




This article has been accepted for publication in Monthly Notices of the Royal Astronomical Society ©: 2019 The Authors. Published by Oxford University Press on behalf of the Royal Astronomical Society. All rights reserved.

The star formation activity of IllustrisTNG galaxies: main sequence, *UVJ* diagram, quenched fractions, and systematics

Martina Donnari ¹★, Annalisa Pillepich,¹ Dylan Nelson,² Mark Vogelsberger,³ Shy Genel ^{4,5}, Rainer Weinberger,⁶ Federico Marinacci ³, Volker Springel² and Lars Hernquist⁶

¹Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany

²Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Str 1, D-85741 Garching, Germany

³Kavli Institute for Astrophysics and Space Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

⁴Center for Computational Astrophysics, Flatiron Institute, 162 Fifth Avenue, New York, NY 10010, USA

⁵Columbia Astrophysics Laboratory, Columbia University, 550 West 120th Street, New York, NY 10027, USA

⁶Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Accepted 2019 March 6. Received 2019 March 6; in original form 2018 December 17

ABSTRACT

We select galaxies from the IllustrisTNG hydrodynamical simulations ($M_{\text{stars}} > 10^9 M_{\odot}$ at $0 \leq z \leq 2$) and characterize the shapes and evolutions of their *UVJ* and star formation rate–stellar mass (SFR– M_{stars}) diagrams. We quantify the systematic uncertainties related to different criteria to classify star-forming versus quiescent galaxies, different SFR estimates, and by accounting for the star formation measured within different physical apertures. The TNG model returns the observed features of the *UVJ* diagram at $z \leq 2$, with a clear separation between two classes of galaxies. It also returns a tight star-forming main sequence (MS) for $M_{\text{stars}} < 10^{10.5} (M_{\odot})$ with a ~ 0.3 dex scatter at $z \sim 0$ in our fiducial choices. If a *UVJ*-based cut is adopted, the TNG MS exhibits a downwardly bending at stellar masses of about $10^{10.5-10.7} M_{\odot}$. Moreover, the model predicts that ~ 80 (50) per cent of $10^{10.5-11} M_{\odot}$ galaxies at $z = 0$ ($z = 2$) are quiescent and the numbers of quenched galaxies at intermediate redshifts and high masses are in better agreement with observational estimates than previous models. However, shorter SFR-averaging time-scales imply higher normalizations and scatter of the MS, while smaller apertures lead to underestimating the galaxy SFRs: overall we estimate the inspected systematic uncertainties to sum up to about 0.2–0.3 dex in the locus of the MS and to about 15 percentage points in the fraction of quenched galaxies. While TNG colour distributions are clearly bimodal, this is not the case for the SFR logarithmic distributions in bins of stellar mass (SFR $\gtrsim 10^{-3} M_{\odot} \text{yr}^{-1}$). Finally, the slope and $z = 0$ normalization of the TNG MS are consistent with observational findings; however, the locus of the TNG MS remains lower by about 0.2–0.5 dex at $0.75 \leq z < 2$ than the available observational estimates taken at face value.

Key words: methods: numerical – galaxies: evolution – galaxies: formation – galaxies: star formation – cosmology: theory.

1 INTRODUCTION

The observed distribution of many galaxy properties, such as colour, morphology, and star formation rate (SFR), are thought to be bimodal in the nearby Universe and for relatively massive galaxies. These properties correlate with galaxy density and stellar mass, suggesting that the processes that regulate the star formation

activity of galaxies depend both on a galaxy’s stellar mass and the environment in which galaxies evolve (Peng et al. 2010; Vulcani et al. 2010, 2012; Paccagnella et al. 2016; Darvish et al. 2017).

Broadly speaking, galaxies in the local Universe can be mainly divided into two categories: disc-like ‘star-forming’ galaxies that preferentially populate the low-density regions of the Universe, and ‘passive’ or ‘quenched’ galaxies, with older stellar populations and little-to-no star formation, which are mostly observed at high stellar masses and/or populating the high-density environments. The latter type is comprised mainly of spheroidal

* E-mail: donnari.m@gmail.com

or elliptical galaxies (Kauffmann et al. 2003; Baldry et al. 2004).

As suggested by observations, star-forming galaxies show a tight correlation between their stellar mass and star-formation rate (SFR), this relation being commonly called ‘star-forming main sequence’ (hereafter MS or SFMS; e.g. Noeske et al. 2007; Rodighiero et al. 2011; Wuyts et al. 2011; Whitaker et al. 2012; Speagle et al. 2014; Renzini & Peng 2015; Schreiber et al. 2015; Shivaei et al. 2015; Tasca et al. 2015). The existence of an MS provides an easy and powerful way to discern between star-forming and quenched galaxies, representing, indeed, a suitable upper-limit below which a galaxy of a given stellar mass can be considered quenched.

Although the MS is widely recovered from both observations and theory, two different schools of thought exist over e.g. the shape and locus of the MS. Some observational works point out that at $M_{\text{stars}} \gtrsim 10^{10.5} M_{\odot}$ there may be a turnover, with the slope of the MS being shallower above this stellar mass value than below (Reddy et al. 2008; Karim et al. 2011; Whitaker et al. 2012, 2014; Gavazzi et al. 2015; Lee et al. 2015, 2018; Schreiber et al. 2016; Tomczak et al. 2016). On the other side, some studies find the MS can be simply described by a single power law, with the logarithm of the SFR increasing with the logarithm of the stellar mass across the whole studied mass range (see Speagle et al. 2014, and references therein for a collection of observational works).

Additionally, observations show that the value of the normalization increases at increasing redshift, with a redshift-independent slope measured in the range (0.4, 1) and the normalization measured to be in the range of $(-1.0, 0.25) M_{\odot} \text{yr}^{-1}$ at $\sim 10^9 M_{\odot}$ and $z = 0-2$, likely depending on sample selection and on the indicator adopted to estimate the SFR of observed galaxies (see e.g. Rodighiero et al. 2014; Speagle et al. 2014; Santini et al. 2017; Jian et al. 2018; Pearson et al. 2018, and references therein). In fact, non-negligible systematics potentially affect the observational estimates of the SFRs of galaxies across masses and times, owing to a variety of different assumptions on e.g. the effects of dust attenuation, the choices of the stellar initial mass function, on the conversions from observed luminosity to SFR and on the levels of contamination (see e.g. Speagle et al. 2014; Leja et al. 2018; Theios et al. 2018).

Different observational indicators adopted to infer galaxy SFRs, such as $H\alpha$ emission lines, X-ray, UV, and IR bands, and the associated definition of star-forming or passive galaxies, can lead to different quantitative results (Davies et al. 2017, 2018; Katsianis et al. 2017). Consequently, any comparison between observational and theoretical works should take into account the effects of different observational tracers, since the time-scales over which different SFR indicators are thought to be sensitive to can vary from a few – 5 or 10 Myr, usually associated with $H\alpha$ and UV emission lines – to hundreds of megayears – for 24 μm , IR, and FIR bands.

Furthermore, the absence of a standard approach in dealing with SFR in galaxies is outlined by several authors who have highlighted the importance of quantifying the effect of apertures in measuring galaxies’ SFRs (Richards et al. 2016), both in the observational data (Kewley, Jansen & Geller 2005) and also in comparison with the predictions of theoretical model (Guidi et al. 2016). For example, the Sloan Digital Sky Survey (SDSS) fibre aperture covers only a limited region of a galaxy at low redshift (~ 5 physical kpc), and thus it is sensitive to only a fractional amount of its $H\alpha$ emission.

From a theoretical point of view, the whole SFR– M_{star} plane and the MS of star-forming galaxies have been investigated by

several authors in analytical galaxy formation models (e.g. Bouché et al. 2010; Davé, Finlator & Oppenheimer 2012; Lilly et al. 2013; Mitra, Davé & Finlator 2015; Mitra et al. 2017), in semi-analytical models (SAMs; Kang & van den Bosch 2008; De Lucia et al. 2012; Hirschmann et al. 2014; Henriques et al. 2017) and hydrodynamical simulations (e.g. Torrey et al. 2014; Bahé & McCarthy 2015; Sparre et al. 2015; Davé, Rafieferantsoa & Thompson 2017; Feldmann 2017). Although the MS is recovered in both the aforementioned techniques, its evolution with redshift is still a controversial issue, making the comparison among theoretical models and observations still unsatisfactory. A general consensus is that the fraction of passive galaxies increases with increasing stellar mass and decreases with increasing redshift. However, an optimal match between theory and observations on how SFR relates to the environment remains elusive, although the use of the state-of-art simulations provides nowadays a powerful tool to investigate this question.

Indeed, the SFR– M_{star} plane is not immune to different interpretations: this is due to the lack of agreement on a universal definition of ‘quenched’ galaxies from both theoretical and observational sides and to the limitations on the minimum SFR values that can be effectively detected in observations and sampled in simulations. Because of the broadly recognized galaxy colour bimodality that is present at least up to $z \sim 3$, a considerable number of observational studies use a colour–colour-based classification to distinguish between blue/star-forming and red/quenched galaxies (e.g. Williams et al. 2009, 2010; Whitaker et al. 2010, 2011; Wuyts et al. 2011; Patel et al. 2012; Quadri et al. 2012; Ilbert et al. 2013; Muzzin et al. 2013; Tomczak et al. 2014; Darvish et al. 2016). Besides a colour bimodality, the data from large-scale surveys like SDSS have been shown to exhibit a double peak in the specific star-formation rate (sSFR = SFR/ M_{stars}) distribution of galaxies (Ilbert et al. 2015), with a depression or break that likely depend on the observed galaxy sample but appears independent of galaxy and host halo mass, at least in the $\gtrsim 10^{10} M_{\odot}$ range (Brinchmann et al. 2004; Kauffmann et al. 2004; Wetzel, Tinker & Conroy 2012). However, recently Feldmann (2017) and Eales et al. (2017, 2018) have argued that the observed SFR bimodality and the existence of a ‘green valley’ region in SFR– M_{stars} plane, that separates star forming from quenched populations, are likely connected to uncertainties and observational biases in the SFR measurements, hence challenging a foundational galaxy-evolution paradigm.

In this paper we aim at providing a thorough characterization of the star formation activity of galaxies produced in the state-of-the-art cosmological magneto-hydrodynamical simulations of *The Next Generation Illustris* project (hereafter IllustrisTNG or TNG; Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018a; Pillepich et al. 2018b; Springel et al. 2018). Particularly, we study galaxies from the TNG100 and TNG300 simulations at redshifts $0 \leq z \leq 2$, by extending to higher redshifts and new diagnostics previous TNG analyses of the TNG $z = 0$ galaxy colours (Nelson et al. 2018a), $z = 0$ SFRs (Weinberger et al. 2018), and $z = 1-2$ SFRs and black hole (BH) populations (Habouzit et al. 2018). For instance, in Nelson et al. (2018a) we have shown that the IllustrisTNG model returns a remarkably good agreement in terms of the shape and width of the red sequence and the blue cloud of galaxies in comparison to SDSS data, by improving upon the galaxy optical-colour bimodality at the high-mass end ($\sim 10^{10}-10^{11} M_{\odot}$) of the original Illustris model (Genel et al. 2014; Vogelsberger et al. 2014a) and thus providing an indirect confirmation of the adopted underlying recipes for stellar and BH feedback.

The IllustrisTNG simulations therefore provide an interesting starting point to further investigate some of the open-ended questions that mine our observational and theoretical understanding of the evolution of galaxies: how does the MS evolve in time? Is there a bending of the star-forming MS at the high-mass end? How do galaxy selection criteria affect the shape of the MS and the estimates of the fraction of quenched galaxies? What is the quantitative connection between colour–colour and SFR-based criteria to select for quenched galaxies? And can a colour bimodality be in place even without a bimodality in SFR or sSFR?

The structure of our paper is as follows. In Section 2, we present the simulations used in this work and we introduce all of the criteria, definitions, and measurement choices adopted throughout the analysis. Our IllustrisTNG results are given in Section 3: we characterize the *UVJ* diagram of TNG galaxies, quantify the relative number of quenched galaxies as a function of mass and time. We also discuss the morphology of the SFR– M_{stars} plane and quantify the effects that different measurement choices have on the locus and shape of the star-forming MS and its scatter. A discussion of the main findings, together with an extensive study of the distributions of galaxy colours and SFRs and of the shape of the star-forming MS at the high-mass end, are presented in Section 4. There we attempt a comparison to a selection of observational data taken at face value. Our summary and conclusions are given in Section 5.

2 METHODS

2.1 The IllustrisTNG simulations

The IllustrisTNG simulations¹ (Marinacci et al. 2018; Naiman et al. 2018; Nelson et al. 2018a; Pillepich et al. 2018b; Springel et al. 2018, TNG hereafter) are a series of cosmological magneto-hydrodynamical simulations that model a range of physical processes considered relevant for the formation of galaxies. They build and improve upon the original Illustris project (Vogelsberger et al. 2013, 2014b; Genel et al. 2014; Torrey et al. 2014; Nelson et al. 2015; Sijacki et al. 2015) by including, among others, a new black hole feedback model, magneto-hydrodynamics, and a new scheme for galactic winds (see Pillepich et al. 2018a; Weinberger et al. 2018, for details on the TNG model).

The simulations are performed with the AREPO code (Springel 2010), which solves Poisson’s equation for gravity by employing a tree-particle-mesh algorithm. For magneto-hydrodynamics, the code uses the finite volume method on an unstructured, moving, Voronoi tessellation of the simulation domain.

The initial conditions of the TNG runs have been initialized at redshift $z = 127$ assuming a matter density $\Omega_m = \Omega_{\text{dm}} + \Omega_b = 0.3089$, baryonic density $\Omega_b = 0.0486$, cosmological constant $\Omega_\Lambda = 0.6911$, Hubble constant $H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}$ with $h = 0.6774$, normalization $\sigma_8 = 0.8159$ and spectral index $n_s = 0.9667$ (Planck Collaboration XIII 2016).

A number of outcomes from the IllustrisTNG simulations have been shown to be consistent with observations: e.g. the shape and width of the red sequence and the blue cloud of $z = 0$ galaxies (Nelson et al. 2018a), the distributions of metals in the intracluster medium at low redshifts (Vogelsberger et al. 2018), the evolution of the galaxy mass–metallicity relation (Torrey et al. 2018) and of the galaxy size–mass relation for star-forming and quiescent galaxies

Table 1. Properties of the TNG100 and TNG300 simulations. Columns read: (1) simulation name; (2) box-side length; (3) initial number of gas cells; (4) number of dark matter particles; (5) target baryonic mass; (6) dark-matter particle mass.

Run	L_{box} (Mpc h^{-1})	N_{GAS}	N_{DM}	m_{baryon} ($10^6 M_\odot$)	m_{DM} ($10^6 M_\odot$)
TNG100	75	1820 ³	1820 ³	1.4	7.5
TNG300	205	2500 ³	2500 ³	11	59

at $0 \leq z \leq 2$ (Genel et al. 2018), the amount and distribution of highly ionized oxygen around galaxies (Nelson et al. 2018b), and the dark matter fractions within the bodies of massive galaxies at $z = 0$ (Lovell et al. 2018). All these validations of the model make the IllustrisTNG simulations an ideal laboratory to further investigate galaxy formation and evolution. Of relevance for the analysis at hand, the new physics ingredients included in the TNG model affect the star formation activity of galaxies, thus potentially producing important differences in the locus of the star-forming main sequence and the number of quenched galaxies in comparison to the original Illustris simulation (Sparre et al. 2015).

In this work, we use the simulations called TNG100 and TNG300, that model two cosmological boxes with a side length of $75 h^{-1} \approx 100 \text{ Mpc}$ and $205 h^{-1} \approx 300 \text{ Mpc}$, respectively. A third simulation, with a side length box of $35 h^{-1} \approx 50 \text{ Mpc}$ and increased numerical resolution (TNG50), is still running and will not be discussed in this paper. The main properties of TNG100 and TNG300 are listed in Table 1 (see Pillepich et al. 2018b; Nelson et al. 2018a, for more details).

2.2 The galaxy sample

All of the properties of haloes and subhaloes used in this work are obtained taking advantage of the Friends-of-Friends and SUBFIND algorithms (Davis et al. 1985; Springel et al. 2001), which are used to identify substructures, and hence galaxies, across the simulated volumes.

For the purpose of this work, we do not distinguish between central and satellite galaxies and select all subhaloes with stellar mass $M_{\text{stars}} > 10^9 M_\odot$. This minimum mass corresponds to have only well-resolved galaxies in our sample, namely, with at least one thousand star particles in total in TNG100 and more than 100 in TNG300 (see the Appendix A for more details).

The stellar mass is defined, throughout this work, as the sum of all stellar particles within twice the stellar half-mass radius, $R_{\text{star,h}}$.

With these prescriptions, we gather, at $z = 0$, more than 18 800 galaxies in TNG100 and more than 253 000 galaxies in TNG300. We study galaxy populations between $z = 0$ and 2.

2.3 Star formation rates in IllustrisTNG

In the TNG model, the star formation is treated following Springel & Hernquist (2003): the gas is stochastically converted into star particles once its density exceeds $n_{\text{H}} = 0.1 \text{ cm}^{-3}$ and on a time-scale determined so that it empirically reproduces the Kennicutt–Schmidt relation (Kennicutt 1989). Star particles hence represent single-age stellar populations with a Chabrier stellar initial mass function (IMF Chabrier 2003). Gas cells are dubbed *star forming* when their density is larger than the above-mentioned threshold and their intrinsic SFR is thus non-vanishing. The SFR of a gas

¹<http://www.tng-project.org>

cell is inversely proportional to a density-dependent time-scale for star formation and is proportional to the subgrid estimate of cold gas mass. Such an SFR is *instantaneous*. Therefore, the SFR of a galaxy can be measured by summing up the ‘instantaneous’ SFRs of any given subset of its star-forming gas cells.

However, from an observational point of view, the SFR of galaxies is generally derived through several diagnostics and adopting the luminosity at different observational bands as an indicator for star formation. As widely discussed in the literature, such star formation tracers span the full electromagnetic spectrum, from X-ray and ultraviolet (UV) to radio wavelengths, via optical and infrared (IR) bands, and can be based on both continuum and line emissions (see the most recent review on the subject by Kennicutt & Evans 2012). Furthermore, different star formation indicators are understood to be sensitive to SFRs that are averaged over different time-scales, ranging from a few to hundreds of million years. Consequently, any comparison between observational and theoretical works, and even among observations, must take into account possible systematic effects due to the different values of SFRs derived via different tracers.

Finally, we note that, even if all galaxies taken into account in this work are resolved in terms of stellar and gas element counts, this is not the case for their SFR values. The finite numerical resolution of the simulation implies a minimum resolvable value for SFR (SFR_{\min}). Galaxies with $\log \text{SFR} < \log \text{SFR}_{\min}$ have unresolved SFR values and are artificially labelled with $\text{SFR} \equiv 0$ in the simulation output. The exact value of SFR_{\min} depends on the model, numerical resolution, and averaging time-scales over which the SFR is measured (see Appendix A for more details).

