

## MIT Open Access Articles

*Quenched fractions in the IllustrisTNG simulations: the roles of AGN feedback, environment, and pre-processing*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

**Citation:** Donnari, Martina, Pillepich, Annalisa, Joshi, Gandhali D, Nelson, Dylan, Genel, Shy et al. 2020. "Quenched fractions in the IllustrisTNG simulations: the roles of AGN feedback, environment, and pre-processing." Monthly Notices of the Royal Astronomical Society, 500 (3).

**As Published:** 10.1093/MNRAS/STAA3006

**Publisher:** Oxford University Press (OUP)

**Persistent URL:** <https://hdl.handle.net/1721.1/142405>

**Version:** Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

**Terms of use:** Attribution-NonCommercial-ShareAlike 4.0 International



# Quenched fractions in the IllustrisTNG simulations: the roles of AGN feedback, environment, and pre-processing

Martina Donnari<sup>1\*</sup>, Annalisa Pillepich<sup>1</sup>, Gandhali D. Joshi<sup>1</sup>, Dylan Nelson<sup>2</sup>, Shy Genel<sup>3</sup>, Federico Marinacci<sup>4</sup>, Vicente Rodriguez-Gomez<sup>5</sup>, Rüdiger Pakmor<sup>2</sup>, Paul Torrey<sup>6</sup>, Mark Vogelsberger<sup>7</sup>, and Lars Hernquist<sup>8</sup>

<sup>1</sup>Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

<sup>2</sup>Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str. 1, D-85748, Garching, Germany

<sup>3</sup>Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA

<sup>4</sup>Department of Physics and Astronomy, University of Bologna, Via Gobetti 93/2, I-40129, Bologna, Italy

<sup>5</sup>Instituto de Radioastronomía y Astrofísica, Universidad Nacional Autónoma de México, A.P. 72-3, 58089 Morelia, México

<sup>6</sup>Department of Astronomy, University of Florida, 211 Bryant Space Science Center, Gainesville, FL 32611, USA

<sup>7</sup>Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

<sup>8</sup>Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

## ABSTRACT

We use the IllustrisTNG hydrodynamical simulations to show how the fractions of quenched galaxies vary across different environments and cosmic time, and to quantify the role AGN feedback and pre-processing play in quenching group and cluster satellites. At  $z = 0$ , we select galaxies with  $M_{\text{stars}} = 10^{9-12} M_{\odot}$  residing within ( $\leq R_{200c}$ ) massive groups and clusters of total host mass  $M_{200c} = 10^{13-15.2} M_{\odot}$  in TNG100 and TNG300. The model predicts a quenched fraction of  $\sim 70$ -90 per cent (on average) for centrals and satellites of mass  $\gtrsim 10^{10.5} M_{\odot}$ , regardless of host mass, cosmic time ( $0 \leq z \leq 0.5$ ), cluster-centric distance and time since infall in the  $z = 0$  host. Low-mass central galaxies ( $\lesssim 10^{10} M_{\odot}$ ), on the other hand, are rarely quenched unless they become members of groups ( $10^{13-14} M_{\odot}$ ) or clusters ( $\geq 10^{14} M_{\odot}$ ), where the quenched fraction rises to  $\sim 80$  per cent. Typically, the fraction of low-mass passive galaxies is higher closer to the host center and for progressively more massive hosts. The population of low-mass satellites accreted more than  $\sim 4$ -6 Gyr ago in massive hosts is almost entirely passive, thus suggesting an upper limit for the time needed for environmental quenching to occur. In fact, about 30 per cent of group and cluster satellites that are quenched at  $z = 0$  were already quenched before falling into their current host, and the bulk of them quenched as early as 4 to 10 billion years ago. For low-mass galaxies ( $M_{\text{stars}} \lesssim 10^{10-10.5} M_{\odot}$ ), this is due to *pre-processing*, whereby current satellites may have been members of other hosts, and hence have undergone environmental processes, before falling into their final host, this mechanism being more common and more effective for the purposes of quenching for satellites found today in more massive hosts. On the other hand, massive galaxies quench on their own and because of AGN feedback, regardless of whether they are centrals or satellites.

**Key words:** methods: numerical — galaxies: formation — galaxies: evolution — galaxies: haloes — galaxies: groups — galaxies: clusters

## 1 INTRODUCTION

Galaxies in the Universe can be broadly classified into two main categories: “passive” or “quenched” galaxies, in which the star formation has ceased, and “star-forming” galaxies characterized by ongoing star formation. Typically, the first population exhibits spheroidal or elliptical morphologies and is mostly observed at high stellar masses or populating the high-density environments, while

the latter preferentially shows disk-like morphologies and resides in the low-density regions of the Universe (Strateva et al. 2001; Kauffmann et al. 2003; Baldry et al. 2004).

Observations and theoretical models agree on the bimodal distribution in the nearby Universe of many galaxy properties such as color, morphology, and star formation rate (SFR), although the bimodal nature of the latter is still a matter of debate (see e.g. Feldmann 2017; Eales et al. 2018). Furthermore, it is widely accepted that, given the hierarchical nature of the growth of struc-

\* E-mail: donnari@mpia-hd.mpg.de

tures, galaxy properties are tightly linked to whether a galaxy is a central or a satellite of its underlying dark matter host halo.

In the last few decades, several authors have proposed a wide variety of quenching mechanisms capable of describing the observed properties of galaxies. Although these mechanisms are still not completely understood, the cessation of star production in galaxies can be divided into two main scenarios, widely known as “mass quenching” and “environmental quenching”, since both the galaxy stellar mass and the environment in which galaxies reside and evolve have been shown to play a crucial role in the star formation activity of galaxies (e.g. Peng et al. 2010; Vulcani et al. 2012; Darvish et al. 2017; Paccagnella et al. 2016).

From observations, galaxies more massive than a few  $10^{10} M_{\odot}$  are likely to be quenched even if they reside in low-density regions. In this case, the quenching mechanisms are generally associated to the internal evolution of these galaxies, although galaxy mergers are also invoked as indirect quenching mechanisms (van der Wel et al. 2009), leading, for example, to gas consumption through starbursts. Among the “secular” processes, from a theoretical perspective, it is widely accepted that in massive galaxies, feedback from Active Galactic Nuclei (AGN) plays an important role in regulating star formation and in quenching galaxies (e.g. Di Matteo et al. 2005; Hopkins et al. 2005; Silk 2013; Hopkins et al. 2014; Somerville & Davé 2015; Beckmann et al. 2017; Penny et al. 2018). In fact, numerical cosmological simulations and semi-analytical models that follow galaxy evolution with and without AGN feedback (e.g., McCarthy et al. 2011; Dubois et al. 2016; Kaviraj et al. 2017; Weinberger et al. 2018) have demonstrated that broad consistency with observations with respect to the fractions of quenched galaxies and galaxy color distributions is only possible when some form of AGN feedback is included.

Conversely, a trend of lower SFRs in higher density environments, such as groups and clusters of galaxies, has emerged from observations of low-mass galaxies ( $\lesssim 10^{10} M_{\odot}$ ), while the latter are rarely quenched if isolated (see e.g. Geha et al. 2012; Medling et al. 2018). For example, recently, Schaefer et al. (2019) have shown that the specific Star Formation Rate (sSFR =  $\text{SFR}/M_{\text{stars}}$ ) of SAMI (Sydney AAO Multi-object Integral Field Galaxy Survey) galaxies drops as the local environment density increases around them. The effect of group environment has also been extensively investigated in SDSS data; for example, more recently in Crossett et al. 2017, by comparing (NUV-r) colors of both satellites and centrals: the authors find that a residual SFR is preferentially seen in field galaxies compared to satellites in the same stellar mass range. Various physical processes are commonly invoked to explain quenching driven by the local environment, where galaxies might be subjected to many mechanisms, some of them capable of reducing their gas reservoirs. Among others, these include strangulation or starvation (Larson et al. 1980; Moore et al. 1999; Nichols & Bland-Hawthorn 2011; Peng et al. 2015), tidal stripping (Merritt 1983), galaxy harassment (Moore et al. 1996). Furthermore, one of the processes often proposed for satellite galaxies is ram pressure stripping (Gunn & Gott 1972; Moore et al. 1999; Poggianti et al. 2017; Barsanti et al. 2018): as the gas is lost into the intra cluster-medium (ICM), the amount available to produce stars inside the galaxy gradually decreases, eventually leading the star formation to cease. This mechanism has been declared as the main contributor to quench observed galaxies in massive clusters, such as in Virgo (see e.g. Boselli et al. 2014, 2016, and references therein), and can now also be investigated in great detail in modern cosmological hydrodynamical galaxy simulations (see e.g. Yun et al. 2019, for a census of jellyfish galaxies in the IllustrisTNG simulations).

The emergence of two distinct quenching scenarios is expected to produce different fractions of quenched galaxies when taking into account galaxies in isolation rather than satellites in groups and clusters. The fraction of quenched galaxies (all together, i.e. without distinguishing between centrals and satellites) has been shown to increase with increasing galaxy stellar mass at any redshift ( $z \lesssim 3$ , see e.g. Muzzin et al. 2013; Darvish et al. 2016; Jian et al. 2018; Fossati et al. 2017; Fang et al. 2018). When only satellite galaxies are considered, the quenched fraction is found to be related to the host mass in which satellites reside, increasing with increasing host-to-satellite mass ratio, in observations (Wetzel et al. 2012, 2013; Davies et al. 2019) as well as in simulations (Furlong et al. 2015; Bahé et al. 2017; Tremmel et al. 2019) and in semi-analytical models (SAMs; De Lucia et al. 2012, 2019; Henriques et al. 2017). Furthermore, satellites are more likely to be quenched if they were accreted early, reinforcing the idea of an environmentally-driven quenching scenario (see e.g. Pasquali et al. 2019; Rhee et al. 2017). Finally, the role of the environment is also found to be strongly correlated with the position of galaxies within their hosts: satellites are more likely to be quenched if they reside towards the cluster center rather than in the outskirts, this trend being stronger for lower-mass satellites (Presotto et al. 2012; Wetzel et al. 2013; Barsanti et al. 2018).

Besides the most massive clusters (total mass  $\gtrsim 10^{14} M_{\odot}$ ), lower-mass groups are widely recognized to play a fundamental role in shaping the evolution of their satellite galaxies, as they are the building blocks of hierarchical cosmic structure formation in the framework of the  $\Lambda$ CDM model. In fact, it is widely accepted that a fraction of galaxies that today reside in massive groups and clusters have been orbiting in smaller subgroups prior to their accretion onto their current host, experiencing what is generally known as “pre-processing” (Zabludoff et al. 1996; Fujita 2004; McGee et al. 2009; Hou et al. 2014; Joshi et al. 2019; Bahé et al. 2019). However, albeit this mechanism has been widely investigated in observations (e.g. in the LoCuSS survey and SDSS: Haines et al. 2015; Bianconi et al. 2018; Wetzel et al. 2013, and references therein) as well as in theoretical models (analytical models, semi-analytical models and cosmological hydrodynamical zoom-in simulations of galaxy clusters: McGee et al. 2009; Bahé et al. 2013, 2019; Rhee et al. 2017; Joshi et al. 2017; Han et al. 2018), its importance in determining galaxy quenched fractions is still unclear.

For example, observationally, Hou et al. (2014) have investigated the quenched fractions of  $0.01 < z < 0.045$  satellites in SDSS groups and clusters. By means of a combination of simulations and mock catalogues, they inferred that pre-processing plays a significant role in establishing the observed quiescent fractions, but only when considering the most massive clusters in their sample ( $> 10^{14.5} M_{\odot}$ ). These results are somewhat consistent with results from SAMs described in De Lucia et al. (2012) and McGee et al. (2009), in which  $\sim 25$ -45 per cent of their simulated cluster galaxies were first accreted by subgroups and then fell into the current host. Similar conclusions have been outlined in Bahé et al. (2013, 2019) by using GIMIC (a set of re-simulations of approximately spherical regions, extracted from the Millennium Simulation, which include  $\sim 100$  hosts with masses in the range  $M_{200c} = 10^{13-15.2} M_{\odot}$ ) and Hydrangea (a suite of cosmological hydrodynamical zoom-in simulations of massive galaxy clusters,  $M_{200c} = 10^{14-15} M_{\odot}$ ), respectively, arguing that the frequency and impact of pre-processing strongly depends on the host mass at the time of observation: more massive clusters have higher fractions of pre-processed galaxies compared to lower-mass groups. Similar conclusions on the dependence on host final mass of the fraction

of pre-processed galaxies have been obtained via an independent analysis of the EAGLE simulation by [Pallero et al. \(2019\)](#). On the other hand, [Berrier et al. \(2009\)](#), using cosmological  $\Lambda$ CDM N-body simulations, have previously argued that pre-processing cannot significantly contribute to the present-day quenched fractions in clusters, since the majority of their satellites fell directly into their current host.

Differently from observations, the possibility of actually following the evolution of galaxies across cosmic time back to their infall into clusters represents the power of theoretical models. Yet, the fact that cluster galaxies might have been orbiting in smaller groups prior to being accreted in their final host does not automatically imply that their star formation activity has been impacted in such earlier stages to the point of quenching: this connection can be probed only with numerical models that self-consistently follow all the mechanisms that are thought to be relevant for galaxy evolution within the full cosmological context.

One such example is represented by the state-of-the-art cosmological magneto-hydrodynamical simulations of *The Next Generation* Illustris project (hereafter IllustrisTNG or TNG; [Naiman et al. 2018](#); [Marinacci et al. 2018](#); [Springel et al. 2018](#); [Pillepich et al. 2018](#); [Nelson et al. 2018a](#); [Pillepich et al. 2019](#); [Nelson et al. 2019b](#)). These sample tens of thousands of galaxies orbiting in hundreds of groups and clusters, thus providing one of the most reliable laboratories to study the evolution of galaxies across different environments and cosmic time. Importantly, unlike in most zoom-in suites, galaxies and halo hosts in the TNG volumes represent unbiased samples of the average galaxy and halo mass distributions in the Universe, thus allowing us to derive statistically-robust and unbiased relationships across the mass spectrum.

In this paper we aim to leverage the outcome of the IllustrisTNG simulations to pinpoint all the various pathways towards galaxy quenching by accounting for the effects of the hierarchical growth of structure, and hence by establishing the role of environment and pre-processing in determining the star-formation activity of galaxies as a function of galaxy and host mass. In practice, we will address the following questions: What is the role of pre-processing in quenching galaxies prior to their accretion onto their  $z = 0$  host? What is the distribution of the host masses where they actually quench? And since when are  $z = 0$  satellites quenched? Do high-mass groups and clusters quench their satellites more efficiently with respect to lower-mass groups even when pre-processing is taken into account? And do massive satellites quench because of environmental or internal processes?

Several outcomes from the IllustrisTNG simulations have validated the model against observational constraints, making them suitable for the tasks at hand. Among them we highlight: the shape and width of the red sequence and the blue cloud of  $z = 0$  galaxies ([Nelson et al. 2018a](#)), the existence and locus of the star formation main sequence at low redshifts ([Donnari et al. 2019](#)), the distribution of stellar mass across galaxy populations at  $z \lesssim 4$  ([Pillepich et al. 2018](#)), the galaxy size-mass relation for star-forming and quiescent galaxies at  $0 \leq z \leq 2$  ([Genel et al. 2018](#)), the evolution of the galaxy mass-metallicity relation ([Torrey et al. 2018](#)), and quantitatively consistent optical morphologies in comparison to Pan-STARRS data ([Rodriguez-Gomez et al. 2019](#)) as far as galaxy properties are concerned. These come in addition to observationally-consistent results in relation to the properties of massive hosts and their intra-halo gas: e.g. the X-ray signals of the hot gaseous atmospheres ([Truong et al. 2020](#); [Davies et al. 2020](#)), the amount and distribution of highly-ionized Oxygen around galaxies ([Nelson et al. 2018b](#)), and the distributions of metals in the intra-cluster

medium at low redshifts ([Vogelsberger et al. 2018](#)). Importantly, the new physical mechanisms included in the TNG model have been shown to return galaxy populations whose star formation activity is in better agreement with observations than previous calculations like Illustris ([Vogelsberger et al. 2014](#); [Genel et al. 2014](#); [Sijacki et al. 2015](#)), in galaxy colors ([Weinberger et al. 2018](#); [Nelson et al. 2018a](#)), atomic and molecular gas content of satellite galaxies ([Stevens et al. 2019](#)), and quenched fractions of central and satellite galaxies taken together with no distinction ( $z \lesssim 2$ , [Donnari et al. 2019](#)). In a companion paper ([Donnari et al. \(2020\)](#)), we discuss in detail the level of agreement between TNG results and observations when centrals and group and cluster satellites are considered separately and show that the quenched fractions of TNG galaxies are overall consistent with observational constraints and hence the results of this paper trustworthy.

