

RESEARCH

Open Access



The IllustrisTNG simulations: public data release

Dylan Nelson^{1*}, Volker Springel^{1,6,7}, Annalisa Pillepich², Vicente Rodriguez-Gomez⁸, Paul Torrey^{9,5}, Shy Genel⁴, Mark Vogelsberger⁵, Ruediger Pakmor^{1,6}, Federico Marinacci^{10,5}, Rainer Weinberger³, Luke Kelley¹¹, Mark Lovell^{12,13}, Benedikt Diemer³ and Lars Hernquist³

Abstract

We present the full public release of all data from the TNG100 and TNG300 simulations of the IllustrisTNG project. IllustrisTNG is a suite of large volume, cosmological, gravo-magnetohydrodynamical simulations run with the moving-mesh code AREPO. TNG includes a comprehensive model for galaxy formation physics, and each TNG simulation self-consistently solves for the coupled evolution of dark matter, cosmic gas, luminous stars, and supermassive black holes from early time to the present day, $z = 0$. Each of the flagship runs—TNG50, TNG100, and TNG300—are accompanied by halo/subhalo catalogs, merger trees, lower-resolution and dark-matter only counterparts, all available with 100 snapshots. We discuss scientific and numerical cautions and caveats relevant when using TNG.

The data volume now directly accessible online is ~ 750 TB, including 1200 full volume snapshots and $\sim 80,000$ high time-resolution subbox snapshots. This will increase to ~ 1.1 PB with the future release of TNG50. Data access and analysis examples are available in IDL, Python, and Matlab. We describe improvements and new functionality in the web-based API, including on-demand visualization and analysis of galaxies and halos, exploratory plotting of scaling relations and other relationships between galactic and halo properties, and a new JupyterLab interface. This provides an online, browser-based, near-native data analysis platform enabling user computation with local access to TNG data, alleviating the need to download large datasets.

Keywords: Methods: data analysis; Methods: numerical; Galaxies: formation; Galaxies: evolution; Data management systems; Data access methods; Distributed architectures

1 Introduction

Some of our most powerful tools for understanding the origin and evolution of large-scale cosmic structure and the galaxies which form therein are cosmological simulations. From pioneering beginnings (Press and Schechter 1974; Davis et al. 1985), dark matter, gravity-only simulations have evolved into cosmological hydrodynamical simulations (Katz et al. 1992). These aim to self-consistently model the coupled evolution of dark matter, gas, stars, and black holes at a minimum, and are now being extended to also include magnetic fields, radiation, cosmic rays, and other fundamental physical components. Such

simulations provide foundational support in our understanding of the Λ CDM cosmological model, including the nature of both dark matter and dark energy.

Modern large-volume simulations now capture cosmological scales of tens to hundreds of comoving megaparsecs, while simultaneously resolving the internal structure of individual galaxies at $\lesssim 1$ kpc scales. Recent examples reaching $z = 0$ include Illustris (Vogelsberger et al. 2014a; Genel et al. 2014), EAGLE (Schaye et al. 2015; Crain et al. 2015), Horizon-AGN (Dubois et al. 2014), Romulus (Tremmel et al. 2017), Simba (Davé et al. 2019), Magnetium (Dolag et al. 2016), among others. These simulations produce observationally verifiable outcomes across a diverse range of astrophysical regimes, from the stellar and gaseous properties of galaxies, galaxy populations, and the

*Correspondence: dnelson@mpa-garching.mpg.de

¹Max-Planck-Institut für Astrophysik, Garching, Germany

Full list of author information is available at the end of the article

supermassive black holes they host, to the expected distribution of molecular, neutral, and ionized gas tracers across interstellar, circumgalactic, and intergalactic scales, in addition to the expected distribution of the dark matter component itself.

Complementary efforts, although not the focus of this data release, include high redshift reionization-era simulations such as BlueTides (Feng et al. 2016), Sphinx (Rosdahl et al. 2018), and CoDa II (Ocvirk et al. 2018), among others. In addition, ‘zoom’ simulation campaigns include NIHAO (Wang et al. 2015), FIRE-2 (Hopkins et al. 2018), and Auriga (Grand et al. 2017), in addition to many others. These have provided numerous additional insights into many questions in galaxy evolution (recent progress reviewed in Faucher-Giguère 2018). For instance, reionization simulations may be able to include explicit radiative transfer, and zoom simulations may be able to reach higher resolutions and/or more rapidly explore model variations, in comparison to large cosmological volume simulations.

Observational efforts studying the properties of galaxies across cosmic time provide ever richer datasets. Surveys such as SDSS (York et al. 2000), CANDELS (Grogin et al. 2011), 3D-HST (Brammer et al. 2012), LEGA-C (van der Wel et al. 2016), SINS/zC-SINF and KMOS3D (Genzel et al. 2014; Wisnioski et al. 2015), KBSS (Steidel et al. 2014), and MOSDEF (Kriek et al. 2015) provide local and high redshift measurements of the statistical properties of galaxy populations. Complementary, spatially-resolved data has recently become available from large, $z = 0$ IFU surveys such as MANGA (Bundy et al. 2015), CALIFA (Sánchez et al. 2012) and SAMI (Bryant et al. 2015).

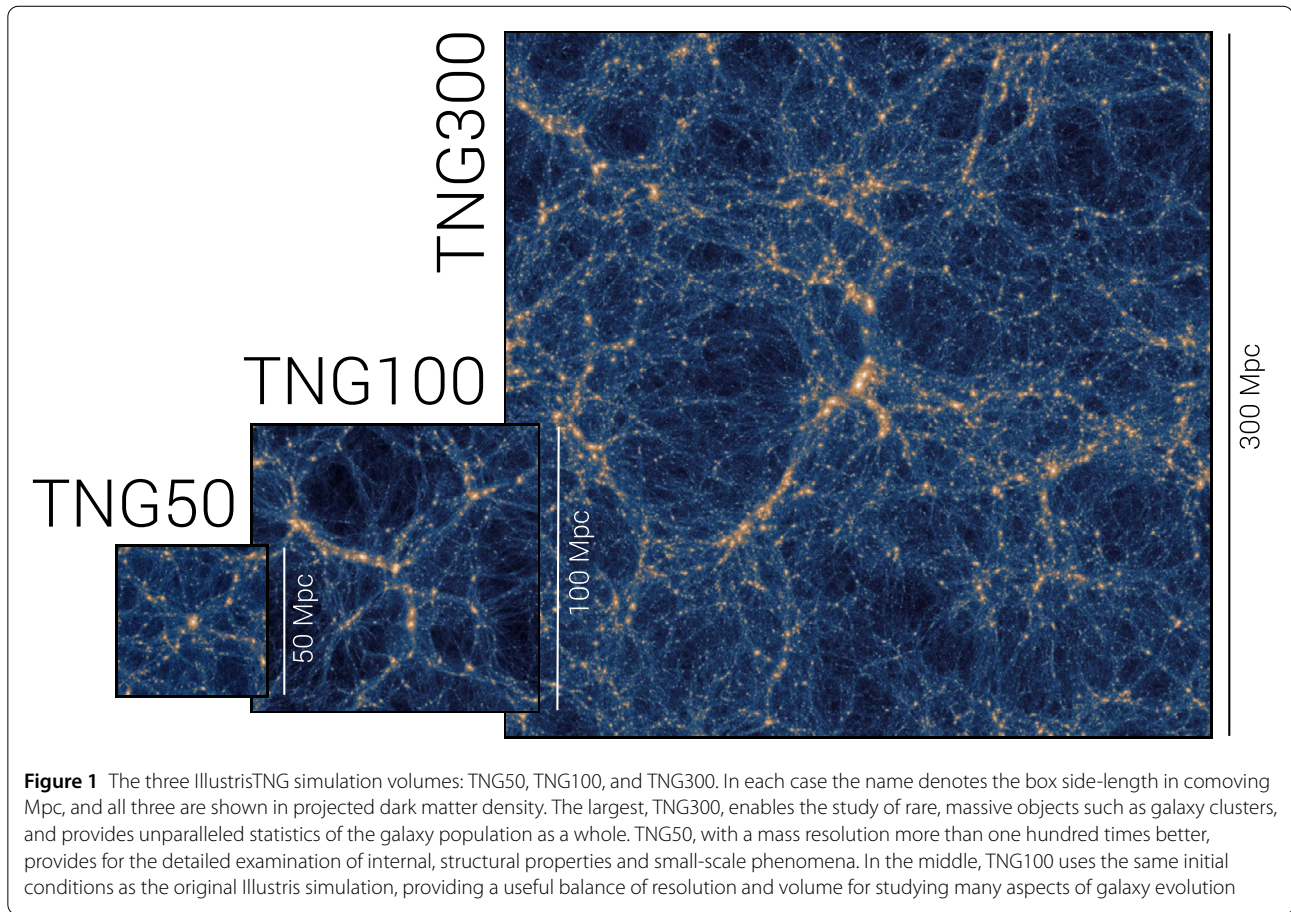
In order to inform theoretical models using observational constraints, as well as to interpret observational results using realistic cosmological models, public data dissemination from both observational and simulation campaigns is required. Observational data release has a successful history dating back at least to the SDSS SkyServer (Szalay et al. 2000, 2002), which provides tools for remote users to query and acquire large datasets (Gray et al. 2002; Szalay et al. 2002). The still-in-use approach is based on user written SQL queries, which provide search results as well as data acquisition. From the theoretical community, the public data release of the Millennium simulation (Springel et al. 2005) was the first attempt of similar scale. Modeled on the SDSS approach, data was stored in a relational database, with an immediately recognizable SQL-query interface (Lemson and VirgoConsortium 2006). It has since been extended to include additional simulations, including Millennium-II (Boylan-Kolchin et al. 2009; Guo et al. 2011), and a first attempt at the idea of a “virtual observatory” (VO) was realized (Overzier et al. 2013). The Theoretical Astrophysical Observatory (TAO; Bernyk et al. 2014) was similarly focused around mock observations of simulated galaxy and galaxy survey data.

Explorations continue on how to best deliver theoretical results within the existing VO framework (Lemson and Zuther 2009; Lemson et al. 2014).

Other dark-matter only simulations have released data with similar approaches, including Bolshoi and MultiDark (CosmoSim; Klypin et al. 2011; Riebe et al. 2013), DEUS (Rasera et al. 2010), and MICE (Cosmohub; Crocce et al. 2010). In contrast, some recent simulation projects have made group catalogs and/or snapshots available for direct download, including MassiveBlack-II (Khandai et al. 2014), the Dark Sky simulation (Skillman et al. 2014), ν^2 GC (Makiya et al. 2016), and Abacus (Garrison et al. 2018). Skies and Universes (Klypin et al. 2017) organizes a number of such data releases. With respect to Illustris, the most comparable in simulation type, data complexity, and scientific scope is the recent public data release of the Eagle simulation, described in (McAlpine et al. 2016) (see also Camps et al. 2018). The initial group catalog release was modeled on the Millennium database architecture, and the raw snapshot data was also subsequently made available (The EAGLE team 2017). More recently, significant infrastructure research and development has focused on providing remote computational resources, including the NOAO Data Lab (Fitzpatrick et al. 2014) and the SciServer project (Medvedev et al. 2016; Raddick et al. 2017). Web-based orchestration projects also include (Ragagnin et al. 2017), Tangos (Pontzen et al. 2018), and Jovial (Araya et al. 2018).

The public release of IllustrisTNG (hereafter, TNG) follows upon and further develops tools and ideas pioneered in the original Illustris data release. We offer direct online access to all snapshot, group catalog, merger tree, and supplementary data catalog files. In addition, we develop a web-based API which allows users to perform many common tasks without the need to download any full data files. These include searching over the group catalogs, extracting particle data from the snapshots, accessing individual merger trees, and requesting visualization and further data analysis functions. Extensive documentation and programmatic examples (in the IDL, Python, and Matlab languages) are provided.

This paper is intended primarily as an overview guide for TNG data users, describing updates and new features, while exhaustive documentation will be maintained online. In Sect. 2 we give an overview of the simulations. Section 3 describes the data products, and Sect. 4 discusses methods for data access. In Sect. 5 we present some scientific remarks and cautions, while in Sect. 6 we discuss community considerations including citation requests. Section 7 describes technical details related to the data release architecture, while Sect. 8 summarizes. Appendix 1 provides a few additional data details, while Appendix 2 shows several examples of how to use the API.



2 Description of the simulations

IllustrisTNG is a suite of large volume, cosmological, gravo-magnetohydrodynamical simulations run with the moving-mesh code AREPO (Springel 2010). The TNG project is made up of three simulation volumes: TNG50, TNG100, and TNG300, shown in Fig. 1. The first two simulations, TNG100 and TNG300, were recently introduced in a series of five presentation papers (Springel et al. 2018; Pillepich et al. 2018a; Naiman et al. 2018; Nelson et al. 2018a; Marinacci et al. 2018), and these are here publicly released in full. The third and final simulation of the project is TNG50 (Pillepich et al. 2019; Nelson et al. 2019a) which will also be publicly released in the future. TNG includes a comprehensive model for galaxy formation physics, which is able to realistically follow the formation and evolution of galaxies across cosmic time (Weinberger et al. 2017; Pillepich et al. 2018b). Each TNG simulation solves for the coupled evolution of dark matter, cosmic gas, luminous stars, and supermassive black holes from a starting redshift of $z = 127$ to the present day, $z = 0$.

The IllustrisTNG project^a is the successor of the original Illustris simulation^b (Genel et al. 2014; Sijacki et al. 2015; Vogelsberger et al. 2014b, 2014a) and its associated galaxy formation model (Vogelsberger et al. 2013; Torrey

et al. 2014). Illustris was publicly released in its entirety roughly three and a half years ago (Nelson et al. 2015). TNG incorporates an updated ‘next generation’ galaxy formation model which includes new physics and numerical improvements, as well as refinements to the original model. TNG newly includes a treatment of cosmic magnetism, following the amplification and dynamical impact of magnetic fields, as described below.

The three flagship runs of IllustrisTNG are each accompanied by lower-resolution and dark-matter only counterparts. Three physical simulation box sizes are employed: cubic volumes of roughly 50, 100, and 300 Mpc side length, which we refer to as TNG50, TNG100, and TNG300, respectively. The three boxes complement each other by enabling investigations of various aspects of galaxy formation. The large physical volume associated with the largest simulation box (TNG300) enables, for instance, the study of galaxy clustering, the analysis of rare and massive objects such as galaxy clusters, and provides the largest statistical galaxy sample. In contrast, the smaller physical volume simulation of TNG50 enables a mass resolution which is more than a hundred times better than in the TNG300 simulation, providing a more detailed look at, for example, the structural properties of galaxies, and small-scale gas

Table 1 Table of physical and numerical parameters for each of the resolution levels of the three flagship TNG simulations. The physical parameters are: the box volume, the box side-length, the initial number of gas cells, dark matter particles, and Monte Carlo tracer particles. The target baryon mass, the dark matter particle mass, the $z = 0$ Plummer equivalent gravitational softening of the collisionless component, the same value in comoving units, and the minimum comoving value of the adaptive gas gravitational softenings. Additional characterizations of the gas resolution, measured at redshift zero: the minimum physical gas cell radius, the median gas cell radius, the mean radius of $SFR > 0$ gas cells, the mean hydrogen number density of star-forming gas cells, and the maximum hydrogen gas density

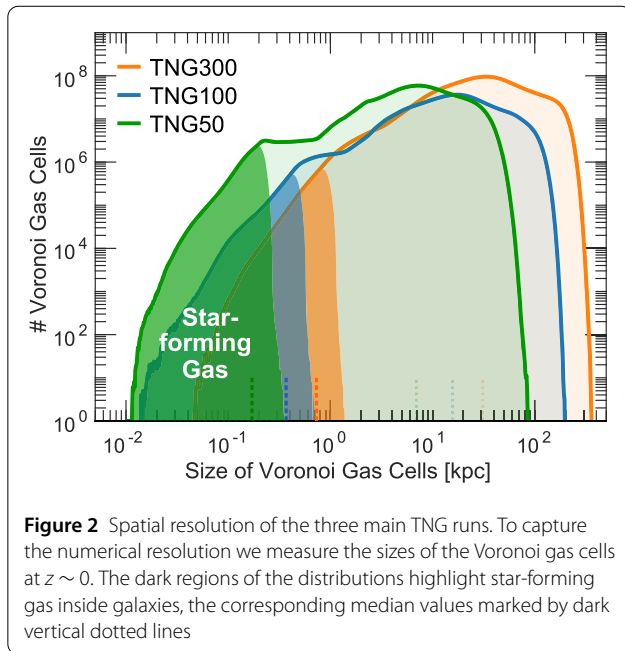
Run	Volume [cMpc ³]	L_{box} [cMpc/h]	$N_{\text{GAS,DM}}$ –	N_{TRACER} –	m_{baryon} [M_{\odot}/h]	m_{DM} [M_{\odot}/h]	m_{baryon} [$10^6 M_{\odot}$]	m_{DM} [$10^6 M_{\odot}$]
TNG50-1	51.7 ³	35	2160 ³	1 × 2160 ³	5.7 × 10 ⁴	3.1 × 10 ⁵	0.08	0.45
TNG50-2	51.7 ³	35	1080 ³	1 × 1080 ³	4.6 × 10 ⁵	2.5 × 10 ⁶	0.68	3.63
TNG50-3	51.7 ³	35	540 ³	1 × 540 ³	3.7 × 10 ⁶	2.0 × 10 ⁷	5.4	29.0
TNG50-4	51.7 ³	35	270 ³	1 × 270 ³	2.9 × 10 ⁷	1.6 × 10 ⁸	43.4	232
TNG100-1	106.5 ³	75	1820 ³	2 × 1820 ³	9.4 × 10 ⁵	5.1 × 10 ⁶	1.4	7.5
TNG100-2	106.5 ³	75	910 ³	2 × 910 ³	7.6 × 10 ⁶	4.0 × 10 ⁷	11.2	59.7
TNG100-3	106.5 ³	75	455 ³	2 × 455 ³	6.0 × 10 ⁷	3.2 × 10 ⁸	89.2	478
TNG300-1	302.6 ³	205	2500 ³	1 × 2500 ³	7.6 × 10 ⁶	4.0 × 10 ⁷	11	59
TNG300-2	302.6 ³	205	1250 ³	1 × 1250 ³	5.9 × 10 ⁷	3.2 × 10 ⁸	88	470
TNG300-3	302.6 ³	205	625 ³	1 × 625 ³	4.8 × 10 ⁸	2.5 × 10 ⁹	703	3760
TNG50-1-Dark	51.7 ³	35	2160 ³	–	–	3.7 × 10 ⁵	–	0.55
TNG50-2-Dark	51.7 ³	35	1080 ³	–	–	2.9 × 10 ⁶	–	4.31
TNG50-3-Dark	51.7 ³	35	540 ³	–	–	2.3 × 10 ⁷	–	34.5
TNG50-4-Dark	51.7 ³	35	270 ³	–	–	1.9 × 10 ⁸	–	275
TNG100-1-Dark	106.5 ³	75	1820 ³	–	–	6.0 × 10 ⁶	–	8.9
TNG100-2-Dark	106.5 ³	75	910 ³	–	–	4.8 × 10 ⁷	–	70.1
TNG100-3-Dark	106.5 ³	75	455 ³	–	–	3.8 × 10 ⁸	–	567
TNG300-1-Dark	302.6 ³	205	2500 ³	–	–	7.0 × 10 ⁷	–	47
TNG300-2-Dark	302.6 ³	205	1250 ³	–	–	3.8 × 10 ⁸	–	588
TNG300-3-Dark	302.6 ³	205	625 ³	–	–	3.0 × 10 ⁹	–	4470

Run	$\epsilon_{\text{DM},\star}^{z=0}$ [kpc]	$\epsilon_{\text{DM},\star}$ [ckpc/h]	$\epsilon_{\text{gas,min}}$ [ckpc/h]	$r_{\text{cell,min}}$ [pc]	\bar{r}_{cell} [kpc]	$\bar{r}_{\text{cell,SF}}$ [pc]	$\bar{n}_{\text{H,SF}}$ [cm ⁻³]	$\eta_{\text{H,max}}$ [cm ⁻³]
TNG50-1	0.29	0.39 → 0.195	0.05	8	5.8	138	0.8	650
TNG50-2	0.58	0.78 → 0.39	0.1	19	12.9	282	0.7	620
TNG50-3	1.15	1.56 → 0.78	0.2	65	25.0	562	0.6	80
TNG50-4	2.30	3.12 → 1.56	0.4	170	50.1	1080	0.5	35
TNG100-1	0.74	1.0 → 0.5	0.125	14	15.8	355	1.0	3040
TNG100-2	1.48	2.0 → 1.0	0.25	74	31.2	720	0.6	185
TNG100-3	2.95	4.0 → 2.0	0.5	260	63.8	1410	0.5	30
TNG300-1	1.48	2.0 → 1.0	0.25	47	31.2	715	0.6	490
TNG300-2	2.95	4.0 → 2.0	0.5	120	63.8	1420	0.5	235
TNG300-3	5.90	8.0 → 4.0	1.0	519	153	3070	0.4	30

phenomena in and around galaxies. Situated in the middle, the TNG100 simulation uses the same initial conditions (identical phases, adjusted for the updated cosmology) as the original Illustris simulation. This facilitates robust comparisons between the original Illustris results and the updated TNG model. For many galaxy evolution analyses, TNG100 provides an ideal balance of volume and resolution, particularly for intermediate mass halos. Despite these strengths, each volume still has intrinsic physical and numerical limitations—for instance, TNG300 is still small compared to the scale of the BAO for precision cosmology, and lacks statistics for the most massive halos at $\sim 10^{15} M_{\odot}$, while TNG50 is still too low-resolution to resolve ultra-faint dwarf galaxies with $M_{\star} \lesssim 10^5 M_{\odot}$, globular clusters, or small-scale galactic features such as

nuclear star clusters. We provide an overview and comparison between the specifications of all the TNG runs in Table 1.