2.4 Identification of the star-forming main sequence and its scatter

Once a method has been adopted to distinguish between star-forming and quiescent galaxies (see next section) and irrespective of the definition of a galaxy’s SFR, we define the star-forming main sequence (MS) of the galaxy population on the $\text{SFR}-M_{\text{stars}}$ plane in two ways. In the first one, we perform a fit to the median SFR of star-forming galaxies in 0.2-dex bins of galaxy stellar mass in the mass range $M_{\text{stars}} = 10^9-10^{10.2} M_{\odot}$, where the MS can be considered linear, and we therefore linearly extrapolate the locus of the MS to larger galaxy masses. The following fitting equation is adopted throughout:

$$\log \left(\frac{\langle \text{SFR} \rangle_{\text{sf-ing galaxies}}}{M_{\odot} \text{ yr}^{-1}} \right) = \alpha(z) \log \left(\frac{M_{\text{stars}}}{M_{\odot}} \right) + \beta(z), \quad (1)$$

where α and β denote the logarithmic slope and normalization, respectively, of the MS and are both redshift dependent (see e.g. Speagle et al. 2014). We refer to this choice as ‘linearly extrapolated’ MS.

In the second way, we record the median SFR of star-forming galaxies as a function of galaxy stellar mass (in 0.2-dex logarithmic bins) across the whole range of available stellar masses. In this case, the MS is allowed to deviate from a linear trend, as in fact it does by exhibiting a turnover at about $M_{\text{stars}} \gtrsim 10^{10.5} M_{\odot}$, as suggested by several observations (Karim et al. 2011; Whitaker et al. 2014; Lee et al. 2015, 2018; Tomczak et al. 2016, see Section 4.2). We refer to this choice as ‘bending’ MS and we fit it with a fourth-order polynomial formula, where appropriate.

Throughout this work, the 1σ scatter of the MS is measured by stacking star-forming galaxies in stellar mass bins of 0.2 dex:

$$\sigma_{\text{MS}} = \sqrt{\frac{\sum_{i=1}^N (\text{SFR}_i - \langle \text{SFR} \rangle)^2}{N}}, \quad (2)$$

where N is the total number of star-forming galaxies in the considered mass bin and $\langle \text{SFR} \rangle$ is their median SFR. This is an intrinsic scatter, i.e. it does not take into account measurement uncertainties. As we demonstrate in Section 3.5, although the general features of the 1σ scatter of the TNG MS do not depend on the measurement choices, its values changes whenever different time-scales, apertures, or quenching definitions are considered.

2.5 Systematics and measurement choices

In this section we introduce and describe all the measurement choices used throughout this work: for the purposes of measuring the SFRs of simulated galaxies, we introduce SFR estimates based on the mass of stars in a galaxy formed over the last N Myr, with N equals to 10, 50, 200, and 1000 Myr, four physical galaxy apertures, and several selection criteria to separate quenched versus star-forming galaxies.

(i) **Averaging time-scales for SFR:** As already mentioned in the previous section, the instantaneous SFR measured from the gas is not an observable. Therefore, we additionally derive a galaxy’s SFR from the amount of stellar mass formed over the last 10, 50, 200, and 1000 Myr, based on the output formation times of individual stellar particles, with the goal of connecting to the time-scales associated with different SFR indicators. A systematic discussion about the differences across these measures and of how numerical resolution affects their values is given in Appendix A.

(ii) **Apertures:** With the aim of bracketing the typical galaxy apertures used in observational surveys, we measure the galaxy SFRs in different ways: (i) by considering all gas cells and/or star particles that are gravitationally bound to the galaxy according to the adopted SUBFIND algorithm; or by additionally restricting the galaxy measurements to the gas cells and/or star particles that are found within the following 3D spherical galactocentric distances: (ii) within twice the stellar half-mass radius; (iii) within 30 physical kpc (hereafter pkpc), and (iv) within 5 pkpc, the latter meant to mock an SDSS fibre aperture.

(iii) **Quenched versus star-forming galaxies:** From both theoretical and observational sides, the separation between ‘quenched’ versus ‘star-forming’ galaxies is ambiguous and somewhat arbitrary. For this reason, in this work we explore and contrast a number of different classification criteria to separate these two populations of galaxies. Details about our definitions are provided in the following, together with a schematic summary in Table 2. Throughout this work, we use the adjectives quenched, quiescent, and passive interchangeably and we do not distinguish between star-forming and star-burst galaxies or between quiescent and green-valley galaxies. All the following classifications can be adopted for any averaging time-scales for SFR.

2.5.1 Specific star formation rate threshold

As often adopted in the literature, we define as ‘quenched’ to be those galaxies whose logarithmic sSFR is lower than a given fixed value at any redshift, namely: $\text{sSFR} \leq 10^{-11} \text{ yr}^{-1}$. Star-forming galaxies are thus those with sSFR larger than this threshold. This fixed criterion does not account for the evolution in time of the

Table 2. Description of the different criteria used to separate star-forming versus quenched galaxies.

Classification criteria	Definition of ‘Quenched’ galaxies	Notes and comments
Population-based	$\log \text{sSFR} (\text{yr}^{-1}) < -11$ $\log \text{SFR} < \text{MS} - 2\sigma$ $\log \text{SFR} < \text{MS} - 1 \text{ dex}$ Colour–colour diagram	Logarithmic sSFR below -11 yr^{-1} . SFR below 2σ from the star-forming main sequence. SFR below 1 dex from the star-forming main sequence. UVJ colours satisfy equations (3)–(4).
Galaxy-based	$\text{SFR} < \overline{\text{SFR}}(t)$	SFR below the lifetime-average SFR.

SFR– M_{stars} plane (Whitaker et al. 2012, 2014; Lee et al. 2015, 2018; Tomczak et al. 2016), and it well describes the sSFR distributions only in a relatively narrow range of redshift. Yet, such threshold-based separation is often used in observations (McGee et al. 2011; Wetzel et al. 2013; Lin et al. 2014; Jian et al. 2018), albeit with somewhat different choices for the threshold value, since it roughly corresponds to the minimum of the bimodal distribution of galaxies’ sSFRs (see Kauffmann et al. 2004; Wetzel et al. 2012, and Section 4.1 of this work for a more detailed discussion).

2.5.2 Distance from the main sequence

A more robust alternative to separate ‘quenched’ versus ‘star-forming’ galaxies, and thus to identify the MS, involves an iterative algorithm. In practice, we stack all galaxies in 0.2-dex bins of stellar mass, measure the median SFR in the bin, and label ‘quenched’ (‘star-forming’) those galaxies whose SFR value falls below (above) a certain relative distance from the median SFR at the corresponding mass. We repeat this procedure by retaining at every step only star-forming galaxies until the median SFR in the bin is converged to a certain accuracy, thus defining the MS of star-forming galaxies. Unless otherwise stated, this is applied to galaxies up $M_{\text{stars}} = 10^{10.2} M_{\odot}$, otherwise the locus of the MS at every step is obtained by linearly extrapolating equation (1) from lower masses. We adopt two relative, mass-dependent distances from the MS to separate between quenched and star-forming galaxies: 2σ and 1 dex from the median SFR value, following e.g. Fang et al. (2018). Namely, quenched galaxies are those whose SFR (log SFR) is below 2σ (1 dex) from the median SFR at the corresponding galaxy stellar mass.

2.5.3 Time-averaged SFR threshold

The star formation histories of galaxies exhibit large variations across their lifetimes. For this reason, similarly to Mistani et al. (2016), we estimate the time-averaged SFRs of individual galaxies across their merger-tree main branch – $\overline{\text{SFR}}(t)$ – and define a galaxy to be ‘quenched’ if its SFR drops below this value at the time of observation. Therefore we use this method to mostly compare to previous theoretical works, without arranging for any comparisons with observations.

2.5.4 Colour–colour UVJ diagram

Most of the selection criteria described above are based on the position of galaxies on the SFR– M_{stars} plane. A powerful tool for selecting quiescent galaxies that is extensively used in observations is the colour–colour diagram, particularly the rest-frame $U-V$ versus $V-J$ plane (hereafter UVJ diagram; Williams et al. 2009, 2010; Whitaker et al. 2010; Wuyts et al. 2011; Patel et al. 2012; Quadri et al. 2012).

An empirical criterion based on the number density distribution of galaxies on the UVJ diagram is usually adopted to divide the sample into red/quiescent versus blue/star-forming galaxies: galaxies located approximately on the top-left portion of the UVJ diagram are dubbed ‘quenched’, and star-forming otherwise. The slope and the position of the separating diagonal cut in the UVJ diagram are fine-tuned to fall on the break of the bimodal number distribution of galaxies (Williams et al. 2009). Since the rest-frame colour distribution could be slightly different from one galaxy sample to another, some authors adjust arbitrarily the cut to better separate the two peaks of the galaxy number density in their sample (Whitaker et al. 2011; Muzzin et al. 2013; Tomczak et al. 2014). We adopt a similar strategy for TNG galaxies: see next section.

Stellar light of simulated galaxies is determined following Nelson et al. (2018a), i.e. adopting the FSPS stellar population synthesis code (Conroy, Gunn & White 2009; Conroy & Gunn 2010; Foreman-Mackey, Sick & Johnson 2014) with the Padova isochrones, MILES stellar library, and assuming a Chabrier initial mass function (Chabrier 2003). The effects of dust attenuation are accounted for by including unresolved and resolved dust (Model C of Nelson et al. 2018a, with the redshift dependences outlined there and adopted across the $z = 0-2$ range), in one random galaxy projection in the simulated volume. A galaxy’s spectrum is the sum of the stellar population spectra of all star particles in a given aperture and is convolved with the rest-frame broad-band filters U , V , and J centred at 365, 551, and 1220 nm, respectively, to obtain $U-V$ and $V-J$ colours.

2.5.5 Fiducial measurement choices

Unless otherwise stated, we elect to use as fiducial the following measurement choices: we use SFRs averaged over 200 Myr, obtained from all stars within twice the 3D spherical stellar half-mass radius, and with a separation of quiescent versus star-forming galaxies based on the TNG UVJ diagram.

3 RESULTS

3.1 The UVJ diagram of TNG galaxies

We start the presentation of the TNG results by examining the simulated colour–colour diagram of TNG galaxies. Fig. 1 shows the UVJ diagram at $z = 0$ (left-hand panel) and $z = 2$ (right-hand panel) by modelling the effects of dust extinction as discussed in Nelson et al. (2018a). We divide all galaxies in 150×150 bins on the $(V-J)$ – $(U-V)$ plane: the colour scale indicates the number of galaxies in each bin and the solid contours encompass 20, 50, 70, 90, and 99 per cent of the simulated galaxies with $M_{\text{stars}} \geq 10^9 M_{\odot}$.

In both panels, a clear separation between two groups of galaxies is visible. This is the case in terms of galaxy number density and it appears more emphasized at high redshift. The existence of two classes of galaxies in the TNG colour–colour is a validation of the

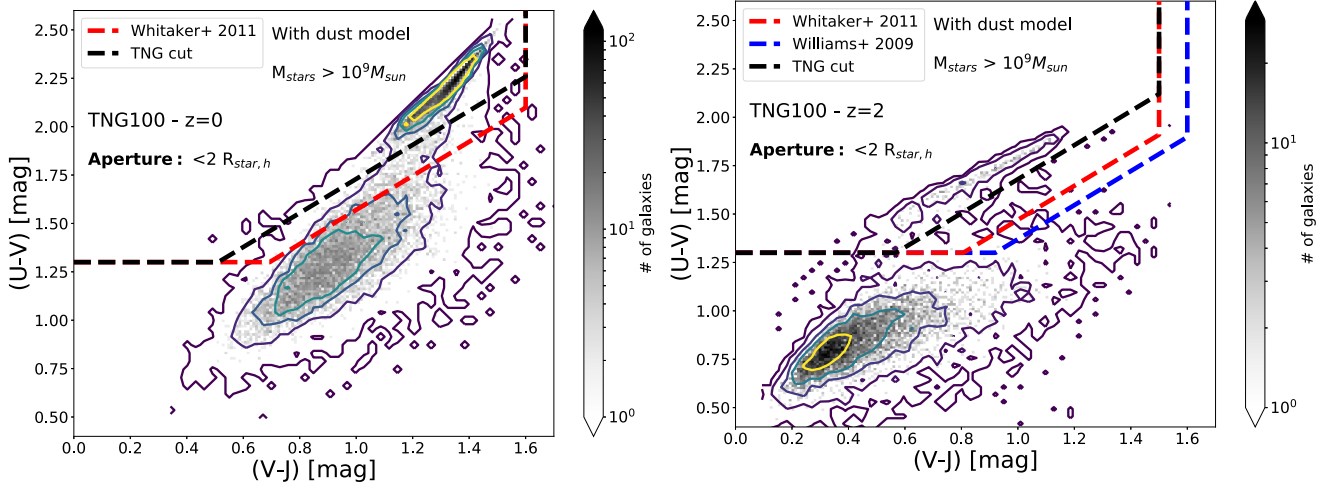


Figure 1. *UVJ* diagram for TNG100 galaxies at $z = 0$ (left-hand panel) and $z = 2$ (right-hand panel). We divide all galaxies in 150×150 bins of the plane $(V - J) - (U - V)$. The colour scale indicates the number of galaxies in each bin and the solid contours encompass 20, 50, 70, 90, and 99 per cent of the galaxies. The black dashed line is the cut based on the number density distribution of TNG galaxies. For comparison, two observational thresholds are shown, defined by Whitaker et al. (2011) (red dashed lines) and by Williams et al. (2009) (blue dashed line at $z = 2$).

model in comparison to observational findings. It provides further support to the separation into star-forming versus quiescent galaxies based on *UVJ* cuts that we adopt throughout the paper.

In fact, the red dashed line in Fig. 1 is the separating threshold to select star-forming versus quenched galaxies adopted by Whitaker et al. (2011). The blue dashed line denotes the analogue separation adopted by Williams et al. (2009), which deviates from the red one only at high redshifts. Fig. 1 demonstrates that, despite the separating red and blue dashed lines are based on observed galaxy samples and observed colours, they can be reasonably well applied also to TNG galaxies on the *UVJ* diagram and could be adopted to separate quite successfully red and quenched galaxies (in the upper part) from blue and star-forming ones (in the lower portion of the diagram).

Yet, since the TNG galaxies populate the *UVJ* diagram in a manner which is broadly consistent with but not identical to the observed ones, to distinguish between star-forming and quenched galaxies, we use the same slope of the separating threshold provided by Whitaker et al. (2011) and we adjust its position to better separate the bimodal number-density distribution of TNG galaxies.

The black dashed line in Fig. 1 denotes the best estimate to separate TNG galaxies in the *UVJ* diagram based on the depicted number density distributions and is described by the following equations:

$$\begin{aligned} (U - V) &> 0.88(V - J) + 0.85 & [z \leq 1] \\ (U - V) &> 0.88(V - J) + 0.80 & [1 < z < 2]. \end{aligned} \quad (3)$$

In addition, for the limits on $U - V$ and $V - J$ we arbitrarily use the ones defined in Whitaker et al. (2011):

$$\begin{aligned} (U - V) &> 1.3; & (V - J) < 1.6 & [0 < z \leq 1.5] \\ (U - V) &> 1.3; & (V - J) < 1.5 & [1.5 < z \leq 2] \\ (U - V) &> 1.2; & (V - J) < 1.4 & [2 < z < 3.5]. \end{aligned} \quad (4)$$

Throughout this work, we use the TNG *UVJ* cut as our fiducial choice and compare our results, whenever necessary, to those obtained with other colour cuts widely used in the literature. We note that, while Fig. 1 reports the distribution of TNG100 galaxies,

a consistent picture emerges from the distribution of TNG300 galaxies, with enhanced statistics but lower numerical resolution. Moreover, it may be useful to point out that the effect of dust in the *UVJ* plane is to shift galaxies more or less along the diagonal portion of the threshold dashed lines. However we note that, in both TNG100 and TNG300, we do not reproduce the full range of colours that is seen in real galaxies at intermediate redshifts. Specifically, e.g. at $z = 2$, the TNG model does not produce the very dusty star-forming galaxies that occupy the upper right corner of the *UVJ* diagram in the observed samples and for which the vertical boundary is needed. This is similar to what found in the MUFASA simulation (Davé et al. 2017) and requires future and further analysis.

3.2 TNG quenched fractions across times, masses, and operational definitions

The majority of the star-forming versus quenched classification criteria adopted in this work can be grouped into two general classes: some are based on the position of each galaxy on the $\text{SFR} - M_{\text{stars}}$ plane and hence its distance from the star-forming MS; others are based on the position of each galaxy on the colour-colour diagram. Here, we present a quantitative comparison among these two operational definitions of quenched galaxies. The goal is to understand whether or not SFR-based and *UVJ*-based classifications lead to quantitatively consistent results in terms of the shape and evolution of the MS and of the amount and evolution of quenched galaxies. Specifically, in Fig. 2 we examine how our colour-colour selection for quenched galaxies compares to other SFR-based selections, and vice versa.

The top left panel of Fig. 2 shows the $\text{SFR} - M_{\text{stars}}$ plane of TNG100 at $z = 0$ in which galaxies are colour-coded according to their SFR: red/quenched if they are below 1 dex from the MS, blue/star-forming if they are above.

As we discuss more in detail in Appendix A, all galaxies with SFR below the minimum resolved values will be assigned to the range $10^{-5} - 10^{-4} M_{\odot} \text{ yr}^{-1}$. Even if their actual distribution is unknown, those galaxies would distribute well below the MS, thus populating the quenched region.

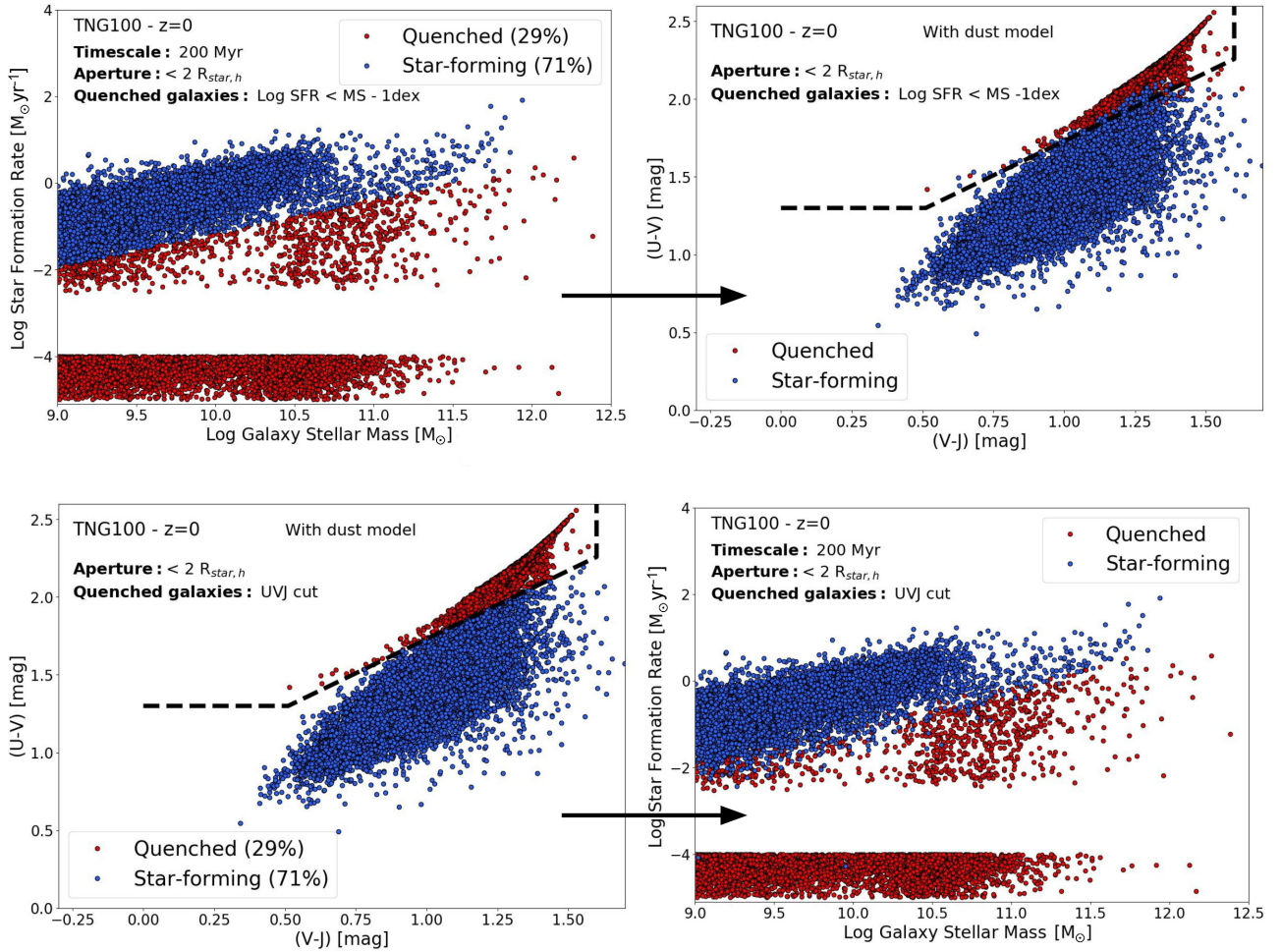


Figure 2. Comparison between SFR-based and UVJ colour-based criteria to distinguish between quiescent versus star-forming galaxies. Top: SFR as a function of galaxy stellar mass (left-hand panel) and UVJ diagram (right-hand panel) of TNG100 galaxies at $z = 0$. In both panels, we identify quenched galaxies (red circles) if their SFR is below 1 dex from the MS. All the other galaxies are identified as star-forming (blue circles). Bottom: the same as in the top but for galaxies selected using the TNG UVJ cut indicated with a black dashed line.