This paper is organized as follows. In Section 2 we introduce the IllustrisTNG simulations and describe the criteria and definitions used throughout. In Section 3 we demonstrate the main results from the TNG model: we investigate the fraction of quenched galaxies versus galaxy stellar mass of both centrals and satellites, across cosmic time. We also present our findings across environments, i.e. host masses for satellite galaxies, and explore how the quenched fractions are correlated with the  $z = 0$  cluster-centric distance, in concert with its distribution in phase space at  $z = 0$  and its dependence on the amount of time spent in group and cluster environments. In Section 4 we introduce the concept of pre-processing and we discuss its role in shaping the  $z = 0$  quenched fraction of satellites, show the distributions of the mass of their quenching host and the time since quenching, and draw a picture of all the different pathways galaxies quench according to the IllustrisTNG simulations. Our summary is given in Section 5.

## 2 METHODS AND DEFINITIONS

### 2.1 The IllustrisTNG simulations

In this work, we use two simulations modelling two cosmological volumes with, respectively, a side length of  $75 h^{-1} \approx 100$  Mpc and a baryonic target mass of  $m_{\text{baryon}} \sim 1.4 \times 10^6 M_{\odot}$  (TNG100) and  $205 h^{-1} \approx 300$  Mpc a side and a baryonic target mass of  $m_{\text{baryon}} \sim 1.1 \times 10^7 M_{\odot}$  (TNG300). These are part of the IllustrisTNG project <sup>1</sup> ([Naiman et al. 2018](#); [Marinacci et al. 2018](#); [Springel et al. 2018](#); [Pillepich et al. 2018](#); [Nelson et al. 2018a](#); [Pillepich et al. 2019](#); [Nelson et al. 2019b](#)). The highest resolution simulation of the project, with a side length box of  $35 h^{-1} \approx 50$  Mpc and a baryonic target mass of  $m_{\text{baryon}} \sim 8.5 \times 10^4 M_{\odot}$  (TNG50; [Pillepich et al. 2019](#); [Nelson et al. 2019b](#)), will not be discussed in this paper.

The TNG100 and TNG300 runs are cosmological magneto-hydrodynamical simulations performed with the AREPO code ([Springel 2010](#)) and including various physical ingredients, such as a novel black hole feedback scheme. All details of the underlying galaxy formation model can be found in the two methods papers by [Weinberger et al. \(2017\)](#) and [Pillepich et al. \(2018\)](#); further details about the numerical resolution achieved in TNG100 and TNG300 can be found for example in [Pillepich et al. \(2018\)](#); [Nelson et al. \(2018a\)](#). All the results presented in this paper are the direct outcomes of TNG100 and TNG300: we do not rescale any property

<sup>1</sup> <http://www.tng-project.org>

of the simulated galaxies to account for resolution effects, as proposed in previous works (Pillepich et al. 2018; Engler et al. 2020). Instead, we will comment throughout and in Appendix A on how numerical resolution affects the presented results.

## 2.2 Galaxy and cluster samples: centrals versus satellites

All the properties of haloes and subhaloes used in this work are obtained via the Friends-of-Friends (FoF) (Davis et al. 1985) and SUBFIND algorithms (Springel et al. 2001; Dolag et al. 2009), which are used to identify substructures across the simulations.

In order to investigate the effect of environment on galaxy evolution, we differentiate galaxies residing in groups and clusters from the whole population. Unless otherwise stated, we define a “virial” radius  $R_{200c}$  as the radius within which the mean enclosed mass density is 200 times the critical density of the Universe. We will refer to the total mass enclosed within this radius as  $M_{200c}$ . In general, all haloes with  $M_{200c} \geq 10^{13} M_{\odot}$  will be named “hosts”. More specifically, throughout the text we will refer to hosts less massive than  $10^{14} M_{\odot}$  as “groups”, and more massive ones as “clusters”.

Furthermore, at  $z = 0$ , all galaxies residing within one virial radius of the host center are dubbed as “satellites” but for the galaxy settled at the minimum of the potential well of its FoF halo, which is typically the most massive and will be named the “central”. Galaxies in hosts with  $M_{200c} \geq 10^{13} M_{\odot}$  are generally called group or cluster galaxies. Moreover, we neglect all those subhalos that are not thought to be of cosmological origin, i.e. they are possibly baryonic fragments of other galactic structures identified by the Subfind algorithms (see Section 5.2 in Nelson et al. 2019a). We note here that, for the purposes of the paper, the adjective “satellite” assumes a more general meaning at  $z > 0$ : namely, when we track cluster galaxies back in time, we also define as satellites those galaxies orbiting outside the virial radius  $R_{200c}$  of their  $z = 0$  host.

At  $z = 0$ , TNG300 (TNG100) samples 3733 (182) among groups and clusters of mass  $M_{200c} \geq 10^{13} M_{\odot}$  and a total of more than 39,000 (2,800) satellites with stellar mass  $M_{\text{stars}} \geq 10^9 M_{\odot}$ . The most massive cluster in TNG300 (TNG100) reaches a total host mass of  $10^{15.2} M_{\odot}$  ( $10^{14.6} M_{\odot}$ ) – see Tab. 1 for more details.

The galaxy stellar mass  $M_{\text{stars}}$  adopted in this work is the sum of all gravitationally bound stellar particles within twice the stellar half-mass radius,  $R_{\text{star,h}}$ . All galaxies in our samples have a stellar mass larger than  $M_{\text{stars}} \geq 10^9 M_{\odot}$ , that corresponds to only having galaxies in our sample with at least one thousand star particles in total in TNG100 and more than one hundred in TNG300. Throughout the text and in all the figures presented in this work, we will clarify whether we take into account satellites, centrals or both.

## 2.3 Star formation rates in IllustrisTNG galaxies

For the purposes of this work we characterize galaxies by their ‘instantaneous’ SFR, i.e. by considering the sum of the individual SFR of all gas cells within  $2 \times R_{\text{star,h}}$ . The effects of different averaging timescales and of the aperture choice for the SFR of TNG galaxies has been extensively investigated in Donnari et al. (2019), where we have shown that the former (latter) are (not) negligible for the identification of the locus of the star-forming main sequence (MS). In the companion paper, Donnari et al. (2020), we emphasize that the aperture choice has non-negligible effects on the values of galaxies’ SFR, particularly when galaxies are being quenched, and

hence any quantitative comparison to the results of this paper must take this notion into account. We note that even if all galaxies in our sample are well-resolved in terms of stellar and gas resolution elements, this is not the case for their SFR values. We assign an SFR value extracted randomly in the range  $(10^{-5} - 10^{-4}) M_{\odot} \text{yr}^{-1}$  to those galaxies with Log SFR below the resolution limit of the simulation, which otherwise would have  $\text{SFR} \equiv 0$ .

## 2.4 Quenched versus star-forming galaxies

Throughout the text, we consider the adjectives “quenched”, “quiescent”, and “passive” as having interchangeable meanings and we do not distinguish between star-forming and star-burst galaxies, nor between quiescent and green-valley galaxies. In Donnari et al. (2019) and Donnari et al. (2020), we explore and compare a number of different criteria to separate quenched versus star-forming galaxies; here we use the following definition throughout the study.

*Quenched galaxies are those whose SFR is one dex below the star-forming main sequence.* Following the same procedure introduced in Donnari et al. (2019) and Pillepich et al. (2019), we stack all galaxies in 0.2 dex bins of stellar mass, measure the median SFR in the bin, and name as “quenched” those galaxies whose SFR falls below 1 dex from the median SFR at the corresponding mass. We then remove the quenched galaxies, calculate a new median, and repeat these steps iteratively until the median SFR in the mass bin is converged to a given accuracy. For this paper we have chosen a convergence criterion where the median SFR in successive iteration steps differ by less than 1 per cent. This allows us to find the locus of the star-forming Main Sequence (MS), that in TNG is linear in the  $M_{\text{stars}} = 10^9 - 10^{10.2} M_{\odot}$  range and can be fitted using Eq. (1) in Donnari et al. (2019). Galaxies whose SFR is 1 dex lower than the fitted MS at the corresponding stellar mass are flagged as quenched; for galaxies more massive than  $M_{\text{stars}} = 10^{10.2} M_{\odot}$  the linear extrapolation of the MS is used as reference.

We emphasize that the choice of extrapolating the locus of the MS beyond a certain mass and the precise stellar-mass range considered to fit the MS may have a non-negligible impact in the quantitative assessments of the quenched fractions, depending on redshift and particularly in regimes with low number statistics: we estimate the systematic errors associated with the latter choice to be as large as 10-20 percentage points (see Fig. 3 in Donnari et al. 2019 and Donnari et al. 2020). On the other hand, we note that, at low redshift, the separating criterion based on the MS used in this work is equivalent (to better than a few percentage points in quenched fraction) to a fixed  $s\text{SFR} \leq 10^{-11} \text{yr}^{-1}$  cut which is often used in low- $z$  environmental observational studies (Fig. 1 of Donnari et al. 2020) and at  $z \leq 2$  to UVJ cuts (Fig. 3 bottom panels of Donnari et al. 2019). Therefore, we estimate that the majority of the results in the following (those at  $z \lesssim 1$ ) are quantitatively robust to better than 10 per cent against quenched definition and we comment accordingly otherwise.

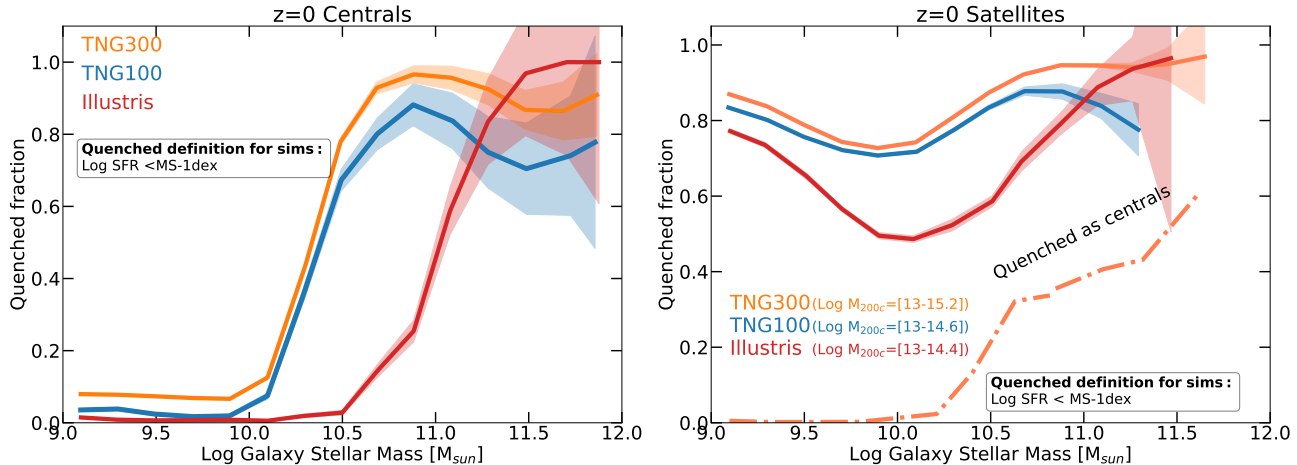
## 3 RESULTS FROM THE TNG SIMULATIONS

### 3.1 Fractions of quenched galaxies: centrals versus satellites at $z = 0$

In this section, we investigate the fractions of quenched galaxies as a function of galaxy stellar mass at  $z = 0$ , according to the TNG model. Our findings are presented in Fig. 1.

	TNG100		TNG300	
	All galaxies	Cluster galaxies	All galaxies	Cluster galaxies
	$M_{\text{stars}} > 10^9 M_{\odot}$	$M_{\text{stars}} > 10^9 M_{\odot}$ $M_{200c} = 10^{13-14.6} M_{\odot}$	$M_{\text{stars}} > 10^9 M_{\odot}$	$M_{\text{stars}} > 10^9 M_{\odot}$ $M_{200c} = 10^{13-15.2} M_{\odot}$
$z = 0$ Centrals	10942	182	149006	3733
$z = 0$ Satellites ( $< R_{200c}$ )	4595	2857	60184	39540

**Table 1.** Number of satellites and centrals in TNG100 and TNG300 at  $z = 0$ . The label “all galaxies” denotes the whole population while “cluster galaxies” indicates only satellites residing in groups and clusters, as explained in the text. In both TNG100 and TNG300, the number of central galaxies in groups and clusters is equal to the number of hosts contained in each simulation. The indicated stellar mass  $M_{\text{stars}}$  and the virial host mass  $M_{200c}$  are the minimum masses above which we gather our samples.



**Figure 1.** IllustrisTNG quenched fractions at  $z = 0$ . Fraction of quenched central (left panel) and satellite (right panel) galaxies at  $z = 0$  for TNG300 (orange), TNG100 (blue) and Illustris (red), in bins of 0.2 dex in stellar mass. Satellites are galaxies within the 3D  $R_{200c}$  of their host, excluding the central. We use the quenched definition of  $\text{Log SFR} < \text{MS} - 1\text{dex}$  here. The dash-dotted curve in the right panel represents the quenched fraction for satellites prior to any infall. At the low-mass end, the quenched fractions are significantly higher for satellites in groups and clusters compared to centrals of similar stellar mass.

We evaluate the  $z = 0$  quenched fraction separately for centrals (left panel) and satellites (right panel) in TNG300 (orange curve) and TNG100 (blue curve), and we compare the results with the original Illustris simulation (red curve), to better highlight the different efficiencies of the quenching mechanisms among these two numerical models. We consider all satellites in the simulated volumes that reside in hosts at least as massive as  $10^{13} M_{\odot}$ : hence the TNG100 and TNG300 curves represent somewhat different satellite populations as in the latter, satellites can be hosted by more massive haloes, up to  $\sim 10^{15.2} M_{\odot}$ . We plot the raw data, i.e. galaxy stellar masses and SFR values taken directly from the simulation output without any correction. We define quenched galaxies as those having SFR 1 dex below the MS. From now on, we only present results if there are at least 10 galaxies per bin. Shaded areas denote the Poissonian error per stellar mass bin.

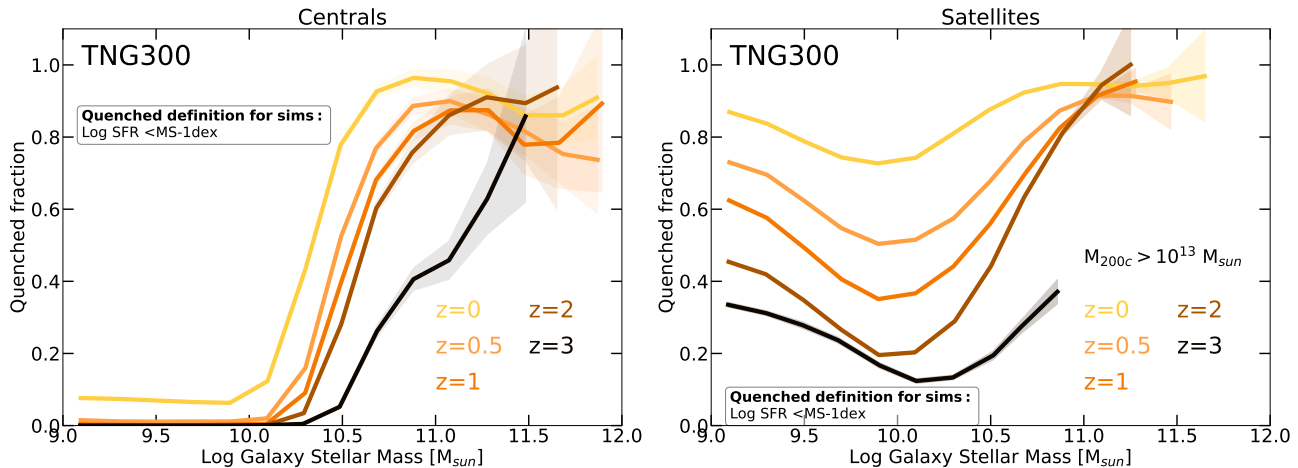
A first comparison between the two panels reveals a similar percentage of quenched galaxies more massive than  $\sim 10^{10.5} M_{\odot}$ , about 70-80 per cent in TNG (on average), whether they are satellites or centrals. This result is consistent with the mass-quenching scenario discussed in Section 1. In the TNG model, the halting of star formation and the reddening of galaxy colors in massive (isolated) galaxies is due to, or at least mediated by, AGN feedback and particularly by black hole driven winds at low-accretion rates (see Weinberger et al. 2017; Nelson et al. 2018a; Terrazas et al. 2020), which are also responsible for setting the physical and thermodynamical properties of gas within and around galaxies (see Truong et al. 2020; Davies et al. 2020; Zinger et al. 2020). Within

the framework of the TNG model, populations of massive central galaxies would *not* quench or redden at the levels expected by observational constraints (or at all) if the black hole driven winds at low-accretion rates were not in place (see e.g. Weinberger et al. 2017). In other words, within the TNG model, there are no secular or internal, population-wide quenching mechanisms other than AGN feedback.

Conversely, centrals at the low mass end ( $\lesssim 10^{10} M_{\odot}$ ) are rarely quenched, in contrast with group and cluster satellites of the same stellar mass: in TNG, the quenched fractions at the low-mass end read less than  $\sim 5$ -8 per cent<sup>2</sup> for centrals against about 80 per cent for satellites in groups and clusters, thus pointing to the fundamental role played by a dense environment in galaxy quenching in this mass range.

