the GV fraction on the outskirts even surpasses that of red galaxies. For more massive systems and more massive galaxies, although the effect is still present, it is less pronounced. The cross-correlation analysis based on SDSS data reveals qualitatively similar trends to those observed in the simulations, although with a weaker signal. Notably, in the SDSS data, the relative fraction of GV galaxies never exceeds that of red galaxies at any spatial scale. We examined the fraction of central and satellite GV galaxies on the outskirts of galaxy systems and found that the majority are satellites already gravitationally bound to the central system, residing in the peripheral regions around R_{200} . These results support the scenario in which GV galaxies are predominantly asso-

ciated with backplash or infalling populations likely undergoing environment-driven quenching. This interpretation is consistent with the findings of Levis et al. (2025), who, using an independent methodology, report that most backplash galaxies are located in the GV.

We explored the causes of the differences in the signal between simulations and observations. We constructed a mock catalogue from the IllustrisTNG simulation data, applying the same selection functions and observational constraints as those used in the SDSS. The mock results closely reproduce the observational findings, suggesting that the discrepancy is not due to intrinsic differences between simulated and real galaxy populations, but rather to projection effects inherent to redshift-space observations. However, it is worth noting that cluster properties in the SDSS and mock samples are derived through the abundance matching method, and thus should be considered as proxies rather than direct measurements. The fact that the results from the mock sample reproduce the trends seen in real data supports two main conclusions: first, that the signal detected is genuine and robust; and second, that the identification of systems and the method of assigning cluster properties are not introducing significant biases relative to the fully three-dimensional information available in the simulation, beyond the expected limitations from projection and sample selection.

Taken together, these findings show the important role of environment in quenching star formation in transition galaxies. The GV galaxies preferentially reside on the outskirts of galaxy systems and are predominantly satellite members of groups and clusters, highlighting that environment-driven processes are key to their evolution. Specifically, our results point out to:

1. Green valley galaxies are found mostly on the outskirts of galaxy systems (in the regions around the virial radius, R_{200}), particularly for lower-mass galaxies and less massive haloes. This suggests that they may predominantly correspond to galaxies on their first infall into the system or to backplash galaxies. This is in agreement with the results of Owers et al. (2019), who observationally find that recently quenched cluster galaxies are concentrated around $\sim 0.5-1$, R_{200} , consistent with ram-pressure stripping quenching star formation on short (~ 1 Gyr) timescales.
2. Most GV galaxies in these outskirts are satellites (either infalling newcomers or backplash objects) already bound to the central system. This implies that many GV galaxies are in the process of quenching due to environmental effects during or shortly after their infall. Our results therefore support a scenario in which transitional GV galaxies are often on their first infall into a cluster or are recent departures that have passed through the pericentre, linking their colour transformation to cluster-related processes. Consistently, cosmological simulations show that backplash galaxies exhibit the highest fraction of green galaxies (Levis et al. 2025).
3. The excess of GV galaxies around clusters is a genuine environmental signal rather than an artefact of observation or simulation. The qualitative agreement between the IllustrisTNG predictions and SDSS measurements, together with the mock catalogue test, demonstrates that the spatial trend we detect is robust.

In summary, the detection of a relative excess of GV galaxies on the outskirts of systems – consistent across simulations, mock data, and SDSS observations – highlights the importance of the environment in regulating star formation during the transition from the BC to the RS.

Acknowledgements. This work has been partially supported with grants from Agencia Nacional de Promoción Científica y Tecnológica, the Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET, PIP-2022-11220210100064CO), from the Agencia Nacional de Promoción de la Investigación, el Desarrollo Tecnológico y la Innovación de la República Argentina (PICT-2020-3690), from the Secretaría de Ciencia y Tecnología de la Universidad Nacional de Córdoba (SECYT-UNC, Res. 258/53). FR would like to acknowledge support from the ICTP through the Junior Associates Programme 2023-2028. The TNG 300-1 simulations used in this work are part of the IllustrisTNG project, which were run on the HazelHen Cray XC40 system at the High-Performance Computing Center Stuttgart, as part of the GCS-ILLU project of the Gauss Centers for Supercomputing (GCS).