This data release includes the TNG100 and TNG300 simulations in full. It will, in the future, also include the final TNG50 simulation. For each, snapshots at all 100 available redshifts, halo and subhalo catalogs at each snapshot, and two distinct merger trees are released. This includes three resolution levels of the 100 and 300 Mpc volumes, and four resolution levels of the 50 Mpc volume, decreasing in steps of eight in mass resolution (two in spatial resolution) across levels. The highest resolution realizations, TNG50-1, TNG100-1 and TNG300-1, include 2×2160^3 , 2×1820^3 and 2×2500^3 resolution elements, respectively (see Table 1). As the actual spatial resolution of cosmo-



logical hydrodynamical simulations is highly adaptive, it is poorly captured by a single number. Figure 2 therefore shows the distribution of Voronoi gas cell sizes in these three simulations, highlighting the high spatial resolution in star-forming gas—i.e., within galaxies themselves. In contrast, the largest gas cells occur in the low-density intergalactic medium.

All ten of the baryonic runs invoke, without modification and invariant across box and resolution, the fiducial “full” galaxy formation physics model of TNG. All ten runs are accompanied by matched, dark matter only (i.e. gravity-only) analogs. In addition, there are multiple, high time-resolution “subboxes”, with up to 8000 snapshots each and time spacing down to one million years.

This paper serves as the data release for IllustrisTNG as a whole, including the future public release of TNG50.

2.1 Physical models and numerical methods

All of the TNG runs start from cosmologically motivated initial conditions, assuming a cosmology consistent with the Planck Collaboration (2016) results ($\Omega_{\Lambda,0} = 0.6911$, $\Omega_{m,0} = 0.3089$, $\Omega_{b,0} = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$ and $h = 0.6774$), with Newtonian self-gravity solved in an expanding Universe. All of the baryonic TNG runs include the following additional physical components: (1) Primordial and metal-line radiative cooling in the presence of an ionizing background radiation field which is redshift-dependent and spatially uniform, with additional self-shielding corrections. (2) Stochastic star formation in dense ISM gas above a threshold density criterion. (3) Pressurization of the ISM due to unresolved supernovae using an effective equation of state model for the

two-phase medium. (4) Evolution of stellar populations, with associated chemical enrichment and mass loss (gas recycling), accounting for SN Ia/II, AGB stars, and NS-NS mergers. (5) Stellar feedback: galactic-scale outflows with an energy-driven, kinetic wind scheme. (6) Seeding and growth of supermassive black holes. (7) Supermassive black hole feedback: accreting BHs release energy in two modes, at high-accretion rates (‘quasar’ mode) and low-accretion rates (‘kinetic wind’ mode). Radiative proximity effects from AGN affect nearby gas cooling. (8) Magnetic fields: amplification of a small, primordial seed field and dynamical impact under the assumption of ideal MHD.

For complete details on the behavior, implementation, parameter selection, and validation of these physical models, see the two TNG methods papers: Weinberger et al. (2017) and Pillepich et al. (2018b). Table 2 provides an abridged list of the key differences between Illustris and TNG. We note that the TNG model has been designed (i.e. ‘calibrated’, or ‘tuned’) to broadly reproduce several basic, observed galaxy properties and statistics. These are: the galaxy stellar mass function and the stellar-to-halo mass relation, the total gas mass content within the virial radius (r_{500}) of massive groups, the stellar mass—stellar size and the BH–galaxy mass relations all at $z = 0$, in addition to the overall shape of the cosmic star formation rate density at $z \lesssim 10$ (see Pillepich et al. 2018b, for a discussion).

The TNG simulations use the moving-mesh AREPO code (Springel 2010) which solves the equations of continuum magnetohydrodynamics (MHD; Pakmor et al. 2011; Pakmor and Springel 2013) coupled with self-gravity. The latter is computed with the Tree-PM approach, while the fluid dynamics employs a Godunov (finite-volume) type method, with a spatial discretization based on an unstructured, moving, Voronoi tessellation of the domain. The Voronoi mesh is generated from a set of control points which move with the local fluid velocity modulo mesh regularization corrections. Assuming ideal MHD, an 8-wave Powell cleaning scheme maintains the zero divergence constraint. The previous MUSCL-Hancock scheme has been replaced with a time integration approach following Heun’s method, and the original Green-Gauss method for gradient estimation of primitive fluid quantities has been replaced with a least-squares method, obtaining second order convergence in the hydrodynamics (Pakmor et al. 2016). The long-range FFT calculation employs a new column-based MPI-parallel decomposition, while the gravity solver has been rewritten based on a recursive splitting of the N-body Hamiltonian into short- and long-timescale systems (as in GADGET-4, Springel in prep.). The code is second order in space, and with hierarchical adaptive time-stepping, also second order in time. Of order 10 million individual timesteps are required to evolve the high-resolution runs to redshift zero.

Table 2 Comparison of key model changes between Illustris and IllustrisTNG. For full details and a more comprehensive comparison including numerical parameter differences, see Table 1 of Pillepich et al. (2018b) and the two TNG methods papers in general

Simulation Aspect	Illustris	TNG (50/100/300)
Magnetic Fields	no	ideal MHD (Pakmor et al. 2011)
BH Low-State Feedback	‘Radio’ Bubbles	BH-driven wind (kinetic kick)
BH Accretion	Boosted Bondi-Hoyle ($\alpha = 100$)	Un-boosted Bondi-Hoyle
BH Seed mass	$10^5 M_{\odot}/h$	$8 \times 10^5 M_{\odot}/h$
Winds (Directionality)	bi-polar ($\vec{v}_{\text{gas}} \times \nabla\phi_{\text{grav}}$)	isotropic
Winds (Thermal Content)	cold	warm (10%)
Winds (Velocity)	$\propto \sigma_{\text{DM}}$	+ scaling with $H(z)$, and v_{min}
Winds (Energy)	constant per unit SFR	+ metallicity dependence in η
Stellar Evolution	Illustris Yields	TNG Yields
Metals Tagging	–	SNla, SNIi, AGB, NSNS, FeSNla, FeSNIi
Shock Finder	no	yes (Schaal and Springel 2015)

During the simulation we employ a Monte Carlo tracer particle scheme (Genel et al. 2013) to follow the Lagrangian evolution of baryons. An on-the-fly cosmic shock finder is coupled to the code (Schaal and Springel 2015; Schaal et al. 2016). Group catalogs are computed during the simulations using the FOF and SUBFIND (Springel et al. 2001) substructure identification algorithms.

2.2 Model validation and early findings

TNG has been shown to produce observationally consistent results in several regimes beyond those adopted to calibrate the model. Some examples regarding galaxy populations, galactic structural and stellar population properties include: the shapes and widths of the red sequence and blue cloud of SDSS galaxies (Nelson et al. 2018a); the shapes and normalizations of the galaxy stellar mass functions up to $z \sim 4$ (Pillepich et al. 2018a); the spatial clustering of red vs. blue galaxies from tens of kpc to tens of Mpc separations (Springel et al. 2018); the spread in Europium abundance of metal-poor stars in Milky Way like halos (Naiman et al. 2018); the emergence of a population of quenched galaxies both at low (Weinberger et al. 2018) and high redshift (Habouzit et al. 2018); stellar sizes up to $z \sim 2$, including separate star-forming and quiescent populations (Genel et al. 2018); the $z = 0$ and evolution of the gas-phase mass-metallicity relation (Torrey et al. 2017); the dark matter fractions within the extended bodies of massive galaxies at $z = 0$ in comparison to e.g. SLUGGS results (Lovell et al. 2018); and the optical morphologies of galaxies in comparison to Pan-STARRS observations (Rodriguez-Gomez et al. 2019).

The IllustrisTNG model also reproduces a broad range of unusual galaxies, tracing tails of the galaxy population, including low surface brightness galaxies (Zhu et al. 2018) and jellyfish, ram-pressure stripped galaxies (Yun et al. 2018). The large-volume of TNG300 helps demonstrate reasonable agreement in several galaxy cluster, intra-cluster and circumgalactic medium properties—for example, the scaling relations between total radio power and

X-ray luminosity, total mass, and Sunyaev–Zel’dovich parameter of massive haloes (Marinacci et al. 2018); the distribution of metals in the intra-cluster plasma (Vogelsberger et al. 2018); the observed fraction of cool core clusters (Barnes et al. 2018); and the OVI content of the circumgalactic media around galaxies from surveys at low redshift including COS-Halos and eCGM (Nelson et al. 2018b).

IllustrisTNG is also producing novel insights on the formation and evolution of galaxies. For instance, halo mass alone is a good predictor for the entire stellar mass profile of massive galaxies (Pillepich et al. 2018a); the metal enrichment of cluster galaxies is higher than field counterparts at fixed mass and this enhancement is present pre-infall (Gupta et al. 2018); star-forming and quenched galaxies take distinct evolutionary pathways across the galaxy size-mass plane (Genel et al. 2018) and exhibit systematically different column densities of OVI ions (Nelson et al. 2018b) and different magnetic-field strengths (Nelson et al. 2018a) at fixed galaxy stellar mass, as well as different magnetic-field topologies (Marinacci et al. 2018). Galaxies oscillate around the star formation main sequence and the mass-metallicity relations over similar timescales and often in an anti-correlated fashion (Torrey et al. 2018); the presence of jellyfish galaxies is signaled by large-scale bow shocks in their surrounding intra-cluster medium (Yun et al. 2018); baryonic processes affect the matter power spectrum across a range of scales (Springel et al. 2018) and steepen the inner power-law total density profiles of early-type galaxies (Wang et al. 2018); a significant number of OVII, OVIII (Nelson et al. 2018b) and NeIX (Martizzi et al. 2018) absorption systems are expected to be detectable by future X-ray telescopes like ATHENA.

IllustrisTNG has also been used to generate mock 21-cm maps (Villaescusa-Navarro et al. 2018) and estimates of the molecular hydrogen budget (Diemer et al. 2018) in central and satellite galaxies in the local (Stevens et al. 2018) as well as in the high-redshift Universe as probed by ALMA (Popping et al. 2019). Finally, TNG provides a test bed to ex-

plore future observational applications of machine learning techniques: for example, the use of Deep Neural Networks to estimate galaxy cluster masses from Chandra X-ray mock images (Ntampaka et al. 2018) or optical morphologies versus SDSS (Huertas-Company et al. 2019).

See the up to date list of results^c for additional references. Please note that on this page we provide, and will continue to release, data files accompanying published papers as appropriate. For instance, electronic versions of tables, and data points from key lines and figures, to enable comparisons with other results. These are available with small [DATA] links next to each paper.

2.3 Breadth of simulated data

All of the observational validations and early results from TNG100 and TNG300 demonstrate the broad applications of the IllustrisTNG simulations. To give a sense of the expansive scope, the richness of the resulting data products, and the potential for wide applications across many areas of galaxy formation and cosmology, Fig. 3 visualizes the TNG100 simulation at redshift zero. Each slice reveals a view into the synthetic IllustrisTNG universe. Together, they range from purely theoretical quantities to directly observable signatures, spanning across the baryonic and non-baryonic matter components of the simulation: dark matter, gas, stars, and black holes.

The wealth of available information in the simulation outputs translates directly into the wide range of astrophysical phenomena which can be explored with the TNG simulations.

3 Data products

We release all 100 snapshots of the IllustrisTNG cosmological volumes. These include up to five types of resolution elements (dark matter particles, gas cells, gas tracers, stellar and stellar wind particles, and supermassive black holes). The same volumes are available at multiple resolutions: high (-1 suffix, e.g. TNG100-1), intermediate (-2 suffix), and low (-3 suffix), always separated by a factor of two (eight) in spatial (mass) resolution. At each resolution, these ‘baryonic’ runs include the fiducial TNG model for galaxy formation physics. Each baryonic run is matched to its dark matter only analog (-Dark suffix).

For all runs, at every snapshot, two types of group catalogs are provided: friends-of-friends (FoF) halo catalogs, and SUBFIND subhalo catalogs. In postprocessing, these catalogs are used to generate two distinct merger trees, which are both released: SUBLINK, and LHALOTREE. Finally, supplementary data catalogs containing additional computations and modeling, and focusing on a variety of topics, are being continually created and released. All these data types are described below.

3.1 Snapshots

3.1.1 Snapshot organization

There are 100 snapshots stored for every run. These include all particles/cells in the whole volume. The complete snapshot listings, spacings and redshifts can be found online. Note that, unlike in Illustris, TNG contains two different types of snapshots: ‘full’ and ‘mini’. While both encompass the entire volume, ‘mini’ snapshots only have a subset of particle fields available (detailed online). In TNG, twenty snapshots are full, while the remaining 80 are mini. The 20 full snapshots are given in Table 3. Every snapshot is stored on-disk in a series of ‘chunks’, which are more manageable, smaller HDF5 files—additional details are provided in Table A.1 of the Appendix.

Note that, just as in Illustris, the snapshot data is not organized according to spatial position. Rather, particles within a snapshot are sorted based on their group/subgroup memberships, according to the FoF or SUBFIND algorithms. Within each particle type, the sort order is: group number, subgroup number, and then binding energy. Particles/cells belonging to the group but not to any of its subhalos (“inner fuzz”) are included after the last subhalo of each group. In Fig. 4 we show a schematic of the particle organization (as in Nelson et al. 2015), for *one particle type*. Note that halos may happen to be stored across multiple, subsequent file chunks, and different particle types of a halo are in general stored in different sets of file chunks.

3.1.2 Snapshot contents

Each HDF5 snapshot contains several groups: ‘Header’, ‘Parameters’, ‘Configuration’, and five additional ‘PartTypeX’ groups, for the following particle types (DM only runs have a single PartType1 group):

- PartType0—GAS
- PartType1—DM
- PartType2—(unused)
- PartType3—TRACERS
- PartType4—STARS & WIND PARTICLES
- PartType5—BLACK HOLES

The ‘Header’ group contains a number of attributes giving metadata about the simulation and snapshot. The ‘Parameters’ and ‘Configuration’ groups provide the complete set of run-time parameter and compile-time configuration options used to run TNG. That is, they encode the fiducial “TNG Galaxy Formation Model”. Many will clearly map to Table 1 of Pillepich et al. (2018b), while others deal with more numerical/technical options. In the future, together with the release of the TNG initial conditions and the TNG code base, this will enable any of the TNG simulations to be reproduced.

The complete snapshot field listings of the ‘PartTypeX’ groups, including dimensions, units and descriptions, are given online. The general system of units is kpc/h for

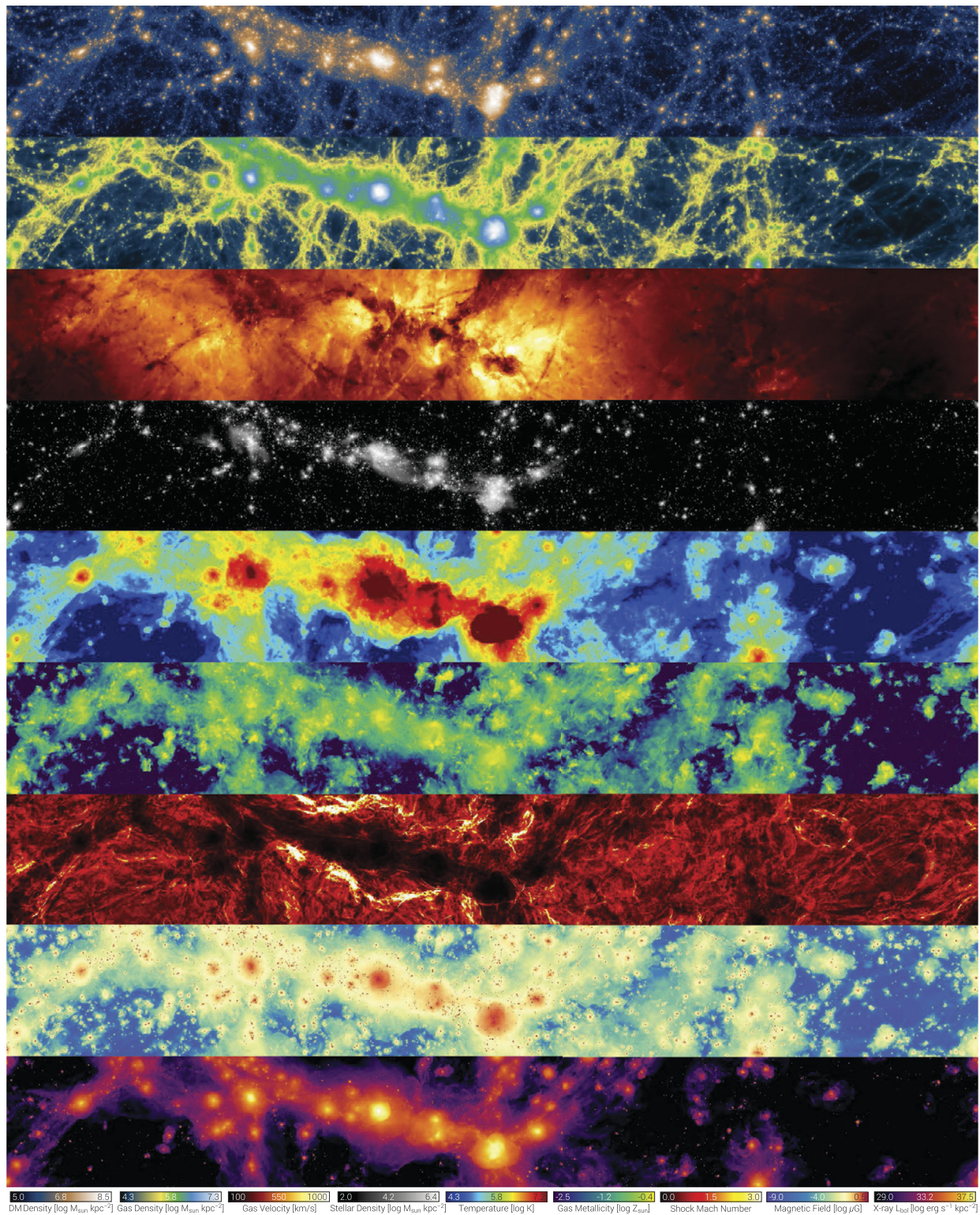
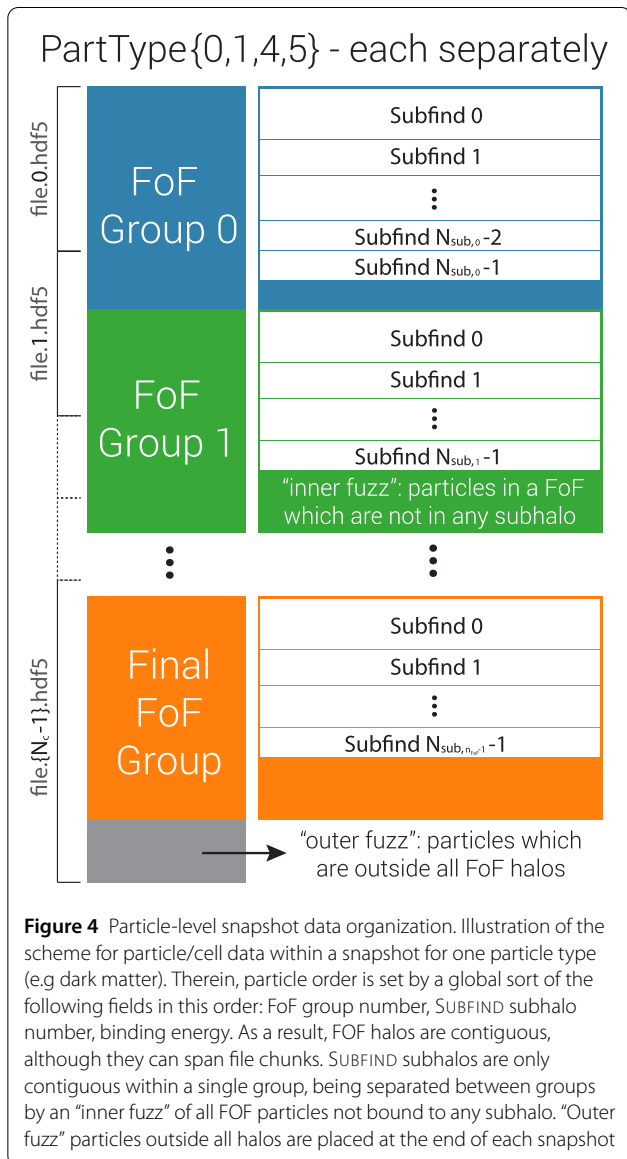


Figure 3 The variety of physical information accessible across different matter components of the TNG simulations. From top to bottom: dark matter density, gas density, gas velocity field, stellar mass density, gas temperature, gas-phase metallicity, shock mach number, magnetic field strength, and x-ray luminosity. Each panel shows the same $\sim 110 \times 14 \times 37$ Mpc volume of TNG100-1 at $z = 0$

Table 3 Abridged snapshot list for TNG runs: snapshot number together with the corresponding scalefactor and redshift. The twenty snapshots shown here are the ‘full’ snapshots, while the remaining eighty are ‘mini’ snapshots with a subset of fields

Snap	a	z	Snap	a	z
2	0.0769	12	33	0.3333	2
3	0.0833	11	40	0.4	1.5
4	0.0909	10	50	0.5	1
6	0.1	9	59	0.5882	0.7
8	0.1111	8	67	0.6667	0.5
11	0.125	7	72	0.7143	0.4
13	0.1429	6	78	0.7692	0.3
17	0.1667	5	84	0.8333	0.2
21	0.2	4	91	0.9091	0.1
25	0.25	3	99	1	0



lengths, $10^{10}M_{\odot}/h$ for masses, and km/s for velocities. Co-moving quantities can be converted to the corresponding physical ones by multiplying by the appropriate power of the scale factor a . New fields in TNG, not previously available in the original Illustris, are specially highlighted.