The top right panel shows the UVJ diagrams for the same galaxies, with the same colour coding adopted from the definition applied in the left-hand panel. The black dashed line is the separating TNG colour–colour cut (Equations 3 and 4). A good agreement between the two criteria is manifest: nearly all galaxies identified as blue/star-forming according to their SFRs, in the top left panel, settle below the black line in the top right panel.

This statement holds also in reverse. The bottom left panel of Fig. 2 shows the UVJ diagram for TNG100 galaxies, at $z = 0$. In this case, the galaxies are colour-coded according to their position with respect to the separating threshold (dashed black line). Then, with the same colour code, we show the position of these galaxies in the SFR– M_{stars} plane (bottom right panel). As in the previous case, galaxies identified as blue/star-forming accordingly to their colour in the bottom left panel, populate the region above $\log \text{SFR} > \text{MS} - 1$ dex in the bottom right panel.

These findings are a further confirmation that the UVJ colours of galaxies correlate well with their SFRs; practically, the simulated UVJ diagram provides a reliable tool to separate galaxies in star-forming and quenched populations by returning results that are overall consistent with SFR-based selections where quiescent galaxies lie 1 dex below the MS. This is in agreement with observational

findings by e.g. Fang et al. (2018). In fact, although in the upper panels some star-forming galaxies pass the UVJ colour threshold and, vice versa, in the lower panels a fraction of galaxies that are identified as star-forming according to their colours turn out to have a quite low SFR. In the whole mass range considered here ($M_{\text{stars}} \geq 10^9 M_{\odot}$), the fraction of quenched galaxies reads 29 per cent with both criteria. Such an excellent quantitative agreement between the two operational definitions of quenched galaxies supports the choice of a UVJ cut tailored to the number density distribution of TNG galaxies.

In fact, even if not shown in Fig. 2, we have repeated the same comparison using the separating UVJ threshold provided by Whitaker et al. (2011) instead of the TNG one. In this case, we find that, at $z = 0$, the SFR-based selection leads to a smaller fraction of quenched galaxies in comparison to the UVJ selection, with 7 per cent of galaxies in the studied mass range being dubbed differently by the two different UVJ cuts. While such mismatch is not negligible, all the results obtained with the Whitaker et al. (2011) observational cut are still in qualitative agreement with TNGs, providing a further qualitative validation of the model.

In the top panel of Fig. 3, we quantify how different selection criteria impact the assessment of the fraction of quenched galaxies

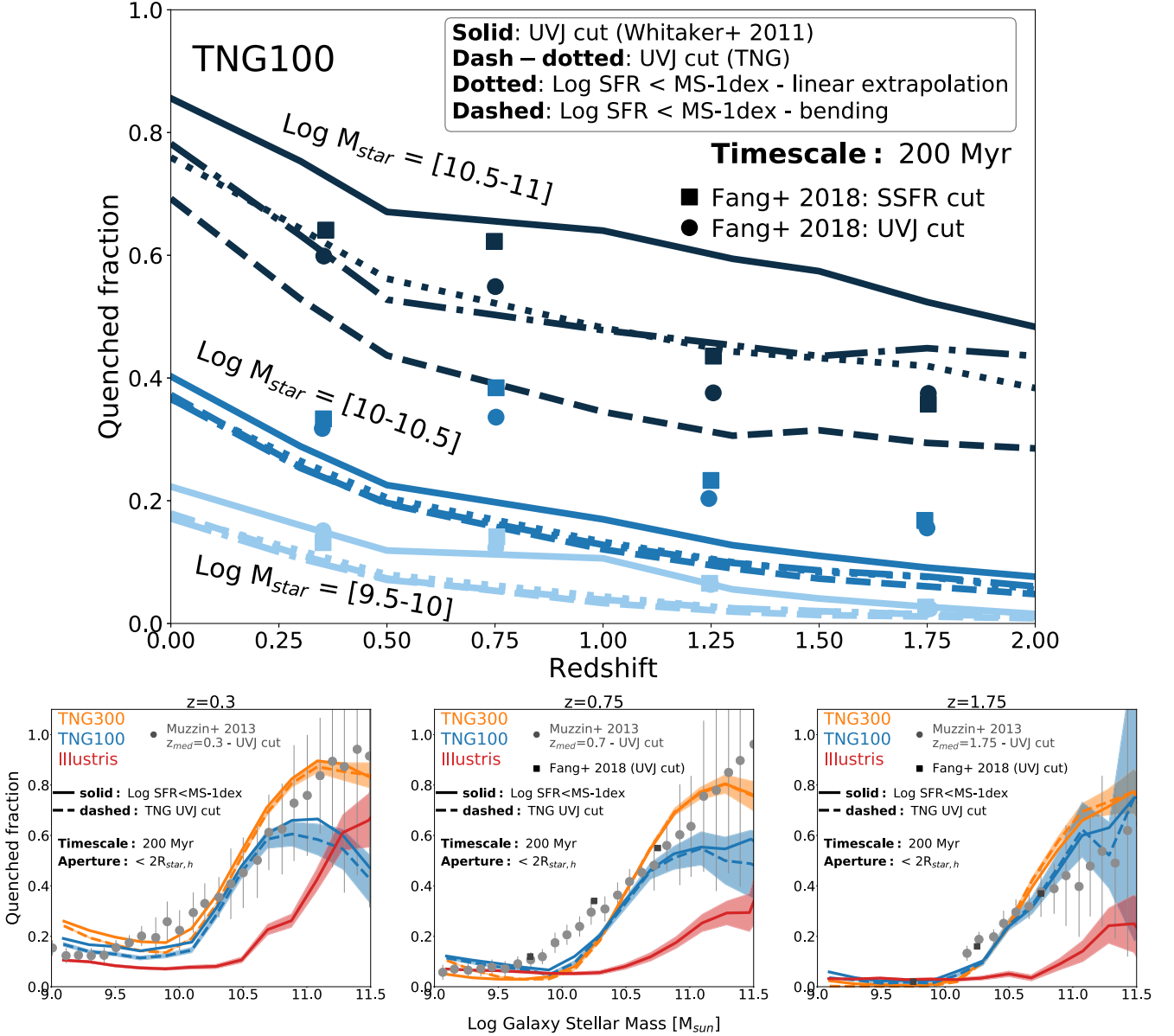


Figure 3. Top panel: fractions of quenched galaxies, in three bins of stellar mass (indicated with different colours), as a function of redshift for TNG100. Quenched galaxies are selected in four ways: via the *UVJ* cut adopted by Whitaker et al. (2011) (solid), via the TNG *UVJ* cut (dash-dotted) and according to their SFR values with respect to the locus of the star-forming main sequence, i.e. by both extrapolating the linear MS to the high-mass end (dotted) and by using a possibly bending MS (dashed). Observed quenched fractions of *CANDELS* galaxies (Fang et al. 2018) are indicated as filled symbols (squares for the sSFR selection and circles for the *UVJ* selection). The colour code indicates the stellar mass bins, as in TNG. At the highest mass end probed here, different quenched definitions may return quenched fractions that differ by up to 10–20 percentage points. Bottom panels: fractions of quenched galaxies (centrals and satellites) as a function of stellar mass at $z = 0$ (bottom left), $z = 0.75$ (bottom central), $z = 1.75$ (bottom right), for TNG300 (orange curves), TNG100 (blue curves), and Illustris (red curves). Quenched galaxies are selected in two ways: via the TNG *UVJ* cut (dashed) and via the linear extrapolation of the MS at the high-mass end (solid). For all curves, we include the effect of 0.2 and 0.3 dex uncertainties on M_{stars} and SFRs, respectively. Shaded areas in all panels indicate the Poissonian error associated with the *UVJ* selection for TNG models and to SFR selection for Illustris. A selection of observational data are indicated as grey symbols. TNG models return a higher quenched fraction than Illustris, at any redshift.

across galaxy stellar mass and redshifts. We provide quenched fractions in bins of stellar mass, (indicated with different colours, light blue for low-mass galaxies, $10^{9.5} - 10^{10} M_{\odot}$, blue for intermediate galaxies, $10^{10} - 10^{10.5} M_{\odot}$ and dark blue for the more massive ones, $10^{10.5} - 10^{11} M_{\odot}$), in the redshift range $0 \leq z \leq 2$. We select quenched galaxies according to their *UVJ* colours and their SFR values. For the *UVJ* cut, we assume both the one provided by Whitaker et al. (2011) (solid curves) and the one based

on the TNG diagram i.e. equations (3) and (4) (dash-dotted curves). For the SFR-based selection where galaxies are quiescent if their SFR is 1 dex below the star-forming MS, we use both the linear extrapolation of the MS beyond $10^{10.2} M_{\odot}$ (dotted curves) as well as we allow the MS to bend at the high-mass end (dashed curves: see Section 2.4 for the definitions and Section 4.2 for a detailed discussion on the bending of the MS). In all cases, we account for all galaxies that are resolved according to the adopted stellar mass

limit ($M_{\text{stars}} \geq 10^{9.0} M_{\odot}$), hence including in the galaxy counts also those with very low values of SFR.

In general, we find that the fraction of quenched galaxies increases from $z = 1.75$ to $z = 0.3$, with a larger increment of TNG300 quenched galaxies at larger masses, being about 70–80 per cent at $z = 1.75$ and reaching more than 90 per cent at $z = 0.3$. Moreover, at all redshifts, the quenched fractions are higher for galaxies in the $10^{10.5} < M_{\text{stars}} (M_{\odot}) < 10^{11}$ mass range with respect to the less massive ones, the latter featuring quenched fractions below about 20 per cent at all times.

For galaxies in the stellar mass range $M_{\text{stars}} = 10^{9.5} - 10^{10.5} M_{\odot}$, the fractions of quenched galaxies via the two SFR-based criteria are perfectly consistent with one another: dashed versus dotted curves. As we will see in Section 3.3, in this mass range the MS is approximately linear. Interestingly, in the same stellar mass range, the excellent agreement previously mentioned between the *UVJ*-based selection – with the TNG cut – and the SFR-based one (Fig. 2), holds even at high redshifts, as suggested by the overlap among dash-dotted, dotted, and dashed curves. Instead, the *UVJ* cut of Whitaker et al. (2011) leads to a systematically higher quenched fraction, even though the amplitude of such a discrepancy is less than 0.05 at all redshifts.

Let us now focus on more massive galaxies: $10^{10.5} \leq M_{\text{stars}} (M_{\odot}) \leq 10^{11}$. In this mass regime, the TNG *UVJ* selection leads to fractions of quenched galaxies that are very similar to the ones found with the linear extrapolation of the MS: ~ 80 , 50, and 40–45 per cent at $z = 0$, 1, and 2, respectively. A fraction that is systematically lower by 10 percentage points is instead recovered when quiescent galaxies are selected by using a curved MS. As in the case of low-mass galaxies, the *UVJ* cut provided by Whitaker et al. (2011) turns into higher quenched fractions with respect to the TNG, of about ~ 10 , 20, and < 10 percentage points at $z = 0$, 1, and 2, respectively.

A similar comparison between *UVJ* and SFR selections has been made by Fang et al. (2018) (filled symbols in the top panel of Fig. 3), who have examined the fraction of quenched *CANDELS* galaxies, in bins of stellar mass, at $0.2 < z < 2.5$. A qualitative and a somewhat quantitative agreement in terms of general redshift and mass trends and results is manifest, barring perhaps the $10^{10} - 10^{10.5} M_{\odot}$ mass range. However we do not push further for any conclusion, since *CANDELS* SFRs are obtained via UV-optical spectral energy distribution fitting including specific light-to-SFR calibrations we do not mock.²

In the bottom panel of Fig. 3 we show the fractions of quenched galaxies (without distinguishing between centrals and satellites) as a function of stellar mass at $z = 0.3$ (bottom left), $z = 0.75$ (bottom central), and $z = 1.75$ (bottom right), for TNG300 (orange), TNG100 (blue), and Illustris (red). For the TNG models we adopt both the TNG *UVJ*-based selection (dashed) and the SFR-based selection (linearly extrapolated MS, solid) to identify quenched galaxies, whereas for Illustris we only use the SFR-based criterion. In order to take into account observational uncertainties on galaxy stellar masses and SFRs, we add a random Gaussian error to the

simulation data, with a width of 0.2 dex for the stellar mass and 0.3 dex for the SFRs.

At any redshift, we find that the fractions of quenched galaxies is systematically higher in TNG than in Illustris: this actually brings the TNG model into better agreement with observational constraints. Observed quenched fractions of COSMOS/UltraVISTA galaxies (Muzzin et al. 2013) are indicated as grey filled circles in the bottom panels of Fig. 3, where quiescent galaxies are defined via the *UVJ* cut provided by Whitaker et al. (2011). For comparison we additionally show the results from Fang et al. (2018) as black filled squares at $z = 0.75$ and $z = 1.75$ consistently to what is reported on the top panel. As seen above for the simulations, also for observational results different quenching definitions might cause a variation in the values of the quenched fractions at different redshifts. For example, in Darvish et al. (2016) quiescent galaxies are selected via an *NUV*–*r*–*J* cut: although at lower redshift the quenched fractions are in a quite good agreement with other observations (e.g. Moustakas et al. 2013; Muzzin et al. 2013), at higher redshift – namely $z \sim 1.75$ – Darvish et al. (2016) provide a fraction of about 15–20 per cent lower than Muzzin et al. (2013).

Because of the new AGN feedback recipes implemented in the TNG model,³ quenching mechanisms are more efficient in TNG than in Illustris, especially for galaxies more massive than $10^{10} M_{\odot}$, for which the fraction of quenched galaxies can be different by up to about 50 percentage points between the two simulations. This is more apparent at higher redshift, $z > 0$: the quenched fractions in Illustris reach $\gtrsim 20$ per cent only for galaxies more massive than $10^{11} M_{\odot}$, whereas both TNG100 and TNG300 return fractions in the range 30–80 per cent in the stellar mass range $M_{\text{stars}} = 10^{10.5} - 10^{11.5} M_{\odot}$.

However, because of resolution effects, differences between TNG100 and TNG300 are clearly noticeable: at lower resolution the fractions of quenched galaxies are larger. Particularly, at $z = 0$ we recover a fraction of about 10–20 per cent higher in TNG300 with respect to TNG100, in the whole mass range, whereas at $z = 0.75$, a difference of about 20 percentage points is apparent only for galaxies more massive than $10^{10.5} M_{\odot}$. Besides the mass resolution effects, the volume of the simulated box could play a crucial role. Indeed, TNG300 samples much denser environments with respect to TNG100, and many of the most massive object will be quite different.

In comparison to previous results from other cosmological models, e.g. EAGLE (Furlong et al. 2015) and L-Galaxies (Henriques et al. 2017), TNG returns similar quenched fractions at low redshifts, but systematically higher ones at higher redshifts. As opposed to the TNG model, the fraction of quenched galaxies more massive than $10^{10.5} M_{\odot}$ predicted by EAGLE (where passive galaxies are selected via sSFR cut) at $z \sim 1$ and 2, falls below the observed ones (Muzzin et al. 2013, and others), by ~ 10 –15 per cent. Similarly, the passive fractions predicted by L-Galaxies at $z = 0.86$ (where

²Note that in the Fang et al. (2018) analysis, selection effects are necessarily in place: the *CANDELS* sample here utilized is selected based on a magnitude limit ($H = 24.5$) and on the quality of the GALFIT fits. None of such cuts implies rejection of objects based on their SFR values and in fact none of them biases the values of quenched fractions used for comparison in this paper. Similar arguments apply to the samples of Muzzin et al. (2013).

³The SMBH model in IllustrisTNG has been significantly updated in comparison to the original Illustris simulation. In particular, in the low-accretion mode regime, the bubble thermal heating feedback implemented in Illustris has been replaced with a kinetic wind. Moreover, SMBHs are seeded at larger masses and a BH mass-dependent threshold is used to prevent low-mass SMBHs to enter the low accretion state (see Weinberger et al. 2018, for more details). The new TNG kinetic feedback at low accretion rates has been demonstrated to be responsible for the quenching of galaxies residing in massive haloes (Weinberger et al. 2018) and to improve the galaxy colour bimodality in comparison to SDSS (Nelson et al. 2018a).