These results are qualitatively consistent with a series of observational findings. We refer the reader to the companion paper by Donnari et al. (2020) for a detailed comparison between the TNG results and observations at low redshift. There we show that the TNG quenched fractions at low redshift ( $z \lesssim 0.6$ ) are in the ballpark of a number of observational estimates for both centrals and satellites (their Fig. 8). More specifically, by comparing to SDSS data at  $z = 0.1$  and by creating detailed mocks via matching a number of analysis choices as in Wetzel et al. (2012, 2013), we demonstrate that the quenched fractions from TNG are in striking quan-

<sup>2</sup> We remind the reader that some of those low-mass centrals could be backplash galaxies.



**Figure 2. Trends with redshift.** Quenched fraction versus galaxy stellar mass for TNG300 centrals (right panel) and satellites (left panel) at five different redshifts:  $z = 0, 0.5, 1, 2,$  and  $3$ . The definition adopted to separate quenched versus star forming galaxies is  $\text{Log SFR} < \text{MS} - 1\text{dex}$ . Shaded areas in both panels indicate the Poissonian error. For both centrals and satellites the quenched fraction is lower at higher redshift.

titative agreement with [Wetzel et al. \(2012\)](#) for both centrals and satellites in the stellar mass range  $10^{9.5-11.5} M_{\odot}$ , when the latter are taken together for hosts with total mass between  $10^{12-15} M_{\odot}$  (their Fig. 9, top row).

The left panel of Fig. 1 shows a slight difference of about 10 percentage points between TNG100 and TNG300 at both low and high stellar masses. This offset is due to a combination of statistical sampling variance owing to the significant differences simulation box sizes as well as the impact of numerical resolution. We comment on these points in more detail in Appendix A. As a result of this offset, we attribute a systematic uncertainty of about 10 percentage points in the interpretation of the quenched fractions from the TNG100/TNG300 simulations. However, within the TNG framework, an even higher numerical resolution and poorer sampling statistics can impact these quenched fraction predictions with systematic uncertainties by up to 10-40 percentage points, depending on galaxy mass, host mass, and redshift (see Appendix A and [Donnari et al. 2020](#), for a discussion).

Interestingly, while the quenched fractions of centrals rise steeply from 5-8 to 80-90 per cent across the investigated stellar mass range, the quenched fractions of satellites are always as high as 80-90 per cent, although with a non-monotonic trend. We relate the slight depression at intermediate mass (more pronounced in the original Illustris, red curves) to a combination of both external (environmental) and internal (mass) quenching processes. Indeed, for low-mass satellites, the hosts in which they reside are likely the culprit for their quenching. However, at progressively higher stellar masses, the role of the environment, although still fundamental, seems to be less efficient, and as soon as the stellar mass reaches and exceeds a few  $10^{10} M_{\odot}$ , the quenched fractions rise again, mainly driven by AGN feedback (at least in the TNG model where no other internal quenching mechanisms are in place).

The orange dot-dashed curve in the right panel represents the fraction of satellites quenched prior to their infall into any host, i.e. before becoming satellites. For galaxies less massive than  $\sim 10^{10} M_{\odot}$ , this fraction is negligible, thus suggesting that they were mostly still star-forming when they were centrals. On the other hand, a non-negligible fraction of satellites that today are more massive than  $10^{10.5} M_{\odot}$ , have fallen into any host already quenched. In this case, the quenched fraction rises up to  $\sim 50$

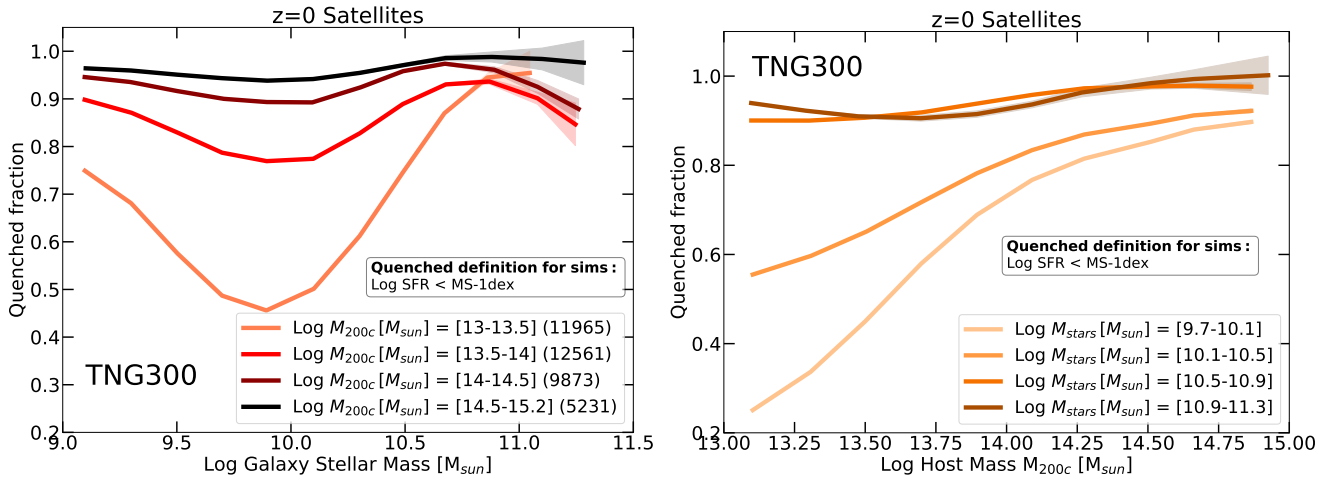
per cent. In Section 4.2, we discuss in more detail the role of pre-processing and internal processes at early times in shaping the  $z = 0$  quenched fraction of satellites.

While the qualitative trend with stellar mass is similar between the TNG and Illustris models, we note a quantitative difference between the two. As already highlighted in [Donnari et al. 2019](#), Illustris returns quenched fractions at fixed galaxy mass that are lower than TNG by up to 0.6 for centrals and 0.3 for satellites at the transitional mass regime of around  $10^{10-11} M_{\odot}$ . For central or high-mass galaxies, this difference could be interpreted as a shift towards higher galaxy masses for the onset of the quenching mechanism and is connected to the new and more effective AGN feedback recipes implemented in the TNG model ([Weinberger et al. 2017](#); [Pillepich et al. 2018](#); [Donnari et al. 2019](#)). For satellite galaxies, we speculate the difference to be due to two concurrent factors: firstly, low-mass Illustris galaxies are generally richer in both halo ([Pillepich et al. 2018](#)) and HI/cold ([Diemer et al. 2019](#)) gas than TNG galaxies because of the different implementation of stellar feedback; secondly, Illustris group-mass hosts are gas-poor compared to TNG and observations because of the strongly ejective AGN feedback implemented in the former ([Pillepich et al. 2018](#); [Terrazas et al. 2020](#)). These can arguably make Illustris satellite galaxies more resilient to gas loss and Illustris hosts less effective at ram-pressure stripping.

### 3.2 Quenched fractions across cosmic time

The evolution of the SFR-stellar mass plane and the increase of the total quiescent fractions with cosmic time are well established observationally (see e.g. [Speagle et al. 2014](#); [Whitaker et al. 2014](#); [Fang et al. 2018](#)) and theoretical results (e.g. [Donnari et al. 2019](#), for the TNG and Illustris simulations).

In Fig. 2 we show the quenched fractions of TNG300 centrals (left panel) and satellites (right panel), across five different redshifts:  $z = 0, 0.5, 1, 2,$  and  $3$  (from yellow curves, same as Fig. 1, to black curves). TNG100 (not shown) returns very similar results. We note here that the MS used to separate star-forming and quiescent galaxies is evaluated at each redshift, using the procedure described in Section 2.4 and always extrapolated for stellar masses larger than  $M_{\text{stars}} = 10^{10.2} M_{\odot}$ . Moreover, as demon-



**Figure 3. Trends with stellar mass and host mass.** Left: quenched fraction as a function of galaxy stellar mass for TNG300 satellites, stacked in four host mass bins, as indicated in the legend. Right: quenched fraction as a function of  $z = 0$  host mass, stacked in four stellar mass bins, as indicated in the legend. In both panels, the definition adopted to separate quenched versus star forming galaxies is  $\text{Log SFR} < \text{MS} - 1\text{dex}$  and shaded areas indicate the Poissonian error. The quenched fraction of satellites is higher in massive clusters than in groups and for massive galaxies compared to the low-mass ones.

strated in Figures 1-3 in [Donnari et al. \(2020\)](#), we warn the reader that different quenched definitions, SFR average timescales, and/or aperture choices can impact the quenched fraction measurements at  $z = 2 - 3$  by up to 70 percentage points.

Within the choices elected in this analysis, as expected, quenched fractions in TNG are generally lower at higher redshifts than at lower ones: while the trends with stellar mass are similar across redshifts, the fraction of quenched galaxies increases by up to 80 percentage points between  $z = 3$  and  $z = 0$  at fixed stellar mass, in both the central and satellite populations below  $\sim 10^{11} M_{\odot}$  ([Darvish et al. 2016](#); [Fossati et al. 2017](#)). A strong redshift evolution is in place even when a narrow range of host mass is considered for the satellites: even if not shown, we have verified that for satellites residing in hosts of  $10^{13-13.5} M_{\odot}$  at any redshift, the quenched fraction is lower by up to 40 percentage points at  $z = 3$  than at  $z = 0$ . Moreover we note that the steep rise in the quenched fraction for central galaxies occurs at lower stellar masses with decreasing redshift. On the other hand, at the very high mass end ( $\gtrsim 10^{11} M_{\odot}$ ), the quenched fraction is about 80-90 per cent regardless of the redshift.

Importantly, in TNG, a significant fraction ( $\gtrsim 40 - 50$  per cent) of galaxies is already quenched at  $z \sim 2$ , and more interestingly  $\sim 30 - 40$  per cent is quenched already at  $z = 3$ , both among centrals ( $\sim 10^{10-11} M_{\odot}$ ) and satellites, suggesting that both AGN feedback and environmental processes are already in action when the Universe was a few billion years old. This is, at least at face value, in remarkably better agreement than in previous models with the recent findings of ever more examples of very massive quiescent galaxies at  $z \gtrsim 3$  (see e.g. the observational findings by [Glazebrook et al. 2017](#); [Forrest et al. 2020](#) and Fig. 10 of [Donnari et al. 2020](#)).

### 3.3 Quenched fractions across host masses

In this Section, we proceed to investigate how the quenched fractions vary with environment; specifically, we focus on how the quenched fractions of satellites change across host masses.

Fig. 3 shows the quenched fraction of TNG300 satellites at

$z = 0$  versus galaxy stellar mass (left panel) and versus host mass (right panel), in bins of host and galaxy stellar mass, respectively.

We find that the non-monotonic trend of quenched fraction with stellar mass for satellites seen in the right panel of Fig. 1 is more pronounced in groups (orange and red curves in the left panel) than in clusters (brown and black curves), for which the quenched fraction is almost constant at 95-100 per cent, at any stellar mass. Moreover, as shown in the right panel, TNG returns a strongly rising trend in quenched fraction with host mass for low-mass satellites of  $M_{\text{stars}} \sim 10^{9.5-10.5} M_{\odot}$ , while – importantly – for the most massive satellites ( $M_{\text{stars}} \gtrsim 10^{10.5} M_{\odot}$ ) the quenched fraction is 90-100 per cent at any host mass (see e.g. [Davies et al. 2019](#), for similar results).

A combination of multiple physical mechanisms can be seen in these plots. Firstly, the independence on host mass of the quenched fractions for massive galaxies (orange and brown curves in the right panel of Fig. 3) suggests that either more massive satellites are less susceptible to environmental processes or are likely to be quenched mostly due to their internal secular processes, or a combination thereof. The fact that, at least in TNG, the quenched fractions of massive galaxies ( $M_{\text{stars}} \gtrsim 10^{10.5-11} M_{\odot}$ ) are beyond 70-80 per cent whether they are centrals or satellites (see also Fig. 1) strongly points to internal quenching processes being the main actors also in high-density environments: in the TNG model, the only internal mechanism that is effective at quenching is feedback from super massive black holes, as already discussed in Section 3.1. Secondly, more massive hosts are likely to impart stronger satellite quenching effects, thus quenching their galaxies more efficiently. However, as we will discuss in more detail in Section 4.2, one must take into account that a satellite might be orbiting in other groups before being accreted into the current one. As a result of this group pre-processing, these satellites might fall into the  $z = 0$  host already quenched.

All the results of Fig. 3 are qualitatively consistent with SDSS-based results compiled by [Wetzel et al. \(2012, 2013\)](#), who focus on galaxies more massive than  $10^{9.7} M_{\odot}$ . In our companion paper ([Donnari et al. 2020](#)) we demonstrate that the agreement with SDSS in quenched fractions is best (and really good, within 5 percentage points) at intermediate host masses



( $10^{13-14} M_{\odot}$ ) and galaxy stellar masses ( $10^{10-11} M_{\odot}$ ), while the TNG quenched fractions of  $10^{9.5-10} M_{\odot}$  satellites in more massive hosts ( $10^{14.5-15} M_{\odot}$ ) are higher by up to 20 percentage points than those inferred from SDSS.

### 3.4 Satellite phase-space diagram, $z = 0$ radial distribution of satellites within $R_{200c}$ and connection to the time since last infall

Observations suggest that the satellite quenched fractions also depend on location within a host (see Introduction). The top panel of Fig. 4 shows the median fraction of quenched galaxies in TNG300 at  $z = 0$  in the plane of satellite radial velocities (normalized to the host virial velocity measured as  $\sqrt{G(M_{200c}/R_{200c})}$  versus their 3D distance to the cluster center (normalized by the virial radius  $R_{200c}$ ). Averages are computed in 2D bins of size 0.1 dex in distance and 0.3 dex in velocity space.

A clear trend is manifest at any fixed value of  $V_{\text{radial}}/V_{\text{virial}}$ : the median quenched fraction decreases at increasing distance to the host center. As we demonstrate later when discussing the middle panels, this trend is mostly due to low-mass satellites, while the fraction of the massive satellites is about 90 per cent regardless of the distance from the host centre. Additionally, at a fixed small distance from the cluster center ( $\lesssim 0.6R_{200c}$ ), infalling ( $V_{\text{radial}} < 0$ ) and outgoing ( $V_{\text{radial}} > 0$ ) satellites exhibit nearly identical quenched fractions. On the other hand, infalling satellites residing at a distance larger than  $\sim 0.6R_{200c}$  show a quenched fraction lower than the outgoing satellites, at the same cluster-centric distance. We speculate that satellites that have already orbited close to the host center and have spent more time in the hosts – as those in outbound orbits and positive radial velocity – are more likely to be quenched compared to those that are still infalling or on their first infalling trajectory.

An analogue trend, in overall qualitative agreement with our results, is described in Pasquali et al. (2019), where the authors combine SDSS group catalogues and the YZiCS hydrodynamical simulations (Choi & Yi 2017). Their results reveal that the passive fraction increases – from 50 per cent to more than 90 per cent – at decreasing projected cluster-centric distance.

In the middle panels of Fig. 4 we therefore show quenched fractions versus 3D distance by stacking satellites in four host mass bins (right panel) and four stellar mass bins (left panel). In general, we find that the fraction of low-mass passive galaxies is higher closer to the host center, consistent with the top panel results just discussed. However, the trend with cluster-centric distance is stronger the lower the host mass is. Importantly, this trend is very weak, if not nonexistent, in the case of massive satellites ( $\gtrsim 10^{10.5} M_{\odot}$ ) and satellites in massive hosts ( $\gtrsim 10^{14} M_{\odot}$ ), for which the quenched fractions are larger than 80-90 per cent irrespective of distance from the host center. Overall, as shown in Donnari et al. (2020), the trend with halocentric distance is well consistent with SDSS estimates (Wetzel et al. 2012), with the level of (dis)agreement depending on host and galaxy stellar mass ranges, being best (and in fact, excellent) for intermediate hosts of  $10^{13-14} M_{\odot}$  and worst for the lowest mass satellites in either the lowest ( $\sim 10^{12} M_{\odot}$ ) or highest mass hosts ( $10^{14.5-15} M_{\odot}$ ).

This confirms, once again, the fundamental role played by the environment in quenching low-mass satellites, while the quenching of massive ones is almost independent of the host in which they reside and hence, arguably mainly related to their internal processes.

A satellite’s location within a host is correlated with the time since accretion: satellites that fell into their current hosts a longer

time ago are typically found closer to their host centers (Rhee et al. 2017). Therefore the trends of the top panels of Fig. 4 can be attributed to two concurrent facts: smaller clustercentric distances correspond to denser regions where environmental effects can be larger but also, even though individual orbits of satellites can span a broad range of radii, satellites orbiting closer to the cluster center have typically spent a longer time within their group or cluster environments: if the fraction of quenched satellites depends on how long they have been orbiting in their  $z = 0$  host (a plausible assumption, see also Li et al. 2020, for similar results), then the probability of being quenched can be higher closer to the host centers than in the outskirts, irrespective of environmental density.