With respect to Illustris, the following new fields are generally available in the snapshots: (i) EnergyDissipation and Machnumber, giving the output of the on-the-fly shock finder, (ii) GFM_Metals, giving the individual element abundances of the nine tracked species (H, He, C, N, O, Ne, Mg, Si, Fe), (iii) GFM_MetalsTagged, metal tracking as described below, (iv) MagneticField and MagneticFieldDivergence, providing the primary result of the MHD solver.

3.1.3 Tagged metals

The units of all the entries of GFM_MetalsTagged field, except for NSNS, are the same as GFM_Metals: dimensionless mass ratios. Summing all elements of GFM_Metals heavier than Helium recovers the sum of the three tags SNIa + SNII + AGB. Likewise, the Fe entry of GFM_Metals roughly equals the sum of FeSNIa + FeSNII, modulo the small amount of iron consumed (i.e. negative contribution) by AGB winds. The fields are (in order):

- SNIa (0): The total metals ejected by Type Ia SN.
- SNII (1): The total metals ejected by Type II SN.
- AGB (2): the total metals ejected by stellar winds, which is dominated by AGB stars.
- NSNS (3): the total mass ejected from NS-NS merger events, which are modeled stochastically (i.e. no fractional events) with a DTD scheme similar to that used for SNIa, except with a different τ value. Note that the units of NSNS are arbitrary. To obtain physical values in units of solar masses, this field must be multiplied by α/α_0 where α is the desired mass ejected per NS-NS merger event, and α_0 is the base value (arbitrary) used in the simulation, e.g. Shen et al. (2015) take $\alpha = 0.05M_{\odot}$. The value of α_0 varies by run, and it is 0.05 for all TNG100 runs, and 5000.0 for all TNG300 and TNG50 runs. See Naiman et al. (2018) for more details.
- FeSNIa (4): The total iron ejected by Type Ia SN.
- FeSNII (5): The total iron ejected by Type II SN.

Note a somewhat subtle but fundamental detail: these tags do not isolate where a given heavy element was created, but rather identify the last star it was ejected from. This can be problematic since, for example, AGB winds create little iron, but eject a significant amount of iron which was previously created by SNIa and SNII at earlier epochs. The FeSNIa field is, for example, more accurately described as ‘the total iron ejected by type Ia supernovae not yet consumed and re-ejected from another star’

Table 4 Configuration of each of the subboxes for all three TNG volumes, including position and size within the periodic parent simulation, and description of the environment contained within

Subbox	Environment	Center Position [Code Units]	Box Size	f_{vol} [%]
TNG100 Subbox-0	Crowded, including a $5 \times 10^{13} M_{\odot}$ halo	(9000, 17,000, 63,000)	7.5 cMpc/h	0.1
TNG100 Subbox-1	Less crowded, including several $> 10^{12} M_{\odot}$ halos	(37,000, 43,500, 67,500)	7.5 cMpc/h	0.1
TNG300 Subbox-0	Massive cluster ($\sim 2 \times 10^{15} M_{\odot}$) merging at $z = 0$	(44, 49, 148) * 1000	15 cMpc/h	0.04
TNG300 Subbox-1	Crowded, above average # of halos above $10^{13} M_{\odot}$	(20, 175, 15) * 1000	15 cMpc/h	0.04
TNG300 Subbox-2	Semi-underdense, one local group analog at $z = 0$	(169, 97.9, 138) * 1000	10 cMpc/h	0.01
TNG50 Subbox-0	Somewhat-crowded (~ 6 MWs)	(26,000, 10,000, 26,500)	4.0 cMpc/h	0.15
TNG50 Subbox-1	Low-density, many dwarfs, no halos $> 5 \times 10^{10} M_{\odot}$	(12,500, 10,000, 22,500)	4.0 cMpc/h	0.15
TNG50 Subbox-2	Most massive cluster ($2 \times 10^{14} M_{\odot}$ at $z = 0$)	(7300, 24,500, 21,500)	5.0 cMpc/h	0.3

Table 5 Details of the subbox snapshots: the number and approximate time resolution Δt at three redshifts: $z = 6$, $z = 2$, and $z = 0$. Every subbox for a given volume and resolution combination has the same output times

Run	N_{snap}	$\Delta t_{(z=6)}$	$\Delta t_{(z=2)}$	$\Delta t_{(z=0)}$
TNG100-3	2431	~ 4 Myr	~ 7 Myr	~ 19 Myr
TNG100-2	4380	~ 2 Myr	~ 4 Myr	~ 10 Myr
TNG100-1	7908	~ 1 Myr	~ 1.5 Myr	~ 2.5 Myr
TNG300-3	2050	~ 6 Myr	~ 11 Myr	~ 8 Myr
TNG300-2	3045	~ 3 Myr	~ 6 Myr	~ 4 Myr
TNG300-1	2449	~ 1.5 Myr	~ 4 Myr	~ 6 Myr
TNG50-4	2333	~ 7 Myr	~ 6 Myr	~ 8 Myr
TNG50-3	4006	~ 2 Myr	~ 3 Myr	~ 4 Myr
TNG50-2	1895	~ 3 Myr	~ 6 Myr	~ 8 Myr
TNG50-1	~ 3600	~ 3 Myr	~ 2 Myr	~ 2 Myr

3.1.4 Subboxes

Separate ‘subbox’ cutouts exist for each baryonic run. These are spatial cutouts of fixed comoving size and fixed comoving coordinates, and the primary benefit is that their time resolution is significantly better than that of the main snapshots—details are provided in Tables 4 and 5. These snapshots are useful for some types of analysis and science questions requiring high time-resolution data, and for creating time-evolving visualizations. There are two subboxes for TNG100 (corresponding to the original Illustris subboxes #0 and #2, the latter increased in size), and three subboxes for TNG50 and TNG300. Note that subboxes, unlike full boxes, are not periodic.

The subboxes sample different areas of the large boxes, roughly described by the environment column in Table 4. The particle fields are all identical to the main snapshots, except that the particles/cells are not sorted by their group membership, since no group catalogs exist for subbox snapshots.

3.2 Group catalogs

Group catalogs give the results of substructure identification, and broadly contain two types of objects: dark matter halos (either FoF halos or central subhalos) and galaxies themselves (the inner stellar component of subhalos, either centrals or satellites). There is one group catalog

produced for each snapshot, which includes both FoF and SUBFIND objects. The group files are split into a small number of sub-files, just as with the raw snapshots. In TNG, these files are called FOF_SUBHALO_TAB_*, whereas in original Illustris they were called GROUPS_* (they are otherwise essentially identical). Every HDF5 group catalog contains the following groups: Header, Group, and Subhalo. The IDs of the member particles of each group/subgroup are not stored in the group catalog files. Instead, particles/cells in the snapshot files are ordered according to group membership.

In order to reduce confusion, we adopt the following terminology when referring to different types of objects. ‘Group’, ‘FoF Group’, and ‘FoF Halo’ all refer to halos. ‘Subgroup’, ‘Subhalo’, and ‘Subfind Group’ all refer to subhalos. The first (most massive) subgroup of each halo is the ‘Primary Subgroup’ or ‘Central Subgroup’. All other following subgroups within the same halo are ‘Secondary Subgroups’, or ‘Satellite Subgroups’.

FoF Groups. The Group fields are derived with a standard friends-of-friends (FoF) algorithm with linking length $b = 0.2$. The FoF algorithm is run on the dark matter particles, and the other types (gas, stars, BHs) are attached to the same groups as their nearest DM particle. **Subfind Groups.** The Subhalo fields are derived with the SUBFIND algorithm. In identifying gravitationally bound substructures the method considers all particle types and assigns them to subhalos as appropriate.

Complete documentation for the TNG group catalogs, comprising FoF halos as well as Subfind subhalos, is available online. Differences and additions with respect to original Illustris are highlighted.

3.3 Merger trees

Merger trees have been created for the TNG simulations using SUBLINK (Rodríguez-Gomez et al. 2015) and LHALOTREE (Springel et al. 2005). In the population average sense the different merger trees give similar results. In more detail, the exact merger history or mass assembly history for any given halo may differ. For a particular science goal, one type of tree may be more or less useful, and users are free to use whichever they prefer. We generally

recommend use of the `SUBLINK` trees as a first option, as they are more efficiently stored and accessible.

Trees can be ‘walked’, i.e. the descendants or progenitors of a given subhalo can be determined, thus linking objects across snapshots saved at different points in time. Main branches, such as the main progenitor branch (MPB), as well as full trees can be extracted. Examples of walking the tree are provided in the example scripts. For the technical details, algorithmic descriptions, and storage structures of the trees, please refer to Nelson et al. (2015) and the online documentation—we omit these details here.

3.3.1 *SubLink*

The `SUBLINK` merger tree is one large data structure split across several sequential HDF5 files named `TREE_EXTENDED.[FILENUM].HDF5`, where `[FILENUM]` goes from e.g. 0 to 19 for the TNG100-1 run, and 0 to 125 for the TNG300-1 run.

3.3.2 *LHaloTree*

The `LHALOTREE` merger tree is one large data structure split across several HDF5 files named `TREES_SF1_99.[CHUNKNUM].HDF5`, where TNG100-1 has for instance 80 chunks enumerated by `[CHUNKNUM]`, while TNG300-1 has 320. Within each file there are a number of HDF5 groups named “TreeX”, each of which represents one disjoint merger tree.

3.3.3 *Offsets files*

As described above, snapshot particle data is ordered by the subhalo each particle belongs to. To facilitate rapid loading of snapshot data, particle ‘offset’ numbers provide the location where particles belonging to each subhalo begin. Most simply, offsets describe where in the group catalog files to find a specific halo/subhalo, and where in the snapshot files to find the start of the particles of a given halo/subhalo.

To use the helper scripts (provided online) for working with the actual data files (snapshots or group catalogs) on a local machine, then it *is required* to download the offset file(s) for the snapshot(s) of interest. The offsets are *not required* when using the web-based API or analyzing the particle cutouts it provides, for instance.

Note that in Illustris, offsets were embedded inside the group catalog files for convenience. In TNG however, we have kept offsets as separate files called `OFFSETS_*.HDF5` (one per snapshot), which must be downloaded as well.

3.3.4 *The ‘simulation.hdf5’ file*

Each run has a single file called ‘simulation.hdf5’ which is purely optional, for convenience, and not required by any of the public scripts. Its purpose is to encapsulate all data of an entire simulation into a single file.

To accomplish this, we make advantage of a new feature of the HDF5 library called “virtual datasets”. A virtual

dataset is a collection of symbolic links to one or more datasets in other HDF5 file(s), where these symlinks can refer to subsets of a dataset, in either the source or target of the link. The `simulation.hdf5` is thus a large collection of links, which refer to other files which actually contain data. In order to use it, the corresponding files must also be downloaded (e.g. of snapshots, group catalogs, or supplementary data catalogs).

Using this resource, the division of snapshots and group catalogs over multiple file chunks is no longer relevant. Loading particle data from snapshots or subhalo or halo fields from group catalogs become one line operations. It also makes loading the particles of a given halo or subhalo using the offset information trivial. Finally, supplementary data catalogs (either those we provide, or similar user-run computations) can be ‘virtually’ inserted as datasets in snapshots or group catalogs. This provides a clean way to organize post-processing computations which produce additional values for halos, subhalos, or individual particles/cells. Such data can then be loaded with the same scripts (and same syntax) as ‘original’ snapshot/group catalog fields.

We refer to the online documentation for examples of each use case as well as technical requirements, namely a relatively new version (1.10+) of the HDF5 library.

3.4 Initial conditions

We provide as part of this release the initial conditions for all TNG volumes as well as the original Illustris volumes. These were created with the Zeldovich approximation and the N-GENIC code (Springel 2015). Each particular realization was chosen from among tens of random realizations of the same volume as the most average, based on inspection of the $z = 0$ power spectrum and/or dark matter halo mass function—see Vogelsberger et al. (2014a) and Pillepich et al. (in prep) for details. Each IC is a single HDF5 file with self-evident structure: the coordinates, velocities, and IDs of the set of total matter particles at $z = 127$, the starting redshift for all runs. These ICs were used as is for dark-matter only simulations, while for baryonic runs total matter particles were split upon initialization in the usual way, into dark matter and gas, according to the cosmic baryon fraction and offsetting in space by half the mean interparticle spacing. These ICs can be run by e.g. `GADGET` or `AREPO` as is, or easily converted into other data formats.

3.5 Supplementary data catalogs

Many additional data products have been computed in post-processing, based on the raw simulation outputs. These are typically in support of specific projects and analysis in a published paper, after which the author makes the underlying data catalog public. Many such catalogs have been made available for the original Illustris simulation, and the majority of these will also be recalculated

for TNG. We provide a list of TNG supplementary data catalogs which are now available or which we anticipate to release in the near future:

- (A) Tracer Tracks—time-evolution of Monte Carlo tracer properties for TNG100 (Nelson et al. in prep).
- (B) Stellar Mass, Star Formation Rates—multi-aperture and resolution corrected masses, time-averaged SFRs (Pillepich et al. 2018a).
- (C) Stellar Circularities, Angular Momenta, and Axis Ratios—for the stellar components of galaxies, as for Illustris (Genel et al. 2015).
- (D) Subhalo Matching Between Runs—cross-matching subhalos between baryonic and dark-matter only runs, between runs at different resolutions, and between TNG100 and Illustris (Lovell et al. 2018; Nelson et al. 2015; Rodriguez-Gomez et al. 2015; Rodriguez-Gomez et al. 2017).
- (E) Stellar Projected Sizes—half-light radii of TNG100 galaxies (Genel et al. 2018).
- (F) Black hole Mergers and Details—records of BH-BH mergers and high time-resolution BH details, as for Illustris (Kelley et al. 2017; Blecha et al. 2016), and with an updated approach (Katz et al. in prep).
- (G) Stellar Assembly—in-situ versus ex-situ stellar growth, as for Illustris (Rodriguez-Gomez et al. 2016, 2017).
- (H) Subbox Subhalo List—record of which subhalos exist in what subboxes across particular redshift ranges, and interpolated properties (Nelson et al. 2019a).
- (I) Molecular and Atomic Hydrogen (HI + H₂)—decomposition of the neutral hydrogen in gas cells and galaxies into HI/H₂ masses (Diemer et al. 2018; Stevens et al. 2018).
- (J) Halo/galaxy angular momentum and baryon content—measurements of spherical overdensity values, as for Illustris (Zjupa and Springel 2017).
- (K) SDSS Photometry and Mock Fiber Spectra—broadband colors and spectral mocks including dust attenuation effects (Nelson et al. 2018a).
- (L) SKIRT Synthetic Images and Optical Morphologies—dust radiative-transfer calculations using SKIRT to obtain broadband images, automated morphological measurements (Rodriguez-Gomez et al. 2019).
- (M) DisPerSE Cosmic Web—topological classification of the volume into sheets, filaments, nodes, and voids (Duckworth et al. in prep).
- (N) Particle-level lightcones—in a variety of configurations, from small field of view ‘deep fields’ to all-sky projections, across the different matter components, to facilitate lensing, *x*-ray, Sunyaev-Zeldovich, and related explorations (Giocoli et al. in prep).

Several of these were previously available for the original Illustris simulation and will be re-computed for TNG. We would plan to provide a number of ‘pre-defined’ galaxy samples, particularly with respect to common observational selection techniques, current and/or upcoming surveys, and other distinct classes of interest. This can include, for example, red versus blue galaxies, luminous red galaxies (LRGs) and emission-line galaxies (ELGs) of SDSS, damped Lyman-alpha (DLA) host halos, and ultra-diffuse or low surface brightness (LSB) galaxies. Such samples would facilitate rapid comparisons to certain types of observational samples, and can be included as supplementary data catalogs as they become available.

4 Data access

There are three complementary ways to access and analyze TNG data products.

- 1 (*Local data, local analysis*). Raw files can be directly downloaded, and example scripts are provided as a starting point for local analysis.
- 2 (*Remote data, local analysis*). The web-based API can be used, either through a web browser or programmatically in a script, to perform search, data extraction, and visualization tasks, followed by local analysis of these derivative products.
- 3 (*Remote data, remote analysis*). A web-based JupyterLab (or Jupyter notebook) session can be instantiated to explore the data, develop analysis scripts with persistent storage, run data-intensive and compute-intensive tasks, and make final plots for publication.

These different approaches can be combined. For example, by downloading the full redshift zero group catalog to perform a complex search which cannot be easily done with the API. After determining a sample of interesting galaxies (i.e. a set of subhalo IDs), one can then extract their individual merger trees (and/or raw particle data) without needing to download the full simulation merger tree (or a full snapshot).

These approaches are described below, while “getting started” tutorials for several languages (currently: Python, IDL, and Matlab) can be found online.

4.1 Direct file download and example scripts

Local data, local analysis. All of the primary outputs of the TNG simulations are released in HDF5 format, which we use universally for all data products. This is a portable, self-describing, binary specification (similar to FITS), suitable for large numerical datasets. File access libraries, routines, and examples are available in all common computing languages. We typically use only basic features of the format: attributes for metadata, groups for organization, and large datasets containing one and two dimensional numeric arrays. To maintain reasonable filesizes for transfer,

most outputs are split across multiple files called “chunks”. For example, each snapshot of TNG100-1 is split into 448 sequentially numbered chunks. Links to the individual file chunks for a given simulation snapshot or group catalog are available under their respective pages on the main data release page.

The provided example scripts (in IDL, Python, and Matlab) give basic I/O functionality, and we expect they will serve as a useful starting point for writing any analysis task, and intend them as a ‘minimal working examples’ which are short and simple enough that they can be quickly understood and extended. For a getting-started guide and reference, see the online documentation.

4.2 Web-based API

Remote data, local analysis. For TNG we enhance the web-based interface (API) introduced with the original Illustris simulation, augmented by a number of new features and more sophisticated functionality. At its core, the API can respond to a variety of user requests and queries. It provides a well-defined interface between the user and simulation data, and the tools it provides are independent, as much as possible, from any underlying details of data structure, heterogeneity, storage format, and so on. The API can be used as an alternative to downloading large data files for local analysis. Fundamentally, the API allows a user to *search*, *extract*, *visualize*, or *analyze* a simulation, a snapshot, a group catalog, or a particular galaxy/halo. By way of example, the following requests can be handled by the current API:

- Search across subhalos with numeric range(s) over any field(s) present in the Subfind catalogs.
- Retrieve a snapshot cutout for all the particles/cells within a given subhalo/halo, optionally restricted to a subset of specified particle/cell type(s) and fields(s).
- Retrieve the complete merger history or main branches for a given subhalo.
- Download subsets of snapshot files, containing only specified particle/cell type(s), and/or specific field(s) for each type.
- Traverse links between halos and subhalos, for instance from a satellite galaxy, to its parent FoF halo, to the primary (central) subhalo of that group, as well as merger tree progenitor/descendant connections.
- Render visualizations of any field(s) of different components (e.g. dark matter, gas, stars) of a particular halo/subhalo.^[d]
- Download actual data from such a halo/subhalo visualization, e.g. maps of projected gas density, O VI column density, or stellar luminosity in a given band.^[d]
- Render a static visualization of the complete merger tree (assembly history) of any subhalo.^[d]
- Plot the relationship between quantities in the group catalogs, e.g. fundamental scaling relations such as the star-forming main sequence of TNG.^[d]

- Plot tertiary relationships between group catalog quantities, e.g. the dependence of gas fraction on offset from the main sequence.^[d]

The IllustrisTNG data access API is available at the following permanent URL:

<http://www.tng-project.org/api/>

For a getting-started guide for the API, as well as a cookbook of common examples and the complete reference, see the online documentation.

4.3 Remote data analysis

Remote data, remote analysis. Coincident with the TNG public data release we introduce a new, third option for working with and analyzing large simulation datasets. Namely, an online, browser-based scientific computing environment which enables researchers’ computations to “be brought to” the data. It is similar in spirit to the NOAO Data Lab (Fitzpatrick et al. 2014) and SciServer services (Raddick et al. 2017), i.e. simultaneously hosting petabyte-scale datasets as well as a full-featured analysis platform and toolset. This alleviates the need to download any data, or run any calculations locally, thereby facilitating broad, universal, open access to large cosmological simulation datasets such as TNG.

To enable this functionality we make use of extensive development on Jupyter and JupyterLab over the last few years. JupyterLab is the evolution of the Jupyter Notebook (Kluyver et al. 2016), previously called IPython (Pérez and Granger 2007). It is a next-generation, web-based user interface suitable for scientific data analysis. In addition to the previous ‘notebook’ format, JupyterLab also enables a traditional workflow based around a collection of scripts on a filesystem, text editors, a console, and command-line execution. It provides an experience nearly indistinguishable from working directly on a remote computing cluster via SSH.

Computation is language agnostic, as ‘kernels’ are supported in all common languages, including Python 2.x, Python 3.x, IDL, Matlab, R, and Julia. Development, visualization, and analysis in any language or environment practically available within a Linux environment is possible, although we focus at present on Python 3.x support.

Practically, this service enables direct access to one of the complete mirrors of the Illustris[TNG] data, which is hosted at the Max Planck Computing and Data Facility (MPCDF) in Germany. Users can request a new, on-demand JupyterLab instance, which is launched on a system at MPCDF and connected to the user web browser. All Illustris[TNG] data is then directly available for analysis. A small amount of persistent user storage is provided, so that under-development scripts, intermediate analysis outputs, and in-progress figures for publication all persist across sessions. Users can log out and pick up later where they left off. A base computing environment is provided,

which can be customized as needed (e.g. by installing new python packages with either `pip` or `conda`). Users can synchronize their pre-existing tools, such as analysis scripts, with standard approaches (`git`, `hg`, `rsync`) or via the JupyterLab interface. Results, such as figures or data files, can be viewed in the browser or downloaded back to the client machine with the same tools.

For security and resource allocation, users must specifically request access to the JupyterLab TNG service. At present we anticipate providing this service on an experimental (beta) basis, and only to active academic users.