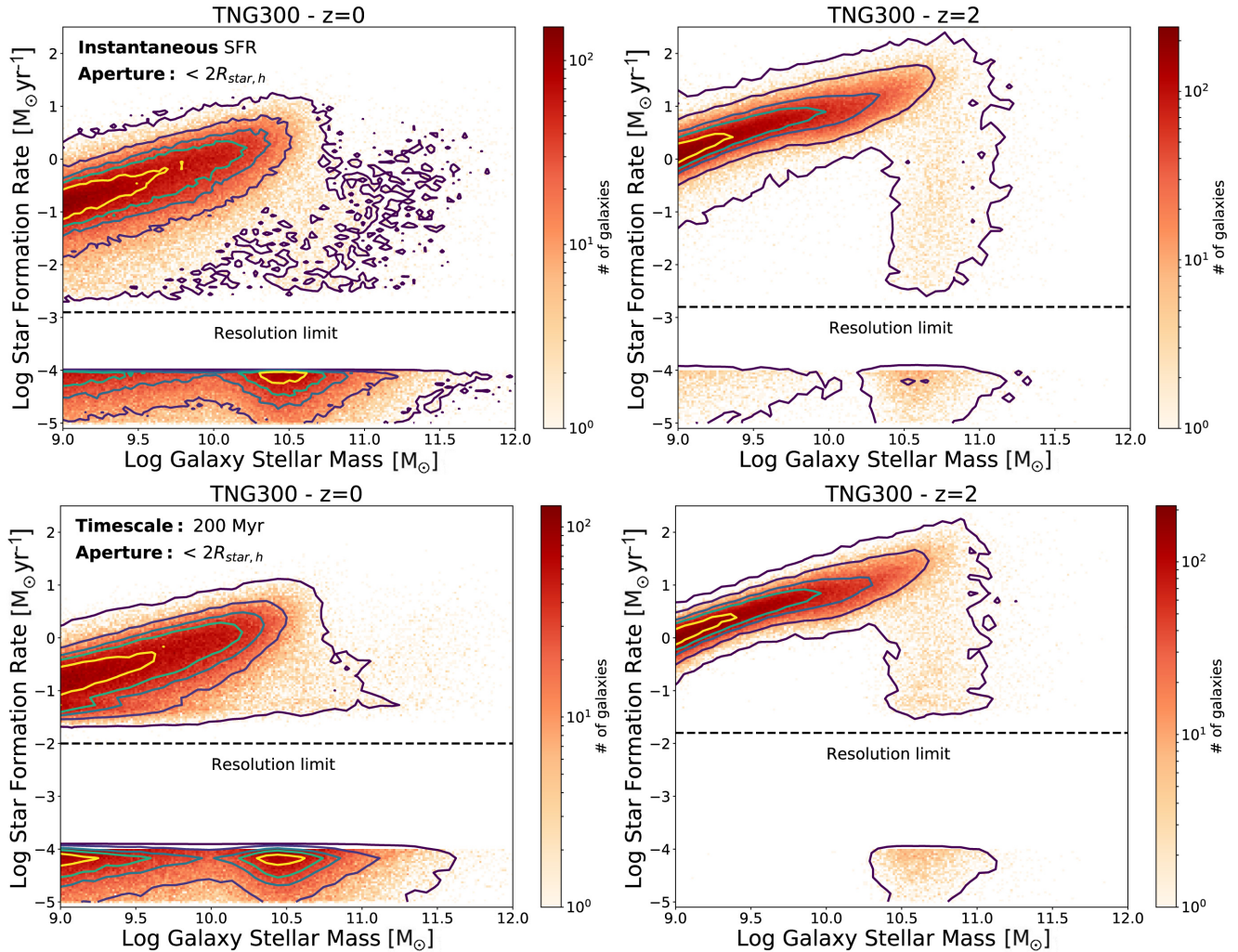


Figure 4. SFR as a function of stellar mass for TNG300 galaxies, at $z = 0$ (left-hand panels) and $z = 2$ (right-hand panels). The colour scale indicates the number of galaxies and the solid contours encompass 20, 50, 70, 90, and 99 per cent of the studied galaxies. To all galaxies with unresolved SFR values, we assign a random value in the range $\text{SFR} = 10^{-5} - 10^{-4} M_{\odot} \text{yr}^{-1}$. Galaxy SFRs are measured within $2 \times R_{\text{star,h}}$ and are estimated from the instantaneous SFRs of the gas (top panels) and from the stars formed in the last 200 Myr (bottom panels).

passive galaxies are identified with a $U - B$ versus M_{stars} cut) fall below constraints from COSMOS (compilation by Peng et al. 2010) by 15–20 per cent too, suggesting that quenching mechanisms are less effective in those models than what observations imply.

3.3 The SFR– M_{stars} plane of TNG galaxies and their main sequence

In this section, we present the general properties of the SFR– M_{stars} plane and characterize the MS of IllustrisTNG galaxies in the redshift range $0 \leq z \leq 2$.

In Fig. 4, we show the SFR– M_{stars} plane at $z = 0$ (left column) and $z = 2$ (right column) for TNG300 galaxies. Here, the SFRs are measured from the instantaneous SFRs of the galaxies’ gas cells (top panels) and from the stellar mass formed in the last 200 Myr (bottom panels). In both cases, the physical aperture is $2 \times R_{\text{star,h}}$. We divide all galaxies in 200×200 bins on the $\log \text{SFR} \times \log M_{\text{stars}}$ plane. The colour scale in Fig. 4 denotes the number of galaxies in

each bin whereas the solid contours encompass 20, 50, 70, 90, and 99 per cent of our galaxy sample.

In the stellar mass range $10^9 - 10^{10.5} M_{\odot}$, the SFR– M_{stars} plane exhibits a densely populated region in which galaxies settle: these form the well-known star-forming main sequence that is clearly visible in all four panels, i.e. at low and high redshifts, and for both the instantaneous and averaged SFRs. A smaller number of galaxies appears to fall off from the star-forming main sequence, by exhibiting lower SFRs than MS galaxies at the same stellar mass. Moreover, we find that ~ 32 per cent of galaxies have unresolved SFR values, due to the mass resolution limits of the simulation. As we discuss more in detail in Appendix A, the minimum resolved SFR value is $\log \text{SFR} (M_{\odot} \text{yr}^{-1}) \sim -2$, meaning that all galaxies with SFR below this limit, will be assigned to the range $10^{-5} - 10^{-4} M_{\odot} \text{yr}^{-1}$.

Furthermore, we find that the number of these galaxies with very low SFRs decreases to 3 per cent at $z = 2$ (right-hand panel), when they are mainly more massive than $> 10^{10} M_{\odot}$. Indeed, as suggested by observations, at high redshifts, although a quenched population is recovered, the SFR– M_{stars} plane is dominated by star-

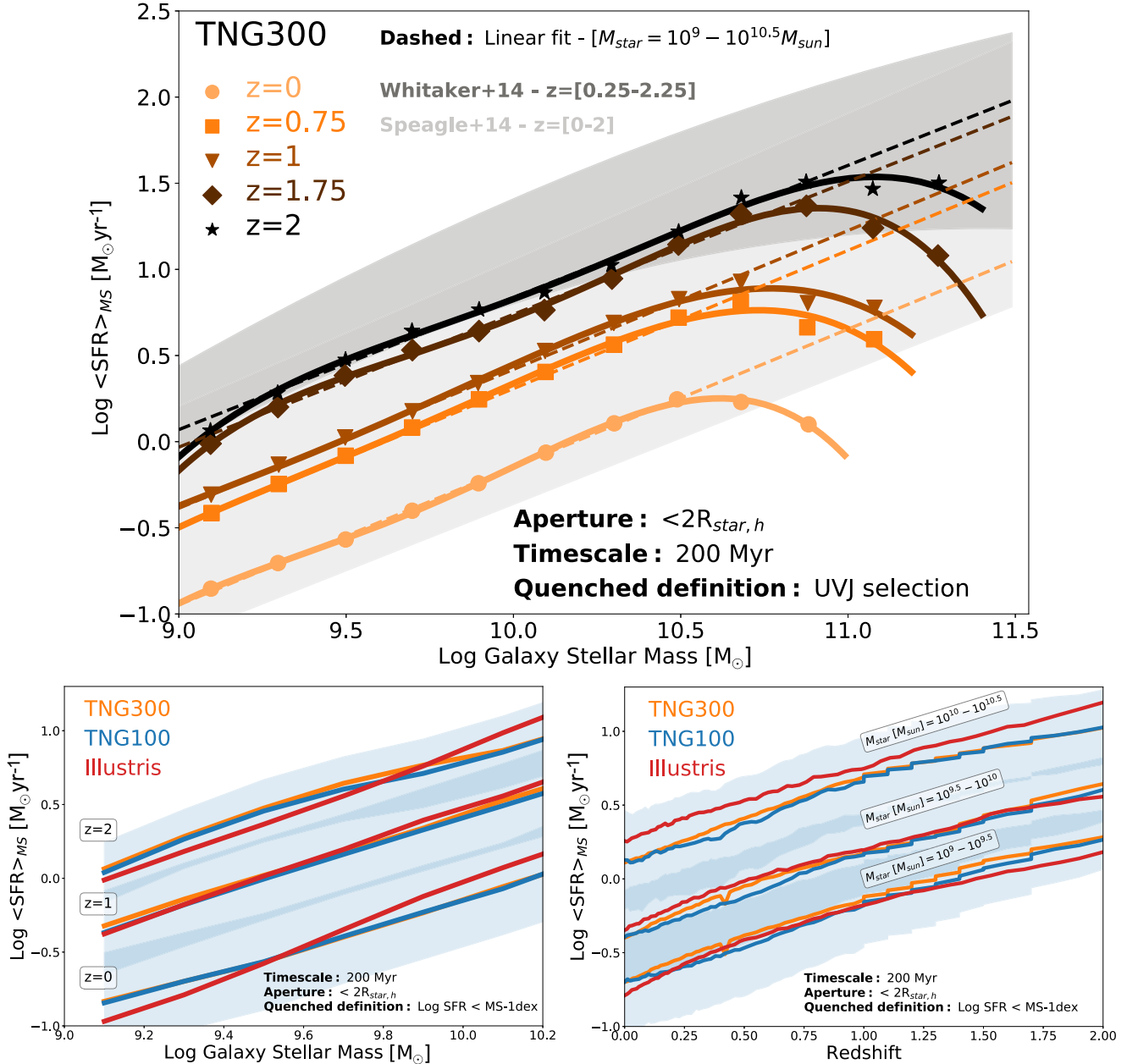


Figure 5. Main sequence of TNG star-forming galaxies, computed with the following measurement choice: the SFR from stars is averaged over 200 Myr and measured within $2 \times R_{\text{star},h}$. Quenched galaxies are neglected in two ways: with a TNG *UVJ* cut (top panel) and with $\log \text{SFR} < \text{MS}-1$ dex (bottom panels). Colours indicate different redshifts: $z = 0, 0.75, 1, 1.75, 2$ (from light orange to black symbols in the top panel). Dashed lines are the linear fit performed in the mass range $M_{\text{stars}} = 10^9 - 10^{10.5} M_{\odot}$ (equation 1). A clear deviation from a linear trend is evident for galaxies more massive than $10^{10.5} M_{\odot}$. In the stellar mass range $M_{\text{stars}} = 10^9 - 10^{10.2} M_{\odot}$, the MS of TNG300 galaxies is compared to TNG100 (blue line) and Illustris (red line) at $z = 0, 1, 2$ (bottom left) and by stacking galaxies in three stellar mass bins (bottom right). The shaded areas indicate the 1σ scatter of the MS in TNG100. An excellent agreement is noticeable between TNG models, whereas a steeper slope and a higher normalization at the high-mass end is recovered in the MS of Illustris galaxies.

forming galaxies at all stellar masses (see e.g. Whitaker et al. 2010).

These general features are confirmed even when the mean SFR is measured over longer time-scales (bottom panels). In this case, at $z = 0$, the number of galaxies with unresolved SFR values is larger with respect to the top panel (~ 37 per cent). We postpone a comparison among different time-scales to Section 3.4 and a detailed discussion on the effects of numerical resolution to Appendix A.

We now focus on the properties of the main sequence, its evolution with time and with galaxy stellar mass. Henceforth, to identify the MS, we refer to our *fiducial choices* (see Section 2.5.1): SFRs are averaged over 200 Myr and measured within a physical aperture of $2 \times R_{\text{star},h}$; quenched and star-forming galaxies are separated using the TNG *UVJ* cut, or alternatively using the SFR-based selection ($\log \text{SFR} < \text{MS}-1$ dex), depending on the cases: throughout the paper we specify whether we adopt one criterion rather than the other.

Table 3. Values of the coefficients from the linear fit for TNG300 MS (Fig. 5), performed in the stellar mass range $M_{\text{stars}} = 10^9\text{--}10^{10.5} M_{\odot}$ (equation 1). The MS is computed with the SFR from stars averaged over 200 Myr, measured within $2 \times R_{\text{star,h}}$. Quenched galaxies are neglected if they are below 1 dex from the MS. Columns read: (1) redshift; (2) and (3) coefficients of the linear fit.

Redshift	α	β
0	0.80 ± 0.01	-8.15 ± 0.11
0.75	0.80 ± 0.01	-7.72 ± 0.03
1	0.81 ± 0.02	-7.64 ± 0.17
1.75	0.77 ± 0.02	-6.97 ± 0.24
2	0.77 ± 0.03	-6.83 ± 0.25

The top panel of Fig. 5 shows the MS of TNG300 galaxies at $z = 0, 0.75, 1, 1.75, 2$, computed using the fiducial choices and selecting star-forming galaxies with the *UVJ* selection. In addition, we show the results of a fourth-order polynomial fit (solid line) and of a linear fit (dashed line) performed in the stellar mass range $M_{\text{stars}} = 10^9\text{--}10^{10.5} M_{\odot}$ (equation 1), to better highlight the bending of the MS at the high-mass end.

The relevant observational features of the MS are also recovered for TNG star-forming galaxies. Indeed we find that the normalization of the MS increases towards higher redshifts (from light orange to black line), regardless of stellar mass. Moreover, while for galaxies with $M_{\text{stars}} < 10^{10.5} M_{\odot}$, the MS can be considered linear, at the high-mass end it exhibits a bending, with a shallower slope. The results of the linear fit as per equation (1) are given in Table 3, for reproducibility.

The shaded grey areas in Fig. 5 represent the MSs of Whitaker et al. (2014) in the redshift range $0.25 < z < 2.25$ (dark grey area) and of a compilation of observational works taken by Speagle et al. (2014) (light grey area) in the redshift range $0 \leq z \leq 2$. The MS of TNG300 galaxies falls within the observational ball park: however, making a comparison at face-value we find a systematic discrepancy between the MS in our model with respect to the observed values, particularly at high redshifts (see Section 4.3 for more details).

In the bottom panels of Fig. 5 we show the MS as a function of galaxy stellar mass (left-hand panel) and redshift (right-hand panel), for TNG300 (orange line), TNG100 (blue line), and Illustris (red line) galaxies, in the stellar mass range $M_{\text{stars}} = 10^9\text{--}10^{10.2} M_{\odot}$. Here, the MS is measured using the fiducial choices and neglecting galaxies with SFR values below 1 dex from the MS. The shaded areas in both panels represent the 1σ scatter of the MS computed for TNG100. In both panels an excellent agreement between TNG models is manifest: this demonstrates that the different mass resolution and statistical sampling of the two runs affect neither the slope nor the normalization of the MS, at any redshift.

On the other hand, the MS exhibits a steeper slope in Illustris than in the TNG models, particularly at $z = 0$ and $z = 2$ and at the highest mass end (left bottom panel). These results expand upon the analysis provided by Sparre et al. (2015) for the Illustris simulation. The inability of Illustris feedback to quench the star formation in massive galaxies ($M_{\text{stars}} > 10^{10} M_{\odot}$) turns into a ~ 0.1 dex higher normalization with respect to TNG at all times (right-hand panel). These results are in line with the findings of Hayward et al. (in preparation) who study the populations of submm galaxies in both TNG and Illustris at intermediate redshifts. The higher SFRs of Illustris galaxies at the high-mass end are most probably due to

the different AGN feedback recipes implemented in Illustris and TNG and subsequently produce different fractions of quenched galaxies as a function of stellar mass, as we demonstrate in Fig. 3 (see Donnari et al. in preparation for a detailed discussion on this topic).

In summary, the general properties of the star-forming MS of TNG galaxies and its trends with the stellar mass and redshift are in qualitative agreement with observational findings. We postpone a more quantitative analysis to Section 4.3, after we quantify the effects of different measurement choices in the next section.

3.4 The effects of different measurement choices on the MS

As mentioned in Section 1, different observational techniques adopted to infer stellar masses and SFRs of galaxies, along with uncertainties on the IMF and diverse galaxy sample selection functions, can lead to a difficult quantitative comparison between theoretical models and observations and, not rarely, also among observational results themselves (see e.g. Speagle et al. 2014, and references therein). In this section, we quantify how the different measurement choices introduced in Section 2.5 affect the locus of the MS of TNG galaxies and hence provide plausible guidelines for comparison across different datasets.

In Fig. 6, we show the MS as a function of galaxy stellar mass in TNG100 at $z = 0$ (left-hand column) and $z = 2$ (right-hand column). Each row of the figure reports the study of a measurement choice: different time-scales (top panels), apertures (central panels), definitions of quenched versus star-forming galaxies (bottom panels).

At $z = 0$ we find that, while the slope of the MS changes only slightly when different averaging time-scales are adopted to determine SFR (different colours), its normalization shows a remarkable variation among them, being up to $\sim 0.1\text{--}0.15$ dex lower for 1000 Myr than for 10 Myr at $10^{9.5} M_{\odot}$.⁴ At $z = 2$, the offset of the MS for the shortest SFR time-scales is more than 0.2 dex higher with respect to the longest one ($10^9 M_{\odot} < M_{\text{stars}} < 10^{9.6} M_{\odot}$), the discrepancy being smaller (≤ 0.1 dex) at higher masses (top right panel).

In order to quantify the effects of different apertures (central panels), we measure the SFR of the whole galaxy – by summing the SFRs of all gravitationally bound star particles – (dashed curves), within $2 \times R_{\text{star,h}}$ (dotted curves) and within 5 pkpc (solid curves). Moreover, the median SFR is measured from stars formed in the last 10 (grey curves) and 1000 Myr (black curves), to bracket the variations identified in the upper panels. At $z = 0$ and $z = 2$ we find that, at fixed time-scales, the normalization of the MS can be up to $\sim 0.1\text{--}0.15$ dex lower for smaller apertures at the low-mass end. At $z = 0$ and for galaxies with $M_{\text{stars}} > 10^{9.5} M_{\odot}$, this difference rises up to ~ 0.3 dex whereas it remains ~ 0.1 dex for $z = 2$, regardless of galaxy stellar mass. In all cases, we find that the slope of the MS is not affected by different apertures. Interestingly, the differences between the SFRs within $2 \times R_{\text{star,h}}$ and within 5 pkpc is more pronounced at $z = 0$ ($\sim 0.05\text{--}0.1$ dex for both 10 and 1000 Myr) than at $z = 2$ (almost overlapping curves for both time-scales). These findings suggest that the star-formation regions of TNG galaxies extend beyond the nominal stellar mass or stellar light sizes, and more so for more massive galaxies

⁴The larger deviation at lower masses is, we think, an effect of numerical resolution: see Appendix A for details.