In the bottom panels of Fig. 4 we show the fraction of quenched satellites as a function of the time since their last infall into the current host (see Section 4.2 for a detailed definition of infall times), by binning galaxies by their present-day host mass (left panel) and their present-day stellar mass (right panel). In both panels, the fraction of galaxies is evaluated including (dashed curves) or excluding (solid curves) satellites that were already quenched at the time of their accretion onto the current host – we postpone to Section 4.2 for a more detailed analysis on pre-processing.

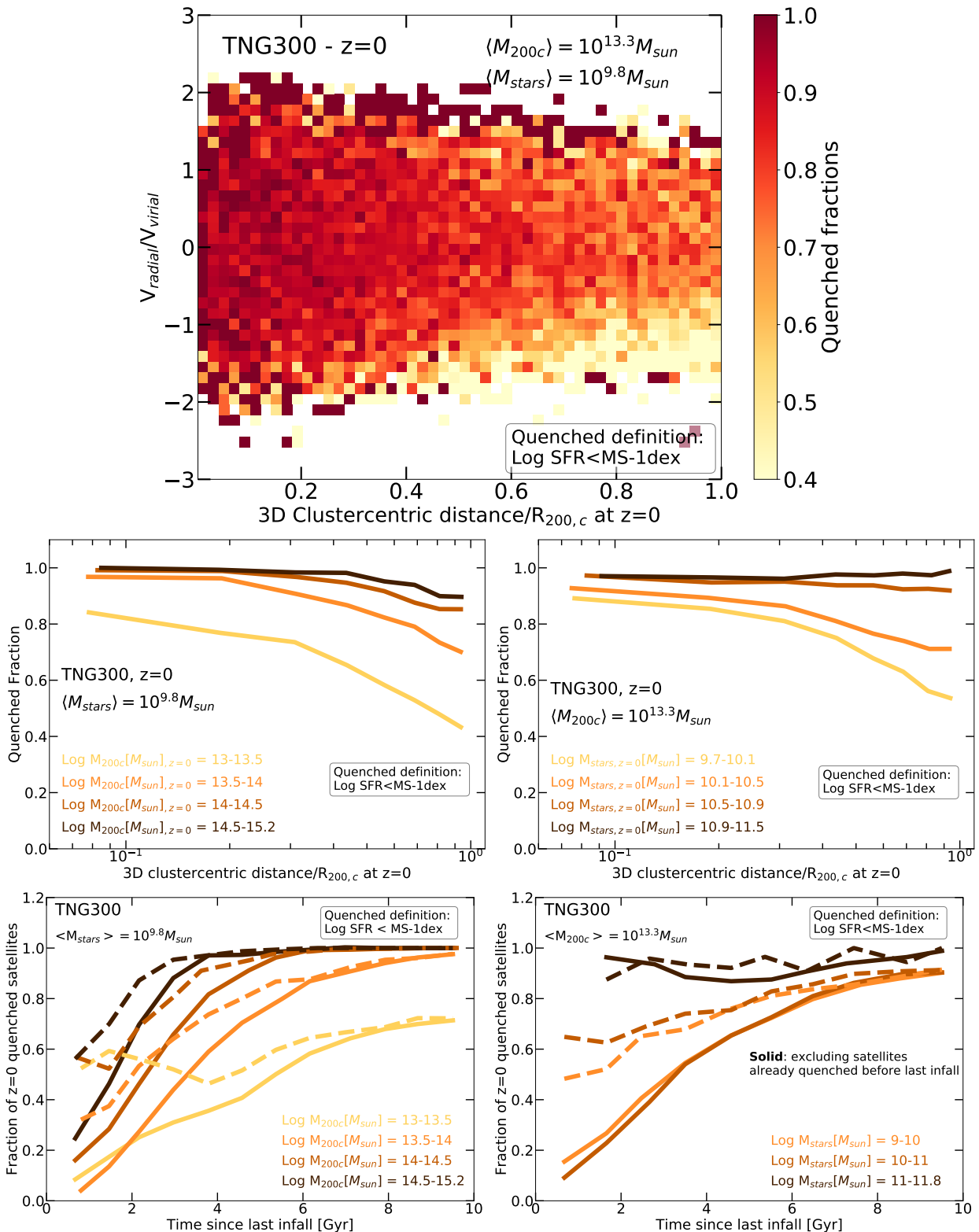
A clear result is evident: the fraction of  $z = 0$  quenched satellites is higher for those accreted earlier into the current host, this correlation being stronger if we remove galaxies already quenched at infall (solid versus dashed curves) and for more massive hosts and lower-mass satellites. Interestingly, the trend flattens completely at 80-100 per cent for satellites that were accreted longer than 4-6 Gyr ago in hosts more massive than  $10^{13.5} M_{\odot}$ : this suggests an upper limit of some billion years for the time needed for the group and cluster environment to have an effect and to induce quenching. Allowing for some delay, this estimate is in line with the results of Yun et al. 2019, where we found that jellyfish galaxies (i.e. galaxies with markedly visible indications of ram-pressure stripping) are recent infallers according to the TNG simulations, i.e. they fell no more than 2-3 billion years before being identified as jellyfish galaxies.

## 4 DISENTANGLING QUENCHING PATHWAYS

In Section 3, we have shown that in TNG, massive galaxies ( $M_{\text{stars}} \gtrsim$  a few  $10^{10} M_{\odot}$ ) display high quenched fractions, regardless of whether they are centrals or satellites and in the latter case, regardless of host mass, cosmic time, cluster-centric distance and time since infall. This would suggest that the massive satellites observed to be quenched at  $z = 0$  could have halted their star formation before ever becoming satellites, or even if they were quenched as satellites, could have done so due to internal or secular quenching mechanisms<sup>3</sup> rather than any environmental effects.

Similarly, in the previous Section we have shown that regardless of galaxy stellar mass and cluster-centric distance, the fraction of quenched satellites orbiting in massive clusters of  $M_{200c} \gtrsim 10^{14} M_{\odot}$  is higher than 80-90 per cent. A large body of work (see e.g. Henriques et al. 2017, and references therein), mostly based on analytical arguments and controlled experiments, supports the idea that would be deduced from this finding, namely that more massive

<sup>3</sup> As discussed in the previous Sections, in the framework of the TNG model there are no emerging or explicitly-modeled internal or secular quenching mechanisms other than the feedback from super massive black holes. Therefore, in the following discourse, we will associate secular/internal quenching mechanisms with AGN feedback.



**Figure 4. Dependence on cluster-centric distance and time since last infall.** Top: phase-space diagram for TNG300 satellites ( $M_{\text{stars}} \geq 10^9 M_{\odot}$ ) in a range of different hosts ( $M_{200c} \geq 10^{13} M_{\odot}$ ) at  $z = 0$  where the color indicates the median of the quenched fractions in each pixel of the grid. Middle row: fraction of quenched galaxies in TNG300 at  $z = 0$  as a function of their 3D distance from the cluster center, scaled to the host virial radius  $R_{200c}$  with galaxies stacked according to their present-day halo mass  $M_{200c}$  (left panel) and their present-day satellite mass (right panel). Bottom: Fraction of quenched galaxies (at  $z = 0$ ) in TNG300 versus the time since their last infall, stacked in four host mass bins (left panel) and three stellar mass bins (right panel). In both panels, masses are measured at  $z = 0$  and the quenched fraction is evaluated including (dashed curves) or excluding (solid curves) satellites already quenched at the time of their last infall. In all plots, the quenched definition is  $\text{Log SFR} < \text{MS}-1 \text{dex}$ . The fraction of passive galaxies is higher closer to the host center. This is not the case for massive satellites, for which the quenched fractions are  $\sim 90$  per cent irrespective of their distance from the host. Early infallers are more likely to be quenched compared to recent infallers.

clusters are more efficient than lower-mass groups at quenching their satellite galaxies. This would be because they provide higher-density environments, whether in terms of deeper potential wells for the purposes of tidal stripping and galaxy encounters, higher galaxy number density for the purposes of harassment, or higher ambient gas density for the purposes of ram-pressure stripping.

However, the hierarchical growth of structure complicates the straightforward interpretation of e.g. Figs. 1 and 3, because the most massive clusters in today’s Universe formed from the merging of smaller groups that in turn have hosted their own satellite populations. Furthermore, hosts of different total mass exhibit different infall time distributions of their galaxy members (Engler et al. 2020) and the longer the exposure to environmental processes, the higher the probability that a galaxy quenches (Section 3.4).

In this section we therefore proceed to address the following questions: do high-density environments quench their satellites more efficiently with respect to lower-mass groups even after taking into account pre-processing? And what is the role of lower-mass groups in quenching galaxies prior to their accretion onto the  $z = 0$  host? Theoretical models offer a unique opportunity to shed light on the timing and circumstances whereby quenching occurs, because of their capability to follow galaxies back in time i.e. for example to follow satellite galaxies until their infall into any host. We hence use the TNG simulations to systematically address these questions in the following subsections.

#### 4.1 The concept of pre-processing

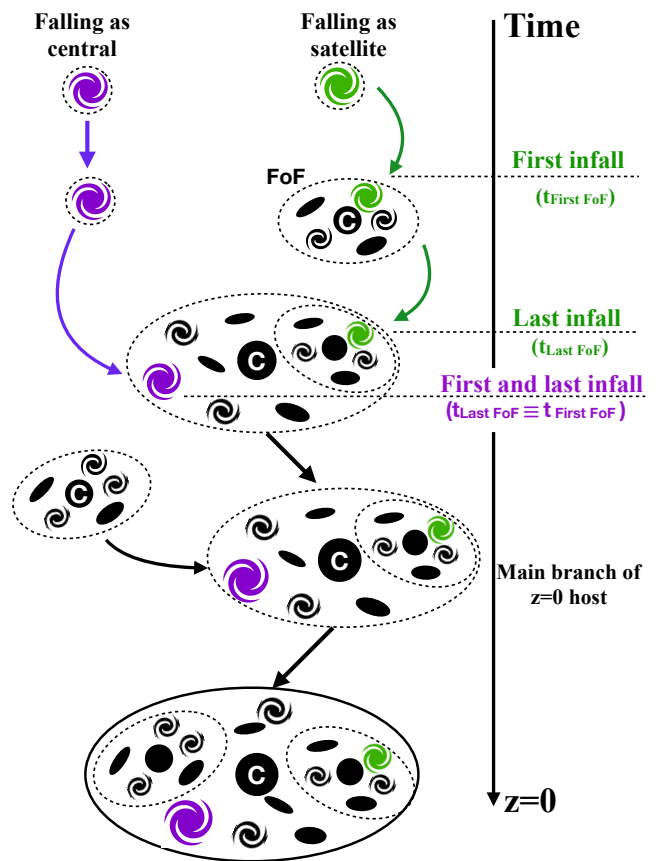
As mentioned in Section 1 and already alluded to in the above TNG-based analyses, both observations and models recognize that environmentally-driven transformations may occur in satellites even before they were accreted into their current host. Because of the nature of the hierarchical growth of structure, we expect the  $z = 0$  population of satellites residing today in groups and clusters to also include galaxies that orbited in lower-mass hosts before falling into (the progenitor of) their current host: this mechanism is commonly referred to as *pre-processing*.

Particularly, in this paper we describe as *pre-processed* all those satellites that have been residing in “subgroups” of total mass of at least  $10^{12} M_{\odot}$  before being accreted by the progenitor of their host at  $z = 0$  (or at the time of observation). The choice of a minimum pre-processing host mass is required to carry out an even-handed analysis between simulations with varied resolutions or halo identification definitions. Here, we have employed a minimum halo mass of  $10^{12} M_{\text{stars}}$  because we are interested in identifying galaxies that may have experienced some levels of environmental processes in the past, e.g. environmental processes that may be capable of quenching galaxies.

In fact, galaxies that are found today in groups and clusters may have followed a diverse range of pathways to become satellites. For simplicity, in Fig. 5 we illustrate the two main possible scenarios. A galaxy may directly fall into the main branch of its  $z = 0$  host and remain there until  $z = 0$  (purple galaxy). We name this category **falling as centrals**. On the other hand, a satellite may be accreted by a subgroup at  $z > 0$  that, due to the hierarchical growth of structure, might in turn merge into the main progenitor of a  $z = 0$  group or cluster. We name these satellites **falling as satellites** or, indeed, **pre-processed**.

Note that our definition of pre-processing so far does not yet assert anything as to whether the galaxy properties are affected by their pre-processing host. In fact, how much the pre-processing contributes to the present-day quenched fractions of cluster galax-

### The diverse pathways galaxies become satellites of groups and clusters



**Figure 5. The concept of pre-processing and the evolutionary pathways of  $z = 0$  satellites.** The schematic cartoon describes two possible evolutionary tracks for satellites inhabiting groups or clusters at  $z = 0$ . A galaxy might be directly accreted into the main branch of the  $z = 0$  host and thus have only one infall time (falling as centrals, purple). On the other hand, a galaxy might first fall into a subgroup rather than into the  $z = 0$  one, which in turn will merge in the final  $z = 0$  host (falling as satellite, green). We call these satellites “pre-processed”. Clearly, for some galaxies the situation might be more complex, as we discuss in the text.

ies is still a matter of debate (see Section 1) and will be addressed in Section 4.4.

Furthermore, the analysis of cosmological simulations clearly shows that the situation is generally more complex than the one depicted in Fig. 5 and other pathways are possible. For example, a galaxy, after being captured by a subgroup, might have an orbit sufficiently wide to leave the host after a few hundreds of Myr, or in some cases after a few Gyr, thus becoming a central again. It could then be accreted by a subgroup or directly onto the main branch of the final host. Alternatively, a galaxy, after being accreted by its final host, can escape again for a non-negligible amount of time and then come back into the main branch of the  $z = 0$  host. Even though these further situations are plausible, throughout the paper we consider those scenarios as part of the “pre-processed” one for simplicity, without making any distinction in our plots unless otherwise indicated.

Infall time	Description
$t_{\text{FirstFoF}}$	The time when a galaxy becomes satellite of any FoF (of $M_{200c} > 10^{12} M_{\odot}$ ) for the first time.
$t_{\text{LastFoF}}$	The time when a galaxy is accreted onto its last host, remaining there until $z = 0$ .

**Table 2.** Description of the two infall times used in this work (see also Fig. 5 for a visual schematic).

#### 4.1.1 Infall times for galaxies falling as satellites and as centrals

The search and labeling of pre-processed satellites is practically implemented in our analysis via the definition of two different infall times. As a reminder, our sample of satellites is selected at  $z = 0$  and includes galaxies with stellar mass  $\geq 10^9 M_{\odot}$  residing within one virial radius  $R_{200c}$  of their host with  $M_{200c} \geq 10^{13} M_{\odot}$ . The so-called backsplash galaxies, i.e. satellites found today outside the virial radius of the host where they were previously orbiting, are hence neglected. Furthermore, the time evolution of each satellite in the simulations is followed using the `SUBLINK` algorithm, which allows us to reconstruct the merger trees at the subhalo level and to study the properties of each galaxy and its host progenitor by following the main branch of the merger tree (Rodríguez-Gomez et al. 2015).

Tracing a galaxy back in time, we can define (at least) two different infall times<sup>4</sup>, as summarized in Tab. 2 and depicted in the schematic of Fig. 5:

(i) The time when a galaxy becomes a satellite of a FoF for the first time: we will refer to this time as  $t_{\text{FirstFoF}}$ . We note here that, in order to consider a galaxy as part of a (different) halo, we require a minimum of three consecutive snapshots in which it is a satellite and a minimum host mass (at the time of infall) of  $10^{12} M_{\odot}$ .

(ii) The time when a galaxy falls – for the last time – into the main branch of its current host and remains there until  $z = 0$ : we will refer to this time as  $t_{\text{LastFoF}}$ .

Therefore, in the case of satellites falling as centrals, the two infall times described above coincide:  $t_{\text{FirstFoF}} \equiv t_{\text{LastFoF}}$ . On the other hand, for galaxies falling as satellites (i.e. pre-processed) we can identify two distinct infall times, with the accretion into a first host occurring at an earlier cosmic time with respect to the accretion into the current host:  $t_{\text{FirstFoF}} < t_{\text{LastFoF}}$ .

## 4.2 Demographics of pre-processed satellites in the TNG simulations

With our definition of pre-processing and of satellites falling as centrals or as satellites in hand, we now proceed to study the demographics of these two populations in the TNG simulations. We first aim to quantify how *frequent* pre-processing is, based on the histories and orbits of TNG satellites and hosts, and with no regard for the internal properties of galaxies. In the next Section, we will quantify how *effective* pre-processing is at determining the  $z = 0$  quenched fractions.

<sup>4</sup> In principle, the definition of infall time can be different from work to work, often invoking the crossing of the virial radius. For clarity, in this paper we only use the two defined in the text and we postpone to a future study the task of investigating in more detail the implication of different choices.

In the top panel of Fig. 6, we show the fraction of  $z = 0$  TNG300<sup>5</sup> galaxies falling as satellites (solid curves), as a function of  $z = 0$  galaxy stellar mass, stacked into four  $z = 0$  host mass bins, as labeled in the legend. The shaded areas denote the Poissonian error in each mass bin. For completeness and to guide the eyes, we also show the fractions of galaxies directly falling as centrals into the main branch of their current host (dashed curves), these curves being complementary to the solid ones.

