PhD Thesis

Milky Way and M31 analogues: insights from the cosmological simulation TNG50

Max-Planck-Institut für Astronomie
Ruprecht-Karls-Universität Heidelberg

Diego Sotillo Ramos

Dissertation

submitted to the
Combined Faculties of Mathematics, Engineering and Natural Sciences
of Heidelberg University, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Diego Sotillo Ramos
born in: Zamora, Spain
Oral examination: July 19, 2024

Milky Way and M31 analogues: insights from the cosmological simulation TNG50

Referees:
Dr. Annalisa Pillepich
Prof. Dr. Andreas Just

Abstract

In this thesis, I study key facets of Milky Way- and Andromeda-like (MW/M31-like) galaxies' formation, evolution and structure in the cosmological context. To this aim, I use TNG50, the highest resolution run of the cutting-edge state-of-the-art IllustrisTNG suite of cosmological mangeto-hydrodynamical simulations. TNG50 simulates 198 MW/M31 analogs, in an unprecedented blend of high numerical resolution and sample size. I study the impact of major mergers on galaxy stellar disks, revealing that a significant fraction of the analyzed galaxies undergo a recent major merger (in the last 5 billion years) and are still disky at $z = 0$. Among these galaxies, for two-thirds of the cases, the merger destroys the disk, but a new one is able to form until $z = 0$. In the remaining galaxies, the disk survives the merger. I analyze and quantify stellar disk flaring, i.e., the increase of vertical stellar disk height with galactocentric distance, showing a great diversity of types and values across the galaxies of the sample. But it is a complex phenomenon, difficult to predict for an individual galaxy according to its $z = 0$ global structural properties or merger history. Finally, I also investigate the presence of very metal-poor stars in the stellar disks and other morphological components of MW-like galaxies, yielding valuable insights into their origin and development: according to TNG50, there is a non-negligible fraction of these stars that populate the disk, and that are also very old, challenging therefore the notion that the stellar halo is the oldest component of the MW. Therefore, this thesis leverages the power of numerical tools like TNG50 to uncover pivotal aspects of the evolution and structure of MW/M31-like galaxies as well as their formation, providing valuable insights that should be tested with future observational studies.

Zusammenfassung

In dieser Dissertation untersuche ich wesentliche Aspekte der Entstehung, Evolution und Struktur von Galaxien, die der Milchstraße und Andromeda ähneln (MW/M31-ähnliche Galaxien), im kosmologischen Kontext. Zu diesem Zweck verwende ich TNG50, den höchstaufgelösten Lauf der modernen IllustrisTNG-Suite für kosmologische Magnetohydrodynamik-Simulationen. TNG50 simuliert 198 MW/M31-Analoga in einer beispiellosen Kombination aus hoher numerischer Auflösung und Stichprobengröße.

Ich untersuche die Auswirkungen von größeren Verschmelzungen auf die stellaren Scheiben von Galaxien und zeige auf, dass ein erheblicher Anteil der analysierten Galaxien eine jüngste größere Verschmelzung (in den letzten 5 Milliarden Jahren) durchläuft und dennoch bei $z = 0$ eine scheibenförmige Struktur aufweist. In zwei Dritteln der Fälle zerstört die Verschmelzung jedoch die Scheibe, aber eine neue kann sich bis $z = 0$ bilden. Bei den übrigen Galaxien überlebt die stellare Scheibe die Verschmelzung.

Des Weiteren analysiere und quantifiziere ich das Aufwölben stellaren Scheiben, d.h. die Zunahme der vertikalen Höhe der stellaren Scheibe in Abhängigkeit vom galaktischen Abstand, und zeige eine große Vielfalt von Typen und Werten innerhalb der Galaxien der Stichprobe. Es handelt sich jedoch um ein komplexes Phänomen, das sich schwer für eine einzelne Galaxie anhand ihrer globalen strukturellen Eigenschaften bei $z = 0$ oder ihrer Verschmelzungsgeschichte vorhersagen lässt.

Schließlich untersuche ich auch das Vorhandensein von sehr metallarmen Sternen in den stellaren Scheiben und anderen morphologischen Komponenten von MW-ähnlichen Galaxien und gewinne wertvolle Erkenntnisse über ihren Ursprung und ihre Entwicklung. Gemäß TNG50 gibt es einen nicht unerheblichen Anteil dieser Sterne, die die Scheibe bevölkern und ebenfalls sehr alt sind, was die Vorstellung herausfordert, dass der stellare Halo das älteste Element der Milchstraße ist.

Daher nutzt diese Dissertation die Leistung numerischer Werkzeuge wie TNG50, um entscheidende Aspekte der Evolution und Struktur von MW/M31-ähnlichen Galaxien sowie ihrer Entstehung aufzudecken und wertvolle Erkenntnisse zu liefern, die in zukünftigen Beobachtungsstudien getestet werden sollten.

Contents

List of Figures

List of Tables

Acknowledgements

I would like to take this opportunity to thank all the people, not few, who have supported me during all the time invested in the work summarized in this thesis. If I have forgotten anyone, please forgive me.

To my supervisor Annalisa, because you have taught me over the years how to conduct research and delve into finding explanations for scientific problems. Because you have been an ocean of wisdom, but also of calm, when required.
To Andreas and María, for agreeing to be part of the thesis committee and for the valuable advice that emerged in each of the meetings. Und danke an Matthias und Hans-Walter, dass ihr meiner Disputation als Komiteemitglieder beigewohnt habt.

Meiner Reisegefährtin Julia, weil du Geduld hattest, sowohl vor als auch während der Durchführung dieser Arbeit. Dafür, dass du an meiner Seite warst, obwohl du wusstest, dass die für diese Arbeit aufgewendete Zeit uns gestohlen war. Aaron, für deine unerschöpfliche Fröhlichkeit, obwohl du mich und meine kulinarischen Experimente fast jede Woche ertragen musstest. Y a Idun, por llegar a nuestras vidas en el momento adecuado, aunque quizá un poco en medio de tiempos convulsos que ya serán mejores. Por haber estado tan risueña todos y cada uno de los días de estos once meses que llevas con nosotros, regalando energía y buen humor a todos los que te contemplamos.

A mamá, papá y Adrián, por el apoyo en la distancia, que se hace notar a pesar de los miles de kilómetros, y por haber estado tan convencidos como yo de que mis cambios de opinión y de rumbo eran aceptables y razonables.