References

- Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, *ApJS*, **182**, 543
- Alfaro, I. G., Rodríguez, F., Ruiz, A. N., Luparello, H. E., & Lambas, D. G. 2022, *A&A*, **665**, A44
- Almeida, A., Anderson, S. F., Argudo-Fernández, M., et al. 2023, *ApJS*, **267**, 44
- Anthopo, J., Ferreras, I., & Silk, J. 2019, *MNRAS*, **488**, L99
- Bagge, R. S., Foster, C., Brough, S., et al. 2025, arXiv e-prints [arXiv:2508.07589]
- Bait, O., Barway, S., & Wadadekar, Y. 2017, *MNRAS*, **471**, 2687
- Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, *ApJ*, **600**, 681
- Baldry, I. K., Robotham, A. S. G., Hill, D. T., et al. 2010, *MNRAS*, **404**, 86
- Barsanti, S., Owers, M. S., Brough, S., et al. 2018, *ApJ*, **857**, 71
- Behroozi, P. S., Conroy, C., & Wechsler, R. H. 2010, *ApJ*, **717**, 379
- Belfiore, F., Maiolino, R., Bundy, K., et al. 2018, *MNRAS*, **477**, 3014
- Blanton, M. R., Brinkmann, J., Csabai, I., et al. 2003, *AJ*, **125**, 2348
- Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2009, *ApJ*, **706**, L173
- Bremer, M. N., Phillipps, S., Kelvin, L. S., et al. 2018, *MNRAS*, **476**, 12
- Bryant, J. J., Owers, M. S., Robotham, A. S. G., et al. 2015, *MNRAS*, **447**, 2857
- Bundy, K., Bershady, M. A., Law, D. R., et al. 2015, *ApJ*, **798**, 7
- Coccolo, L., Fraser-McKelvie, A., Jaffé, Y. L., et al. 2022, *MNRAS*, **515**, 201
- Coenda, V., Martínez, H. J., & Muriel, H. 2018, *MNRAS*, **473**, 5617
- Coil, A. L., Newman, J. A., Croton, D., et al. 2008, *ApJ*, **672**, 153
- Comparat, J., Merloni, A., Ponti, G., et al. 2025, *A&A*, **697**, A173
- Conroy, C., Wechsler, R. H., & Kravtsov, A. V. 2006, *ApJ*, **647**, 201
- Das, A., Pandey, B., & Sarkar, S. 2021, *JCAP*, **2021**, 045
- Davies, L. J. M., Lagos, C. D. P., Katsianis, A., et al. 2019, *MNRAS*, **483**, 1881
- Davis, M., & Peebles, P. J. E. 1983, *ApJ*, **267**, 465
- Dressler, A. 1980, *ApJ*, **236**, 351
- Driver, S. P., Allen, P. D., Graham, A. W., et al. 2006, *MNRAS*, **368**, 414
- Eales, S. A., Baes, M., Bourne, N., et al. 2018, *MNRAS*, **481**, 1183
- Estrada-Carpenter, V., Papovich, C., Momcheva, I., et al. 2023, *ApJ*, **951**, 115
- Gonzalez, E. J., Rodríguez, F., Merchán, M., et al. 2021, *MNRAS*, **504**, 4093
- Heinis, S., Budavári, T., Szalay, A. S., et al. 2009, *ApJ*, **698**, 1838
- Huchra, J., & Geller, M. 1982, *ApJ*, **257**, 423
- Joy, K., Sureshkumar, U., Narayanan, A., et al. 2025, ArXiv e-prints [arXiv:2505.10429]
- Kakos, J., Rodríguez-Puebla, A., Primack, J. R., et al. 2024, *MNRAS*, **533**, 3585
- Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, **341**, 54
- Krause, E., Hirata, C. M., Martin, C., Neill, J. D., & Wyder, T. K. 2013, *MNRAS*, **428**, 2548
- Kravtsov, A. V., Berlind, A. A., Wechsler, R. H., et al. 2004, *ApJ*, **609**, 35
- Lee, G.-H., Hwang, H. S., Lee, M. G., et al. 2015, *ApJ*, **800**, 80
- Levis, S., Coenda, V., Muriel, H., et al. 2025, *A&A*, **698**, A57
- Lin, X., Fang, G., Cai, Z.-Y., et al. 2019, *ApJ*, **875**, 83
- Loh, Y.-S., Rich, R. M., Heinis, S., et al. 2010, *MNRAS*, **407**, 55
- McNab, K., Balogh, M. L., van der Burg, R. F. J., et al. 2021, *MNRAS*, **508**, 157
- Medling, A. M., Cortese, L., Croom, S. M., et al. 2018, *MNRAS*, **475**, 5194
- Merchán, M. E., & Zandivarez, A. 2005, *ApJ*, **630**, 759
- Mishra, P. K., Wadadekar, Y., & Barway, S. 2018, *MNRAS*, **478**, 351
- Nelson, D., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **475**, 624
- Nelson, D., Springel, V., Pillepich, A., et al. 2019, *Comput. Astrophys. Cosmol.*, **6**, 2
- Owers, M. S., Hudson, M. J., Oman, K. A., et al. 2019, *ApJ*, **873**, 52
- Parente, M., Ragone-Figueroa, C., Granato, G. L., et al. 2025, *A&A*, **697**, A231
- Phillipps, S., Bremer, M. N., Hopkins, A. M., et al. 2019, *MNRAS*, **485**, 5559
- Planck Collaboration XIII. 2016, *A&A*, **594**, A13
- Rodríguez, F., & Merchán, M. 2020, *A&A*, **636**, A61
- Rodríguez, F., Merchán, M., & Sgró, M. A. 2015, *A&A*, **580**, A86
- Rodríguez, F., Montero-Dorta, A. D., Angulo, R. E., Artale, M. C., & Merchán, M. 2021, *MNRAS*, **505**, 3192
- Rodríguez, F., Merchán, M., & Artale, M. C. 2022, *MNRAS*, **514**, 1077
- Rodríguez-Medrano, A. M., Paz, D. J., Stasyszyn, F. A., et al. 2023, *MNRAS*, **521**, 916
- Salim, S. 2014, *Serb. Astron. J.*, **189**, 1
- Sampaio, V. M., de Carvalho, R. R., Aragón-Salamanca, A., et al. 2024, *MNRAS*, **532**, 982
- Schawinski, K., Urry, C. M., Simmons, B. D., et al. 2014, *MNRAS*, **440**, 889
- Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, *ApJ*, **500**, 525
- Schmidt, M. 1968, *ApJ*, **151**, 393
- Smith, D., Habertzell, L., Porter, L. E., et al. 2022, *MNRAS*, **517**, 4575
- Springel, V., White, S. D. M., Tormen, G., & Kauffmann, G. 2001, *MNRAS*, **328**, 726
- Strateva, I., Ivezić, Ž., Knapp, G. R., et al. 2001, *AJ*, **122**, 1861
- Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, *AJ*, **124**, 1810
- Tempel, E., Tuvikene, T., Kipper, R., & Libeskind, N. I. 2017, *A&A*, **602**, A100
- Trayford, J. W., Theuns, T., Bower, R. G., et al. 2015, *MNRAS*, **452**, 2879
- Vale, A., & Ostriker, J. 2004, *MNRAS*, **353**, 189
- Walker, L. M., Butterfield, N., Johnson, K., et al. 2013, *ApJ*, **775**, 129
- Wyder, T. K., Martin, D. C., Schiminovich, D., et al. 2007, *ApJS*, **173**, 293
- Yang, X., Mo, H., Van Den Bosch, F. C., & Jing, Y. 2005, *MNRAS*, **356**, 1293
- Yang, X., Mo, H. J., van den Bosch, F. C., Zhang, Y., & Han, J. 2012, *ApJ*, **752**, 41
- York, D. G., Adelman, J., Anderson, J. E., et al. 2000, *AJ*, **120**, 1579
- Zehavi, I., Blanton, M. R., Frieman, J. A., et al. 2002, *ApJ*, **571**, 172
- Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2005, *ApJ*, **630**, 1
- Zehavi, I., Zheng, Z., Weinberg, D. H., et al. 2011, *ApJ*, **736**, 59