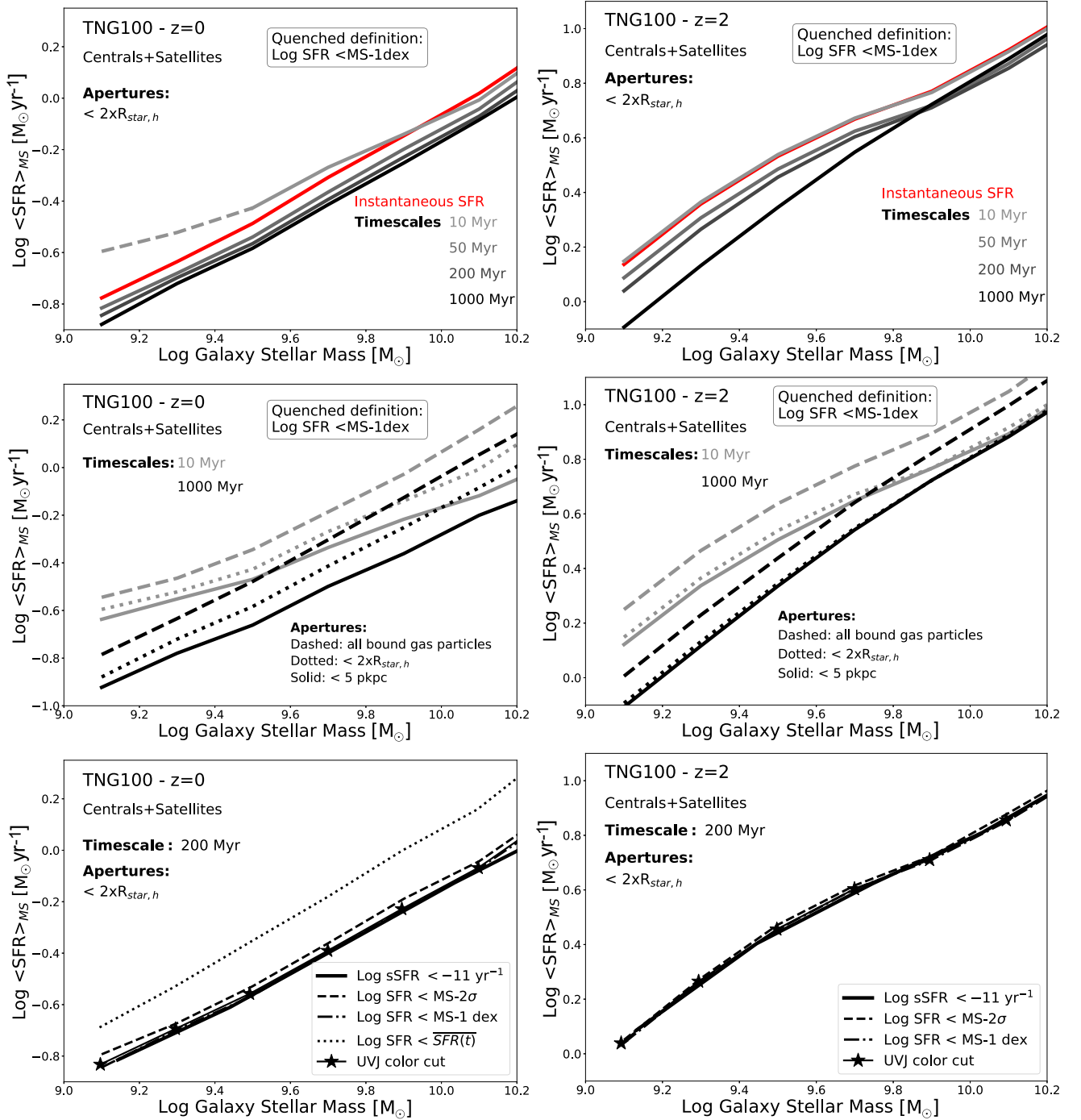


Figure 6. Median of the MS in TNG100 at $z=0$ (left-hand column) and $z=2$ (right-hand column). Here we consider different systematics, as explained in the following. Top panels: comparison among SFRs measured over different averaging time-scales (from 10 Myr, grey curves, to 1000 Myr, black curves) and the instantaneous one (red curves). Here quenched galaxies have been removed if their SFR is below 1 dex from the MS. The larger deviation at the high-mass end in the top right panel (dashed part of 10 Myr curve) is likely an effect of numerical resolution. Central panels: comparison across two SFR-averaging time-scales (10 Myr, grey curves, and 1000 Myr, black curves) and three different apertures: all gravitationally bound elements (dashed), within twice the stellar half-mass radius ($R_{\text{star},h}$, dotted), and within 5 pkpc (solid curves). Bottom panels: comparison among different classifications of quenched versus star-forming galaxies, as labelled in the legend.

at lower redshifts (see also Nelson et al. 2019; Pillepich et al. 2019).

Finally, we compare the MS measured by selecting star-forming galaxies with different quenched definitions (bottom panels). Here, we average the SFRs over 200 Myr, within an aperture of $2 \times$

$R_{\text{star},h}$. The MS exhibits an offset 0.1 dex higher for the $\overline{\text{SFR}(t)}$ -based selection in comparison to all other definitions (dotted curves in the bottom left panel). Albeit this difference is not negligible, we stress here that this selection based on the star formation histories of galaxies is purely theoretical, and we use it only at

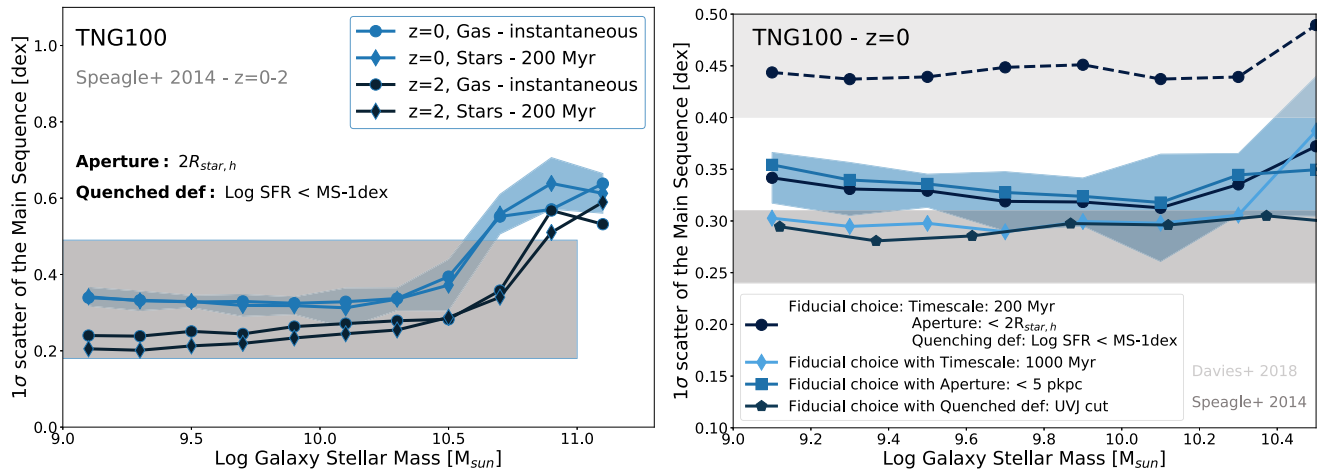


Figure 7. Left-hand panel: 1σ scatter of the MS of TNG100 galaxies as a function of galaxy stellar mass, at $z = 0$ (blue lines) and $z = 2$ (dark blue lines). The SFRs are measured from the instantaneous SFRs of the gas (filled circles) and from the stars formed over the last 200 Myr (filled diamonds). A selection of observational estimates of the intrinsic scatter (see table 4 of Speagle et al. 2014) is indicated in the grey box at $0 \leq z \leq 2$. Right-hand panel: scatter of the MS in the stellar mass range $10^9 < M_{\text{stars}} (M_{\odot}) < 10^{10.5}$. Once a fiducial choice (filled circles) is defined, we change one systematic at a time: time-scale (filled pentagons), aperture (filled squares), and quenched definition (filled diamonds). In addition, the scatter computed with the fiducial choice is shown in concert with the effect of 0.2 and 0.3 dex uncertainties on the stellar mass and SFR, respectively (dashed, dark blue curve). The shaded regions in both panels indicate the additional uncertainty in TNG100 due to cosmic variance, computed at $z = 0$ for our fiducial choices. The dark grey boxes indicate the observed scatter by Speagle et al. (2014) (dark grey) and by Davies et al. (2018) (light grey).

$z = 0$ to compare with the other quenched definitions and not in comparison to observations. On the other hand, the variations in the locus of the MS measure less than 0.05 dex when we consider the population-based criteria, thus suggesting that the MS in TNG100 is well captured even when using different classification criteria. This is even more evident at $z = 2$, when the MSs overlap one another regardless of the underlying adopted criterion (bottom right panel).

The results presented in this Section and visualized in Fig. 6 can be used as practical guidelines to quantify biases and systematics also across observational datasets with differently derived measures.

3.5 The scatter of the TNG main sequence

Finally, we conclude our quantification of the TNG star-formation activities by measuring the scatter of the main sequence as a function of galaxy stellar mass for TNG100 galaxies.

We follow the operational definitions given in equation (2) and Section 2.4. Here, we adopt the set of fiducial choices (see Section 2.5.1) where all galaxies with $\log \text{SFR} > \text{MS}-1$ dex are classified as star-forming and hence determine the width of the star-forming MS. In Fig. 7, the intrinsic scatter of the TNG MS is given as a function of galaxy stellar mass and we quantify the effects of the measurement choices by changing them one at a time.

In the stellar mass range $M_{\text{stars}} = 10^9 - 10^{10.5} M_{\odot}$, the scatter of the MS is constant with stellar mass and slightly decreases with increasing redshift, reading ~ 0.35 dex at $z = 0$ and ~ 0.2 dex at $z = 2$ (light and dark blue lines in the left-hand panel). For TNG300, these values read ~ 0.4 and ~ 0.25 dex, respectively. Moreover, the difference between the instantaneous and 200-Myr SFRs is negligible at low redshift (overlapping light blue curves) but becomes more significant at $z = 2$, albeit still smaller than 0.05 dex.

On the other hand, at the high-mass end, we find that the scatter significantly rises up to 0.6 dex for $M_{\text{stars}} \sim 10^{11} M_{\odot}$, likely due to an ill-defined main sequence in this mass range, as also outlined for the original Illustris simulation by Sparre et al. (2015).

The effect of different measurement choices are shown in the right-hand panel of Fig. 7. With our fiducial choices as a reference (filled circles with solid curves), we find that the scatter decreases by up ~ 0.05 dex for longer time-scales (filled diamonds) and it slightly increases if the SFR is measured within smaller apertures (filled squares). Selecting star-forming galaxies using a TNG UVJ cut instead of SFR-based cut, we find that the scatter is ~ 0.05 dex lower (filled pentagons) than our fiducial choice. Due to the limited simulation volume, we account for the cosmic variance by measuring the scatter (at $z = 0$ for our fiducial choices) in eight sub-boxes of ~ 50 Mpc on a side and computing the *jackknife* error (shaded areas in both panels).

All these are estimates for the *intrinsic* scatter. In Fig. 7, right-hand panel, a dashed black curve denotes the 1σ scatter when measurement uncertainties are taken into account: 0.2 dex uncertainties on the galaxy stellar mass and 0.3 dex on the SFRs. The effect of the measurement uncertainties necessarily turns into an increase of the scatter by about ~ 0.1 dex with respect the one computed with the raw data. Even if not shown, we find that the consequences of different choices quantified above hold also when mass and SFR uncertainties are taken into account, and in general all these trends are preserved and slightly enhanced at high redshifts.

Observational estimates place the intrinsic scatter in the range 0.2–0.5 dex (see e.g. Whitaker et al. 2012; Speagle et al. 2014; Schreiber et al. 2015; Davies et al. 2018; Pearson et al. 2018, and references therein), depending on the SFR measures and on the galaxy sample. For the sake of clarity, we stress here that for the comparison to Davies et al. (2018) (right-hand panel), since the authors use different methods to isolate star-forming and quiescent populations, we choose their scatter values measured using a $u - r$

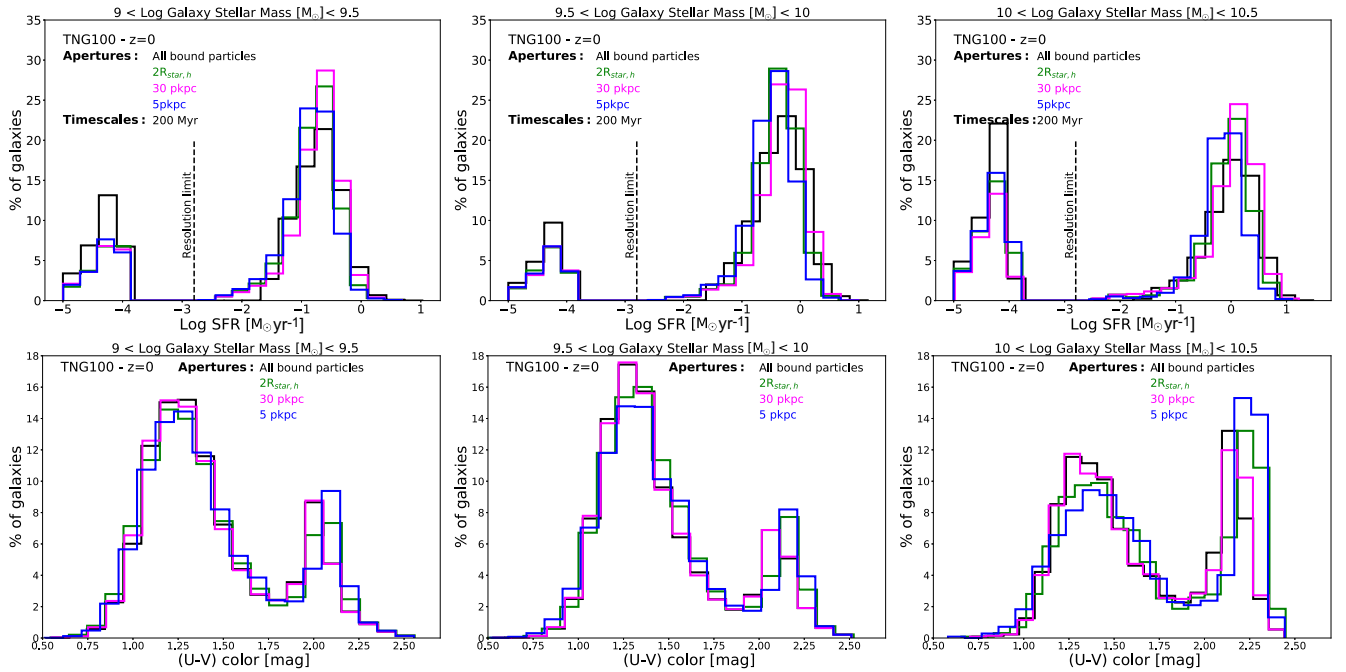


Figure 8. Distribution of the SFR (top panels) and $U - V$ colours (bottom panels) of TNG100 galaxies at $z = 0$ for three stellar mass bins, as labelled in each panel. For the SFR distributions, all galaxies with $\log \text{SFR}$ between -4 and $-5 \text{ M}_{\odot} \text{ yr}^{-1}$ are placed there arbitrarily, since they have unresolved SFR values and their actual distribution of SFRs is unknown (albeit firmly placed below the indicated resolution limit). In all panels, histograms are normalized to the total number in each stellar mass bins. A clear bimodality in colour is associated with a distribution in SFR that, if galaxies with unresolved values of SFR are ignored, rather than being bimodal in any sense, is unimodal and asymmetric.

colour selection and the selection based on the sSFRs. We find that, in the redshift range $0 \leq z \leq 2$ the scatter of the MS in TNG100 is in good agreement with observations, indicated as grey boxes in both panels of Fig. 7 but a more careful comparison once again requires a proper match or mocking of the adopted measurement choices.

4 DISCUSSION AND IMPLICATIONS

4.1 On the bimodality: colour versus SFR

As mentioned in Section 1, several galaxy properties are observed to be bimodal in narrow bins of galaxy stellar mass, including colour, SFR, morphology. As a result, galaxies are traditionally classified into two main categories: those with red colours, which are mostly quiescent and exhibit early-type, i.e. elliptical-like morphologies; on the other side, blue-cloud galaxies that are on the star-forming main sequence and with late-type, i.e. disc-like morphologies (Strateva et al. 2001; Kauffmann et al. 2003; Baldry et al. 2004; Balogh et al. 2004).

For example, Kauffmann et al. (2004) and Wetzel et al. (2012) have examined the number density distribution of SDSS galaxies in specific star-formation rate and found a break, a local minimum at $\text{sSFR} = 10^{-11} \text{ yr}^{-1}$ at $0.04 < z < 0.06$ whose value is independent of galaxy stellar mass and host halo mass (at least for stellar masses around $10^{10} \text{ M}_{\odot}$ and above). Interestingly, as pointed out by the authors themselves, albeit the rise of the distribution below the break is real, the strong peak near $\text{sSFR} = 10^{-12} \text{ yr}^{-1}$ is artificial, since all galaxies with undetectable SFR have been assigned arbitrary values in that range. Other authors over the years have questioned whether the galaxy distribution in SFR is effectively bimodal (e.g. Elbaz et al. 2007; McGee et al. 2011). Yet, the idea that low-redshift

galaxies are separable into two groupings of star-forming versus quiescent galaxies *because* their distribution of sSFR or SFR is bimodal in bins of stellar mass remains a foundational motif in galaxy evolution.

Recently, Feldmann (2017) has questioned the existence of a bimodality of the galaxy population in SFR and argued that, if one neglects ‘dead’ galaxies in numerical models (i.e. galaxies with unresolved SFR values), the logarithmic distribution of SFR is unimodal and that the observed bimodality is likely due to uncertainties in the SFR measurements. In the same spirit, Eales et al. (2017) have stressed again that the tight galaxy red sequence in the optical colour–magnitude diagram is the result of optical colours depending only very weakly on specific star-formation rate below a certain value. Therefore the red sequence is the result of the pile up at the same optical colour of galaxies with very low sSFR rather than the representation of a distinct class of galaxies. Even more fundamentally, Eales et al. (2018) have argued that submm and far-IR Herschel data favour an sSFR versus galaxy mass plane that is populated by a single ‘Galaxy Sequence’ that smoothly extends towards very low values of sSFR rather than by two distinct, well-separated classes of galaxies. They argue that the existence of a less densely populated region denoted the ‘green valley’ and residing between the star-forming and quenched populations is likely due to observational biases.

Here we show that numerical models like IllustrisTNG (and in fact also Illustris) naturally return a bimodal distribution in colour even though a clear bimodality in the logarithm of SFR or sSFR in bins of galaxy stellar mass is not present.

In Fig. 8, we show the distribution of the logarithmic SFR averaged over 200 Myr (top) and $U - V$ colour (bottom) of TNG100 galaxies at $z = 0$, stacked in three stellar mass bins, as labelled at the top of each panel. Histograms are normalized to the

total number of object in each stellar mass bins. In all cases, we show the distribution accounting for four different physical apertures, denoted by different colours. As already discussed in Section 2.5, small apertures shift the locus of the MS to lower values (see Fig. 6, central panels). Indeed, the top panels of Fig. 8 show that, while the tails of the distribution are not affected by different apertures, the peak – which marks the normalization of the MS – for 5 pkpc aperture is moved to lower values of the log SFRs (blue histograms). As done throughout, all galaxies with log SFR between -4 and $-5 M_{\odot} \text{ yr}^{-1}$ are placed there artificially, since they have unresolved SFR values, i.e. values of SFR that fall below the resolution limit of the simulation: the latter is indicated with vertical dashed lines in each panel. It is important to notice that such resolution limit is one or two orders of magnitude lower than observational ones.

Once the galaxies with unresolved SFR values are ignored, the distribution of log SFR is clearly unimodal, albeit asymmetric and skewed towards low SFR values. All these findings are consistent with the aforementioned arguments proposed by e.g. Feldmann (2017) and Eales et al. (2018). In fact, the SFR distributions of Fig. 8, top panels, could be characterized using a Gaussian mixture modelling, as recently proposed for e.g. the Illustris, EAGLE and Mufasa results by Hahn et al. (2018), but it is unclear whether the medians of such multiple Gaussians would identify distinct classes of galaxies or simply represent a numerical overfitting. In order to quantify the asymmetry, we fit the log SFR distribution (above the resolution limit) with a zero-inflated gamma distribution proposed by Feldmann (2017). We find that the asymmetric log SFR distribution of Fig. 8 is well characterized by this description, more reliably than those based on the standard log-normality assumption.