To my office mates, especially Ismael and Nick, for sharing so many hours of work (even during the pandemic), countless lunch times, and probably too many conversations about a wide range of bizarre topics. And of course, also for your attentive plant-caretaking efforts.
Many thanks also to all my colleagues from the GC Theory group (Martina, Eric, Maxime, Chris, Gandhali, Lukas, Elad, Matthew, Noa, María-Renée, Martin, etc.), as well as from satellite groups (Dylan, Neige, Anna, etc.), for all the interesting contributions and scientific debates that have arisen week after week and that on several occasions ended up providing crucial ideas for tackling the work presented in this thesis.

Y, finalmente, a Ana, Ricardo, Maritza y Mélanie, por dejarme bien situado en la rampa de lanzamiento para llegar hasta aquí. Por haber pensado, muy razonablemente, que la ingeniería sanabresa también podría aportar mucho a este campo.

Para mis dos familias

Preface

As humans, we have been lucky enough to live in the best time in the Universe (within a few million years) to study cosmology and galaxy formation and evolution. If our cosmological model is correct, we believe we can understand what the history of the Universe has been like so far, and predict how the Universe will evolve in future eons.
If we had observed the Universe a few hundred million years ago, we could not have discovered that there is a major component, which we call 'dark energy', that dominates the expansion of the Universe and causes it to accelerate. Its origin and even its nature is not known to us, but its effects are predictable.
If we had witnessed the Universe some (many) hundreds of millions of years in the future, many of the galaxies would have escaped our field of view (technically, they would have crossed our event horizon), so that the view of the Universe then is that it is composed simply of a handful of galaxies (the few ones to which the Milky Way is gravitationally bounded). Without distant galaxies to observe, it is not possible to determine that the Universe is expanding. Furthermore, the cosmic microwave background radiation would be so faint that it would be difficult to detect, so all the clues it is giving us today to the origin of the Universe would be out of reach.

This thesis has the humble aim of helping us to understand a little better some of the processes that determine the formation and evolution of galaxies in the Universe. To do so, I will focus on the galaxy we know best, the Milky Way (MW), and also its neighbour Andromeda (M31), and I will explain some of its characteristics and also the physical processes that have been involved in its formation and evolution. Going from the particular to the general, we consider that such mechanisms also apply to similar galaxies and therefore we can extrapolate this knowledge to better understand the universal physical processes that drive galactic evolution.
The approach to this task will be mainly through the use of cosmological simulations of galaxy formation. In these we apply most of the known physics to simulate a fragment of the Universe: using cosmological initial conditions, and including gas and dark matter, ingredients that are thought to have been present in the early Universe, we evolve them in a self-consistent way up to a time equivalent to the present. Surprisingly, the results obtained reflect, statistically, the observed galactic populations and their properties. The simulated portion of the Universe is large enough to contain about two hundred galaxies resembling MW and M31.
In the following section I briefly describe the main physical processes involved in the formation of structure and galaxies. I mention how physical laws (gravity) that determine the evolution of the Universe as a whole also determine the formation of structures on smaller scales (but large compared to the size of galaxies). Further ingredients will be necessary to explain how galaxies can take shape. I then describe the main types of galaxies that have been observed, with an emphasis on spiral galaxies and describing in great detail what we know about the MW. In each section I add the most relevant observational data that have been necessary to increase our understanding.
After that, I describe the basic fundamentals of cosmological simulations. I mention in what concrete way the physical processes described in the first section can be solved numerically. With an overview of the types of cosmological simulations, I go on to describe the tool I have used in this thesis, the cosmological simulation TNG50, and the scientific results I have extracted from it.

I. INTRODUCTION

1 Theoretical background and context: cosmological model, galaxy formation and morphological classification, and the Milky Way

Although this thesis aims to study MW- and M31-like galaxies, it intends to do so not by considering them as isolated entities, but rather taking into account the cosmological context, by means of cosmological hydrodynamical simulations. Therefore, I will begin by explaining the cosmological model that we think best describes the evolution of the Universe as a whole and how the necessary initial conditions of the elements that will later evolve to form galaxies are established.

1.1 Formation and evolution of galaxies in a cosmological context

Galaxies form via distinct physical phenomena that operate on various spatial and time scales. These include the overarching evolution of the Universe on large scales, the gravitational collapse of matter, the in-fall and thermal processes (cooling and heating) of the gas, the formation of stars, the influence of stellar feedback through mechanisms such as supernovae or stellar winds, and the enrichment of chemical elements.
The present framework for understanding how matter is distributed in the Universe assumes that galaxies grow hierarchically (White & Rees, 1978) within a Universe whose evolution and whose large scale structures are described according to the LCDM model.

1.1.1 The standard model of big bang cosmology

The Standard Model of Cosmology (SMC), or **Lambda Cold Dark Matter** (LCDM) model, is a mathematical parameterization of the hot big bang cosmological model, that has been established as the most successful framework for explaining and predicting the evolution of the Universe at its largest scales. It assumes that General Relativity (GR) is the correct theory of gravity and considers that the Universe is constituted by the following main ingredients: photons, neutrinos, ordinary matter (electrons and baryons), cold 'dark' matter (CDM) and 'dark' energy. CDM is a hypothetical form of non-baryonic matter that interacts only gravitationally, does not undergo electromagnetic interactions (therefore 'dark'), and moves at non-relativistic speeds (therefore 'cold'). Dark energy is an unkown form of energy that behaves like a positive cosmological constant in the field equations of general relativity, Λ as introduced by Einstein 1917. The energy densities of photons and neutrinos are negligible at present day, while the ordinary matter, CDM and dark energy represent, respectively, 5, 27 and 68 per cent of the total energy density of the Universe (Planck Collaboration et al., 2020). LCDM also assumes that the 'cosmological principle' is valid: when averaged on large scales (effectively, $\gtrapprox$ 100 Mpc), the matter distribution in the Universe can be considered isotropic and homogeneous. In GR this means also that the space itself is isotropic and homogeneous. In this picture, it is possible to describe the evolution of the Universe as a time-ordered succession of 3-spaces where the