4.4 Further online tools

4.4.1 Subhalo search form

We provide the same, simple search form to query the subhalo database as was available in the Illustris data release. It exposes the search capabilities of the API in a user-friendly interface, enabling quick exploration without the need to write a script or URL by hand. As examples, objects can be selected based on total mass, stellar mass, star formation rate, or gas metallicity. The tabular output lists the subhalos which match the search, along with their properties. In addition, each result contains links to a common set of API endpoints and web-based tools for inspection and visualization.

4.4.2 Explore: 2D and 3D

The 2D Explorer and 3D Explorer interfaces are experiments in the *interactive* visualization and exploration of large data sets such as those generated by the IllustrisTNG simulations. They both leverage the approach of thin-client interaction with derived data products. The 2D Explorer exposes a Google Maps-like tile viewer of pre-computed imagery from a slice of the TNG300-1 simulation at redshift zero, similar to the original Illustris explorer. Multiple views of different particle types (gas, stars, dark matter, and black holes) can be toggled and overlaid, which is particularly useful in exploring the spatial relationships between different phenomena of these four matter components.

The 3D Explorer introduces a new interface, showing a highly derivative (although spatially complete) view of an entire snapshot. That is, instead of particle-level information, we facilitate interactive exploration of the group catalog output in three-dimensional space. This allows users to rotate, zoom, and move around the cubic box representing the simulation domain, where the largest dark matter halos are represented by wireframe spheres of size equal to their virial radii, while the remaining smaller halos are represented by points. User selection of a particular halo, via on-click ray cast and sphere intersection testing, launches an API query and returns the relevant halo information and further introspection links. At present, both Explorers remain largely proof of concept interfaces for how tighter in-

tegration of numeric, tabular, and visual data analysis components may be combined in the future for the effective exploration and analysis of large cosmological datasets (see also Dykes et al. 2018, and the Dark Sky simulation).

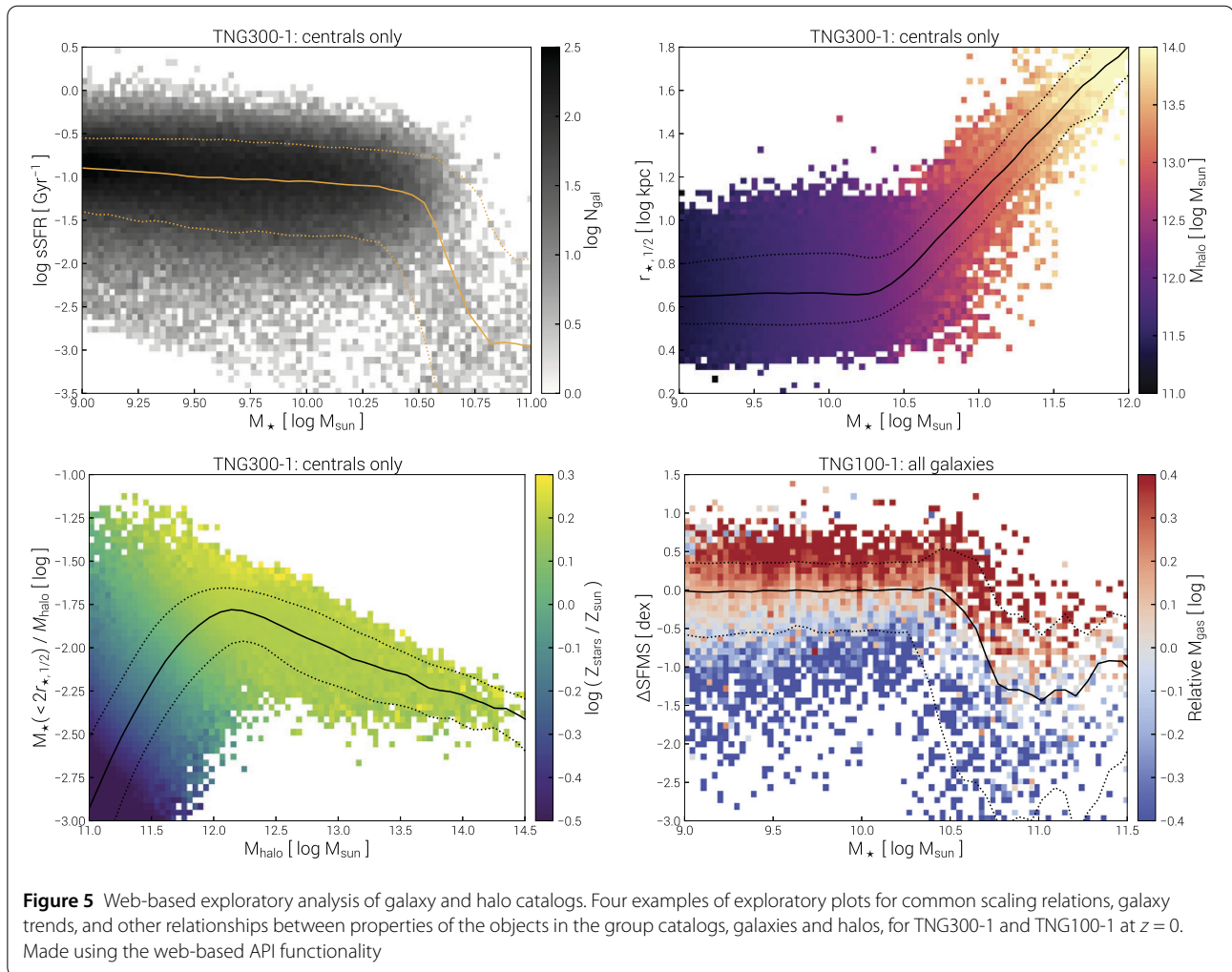
4.4.3 Merger tree visualization

In the Illustris data release we demonstrated a rich-client application built on top of the API, in the form of an interactive visualization of merger trees. The tree is vector based, and client side, so each node can be interacted with individually. The informational popup provides a link, back into the API, where the details of the selected progenitor subhalo can be interrogated. This functionality is likewise available for all new simulations. Furthermore, we have added a new, static visualization of the complete merger tree of a subhalo. This allows a quick overview of the assembly history of a given object, particularly its past merger events and its path towards quiescence. In the fiducial configuration, node size in the tree is scaled with the logarithm of total halo mass, while color is mapped to instantaneous sSFR.

4.4.4 Plot group catalog

The first significant new feature of the API for the TNG public data release is a plotting routine to examine the group catalogs. Since the objects in the catalogs are either galaxies or dark matter halos, plotting the relationships among their various quantities is one of most fundamental explorations of cosmological simulations. Classically observed scaling relations, such as Tully–Fisher (rotation velocity vs. stellar mass), Faber–Jackson (stellar velocity dispersion vs. luminosity), the stellar size-mass relation, the star-formation main sequence, or the Magorrian relation (black hole mass versus bulge mass) are all available herein. Such relations can be used to assess the outcome of the simulations by comparison to observational data. More complex relations, those involving currently unobserved properties of galaxies/halos, and/or those only currently observed with very limited statistics or over limited dynamic range, represent a powerful discovery space and predictive regime for simulations such as TNG. At the level of the galaxy (or halo) population, i.e. with tens to hundreds of thousands of simulated objects, many such relationships reveal details of the process of galaxy formation and evolution, as well as the working mechanisms of the physical/numerical models.

The ‘group catalog plotter’ is an API endpoint which returns publication quality figures (e.g. PNG or PDF outputs). In Fig. 5 we show several examples of its output, taken from TNG300-1 and TNG100-1 at $z = 0$. Many options exist to control the behavior and structure of the plots, all of which are detailed in the online documentation. As for the subhalo search form, we also provide



a new web-based interface to assist in interactively constructing plots from this service. Fundamentally, the quantities to be plotted against each other on the x - and y -axes can be selected. In this case, a two-dimensional histogram showing the density of subhalos in this space is overlaid with the median relation and bounding percentiles. Optionally, a third quantity can be added, which is then used to color every pixel in the histogram according to a user-defined statistic (e.g. median) of all the objects in that bin. For example, plotting the stellar-mass halo-mass relation, colored by stellar metallicity, reveals one reason for the scatter in this relation. This third quantity can optionally be normalized relative to the median value at each x -axis value (e.g. as a function of stellar mass), highlighting the ‘relative’ deviation of that property compared to its evolving median value. The types of subhalos included can be chosen, for example selecting only centrals or only satellite galaxies, and the subhalos to be included can be filtered based on numeric range selections on a fourth quantity. We expect that this tool will enable rapid, ini-

tial exploration of interesting relationships among galaxy (or halo) integral properties, and serve as a starting point for more in-depth analysis (see also de Souza and Ciardi 2015). Complete usage documentation is available online.

4.4.5 Visualize galaxies and halos

The second significant new feature of the API for the TNG public data release is an on-demand visualization service. Primarily, this API endpoint renders projections of particle-level quantities (of gas cells, dark matter particles, or stellar particles) for a given subhalo or halo. For example, it can produce gas column density projections, gas temperature projections, stellar line-of-sight velocity maps, or dark matter velocity dispersion maps. Its main rendering algorithm is based on the standard SPH kernel projection technique, with adaptive kernel sizes for all particle types, although alternatives are available. In Fig. 6 we show several examples of output, at both the halo-scale (circle indicating virial radius), and the galaxy

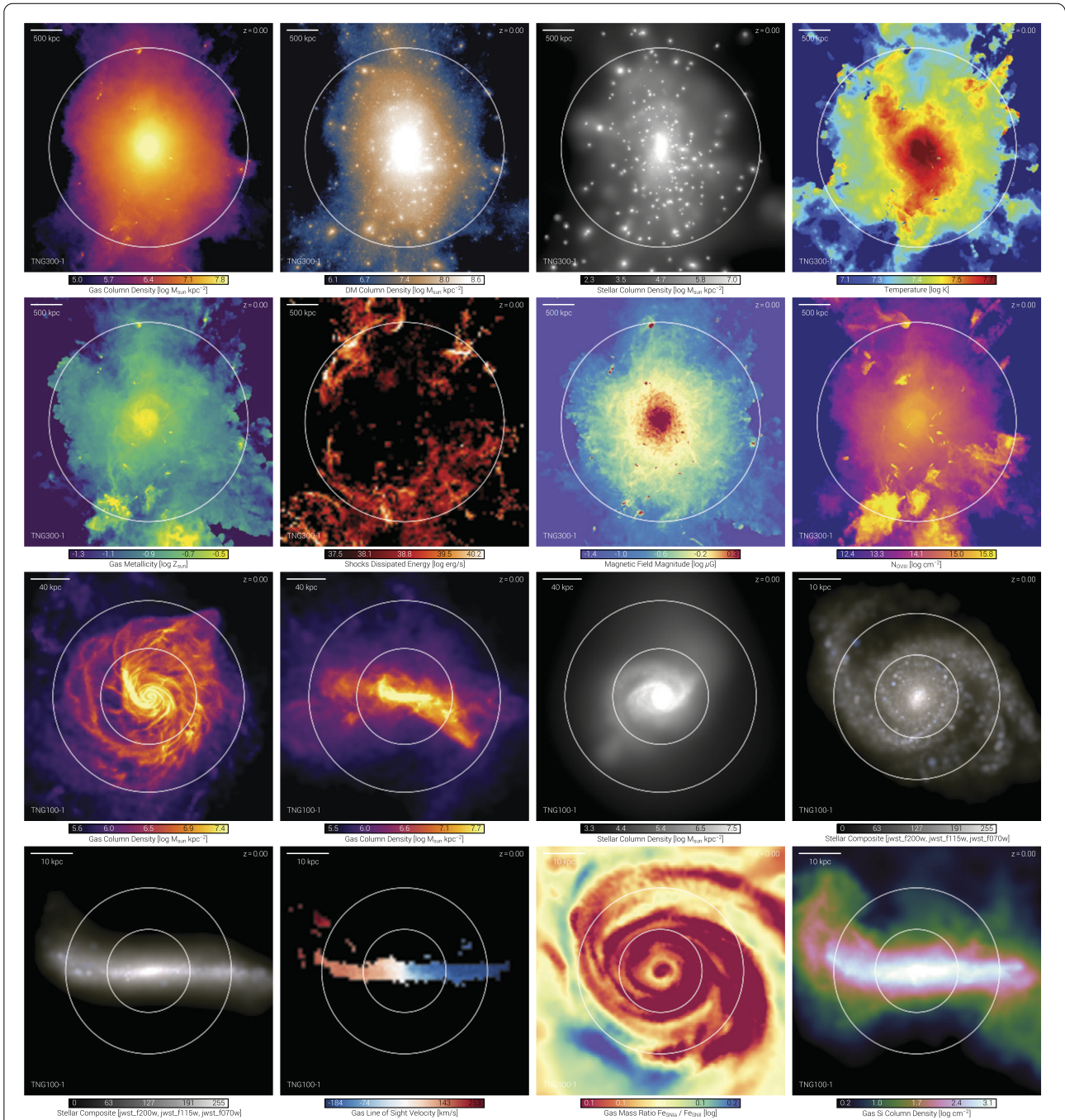


Figure 6 Web-based visualization of galaxies and their halos. Example of some possible halo-scale and galaxy-scale visualizations from TNG300-1 and TNG100-1, made using the web-based API functionality, viewing the dark matter, gas, and stars. The top eight panels show the 20th most massive halo of TNG300-1 at $z = 0$ (circle indicating r_{vir}). The bottom eight panels show face-on and edge-on views of subhalo 468590 of TNG100-1 at $z = 0$ (circles indicating $r_{1/2,\star}$ and $2r_{1/2,\star}$)

scale (outer circle showing twice the stellar half mass radius).

The visualization service returns publication quality figures (e.g. PNG or PDF outputs). It can also return the raw data used to construct any image, in scientifically accurate

units (HDF5 output). For instance, a user can request not only an image of the gas density projection of an ongoing galaxy merger, but also the actual grid of density values in units of e.g. $M_{\odot} \text{kpc}^{-2}$. Numerous options exist to control the behavior of the rendered projections, as well as the out-

put style, all of which are detailed in the online documentation. All parameters of the rendering can be specified—as an example, the view direction can be a rotation into face-on or edge-on orientations. Most properties available in the snapshots can be visualized, for any galaxy/halo, at any snapshot, for any run.

Beyond snapshot level information, the visualization service currently has two more advanced features. First, it is coupled to the CLOUDY photoionization code (Ferland et al. 2017), following Nelson et al. (2018b). This enables ionic abundance calculations for gas cells on the fly. For example, a user can request a column density map of the O VI or C IV ions. All relevant atoms are supported, assuming solar abundances for non-tracked elements, typically up to the tenth ionization state (Al, Ar, Be, B, Ca, C, Cl, Cr, Co, Cu, F, He, H, Fe, Li, Mg, Mn, Ne, Ni, N, O, P, K, Sc, Si, Na, S, Ti, V, Zn). Emission line luminosities are also available—a surface brightness map of metal-line emission from O VIII at 22.1012\AA , for example. Secondly, this service is also coupled to the FSPS stellar population synthesis code (Conroy et al. 2009; Conroy and Gunn 2010) through PYTHON-FSPS (Foreman-Mackey et al. 2014), following Nelson et al. (2018a). This enables emergent stellar light calculations for stellar population particles on the fly, with optional treatments of dust attenuation effects. For example, a user can request a map of stellar surface brightness, or luminosity, either rest frame or observed frame, for any of the ~ 140 available bands, including common filters on surveys/instruments such as SDSS, DES, HST, and JWST.

We expect that this tool will enable rapid, initial exploration of many interesting facets of galaxies and halos across the simulations, and serve as a starting point for more in-depth analysis. We caution that, used improperly, this tool can easily return nonsensical results (e.g. requesting OI emission properties from ISM gas), and users should understand the relevant scientific limitations. In this particular case, we refer to the effective two-phase ISM model used in TNG (Springel and Hernquist 2003) which intentionally avoids resolving the cold, dense phases of the ISM. Complete usage documentation is available online.

5 Scientific remarks and cautions

In the original Illustris simulation we identified a number of non-trivial issues in the simulated galaxy and halo populations in comparison to observational constraints (see Nelson et al. 2015, for a summary). These disagreements motivated a series of important caveats against drawing certain strong scientific conclusions in a number of regimes.

In contrast, our initial explorations of TNG (specifically, of the TNG100-1 and TNG300-1 simulations) have revealed no comparably significant tensions with respect

to observable comparisons. With this data release we invite further detailed observational comparisons and scrutiny. The TNG simulations have been shown to realistically resolve numerous aspects of galactic structure and evolution, including many internal properties of galaxies (though, clearly, not all) as well as their co-evolution within the cosmic web of large-scale structure (see Sect. 2.2). TNG reproduces various observational details and scaling relations of the demographics and properties of the galaxy population, not only at the present epoch ($z = 0$), but also at earlier times (see likewise Sect. 2.2). This has been achieved with a physically plausible although necessarily simplified galaxy formation model. The TNG model is intended to account for most, if not all, of the primary processes that are believed to be important for the formation and evolution of galaxies.

5.1 IllustrisTNG: possible observational tensions

We therefore do not specifically caution against the use of TNG in any of the regimes where the original Illustris simulation was found to be less robust. However, the enormous spatial and temporal dynamic range of these simulations, as well as the multi-scale, multi-physics nature of the complex physical phenomena involved, implies modeling approximations and uncertainties. Early comparisons of TNG against observations have identified a number of interesting regimes in which possible tensions exist.

Our ability to make any stronger statement is frequently limited by the complexity of the observational comparison, i.e. the need to accurately reproduce (or ‘mock’) the observational measurement closely and with care. In the qualitative sense, however, these regimes may plausibly indicate areas where the TNG model has shortcomings or is less physically realistic. It will be helpful for any user of the public data to be aware of these points, which should be carefully considered when advancing strong scientific conclusions or making claims based on observational comparisons. Possible tensions of interest include the following:

- (I) The simulated stars in Milky Way-like galaxies are too alpha-enhanced in comparison to observations of the Milky Way (Naiman et al. 2018).
- (II) The Eddington ratio distributions of massive black holes ($>10^9 M_{\odot}$) at $z = 0$ are dominated in TNG by low accretion rates, generically far below the Eddington limit; recent observations favor at least some fraction of higher accretion rate massive black holes. This is reflected in a steeper hard X-ray AGN luminosity function at $1 \lesssim z \lesssim 4$ (Habouzit et al. 2018).

- (III) TNG galaxies may have a weaker connection between galaxy morphology and color than observed at $z \sim 0$, reflected in a possible excess of red disk-like galaxies in the simulations (Rodríguez-Gomez et al. 2019), although see Tachella et al. (in prep).
- (IV) TNG galaxies exhibit a somewhat sharper trend than observations in quenched fraction vs. galaxy stellar mass for $M_* \in 10^{10-11} M_\odot$ (Donnari et al. 2018), and similarly in the relation between sSFR and M_{BH} at low redshift (Terrazas et al. in prep).
- (V) The locus of the galaxy star-forming main sequence is below the face-value observed SFMS at $1 \lesssim z \lesssim 2$, modulo known inconsistencies with e.g. the observed stellar mass function (Donnari et al. 2018).
- (VI) Similarly, the H_2 mass content of massive TNG galaxies at $z = 1 - 3$ may be lower than implied by ALMA observations (Popping et al. 2019) and sub-mm galaxy demographics (Hayward et al. in prep).
- (VII) The DM fractions within massive elliptical galaxies at $z = 0$ are consistent with observations at large galactocentric distances, but may be too high within their effective radii (Lovell et al. 2018) and likewise are tentatively higher than values inferred from observations of massive $z = 2$ star forming galaxies (Lovell et al. 2018 and Übler et al. in prep).

With respect to points (III)–(IV) there is, in general, an interesting transitional mass regime (galaxy stellar mass $\sim 10^{10.5} M_\odot$) where central blue vs. red galaxies or star-forming vs. quiescent galaxies co-exist: this reflects the effective quenching mechanism of the TNG model based on SMBH feedback (Nelson et al. 2018a; Weinberger et al. 2018) but how precisely such transitional galaxies differ also in other structural and kinematical properties still requires careful examination and consideration.

Note that for the items in this list we have not included more specific quantification of observed tension (i.e. χ^2 or fractional deviation values)—the referenced papers contain more discussion. On the one side, not all observational results are in agreement among each other, making quantitative statements necessarily partial; nor observational statements of different galaxy properties are necessarily consistent within one another, especially across cosmic times. On the other side, excruciating care is often necessary to properly map simulated variables into observationally-derived quantities.

5.2 Numerical considerations and issues

To better inform which features of the simulations are robust when making science conclusions, we note the following points related to numerical considerations:

1. SubhaloFlag. Not all objects in the Subfind group catalogs should be considered ‘galaxies’. In particular, not all satellite subhalos have a cosmological origin, in the sense that they may not have formed due to the process of structure formation and collapse. Rather, some satellite subhalos will represent fragments or clumps produced through baryonic processes (e.g. disk instabilities) in already formed galaxies, and the Subfind algorithm cannot *a priori* differentiate between these two cases. Such non-cosmological objects are typically low mass, small in size, and baryon dominated (i.e. with little or no dark matter), residing at small galactocentric distances from their host halos, preferentially at late times ($z < 1$). These objects may appear as outliers in scatter plots of typical galaxy scaling relations, and should be considered with care.

We have added a SUBHALOFLAG field to the group catalogs to assist in their identification, which was constructed as follows. First, a variant of the SubLink merger tree was used which tracks baryonic, rather than dark matter, particles—namely, star-forming gas cells and stars (Rodríguez-Gomez et al. 2015). The algorithm is otherwise the same, with the same weighting scheme for determining descendants/progenitors, except that this “SubLink-Gal” tree allows us to track subhalos which contain little or no dark matter.

Then, we flag a subhalo as non-cosmological if all the following criteria are true: (i) the subhalo is a satellite at its time of formation, (ii) it forms within 1.0 virial radii of its parent halo, and (iii) its dark matter fraction, defined as the ratio of dark matter mass to total subhalo mass, at the time of formation of the subhalo, is less than 0.8.