The lower panels of Fig. 8 extend the analysis on TNG presented by Nelson et al. (2018a), who characterized the bimodal distribution of $(g - r)$ colours of TNG galaxies, in the stellar mass range $10^9 < M_{\text{stars}} (M_{\odot}) < 10^{12.5}$. There, a quantitative comparison to SDSS data has shown TNG galaxies to be in excellent agreement with $z \sim 0$ observations, providing an empirical indirect validation to the underlying choices for AGN and stellar feedback adopted in TNG in comparison to the original Illustris. Here, in addition to such previous analysis, we explicitly examine the $(U - V)$ colour distribution of TNG galaxies in three bins of stellar mass (bottom panels of Fig. 8). In the previous Sections, we have demonstrated that the UVJ diagram of TNG galaxies is populated by two groups of galaxies that in terms of number density and SFR values clearly separate on the UVJ plane. Consistent with this, a pronounced bimodality in $(U - V)$ at fixed galaxy stellar mass is apparent, with a depression around 1.75 mag for low-mass galaxies between 10^9 and $10^{9.5} M_{\odot}$ and at slightly redder colours for the more massive ones. This again confirms that a strong colour bimodality may coexist with a unimodal distribution of SFRs.

Even if not shown, we have repeated the bottom panel of Fig. 8, by excluding galaxies with unresolved SFR values from the sample: in this case, the colour bimodality is not in place. This could be interpreted similarly to what have been seen for the SFR distributions, that look like bimodal only when very low SFR galaxies are accounted before (i.e. by accumulating them at some arbitrary low SFR value). In fact, this result enhances and reinforces what stated above: colour distributions can be bimodal even if SFR distributions are not. Galaxies with $\text{SFR} = 0$ would distribute below the resolution limit but not necessarily creating a bimodality. Instead, colours are bimodal whenever well defined and realistic colour values are considered. In the simulations, the galaxy colours are well captured also for those galaxies whose SFR values are below the resolution limit. Even though a fraction of galaxies have artificially null SFR

values, the same galaxies have well defined and observationally consistent $(g - r)$ and $(U - V)$ colours, making the colour bimodality of Fig. 8 robust.

Finally, although we do not show it, we have examined the colour distribution of TNG100 galaxies also at higher redshifts, e.g. $z = 1$ and $z = 2$. Differently from the bottom panels of Fig. 8, where the colour bimodality is manifest for low-mass as well as high-mass galaxies, the colour bimodality is still present at higher redshifts but only at higher galaxy stellar masses. Specifically, at $z = 1$, two peaks are clearly visible for galaxies more massive than $10^{10} M_{\odot}$ whereas at $z = 2$ such bimodality starts to be apparent only for galaxies with stellar mass $> 10^{10.5} M_{\odot}$.

4.2 On the bending of the MS at the high-mass end

As demonstrated in this work and in several previous observational and theoretical studies, the MS of star-forming galaxies with stellar mass $M_{\text{stars}} \lesssim 10^{10} M_{\odot}$ follows a single power law, whose slope and normalization values depend somewhat on the adopted IMF, galaxy sample selection, and SFR tracer (see Speagle et al. 2014, and references therein).

On the other hand, at higher galaxy masses, several authors have observed a more complex scenario. From the local Universe to $z \sim 4$, some studies indicate the presence of a turnover, namely of a downwardly bending of the $\text{SFR} - M_{\text{stars}}$ relation for galaxies with $M_{\text{stars}} \gtrsim 10^{10} - 10^{10.5} M_{\odot}$, hence proposing a stellar mass-dependent slope as best fit instead of a single power law (Karim et al. 2011; Whitaker et al. 2012, 2014; Gavazzi et al. 2015; Lee et al. 2015, 2018; Tomczak et al. 2016). Whether the star-forming MS bends or not, and by how much, is still a matter of debate in the observational literature. Moreover, its physical reason and the existence of a turnover mass are still unclear. Morphological studies of galaxies have been proposed to explain the bending (Abramson et al. 2014; Schawinski et al. 2014), albeit the most likely scenario is related to secular internal processes, like AGN feedback, that drive the quenching of massive galaxies.

As shown in Section 3.1 and Fig. 5, the main sequence of TNG galaxies exhibits a bending if the classification of galaxies into star-forming and quiescent is done by means of a UVJ cut. There we found that the bending is somewhat more pronounced at lower redshifts and that the turnover occurs at large stellar masses ($M_{\text{stars}} \gtrsim 10^{10.5} M_{\odot}$), clearly revealing that massive galaxies classified as star-forming through a UVJ cut have lower SFRs in comparison to lower mass counterparts.

We now show that, given the same galaxy population, the star-forming MS may bend or not according to the criterion adopted to separate star-forming versus quiescent galaxies. Fig. 9 shows the median of the MS of TNG300 galaxies in the stellar mass range $M_{\text{stars}} = 10^{9.5} - 10^{11} M_{\odot}$ at $z = 0, 0.75, 1.75$. Star-forming galaxies are selected in three ways: according to their colours by using the UVJ cut proposed by Whitaker et al. (2011) (black curves), and the cut based on the TNG diagram itself (equations 3 and 4: red curves), and according to their SFR values, $\log \text{SFR} > \text{MS}-1$ dex (blue curves), without extrapolating the linear MS from lower masses. Additionally, we show the linear extrapolation of the MS (dashed curves), for reference.

For galaxies more massive than $\sim 10^{10.5} M_{\odot}$ a bending of the MS is clearly noticeable if star-forming galaxies are identified via the SFR-based criterion and via the TNG UVJ cut. While a steep decline of the MS for massive galaxies is manifest at all redshifts, such a turnover is less prominent in the case of the TNG UVJ cut when compared to the SFR-based selection. In fact, when we adopt

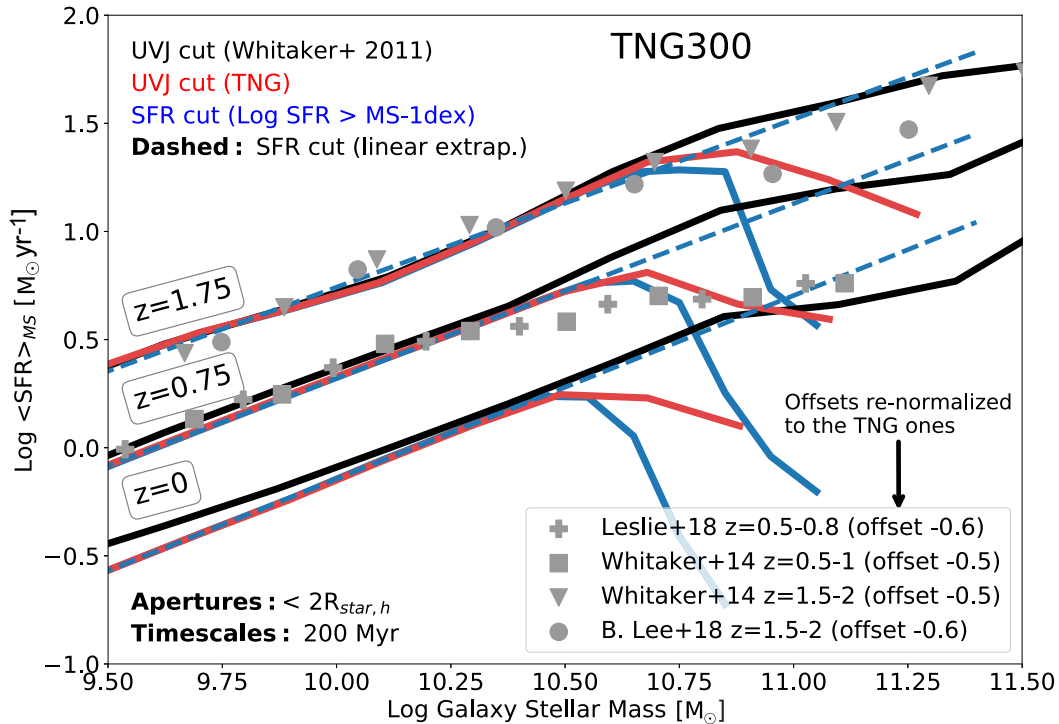


Figure 9. Median SFR on the MS of TNG300 galaxies at $z = 0, 0.75, 1.75$. The aperture is $2 \times R_{\text{star},h}$ and the galaxy SFR is averaged over 200 Myr. Star-forming galaxies are identified in different ways: with the *UVJ* selection proposed by Whitaker et al. (2011) (solid black lines), with the TNG *UVJ* selection (solid red lines), and with an SFR cut using $\log \text{SFR} > \text{MS}-1$ dex (dashed blue lines). Additionally, the linear extrapolation from low masses of the MS obtained via the SFR cut is shown (dotted line), for reference. Grey symbols are observational measurements, arbitrarily normalized to the TNG ones (as indicated in the legend), to highlight the turnover of the MS. Given the same TNG300 galaxy sample, the star-forming MS may or may not bend according to the criterion adopted to separate star-forming versus quiescent galaxies.

instead the *UVJ* cut defined in Whitaker et al. (2011), we find that the turnover of the MS is much less pronounced than the other two methods, even though a deviation from the linear trend is recovered also in this case at the highest mass end.

For comparison, a selection of observational data points at $z = 0.75$ and $z = 1.75$ is included as grey symbols. They all adopt a *UVJ* cut to select for star-forming galaxies but for Leslie et al. (in preparation), who use $NUV - r - J$ selection. Here the normalization of the observationally derived MSs is artificially adjusted to match the one from TNG, in order to facilitate the comparison of the bending. In the linear, low-mass regime, irrespective of the adopted selection criteria, the slope of the MS in TNG is in good agreement with the observational data at both $z = 0.75$ and $z = 1.75$. Moreover, the TNG *UVJ* cut (red curve) reproduces quite well some of the observational results in terms of bending, at least at $z = 0.75$. However, the level of bending at $z \sim 1.75$ is also unclear across observational datasets, while we find that the TNG model returns a weak trend with redshift of the turnover mass, that moves from $M_{\text{stars}} \simeq 10^{10.5} M_{\odot}$ at $z = 0$ to about $10^{10.7} M_{\odot}$ at $z = 1.75$.

In conclusion, the characterization of the shape of the MS at the highest mass end depends on the criterion adopted to separate between star-forming and quiescent galaxies. This is the case at least for TNG galaxies but could also apply to the observed galaxy population. *UVJ* selections that minimize the number of quenched galaxies (and hence classify a larger number of galaxies as star forming: e.g. the TNG *UVJ* cut in Fig. 1 in comparison to the Whitaker ones) are destined to return an MS that is more bending than those that are more conservative in classifying galaxies as star

forming. In the case of TNG galaxies, this ambiguity is strictly connected to the way galaxies occupy the $\text{SFR}-M_{\text{stars}}$ plane at the high-mass end.⁵ As can be seen in Figs 2, 4, and 8, TNG galaxies populate the low-SFR regions below the high-density star-forming MS without producing an obvious depression or separation or green valley: see previous section and the absence of a clear galaxy bimodality in SFR in TNG. This in turn makes the classification of galaxies into two classes somewhat arbitrary, albeit operationally still viable, and the characterization of the star-forming main sequence at the highest mass end somewhat ill defined. We postpone to future work the task to connect the shape of the $\text{SFR}-M_{\text{stars}}$ plane to the time-scales of the underlying quenching mechanisms in the TNG model.

4.3 Comparison to observations

As demonstrated in Section 2.5, different measurement choices can influence to various degrees the locus of the star-forming MS. Now, with this awareness, we attempt to compare the MS of TNG100 galaxies to selected observational works.

As noticeable from Table 4, different observational analyses adopt different methods or indicators to derive the SFRs of observed galaxies ($H\alpha$, UV, FUV, IR, etc.), different IMFs and, unavoidably different implicitly adopted physical apertures within the luminous

⁵For the sake of clarity, we stress here that even if not shown, we have computed the main sequence also by removing galaxies with unresolved SFR values. Our findings suggest that neither the locus of the ‘linear’ MS nor the bending is quantitatively affected by this choice.

Table 4. Summary of the observational datasets adopted for comparisons in Figs 9 and 10. Columns read: (1) reference; (2) redshift range; (3) SFR indicator; (4) assumed stellar initial mass function; (5) criterion used to separate quenched from star-forming galaxies; (6) shape of the MSs; (7) mass range observed in each sample.

References	Redshift	SFR tracers	IMF	Galaxies selection criteria	Main-sequence shape	$\log M_{\text{stars}} (M_{\odot})$ range
Salim et al. (2007)	0.005–0.22	UV-SED	Chabrier (2003)	$NUV - r$	Linear	9–11.1
Elbaz et al. (2007)	0.015–0.1	$H\alpha$	Salpeter (1955)	$U - g - M_B$	Linear	9.1–11.3
Zahid et al. (2012)	0.04–0.1	$H\alpha$	Chabrier (2003)	?	Linear	8.5–10.4
Oliver et al. (2010)	0.0–0.2	IR	Salpeter (1955)	Optical/NIR SED	Linear	9.1–11.6
Chang et al. (2015)	0.0–0.2	12–22 μm	Chabrier (2003)	$u - r - z$	Linear	9–11.1
Karim et al. (2011)	0.6–0.8	Radio	Chabrier (2003)	$NUV - r - J$	Linear	9.1–11.1
Whitaker et al. (2014)	0.5–1	UV + IR	Chabrier (2003)	UVJ	Bending	8.4–11.1
Leslie et al. (in preparation)	0.5–1	Radio-IR	Chabrier (2003)	$NUV - r - J$	Bending	8.8–11.2
Lee et al. (2015)	0.8–1	IR	Chabrier (2003)	$NUV - r - J$	Bending	8.5–11
Leslie et al. (in preparation)	0.8–1.1	Radio-IR	Chabrier (2003)	$NUV - r - J$	Bending	8.8–11.2
Lee et al. (2018)	1.5–2	FUV + IR	Chabrier (2003)	UVJ	Bending	9–11.5
Whitaker et al. (2014)	1.5–2	UV + IR	Chabrier (2003)	UVJ	Bending	9.2–11.5
Santini et al. (2009)	1.5–2.5	UV + IR	Salpeter (1955)	?	Linear	8.7–11.5
Rodighiero et al. (2011)	1.5–2.5	FUV	Salpeter (1955)	BzK	Linear	8.7–11.3
Pannella et al. (2009)	1–3	Radio	Salpeter (1955)	?	Linear	10.1–11.2
Whitaker et al. (2014)	2–2.5	UV + IR	Chabrier (2003)	UVJ	Bending	8.4–11.1
Tomczak et al. (2016)	2–2.5	UV + IR	Chabrier (2003)	UVJ	Bending	8.5–11.5
Speagle et al. (2014)	0, 0.75, 1.75, 2	Multiple tracers	Kroupa (2001)	Multiple criteria	Linear	9.7–11.1

bodies of galaxies. As discussed at length throughout the paper, such a diversity of methodologies can lead to non-negligible apparent discrepancies in the quantification of the normalization of the MS (see e.g. Fig. 6 and Speagle et al. 2014; Theios et al. 2018, and references therein). At $z = 0$, different observed MS normalizations taken at face value differ by up to ~ 0.5 – 0.7 dex; this is the case also at $z = 2$. Until a robust understanding of the observational systematics is available, we would be tempted to treat the comparison between observations and theoretical models by only focusing on the slope and mass trends of the MS and by ‘arbitrarily’ matching its normalization, as we have done in Fig. 9 and has been done with the evolution of the TNG galaxy mass–metallicity relation by e.g. Torrey et al. (2018). Indeed, an insightful comparison between simulated and observational results demands a deep understanding of how different observational SFR indicators compare to one another and effectively map into un-biased estimates of the SFRs of galaxies as directly available from simulations. In fact, here we can only proceed with a face-value comparison, by which we mean that we keep unchanged the observational data points and only adjust them to be consistent with the Chabrier IMF adopted in the simulations ($\log M_{\text{stars}}^{\text{Chabrier}} = \log M_{\text{stars}}^{\text{Salpeter}} - 0.24$; $\log \text{SFR}^{\text{Chabrier}} = \log \text{SFR}^{\text{Salpeter}} - 0.15$, see e.g. Davé 2008; Santini et al. 2014). We focus on a selection of observational dataset based on commonly available redshift ranges and median redshifts.

In Fig. 10, we show the TNG100 MS in the stellar-mass range $M_{\text{stars}} = 10^9$ – $10^{10.5} M_{\odot}$. The SFRs are averaged over 200 Myr (dark blue curves), 50 Myr (blue) and 10 Myr (light blue), and measured within $2 \times R_{\text{star,h}}$ (solid) or by considering all the star particles gravitationally bound to the galaxy (dashed). All these variations are included to bracket the largest uncertainties due to the possible measurement choices. Star-forming galaxies are selected using the colour cut based on the UVJ diagram of TNG100 galaxies discussed in Section 2. Each panel denotes a different redshift: $z = 2$ (top),

$z = 1.75$ (top central), $z = 0.75$ (bottom central), and $z = 0$ (bottom). The selection of observationally derived MS is indicated in grey symbols.

At first glance, a qualitative and quantitative agreement in the slope of the MSs between TNG and observed data is manifest. Indeed, by fitting equation (1), we measured the MS slope of TNG galaxies in the range $\alpha_{\text{TNG}} \sim 0.7$ – 0.8 (see Table 3). This settles within the observational range, $\alpha_{\text{OBS}} \sim 0.6$ – 1 , independent of redshift.

On the other hand, statements concerning the normalization of the star-forming MS and its comparison to observational data depend significantly on redshift. At $z = 0$, the MS of TNG galaxies lies in the ballpark of the observations, falling in between Oliver et al. (2010) and Zahid et al. (2012). At intermediate redshifts – $z = 0.75$ and $z = 1.75$ –, on the other hand, the TNG normalization is lower than any available observational result, by up to 0.5 dex and by at least 0.2 dex, even considering different measurement choices. This discrepancy holds even at higher redshift ($z = 2$, top panel), although to a lesser degree: the TNG normalization is ~ 0.25 dex lower than the majority of the selected datasets. By allowing for larger apertures and smaller averaging time-scales (light blue dashed curves), the TNG MS becomes progressively more consistent with the results by Santini et al. (2009) and Pannella et al. (2009), in which galaxy SFRs are derived from UV + IR and from radio, respectively.

We stress here that the discrepancies between the TNG outcome and observations at intermediate redshifts persists even when the calibrated data from Speagle et al. (2014) are considered – the latter being a recompilation of many observational results all converted to a common absolute calibration to overcome the limitations of the different adopted star-formation indicators. For reference, we include these data as black stars in Fig. 10, by using their best MS fit (equation 28 of Speagle et al. 2014) and, as for the other observational data, by adjusting the IMF to be con-

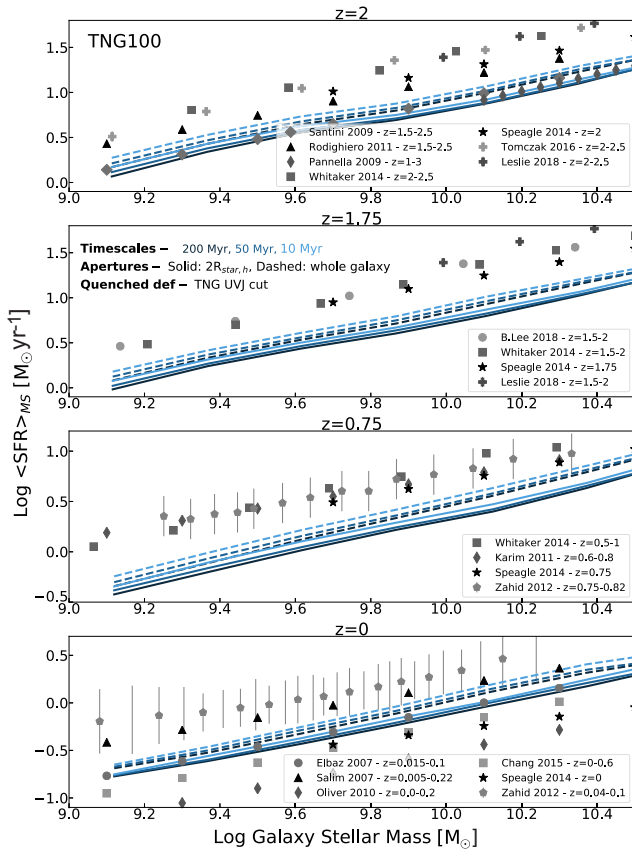


Figure 10. Star-forming MS in TNG100 at $z = 0, 0.75, 1.75, 2$ (from bottom to top panels). The SFR is measured within $2 \times R_{\text{star,h}}$ (solid lines) and by summing the individual SFR of all stars gravitationally bound to the galaxy (dashed lines). The shades of blue curves denote three different timescales: 200 Myr (dark blue), 50 Myr (blue), and 10 Myr (light blue). The star-forming galaxies are selected using the TNG UVJ cut, as explained in Section 2.5. A selection of observational data is indicated in grey symbols.

sistent with the Chabrier one ($\log M_{\text{stars}}^{\text{Chabrier}} = \log M_{\text{stars}}^{\text{Kroupa}} - 0.025$; $\log \text{SFR}_{\text{Chabrier}} = \log \text{SFR}_{\text{Kroupa}} - 0.15$).