Cosmological Principle holds (the 'cosmic time' is used to label this succession). The metric of this type of spacetimes can be written in the form

$$ds^2 = -(cdt)^2 + a^2(t)\left(\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right) \quad (1)$$

and is commonly know as Friedmann-Lemaître-Robertson-Walker (FLRW) metric. (r, θ, ϕ) are spherical comoving coordinates and t is the proper time; $a(t)$ represents the scale factor, proportional to the size of the Universe and normalized at present time to $a(t_0) = 1$. The curvature parameter, k, takes different values according to the spatial curvature: closed (< 0), flat (0) or open (> 0). The LCDM model assumes that the Universe is flat.
From the Einstein field equations of relativity, with a FLRW metric and the assumption that the content of the Universe is a perfect fluid, it is possible to obtain the Friedmann equations (Friedmann, 1922) that govern the evolution of the scale factor:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3} \quad (2)$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \quad (3)$$

where Λ is the cosmological constant.
The great achievement of this model is that it is able to explain and predict, using only six free parameters (the specific set of six parameters is open to choice), the following properties of the Universe: the accelerated expansion of the cosmos (first observed by Riess et al., 1998), the existence of a cosmic microwave background (CMB; first predicted by Gamow, 1946) and the statistical properties of its anisotropies, the primordial abundances of the light elements, H, He, Li (first predicted by Alpher et al., 1948), and the spectrum and statistical properties of the distribution of the large scale structure.
For the Universe to have the properties observed today, if we consider the physics of the LCDM model, we come to the conclusion that some of its initial conditions must have been special, that is, only certain very specific values of them allow the Universe to evolve to the actual present state (they are so-called 'fine-tuning' problems). We refer to the 'flatness problem' and the 'horizon problem'. The flatness problem deals with the fact that the geometry of space today is extremely flat, and for this to be so, it had to be in primordial times even orders of magnitude flatter (it can be described analogously in that only an energy density extremely close to a specific value (critical density) is compatible with this geometry). The horizon problem refers to how distant and (apparently) causally disconnected regions of the ancient Universe (for example, as observed in the CMB) have properties that are quantitatively very similar (in the CMB, temperature fluctuations of the order of $\Delta T/T \approx 10^{-5}$), even though they did not have time for thermal equilibrium to be reached between them. Therefore the only possibility, within LCDM, is that the initial conditions at widely separated locations also had to be very similar. Both of these problems can be solved if the Universe expanded exponentially at a very early epoch over a very short period of time (Starobinsky, 1979, Guth, 1981). This phenomenon is called 'cosmic inflation' and is now widely accepted as an addition to the LCDM model.
Presently, the estimation of the free parameters of the SMC is accomplished through analysis of the CMB anisotropies (emitted ≈380,000 years after the Big Bang, see Fig. 1) in addition to the

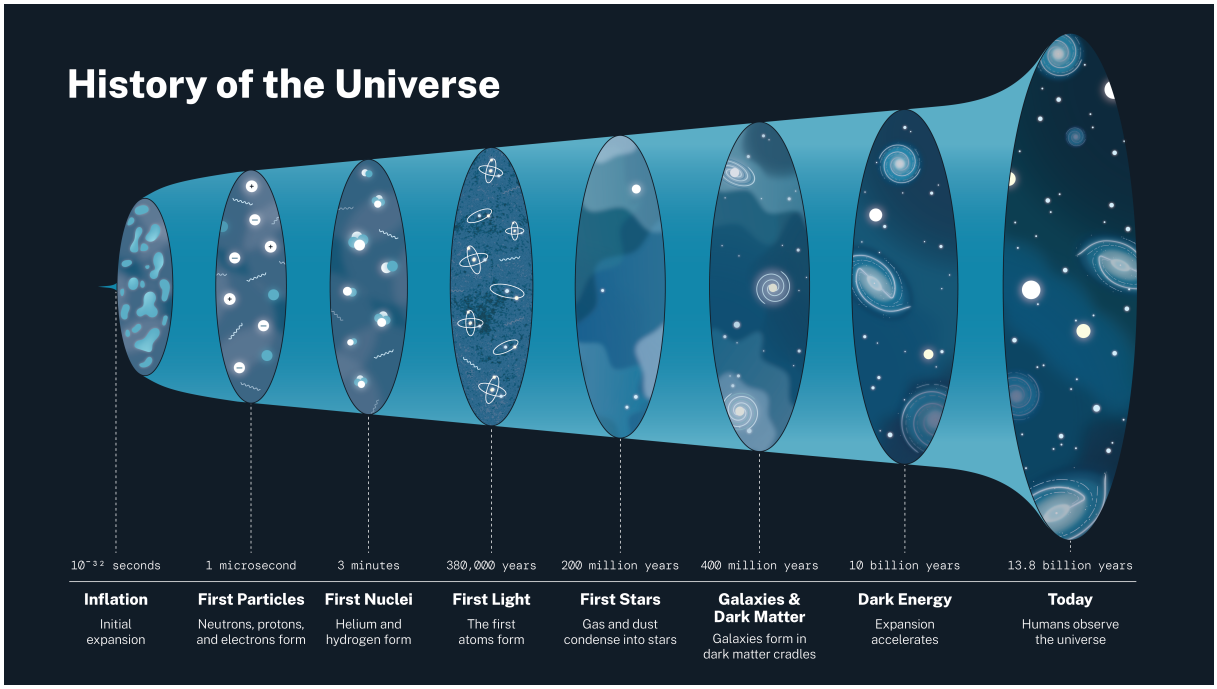

Figure 1: **History of the Universe.** Chronological representation of some of the main events in the history of the Universe. The first stars formed between 100 and 200 million years after the Big Bang. The first galaxies formed some 100-200 million years after the first stars formed. (Credit: NASA)

large-scale structure at later times. The most accurate measurements in this regard have been derived from data gathered by the Planck satellite, that scanned the microwave and submillimetre radiation, giving as a result all-sky maps in nine frequency bands from 30 to 857 GHz (being the last analyses presented as the series 'Planck 2018 results', Planck Collab. et al., 2020).

1.1.2 Formation and growth of structures

On scales smaller than ≈100 Mpc, the Universe exhibits structures that break the isotropy and homogeneity: galaxies and groups, clusters and superclusters of galaxies. It is presently assumed that the structure in the Universe has its origin in minuscule density fluctuations in the primordial stages of its development. Detailed studies on CMB temperature fluctuations (of relative order of $\Delta T/T \sim 10^{-5}$ and directly related to the density fluctuations), have shown that these are almost Gaussian and nearly scale-invariant. The predictions derived from inflation are in good agreement with these observations. The evolution of these density fluctuations and the subsequent formation and growth of structure is mainly governed by gravity. Under the influence of gravity, over-dense regions attract more matter, while under-dense regions become even emptier.