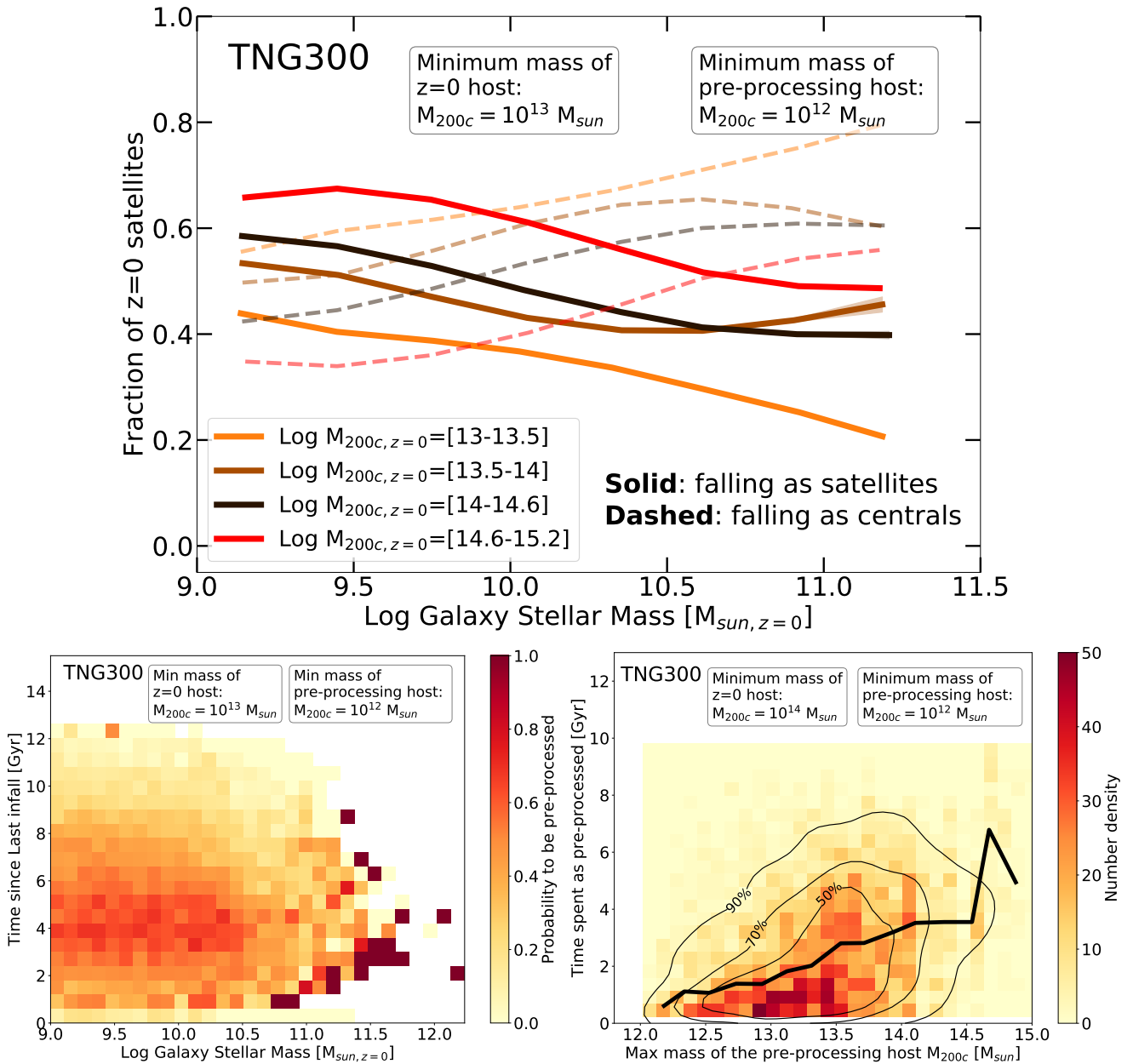
Overall, pre-processing in TNG is predominant in clusters ( $\gtrsim 10^{14} M_{\odot}$ ) compared to low-mass groups ( $< 10^{13.5} M_{\odot}$ ), by about 20 percentage points at any stellar mass, and slightly decreases with increasing galaxy stellar mass. We ascribe this host-mass dependence of the fraction of pre-processed satellites to the nature of the hierarchical growth of structures. About 65 (40) per cent of  $10^{9-9.5} M_{\odot}$  satellites residing in hosts of mass  $> 10^{14.6} M_{\odot}$  ( $10^{13-13.5} M_{\odot}$ ) have previously been part of groups of at least  $10^{12} M_{\odot}$  and then fallen into (the progenitor of) their final host. For more massive satellites of  $10^{11} M_{\odot}$ , this fraction drops to 50 per cent in massive clusters and to 20 per cent in low-mass groups. These results are consistent with those by Bahé et al. (2013, 2019), which are also based on cosmological hydrodynamical simulations of large sets of hosts.

The bottom left panel of Fig. 6 shows the average probability of TNG satellites having been pre-processed, i.e. orbiting in other previous hosts of at least  $10^{12} M_{\odot}$ , as a function of the time since last infall and as a function of  $z = 0$  galaxy stellar mass. Interestingly, we find a non-monotonic correlation between the probability of being pre-processed and time since last infall, with the highest probability occurring for satellites that were accreted into their final host between 3 and 6 Gyr ago. As sensible, satellites that fell in more than 8 Gyr ago have a low probability of being pre-processed by other groups – less than 30 per cent – and this is also the case for satellites accreted between 1 and 2 Gyr ago. Additionally, the few satellites that today are more massive than  $10^{11.5} M_{\odot}$  were almost certainly part of other subgroups, falling into their current host between 2 and 8 Gyr ago.

It is likely that the time the galaxies have spent in their pre-processing hosts and the mass of the preprocessing hosts are both key factors in determining how much the galaxy is affected by pre-processing (Han et al. 2018). Therefore, we show the distributions of both these properties in the bottom right panel of Fig. 6. The color denotes the galaxy number density in the depicted plane while the contours encompass different fractions of satellites, divided into 2D bins of size 0.75 dex in maximum mass of the preprocessing host and 0.25 Gyr in time spent as preprocessed satellite. The running median is represented by a solid black line.

The maximum mass of the host is evaluated as the peak mass of the group in which a satellite resides, right before being captured by the  $z = 0$  one. Since pre-processed satellites may have been part of more than one previous host and since any such hosts will also increase in mass with time, we must choose a representative definition for the mass of the preprocessing hosts. We focus on the maximum mass of the host to gauge the maximal effect of pre-processing.

<sup>5</sup> Although not shown, the same qualitative trends with both stellar and host mass also hold in TNG100. However, we note that due to a combination of different sampling of differently-sized volumes (chiefly) and different numerical resolution (to a lesser degree), the fraction of preprocessed satellites residing in low-mass hosts of  $10^{13-13.5} M_{\odot}$  in TNG300 is about 10 per cent higher with respect to TNG100.



**Figure 6. Pre-processing statistics for group and cluster satellites.** Top: fraction of TNG300  $z = 0$  satellites that have been pre-processed (i.e. fallen in as satellites, solid) or have fallen in as centrals (dashed), as a function of galaxy stellar mass. The fraction of pre-processed satellites is higher in higher mass hosts at  $z = 0$ . Bottom left: time since last infall of TNG300 satellites (into the  $z = 0$  host) as a function of galaxy stellar mass. The color represents the probability of having been pre-processed by other hosts with a minimum mass of  $M_{200c} > 10^{12} M_{\odot}$ , before falling into the  $z = 0$  host. Bottom right: time spent as a pre-processed satellite as a function of the peak host mass. The color denotes the number density of galaxies while contours encompass 50, 70 and 90 percent of the galaxies sample. The higher the mass of the pre-processing host, the longer the time a satellite spent there.

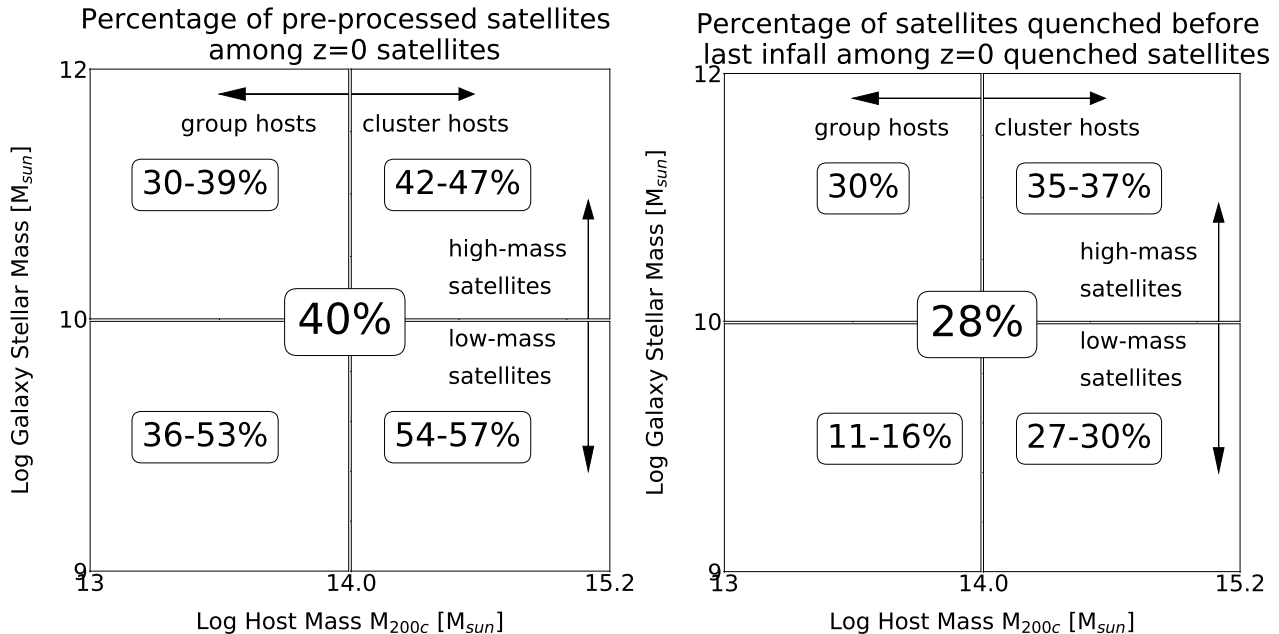
Finally, we show only those galaxies that remain satellites until  $z = 0$ , never leaving their host, so that the y-axis is effectively the time spent in any host (of at least  $10^{12} M_{\odot}$ ) before the  $z = 0$  one. Additionally, it is worth noticing that the y-axis of Fig. 6 has to be intended as the “total” amount of time that a satellite may have spent in other hosts – even if there were more than one such hosts.

As expected, Fig. 6 shows that the more massive the host (peak mass) in which a satellite is pre-processed, the longer the time it spent there. Of more relevance is the fact that galaxies can be pre-processed for a long time: for as long as 3-4 Gyr on average in hosts more massive than  $10^{14.5} M_{\odot}$ , and up to 5-6 Gyr in the tails

of the distributions. On the other hand, satellites have spent on average less than 2 Gyr in pre-processing hosts of maximum mass  $10^{12-13} M_{\odot}$ .

### 4.3 Fractions of pre-processed and pre-infall quenched satellites

In the top panel of Fig. 6 we have demonstrated that pre-processing is more frequent among satellites currently residing in clusters than in groups. The left panel of Fig. 7 shows an integrated version of



**Figure 7. Percentages of pre-processed satellites (left) and of satellites quenched before last infall (right).** The fractions given in the left represent the number of  $z = 0$  satellites that have been pre-processed, i.e. have orbited in different subgroup prior to infall into their  $z = 0$  host. The right panel depicts the fractions of satellites that quenched before falling into their current host over the  $z = 0$  quenched population. Each panel is divided into four tiles, denoting low and high stellar masses (below or above  $\sim 10^{10} M_{\odot}$ ) and residing in groups and clusters (below or above  $\sim 10^{14} M_{\odot}$ ). The numbers in each tile bracket the results from TNG100 and TNG300, respectively. Pre-processing occurs more frequently in clusters than in groups, regardless of stellar mass. Among  $z = 0$  quenched satellites, about 30 per cent are already quenched at the time of their last infall.

Fig. 6: the percentage of TNG<sup>6</sup> satellites residing in hosts of at least  $10^{12} M_{\odot}$  before the current host, which is about 40 per cent of the whole population (central box) (see also McGee et al. 2009; Han et al. 2018, for similar results). We further divide the  $z = 0$  satellite sample into four tiles – low and high-mass satellites residing in group or cluster hosts, all evaluated at  $z = 0$  – and show the percentage of pre-processed galaxies in each of them. This picture clearly shows that pre-processing is more common in clusters than in groups, for both low and high-mass satellites, as shown above. Additionally, the frequency with which this occurs is higher for satellites below  $10^{10} M_{\odot}$ , for which TNG returns about 54-57 per cent of pre-processed satellites residing today in clusters.

The fractions of Fig. 7, left, do not tell us yet whether pre-processing has had an effect on the internal properties of satellite galaxies. Nevertheless, due to the high percentage of pre-processed satellites (Figs. 6 and 7, left) and the long amounts of time that they might have spent in other subgroups (Fig. 6, bottom), it is plausible that a non-negligible fraction of satellites fall into their current host already quenched and that such a fraction may be a significant contribution to the overall  $z = 0$  quenched fractions. We have discussed this phenomenon in Section 1 and already shown its impact on quenched fractions for TNG in Fig. 4. In Fig. 7, we explicitly provide the numbers of TNG satellites that were already quenched at the time of infall into in their current host – either as centrals

(before any infall) or as satellites – normalized by the number of  $z = 0$  quenched satellites.

Among the  $z = 0$  quenched satellites, about 28 per cent are already quenched at the time of last infall. The percentages in the four tiles (representing the same satellites population of the left panel) demonstrate that  $\sim 35$ -37 (30) per cent of satellites more massive than  $10^{10} M_{\odot}$  residing in clusters (groups) quench before falling into their current host, compared to  $\sim 27$ -30 (11-16) per cent of low-mass satellites.

It is important to realize that the (higher) percentages of massive satellites that were already quenched before their last infall also include those that quenched when they were centrals, likely because of their AGN feedback, and not necessarily because of pre-processing, as already discussed in Section 3.1. Hence, in the next section we proceed to better investigate the diverse pathways towards the quenching of satellites, detailing what occurs at different stellar masses.

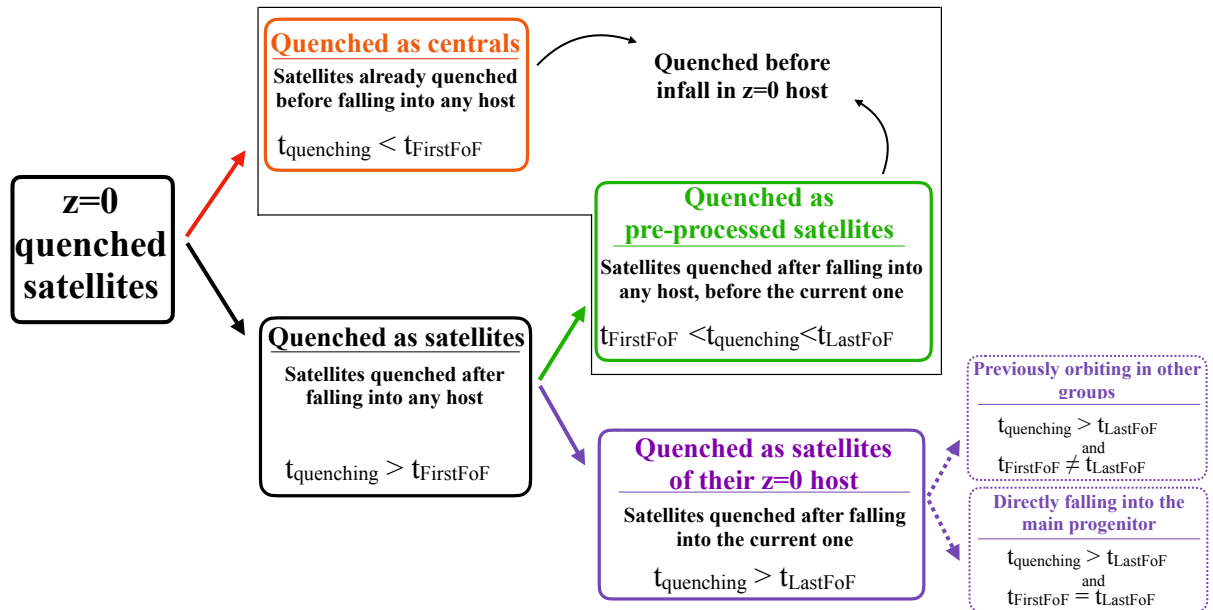
#### 4.4 The effect of pre-processing and internal processes on the $z = 0$ quenched fractions of satellites

A detailed inspection of the histories of TNG galaxies reveals that satellites found at  $z = 0$  in groups and clusters may have followed diverse pathways to quenching: these pathways are schematically summarised in Fig. 8.