These are relatively conservative choices, implying a low false positive rate. On the other hand, some spurious subhalos may not be flagged under this definition. A much more aggressive criterion would be to flag a subhalo if its instantaneous dark matter fraction is low, e.g. less than 10% (as used in e.g. Genel et al. 2018; Pillepich et al. 2018b). Such a selection will result in a purer sample, with less contaminating subhalos, but will also exclude more genuine galaxies, such as those which have undergone extensive (i.e. physical) stripping of their dark matter component during infall. We encourage users to enforce the provided SubhaloFlag values as a default, but to carefully consider the implications and details, particularly for analyses focused on satellite galaxy populations or dark-matter deficient systems.

2. Gas InternalEnergy Corrections. In all TNG simulations, the time-variable UV-background radiation field (Faucher-Giguère et al. 2009, FG11 version) is enabled only for $z < 6$. Therefore, the ionization state of the IGM above redshift six should be studied with caution, as the usual density-temperature relation will not be present.

Two further technical issues exist for the original INTERNALENERGY field (i.e. gas temperature) of all TNG simulations. These have been corrected in post-processing, as described below, and the fiducial INTERNALENERGY field of all snapshots in all TNG simulations has been rewritten with updated values. The original dataset has been renamed to INTERNALENERGYOLD for reference, although we do not recommend its use for any purpose.

The first issue is seen in the low-density, low-temperature regime of the intergalactic medium (IGM). Here, due to a numerical issue in the TNG codebase related to the Hubble flow across gas cells, spurious energy injection could occur in underdense gas. In practice, this only affects extremely low density IGM gas in equilibrium between adiabatic cooling and photoheating from the background. The result is a slight upwards curvature in the usual (ρ, T) phase diagram. To correct this issue, we have used one of the TNG model variant boxes (with side length 25 Mpc/h and 512^3 resolution) which includes the fix for this issue. The adiabat was then identified in all TNG runs as well as in the corrected simulation by binning the density-temperature phase diagram and locating the temperature of peak gas mass occupancy as a function of density. A multiplicative correction f_{corr} , taken as the ratio between the corrected and uncorrected linear gas temperatures, is then defined and applied as a function of density, for gas with physical hydrogen number density $< 10^{-6} a^{-3} \text{ cm}^{-3}$. We further restrict the correction to the low-temperature IGM by smoothly damping f_{corr} to unity by $10^{5.0} \text{ K}$ as $\log T_{\text{corr}} = \log T_{\text{orig}} + \log(f_{\text{corr}})w(T_{\text{orig}})$ with the window function $w(T_{\text{orig}}) = 1 - [\tanh(5(T_{\text{orig}} - 5.0)) + 1]/2$. This issue manifests only towards low redshift, and for simplicity and clarity we apply this correction only for $z \leq 5$ (snapshots 17 and later).

The second issue arises for a very small fraction of low-temperature gas cells with $T < 10^4 \text{ K}$, the putative cooling floor of the model. Here, due to a numerical issue in the TNG codebase related to the cosmological scaling of the energy source term in the Powell divergence cleaning of the MHD scheme (right-most term in Eqn. 21 of Pakmor and Springel 2013), spurious cooling could occur in gas with high bulk velocity and large, local divergence error ($|\nabla \cdot \vec{B}| > 0$). In practice, this affects a negligible number of cells which appear in the usual (ρ, T) phase diagram with temperatures less than 10,000 K and densities between the star-formation threshold and four orders of magnitude lower. To correct this issue we simply update the gas temperature values, for all cells in this density range with $\log(T[\text{K}]) < 3.9$, to the cooling floor value of 10^4 K , near the background equilibrium value. As this issue also manifests only towards low redshift (being most

problematic at intermediate redshifts $1 \lesssim z \lesssim 4$), we likewise apply this correction only for $z \leq 5$ (snapshots 17 and later).

Note that for both issues, we have verified in reruns of smaller volume simulations, by applying the fix in correspondingly corrected TNG model variant simulations, that no properties of galaxies or of the galaxy population are noticeably affected by these fixes.

3. *Unresolved ISM.* The multi-phase model of the interstellar medium in TNG (which is the same as in Illustris) is a necessarily coarse approximation of a complex physical regime. In particular, the cold neutral and molecular phases of the ISM are not resolved in the current generation of cosmological simulations like TNG; giant molecular clouds (GMCs) and the individual birth sites of massive star formation and, for example, the resultant nebular excitation is likewise not explicitly resolved. Modeling observables which arise in dense ISM phases (e.g. CO masses) should be undertaken with care.

The modeling of the star formation process is explicitly designed to reproduce the empirical Kennicutt-Schmidt relation, so the correlation between star formation rate and gas density, at the scale where this scaling is invoked, is not a predictive result. Star formation, as in all computational models of galaxy formation, proceeds at a numerical threshold density which is many orders of magnitude lower than the true density at which stars form. This threshold is $n_{\text{H}} \simeq 0.1 \text{ cm}^{-3}$ in TNG, which may have consequences for the spatial clustering of young stars, as one example (Buck et al. 2018).

4. *Resolution Convergence.* Numerical convergence is a complex issue, and working with simulations at multiple resolutions can be challenging. Analysis which includes more than one TNG box at once (e.g. TNG100 and TNG300 together), or explicitly uses multiple realizations at different resolutions should carefully consider the issue of convergence. The degree to which various properties of galaxies or the simulation as a whole is converged depends on the specific property, as well as the mass regime, redshift, and so on. For example, see Pillepich et al. (2018b) for convergence of the stellar mass functions of TNG100 and TNG300, and details on a simple ‘resolution correction’ procedure which may be desirable to apply, particularly when combining the results of multiple flagship boxes together into a single analysis.

6 Community considerations

6.1 Citation request

To support proper attribution, recognize the effort of individuals involved, and monitor ongoing usage and impact, the following is requested. Any publication making use of data from the TNG100/TNG300 simulations should cite this release paper (Nelson et al. 2019b) as well as the five works from the ‘introductory paper series’ of TNG100/300, the order being random:

- Pillepich et al. (2018a) (stellar contents),
- Springel et al. (2018) (clustering),
- Nelson et al. (2018a) (colors),
- Naiman et al. (2018) (chemical enrichment),
- Marinacci et al. (2018) (magnetic fields).

Any publication making use of the data from the TNG50 simulation should cite this release paper, as well as the two introductory papers of TNG50, the order being random:

- Nelson et al. (2019a) (outflows),
- Pillepich et al. (2019) (structure & kinematics).

Finally, use of any of the supplementary data products should include the relevant citation. A full and up to date list is maintained on the TNG website.

6.2 Collaboration and contributions

The full snapshots of TNG100-1, and especially those of TNG300-1, are sufficiently large that it may be prohibitive for most users to acquire or store a large number. We note that transferring ~ 1.5 TB (the size of one full TNG100-1 snapshot) at a reasonably achievable 10 MB/s will take roughly 48 hours, increasing to roughly five days for a ~ 4.1 TB full snapshot of TNG300-1. As a result, projects requiring access to full simulation datasets, or extensive post-processing computations beyond what are being made publicly available, may benefit from closer interaction with members of the TNG collaboration.

We also welcome suggestions, comments, or contributions to the data release effort itself. These could take the form of analysis code, derived data catalogs, etc. For instance, interesting data products can be released as a “supplementary data catalog”. Fast analysis routines which operate on individual halos/subhalos can be integrated into the API, such that the result can be requested on demand for any object.

6.3 Future data releases

We anticipate to release additional data in the future, for which further documentation will be provided online.

6.3.1 *Rockstar and velociraptor*

We plan to derive and release different group catalogs, based on the ROCKSTAR (Behroozi et al. 2013) and VELOCIRAPTOR (Elahi et al. 2011) algorithms, and will provide further documentation at that time. Such group catalogs will identify different subhalo populations than found by the Subfind algorithm, particularly during mergers. The algorithm used to construct the ‘Consistent Trees’ assembly histories also has fundamental differences to both LHALOTREE and SUBLINK. This can provide a powerful comparison and consistency check for any scientific analysis. We also anticipate that some users will simply be more familiar with these outputs, or need them as inputs to other tools.

6.3.2 *Additional simulations*

The flagship volumes of the IllustrisTNG—TNG50, TNG100, and TNG300—are accompanied by an additional resource: a large number of ‘TNG Model Variation’ simulations. Each modifies exactly one choice or parameter value of the base, fiducial TNG model. The variations cover every physical aspect of the model, including the stellar and black hole feedback mechanisms, aspects of the star formation, as well as numerical parameters. They are invaluable in assessing the robustness of a physical conclusion to model changes, as well as in diagnosing the underlying cause or mechanism responsible for a given feature in the primary simulations. They were first presented in the Pillepich et al. (2018b) TNG methods paper, and used for example in Nelson et al. (2018b) to understand the improvement in OVI column densities, in Lovell et al. (2018) to study the impact of baryons on dark matter fractions, and in Terrazas et al. (in prep) to probe the origin of quenched galaxies in the TNG model.

Each of the ~ 100 TNG model variants simulates the exact same 25 Mpc/h volume at a resolution approximately equivalent to the flagship TNG100-1. Individual halos can also therefore be cross-matched between the simulations, although the statistics is necessarily limited by the relatively small volume. We plan to publicly release these variations in the near future, and encourage those interested to get in touch in the meantime.

Finally, we anticipate that ongoing and future simulation projects will also be released through this platform in the future. Most notably, this includes the high-resolution TNG50 simulation (Pillepich et al. 2019; Nelson et al. 2019a), the third and final volume of the IllustrisTNG project, and potentially other simulations as well.

6.3.3 *API functionality expansion*

There is significant room for the development of additional features in the web-based API. In particular, for (i) on-demand visualization tasks, (ii) on-demand analysis tasks, and (iii) client-side, browser based tools for data exploration and visualization. We have two specific services which are anticipated to be developed in the near-term future and made available.

First, the on-demand generation of ‘zoom’ initial conditions (ICs), for individual galaxies/halos, based on any object of interest selected from any simulation box. This will allow a user to select a sample of galaxies, perhaps in analogy to an observed sample, or with a peculiar type of assembly history, and obtain ICs for resimulation. Such resimulations could use other codes or galaxy formation models, or explore model parameter variations, to assess how such changes affected a particular galaxy/halo, or class of galaxies/halos. As IC generation will take several

minutes at least, it does not fit within our current framework of ‘responses within a few seconds’, and therefore requires a task-based work queue with delayed execution and subsequent notification (e.g. via email) of completion and the availability of new data products for download.

Second, the on-demand execution of longer running analysis tasks, with similar notification upon completion. Specifically, the ability to request SKIRT radiative-transfer calculations for specific galaxies/halos of interest, leveraging the development efforts of Rodriguez-Gomez et al. (2019). Other expensive mocks, such as spectral HI (with MARTINI; Oman et al. 2019) or x -ray datacubes, or intergalactic quasar absorption sightlines, can similarly be generated.

We welcome community input and/or contributions in any of these directions, or comments related to any aspect of the public data release of TNG.

7 Architectural and design details

In the development of the original Illustris public data release, many design decisions were made, including technical details related to the release effort, based on expected use cases and methods of data analysis. Nelson et al. (2015) discusses the architectural goals and considerations that we followed and continue to follow with the IllustrisTNG data release, and contrasts against other methodologies, as implemented in other astrophysics simulation data releases. We refer the reader to that paper and present only a few updates here.

7.1 Usage of the illustris public data release

Since its release, the original Illustris public data release has seen widespread adoption and use. To date, in the three and a half years since launch, 2122 new users have registered and made a total of 269 million API requests, including 2.7 million ‘mock FITS’ file downloads. For the flagship Illustris-1 run, a total of 1390 full snapshots, 6650 group catalogs, and 180 merger trees have been downloaded. 26 million subhalo ‘cutouts’ of particle-level data, and 3.1 million SUBLINK merger tree extractions have been requested. The total data transfer for this simulation to date is $\simeq 2.15$ PB. Roughly 3100 subbox snapshots of Illustris-1 have been downloaded. The next most accessed simulation is Illustris-3, likely because it is included in the getting started tutorials as an easy, lightweight alternative to Illustris-1. Since launch, there has been a nearly constant number of ~ 100 – 120 active users, based on activity within the last 30 days.

To date, 163 publications have directly resulted from, or included analysis results from, the Illustris simulation. While early papers were written largely by the collaboration itself, recent papers typically do not involve members of the Illustris team, representing widespread public use

of the data release. Of the 10 most recent papers published on Illustris, only one was from the team. Given the significantly expanded scope of TNG with respect to Illustris, as well as the relatively more robust and reliable physical model and outcomes, we expect that uptake and usage will be similarly broad.

7.2 New JupyterLab interface

In the original Illustris data release, we promoted two ways to work with the data: either downloading large simulation data files directly (referred to above as ‘local data, local compute’), or by searching and downloading data subsets using functionality in the web-based API (‘remote data, local compute’). Previously, the backend was focused solely on storage and data delivery, and did not have any system in place to allow guest access to compute resources which were local to the datasets themselves. For the TNG data release we have developed this functionality.

We label this newly introduced, third method of working with the data ‘remote data, remote computation.’ Technically, we make use of JupyterHub to manage the instantiation of per-user JupyterLab instances. These are spawned inside containerized Docker instances (Merkel 2014) to isolate the user from the host systems—Singularity (Kurtzer et al. 2017) could be used in the future. Read-only mounts to the parallel filesystems hosting simulation data are provided, while the user home directory within the container is made persistent by volume mapping it onto the host. Resource limits on CPU, memory, and storage are controllable and will be adjusted during the initial phase of this service as needed.

The JupyterLab instances themselves provide a familiar environment for the development and execution of user analysis programs. Over the past few years there has been significant recent development on remote, multi-user, rich interfaces to computational kernels, and JupyterLab (the successor of Jupyter, previously called IPython; Pérez and Granger 2007) is a mature, full-featured solution we deploy. These instances are launched, on demand, inside the sand-boxed containers, through a web-based portal with authentication integrated into the existing user registration system of the data release. We anticipate that this will be a particularly interesting development for researchers who would otherwise not have the computational resources to use the simulation data for their science.

7.3 Retiring the relational database

In the original Illustris data release we noted that the read-only, highly structured nature of simulation output motivates different and more efficient approaches for data search, aggregation, processing, and retrieval tasks. The web-based API uses a representational state transfer architecture (REST, Fielding 2000), and in TNG we continue to

employ a relational database on the backend, although we made a design decision never to expose such a database to direct user query.

Looking forward, instead of bringing the object or group catalog data into a traditional database, one could employ a scheme such as bitmap indexing over HDF5, e.g. FastQuery (Chou et al. 2011; Byna et al. 2012), possibly combined with a SQL-compatible query layer (Wang et al. 2013). In this case, the database would be used only to handle light meta-data—fast index-accelerated search queries would be made directly against structured binary data on disk. This improvement would be largely transparent from the user perspective. Most obviously, it would remove a layer of complexity and the need to ingest of order billions of rows of group catalog data into a database. It would also enable a tighter coupling of search capabilities and on-disk data contents.

More efficient API standards such as GraphQL represent modern alternatives to REST, whereby users make specific, detailed requests to a single endpoint based on a well-defined query language and typed schema, rather than a number of generic requests to a diversity of endpoints. Resolving these declarative queries efficiently and directly on the simulation output data would unify many of these goals—a clear target for future development.

8 Summary and conclusions

We have made publicly available data from the IllustrisTNG simulation project at the permanent URL:

<http://www.tng-project.org/data/>

IllustrisTNG is a series of large-scale, cosmological simulations ideal for studying the formation and evolution of galaxies. The simulation suite consists of three volumes: TNG50, TNG100, and TNG300. Each flagship run is accompanied by lower-resolution realizations, and a dark-matter only analog of every simulation is also available. The current data release includes TNG100 and TNG300 in their entirety, and TNG50 will be publicly released in the future. Full snapshots, group catalogs (both friends of friends halos and SUBFIND subhalos), merger trees, high time-resolution subboxes, and many supplementary data catalogs are made available. The highest resolution TNG300-1 includes more than ten million gravitationally bound structures, and the TNG100-1 volume contains $\sim 20,000$ well-resolved galaxies at $z = 0$ with stellar mass exceeding $10^9 M_{\odot}$. The galaxies sampled in these volumes encompass a broad range of mass, type, environment and assembly history, and realize fully representative synthetic universes within the context of Λ CDM.

The total data volume produced by the Illustris[TNG] project is sizeable, ~ 1.1 PB in total, all directly accessible online. We have developed several tools to make these data accessible to the broader community, without requiring extensive local computational resources. In addition

to direct data download, example scripts, web-based API access methods, and extensive documentation previously developed for the original Illustris simulation, we extend the data access functionality in several ways. Namely, with new on-demand visualization and analysis functionality, and with the remote JupyterLab-based analysis interface. By making the TNG data publicly available, we aim to maximize the scientific return from the considerable computational resources invested in the TNG simulation suite.

Appendix 1: Simulation data details

See Table A.1.

Appendix 2: Web-based API examples

By way of explicit example, the following are absolute URLs for the web-based API which encompass some of its functionality. The type of the request, as well as the data expected in return, should be relatively clear:

- www.tng-project.org/api/Illustris-2/
- www.tng-project.org/api/Illustris-2/snapshots/68/
- www.tng-project.org/api/Illustris-1/snapshots/135/subhalos/73664/
- www.tng-project.org/api/Illustris-1/snapshots/80/halos/523312/cutout.hdf5?dm=Coordinates&gas=all
- www.tng-project.org/api/Illustris-3/snapshots/135/subhalos?mass_gt=10.0&mass_lt=20.0
- www.tng-project.org/api/Illustris-2/snapshots/68/subhalos/50000/sublink/full.hdf5
- www.tng-project.org/api/Illustris-2/snapshots/68/subhalos/50000/sublink/mpb.json
- www.tng-project.org/api/TNG100-2/snapshots/99/subhalos/50000/sublink/mpb.json
- www.tng-project.org/api/TNG300-1/snapshots/99/subhalos?mass_gt=10.0&mass_lt=11.0
- www.tng-project.org/api/Illustris-2/files/ics.hdf5
- www.tng-project.org/api/Illustris-1/files/groupcat-135.5.hdf5
- www.tng-project.org/api/Illustris-2/files/snapshot-135.10.hdf5?dm=all
- www.tng-project.org/api/TNG100-1/snapshots/50/subhalos/plot.png?xQuant=mstar2_log&yQuant=ssfr
- www.tng-project.org/api/TNG300-1/snapshots/7/subhalos/plot.png?xQuant=mstar2&yQuant=delta_sfms&cQuant=Z_gas
- www.tng-project.org/api/TNG100-1/snapshots/99/halos/320/vis.png
- www.tng-project.org/api/TNG100-2/snapshots/67/halos/0/vis.png?partType=gas&partField=temp

A ‘getting started’ guide for the web-based API is available in the online documentation, and this includes a cookbook of common analysis tasks (available in Python, IDL, and Matlab). To give a sense of this method of analyzing TNG data, we include here four short examples, in Python.

Table A.1 Details on the file organization for all twenty TNG runs, both baryonic and dark-matter only. We include the number of file chunks, the average size of a full snapshot and the corresponding group catalog, and an estimate of the total data volume of the simulation

Run	Alternate Name	Total N_{DM}	N_{chunks}	Full Snapshot Size	Avg Groupcat Size	Total Data Volume
L35n270TNG	TNG50-4	19,683,000	11	5.2 GB	20 MB	0.6 TB
L35n270TNG_DM	TNG50-4-Dark	19,683,000	4	1.2 GB	10 MB	0.1 TB
L35n540TNG	TNG50-3	157,464,000	11	44 GB	130 MB	7.5 TB
L35n540TNG_DM	TNG50-3-Dark	157,464,000	4	9.4 GB	50 MB	0.6 TB
L35n1080TNG	TNG50-2	1,259,712,000	128	350 GB	860 MB	18 TB
L35n1080TNG_DM	TNG50-2-Dark	1,259,712,000	85	76 GB	350 MB	4.5 TB
L35n2160TNG	TNG50-1	10,077,696,000	680	2.7 TB	7.2 GB	~320 TB
L35n2160TNG_DM	TNG50-1-Dark	10,077,696,000	128	600 GB	2.3 GB	36 TB
L75n455TNG	TNG100-3	94,196,375	8	27 GB	110 MB	1.5 TB
L75n455TNG_DM	TNG100-3-Dark	94,196,375	4	5.7 GB	40 MB	0.4 TB
L75n910TNG	TNG100-2	753,571,000	56	215 GB	650 MB	14 TB
L75n910TNG_DM	TNG100-2-Dark	753,571,000	8	45 GB	260 MB	2.8 TB
L75n1820TNG	TNG100-1	6,028,568,000	448	1.7 TB	4.3 GB	128 TB
L75n1820TNG_DM	TNG100-1-Dark	6,028,568,000	64	360 GB	1.7 GB	22 TB
L205n625TNG	TNG300-3	244,140,625	16	63 GB	340 MB	4 TB
L205n625TNG_DM	TNG300-3-Dark	244,140,625	4	15 GB	130 MB	1 TB
L205n1250TNG	TNG300-2	1,953,125,000	100	512 GB	2.2 GB	31 TB
L205n1250TNG_DM	TNG300-2-Dark	1,953,125,000	25	117 GB	810 MB	7.2 TB
L205n2500TNG	TNG300-1	15,625,000,000	600	4.1 TB	14 GB	235 TB
L205n2500TNG_DM	TNG300-1-Dark	15,625,000,000	75	930 GB	5.2 GB	57 TB

Task 1 For TNG300-1 at $z=0$, get all the information available for the ID = 0 subhalo, print both its total mass and stellar half mass radius

```
>>> url = "http://www.tng-project.org/api/TNG300-1/snapshots/99/subhalos/0/"
>>> r = get(url)
>>> r['mass']
128335.0

>>> r['halfmassrad_stars']
130.065
```

Task 2 For TNG100-1 at $z=2$, search for all subhalos with total mass $10^{12.1}M_{\odot} < M < 10^{12.2}M_{\odot}$ and print the SUBFIND IDs of the first five results

```
>>> # convert from log solar masses to group catalog units
>>> mass_min = 10**12.1 / 1e10 * 0.6774
>>> mass_max = 10**12.2 / 1e10 * 0.6774
>>>
>>> params = {'mass__gt':mass_min, 'mass__lt':mass_max}
>>>
>>> # make the request
>>> url = "http://www.tng-project.org/api/TNG100-1/snapshots/z=2/subhalos/"
>>> subhalos = get(url, params)
>>>
>>> ids = [ subhalos['results'][i]['id'] for i in range(5) ]
>>> ids
[13845, 16799, 23224, 24400, 12718]
```

Task 11 Download the entire TNG300-1 $z = 0$ snapshot including *only the positions, masses, and metallicities of stars* (in the form of 600 HDF5 files). In this example, since we only need these three fields for stars only, we can reduce the download and storage size from ~4.1 TB to ~20 GB

```
>>> base_url = "http://www.tng-project.org/api/TNG300-1/"
>>> sim_metadata = get(base_url)
>>> params = {'stars': 'Coordinates,Masses,GFM_Metallicity'}
>>>
>>> for i in range(sim_metadata['num_files_snapshot']):
>>>     file_url = base_url + "files/snapshot-99." + str(i) + ".hdf5"
>>>     saved_filename = get(file_url, params)
>>>     print(saved_filename)
```

Each uses of the `GET()` helper function which performs the actual HTTP request, automatically parsing JSON-type returns.