For the case of small fluctuations (the value of the perturbations is small compared to that of the mean density value), it is possible to find approximate analytic solutions using a linearized version of the perturbed equations for the homogeneous case (see, e.g., Mukhanov, 2005, Baumann, 2022, for a detailed description of the cosmological perturbation theory).

The CMB structure can be explained with linear perturbation theory. Also, perturbations on very large scales will remain in the linear regime for longer times. It is possible to extend the range of application using perturbation theory to higher orders, but the equations quickly become very complicated. However, for just a few simplified cases with high symmetry it is possible to find analytic solutions in the non-linear regime.
For the formation of structures at smaller scales, that are completely in the non-linear regime, the most extended and complete approach is the numerical solution of the relevant equations. Because CDM constitutes more than 80 per cent of the total matter and interacts only gravitationally, the formation of the first structures is dominated by the behaviour of this component. DM does not interact electromagnetically with radiation and therefore does not experience the same pressure forces from the photons as ordinary matter does. If it is 'cold', it starts its collapse under self-gravity earlier. Density perturbations grow in the linear regime until they reach a critical density, what makes them detach from the expansion of the Universe and collapse to form virialized dark matter halos. They continue growing initially through the accretion of more DM and hierarchically, through mergers: this is a specific prediction of the CDM paradigm that differs from other types of DM, as 'hot' or 'warm', and is well supported by observations of cosmological large-scale structure and by cosmological simulations. DM haloes have also been studied numerically and exhibit an universal radial density profile first described by Navarro et al. (1997). DM halos are connected with DM overdensities in the form of filaments. The whole structure is usually called the 'cosmic web'. For the rest of the thesis, we are going to refer only to the properties of 'cold' DM when we mention DM.
The effects of baryons will be relevant only later and in scales of galaxies. Therefore, it is possible to study the distribution of structure in a significant range of scales by solving a DM-only universe that interacts only gravitationally. This situation is described by the collisionless Boltzmann equation together with the Poisson equation, solved on a background expanding spacetime. The most efficient way to solve numerically this problem is to discretize the continuous distribution of DM as a bundle of individual particles and solve it as an N-body problem.
To study in detail the formation of galaxies, the inclusion of baryonic elements and additional physical processes is necessary. A description of these processes is given in Section 1.1.3.

Observations of the large-scale structure

The DM cosmic web structure cannot be observed directly. On smaller scales, galaxy are grouped in clusters but on larger scales they are located in knots and filaments of the web: to discover the pattern of the web it is necessary to precisely locate in (3D-)space a large number of galaxies. This is done with galactic redshift surveys, where the radial distance is estimated from the measured redshift.
In the last 25 years several surveys have been completed: the 2-degree Field Galaxy Redshift Survey (2dFGRS Colless et al., 2003), covered more than 220000 galaxies; Sloan Digital Sky Survey (SDSS, Tegmark et al., 2004) provided large-scale power spectra of more than 205,000 galaxies; more recently, the Baryon Oscillation Spectroscopic Survey (BOSS, Dawson et al., 2013), measured 1.5 million galaxies to $z \approx 0.7$ and 100,000 quasars to $z \approx 2.5$; the ground based Dark Energy Spectroscopic Instrument (DESI, DESI Collaboration et al., 2016), plans to cover 30 million galaxies to $z \approx 1.7$ and quasars up to $z \approx 3.5$ when completed in 2026, or the recently launched satellite Euclid (Euclid Collaboration et al., 2022), will obtain the 3D-positions of $z \approx 35$ million galaxies.

The quantification of the large-scale structure aims to describe the matter distribution in the Universe in a statistical way as no model can predict the exact matter distribution of the Universe. The most common method is the galaxy two-point correlation function $\xi(r)$ that describes the degree of clustering as a function of distance. Equivalently, the power spectrum of matter $P(k)$ describes the level of structure as a function of the scale (k is the wave number, related to the scale as $L = 2\pi/k$). When the surveys cover a large enough depth, it is possible to additionally detect the characteristic scale imprinted by baryon acoustic oscillations in the early Universe, that constitute a standard ruler in cosmology. The mentioned statistical values can then be compared with the ones obtained from the LCDM model or with with the ones predicted by from cosmological simulations.

1.1.3 Physical processes in galaxy formation and evolution

The baryonic component in the Universe was, at the early stages, in the form of a hot, dense gas. As mentioned above, the formation of the first structure is dominated by DM (cold in the SMC): the DM haloes act as potential wells for the accretion of primordial gas. It is estimated that the first galaxies formed only 300-400 million years after the Big Bang (see Fig. 1).
The formation of galaxies in DM haloes involves several additional physical processes: the hydrodynamics and thermodynamics of the gas, the formation of stars, the evolution of stars, the chemical cycle, with the interstellar medium (ISM) enrichment from the stars via winds or supernovae, the formation of supermassive black holes (SMBH, with masses $\gtrapprox 10^5$ $M_\odot$) and their effects in the galaxies via feedback. All of these processes are also non-linearly coupled. Many of the details of the involved physical phenomena lack comprehensive understanding, prompting to rely on simulations for valuable insights into the intricate details of these processes. In the following, we succinctly describe the most relevant processes:

Gas hydrodynamics and thermodynamics: the first stars form from the inflow and collapse of gas in the primordial DM potential wells (Eggen et al., 1962). Gas needs to cool to be able to concentrate in dense regions where stars are born. Radiative cooling is the main channel that allows matter to lose energy (Silk, 1977, Rees & Ostriker, 1977). The presence of metals will affect the cooling mechanisms of the gas, as it will be explained below.

Formation and evolution of stars: When the gas is cool and dense enough, it will become self-gravitating. Gas then forms giant molecular clouds (GMC) where stars form. Stars will affect their surrounding across all their lifetimes, in the form of highly energetic radiation, winds and violent mass emissions if they end their lives as supernovae.