Appendix A: Complementary figures

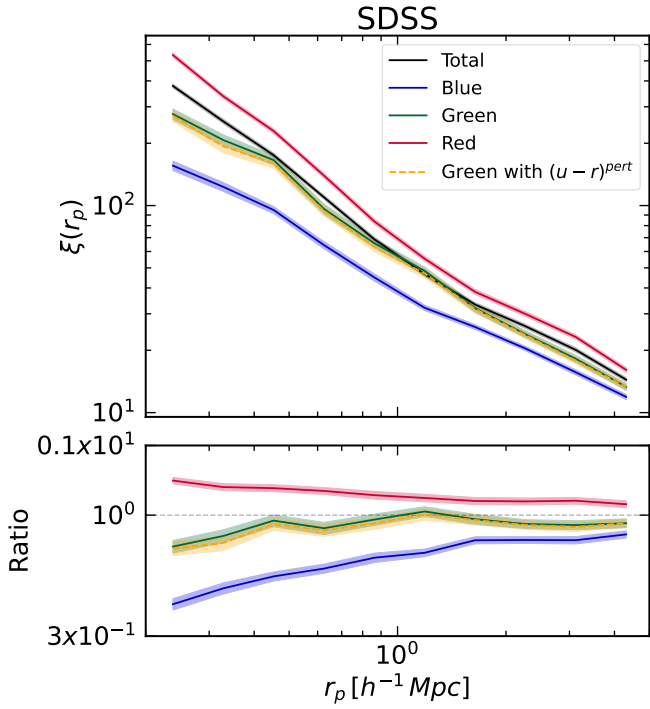


Fig. A.1. Upper panel: CCFs measured from the SDSS DR18 sample. The black curve shows the full sample, with all galaxies used as tracers and all galaxy groups used as centres. The blue, green, and red curves correspond to the CCFs of the respective colour subsamples. The orange dashed curve shows the CCF for the galaxy subsample of green galaxies constructed by perturbed $(u-r)$ colour applied to the CMD as mentioned in the Sect. 2.2. Lower panels: Ratios of the CCFs of the red, green, blue, and orange subsamples to that of the total sample. Shaded regions indicate uncertainties estimated via jackknife resampling.