In summary, despite a reasonable consistency between the TNG and observed MS at $z = 0$ and despite controlling for a number of measurement choices or systematics, we find non-negligible tension at intermediate redshifts ($0.75 \leq z < 2$) in terms of normalization of the MS, with the simulated one being lower than observed. These at-face-value findings are in agreement with what outlined also for the original Illustris model (see e.g. Genel et al. 2014; Sparre et al. 2015; Davidzon et al. 2018), even though the Illustris’ MS is higher than the TNG ones, and by several other works that compared the observed MS with theoretical models of galaxy formation (Daddi et al. 2007; Davé 2008; Somerville & Davé 2015). This could point towards some fundamental limitations of theoretical models. However, it is also important to keep in mind that, in observations, possible contamination from non-star-forming sources may have a notable impact on the derivation of the SFRs of galaxies, biasing high their estimates. For instance, the SFR of a galaxy inferred using $H\alpha$ emission can be contaminated by non-star-forming sources. Indeed, this emission is due to ionized hydrogen gas and its ionization could be the result of different processes and not only due to the star formation regions. Moreover, in SDSS, the largest source of uncertainty in evaluating the galaxies

SFR within the fibre is due to the possible contamination of their spectra by other sources of ionizing radiation (Brinchmann et al. 2004). For the emission in the IR band, the major contamination is thought to be the contribution of synchrotron radiation to the long-wavelength thermal dust emission: if not well-constrained, the latter can lead to an overestimation of the galaxies’ SFR (Archibald et al. 2001; Falkendal et al. 2019). In addition, galaxy SFRs can also be biased high by dust-heating from older stellar populations (see e.g. Bendo et al. 2010; Groves et al. 2012). Finally, in the UV band, especially for systems with very low SFRs, another source of contamination is from old post-AGB stars and the derived SFR should be intended only as an upper limit (Kennicutt & Schweizer 1998).

5 SUMMARY AND CONCLUSIONS

A tight relation between the star-formation rate and the stellar mass of star-forming galaxies has been widely observed in the past decade, known as the star-forming *main sequence*. However, debates about how it evolves in time, what drives its shape, about the existence of a downwardly bending at the high-mass end, and whether two well-distinct classes of galaxies effectively exist based on their SFRs (star-forming versus quiescent) are still open.

In this paper, we have used the two highest resolution simulations currently available from the IllustrisTNG project (TNG100 and TNG300) to verify whether the underlying galaxy-physics hydrodynamical model naturally returns such a fundamentally observed property of the galaxy population as the relationship between star-formation activity and stellar mass. In particular, we have selected all galaxies with $M_{\text{stars}} > 10^9 M_{\odot}$, sampling tens of thousands of objects between satellites and centrals, and we have quantified the SFR– M_{stars} plane, its star-forming MS, the fraction of quenched galaxies, and the UVJ diagram of the TNG simulated galaxies from $z = 0$ to $z \sim 2$.

Particular focus has been placed in quantifying some of the most notable measurement choices, i.e. systematics that affect the derivation of star-formation rates from various tracers, and hence possibly bias the comparison between theoretical and observational results. We have explored a number of ways to separate between star-forming and quiescent galaxies: colour–colour cuts in the UVJ diagram as well as selections based on the position of individual galaxies on the SFR– M_{stars} plane (see Table 2). We have measured the SFR of galaxies based on the instantaneous SFR of their gas cells as well as from stars formed over the last 10, 50, 200, and 1000 Myr, in order to connect with the time-scales probed by observational star-formation indicators. Moreover, we have accounted for the star formation occurring within different portions of the stellar bodies of galaxies: within 5 physical kpc, 30 physical kpc, within twice the stellar half-mass radius and throughout the whole gravitationally bound matter: see Section 2.5 for details. We have therefore attempted some ‘face-value’ comparisons with observational measurements.

Our main findings can be summarized as follows:

(i) The TNG model reproduces the general qualitative features of the observed UVJ diagram at both $z = 0$ and 2 (Fig. 1), with a tight clump of low-SFR galaxies residing in the top region at redder ($U - V$) colours and a broader cloud of star-forming galaxies that populate the bluer ($U - V$) colours. A low-density region, where fewer galaxies lie, separates the two populations, thus allowing us to adopt a colour–colour cut based on the number density distribution of TNG galaxies on the UVJ plane (equations 3–4). In fact, also

observationally proposed *UVJ* cuts (e.g. by Whitaker et al. 2011 and Williams et al. 2009) reasonably intersect this depression, providing a validation of the TNG model in comparison to observations (see Section 3.1).

(ii) TNG galaxies populate the $\text{SFR}-M_{\text{stars}}$ plane also in ways that are qualitatively consistent with observations: a dense region exists, where the SFR of galaxies tightly correlates with their stellar mass in the range $M_{\text{stars}} = 10^9-10^{10.5} M_{\odot}$. Such a well-defined star-forming main sequence is present at both low ($z = 0$) and high ($z = 2$) redshifts (see Fig. 4) and independently from the method adopted to measure the SFRs (see Appendix A and Fig. A1).

(iii) A number of galaxies seem to *fall off* from the star-forming MS, by asymmetrically and continuously broadening it towards lower values of SFR. However, because of the finite numerical resolution of the simulations, galaxy SFRs are resolved only down to a minimum value, depending on the averaging time-scales. For example, for 1000 Myr, the minimum resolved SFR value reads about $10^{-2.6} M_{\odot} \text{yr}^{-1}$ for TNG300 and $10^{-3.2} M_{\odot} \text{yr}^{-1}$ for TNG100, see Appendix A. Below this limit, all simulated galaxies are labelled with $\text{SFR} \equiv 0$. Such ‘completely quenched’ galaxies are distinctly detached from the star-forming MS but feature as a distinct class of objects on the simulated $\text{SFR}-M_{\text{stars}}$ only because of the simulation resolution limit.

(iv) The relative fractions of quenched galaxies depend on stellar mass, redshift, and operational definition (see Fig. 3). The fraction of quenched galaxies in general increases from high to low redshifts, and more so at larger masses. At $z = 0$, the TNG model predicts about 80 per cent of galaxies in the $10^{10.5} \leq M_{\text{stars}}/M_{\odot} \leq 10^{11}$ range to be quiescent. This is largely due to the adopted AGN feedback. However, this estimate can vary anywhere between 70 and 85 per cent according to the operational definition of what constitutes quenched galaxies. Interestingly, within the TNG simulations, a properly identified cut in the *UVJ* diagram returns quenched fractions in the redshift range $0 \leq z \leq 2$ that are in agreement at the percentage level with those obtained by labelling quenched those galaxies whose SFR is 1 dex below the locus of the linearly extrapolated MS (see Section 3.2 and Figs 2 and 3, top panel).

(v) At $z = 1.75$, the TNG model returns a fraction of massive quenched galaxies of ≈ 65 per cent ($M_{\text{stars}} = 10^{11}-10^{11.5} M_{\odot}$). In general, we obtain a larger quenched fractions at progressively larger masses and lower redshifts. However, effects of resolution are not negligible in the resulting quenched fractions: at $z = 0$, TNG300 returns a fractions of quenched galaxies of about 10–20 per cent larger than TNG100, for the whole stellar mass range; at higher redshift ($z = 0.75$) the resolution effect is apparent for galaxies more massive than $10^{10.5} M_{\odot}$, being ~ 20 per cent larger for TNG300 with respect to TNG100. In the same mass regime, at $z = 1.75$, the discrepancy between the models is less than 10 per cent. Despite such differences between TNG100 and TNG300, the TNG model returns systematically larger quenched fractions than the original Illustris and other models: this brings the TNG galaxies overall to a better agreement with observational measurements (see Section 3.2 and Fig. 3, bottom panels)

(vi) The TNG model qualitatively reproduces the salient features of the observed star-forming MS: in the mass range $M_{\text{stars}} = 10^9-10^{10.5} M_{\odot}$, the MS can be described by a single power law (equation 1); the normalization increases with redshift by ~ 0.6 dex from $z = 0$ to $z = 1$ and by ~ 0.4 dex from $z = 1$ to $z = 2$, independent of galaxy stellar mass, while the MS slope appears to be quite constant across cosmic times. At the high-mass end, if a *UVJ* cut is adopted to identify star-forming galaxies, the TNG

MS exhibits a bending, with a shallower slope for galaxies with $M_{\text{stars}} > 10^{10.5} M_{\odot}$ (see Section 3.3 and Fig. 5, top panel).

(vii) Differently from the quenched fractions, a comparison between the output of the TNG100 and TNG300 simulations reveals instead an excellent agreement between the two runs in terms of locus and shape of the star-forming MS, despite the different mass resolutions and sizes of the simulated volumes. On the other hand, the MS of Illustris galaxies exhibits a steeper slope with respect to TNG and a normalization that is ~ 0.1 dex higher than that of TNG galaxies with $M_{\text{stars}} > 10^{10} M_{\odot}$. This scenario is related to the different recipes for feedbacks implemented in the two simulation projects, being less efficient at regulating or even halting the star formation in massive galaxies in Illustris than in TNG (see Section 3.3 and Fig. 5, bottom panels).

(viii) The precise locus of the MS depends on the adopted measurement choices. Shorter time-scales for the averaging of the SFR estimates return higher normalizations: the MS normalization obtained by averaging the SFRs over 10 and 50 Myr (that can be associated with UV, $\text{H}\alpha$, and $\text{H}\beta$ observational tracers) is found to be ~ 0.1 dex (~ 0.2 dex) at $z = 0$ (at $z = 2$) higher than that obtained by measuring SFRs over longer time-scales like 200 and 1000 Myr (comparable to those of radio and IR indicators): see Fig. 6, top panels. Smaller physical apertures imply systematically lower SFRs and a lower normalization of the MS, with discrepancies by up to 0.3 dex between e.g. the 5 pkpc aperture and larger ones for massive galaxies at low redshifts (Fig. 6, central panels). Different methods to classify star-forming versus quiescent galaxies, on the other hand, lead to negligible variations in the MS offset, thus suggesting that the MS of TNG galaxies is well captured independently from the adopted selection criterion (see Section 3.4 and Fig. 6, bottom panels).

(ix) The intrinsic scatter of the star-forming MS is constant with galaxy stellar mass ($\lesssim 10^{10.5} M_{\odot}$) and decreases with increasing redshift, reading ~ 0.35 dex at $z = 0$ and ~ 0.2 dex at $z = 2$ in our fiducial choices (Fig. 7, left-hand panel). Longer averaging time-scales lead to smaller levels of scatter (e.g. 0.30 dex versus 0.35 dex for 1000 Myr versus 200 Myr time-scales), while the effects of apertures are negligible and the definition of quenched galaxies affects the quantitative assessment of the MS scatter only at the high-mass end (Fig. 7, right-hand panel). By accounting for measurement uncertainties in stellar mass (~ 0.2 dex) and SFR (~ 0.3 dex), the MS scatter is inflated by an additional 0.1 dex, making the width of the simulated MS overall consistent with observational findings (see Section 3.5).

(x) The TNG model predicts a clear and strong bimodality in the colour distributions of galaxies that is visible in both the ($U - V$) distributions in bins of galaxy stellar mass (see Section 4.1 and Fig. 8, bottom) and in the morphology of the *UVJ* diagram (Fig. 1), at both $z = 0$ and $z = 2$. However, such a colour bimodality is in place even though a bimodality in the log SFR distributions is much less obvious, if not non-existent. Once galaxies with unresolved SFR values are neglected, the logarithmic SFR distributions at $\text{SFR} \gtrsim 10^{-2} M_{\odot} \text{yr}^{-1}$ at $z = 0$ can be well described as unimodal, with a peak at the star-forming MS and a long asymmetric tail extending towards lower SFR values (see Section 4.1 and Fig. 8, top).

(xi) Within our simulated samples of galaxies, whether the MS bends or not at the high-mass end, and by how much, depends on the criterion adopted to separate star-forming versus quiescent galaxies. This is related to the shape of the $\text{SFR}-M_{\text{stars}}$ plane in the TNG simulations, where a clear well-defined MS ceases to be in place at stellar masses in the range of $10^{10.7-11.5} M_{\odot}$. When star-forming galaxies are selected based on their SFR values (e.g. $\text{SFR} >$

MS-1 dex, where the MS is *not* linearly extrapolated), the TNG MS exhibits a well-defined turnover mass: $10^{10.5}-10^{10.8} M_{\odot}$, depending on the redshift. TNG *UVJ* selections also return a downward bending MS, albeit less pronounced than the SFR-based selection. In both cases, we recover a weak trend of larger turnover masses at larger redshifts and overall a broad agreement with observed bending (see Section 4.2 and Fig. 9).

(xii) However, when we compare the TNG results on the MS with observational estimates taken at face value, we find that while the TNG model reproduces, qualitatively and quantitatively, the observed slope of the MS, the redshift evolution of its normalization in the stellar mass range $10^9-10^{10.5} M_{\odot}$ is in tension with observations. While the simulated MS falls in the ballpark of $z = 0$ observations, its normalization is systematically lower than observations imply at intermediate redshifts ($0.75 \leq z \lesssim 2$), by up to 0.2–0.5 dex and despite allowing for a number of measurement choices and systematics (see Section 4.3 and Fig. 10).

In conclusion, in this work we have demonstrated that the TNG model naturally returns the qualitative basic features of the *UVJ* and SFR– M_{stars} planes and the existence of a star-forming MS and of two classes of galaxies (star-forming versus quiescent) that separate clearly at least in their colour distributions. Importantly, we have demonstrated that the TNG model is capable of quenching massive galaxies at intermediate redshifts ($z \sim 1-2$) in fractions that are in better agreement with observational findings than previous numerical model. At the same time, and most probably for different reasons, an apparent underestimation of the locus of the star-forming main sequence is still in place in the TNG model at $0.75 \leq z \lesssim 2$ in comparison to observations. We have quantified the effects of a number of measurement choices in relation to SFR time-scales, galaxy apertures, and quenched definitions, and hence provided guidelines for the systematic biases that may affect comparisons across datasets, whether simulated or observed. Such systematics can affect the locus of the star-forming MS by up to 0.2–0.3 dex. For example, the fact that smaller physical apertures shift the locus of the MS towards lower values is an indication that star formation can occur in non-negligible amounts also in regions that extend far beyond the main bright body of galaxies.

The findings outlined in this paper in relation to the comparison to observations might indicate the following learning point: the TNG model captures reasonably well the effects of the dominant quenching mechanism at the high-mass end (i.e. the AGN feedback) but might require important adjustments in terms of the feedback mechanisms that regulate (without necessarily halting) star formation in $\gtrsim 10^9 M_{\odot}$ galaxies at the cosmic noon (i.e. stellar feedback). However, two important conceptual caveats remain before conclusive messages can be derived. First, in this paper we have not ventured into a thorough discussion of the biases that may affect observational estimates of the SFRs of galaxies. From the theoretical view point, the SFRs we can measure from the simulated galaxies are unbiased, while non-negligible uncertainties affect the mapping between observed light of certain star-formation indicators and SFRs, because of contamination or dust reddening. For example, Theios et al. (2018) have added to a vast body of literature and argued that observed SFR estimates can be afflicted by an additional 0.5 dex of systematic uncertainties, making previously published MS relations possibly biased high (at the intermediate redshifts of interest here). Sorba & Sawicki (2018) showed that observational indicators of galaxy stellar mass may be biased, causing a systematic effect on the SSFR. Just recently, Leja et al. (2018) have re-inferred the SFRs and galaxy stellar masses of 3D-

HST galaxies at $0.5 < z < 2.5$ taking advantage of a new high-dimensional model and have shown that, among other findings, the newly derived galaxy SFRs can be 0.1–1 dex lower than the UV + IR-based ones. In fact, it is also important to notice that, while the SFRs of galaxies on the TNG MS appear low around e.g. $z \sim 1$, a zeroth-order comparison of the TNG galaxy stellar mass function to observational findings does not raise any obvious flag: see fig. 14 of Pillepich et al. (2018b). The one fundamental advantage in simulations is their self-consistency across times: the galaxy mass functions at certain times are the result of the time integral of the SFRs at previous times and smaller masses. If anything, the TNG galaxy mass functions at $0.5 < z < 2$ in certain mass regimes (just left of the knee) are somewhat higher than observational constraints. On the other hand, tensions between the observed evolution of the galaxy mass or luminosity functions and the observed sSFRs – particularly at $z \gtrsim 2$ – have been pointed out by several authors (see e.g. Leja et al. 2015; Yu & Wang 2016; Behroozi et al. 2018, and references therein), albeit the actual reason of these tensions is still unclear. We postpone to future work the task of placing into a coherent and consistent picture all these statements and the quantification of the effects of environment on the star formation activity of TNG galaxies.