Chemical cycle of the stars and the interstellar medium (ISM): elements are synthesized within stars and subsequently transported to the ISM through stellar winds during their lifetimes (Lamers & Cassinelli, 1999). For massive stars, elements are expelled more violently in supernovae explosions. The subsequent presence of heavier elements in the gas alters its thermodynamics and also determines the type of stars that can form from it: as the gas becomes richer in metals, additional cooling channels are allowed, especially metal-line and molecular cooling.

Feedback from supermassive black holes (SMBHs): it is assumed that SMBHs populate the centers of all galaxies (with the exception of some dwarf galaxies) and are responsible for the non-stellar radiation coming from the central regions of certain galaxies, know as active galaxy nuclei (AGN), as the matter is accelerated around the SMBH. The feedback from the SMBH

can affect the gas flows around the galaxy and regulate the star formation in the whole galaxy (Silk & Rees, 1998), as well as its satellites (Martín-Navarro et al., 2021).

To see a more extended description of the topics of this section, see the recent reviews by Somerville & Davé (2015), Naab & Ostriker (2017) and also Vogelsberger et al. (2020).

1.2 The diversity of the galaxy population

According to the standard model of cosmology, CDM haloes and hence galaxies, grow hierarchically, i.e., small galaxies form first and bigger galaxies emerge from the merging of galaxies (White & Rees, 1978). Galaxies have been found across a stellar mass range of several orders of magnitude, going from dwarfs of with stellar masses as low as $\sim 10^3\,M_\odot$ (Belokurov et al., 2009) to the large ellipticals of $\sim 10^{12}\,M_\odot$ of stars (Dullo et al., 2017) in the centre of galaxy clusters. They have diameters ranging from a few thousand kpc to hundreds of thousands kpc and are typically separated by distances of millions of light years.

The environment plays a crucial role in shaping galaxies: galaxy populations in the field and in groups and clusters exhibit fundamental differences in their morphology, color, or star formation rate (Bamford et al., 2009, Tempel et al., 2011). In the fields, galaxies undergo less interactions and therefore maintain their structural integrity and retain more easily their gas to keep forming stars. In galaxy clusters, interactions of between galaxies and especially with the denser environment are more frequent, leading to disruptions and mergers. In these environments the gas is easier removed, by galaxy-galaxy interactions or, more frequently, by ram pressure stripping caused as the galaxies enter or orbit the dense intra-cluster medium (ICM). A clear environmental effect is expressed by the morphology-density relationship (Dressler, 1980, Goto et al., 2003), which shows the effects of local density on galaxy morphology. From the less dense parts of the field to the densest ones in the centers of the clusters, the fraction of spiral galaxies falls steadily, while there is a rise in the fraction of elliptical and S0 galaxies.

If we take into account the distribution of stellar light as observed from our vantage point on Earth, galaxies exhibit a diverse array of shapes. A first attempt at classification was undertaken by Hubble, with his 'tuning fork' classification (Hubble, 1926, Jeans, 1961). Although somewhat simplistic by today's standards, it encapsulates two fundamental ideas: two main groups, mainly spherical and mainly oblate galaxies, with a large number of intermediate forms that encompass what is, in reality, a continuum of morphology. Even with these limitations, galaxies are broadly divided nowadays into three main categories: spiral galaxies, elliptical galaxies, and irregular galaxies.

Spiral or disk galaxies exhibit relatively thin, rotationally supported stellar and gas disks. It is assumed that spiral galaxies form from the accretion onto DM halos of gas with high angular momentum (Fall & Efstathiou, 1980). The disk often exhibits inner structure, like spiral arms or central bars. Spiral galaxies often contain a central spheroidal component, the bulge. They also usually have a high mass fraction of cold gas, that allows star formation. This gives them, in general, a bluer color than elliptical galaxies, although bluer and redder spirals can be found, with a manifest color dependence in the environment (Bamford et al., 2009). They are also the most common type of galaxy at $z \approx 0$ and some of them have been found even as early as at $z \approx 4.5$ (Tsukui & Iguchi, 2021).

Elliptical galaxies have a smooth light distribution, without features and structure. The surface brightness profile can be described by a Sérsic profile (Sérsic, 1963) They are often the conse-

quence of the merger of spiral galaxies. Their stars exhibit larger values of velocity dispersion than in spiral galaxies. Their stellar population is older and therefore their aspect redder, with a color that is luminosity dependent: more luminous elliptical galaxies exhibit a redder color (Sandage & Visvanathan, 1978). Some of them can contain warm or hot gas, but they lack cold gas, therefore star formation is unusual or low. They are found in groups and clusters, where the proximity of galaxies allows more interactions and mergers, and at the highest mass end.

Irregular galaxies are all galaxies that do not fit in the previous groups. They are on average smaller than spiral and elliptical galaxies, and in many cases it is thought that they are galaxies from the two other morphological groups that were deformed by external perturbations.

1.2.1 MW/M31 mass scale as the transitioning scale

The MW and M31 both are thought to reside in DM halos of mass $\sim 10^{12}\mathrm{M}_{\odot}$. Galaxies residing in haloes of this mass have a stellar mass of $10^{10-11}\mathrm{M}_{\odot}$. By studying the MW and M31 we study not only the two main galaxies of the Local Group, or a very common type of galaxy in the Universe, but also a range of stellar and halo masses in which various phenomena and properties of galaxies undergo a clear transition.

- The ratio between stellar and halo mass (M_*/M_{halo}) reaches its maximum (≈ 0.2) at all redshifts (Behroozi et al., 2013), i.e., halos of mass $\sim 10^{12}\mathrm{M}_{\odot}$ have the most efficient integrated star formation history. There must be mechanisms that decrease the SF efficiency at lower and higher masses: for example, feedback from SN and stellar winds, in haloes of lower mass, or AGN feedback in haloes of higher mass (Silk & Mamon, 2012).

- Without taking into account the satellite galaxies, two additional effects are recognizable. Firstly, the fraction of spheroidal galaxies undergoes a sharp transition, from low to high, at this mass range. Secondly, below a stellar mass of $\approx 10^{10}\mathrm{M}_{\odot}$ most of the galaxies are forming stars ($\gtrapprox$ 90 per cent). Above $\approx 10^{11}\mathrm{M}_{\odot}$, most of them are quenched. Within this range, the ratio of quenched and SF galaxies is 50-50. A relation between these two effects was shown by Liu et al. (2019).