Firstly, satellites can be quenched before any infall because of their internal processes (**quenched as centrals**, orange box) or after being accreted into any hosts (**quenched as satellites**, black box). Furthermore, the latter category can be further split into two, based on whether a satellite was quenched as it was orbiting in its current host (**satellites quenched in their  $z = 0$  host**, purple box) or when it was a member of other pre-processing hosts (**satellites**

<sup>6</sup> In all fractions presented here we bracket the outcome from TNG100 (leftmost numbers in each tile) and TNG300 (rightmost numbers in each tile). The differences between TNG100 and TNG300 are mostly driven by the fact that the two simulations represent the evolution of two different volumes, one larger than the other, and with different relative contributions to the galaxy populations of those residing in very massive hosts.

## The diverse pathways towards quenching of satellite galaxies that are found at $z=0$ in groups and clusters



**Figure 8.** The diverse pathways towards the quenching of satellite galaxies that are found at  $z = 0$  in groups and clusters. Satellites residing today within groups and clusters may have followed different pathways to quenching. Some of them could have halted their star formation because of their AGN feedback, prior to their accretion into a dense environment (orange box). Others may have quenched after falling into a host, either within a pre-processing host (green box) or within the current one (purple box), the latter category also including satellites previously orbiting in other hosts (pre-processed but not quenched) or directly falling in the  $z = 0$  one. All these pathways are responsible for the trends in quenched fractions and magnitudes of e.g. Figs. 1, 3, 4.

**quenched as pre-processed satellites**, green box). Satellites that either quenched as centrals or while being pre-processed are both part of the satellites that were quenched before their last infall into the  $z = 0$  host (as depicted by the common black box). Their frequency has been given and discussed with Fig. 7, right panel: about 30 per cent of the  $z = 0$  quenched population in the studied mass ranges. Additionally, even if a satellite was actually quenched in its  $z = 0$  host, this does not rule out the possibility that it was orbiting in a different host before infalling in the current host (purple dashed boxes) and that this pre-processing might have partially contributed to the quenching.

In Fig. 8 we also detail the operational definitions adopted to practically assign  $z = 0$  satellites to these various categories. These are based on the comparison between the infall times, as explained in Section 4.1.1 and Tab. 2, and the quenching time. The latter is defined as the last time the satellite’s SFR falls below 1 dex from the main sequence, evaluated using the procedure described in Section 2.4.

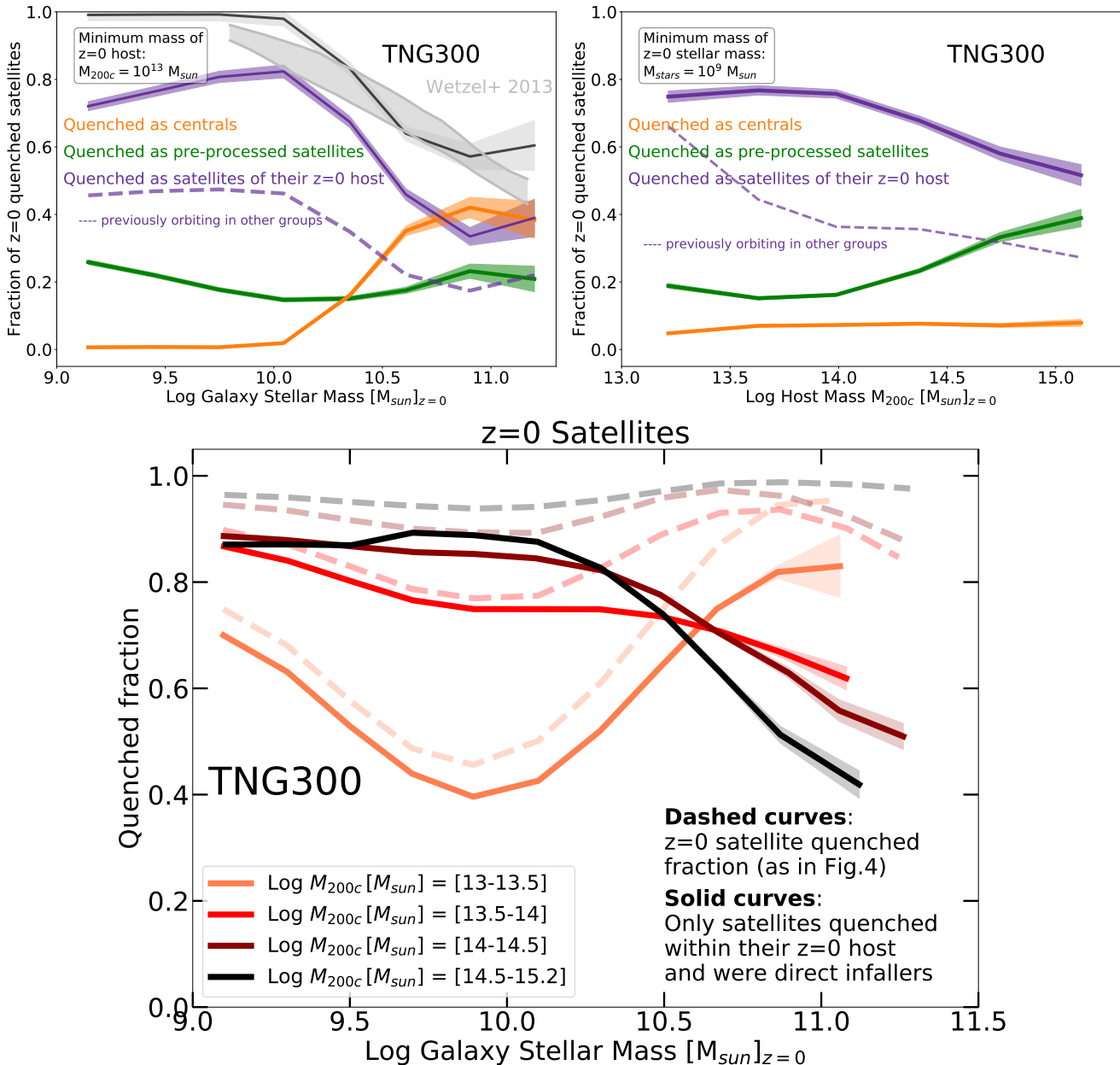
The top panels of Fig. 9 show the fractions of  $z = 0$  quenched satellites in TNG300 that were quenched via the different pathways summarized in Fig. 8, as a function of current galaxy stellar mass (left panel) and current host mass (right panel). Even if not shown, the following results also hold for TNG100. The fractions are evaluated as follows: in each mass bin of 0.2 dex, we count the number of galaxies quenched in one of the aforementioned scenarios and divide by the total number of  $z = 0$  quenched satellites in that bin. The shaded areas represent the Poissonian error and we only show mass bins with more than 10 satellites.

We find that satellites more massive than a few  $10^{10} M_{\odot}$  today are the only ones that could quench when they were centrals (orange curve, left panel), i.e. before any infall: in TNG this is due to their AGN feedback and is in line with our findings already discussed in Section 3.1. At the very high-mass end ( $\gtrsim 10^{10.5} M_{\odot}$ ) this percentage is, on average,  $\sim 40$  per cent of the  $z = 0$  quenched population. Across the considered stellar mass range, this fraction is independent of the mass of the host where they are found today (orange curve, right panel).

On the other hand, quenched satellites that were quenched within their current host (purple curves) exhibit a strong correlation with galaxy stellar mass (left panel): the fraction is  $\sim 80$  per cent for low-mass satellites  $M_{\text{stars}} \lesssim 10^{10} M_{\odot}$  and decreases to  $\sim 40$  per cent for satellites more massive than  $10^{10.5} M_{\odot}$ .

We omit showing a curve representing the fraction of all satellites that were quenched *before* their last infall, for which we have provided and discussed the percentages in the right panel of Fig. 7: this would be given by the sum of satellites quenched as centrals (orange curves) and of those quenched in any hosts other than their current ones (quenched as pre-processed, green curves).

The fraction of  $z = 0$  quenched satellites that quenched as satellites i.e. after any infall in any host is represented by the black curve: this is given by the sum of those quenched as pre-processed satellites (green curves) and those quenched as satellites of their  $z=0$  host (purple curves). The gray shaded area in the left panel of Fig. 9 represents the fraction of galaxies quenched as satellites found by Wetzel et al. (2013), where the width indicates the uncertainty due to their adopted empirical parametrization for the evolu-



**Figure 9.** Top: fraction of  $z = 0$  quenched satellites versus galaxy stellar mass (left panel) and versus host mass (right panel), both measured at  $z = 0$ . In both panels, colors denote the three different sub-samples that constitute the  $z = 0$  quenched population (see Fig. 8): satellites quenched as centrals (orange curves), as pre-processed (green curves) and as satellites in their  $z = 0$  hosts (purple curves). The black curve represents satellites quenched as satellites, i.e. the sum of the green and purple curves. Additionally, we also show the fraction of galaxies quenched as satellites found by [Wetzel et al. \(2013\)](#) (grey shaded area), in broad agreement with TNG (see text for more details). The fraction is evaluated as the number of quenched satellites in each population over the total number of quenched satellites in a specific mass bin. Bottom: quenched fraction of satellites that have fallen as centrals (direct infallers, solid curves) in the current host and were star-forming at accretion. For comparison, we also plot the left panel of Fig. 3 as dashed curves.

tion of SFRs of SDSS group/cluster satellites catalogues. Similarly to our analysis, the shaded area shows what fraction of currently quiescent satellites were quenched as satellites. This is in broad agreement with the black curve from TNG.

This analysis, overall, confirms two important results. 1) The build up of the  $z = 0$  quenched satellite population is strongly dominated at the low-mass end by the environmental effects imparted by group and cluster environments. 2) On the other hand, at the high stellar mass end ( $\gtrsim 10^{10.5-11} M_{\odot}$ ), quenching as a central (orange data) is almost as frequent as quenching as a satellite (pur-

ple + green data). It is important to note, however, that for the latter we have not yet conclusively demonstrated whether quenching occurs because of environmental processes alone, AGN feedback alone or, as is more plausible, a non-obvious interplay between the two – we discuss this further in Section 4.5.

Finally, to specifically address the question of how important pre-processing is in quenching satellites, we emphasize that a) the majority of low-mass satellites that quench after (any) infall, seem to owe the halt of their star formation to their current host: purple versus green curves; and b) satellites quenched in any hosts



other than their current ones (quenched as pre-processed, green curve, left panel) show an almost constant fraction of 20-30 per cent, regardless of galaxy stellar mass, for all hosts above  $10^{13}$ . This would seem to indicate that overall, pre-processing is unimportant, directly quenching at most only 1/3 of the low-mass  $z = 0$  quenched population of group and cluster satellites. There are two caveats to this general statement.

Firstly, even among those satellites quenched in their final host (purple curves), up to  $\sim 50$  have been pre-processed by subgroups (dashed purple curve, left panel), suggesting that, even if the hosts in which the satellites reside today are fundamental in modifying their star-formation activity, they may also have been helped by group pre-processing.

Furthermore, the fraction of satellites quenched in their current host decreases with increasing  $z = 0$  host mass (purple curve, right panel), thus ushering an intriguing possibility: even if the highest quenched fractions are found in massive clusters, this may not mean that they quench their satellites more efficiently compared to groups, once the role of pre-processing and quenching as centrals is taken into account.

Secondly, not only the frequency (Section 4.2), but also the relative importance of pre-processing for quenching is a strong function of current host mass: as shown in the right panel of Fig. 9 where the statistics is dominated by low-mass galaxies, on average, about 50 per cent of the quenched satellites residing in clusters more massive than  $10^{14.5} M_{\odot}$  are actually quenched in these hosts (purple curve), but the remaining half (i.e. not just 20-30 per cent as for lower-mass hosts: green curve) become passive in other subgroups, entering the  $z = 0$  host already quenched. Additionally, among those satellites that were quenched in their  $z = 0$  host, roughly half were previously orbiting in other subgroups (dashed purple curve, right panel). These facts suggest that, if the quenched fractions of  $10^{9.5-10} M_{\odot}$  satellites of massive  $10^{14.5-15} M_{\odot}$  hosts at  $z \sim 0$  are higher in TNG than in the Universe (as we find by 10-20 percentage points when comparing to those inferred from SDSS: see [Donnari et al. 2020b](#)), this may be due to a combination of phenomena, i.e. not only those pertaining to the environmental processes in very massive hosts at recent cosmic epochs. Future and detailed comparisons of satellite quenched fractions in group-mass hosts at intermediate and high redshifts will be of the essence to inform possible modeling issues.

#### 4.5 On the role of massive hosts and of AGN feedback in massive satellites

The knowledge from the previous section allows us to go back to the left panel of Fig. 3, in which we have shown the quenched fraction of  $z = 0$  satellites versus galaxy stellar mass, stacked in four bins of  $z = 0$  host mass, to highlight how this fraction is shaped within (and because of) different environments. We are now able to answer the question: how does this picture change when pre-processing is accounted for or, conversely, once we exclude satellites quenched before the last infall into the  $z = 0$  host?

In order to better isolate the efficiency of hosts as a function of host mass at quenching their satellites, we show in the bottom panel of Fig. 9 the quenched fraction evaluated as the number of satellites falling as centrals (and then not-preprocessed) that were quenched in their  $z = 0$  hosts over the total number of satellites that fell as centrals and that were still star-forming at the time of accretion (solid curves), thereby excluding satellites quenched because of pre-processing or quenched prior to infall because of internal processes. For reference, we also plot the curves from the left panel

of Fig. 3, in dashed. Two key results can be seen in this figure. For satellites with masses below  $10^{10.5} M_{\odot}$ , more massive clusters are indeed more efficient at quenching than lower mass hosts, particularly with respect to lower-mass groups of  $10^{13-13.5} M_{\odot}$ : black and brown curves versus orange curve. However, we have to bear in mind that this result is convolved with the distributions of satellite infall times, which depend on host mass (see e.g. [Engler et al. 2020](#), their Fig.7), and we have seen in Fig. 4 that longer infall times imply higher probabilities of being quenched. We have checked (even if not shown) that the trend with host mass of Fig. 9, bottom, would appear stronger (i.e. more massive clusters would appear progressively more efficient at quenching than lower-mass hosts) once we restrict to sub-populations of satellites with similar infall times, particularly with comparable recent infall times of 1 to 4 billion years ago.

On the other hand, at the high-mass end of the satellite population that was still star-forming at accretion, an inversion of the trend with host mass appears, whereby the quenched fractions are *lower* in clusters than in groups, by up to 40 percentage points. This is extremely intriguing: this inversion is inconsistent with quenching being driven by environment, as environmental processes are expected to manifest themselves as more being effective at quenching in more massive hosts – see previous results.

The “host-rank” inversion of Fig. 9, together with the absence of a cluster-centric distance dependence of the quenched fraction of massive satellites (Fig. 4, middle right panel), imply that the main culprits for quenching massive satellites are in fact secular and internal processes, these being AGN feedback in TNG. Furthermore, such processes are made *less* effective in higher mass hosts. In [Joshi et al. 2020](#), we have shown that, in TNG, gas accretion into black holes (BH) is significantly suppressed for satellite galaxies after their accretion into  $\gtrsim 10^{14} M_{\odot}$  in comparison to galaxies of similar mass “in the field” (their Fig. 14). Lower BH accretion in turn would reduce the overall energy injected as BH feedback throughout the lifetime of massive satellites, ultimately reducing or delaying its quenching effects<sup>7</sup>.

The phenomenology described here and predicted by the TNG simulations, whereby AGN feedback is hindered for satellites in massive environments, is hard to prove with observations because of the degeneracies of the quenching pathways described so far; but it would be very interesting to see whether it is common to other galaxy formation models. In fact, it is important to bear in mind the following: asserting that quenching in massive satellites is driven by internal processes does not exclude the possibility of witnessing signatures of environmental processes, such as ram-pressure stripped gaseous tails, which in fact are clearly visible in massive TNG satellites across cosmic epochs ([Yun et al. 2019](#)). Moreover, a long-lasting and population-averaged environmental obstruction to AGN activity is not necessarily in contrast to previous findings which have shown that higher density cluster environments can actually trigger enhanced BH activity in satellites for a period of time,

<sup>7</sup> Because of the numerical implementation of BH positioning in the TNG model ([Weinberger et al. 2017](#)), a fraction of satellite galaxies may not have a super massive BH. Even if we do not show it here for the sake of brevity, we have checked and we think that the ‘host-rank’ inversion of Fig. 9 is not related to the lack of BHs in some satellites. For satellites more massive than  $\sim 10^{10.5} M_{\odot}$ , the fraction of those with no BHs is small (less than 10 per cent) and has a little-to-no dependence on host mass in between  $10^{13-14.5} M_{\odot}$ . At higher host masses, this fraction grows to about 20 per cent, but it remains subdominant and thus we think it cannot be responsible for the host dependencies we find in the bottom panel of Fig. 9.

which in some instances can in turn act to eventually quench the satellites (e.g. McGee 2013; Poggianti et al. 2017; Ricarte et al. 2020).

#### 4.6 Distribution of the quenching times and the host masses of quenching hosts

We conclude our analysis by exploring the typical mass scale of hosts where satellites quench and how long ago they actually quenched.