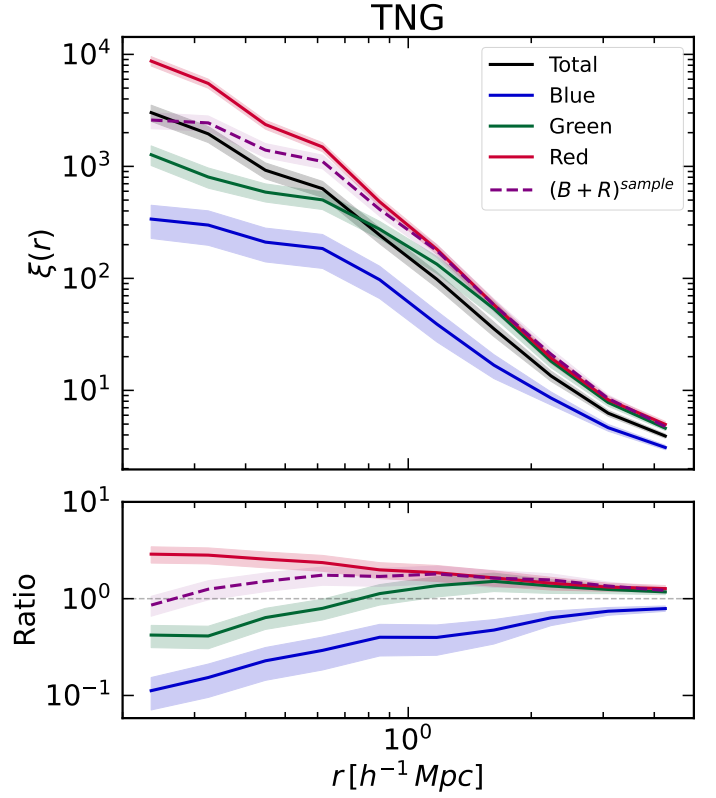


Fig. A.2. Upper panel: CCFs measured from the IllustrisTNG300-1 simulation. The black curve shows the full sample. The blue, green, and red curves correspond to the CCFs of the respective colour subsamples. The violet dashed curve shows the CCF for the $(B+R)$ sample selected as explained in Sect. 4. Lower panels: Ratios of the CCFs of the red, green, blue, and violet subsamples to that of the total sample. Shaded regions indicate uncertainties estimated via jackknife resampling.

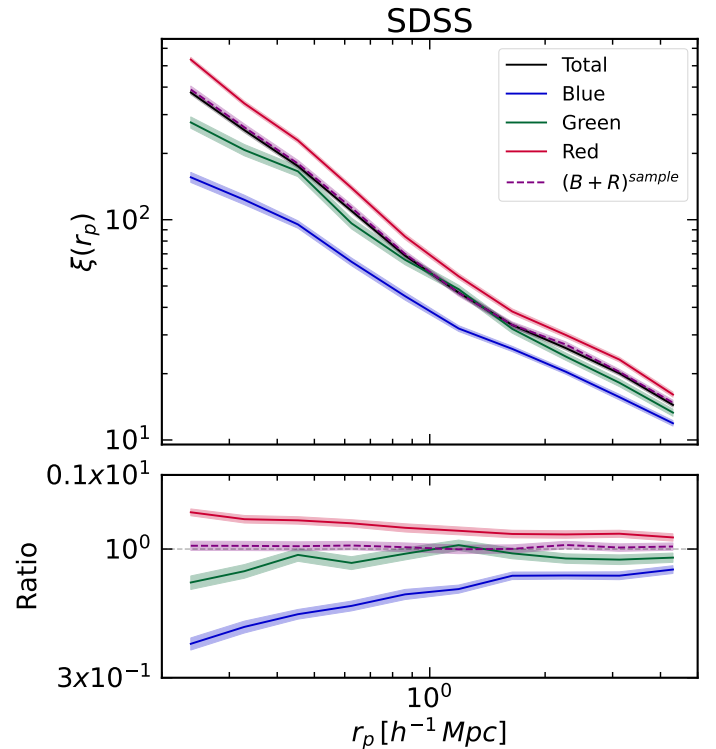


Fig. A.3. Same as Fig. A.2 for the SDSS DR18 sample.