Therefore, a better understanding of WM/M31-like galaxies makes it possible to comprehend the physical phenomena involved in the significant changes that occur at this scale and that are determinant in the process of galaxy evolution in general.

1.2.2 Spiral/disk galaxies and their components

Spiral galaxies are characterized by a flat, rotating disk-like structure. Their main structural components are the following:

The disk is the most significant component, consisting of stars, gas, and dust that are organized in a thin, circular distribution. They often present internal structure like spiral arms or bars. Most of the star formation occurs in the disk of the galaxy, giving rise to the spiral arms and other features. The disk radial stellar surface density profile can be described with an exponential, while the vertical density profile is described with a single or double functional profile (commonly exponentials or hyperbolic secants, Yoachim & Dalcanton 2006, Jurić et al. 2008).

Spiral arms are prominent features within the disks of some galaxies. They are over-densities of gas and stars, and often exhibit ongoing star formation. They are dynamic elements that do not represent fixed structures: stars enter and leave the arms as they orbit across the galaxy. The formation of these structures is believed to be influenced by a combination of gravitational interactions, density waves, and the distribution of matter within the galaxy.

A central bar is found in more than 50 per cent of all spiral galaxies (Mulchaey & Regan, 1997, Masters et al., 2011). Their presence can affect the presence of spiral arms. Like those, the bars are dynamic elements.

The bulge is a roughly ellipsoidal region at the center of most spiral galaxies. In its classical form, it is thought to be constituted by stars that formed in the earlier stages of the galaxy, therefore its stellar age distribution is biased toward older ages than the disk.

The stellar halo is an ellipsoidal component surrounding the galactic disk. It can extend to hundreds of kpc, with a density following roughly a decaying power law (Zibetti et al., 2004). The concentration of stars here is much sparser than in the disk or the bulge. It is thought to be formed mainly from the stars of accreted and disrupted satellite galaxies (Bullock et al., 2001). Most of these stars will follow heated, isotropic orbits, but in some cases they keep a coherent spatial structure as stellar streams (Helmi et al., 1999).
All these components are embedded in a much more massive and extended DM halo of spherical or ellipsoidal shape.

1.3 The Milky Way

The Milky Way can be considered a typical spiral galaxy that also contains a bulge and a bar. With our privileged position in it, it is possible to observe and study in great detail the properties of billions of individual stars. The MW has a stellar mass of $\approx 6 \times 10^{10}\,M_\odot$, and an inferred halo mass of $\approx 0.7 - 2.7 \times 10^{12}\,M_\odot$. It contains in its centre a $\approx 4 \times 10^{6}\,M_\odot$ SMBH, relatively small compared to the ones present in similar galaxies. For a comprehensive and up-to-date overview of MW properties see, e.g. Bland-Hawthorn & Gerhard (2016). In Fig. 2 we find an artistic depiction of the MW in a face-on view.
The MW resides in a low density environment. It is part of the Local Group (LG), that contains around 80 galaxies, most of them being dwarfs. The number of detected dwarfs increased significantly in the recent years, thanks to deeper surveys that had allowed the detection of tens of new ultra-faint objects; Laevens et al. (see, e.g., 2015), Koposov et al. (see, e.g., 2015). The group extends to ≈ 2 Mpc around the MW. The MW is the second most massive in the LG, after the Andromeda galaxy, and they are located at 750 kpc from each other. About 50 satellites are gravitationally bound to the MW (some of which have been recently discovered in the works by Homma et al., 2019, Mau et al., 2020), being the Large Magellanic Cloud (LMC), the Small Magellanic Cloud (SMC) and the Sagittarius Dwarf Galaxy among the closest and most massive ones.

1.3.1 Observations of the MW

The Sun is located at ≈ 8 kpc from the galactic centre (e.g., VERA Collaboration et al., 2020) and ≈ 25 pc above the stellar midplane (see, e.g., Chen et al., 2001, Karim & Mamajek, 2017). From our position within the Galaxy we can study with great precision the stars surrounding the