In Fig. 10, main panel, we plot the 2D histogram of the mass of the “quenching” host (i.e. the mass of the halo host at the time of quenching) against the time since quenching, dividing our sample in 40x40 bins on the plane. The color is the number density of  $z = 0$  satellites and the colored contours encompass 50 and 70 per cent of satellites quenched as centrals (orange), those quenched as pre-processed satellites (green) and those quenched in their current host (purple), for which we include also the contour encompassing 95 per cent. In the main panel, all  $z = 0$  satellites are considered, with stellar mass  $\geq 10^9 M_\odot$  in hosts with total mass exceeding  $10^{13} M_\odot$ .

The top and right histograms represent the distributions of the quenching time and the host mass at quenching, respectively, with the same color code as of the central panel. Additionally, we show results for high-mass ( $> 10^{10} M_\odot$ , solid curves) and low-mass satellites ( $< 10^{10} M_\odot$ , dashed curves) separately.

Satellites quenched prior to their last infall, but after being accreted by a dense environment (pre-processed satellites, green curves), typically became passive between 6 to 8 Gyr ago, the majority of them in groups of  $M_{200c} = 10^{13-14} M_\odot$  (at the time of quenching), with no strong dependence on  $z = 0$  stellar mass.

Satellites quenched within their  $z = 0$  host (purple curves), instead, exhibit a broader distribution of quenching times – between 1 and 7 Gyr ago – regardless of their  $z = 0$  stellar mass. Because of this, the typical mass scale of hosts in which they were actually quenched can be lower, since the progenitors of their  $z = 0$  hosts would have been less massive than the ones in which they currently reside. In fact, we find a narrow distribution around  $M_{200c} = 10^{13.2-14} M_\odot$  (between 16th and 84th percentiles). So overall, across the whole  $z = 0$  satellite population that quenched as satellites, quenching most frequently occurs in hosts of mass  $10^{13-14} M_\odot$  over the last many billion years of cosmic evolution.

Finally, as already shown in Fig. 9, only satellites more massive than  $10^{10-10.5} M_\odot$  quench as centrals (orange curves), before falling into any host. In their case, the label “quenching host” may be misleading, as in fact the depicted mass scale refers to their own underlying dark matter haloes. Satellites that quenched as centrals typically quenched in a narrow halo mass range, at  $10^{12.1-12.6} M_\odot$  (between 16th and 84th percentiles). For massive satellite galaxies that are now found in groups and clusters and quenched as centrals, quenching typically occurred between 6 and 10 Gyr ago (between 16th and 84th percentiles).

## 5 SUMMARY AND CONCLUSIONS

In this paper, we have investigated the star formation activity of  $M_{\text{stars}} = 10^{9-12} M_\odot$  galaxies using two cosmological hydrodynamical simulations from the IllustrisTNG project, TNG100 and TNG300. In particular, we have focused on group and cluster satellites, i.e. galaxies found today within the virial radius of haloes with total mass  $M_{200c} = 10^{13-15.2} M_\odot$ . We have explored how

the quenched fractions differ for central versus satellite galaxies and how they change as a function of galaxy stellar mass, redshift, host mass, cluster-centric distance, and time since infall.

We find that the IllustrisTNG model returns the qualitative features of the two main quenching scenarios widely recognized in both observations and previous theoretical models, namely environmentally-driven and mass quenching. While we refer the reader to the companion paper, Donnari et al. (2020), for a detailed quantitative comparison to observational results, here we have leveraged the outcome of the IllustrisTNG simulations to draw a complete picture of the diverse pathways that galaxies can take towards quenching within the hierarchical growth of structure scenario, and to put forward a plausible framework for the interpretation of observational findings. In particular, we have singled out the role that pre-processing and internal secular mechanisms (chiefly, AGN feedback) play in shaping the  $z = 0$  quenched fraction of satellite populations.

Our key findings are the following:

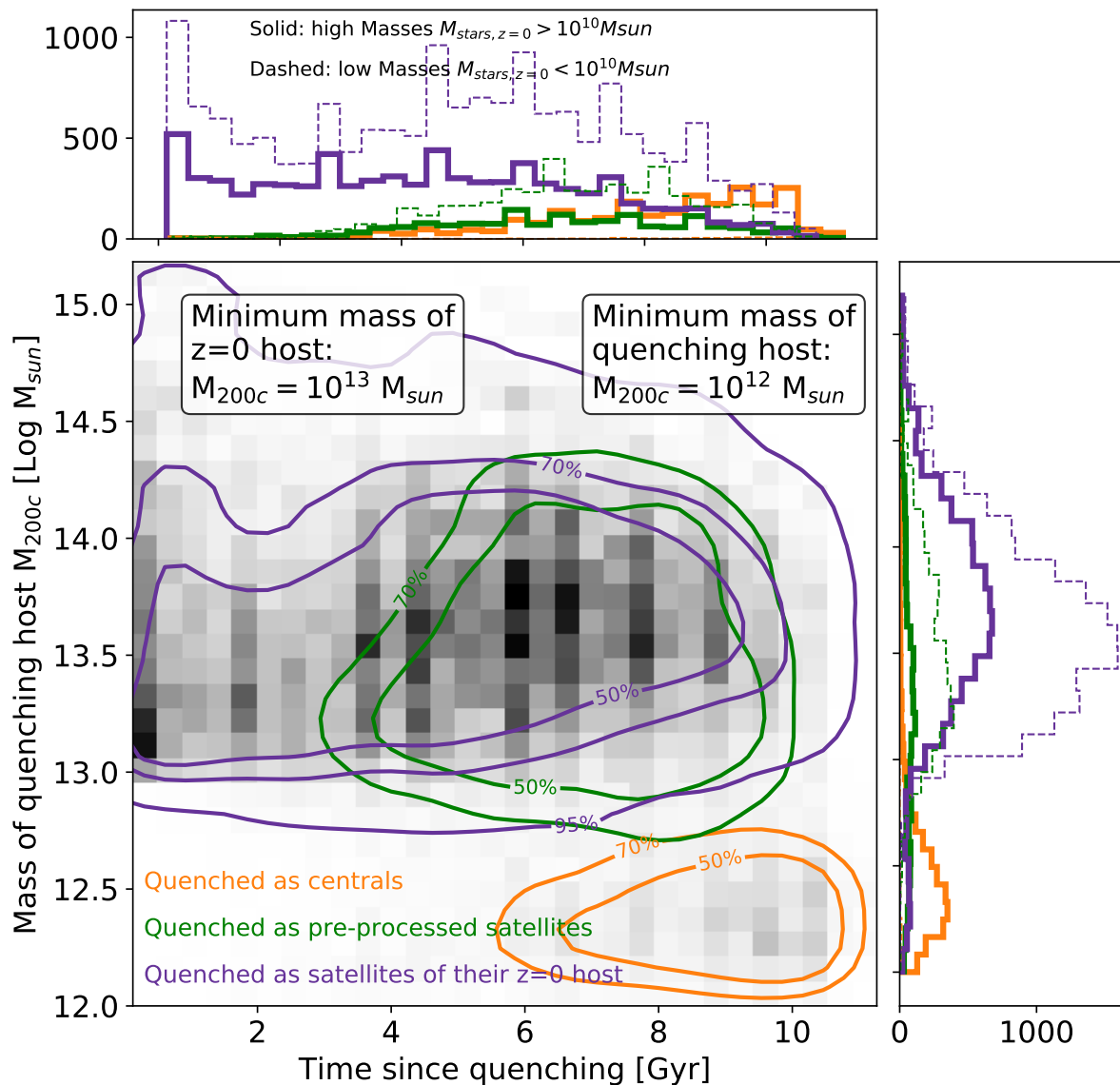
(i) Low-mass galaxies ( $M_{\text{stars}} \lesssim 10^{10} M_\odot$ ) are rarely quenched (5-8 per cent for centrals) unless they are satellites of groups or clusters, consistently with the environmentally-driven quenching scenario. In the case of satellites, at fixed galaxy stellar mass, the fraction of low-mass passive satellites is higher in more massive hosts, closer to the host center, and for those accreted earlier into their current host, with an upper limit of about 4-6 billion years for the time needed for processes in groups and clusters more massive than  $10^{13.5} M_\odot$  to have an effect and hence for quenching to occur.

(ii) Conversely, high-mass galaxies ( $\gtrsim$  a few  $10^{10} M_\odot$ ) quench on their own: internal processes, i.e. AGN feedback in the case of IllustrisTNG, and not environment, is the main culprit for the cessation of star formation in massive galaxies, *regardless of whether they are centrals or satellites*. The massive satellite quenched fractions are high (80-90 per cent) regardless of host mass, cosmic time ( $z \lesssim 0.5$ ), cluster-centric distance, and time since infall.

(iii) A non-negligible subset (about 30 per cent) of the quenched satellite populations of groups and clusters at low-redshift were already quenched before infalling into their last host: for low-mass satellites, this is entirely due to environmental effects during pre-processing, i.e. while orbiting in smaller hosts prior to infall into their final one; for high-mass satellites, quenching occurred when they were centrals, because of their AGN feedback.

(iv) Both environmental and internal star-formation quenching can occur at early cosmic epochs, when the Universe was only a couple billion years old. While, as expected, the quenched fractions of IllustrisTNG central and satellite galaxies are generally lower at higher redshifts (Fig. 2), frequent manifestations of environmental processes in hosts more massive than about  $10^{13} M_\odot$  are already in place at  $z \lesssim 1$  and the bulk of the  $z = 0$  group and cluster quenched satellites ceased their star formation many billion years ago: 6-10, 1-7, and 4-8 Gyrs ago (16th-84th percentiles) for satellites that quenched as centrals, in their current host, or as satellites before falling into it, respectively.

Overall, the IllustrisTNG simulations reveal a great diversity of quenching pathways for satellite galaxies that are found today within groups and clusters of galaxies (see schematic in Fig. 8). We have demonstrated that satellites can quench before any infall in



**Figure 10.** Distribution of the mass of quenching hosts and the time since quenching of  $z = 0$  group and cluster satellites. Mass of quenching host versus time of quenching (lookback time) for the three subsamples discussed in the text: satellites already quenched as centrals (orange), those quenched as pre-processed satellites (green) and those quenched in their current host (purple). Contours represent 50, 70 and 95 per cent of each subsample, while shades of grey denote the number of galaxies in each pixel. The majority of satellites that quenched in their  $z = 0$  host became passive in groups of  $\simeq 10^{13.5-13.8} M_{\odot}$  between 0 and 10 Gyr ago. Only satellites more massive than  $10^{10} M_{\odot}$  quenched as centrals  $\sim 9$  Gyr ago, likely due to secular internal processes.

dense environments (i.e. in our jargon, as centrals) or after being accreted into any hosts (i.e. quenched as satellites) and that among the latter, satellites can quench while being members of pre-processing hosts other than the one where they are found today (see schematic in Fig. 5). The frequency of each quenching pathway and their contribution to the  $z = 0$  satellite quenched fraction depend strongly on the mass of both satellites and hosts under consideration at the time of inquiry (or observation). Therefore, comparisons among models and observational inferences cannot be performed without taking this into account and different satellite and host mass ranges or different satellite and host mass distributions may be partially the cause for the apparently contradicting results in the literature.

Importantly, not only the frequency, but also the relative impact of pre-processing on quenching is a function of current host mass. About 40-60 per cent of satellites that are found today in

massive clusters ( $M_{200c} \gtrsim 10^{14-14.5} M_{\odot}$ ) have been members of other subgroups of at least  $M_{200c} = 10^{12} M_{\odot}$  before falling into the one in which we find them today (i.e. have been pre-processed). These compare to 30-50 per cent of pre-processed satellites that are currently members of lower-mass groups ( $M_{200c} \sim 10^{13-14} M_{\odot}$  – Fig. 6 and left panel of Fig. 7). Furthermore, members of today’s massive clusters have been pre-processed for as long as 3-4 billion years on average and up to 5-6 billion years in the tails of the distributions (Fig. 6, bottom right panel). It is therefore not surprising that pre-processing is responsible for the quenching of about half of the low-mass satellites that inhabit the most massive clusters today ( $M_{200c} \gtrsim 10^{14.5-15.2} M_{\odot}$ ) and not just one third as when averaged across the group and cluster host mass function (Fig. 8, top right panel).

The typical and most frequent host mass where environmen-

tal quenching takes place is  $M_{200c} = 10^{13-14} M_{\odot}$ , by averaging across the  $z = 0$  satellite and host populations and accounting for the last many billion years of cosmic evolution: this finding once more confirms the role of groups (rather than clusters) as fundamental building blocks of the hierarchical cosmic structure formation. Still, after excluding satellites quenched by their AGN feedback and those that have been quenched by pre-processing, massive clusters remain more efficient at quenching their members compared to lower-mass hosts: this is the case for satellites with stellar masses below about  $10^{10.5} M_{\odot}$  (Fig. 9) and is in line with the widely-accepted dependence of environmental quenching processes such as ram-pressure stripping on environmental density and depth of the host gravitational potential well.

On the other hand, for the more massive satellites, the IllustrisTNG model predicts an inversion of the “ranking of host efficiency”, in the sense that the percentage of massive satellites that actually quench in their current host is higher in groups rather than in clusters. According to IllustrisTNG, AGN feedback is hindered for satellites in dense environments or, in other terms, massive hosts reduce on average the effectiveness of quenching internal secular processes. While observationally arduous to be proven, it will be interesting to see whether a long-lasting and population-averaged environmental obstruction to AGN activity is common also to other galaxy formation models.

#### DATA AVAILABILITY

The data from the Illustris and TNG simulations used in this work are publicly available at the websites <https://www.illustris-project.org> and <https://www.tng-project.org>, respectively (Nelson et al. 2015, 2019a).

#### ACKNOWLEDGEMENT

The authors are grateful to the referee for their comments and suggestions that helped to improve an earlier version of this manuscript. MD acknowledges support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 138713538 – SFB 881 (“The Milky Way System”, sub-project A01). FM acknowledges support through the Program “Rita Levi Montalcini” of the Italian MIUR. MV acknowledges support through an MIT RSC award, a Kavli Research Investment Fund, NASA ATP grant NNX17AG29G, and NSF grants AST-1814053, AST-1814259 and AST-1909831. The flagship simulations of the IllustrisTNG project used in this work have been run on the Hazel-Hen Cray XC40-system at the High Performance Computing Center Stuttgart as part of project GCS-ILLU of the Gauss centres for Super-computing(GCS). Ancillary and test runs of the project were also run on the Stampede supercomputer at TACC/XSEDE (allocation AST140063), at the Hydra and Draco supercomputers at the Max Planck Computing and Data Facility, and on the MIT/Harvard computing facilities supported by FAS and MIT MKI.