Acknowledgements

We thank the Max Planck Computing and Data Facility (MPCDF) and the Max Planck Institute for Astrophysics (MPA) in Garching, Germany, for significant computational resources and ongoing support, without which the TNG public data release would not be possible. We also thank the Research Computing group of Harvard University, for significant computational resources and continuous support, without which the original Illustris public data release and ongoing activities would not be possible.

In addition, we thank the Center for Computational Astrophysics and the Simons Foundation for hosting a full mirror of the Illustris[TNG] data, particularly Dylan Simon and the Scientific Computing Core at the Flatiron Institute for orchestration efforts. In addition, we thank Greg Snyder and the Mikulski Archive for Space Telescopes (MAST) at STScI for future plans to host relevant datasets based on TNG.

The authors would like to thank many additional people for contributing to analysis and understanding of the TNG simulations and their results: Maria Celeste Artale, Kelly Blumenthal, David Barnes, Martina Donnari, Elena D'Onghia, Min Du, Catalina Gomez, Hong Guo, Anshu Gupta, Melanie Habouzit, Chris Hayward, Ghandali Joshi, Guinevere Kauffmann, Davide Martizzi, Michelle Ntampaka, Ana-Roxana Pop, Gergö Popping, Malin Renneby, Greg Snyder, Adam Stevens, Sandro Tacchella, Bryan Terrazas, Nhut Truong, Hannah Übler, Francisco Villaescusa-Navarro, Sarah Wellons, Po-Feng Wu, Dandan Xu, Kiyun Yun, Qirong Zhu, Elad Zinger, and Jolanta Zjupa.

In the execution of the primary simulations presented herein, the authors gratefully acknowledge the Gauss Centre for Supercomputing (GCS) for providing computing time for the GCS Large-Scale Projects GCS-ILLU (2014) and GCS-DWAR (2016) on the GCS share of the supercomputer Hazel Hen at the High Performance Computing Center Stuttgart (HLRS). GCS is the alliance of the three national supercomputing centres HLRS (Universität Stuttgart), JSC (Forschungszentrum Jülich), and LRZ (Bayerische Akademie der Wissenschaften), funded by the German Federal Ministry of Education and Research (BMBF) and the German State Ministries for Research of Baden-Württemberg (MWK), Bayern (StMWFK) and Nordrhein-Westfalen (MIWF). Additional simulations were carried out on the Hydra and Draco supercomputers at the Max Planck Computing and Data Facility (MPCDF, formerly known as RZG), as well as on the Stampede supercomputer at the Texas Advanced Computing Center through the XSEDE project AST140063. Some additional computations in this paper were run on the Odyssey cluster supported by the FAS Division of Science, Research Computing Group at Harvard University.

Funding

Not applicable.

Abbreviations

Not applicable.

Availability of data and materials

Permanently available at www.tng-project.org/data. This paper describes an explicit public data release of the relevant material.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

DN has designed and carried out the data release, authored the presentation and data distribution websites, developed and maintains the technical infrastructure, and composed this document. VS, AP, RP and DN have carried out the TNG simulations. VS, LH, RW and AP have developed the TNG model with further input from SG, PT, MV, FM, RP and DN. VRG, LK, ML, and BD have contributed additional data products. All authors read and approved the final manuscript.

Author details

¹Max-Planck-Institut für Astrophysik, Garching, Germany. ²Max-Planck-Institut für Astronomie, Heidelberg, Germany. ³Center for Astrophysics, Harvard & Smithsonian, Cambridge, USA. ⁴Center for Computational Astrophysics, Flatiron Institute, New York, USA. ⁵Kavli Institute for Astrophysics and Space Research, Department of Physics, MIT, Cambridge, USA. ⁶Heidelberg Institute for Theoretical Studies, Heidelberg, Germany. ⁷Zentrum für Astronomie der Universität Heidelberg, Heidelberg, Germany. ⁸Instituto de Radioastronomía y Astrofísica, Universidad Nacional Autónoma de México, Morelia, Mexico. ⁹Department of Physics, University of Florida, Gainesville, USA. ¹⁰Department of Physics and Astronomy, University of Bologna, Bologna, Italy. ¹¹Center for Interdisciplinary Exploration and Research in Astrophysics, Northwestern University, Evanston, USA. ¹²Center for Astrophysics and Cosmology, Science Institute, University of Iceland, Reykjavik, Iceland. ¹³Institute for Computational Cosmology, Durham University, Durham, UK.

Endnotes

- a www.tng-project.org
- b www.illustris-project.org
- c www.tng-project.org/results
- d New feature in the TNG data release.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 15 January 2019 Accepted: 16 April 2019

Published online: 14 May 2019

References

- Araya, M., Osorio, M., Díaz, M., Ponce, C., Villanueva, M., Valenzuela, C., Solar, M.: JOVIAL: notebook-based astronomical data analysis in the cloud. *Astron. Comput.* **25**, 110–117 (2018). <https://doi.org/10.1016/j.ascom.2018.09.001.1812.01477>
- Barnes, D.J., Vogelsberger, M., Kannan, R., Marinacci, F., Weinberger, R., Springel, V., Torrey, P., Pillepich, A., Nelson, D., Pakmor, R., Naiman, J., Hernquist, L.,

- McDonald, M.: A census of cool-core galaxy clusters in IllustrisTNG. *Mon. Not. R. Astron. Soc.* **481**, 1809–1831 (2018). <https://doi.org/10.1093/mnras/sty2078>. 1710.08420
- Behroozi, P.S., Wechsler, R.H., Wu, H.-Y., Busha, M.T., Klypin, A.A., Primack, J.R.: Gravitationally consistent halo catalogs and merger trees for precision cosmology. *Acad. Publ. J.* **763**, 18 (2013). <https://doi.org/10.1088/0004-637X/763/1/18>. 1110.4370
- Bernyk, M., Croton, D.J., Tonini, C., Hodkinson, L., Hassan, A.H., Garel, T., Duffy, A.R., Mutch, S.J., Poole, G.B.: The Theoretical Astrophysical Observatory: cloud-Based Mock Galaxy Catalogues (2014). ArXiv e-prints 1403.5270
- Blecha, L., Sijacki, D., Kelley, L.Z., Torrey, P., Vogelsberger, M., Nelson, D., Springel, V., Snyder, G., Hernquist, L.: Recoiling black holes: prospects for detection and implications of spin alignment. *Mon. Not. R. Astron. Soc.* **456**, 961–989 (2016). <https://doi.org/10.1093/mnras/stv2646>
- Boylan-Kolchin, M., Springel, V., White, S.D.M., Jenkins, A., Lemson, G.: Resolving cosmic structure formation with the millennium-II simulation. *Mon. Not. R. Astron. Soc.* **398**, 1150–1164 (2009). <https://doi.org/10.1111/j.1365-2966.2009.15191.x>. 0903.3041
- Brammer, G.B., van Dokkum, P.G., Franx, M., Fumagalli, M., Patel, S., Rix, H.-W., Skelton, R.E., Kriek, M., Nelson, E., Schmidt, K.B., Bezanson, R., da Cunha, E., Erb, D.K., Fan, X., Förster Schreiber, N., Illingworth, G.D., Labbé, I., Leja, J., Lundgren, B., Magee, D., Marchesini, D., McCarthy, P., Momocheva, I., Muzzin, A., Quadri, R., Steidel, C.C., Tal, T., Wake, D., Whitaker, K.E., Williams, A.: 3D-HST: a wide-field grism spectroscopic survey with the hubble space telescope. *Astrophys. J. Suppl. Ser.* **200**, 13 (2012). <https://doi.org/10.1088/0067-0049/200/2/13>. 1204.2829
- Bryant, J.J., Owers, M.S., Robotham, A.S.G., Croom, S.M., Driver, S.P., Drinkwater, M.J., Lorente, N.P.F., Cortese, L., Scott, N., Colless, M., Schaefer, A., Taylor, E.N., Konstantopoulos, I.S., Allen, J.T., Baldry, I., Barnes, L., Bauer, A.E., Bland-Hawthorn, J., Bloom, J.V., Brooks, A.M., Brough, S., Cecil, G., Couch, W., Croton, D., Davies, R., Ellis, S., Fogarty, L.M.R., Foster, C., Glazebrook, K., Goodwin, M., Green, A., Gunawardhana, M.L., Hampton, E., Ho, I.-T., Hopkins, A.M., Kewley, L., Lawrence, J.S., Leon-Saval, S.G., Leslie, S., McElroy, R., Lewis, G., Liske, J., López-Sánchez, Á.R., Mahajan, S., Medling, A.M., Metcalfe, N., Meyer, M., Mould, J., Obreschcow, D., O’Toole, S., Pracy, M., Richards, S.N., Shanks, T., Sharp, R., Sweet, S.M., Thomas, A.D., Tonini, C., Walcher, C.J.: The SAMI galaxy survey: instrument specification and target selection. *Mon. Not. R. Astron. Soc.* **447**, 2857–2879 (2015). <https://doi.org/10.1093/mnras/stu2635>. 1407.7335
- Buck, T., Dutton, A.A., Macciò, A.V.: A conclusive test for star formation prescriptions in cosmological hydrodynamical simulations (2018). ArXiv e-prints 1812.05613
- Bundy, K., Bershady, M.A., Law, D.R., Yan, R., Drory, N., MacDonald, N., Wake, D.A., Cherinka, B., Sánchez-Gallego, J.R., Weijmans, A.-M., Thomas, D., Tremonti, C., Masters, K., Coccatto, L., Diamond-Stanic, A.M., Aragón-Salamanca, A., Avila-Reese, V., Badenes, C., Falcón-Barroso, J., Belfiore, F., Bizyaev, D., Blanc, G.A., Bland-Hawthorn, J., Blanton, M.R., Brownstein, J.R., Byler, N., Cappellari, M., Conroy, C., Dutton, A.A., Emsellem, E., Etherington, J., Frinchaboy, P.M., Fu, H., Gunn, J.E., Harding, P., Johnston, E.J., Kauffmann, G., Kinemuchi, K., Klaene, M.A., Knapen, J.H., Leauthaud, A., Li, C., Lin, L., Maiolino, R., Malanushenko, V., Malanushenko, E., Mao, S., Maraston, C., McDermid, R.M., Merrifield, M.R., Nichol, R.C., Oravetz, D., Pan, K., Parejko, J.K., Sanchez, S.F., Schlegel, D., Simmons, A., Steele, O., Steinmetz, M., Thanjavur, K., Thompson, B.A., Tinker, J.L., van den Bosch, R.C.E., Westfall, K.B., Wilkinson, D., Wright, S., Xiao, T., Zhang, K.: Overview of the SDSS-IV MaNGA survey: mapping nearby galaxies at apache point observatory. *Acad. Publ. J.* **798**, 7 (2015). <https://doi.org/10.1088/0004-637X/798/1/7>. 1412.1482
- Byna, S., Chou, J., Rübél, O., Prabhat Karimabadi, H., Daughton, W.S., Roytershteyn, V., Bethel, E.W., Howison, M., Hsu, K.-J., Lin, K.-W., Shoshani, A., Useton, A., Wu, K.: Parallel i/o, analysis, and visualization of a trillion particle simulation. In: SC ’12: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, pp. 59–15912. IEEE Comput. Soc., Los Alamitos (2012). <http://dl.acm.org/citation.cfm?id=2388996.2389077>
- Camps, P., Trčka, A., Trayford, J., Baes, M., Theuns, T., Crain, R.A., McAlpine, S., Schaller, M., Schaye, J.: Data release of UV to submillimeter broadband fluxes for simulated galaxies from the EAGLE project. *Astrophys. J. Suppl. Ser.* **234**, 20 (2018). <https://doi.org/10.3847/1538-4365/aa24c>
- Chou, J., Wu, K., Prabhat: Fastquery: a parallel indexing system for scientific data. In: Cluster Computing (CLUSTER), 2011 IEEE International Conference on, pp. 455–464 (2011). <https://doi.org/10.1109/CLUSTER.2011.86>
- Conroy, C., Gunn, J.E.: The propagation of uncertainties in stellar population synthesis modeling. III. Model calibration, comparison, and evaluation. *Acad. Publ. J.* **712**, 833–857 (2010). <https://doi.org/10.1088/0004-637X/712/2/833>. 0911.3151
- Conroy, C., Gunn, J.E., White, M.: The propagation of uncertainties in stellar population synthesis modeling. I. The relevance of uncertain aspects of stellar evolution and the initial mass function to the derived physical properties of galaxies. *Acad. Publ. J.* **699**, 486–506 (2009). <https://doi.org/10.1088/0004-637X/699/1/486>. 0809.4261
- Crain, R.A., Schaye, J., Bower, R.G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C.S., McCarthy, I.G., Helly, J.C., Jenkins, A., Rosas-Guevara, Y.M., White, S.D.M., Trayford, J.W.: The EAGLE simulations of galaxy formation: calibration of subgrid physics and model variations. *Mon. Not. R. Astron. Soc.* **450**, 1937–1961 (2015). <https://doi.org/10.1093/mnras/stv725>. 1501.01311
- Crocce, M., Fosalba, P., Castander, F.J., Gaztañaga, E.: Simulating the universe with MICE: the abundance of massive clusters. *Mon. Not. R. Astron. Soc.* **403**, 1353–1367 (2010). <https://doi.org/10.1111/j.1365-2966.2009.16194.x>. 0907.0019
- Davé, R., Anglés-Alcázar, D., Narayanan, D., Li, Q., Rafieferantsoa, M.H., Appleby, S.: Simba: cosmological simulations with black hole growth and feedback (2019). ArXiv e-prints 1901.10203
- Davis, M., Efstathiou, G., Frenk, C.S., White, S.D.M.: The evolution of large-scale structure in a universe dominated by cold dark matter. *Acad. Publ. J.* **292**, 371–394 (1985). <https://doi.org/10.1086/163168>
- de Souza, R.S., Ciardi, B.: AMADA-analysis of multidimensional astronomical datasets. *Astron. Comput.* **12**, 100–108 (2015). <https://doi.org/10.1016/j.ascom.2015.06.006>. 1503.07736
- Diemer, B., Stevens, A.R.H., Forbes, J.C., Marinacci, F., Hernquist, L., Lagos, C.d.P., Sternberg, A., Pillepich, A., Nelson, D., Popping, G., Villaescusa-Navarro, F., Torrey, P., Vogelsberger, M.: Modeling the atomic-to-molecular transition in cosmological simulations of galaxy formation. *Astrophys. J. Suppl. Ser.* **238**, 33 (2018). <https://doi.org/10.3847/1538-4365/aae387>. 1806.02341
- Dolag, K., Komatsu, E., Sunyaev, R.: SZ effects in the magneticum pathfinder simulation: comparison with the Planck, SPT, and ACT results. *Mon. Not. R. Astron. Soc.* **463**, 1797–1811 (2016). <https://doi.org/10.1093/mnras/stw2035>. 1509.05134
- Donnari, M., Pillepich, A., Nelson, D., Vogelsberger, M., Genel, S., Weinberger, R., Marinacci, F., Springel, V., Hernquist, L.: The star-formation activity of IllustrisTNG galaxies: main sequence, UVJ diagram, quenched fractions, and systematics (2018). ArXiv e-prints 1812.07584
- Dubois, Y., Pichon, C., Welker, C., Le Borgne, D., Devriendt, J., Laigle, C., Codis, S., Pogosyan, D., Arnouts, S., Benabed, K., Bertin, E., Blaizot, J., Bouchet, F., Cardoso, J.-F., Colombi, S., de Lapparent, V., Desjacques, V., Gavazzi, R., Kassim, S., Kimm, T., McCracken, H., Milliard, B., Peirani, S., Prunet, S., Rouberol, S., Silk, J., Slyz, A., Sousbie, T., Teysier, R., Tresse, L., Treyer, M., Vibert, D., Volonteri, M.: Dancing in the dark: galactic properties trace spin swings along the cosmic web. *Mon. Not. R. Astron. Soc.* **444**, 1453–1468 (2014). <https://doi.org/10.1093/mnras/stu1227>. 1402.1165
- Dykes, T., Hassan, A., Gheller, C., Croton, D., Krokos, M.: Interactive 3D visualization for theoretical virtual observatories. *Mon. Not. R. Astron. Soc.* **477**, 1495–1507 (2018). <https://doi.org/10.1093/mnras/sty855>. 1803.11399
- Elahi, P.J., Thacker, R.J., Widrow, L.M.: Peaks above the Maxwellian sea: a new approach to finding substructures in N-body haloes. *Mon. Not. R. Astron. Soc.* **418**, 320–335 (2011). <https://doi.org/10.1111/j.1365-2966.2011.19485.x>. 1107.4289
- Faucher-Giguère, C.-A.: Recent progress in simulating galaxy formation from the largest to the smallest scales. *Nat. Astron.* **2**, 368–373 (2018). <https://doi.org/10.1038/s41550-018-0427-y>. 1805.00014
- Faucher-Giguère, C.-A., Lidz, A., Zaldarriaga, M., Hernquist, L.: A new calculation of the ionizing background spectrum and the effects of He II reionization. *Acad. Publ. J.* **703**, 1416–1443 (2009). <https://doi.org/10.1088/0004-637X/703/2/1416>. 0901.4554
- Feng, Y., Di-Matteo, T., Croft, R.A., Bird, S., Battaglia, N., Wilkins, S.: The BlueTides simulation: first galaxies and reionization. *Mon. Not. R. Astron. Soc.* **455**, 2778–2791 (2016). <https://doi.org/10.1093/mnras/stv2484>. 1504.06619
- Ferland, G.J., Chatzikos, M., Guzmán, F., Lykins, M.L., van Hoof, P.A.M., Williams, R.J.R., Abel, N.P., Badnell, N.R., Keenan, F.P., Porter, R.L., Stancil, P.C.: The 2017 release cloudy. *RMxAA* **53**, 385–438 (2017)
- Fielding, R.T.: Architectural styles and the design of network-based software architectures (2000). AAI9980887