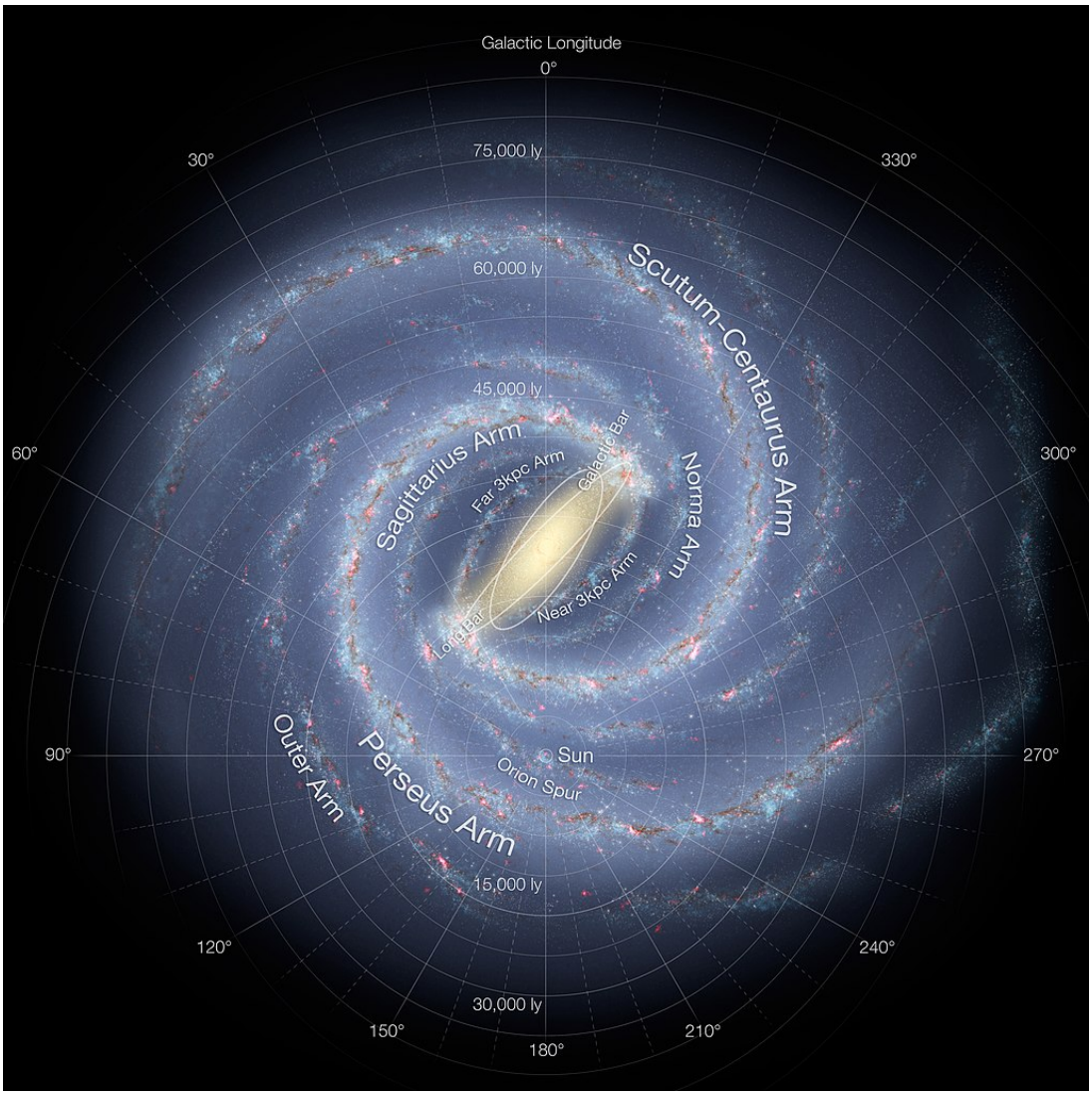

Figure 2: **Milky Way's artistic impression.** Structure of the Galaxy with the location of the spiral arms, bulge and bar. Deduced from data from several surveys. Overlaid is the Galactic longitude, centered at the Sun (Credit: European Southern Observatory)

sun, down to distances of several kpc. The study of a large fraction of individual stars, along the entire main sequence, is something that is only possible for the MW (and, exceptionally, some of its closest satellites).
The stars, apart from giving us information about what the Galaxy is like today, have recorded in their composition, kinematics and age, what the Galaxy was like in the past. The study of the history of the MW through the properties of its stars in the present is often called 'Galactic Archaeology'.
It is possible to obtain various types of information about the stars. Depending on how the observations are made, surveys can be classified into several categories: photometric surveys allow the collection of color information in a small number of band filters. This makes possible the study of a large number of stars in a short time. Astrometric surveys provide an estimate of the distance to the star as well as two components of its velocity. Spectroscopic surveys provide detailed spectra of the stars. This allows estimation of their elemental chemical composition.

Astrometric and photometric surveys, and Gaia

The most recent, successful and complete astrometric data has been obtained with **Gaia** (Gaia Collaboration et al., 2016). The objective of the Gaia mission , a satellite placed at the Lagrange point L2 in the Sun–Earth system, is the creation of a 3D map of 1 per cent ($\sim$ 1 billion) of the stars of the MW. It is complete to 20th magnitude in the Gaia-specific G filter, so it also includes a certain number of non-stellar intra- and extragalactic objects. The astrometry includes positions, proper motions and parallaxes of all MW observed stars. The multi-band photometry is used to estimate luminosity, temperature, mass, age and chemical composition of the stars. The combination of the detailed 6D phase-space map with the photometric information allows to reconstruct the past of the Galaxy and make predictions of its evolution. Additionally, Gaia posses also a spectrometer that allows to obtain high precision radial velocities for stars down to magnitude 17th. Fig. 3 shows a sky map of the density of measured sources in the second data release (GaiaDR2).

Spectroscopic surveys

Many ground-based spectroscopic surveys have been started in the last years to provide high resolution spectra of hundreds of thousands of stars of the MW, or low resolution spectra of even a few million stars, allowing to estimate stellar parameters and elemental abundances and thus complementing Gaia astrometry. We list below some of them, as well as some of their most relevant results:

LAMOST (Large sky Area Multi-Object fibre Spectroscopic Telescope, Deng et al. 2012, Zhao et al. 2012): is a low resolution survey, $R \approx 1800$, in the optical range, providing spectra of more than 9 million stars. With these results, Li & Zhao (2017) for example studied the kinematics of thick disk stars, concluding that the in situ formation scenario is the most likely one.

APOGEE (Apache Point Observatory Galactic Evolution Experiment, Majewski et al. 2017): is a high resolution survey, $R \approx 22500$, in infrared range, that has provided spectra of more than 150,000 stars from all structural components. It has enabled to discover a significant number of very metal-poor stars in the Galactic bulge (García Pérez et al., 2013).

RAVE (RAdial Velocity Experiment, Steinmetz et al. 2006): provides medium resolution spectra ($R \approx 7000$) of $\sim 500,000$ stars in the southern hemisphere. Two of its relevant results are an

estimation of the local escape speed in 544 km/s (Smith et al., 2007) and the discovery that supersolar metallicity stars in the solar neighbourhood must have formed in the inner disk and have migrated outwards (Kordopatis et al., 2015).

GALAH (GALactic Archaeology with HERMES, Zucker et al. 2012): gives high resolution spectra for more than 300,000 stars, mostly from the disk, allowing to calculate the abundances of 30 elements. With data from this survey it has been possible to demonstrate that open clusters and field stars trace the distribution of the chemical elements of the disc differently (Spina et al., 2021).

4MOST (4-metre Multi-Object Spectroscopic Telescope, de Jong et al. 2012): aims to obtain medium and high resolution spectra of a few tens of millions of stars in the southern hemisphere.

1.3.2 The Milky Way's stellar disk

About 70 per cent of the MW stars are located in the disk ($\approx 5 \times 10^{10}\,M_\odot$). Most of the stars are concentrated within a diameter of $\approx$ 30 kpc. The disk radial surface density profile can be characterized by an exponential of scale length of $\approx$ 2-3 kpc and its vertical density profile, close to the Sun position, can be characterized by a double functional, commonly a double exponential or a double hyperbolic secant, with respective scale heights of ~300 and ~900 kpc respectively (see, e.g. Jurić et al., 2008, for a recent estimation). They have been commonly named 'thin' and 'thick' disk and it has been debated whether their stars have different properties: metallicities, kinematics, etc. Based on the kinematics, it is common to decompose the disk into 'cold' and 'warm' disks, according to the vertical velocity dispersion of the stars. A most recent approach decomposes the disk in two components whose stars differ on the abundance in alpha elements. The 'low-α' disk is kinematically cold and has shorter scale length than the 'high-α' disk, that is besides kinematically hotter. Furthermore, stars belonging to the high-α disk tend to be older ($\gtrsim$ 8 Gyr) than stars from the low-α disk.