#### REFERENCES

Bahé Y. M., McCarthy I. G., Balogh M. L., Font A. S., 2013, *MNRAS*, **430**, 3017  
 Bahé Y. M., et al., 2017, *MNRAS*, **470**, 4186  
 Bahé Y. M., et al., 2019, *MNRAS*, **485**, 2287

Baldry I. K., Balogh M. L., Bower R., Glazebrook K., Nichol R. C., 2004, in Allen R. E., Nanopoulos D. V., Pope C. N., eds, American Institute of Physics Conference Series Vol. 743, The New Cosmology: Conference on Strings and Cosmology. pp 106–119 ([arXiv:astro-ph/0410603](https://arxiv.org/abs/astro-ph/0410603)), doi:10.1063/1.1848322  
 Barsanti S., et al., 2018, *ApJ*, **857**, 71  
 Beckmann R. S., et al., 2017, *MNRAS*, **472**, 949  
 Berrier J. C., Stewart K. R., Bullock J. S., Purcell C. W., Barton E. J., Wechsler R. H., 2009, *ApJ*, **690**, 1292  
 Bianconi M., Smith G. P., Haines C. P., McGee S. L., Finoguenov A., Egami E., 2018, *MNRAS*, **473**, L79  
 Boselli A., et al., 2014, *A&A*, **570**, A69  
 Boselli A., et al., 2016, *A&A*, **596**, A11  
 Choi H., Yi S. K., 2017, *ApJ*, **837**, 68  
 Crossett J. P., Pimblett K. A., Jones D. H., Brown M. J. I., Stott J. P., 2017, *MNRAS*, **464**, 480  
 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., Stroe A., Hemmati S., Kartaltepe J., 2017, *ApJ*, **837**, 16  
 Davies L. J. M., et al., 2019, *MNRAS*, **483**, 5444  
 Davies J. J., Crain R. A., Oppenheimer B. D., Schaye J., 2020, *MNRAS*, **491**, 4462  
 Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, *ApJ*, **292**, 371  
 De Lucia G., Weinmann S., Poggianti B. M., Aragón-Salamanca A., Zaritsky D., 2012, *MNRAS*, **423**, 1277  
 De Lucia G., Hirschmann M., Fontanot F., 2019, *MNRAS*, **482**, 5041  
 Di Matteo T., Springel V., Hernquist L., 2005, *Nature*, **433**, 604  
 Diemer B., et al., 2019, *MNRAS*, **487**, 1529  
 Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, **399**, 497  
 Donnari M., et al., 2019, *MNRAS*, **485**, 4817  
 Donnari M., Pillepich A., Nelson D., Marinacci F., Vogelsberger M., Hernquist L., 2020, arXiv e-prints, p. [arXiv:2008.00004](https://arxiv.org/abs/2008.00004)  
 Dubois Y., Peirani S., Pichon C., Devriendt J., Gavazzi R., Welker C., Volonteri M., 2016, *MNRAS*, **463**, 3948  
 Eales S., et al., 2018, *MNRAS*, **473**, 3507  
 Engler C., et al., 2020, arXiv e-prints, p. [arXiv:2002.11119](https://arxiv.org/abs/2002.11119)  
 Fang J. J., et al., 2018, *ApJ*, **858**, 100  
 Feldmann R., 2017, *MNRAS*, **470**, L59  
 Forrest B., et al., 2020, *ApJ*, **890**, L1  
 Fossati M., et al., 2017, *ApJ*, **835**, 153  
 Fujita Y., 2004, *PASJ*, **56**, 29  
 Furlong M., et al., 2015, *MNRAS*, **450**, 4486  
 Geha M., Blanton M. R., Yan R., Tinker J. L., 2012, *ApJ*, **757**, 85  
 Genel S., et al., 2014, *MNRAS*, **445**, 175  
 Genel S., et al., 2018, *MNRAS*, **474**, 3976  
 Glazebrook K., et al., 2017, *Nature*, **544**, 71  
 Gunn J. E., Gott III J. R., 1972, *ApJ*, **176**, 1  
 Haines C. P., et al., 2015, *ApJ*, **806**, 101  
 Han S., Smith R., Choi H., Cortese L., Catinella B., Contini E., Yi S. K., 2018, *ApJ*, **866**, 78  
 Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Wang W., 2017, *MNRAS*, **469**, 2626  
 Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Martini P., Robertson B., Springel V., 2005, *ApJ*, **630**, 705  
 Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, **445**, 581  
 Hou A., Parker L. C., Harris W. E., 2014, *MNRAS*, **442**, 406  
 Jian H.-Y., et al., 2018, *PASJ*, **70**, S23  
 Joshi G. D., Wadsley J., Parker L. C., 2017, *MNRAS*, **468**, 4625  
 Joshi G. D., Parker L. C., Wadsley J., Keller B. W., 2019, *MNRAS*, **483**, 235  
 Joshi G. D., Pillepich A., Nelson D., Marinacci F., Springel V., Rodriguez-Gomez V., Vogelsberger M., Hernquist L., 2020, arXiv e-prints, p. [arXiv:2004.01191](https://arxiv.org/abs/2004.01191)  
 Kauffmann G., et al., 2003, *MNRAS*, **341**, 54  
 Kaviraj S., et al., 2017, *MNRAS*, **467**, 4739  
 Larson R. B., Tinsley B. M., Caldwell C. N., 1980, *ApJ*, **237**, 692

- Li P., Wang H., Mo H. J., Wang E., Hong H., 2020, arXiv e-prints, p. [arXiv:2003.09776](https://arxiv.org/abs/2003.09776)
- Marinacci F., et al., 2018, *MNRAS*, **480**, 5113
- McCarthy I. G., Schaye J., Bower R. G., Ponman T. J., Booth C. M., Dalla Vecchia C., Springel V., 2011, *MNRAS*, **412**, 1965
- McGee S. L., 2013, *MNRAS*, **436**, 2708
- McGee S. L., Balogh M. L., Bower R. G., Font A. S., McCarthy I. G., 2009, *MNRAS*, **400**, 937
- Medling A. M., et al., 2018, *MNRAS*, **475**, 5194
- Merritt D., 1983, *ApJ*, **264**, 24
- Moore B., Katz N., Lake G., Dressler A., Oemler A., 1996, *Nature*, **379**, 613
- Moore B., Lake G., Quinn T., Stadel J., 1999, *MNRAS*, **304**, 465
- Muzzin A., Marchesini D., Stefanon M., Franx M., et al. 2013, *ApJ*, **777**, 18
- Naiman J. P., et al., 2018, *MNRAS*, **477**, 1206
- Nelson D., et al., 2015, *Astronomy and Computing*, **13**, 12
- Nelson D., et al., 2018a, *MNRAS*, **475**, 624
- Nelson D., et al., 2018b, *MNRAS*, **477**, 450
- Nelson D., et al., 2019a, *Computational Astrophysics and Cosmology*, **6**, 2
- Nelson D., et al., 2019b, *MNRAS*, **490**, 3234
- Nichols M., Bland-Hawthorn J., 2011, *ApJ*, **732**, 17
- Paccagnella A., et al., 2016, *ApJ*, **816**, L25
- Pallero D., Gómez F. A., Padilla N. D., Torres-Flores S., Demarco R., Cerulo P., Olave-Rojas D., 2019, *MNRAS*, **488**, 847
- Pasquali A., Smith R., Gallazzi A., De Lucia G., Zibetti S., Hirschmann M., Yi S. K., 2019, *MNRAS*, **484**, 1702
- Peng Y.-j., Lilly S. J., Kovač K., Bolzonella M., et al. 2010, *ApJ*, **721**, 193
- Peng Y., Maiolino R., Cochrane R., 2015, *Nature*, **521**, 192
- Penny S. J., et al., 2018, *MNRAS*, **476**, 979
- Pillepich A., et al., 2018, *MNRAS*, **475**, 648
- Pillepich A., et al., 2019, *MNRAS*, **490**, 3196
- Poggianti B. M., et al., 2017, *ApJ*, **844**, 48
- Presotto V., et al., 2012, *A&A*, **539**, A55
- Rhee J., Smith R., Choi H., Yi S. K., Jaffé Y., Sánchez-Jánssen R., 2017, *ApJ*, **843**, 128
- Ricarte A., Tremmel M., Natarajan P., Quinn T., 2020, arXiv e-prints, p. [arXiv:2003.05950](https://arxiv.org/abs/2003.05950)
- Rodríguez-Gomez V., et al., 2015, *MNRAS*, **449**, 49
- Rodríguez-Gomez V., et al., 2019, *MNRAS*, **483**, 4140
- Schaefer A. L., et al., 2019, *MNRAS*, **483**, 2851
- Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, *MNRAS*, **452**, 575
- Silk J., 2013, *ApJ*, **772**, 112
- Somerville R. S., Davé R., 2015, *ARA&A*, **53**, 51
- Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, *ApJS*, **214**, 15
- Springel V., 2010, *MNRAS*, **401**, 791
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, **328**, 726
- Springel V., et al., 2018, *MNRAS*, **475**, 676
- Stevens A. R. H., et al., 2019, *MNRAS*, **483**, 5334
- Strateva I., et al., 2001, *AJ*, **122**, 1861
- Terrazas B. A., et al., 2020, *MNRAS*, **493**, 1888
- Torrey P., et al., 2018, *MNRAS*, **477**, L16
- Tremmel M., et al., 2019, *MNRAS*, **483**, 3336
- Truong N., et al., 2020, *MNRAS*, **494**, 549
- Vogelsberger M., et al., 2014, *Nature*, **509**, 177
- Vogelsberger M., et al., 2018, *MNRAS*, **474**, 2073
- Vulcani B., et al., 2012, *MNRAS*, **420**, 1481
- Weinberger R., et al., 2017, *MNRAS*, **465**, 3291
- Weinberger R., et al., 2018, *MNRAS*, **479**, 4056
- Wetzel A. R., Tinker J. L., Conroy C., 2012, *MNRAS*, **424**, 232
- Wetzel A. R., Tinker J. L., Conroy C., van den Bosch F. C., 2013, *MNRAS*, **432**, 336
- Whitaker K. E., et al., 2014, *ApJ*, **795**, 104
- Yun K., et al., 2019, *MNRAS*, **483**, 1042
- Zabludoff A. I., Zaritsky D., Lin H., Tucker D., Hashimoto Y., Shectman S. A., Oemler A., Kirshner R. P., 1996, *ApJ*, **466**, 104
- Zinger E., et al., 2020, arXiv e-prints, p. [arXiv:2004.06132](https://arxiv.org/abs/2004.06132)
- van der Wel A., Rix H.-W., Holden B. P., Bell E. F., Robaina A. R., 2009, *ApJ*, **706**, L120

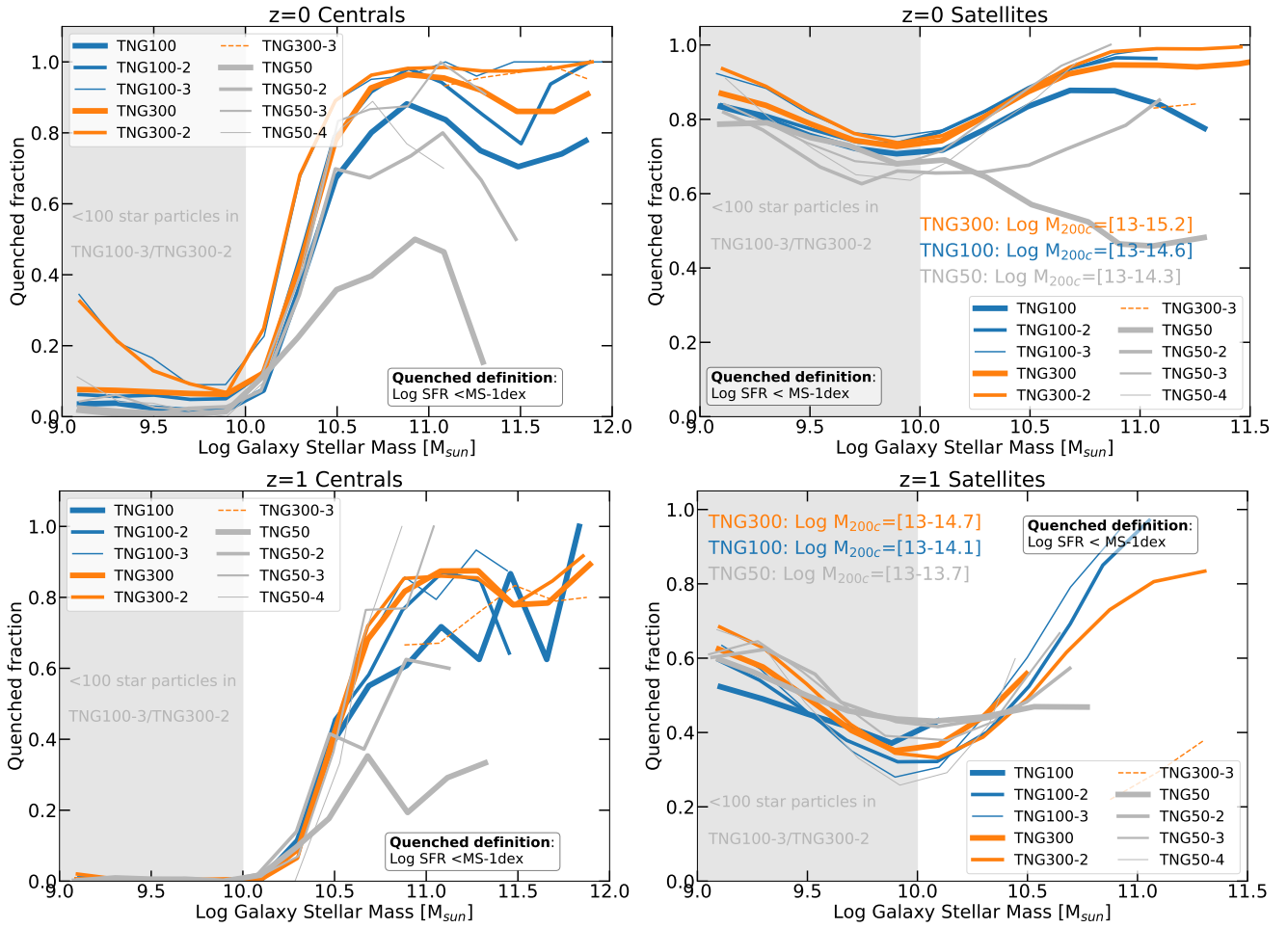
## APPENDIX A: EFFECTS OF NUMERICAL RESOLUTION ON QUENCHED FRACTIONS OF CENTRALS AND SATELLITES

In this Appendix, we quantify the effect of numerical resolution on quenched fractions. Necessarily, the comparisons in this Appendix will also be affected by sample variance, as – as noted throughout the text – TNG100 and TNG300 span different simulation box sizes.

Here, we show results according to TNG100 and TNG300 at their different resolution levels, namely TNG100-2, TNG100-3 and TNG300-2, TNG300-3, respectively. In the TNG100 series, the baryonic mass reads  $1.4 \times 10^6$ ,  $1.1 \times 10^7$ , and  $8.9 \times 10^7 M_\odot$ , respectively, while in the TNG300 series we have  $1.1 \times 10^7$ ,  $8.8 \times 10^7$ , and  $7.0 \times 10^8 M_\odot$  (see also Tab. 1 in Pillepich et al. 2018, for more details), so that the resolutions of e.g. TNG300 and TNG100-2 are identical. For completeness, and albeit not used for this work, we also include in this Appendix the results from the highest-resolution run of the IllustrisTNG series, TNG50, with its lower-resolution realizations: TNG50-2, TNG50-3 and TNG50-4. The different resolution realizations of TNG50 are characterized by a baryonic mass of  $8.5 \times 10^4 M_\odot$ ,  $6.8 \times 10^5 M_\odot$ ,  $5.4 \times 10^6 M_\odot$ , and  $4.3 \times 10^7 M_\odot$ , respectively (see e.g. Pillepich et al. 2019), so that that TNG100 has a mass resolution in between TNG50-2 and TNG50-3, while TNG100-2 (and TNG300) have a mass resolution in between TNG50-3 and TNG50-4. TNG50 encompasses a smaller volume than TNG100 and TNG300 (about 50 comoving Mpc a side, in comparison to about 100 and 300 comoving Mpc of TNG100 and TNG300): it therefore lacks the highest-mass end clusters (above a few  $10^{14} M_\odot$ ) that this work focuses on.

Fig. A1 shows the analog of Fig. 1. We compare quenched fractions of centrals (left panel) and satellites (right panel) at  $z = 0$  (top) and  $z = 1$  (bottom) according to the three resolution levels of TNG300 and TNG100, represented with different curve thickness. Satellites in the right columns include those in hosts with minimum masses of  $10^{13} M_\odot$  and maximum masses labeled in each panel, depending on the run and redshift. To avoid confusion, results from TNG300-3 are presented only above its mass resolution where the samples of centrals and satellites include galaxies with at least one hundred stellar particles, namely  $M_{\text{stars}} \gtrsim 10^{11} M_\odot$  (dashed curve). Moreover, shaded gray area indicate the stellar mass ranges where galaxies are resolved with fewer than 1000 stellar particles in the TNG100-3 and TNG300-2 runs, which have the same numerical resolution.

Typically, the lower resolution runs return larger quenched fractions, both for centrals and satellites, at any stellar masses, with no noticeable redshift trend (at least at  $z \leq 1$ ), but with some inversions of trends for more massive satellites at high redshifts. Focusing on TNG100 and TNG300 and barring the outcome of the lowest resolution run of the series that returns completely disparate results (TNG300-3, with a baryonic mass resolution of  $7.0 \times 10^8 M_\odot$ ), the resolution-induced discrepancy across factors of 8 in particle mass resolution is however rather small, being typically smaller than about 10 percentage points at  $z \leq 1$ , but being somewhat confounded by sampling and statistics, and therefore reaching larger deviations up to 15-20 percentage points in regimes with poor sampling, like at the highest-mass end at higher redshifts. This indicates that at fixed galaxy stellar mass the TNG300/TNG100 results carry a systematic average error of about 10 percentage points across the regimes studied in this paper: this is comparable or even smaller than the differences between inherent simulation results versus observation-mocks needed for comparison to observational results (see Donnari et al. 2020). However, within the IllustrisTNG framework and as further demonstrated and discussed by Donnari et al. 2020, the higher-resolution run TNG50 (not used in this paper) returns  $z = 0$  quenched fractions that are up to another 10-30 percentage points lower than those from TNG100: as we discuss there, and given that TNG50 simulates fewer



**Figure A1. Resolution study of the quenched fractions.** Fraction of quenched galaxies as a function of stellar mass for centrals (left) and satellites (right) at  $z = 0$  (top) and  $z = 1$  (bottom) according to the three different resolution levels of TNG300 (orange) and TNG100 (blue), and the four resolution levels of TNG50 (grey). The grey shaded area denotes the resolution limit of TNG100-3 and TNG300-2, where the sample of centrals includes galaxies with less than one hundred of stellar particles.

high-mass galaxies and hosts due its smaller comoving volume, this is due to a combination of resolution effects and sample variance. Nonetheless, Fig. A1 confirms what was already pointed out in Section 3.1: the differences between TNG100 and TNG300, albeit small, are mainly due to a numerical resolution effect. However, higher numerical resolutions and poorer sampling statistics can impact the predictions of the quenched fractions with systematic uncertainties by up to 10-40 percentage points, depending on galaxy mass, host mass, and redshift (see also Donnari et al. 2020).