- Fitzpatrick, M.J., Olsen, K., Economou, F., Stobie, E.B., Beers, T.C., Dickinson, M., Norris, P., Saha, A., Seaman, R., Silva, D.R., Swaters, R.A., Thomas, B., Valdes, F.: The noao data laboratory: a conceptual overview. *Proc. SPIE* **9149**, 9149–914916 (2014). <https://doi.org/10.1117/12.2057445>
- Foreman-Mackey, D., Sick, J., Johnson, B.: Python-Fsps: Python Bindings to FSPS (V0.1.1) (2014). <https://doi.org/10.5281/zenodo.12157>.
- Garrison, L.H., Eisenstein, D.J., Ferrer, D., Tinker, J.L., Pinto, P.A., Weinberg, D.H.: The abacus cosmos: a suite of cosmological N-body simulations. *Astrophys. J. Suppl. Ser.* **236**, 43 (2018). <https://doi.org/10.3847/1538-4365/aabfd3>. 1712.05768
- Genel, S., Fall, S.M., Hernquist, L., Vogelsberger, M., Snyder, G.F., Rodriguez-Gomez, V., Sijacki, D., Springel, V.: Galactic angular momentum in the illustris simulation: feedback and the hubble sequence (2015). ArXiv e-prints [1503.01117](https://arxiv.org/abs/1503.01117)
- Genel, S., Nelson, D., Pillepich, A., Springel, V., Pakmor, R., Weinberger, R., Hernquist, L., Naiman, J., Vogelsberger, M., Marinacci, F., Torrey, P.: The size evolution of star-forming and quenched galaxies in the IllustrisTNG simulation. *Mon. Not. R. Astron. Soc.* **474**, 3976–3996 (2018). <https://doi.org/10.1093/mnras/stx3078>. 1707.05327
- Genel, S., Vogelsberger, M., Nelson, D., Sijacki, D., Springel, V., Hernquist, L.: Following the flow: tracer particles in astrophysical fluid simulations. *Mon. Not. R. Astron. Soc.* **435**, 1426–1442 (2013). <https://doi.org/10.1093/mnras/stt1383>. 1305.2195
- Genel, S., Vogelsberger, M., Springel, V., Sijacki, D., Nelson, D., Snyder, G., Rodriguez-Gomez, V., Torrey, P., Hernquist, L.: Introducing the Illustris project: the evolution of galaxy populations across cosmic time. *Mon. Not. R. Astron. Soc.* **445**, 175–200 (2014). <https://doi.org/10.1093/mnras/stu1654>. 1405.3749
- Genzel, R., Förster Schreiber, N.M., Lang, P., Tacchella, S., Tacconi, L.J., Wuyts, S., Bandara, K., Burkert, A., Buschkamp, P., Carollo, C.M., Cresci, G., Davies, R., Eisenhauer, F., Hicks, E.K.S., Kurk, J., Lilly, S.J., Lutz, D., Mancini, C., Naab, T., Newman, S., Peng, Y., Renzini, A., Shapiro Griffin, K., Sternberg, A., Vergani, D., Wisnioski, E., Wuyts, E., Zamorani, G.: The SINS/zC-SINF survey of $z \sim 2$ galaxy kinematics: evidence for gravitational quenching. *Acad. Publ. J.* **785**, 75 (2014). <https://doi.org/10.1088/0004-637X/785/1/75>. 1310.3838
- Grand, R.J.J., Gómez, F.A., Marinacci, F., Pakmor, R., Springel, V., Campbell, D.J.R., Frenk, C.S., Jenkins, A., White, S.D.M.: The auriga project: the properties and formation mechanisms of disc galaxies across cosmic time. *Mon. Not. R. Astron. Soc.* **467**, 179–207 (2017). <https://doi.org/10.1093/mnras/stx071>. 1610.01159
- Gray, J., Szalay, A.S., Thakar, A.R., Kunszt, P.Z., Stoughton, C., Slutz, D., vandenBerg, J.: Data mining the SDSS SkyServer database (2002). e-prints arXiv [cs/0202014](https://arxiv.org/abs/cs/0202014)
- Grogin, N.A., Kocevski, D.D., Faber, S.M., Ferguson, H.C., Koekemoer, A.M., Riess, A.G., Acquaviva, V., Alexander, D.M., Almaini, O., Ashby, M.L.N., Barden, M., Bell, E.F., Bournaud, F., Brown, T.M., Caputi, K.I., Casertano, S., Cassata, P., Castellano, M., Challis, P., Chary, R.-R., Cheung, E., Cirasuolo, M., Conselice, C.J., Roshan Cooray, A., Croton, D.J., Daddi, E., Dahlen, T., Davé, R., de Mello, D.F., Dekel, A., Dickinson, M., Dolch, T., Donley, J.L., Dunlop, J.S., Dutton, A.A., Elbaz, D., Fazio, G.G., Filippenko, A.V., Finkelstein, S.L., Fontana, A., Gardner, J.P., Garnavich, P.M., Gawiser, E., Giavalisco, M., Grazian, A., Guo, Y., Hathi, N.P., Häussler, B., Hopkins, P.F., Huang, J.-S., Huang, K.-H., Jha, S.W., Kartaltepe, J.S., Kirshner, R.P., Koo, D.C., Lai, K., Lee, K.-S., Li, W., Lotz, J.M., Lucas, R.A., Madau, P., McCarthy, P.J., McGrath, E.J., McIntosh, D.H., McLure, R.J., Mobasher, B., Moustakas, L.A., Mozena, M., Nandra, K., Newman, J.A., Niemi, S.-M., Noeske, K.G., Papovich, C.J., Pentericci, L., Pope, A., Primack, J.R., Rajan, A., Ravindranath, S., Reddy, N.A., Renzini, A., Rix, H.-W., Robaina, A.R., Rodney, S.A., Rosario, D.J., Rosati, P., Salimbeni, S., Scarlata, C., Siana, B., Simard, L., Smidt, J., Somerville, R.S., Spinrad, H., Straughn, A.N., Strolger, L.-G., Telford, O., Teplitz, H.J., Trump, J.R., van der Wel, A., Villforth, C., Wechsler, R.H., Weiner, B.J., Wiklind, T., Wild, V., Wilson, G., Wuyts, S., Yan, H.-J., Yun, M.S.: CANDELS: the cosmic assembly near-infrared deep extragalactic legacy survey. *Astrophys. J. Suppl. Ser.* **197**, 35 (2011). <https://doi.org/10.1088/0067-0049/197/2/35>. 1105.3753
- Guo, Q., White, S., Boylan-Kolchin, M., De Lucia, G., Kauffmann, G., Lemson, G., Li, C., Springel, V., Weinmann, S.: From dwarf spheroidals to cD galaxies: simulating the galaxy population in a Λ CDM cosmology. *Mon. Not. R. Astron. Soc.* **413**, 101–131 (2011). <https://doi.org/10.1111/j.1365-2966.2010.18114.x>. 1006.0106
- Gupta, A., Yuan, T., Torrey, P., Vogelsberger, M., Martizzi, D., Tran, K.-V.H., Kewley, L.J., Marinacci, F., Nelson, D., Pillepich, A., Hernquist, L., Genel, S., Springel, V.: Chemical pre-processing of cluster galaxies over the past 10 billion years in the IllustrisTNG simulations. *Mon. Not. R. Astron. Soc.* **477**, 35–39 (2018). <https://doi.org/10.1093/mnras/sly037>. 1801.03500
- Habouzit, M., Genel, S., Somerville, R.S., Kocevski, D., Hirschmann, M., Dekel, A., Choi, E., Nelson, D., Pillepich, A., Torrey, P., Hernquist, L., Vogelsberger, M., Weinberger, R., Springel, V.: Linking galaxy structural properties and star formation activity with IllustrisTNG (2018). ArXiv e-prints [1809.05588](https://arxiv.org/abs/1809.05588)
- Hopkins, P.F., Wetzel, A., Kereš, D., Faucher-Giguère, C.-A., Quataert, E., Boylan-Kolchin, M., Murray, N., Hayward, C.C., Garrison-Kimmel, S., Hummels, C., Feldmann, R., Torrey, P., Ma, X., Anglés-Alcázar, D., Su, K.-Y., Orr, M., Schmitz, D., Escala, I., Sanderson, R., Grudic, M.Y., Hafen, Z., Kim, J.-H., Fitts, A., Bullock, J.S., Wheeler, C., Chan, T.K., Elbert, O.D., Narayanan, D.: FIRE-2 simulations: physics versus numerics in galaxy formation. *Mon. Not. R. Astron. Soc.* **480**, 800–863 (2018). <https://doi.org/10.1093/mnras/sty1690>. 1702.06148
- Huertas-Company, M., Rodriguez-Gomez, V., Nelson, D., Pillepich, A., Bernardi, M., Domínguez-Sánchez, H., Genel, S., Pakmor, R., Snyder, G.F., Vogelsberger, M.: The hubble sequence at $z \sim 0$ in the IllustrisTNG simulation with deep learning (2019). ArXiv e-prints [1903.07625](https://arxiv.org/abs/1903.07625)
- Katz, N., Hernquist, L., Weinberg, D.H.: Galaxies and gas in a cold dark matter universe. *Astrophys. J. Lett.* **399**, 109–112 (1992). <https://doi.org/10.1086/186619>
- Kelley, L.Z., Blecha, L., Hernquist, L.: Massive black hole binary mergers in dynamical galactic environments. *Mon. Not. R. Astron. Soc.* **464**, 3131–3157 (2017). <https://doi.org/10.1093/mnras/stw2452>. 1606.01900
- Khandai, N., Di Matteo, T., Croft, R., Wilkins, S.M., Feng, Y., Tucker, E., DeGraf, C., Liu, M.-S.: The MassiveBlack-II simulation: the evolution of halos and galaxies to $z \sim 0$ (2014). Preprint [1402.0888](https://arxiv.org/abs/1402.0888)
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., Willing, C.: Jupyter notebooks—a publishing format for reproducible computational workflows. In: Loizides, F., Schmidt, B. (eds.) *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pp. 87–90. IOS Press, Amsterdam (2016)
- Klypin, A., Prada, F., Comparat, J.: Skies and universes: accessing cosmological simulations and theoretical predictions. (2017). ArXiv e-prints [1711.01453](https://arxiv.org/abs/1711.01453)
- Klypin, A.A., Trujillo-Gomez, S., Primack, J.: Dark matter halos in the standard cosmological model: results from the bolshoi simulation. *Acad. Publ. J.* **740**, 102 (2011). <https://doi.org/10.1088/0004-637X/740/2/102>. 1002.3660
- Kriek, M., Shapley, A.E., Reddy, N.A., Siana, B., Coil, A.L., Mobasher, B., Freeman, W.R., de Groot, L., Price, S.H., Sanders, R., Shivali, I., Brammer, G.B., Momcheva, I.G., Skelton, R.E., van Dokkum, P.G., Whitaker, K.E., Aird, J., Azadi, M., Kassis, M., Bullock, J.S., Conroy, C., Davé, R., Kereš, D., Krumholz, M.: The MOSFIRE deep evolution field (MOSDEF) survey: rest-frame optical spectroscopy for ~ 1500 H-selected galaxies at $1.37 < z < 3.8$. *Astrophys. J. Suppl. Ser.* **218**, 15 (2015). <https://doi.org/10.1088/0067-0049/218/2/15>. 1412.1835
- Kurtzer, G.M., Sochat, V., Bauer, M.W.: Singularity: scientific containers for mobility of compute. *PLoS ONE* **12**(5), 1–20 (2017). <https://doi.org/10.1371/journal.pone.0177459>
- Lemson, G., Bourges, L., Cervino, M., Gheller, C., Gray, N., LePetit, F., Louys, M., Ooghe, B., Wagner, R., Wozniak, H.: IVOA recommendation: simulation data model (2014). ArXiv e-prints [1402.4744](https://arxiv.org/abs/1402.4744)
- Lemson, G., Virgo Consortium, T.: Halo and galaxy formation histories from the millennium simulation: public release of a VO-oriented and SQL-queryable database for studying the evolution of galaxies in the LambdaCDM cosmogony. (2006). ArXiv Astrophysics e-prints [astro-ph/0608019](https://arxiv.org/abs/astro-ph/0608019)
- Lemson, G., Zuther, J.: Theory in the virtual observatory. *MEMSAI* **80**, 342 (2009)
- Lovell, M.R., Pillepich, A., Genel, S., Nelson, D., Springel, V., Pakmor, R., Marinacci, F., Weinberger, R., Torrey, P., Vogelsberger, M., Alabi, A., Hernquist, L.: The fraction of dark matter within galaxies from the IllustrisTNG simulations. *Mon. Not. R. Astron. Soc.* **481**, 1950–1975 (2018). <https://doi.org/10.1093/mnras/sty2339>. 1801.10170
- Makiya, R., Enoki, M., Ishiyama, T., Kobayashi, M.A.R., Nagashima, M., Okamoto, T., Okoshi, K., Oogi, T., Shirakata, H.: The new numerical galaxy catalog (ν^2 GC): an updated semi-analytic model of galaxy and active galactic nucleus formation with large cosmological N-body simulations. *Publ. Astron. Soc. Jpn.* **68**, 25 (2016). <https://doi.org/10.1093/pasj/psw005>. 1508.07215
- Marinacci, F., Vogelsberger, M., Pakmor, R., Torrey, P., Springel, V., Hernquist, L., Nelson, D., Weinberger, R., Pillepich, A., Naiman, J., Genel, S.: First results from the IllustrisTNG simulations: radio haloes and magnetic fields. *Mon.*

- Not. R. Astron. Soc. **480**, 5113–5139 (2018). <https://doi.org/10.1093/mnras/sty2206>. 1707.03396
- Martizzi, D., Vogelsberger, M., Artale, M.C., Haider, M., Torrey, P., Marinacci, F., Nelson, D., Pillepich, A., Weinberger, R., Hernquist, L., Naiman, J., Springel, V.: Baryons in the CosmicWeb of IllustrisTNG - I: gas in knots, filaments, sheets and voids (2018). ArXiv e-prints [1810.01883](https://arxiv.org/abs/1810.01883)
- McAlpine, S., Helly, J.C., Schaller, M., Trayford, J.W., Qu, Y., Furlong, M., Bower, R.G., Crain, R.A., Schaye, J., Theuns, T., Dalla Vecchia, C., Frenk, C.S., McCarthy, I.G., Jenkins, A., Rosas-Guevara, Y., White, S.D.M., Baes, M., Camps, P., Lemson, G.: The EAGLE simulations of galaxy formation: public release of halo and galaxy catalogues. *Astron. Comput.* **15**, 72–89 (2016). <https://doi.org/10.1016/j.ascom.2016.02.004>. 1510.01320
- Medvedev, D., Lemson, G., Rippin, M.: Sciserver compute: bringing analysis close to the data. In: Proceedings of the 28th International Conference on Scientific and Statistical Database Management, SSDBM 2016, Budapest, Hungary, July 18–20, 2016, pp. 27–1274 (2016). <https://doi.org/10.1145/2949689.2949700>.
- Merkel, D.: Docker: lightweight Linux containers for consistent development and deployment. *Linux J.* **2014**, 239 (2014)
- Naiman, J.P., Pillepich, A., Springel, V., Ramirez-Ruiz, E., Torrey, P., Vogelsberger, M., Pakmor, R., Nelson, D., Marinacci, F., Hernquist, L., Weinberger, R., Genel, S.: First results from the IllustrisTNG simulations: a tale of two elements—chemical evolution of magnesium and europium. *Mon. Not. R. Astron. Soc.* **477**, 1206–1224 (2018). <https://doi.org/10.1093/mnras/sty618>. 1707.03401
- Nelson, D., Kauffmann, G., Pillepich, A., Genel, S., Springel, V., Pakmor, R., Hernquist, L., Weinberger, R., Torrey, P., Vogelsberger, M., Marinacci, F.: The abundance, distribution, and physical nature of highly ionized oxygen O_{VI} , O_{VII} , and O_{VIII} in IllustrisTNG. *Mon. Not. R. Astron. Soc.* **477**, 450–479 (2018b). <https://doi.org/10.1093/mnras/sty656>. 1712.00016
- Nelson, D., Pillepich, A., Genel, S., Vogelsberger, M., Springel, V., Torrey, P., Rodriguez-Gomez, V., Sijacki, D., Snyder, G.F., Griffen, B., Marinacci, F., Blecha, L., Sales, L., Xu, D., Hernquist, L.: The Illustris simulation: public data release. *Astron. Comput.* **13**, 12–37 (2015). <https://doi.org/10.1016/j.ascom.2015.09.003>. 1504.00362
- Nelson, D., Pillepich, A., Springel, V., Pakmor, R., Weinberger, R., Genel, S., Torrey, P., Vogelsberger, M., Marinacci, F., Hernquist, L.: First results from the TNG50 simulation: galactic outflows driven by supernovae and black hole feedback (2019a). ArXiv e-prints [1902.05554](https://arxiv.org/abs/1902.05554)
- Nelson, D., Pillepich, A., Springel, V., Weinberger, R., Hernquist, L., Pakmor, R., Genel, S., Torrey, P., Vogelsberger, M., Kauffmann, G., Marinacci, F., Naiman, J.: First results from the IllustrisTNG simulations: the galaxy colour bimodality. *Mon. Not. R. Astron. Soc.* **475**, 624–647 (2018a). <https://doi.org/10.1093/mnras/stx3040>. 1707.03395
- Nelson, D., Springel, V., Pillepich, A., Rodriguez-Gomez, V., Torrey, P., Genel, S., Vogelsberger, M., Pakmor, R., Marinacci, F., Weinberger, R., Kelley, L., Lovell, M., Diemer, B., Hernquist, L.: The IllustrisTNG simulations: public data release. (2019b). ArXiv e-prints [1812.05609](https://arxiv.org/abs/1812.05609)
- Ntampaka, M., Zuhone, J., Eisenstein, D., Nagai, D., Vikhlinin, A., Hernquist, L., Marinacci, F., Nelson, D., Pakmor, R., Pillepich, A., Torrey, P., Vogelsberger, M.: A deep learning approach to galaxy cluster X-ray masses (2018). ArXiv e-prints [1810.07703](https://arxiv.org/abs/1810.07703)
- Ocvirk, P., Aubert, D., Sorce, J.G., Shapiro, P.R., Deparis, N., Dawoodbhoy, T., Lewis, J., Teyssier, R., Yepes, G., Gottlöber, S., Ahn, K., Iliev, I.T., Hoffman, Y.: Cosmic Dawn II (CoDa II): a new radiation-hydrodynamics simulation of the self-consistent coupling of galaxy formation and reionization (2018). ArXiv e-prints [1811.11192](https://arxiv.org/abs/1811.11192)
- Oman, K.A., Marasco, A., Navarro, J.F., Frenk, C.S., Schaye, J., Benítez-Llambay, A.: Non-circular motions and the diversity of dwarf galaxy rotation curves. *Mon. Not. R. Astron. Soc.* **482**, 821–847 (2019). <https://doi.org/10.1093/mnras/sty2687>. 1706.07478
- Overzier, R., Lemson, G., Angulo, R.E., Bertin, E., Blaizot, J., Henriques, B.M.B., Marleau, G.-D., White, S.D.M.: The millennium run observatory: first light. *Mon. Not. R. Astron. Soc.* **428**, 778–803 (2013). <https://doi.org/10.1093/mnras/sts076>. 1206.6923
- Pakmor, R., Bauer, A., Springel, V.: Magnetohydrodynamics on an unstructured moving grid. *Mon. Not. R. Astron. Soc.* **418**, 1392–1401 (2011). <https://doi.org/10.1111/j.1365-2966.2011.19591.x>. 1108.1792
- Pakmor, R., Springel, V.: Simulations of magnetic fields in isolated disc galaxies. *Mon. Not. R. Astron. Soc.* **432**, 176–193 (2013). <https://doi.org/10.1093/mnras/stt428>. 1212.1452
- Pakmor, R., Springel, V., Bauer, A., Mocz, P., Munoz, D.J., Ohlmann, S.T., Schaaf, K., Zhu, C.: Improving the convergence properties of the moving-mesh code AREPO. *Mon. Not. R. Astron. Soc.* **455**, 1134–1143 (2016). <https://doi.org/10.1093/mnras/stv2380>. 1503.00562
- Pérez, F., Granger, B.E.: IPython: a system for interactive scientific computing. *Comput. Sci. Eng.* **9**(3), 21–29 (2007). <https://doi.org/10.1109/MCSE.2007.53>
- Pillepich, A., Nelson, D., Hernquist, L., Springel, V., Pakmor, R., Torrey, P., Weinberger, R., Genel, S., Naiman, J.P., Marinacci, F., Vogelsberger, M.: First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies. *Mon. Not. R. Astron. Soc.* **475**, 648–675 (2018a). <https://doi.org/10.1093/mnras/stx3112>. 1707.03406
- Pillepich, A., Nelson, D., Springel, V., Pakmor, R., Torrey, P., Weinberger, R., Vogelsberger, M., Marinacci, F., Genel, S., van der Wel, A., Hernquist, L.: First results from the TNG50 simulation: the evolution of stellar and gaseous disks across cosmic time pp. 1902–05553 (2019). ArXiv e-prints [1902.05553](https://arxiv.org/abs/1902.05553)
- Pillepich, A., Springel, V., Nelson, D., Genel, S., Naiman, J., Pakmor, R., Hernquist, L., Torrey, P., Vogelsberger, M., Weinberger, R., Marinacci, F.: Simulating galaxy formation with the IllustrisTNG model. *Mon. Not. R. Astron. Soc.* **473**, 4077–4106 (2018b). <https://doi.org/10.1093/mnras/stx2656>. 1703.02970
- Planck Collaboration: Planck 2015 results. XIII. Cosmological parameters. *A&A* **594**, 13 (2016). <https://doi.org/10.1051/0004-6361/201525830>. 1502.01589
- Pontzen, A., Tremmel, M.: TANGOS: the agile numerical galaxy organization system. *Astrophys. J. Suppl. Ser.* **237**, 23 (2018). <https://doi.org/10.3847/1538-4365/aac832>. 1803.00010
- Popping, G., Pillepich, A., Somerville, R.S., Decarli, R., Walter, F., Aravena, M., Carilli, C., Cox, P., Nelson, D., Riechers, D., Weiss, A., Boogaard, L., Bouwens, R., Contini, T., Cortes, P.C., da Cunha, E., Daddi, E., Díaz-Santos, T., Diemer, B., González-López, J., Hernquist, L., Ivison, R., Le Fevre, O., Marinacci, F., Rix, H.-W., Swinbank, M., Vogelsberger, M., van der Werf, P., Wagg, J., Yung, L.Y.A.: The ALMA spectroscopic survey in the HUDF: the molecular gas content of galaxies and tensions with IllustrisTNG and the Santa Cruz SAM (2019). ArXiv e-prints [1903.09158](https://arxiv.org/abs/1903.09158)
- Press, W.H., Schechter, P.: Formation of galaxies and clusters of galaxies by self-similar gravitational condensation. *Acad. Publ. J.* **187**, 425–438 (1974). <https://doi.org/10.1086/152650>
- Raddick, J., Souter, B., Lemson, G., Taghizadeh-Popp, M.: SciServer: An Online Collaborative Environment for Big Data in Research and Education. *American Astronomical Society Meeting Abstracts* **229**, 236.15 (2017)
- Ragagnin, A., Dolag, K., Biffi, V., Cadolle Bel, M., Hammer, N.J., Krukau, A., Petkova, M., Steinborn, D.: A web portal for hydrodynamical, cosmological simulations. *Astron. Comput.* **20**, 52–67 (2017). <https://doi.org/10.1016/j.ascom.2017.05.001>. 1612.06380
- Rasera, Y., Alimi, J.-M., Courtin, J., Roy, F., Corasaniti, P.-S., Füzfa, A., Boucher, V.: Introducing the dark energy universe simulation series (DEUSS). In: Alimi, J.-M., Fuözfa, A. (eds.) *American Institute of Physics Conference Series*. American Institute of Physics Conference Series, vol. 1241, pp. 1134–1139 (2010). <https://doi.org/10.1063/1.3462610>. 1002.4950
- Riebe, K., Partl, A.M., Enke, H., Forero-Romero, J., Gottlöber, S., Klypin, A., Lemson, G., Prada, F., Primack, J.R., Steinmetz, M., Turchaninov, V.: The MultiDark database: release of the bolshoi and MultiDark cosmological simulations. *Astron. Nachr.* **334**, 691–708 (2013). <https://doi.org/10.1002/asna.201211900>
- Rodriguez-Gomez, V., Genel, S., Vogelsberger, M., Sijacki, D., Pillepich, A., Sales, L.V., Torrey, P., Snyder, G., Nelson, D., Springel, V., Ma, C.-P., Hernquist, L.: The merger rate of galaxies in the Illustris simulation: a comparison with observations and semi-empirical models. *Mon. Not. R. Astron. Soc.* **449**, 49–64 (2015). <https://doi.org/10.1093/mnras/stv264>. 1502.01339
- Rodriguez-Gomez, V., Pillepich, A., Sales, L.V., Genel, S., Vogelsberger, M., Zhu, Q., Wellons, S., Nelson, D., Torrey, P., Springel, V., Ma, C.-P., Hernquist, L.: The stellar mass assembly of galaxies in the Illustris simulation: growth by mergers and the spatial distribution of accreted stars. *Mon. Not. R. Astron. Soc.* **458**, 2371–2390 (2016). <https://doi.org/10.1093/mnras/stw456>. 1511.08804
- Rodriguez-Gomez, V., Sales, L.V., Genel, S., Pillepich, A., Zjupa, J., Nelson, D., Griffen, B., Torrey, P., Snyder, G.F., Vogelsberger, M., Springel, V., Ma, C.-P., Hernquist, L.: The role of mergers and halo spin in shaping galaxy morphology. *Mon. Not. R. Astron. Soc.* **467**, 3083–3098 (2017). <https://doi.org/10.1093/mnras/stx305>. 1609.09498