ACKNOWLEDGEMENTS

The authors would like to thank Kate Whitaker, Camilla Pacifici, Sandy Faber, Arjen van der Wel, Bruno Henriques, Adam Muzzin, and Robert Feldmann for useful discussions and input and all the participants of the Lorentz center workshop on ‘A Decade of the Star-Forming Main Sequence’ for inspiring conversations. The authors are also grateful to Martin Sparre, Gandhali Joshi, Gergő Popping, and Manuel Arca Sedda for comments on a draft version of this paper and to Sarah Leslie for sharing her data. SG, through the Flatiron Institute, is supported by the Simons Foundation. MV acknowledges support through an MIT RSC award, a Kavli Research Investment Fund, NASA ATP grant NNX17AG29G, and NSF grants AST-1814053 and AST-1814259. The flagship simulations of the IllustrisTNG project used in this work have been run on the HazelHen 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

- Abramson L. E., Kelson D. D., Dressler A., Poggianti B., Gladders M. D., Oemler Jr. A., Vulcani B., 2014, *ApJ*, 785, L36
 Archibald E. N., Dunlop J. S., Hughes D. H., Rawlings S., Eales S. A., Ivison R. J., 2001, *MNRAS*, 323, 417
 Bahé Y. M., McCarthy I. G., 2015, *MNRAS*, 447, 969
 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, AIP Conf. Proc. Vol. 743, The New Cosmology: Conference on Strings and Cosmology. Am. Inst. Phys., New York, p. 106
 Balogh M. L., Baldry I. K., Nichol R., Miller C., Bower R., Glazebrook K., 2004, *ApJ*, 615, L101
 Behroozi P., Wechsler R., Hearin A., Conroy C., 2018, preprint ([arXiv:1808.07447](https://arxiv.org/abs/1808.07447))
 Bendo G. J. et al., 2010, *A&A*, 518, L65

- Bouché N. et al., 2010, *ApJ*, 718, 1001
- Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151
- Chabrier G., 2003, *PASP*, 115, 763
- Chang Y.-Y., van der Wel A., da Cunha E., Rix H.-W., 2015, *ApJS*, 219, 8
- Conroy C., Gunn J. E., 2010, *ApJ*, 712, 833
- Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486
- Daddi E. et al., 2007, *ApJ*, 670, 156
- 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
- Davé R., 2008, *MNRAS*, 385, 147
- Davé R., Finlator K., Oppenheimer B. D., 2012, *MNRAS*, 421, 98
- Davé R., Rafieferantsoa M. H., Thompson R. J., 2017, *MNRAS*, 471, 1671
- Davidzon I., Ilbert O., Faisst A. L., Sparre M., Capak P. L., 2018, *ApJ*, 852, 107
- Davies L. J. M. et al., 2017, *MNRAS*, 466, 2312
- Davies L. J. M. et al., 2018, *MNRAS*, 480, 768
- 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
- Eales S., de Vis P., Smith M. W. L., Appah K., Ciesla L., Duffield C., Schofield S., 2017, *MNRAS*, 465, 3125
- Eales S. et al., 2018, *MNRAS*, 473, 3507
- Elbaz D. et al., 2007, *A&A*, 468, 33
- Falkendal T. et al., 2019, *A&A*, 621, A27
- Fang J. J. et al., 2018, *ApJ*, 858, 100
- Feldmann R., 2017, *MNRAS*, 470, L59
- Foreman-Mackey D., Sick J., Johnson B., 2014, Technical report, python-fsps: Python bindings to FSPS (v0.1.1) (Version v0.1.1), Zenodo
- Furlong M. et al., 2015, *MNRAS*, 450, 4486
- Gavazzi G. et al., 2015, *A&A*, 580, A116
- Genel S. et al., 2014, *MNRAS*, 445, 175
- Genel S. et al., 2018, *MNRAS*, 474, 3976
- Groves B. et al., 2012, *MNRAS*, 426, 892
- Guidi G., Scannapieco C., Walcher J., Gallazzi A., 2016, *MNRAS*, 462, 2046
- Habouzit M. et al., 2018, *MNRAS*, 484, 4413
- Hahn C. et al., 2018, preprint ([arXiv:1809.01665](https://arxiv.org/abs/1809.01665))
- Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Wang W., 2017, *MNRAS*, 469, 2626
- Hirschmann M., De Lucia G., Wilman D., Weinmann S., Iovino A., Cucciati O., Zibetti S., Villalobos Á., 2014, *MNRAS*, 444, 2938
- Ilbert O. et al., 2013, *A&A*, 556, A55
- Ilbert O. et al., 2015, *A&A*, 579, A2
- Jian H.-Y. et al., 2018, *PASJ*, 70, S23
- Kang X., van den Bosch F. C., 2008, *ApJ*, 676, L101
- Karim A. et al., 2011, *ApJ*, 730, 61
- Katsianis A. et al., 2017, *MNRAS*, 472, 919
- Kauffmann G. et al., 2003, *MNRAS*, 341, 54
- Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, *MNRAS*, 353, 713
- Kennicutt R. C., Jr, 1989, *ApJ*, 344, 685
- Kennicutt R. C., Evans N. J., 2012, *ARA&A*, 50, 531
- Kennicutt R. C., Jr, Schweizer F. B. J. E., 1998, *Galaxies: Interactions and Induced Star Formation*. Springer-Verlag, Berlin Heidelberg
- Kewley L. J., Jansen R. A., Geller M. J., 2005, *PASP*, 117, 227
- Kroupa P., 2001, *MNRAS*, 322, 231
- Lee N. et al., 2015, *ApJ*, 801, 80
- Lee B. et al., 2018, *ApJ*, 853, 131
- Leja J., van Dokkum P. G., Franx M., Whitaker K. E., 2015, *ApJ*, 798, 115
- Leja J. et al., 2018, preprint ([arXiv:1812.05608](https://arxiv.org/abs/1812.05608))
- Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, *ApJ*, 772, 119
- Lin L. et al., 2014, *ApJ*, 782, 33
- Lovell M. R. et al., 2018, *MNRAS*, 481, 1950
- Marinacci F. et al., 2018, *MNRAS*, 480, 5113
- McGee S. L., Balogh M. L., Wilman D. J., Bower R. G., Mulchaey J. S., Parker L. C., Oemler A., 2011, *MNRAS*, 413, 996
- Mistani P. A. et al., 2016, *MNRAS*, 455, 2323
- Mitra S., Davé R., Finlator K., 2015, *MNRAS*, 452, 1184
- Mitra S., Davé R., Simha V., Finlator K., 2017, *MNRAS*, 464, 2766
- Moustakas J. et al., 2013, *ApJ*, 767, 50
- Muzzin A. et al., 2013, *ApJ*, 777, 18
- Naiman J. P. et al., 2018, *MNRAS*, 477, 1206
- Nelson D. et al., 2015, *Astron. Comput.*, 13, 12
- Nelson D. et al., 2018a, *MNRAS*, 475, 624
- Nelson D. et al., 2018b, *MNRAS*, 477, 450
- Nelson D. et al., 2019, preprint ([arXiv:e-print](https://arxiv.org/abs/1905.00016))
- Noeske K. G. et al., 2007, *ApJ*, 660, L43
- Oliver S. et al., 2010, *MNRAS*, 405, 2279
- Paccagnella A. et al., 2016, *ApJ*, 816, L25
- Pannella M. et al., 2009, *ApJ*, 698, L116
- Patel S. G., Holden B. P., Kelson D. D., Franx M., van der Wel A., Illingworth G. D., 2012, *ApJ*, 748, L27
- Pearson W. J. et al., 2018, *A&A*, 615, A146
- Peng Y.-j. et al., 2010, *ApJ*, 721, 193
- Pillepich A. et al., 2018a, *MNRAS*, 473, 4077
- Pillepich A. et al., 2018b, *MNRAS*, 475, 648
- Pillepich A. et al., 2019, preprint ([arXiv:e-print](https://arxiv.org/abs/1905.00016))
- Planck Collaboration et al., 2016, *A&A*, 594, A13
- Quadri R. F., Williams R. J., Franx M., Hildebrandt H., 2012, *ApJ*, 744, 88
- Reddy N. A., Steidel C. C., Pettini M., Adelberger K. L., Shapley A. E., Erb D. K., Dickinson M., 2008, *ApJS*, 175, 48
- Renzini A., Peng Y.-j., 2015, *ApJ*, 801, L29
- Richards S. N. et al., 2016, *MNRAS*, 455, 2826
- Rodighiero G. et al., 2011, *ApJ*, 739, L40
- Rodighiero G. et al., 2014, *MNRAS*, 443, 19
- Salim S. et al., 2007, *ApJS*, 173, 267
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Santini P. et al., 2009, *A&A*, 504, 751
- Santini P. et al., 2014, *A&A*, 562, A30
- Santini P. et al., 2017, *ApJ*, 847, 76
- Schawinski K. et al., 2014, *MNRAS*, 440, 889
- Schreiber C. et al., 2015, *A&A*, 575, A74
- Schreiber C., Elbaz D., Pannella M., Ciesla L., Wang T., Koekemoer A., Rafelski M., Daddi E., 2016, *A&A*, 589, A35
- Shivaei I. et al., 2015, *ApJ*, 815, 98
- Sijacki D., Vogelsberger M., Genel S., Springel V., Torrey P., Snyder G. F., Nelson D., Hernquist L., 2015, *MNRAS*, 452, 575
- Somerville R. S., Davé R., 2015, *ARA&A*, 53, 51
- Sorba R., Sawicki M., 2018, *MNRAS*, 476, 1532
- Sparre M. et al., 2015, *MNRAS*, 447, 3548
- Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, *ApJS*, 214, 15
- Springel V., 2010, *MNRAS*, 401, 791
- Springel V., Hernquist L., 2003, *MNRAS*, 339, 289
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Springel V. et al., 2018, *MNRAS*, 475, 676
- Strateva I. et al., 2001, *AJ*, 122, 1861
- Tasca L. A. M. et al., 2015, *A&A*, 581, A54
- Theios R. L., Steidel C. C., Strom A. L., Rudie G. C., Trainor R. F., Reddy N. A., 2018, preprint ([arXiv:1805.00016](https://arxiv.org/abs/1805.00016))
- Tomczak A. R. et al., 2014, *ApJ*, 783, 85
- Tomczak A. R. et al., 2016, *ApJ*, 817, 118
- Torrey P., Vogelsberger M., Genel S., Sijacki D., Springel V., Hernquist L., 2014, *MNRAS*, 438, 1985
- Torrey P. et al., 2018, *MNRAS*, 477, L16
- Trayford J. W., Theuns T., Bower R. G., Crain R. A., Lagos C. d. P., Schaller M., Schaye J., 2016, *MNRAS*, 460, 3925
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, *MNRAS*, 436, 3031
- Vogelsberger M. et al., 2014a, *MNRAS*, 444, 1518
- Vogelsberger M. et al., 2014b, *Nature*, 509, 177

- Vogelsberger M. et al., 2018, *MNRAS*, 474, 2073
 Vulcani B., Poggianti B. M., Finn R. A., Rudnick G., Desai V., Bamford S., 2010, *ApJ*, 710, L1
 Vulcani B. et al., 2012, *MNRAS*, 420, 1481
 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., van Dokkum P. G., Brammer G., Franx M., 2012, *ApJ*, 754, L29
 Whitaker K. E. et al., 2010, *ApJ*, 719, 1715
 Whitaker K. E. et al., 2011, *ApJ*, 735, 86
 Whitaker K. E. et al., 2014, *ApJ*, 795, 104
 Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, *ApJ*, 691, 1879
 Williams R. J., Quadri R. F., Franx M., van Dokkum P., Toft S., Kriek M., Labbé I., 2010, *ApJ*, 713, 738
 Wuyts S. et al., 2011, *ApJ*, 742, 96
 Yu H., Wang F. Y., 2016, *ApJ*, 820, 114
 Zahid H. J., Dima G. I., Kewley L. J., Erb D. K., Davé R., 2012, *ApJ*, 757, 54

APPENDIX A: SFRs OVER DIFFERENT TIME-SCALES AND EFFECTS OF NUMERICAL RESOLUTION

In this paper we use two simulations of the IllustrisTNG project: TNG100 and TNG300, characterized by a factor of 8 (2) difference in mass (spatial) resolution (see Table 1 and www.tng-project.org). We measure the SFRs of simulated galaxies by averaging the stellar mass production rate over different time-scales. In this section, we discuss how our results are affected by different numerical resolutions and for different SFR estimates.

First, Fig. A1 shows explicitly the SFR– M_{stars} plane for TNG100 (left) and TNG300 (right) for increasingly longer SFR averaging time-scales (from top to bottom). The galaxy SFRs in all panels are measured within $2 \times R_{\text{star,h}}$. As a reminder, TNG100 has the same initial conditions, cosmological box volume (~ 100 Mpc on-a-side) and similar resolution of the original Illustris run, with a baryonic target mass of $m_{\text{baryon}} \sim 1.4 \times 10^6$ and $\sim 1.26 \times 10^6 M_{\odot}$, respectively. On the other hand, the volume of TNG300 is ~ 300 Mpc on-a-side, with the baryonic target mass of $1.1 \times 10^7 M_{\odot}$. The mass of each gas cells is kept within a factor of 2 of the target mass (Pillepich et al. 2018a) and so is the initial mass of the star particles at birth.

All galaxies considered in this work have a minimum stellar mass of $10^9 M_{\odot}$, meaning that each galaxy in TNG100 has about at least 1000 star particles in total. In the extreme case of small galaxies ($M_{\text{stars}} = 10^9 M_{\odot}$) in the TNG300 simulated volume, the number of star particles counted in each galaxy is of the order of 100, forcing us to be careful in using TNG300 in this limiting case.

Moreover, the different mass resolutions affect our estimates for the galaxy SFRs, especially if it is averaged over different time-scales. From the minimum baryonic particle masses of each simulation we can estimate the lowest measurable value of the SFR, SFR_{min} . For TNG100, the smallest non-vanishing SFR values read $\log \text{SFR}_{\text{min}} (M_{\odot} \text{ yr}^{-1}) \simeq -1.15, -1.85, -2.46, -3.15$ for SFRs averaged over 10, 50, 200, and 1000 Myr, respectively. Analogue considerations for the TNG300 run imply a minimum resolved values for SFR of $\log \text{SFR}_{\text{min}} (M_{\odot} \text{ yr}^{-1}) \simeq -0.6, -1.3, -1.9, -2.6$, for the same averaging time-scales.

On the other hand, for the case of the instantaneous SFR of the gas cell, the lowest SFR measurable values are found to be $\log \text{SFR} (M_{\odot} \text{ yr}^{-1}) \simeq -4$ and $\log \text{SFR} (M_{\odot} \text{ yr}^{-1}) \simeq -3$ for TNG100 and TNG300, respectively. Galaxies with $\log \text{SFR} < \log \text{SFR}_{\text{min}}$, have unresolved SFR values, corresponding to $\text{SFR} = 0$ in the simulations and will be assigned to the range $10^{-5} - 10^{-4} M_{\odot} \text{ yr}^{-1}$.

The effect of the mass resolutions on the averaging time-scales is well detectable by-eye in Fig. A1, where galaxies are not found below a certain value of SFR. Indeed, individual star particles have such high masses that low-SFR galaxies will very rarely be able to spawn a new star particle, thus in general there will be few young stars with age less than e.g. 10 Myr. This also corresponds to no young stars being able to generate blue light, which is why it is impossible to get blue enough galaxies at poor resolution (Trayford et al. 2016; Nelson et al. 2018a).

Mass resolution effects are much more pronounced for shorter time-scales: this could potentially bias high the main sequence’s locus. This is the case, for example, for the shortest time-scale (10 Myr) at the low-mass end ($M_{\text{stars}} < 10^{9.5} M_{\odot}$): see top left panel of Fig. 6. For this reason, we have not enunciated statements in the paper that are based on such short time-scales at low galaxy masses.

We have analysed the shape of the SFR– M_{stars} planes and of the star-forming MS also for the lower resolution runs of TNG100, namely TNG100-2 and TNG100-3, that have 8 and 64 times worse mass resolution than TNG100, respectively (TNG100-2 being comparable in resolution to TNG300). We find that the worst mass resolution of TNG100-3 leads to a considerably overestimation of the MS, in terms of both slope and normalization, unless the longest time-scales – 1000 Myr – are considered. On the other hand, the intermediate numerical resolutions of TNG100-2 and TNG300 return results on the star-forming MS that are converged with those of TNG100 if the SFRs are averaged over time-scales larger than 50 Myr at masses above $10^{10} M_{\odot}$.

These considerations are taken into account when the results of the paper are discussed and motivate the usage of the 200-Myr SFR estimates as fiducial choice for both TNG100 and TNG300.

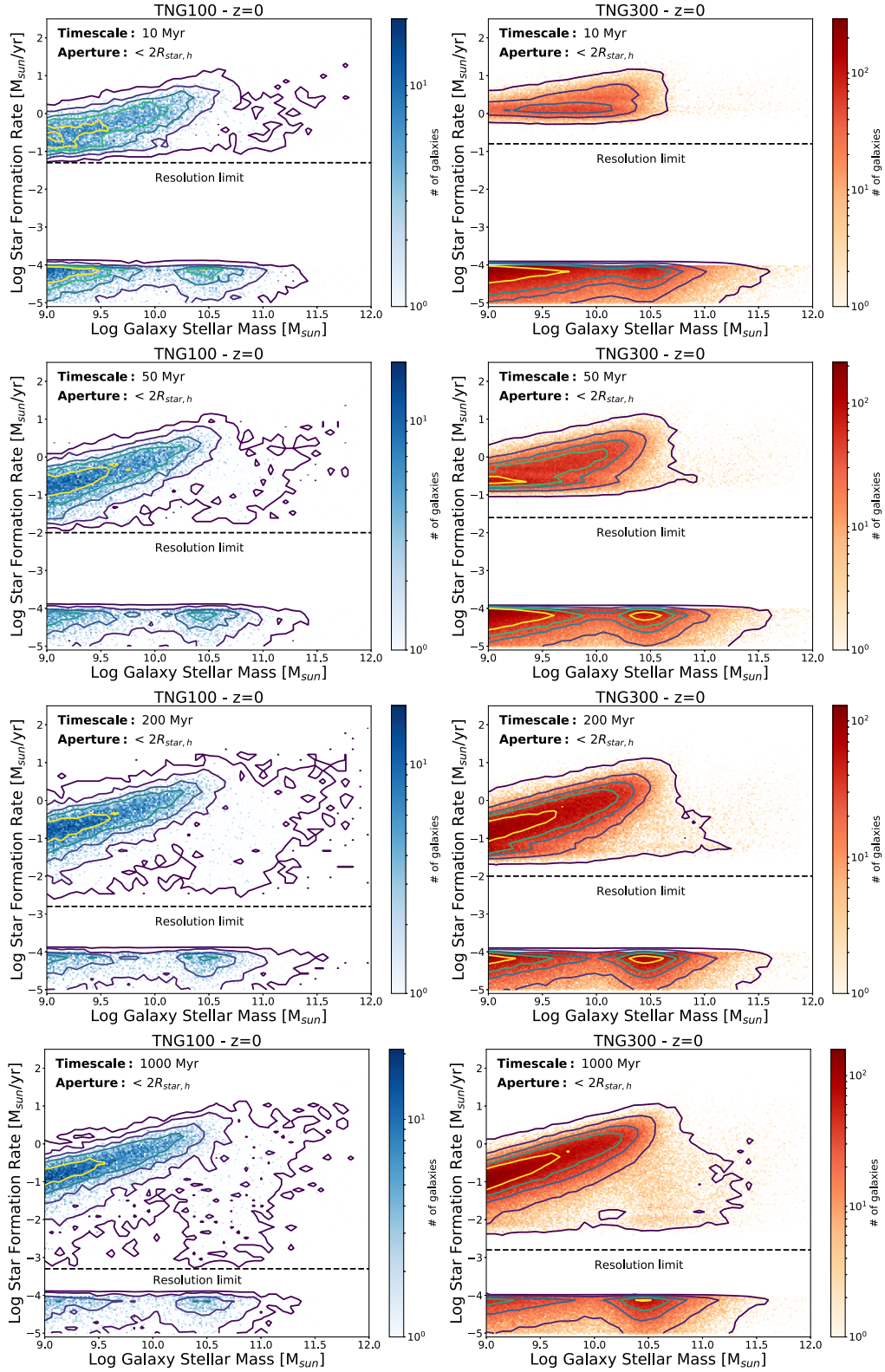


Figure A1. SFR as a function of stellar mass M_{stars} of TNG300 (left-hand column) and TNG100 (right-hand column) galaxies. From top to bottom panels the SFR is averaged over 10, 50, 200, and 1000 Myr and measured within $2 \times R_{\text{star},h}$. The colour scale indicates the number of galaxies and the solid contours encompass 20, 50, 70, 90, and 99 per cent of the galaxies. To all galaxies with unresolved SFR values we assign a random values in the range $\text{SFR} = 10^{-5} - 10^{-4} M_{\odot} \text{yr}^{-1}$.

This paper has been typeset from a \LaTeX file prepared by the author.