- Rodriguez-Gomez, V., Snyder, G.F., Lotz, J.M., Nelson, D., Pillepich, A., Springel, V., Genel, S., Weinberger, R., Tacchella, S., Pakmor, R., Torrey, P., Marinacci, F., Vogelsberger, M., Hernquist, L., Thücker, D.A.: The optical morphologies of galaxies in the IllustrisTNG simulation: a comparison to Pan-STARRS observations. *Mon. Not. R. Astron. Soc.* **483**, 4140–4159 (2019). <https://doi.org/10.1093/mnras/sty3345>. 1809.08239
- Rosdahl, J., Katz, H., Blaizot, J., Kimm, T., Michel-Dansac, L., Garel, T., Haehnelt, M., Ocvirk, P., Teysseier, R.: The SPHINX cosmological simulations of the first billion years: the impact of binary stars on reionization. *Mon. Not. R. Astron. Soc.* **479**, 994–1016 (2018). <https://doi.org/10.1093/mnras/sty1655>. 1801.07259
- Sánchez, S.F., Kennicutt, R.C., Gil de Paz, A., van de Ven, G., Vilchez, J.M., Wisotzki, L., Walcher, C.J., Mast, D., Aguerri, J.A.L., Albiol-Pérez, S., Alonso-Herrero, A., Alves, J., Bakos, J., Bartkóvá, T., Bland-Hawthorn, J., Boselli, A., Bomans, D.J., Castillo-Morales, A., Cortijo-Ferrero, C., de Lorenzo-Cáceres, A., Del Olmo, A., Dettmar, R.-J., Díaz, A., Ellis, S., Falcón-Barroso, J., Flores, H., Gallazzi, A., García-Lorenzo, B., González Delgado, R., Gruel, N., Haines, T., Hao, C., Husemann, B., Iglésias-Páramo, J., Jahnke, K., Johnson, B., Jungwiert, B., Kalinova, V., Kehrig, C., Kupko, D., López-Sánchez, Á.R., Lyubenova, M., Marino, R.A., Mármod-Queraltó, E., Márquez, I., Masegosa, J., Meidt, S., Mendez-Abreu, J., Monreal-Ibero, A., Montijo, C., Mourão, A.M., Palacios-Navarro, G., Papaderos, P., Pasquali, A., Peletier, R., Pérez, E., Pérez, I., Quirrenbach, A., Relaño, M., Rosales-Ortega, F.F., Roth, M.M., Ruiz-Lara, T., Sánchez-Blázquez, P., Sengupta, C., Singh, R., Stanishev, V., Trager, S.C., Vazdekis, A., Viironen, K., Wild, V., Zibetti, S., Ziegler, B.: CALIFA, the calar alto legacy integral field area survey. I. Survey presentation. *A&A* **538**, 8 (2012). <https://doi.org/10.1051/0004-6361/201117353>. 1111.0962
- Schaal, K., Springel, V.: Shock finding on a moving mesh—I. Shock statistics in non-radiative cosmological simulations. *Mon. Not. R. Astron. Soc.* **446**, 3992–4007 (2015). <https://doi.org/10.1093/mnras/stu2386>. 1407.4117
- Schaal, K., Springel, V., Pakmor, R., Pfrommer, C., Nelson, D., Vogelsberger, M., Genel, S., Pillepich, A., Sijacki, D., Hernquist, L.: Shock finding on a moving-mesh—II. Hydrodynamic shocks in the Illustris universe. *Mon. Not. R. Astron. Soc.* **461**, 4441–4465 (2016). <https://doi.org/10.1093/mnras/stw1587>. 1604.07401
- Schaye, J., Crain, R.A., Bower, R.G., Furlong, M., Schaller, M., Theuns, T., Dalla Vecchia, C., Frenk, C.S., McCarthy, I.G., Helly, J.C., Jenkins, A., Rosas-Guevara, Y.M., White, S.D.M., Baes, M., Booth, C.M., Camps, P., Navarro, J.F., Qu, Y., Rahmati, A., Sawala, T., Thomas, P.A., Trayford, J.: The EAGLE project: simulating the evolution and assembly of galaxies and their environments. *Mon. Not. R. Astron. Soc.* **446**, 521–554 (2015). <https://doi.org/10.1093/mnras/stu2058>. 1407.7040
- Shen, S., Cooke, R.J., Ramirez-Ruiz, E., Madau, P., Mayer, L., Guedes, J.: The history of R-process enrichment in the Milky Way. *Acad. Publ. J.* **807**, 115 (2015). 1407.3796
- Sijacki, D., Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Snyder, G.F., Nelson, D., Hernquist, L.: The Illustris simulation: the evolving population of black holes across cosmic time. *Mon. Not. R. Astron. Soc.* **452**, 575–596 (2015). <https://doi.org/10.1093/mnras/stv1340>. 1408.6842
- Skillman, S.W., Warren, M.S., Turk, M.J., Wechsler, R.H., Holz, D.E., Sutter, P.M.: Dark sky simulations: early data release (2014). ArXiv e-prints 1407.2600
- Springel, V.: E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. *Mon. Not. R. Astron. Soc.* **401**, 791–851 (2010). <https://doi.org/10.1111/j.1365-2966.2009.15715.x>. 0901.4107
- Springel, V.: NGenIC: cosmological structure initial conditions. *Astrophysics Source Code Library*, (2015). 1502.003
- Springel, V., Di Matteo, T., Hernquist, L.: Modelling feedback from stars and black holes in galaxy mergers. *Mon. Not. R. Astron. Soc.* **361**, 776–794 (2005). <https://doi.org/10.1111/j.1365-2966.2005.09238.x>. astro-ph/0411108
- Springel, V., Hernquist, L.: Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation. *Mon. Not. R. Astron. Soc.* **339**, 289–311 (2003). <https://doi.org/10.1046/j.1365-8711.2003.06206.x>. astro-ph/0206393
- Springel, V., Pakmor, R., Pillepich, A., Weinberger, R., Nelson, D., Hernquist, L., Vogelsberger, M., Genel, S., Torrey, P., Marinacci, F., Naiman, J.: First results from the IllustrisTNG simulations: matter and galaxy clustering. *Mon. Not. R. Astron. Soc.* **475**, 676–698 (2018). <https://doi.org/10.1093/mnras/stx3304>. arXiv:1707.03397
- Springel, V., White, S.D.M., Jenkins, A., Frenk, C.S., Yoshida, N., Gao, L., Navarro, J., Thacker, R., Croton, D., Helly, J., Peacock, J.A., Cole, S., Thomas, P., Couchman, H., Evrard, A., Colberg, J., Pearce, F.: Simulations of the formation, evolution and clustering of galaxies and quasars. *Nature* **435**, 629–636 (2005). <https://doi.org/10.1038/nature03597>. astro-ph/0504097
- Springel, V., White, S.D.M., Tormen, G., Kauffmann, G.: Populating a cluster of galaxies—I. Results at $z = 0$. *Mon. Not. R. Astron. Soc.* **328**, 726–750 (2001). <https://doi.org/10.1046/j.1365-8711.2001.04912.x>. astro-ph/0012055
- Steidel, C.C., Rudie, G.C., Strom, A.L., Pettini, M., Reddy, N.A., Shapley, A.E., Trainor, R.F., Erb, D.K., Turner, M.L., Konidaris, N.P., Kulas, K.R., Mace, G., Matthews, K., McLean, I.S.: Strong nebular line ratios in the spectra of $z \sim 2$ –3 star forming galaxies: first results from KBSS-MOSFIRE. *Acad. Publ. J.* **795**, 165 (2014). <https://doi.org/10.1088/0004-637X/795/2/165>. 1405.5473
- Stevens, A.R.H., Diemer, B., Lagos, C.d.P., Nelson, D., Pillepich, A., Brown, T., Catinella, B., Hernquist, L., Weinberger, R., Vogelsberger, M., Marinacci, F.: Atomic hydrogen in IllustrisTNG galaxies: the impact of environment paralleled with local 21-cm surveys (2018). ArXiv e-prints 1810.12158
- Szalay, A.S., Gray, J., Thakar, A.R., Kunz, P.Z., Malik, T., Raddick, J., Stoughton, C., vandenBerg, J.: The SDSS SkyServer: public access to the Sloan Digital Sky Server Data (2002). e-prints cs/0202013
- Szalay, A.S., Gray, J., vandenBerg, J.: Petabyte scale data mining: dream or reality? In: Tyson, J.A., Wolff, S. (eds.) *Survey and Other Telescope Technologies and Discoveries*. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, vol. 4836, pp. 333–338 (2002). <https://doi.org/10.1117/12.461427>. cs/0208013
- Szalay, A.S., Kunz, P.Z., Thakar, A.R., Gray, J., Slutz, D.: The Sloan Digital Sky Survey and its archive. In: Manset, N., Veillet, C., Crabtree, D. (eds.) *Astronomical Data Analysis Software and Systems IX*. Astronomical Society of the Pacific Conference Series, vol. 216, p. 405 (2000). astro-ph/9912382
- The EAGLE team: The EAGLE simulations of galaxy formation: public release of particle data. (2017). ArXiv e-prints 1706.09899
- Torrey, P., Vogelsberger, M., Genel, S., Sijacki, D., Springel, V., Hernquist, L.: A model for cosmological simulations of galaxy formation physics: multi-epoch validation. *Mon. Not. R. Astron. Soc.* **438**, 1985–2004 (2014). <https://doi.org/10.1093/mnras/stt2295>. 1305.4931
- Torrey, P., Vogelsberger, M., Hernquist, L., McKinnon, R., Marinacci, F., Simcoe, R.A., Springel, V., Pillepich, A., Naiman, J., Pakmor, R., Weinberger, R., Nelson, D., Genel, S.: Similar star formation rate and metallicity variability time-scales drive the fundamental metallicity relation. *Mon. Not. R. Astron. Soc.* **477**, 16–20 (2018). <https://doi.org/10.1093/mnras/sly031>. 1711.11039
- Torrey, P., Vogelsberger, M., Marinacci, F., Pakmor, R., Springel, V., Nelson, D., Naiman, J., Pillepich, A., Genel, S., Weinberger, R., Hernquist, L.: The evolution of the mass-metallicity relation in IllustrisTNG (2017). ArXiv e-prints 1711.05261
- Tremmel, M., Karcher, M., Governato, F., Volonteri, M., Quinn, T.R., Pontzen, A., Anderson, L., Bellovary, J.: The romulus cosmological simulations: a physical approach to the formation, dynamics and accretion models of SMBHs. *Mon. Not. R. Astron. Soc.* **470**, 1121–1139 (2017). <https://doi.org/10.1093/mnras/stx1160>. 1607.02151
- van der Wel, A., Noeske, K., Bezanson, R., Pacifici, C., Gallazzi, A., Franx, M., Muñoz-Mateos, J.C., Bell, E.F., Brammer, G., Charlot, S., Chauk, P., Labbé, I., Maseda, M.V., Muzzin, A., Rix, H.-W., Sobral, D., van de Sande, J., van Dokkum, P.G., Wild, V., Wolf, C.: The VLT LEGA-C spectroscopic survey: the physics of galaxies at a lookback time of 7 yr. *Astrophys. J. Suppl. Ser.* **223**, 29 (2016). <https://doi.org/10.3847/0067-0049/223/2/29>. 1603.05479
- Villaescusa-Navarro, F., Genel, S., Castorina, E., Obuljen, A., Spergel, D.N., Hernquist, L., Nelson, D., Carucci, I.P., Pillepich, A., Marinacci, F., Diemer, B., Vogelsberger, M., Weinberger, R., Pakmor, R.: Ingredients for 21 cm intensity mapping. *Acad. Publ. J.* **866**, 135 (2018). <https://doi.org/10.3847/1538-4357/aadba0>. 1804.09180
- Vogelsberger, M., Genel, S., Sijacki, D., Torrey, P., Springel, V., Hernquist, L.: A model for cosmological simulations of galaxy formation physics. *Mon. Not. R. Astron. Soc.* **436**, 3031–3067 (2013). <https://doi.org/10.1093/mnras/stt1789>. 1305.2913
- Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Bird, S., Nelson, D., Hernquist, L.: Properties of galaxies reproduced by a hydrodynamic simulation. *Nature* **509**, 177–182 (2014a). <https://doi.org/10.1038/nature13316>. 1405.1418
- Vogelsberger, M., Genel, S., Springel, V., Torrey, P., Sijacki, D., Xu, D., Snyder, G., Nelson, D., Hernquist, L.: Introducing the Illustris project: simulating the coevolution of dark and visible matter in the universe. *Mon. Not. R. Astron. Soc.* **444**, 1518–1547 (2014b). <https://doi.org/10.1093/mnras/stu1536>. 1405.2921
- Vogelsberger, M., Marinacci, F., Torrey, P., Genel, S., Springel, V., Weinberger, R., Pakmor, R., Hernquist, L., Naiman, J., Pillepich, A., Nelson, D.: The uniformity and time-invariance of the intra-cluster metal distribution in galaxy

- clusters from the IllustrisTNG simulations. *Mon. Not. R. Astron. Soc.* **474**, 2073–2093 (2018). <https://doi.org/10.1093/mnras/stx2955>. 1707.05318
- Wang, L., Dutton, A.A., Stinson, G.S., Macciò, A.V., Penzo, C., Kang, X., Keller, B.W., Wadsley, J.: NIHAO project—I. Reproducing the inefficiency of galaxy formation across cosmic time with a large sample of cosmological hydrodynamical simulations. *Mon. Not. R. Astron. Soc.* **454**, 83–94 (2015). <https://doi.org/10.1093/mnras/stv1937>. 1503.04818
- Wang, Y., Su, Y., Agrawal, G.: Supporting a light-weight data management layer over HDF5. *Cluster, Cloud and Grid Computing (CCGrid)* **2013**, 335–342 (2013). <https://doi.org/10.1109/CCGrid.2013.9>
- Wang, Y., Vogelsberger, M., Xu, D., Mao, S., Springel, V., Li, H., Barnes, D., Hernquist, L., Pillepich, A., Marinacci, F., Pakmor, R., Weinberger, R., Torrey, P.: Early-type galaxy density profiles from IllustrisTNG: I. Galaxy correlations and the impact of baryons (2018). ArXiv e-prints 1811.06545
- Weinberger, R., Springel, V., Hernquist, L., Pillepich, A., Marinacci, F., Pakmor, R., Nelson, D., Genel, S., Vogelsberger, M., Naiman, J., Torrey, P.: Simulating galaxy formation with black hole driven thermal and kinetic feedback. *Mon. Not. R. Astron. Soc.* **465**, 3291–3308 (2017). <https://doi.org/10.1093/mnras/stw2944>. 1607.03486
- Weinberger, R., Springel, V., Pakmor, R., Nelson, D., Genel, S., Pillepich, A., Vogelsberger, M., Marinacci, F., Naiman, J., Torrey, P., Hernquist, L.: Supermassive black holes and their feedback effects in the IllustrisTNG simulation. *Mon. Not. R. Astron. Soc.* **479**, 4056–4072 (2018). <https://doi.org/10.1093/mnras/sty1733>. 1710.04659
- Wisnioski, E., Förster Schreiber, N.M., Wuyts, S., Wuyts, E., Bandara, K., Wilman, D., Genzel, R., Bender, R., Davies, R., Fossati, M., Lang, P., Mendel, J.T., Beifiori, A., Brammer, G., Chan, J., Fabricius, M., Fudamoto, Y., Kulkarni, S., Kurk, J., Lutz, D., Nelson, E.J., Momcheva, I., Rosario, D., Saglia, R., Seitz, S., Tacconi, L.J., van Dokkum, P.G.: The KMOS^{3D} survey: design, first results, and the evolution of galaxy kinematics from $0.7 \leq z \leq 2.7$. *Acad. Publ. J.* **799**, 209 (2015). <https://doi.org/10.1088/0004-637X/799/2/209>. 1409.6791
- York, D.G., Adelman, J., Anderson, J.E., Anderson, S.F., Annis, J., Bahcall, N.A., Bakken, J.A., Barkhouser, R., Bastian, S., Berman, E., Boroski, W.N., Bracker, S., Briegel, C., Briggs, J.W., Brinkmann, J., Brunner, R., Burles, S., Carey, L., Carr, M.A., Castander, F.J., Chen, B., Colestock, P.L., Connolly, A.J., Crocker, J.H., Csabai, I., Czarapata, P.C., Davis, J.E., Doi, M., Dombeck, T., Eisenstein, D., Ellman, N., Elms, B.R., Evans, M.L., Fan, X., Federwitz, G.R., Fiscelli, L., Friedman, S., Frieman, J.A., Fukugita, M., Gillespie, B., Gunn, J.E., Gurbani, V.K., de Haas, E., Haldeman, M., Harris, F.H., Hayes, J., Heckman, T.M., Hennessy, G.S., Hindsley, R.B., Holm, S., Holmgren, D.J., Huang, C.-H., Hull, C., Husby, D., Ichikawa, S.-I., Ichikawa, T., Ivezić, Ž., Kent, S., Kim, R.S.J., Kinney, E., Klaene, M., Kleinman, A.N., Kleinman, S., Knapp, G.R., Korienek, J., Kron, R.G., Kunzst, P.Z., Lamb, D.Q., Lee, B., Leger, R.F., Limmongkol, S., Lindenmeyer, C., Long, D.C., Loomis, C., Loveday, J., Lucinio, R., Lupton, R.H., MacKinnon, B., Mannery, E.J., Mantsch, P.M., Margon, B., McGehee, P., McKay, T.A., Meiksin, A., Merelli, A., Monet, D.G., Munn, J.A., Narayanan, V.K., Nash, T., Neilsen, E., Neswold, R., Newberg, H.J., Nichol, R.C., Nicinski, T., Nonino, M., Okada, N., Okamura, S., Ostriker, J.P., Owen, R., Pauls, A.G., Peoples, J., Peterson, R.L., Petravick, D., Pier, J.R., Pope, A., Pordes, R., Protopop, A., Rechenmacher, R., Quinn, T.R., Richards, G.T., Richmond, M.W., Rivetta, C.H., Rockosi, C.M., Ruthmansdorfer, K., Sandford, D., Schlegel, D.J., Schneider, D.P., Sekiguchi, M., Sergey, G., Shimasaku, K., Siegmund, W.A., Smee, S., Smith, J.A., Snedden, S., Stone, R., Stoughton, C., Strauss, M.A., Stubbs, C., SubbaRao, M., Szalay, A.S., Szapudi, I., Szokoly, G.P., Thakar, A.R., Tremonti, C., Tucker, D.L., Uomoto, A., Vanden Berk, D., Vogeley, M.S., Waddell, P., Wang, S.-I., Watanabe, M., Weinberg, D.H., Yanny, B., Yasuda, N.: SDSS collaboration: the sloan digital sky survey: technical summary. *Astron. J.* **120**, 1579–1587 (2000). <https://doi.org/10.1086/301513>. astro-ph/0006396
- Yun, K., Pillepich, A., Zinger, E., Nelson, D., Donnari, M., Joshi, G., Rodriguez-Gomez, V., Genel, S., Weinberger, R., Vogelsberger, M., Hernquist, L.: Jellyfish galaxies with the IllustrisTNG simulations: I. Gas-stripping phenomena in the full cosmological context. *Mon. Not. R. Astron. Soc.* (2018). <https://doi.org/10.1093/mnras/sty3156>. 1810.00005
- Zhu, Q., Xu, D., Gaspari, M., Rodriguez-Gomez, V., Nelson, D., Vogelsberger, M., Torrey, P., Pillepich, A., Zjupa, J., Weinberger, R., Marinacci, F., Pakmor, R., Genel, S., Li, Y., Springel, V., Hernquist, L.: Formation of a malin 1 analogue in IllustrisTNG by stimulated accretion. *Mon. Not. R. Astron. Soc.* **480**, 18–22 (2018). <https://doi.org/10.1093/mnras/sly111>. 1805.09341
- Zjupa, J., Springel, V.: Angular momentum properties of haloes and their baryon content in the Illustris simulation. *Mon. Not. R. Astron. Soc.* **466**, 1625–1647 (2017). <https://doi.org/10.1093/mnras/stw2945>. 1608.01323