The definitions based on the geometric decomposition of the disk, on the kinematic of the stars and on the alpha abundances give as a result roughly similar subpopulations (see, e.g., Haywood et al., 2013, for an analysis of the correspondence between the geometric and the chemical disks), but they are not completely equivalent. It has to be highlighted that only the last one, based on alpha abundances, can assign individual stars to each separate components.

The structure of the MW disk is not axisymmetric: there is a central bar of longitude 5 kpc (Wegg et al., 2015) and a number of spiral arms where most of the current star formation happens. There are four major arms: 'Perseus', 'Norma', 'Crux-Scutum' and 'Carina-Sagittarius'. The Sun is located in the 'Orion-Cygnus' minor arm, close to the Perseus arm. Fig. 2 shows the sun position and the location of the main arms.

1.4 The Andromeda Galaxy

The Andromeda Galaxy, also known as M31, is the biggest galaxy in the Local Group, slightly larger than the MW. It is located approximately 765 kpc away from the MW. With a diameter of around 46 kpc, M31 is more extended than the MW. It contains an estimated stellar mass of $10^{11}\,M_\odot$. Morphologically, M31 exhibits prominent spiral arms, which emanate from a central bar structure. It is also surrounded by a large number satellite galaxies ($\approx$ 37), including the Messier objects M32 and M110.

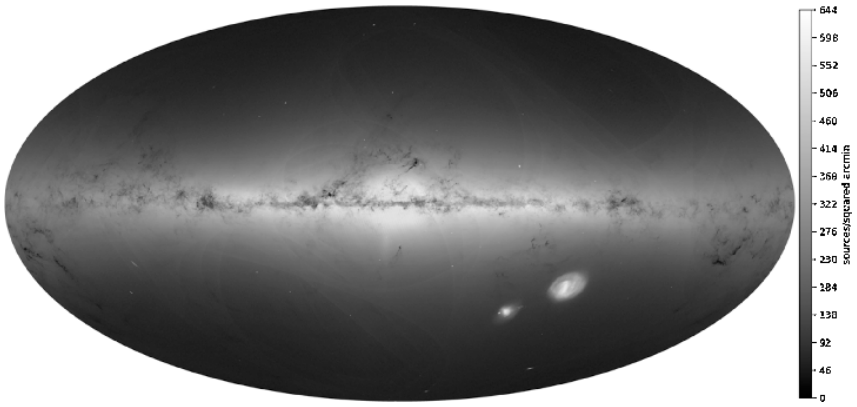

Figure 3: **Gaia DR2 sources.** Sky distribution of all sources in Galactic coordinates in Hammer projection. The logarithmic colour scale represents the number density of sources (From Gaia Collaboration et al. 2018)

In view of the above data (expanded in Table 1), we can be certain that MW and M31 are galaxies with comparable size and mass and very similar morphology. However, there are also significant differences:

- It is estimated that their assembly and merger histories were considerably different, with MW undergoing its last major merger 9-10 Gyr ago and M31 only 2 Gyr ago.
- The stellar disks are much thinner in the MW: scale-heights of ~200 and ~800 pc (for thin and thick disk) for the MW, vs. ~1000 and ~2800 for M31.
- The central SMBHs have very disparate masses, differing in orders of magnitude: $\approx 4 \times 10^6\,M_\odot$ in the MW and $\sim 10^{10}\,M_\odot$ in M31.

Table 1: **Observational properties of MW and M31.** Compilation of estimated values for the MW and M31 based on observations. The displayed value ranges encompass the best values and errors from a series of measurements and procedures. For the Sun to MW-centre distance (McMillan 2011, Rix & Bovy 2013, Russeil et al. 2017), total halo mass (MW: Peñarrubia et al. 2014, Cautun et al. 2019 ; M31: Tamm et al. 2012, Sick et al. 2014), stellar mass (MW: Licquia & Newman 2015, Bland-Hawthorn & Gerhard 2016 ; M31: Tamm et al. 2012, Sick et al. 2014), baryon mass, gas mass (MW: Dame 1993 ; M31: Nieten et al. 2006, Tamm et al. 2012), star formation rate (MW: Bland-Hawthorn & Gerhard 2016, Boardman et al. 2020 ; M31: Boardman et al. 2020), disk length, thin disk scale-length (MW: Ojha 2001, McMillan 2011, Bland-Hawthorn & Gerhard 2016), thick disk scale-length (MW: Ojha 2001, McMillan 2011, Bland-Hawthorn & Gerhard 2016), thin disk scale-height (MW: Siegel et al. 2002, Jurić et al. 2008, Bland-Hawthorn & Gerhard 2016; M31: Collins et al. 2011), thick disk scale-height (MW: Siegel et al. 2002, Bland-Hawthorn & Gerhard 2016; M31: Collins et al. 2011), bar half-length (MW: Wegg et al. 2015; M31: Beaton et al. 2007, Blaña Díaz et al. 2017), SMBH mass (MW: Ghez et al. 2008, Chatzopoulos et al. 2015; M31: Bender et al. 2005)

		Milky Way	Andromeda
Class	Property [unit]	Value	Value
Coordinates	R_0[kpc]	7.9 – 8.3	
Mass budget	Total halo mass[$M_\odot$]	4×10^{10} – 7.3×10^{10}	7.8×10^{10} – 1.5×10^{11}
	Stellar mass[$M_\odot$]	4×10^{10} – 7.3×10^{10}	7.8×10^{10} – 1.5×10^{11}
	Baryonic mass[$M_\odot$]	5.4×10^{10} – 7.3×10^{10}	
	Gas mass[$M_\odot$]	6.5×10^{9} – 1.25×10^{10}	5×10^{9} – 7×10^{9}
	Star formation rate[$M_\odot$]	0.68 – 1.84	0.41 – 0.83
Stellar morphology	Disk scale length [kpc]	1.7 – 2.9	4.8 – 6.8
	Thin disk scale length [kpc]	2.1 – 3.1	
	Tick disk scale length [kpc]	1.8 – 2.2	
	Thin disk scale height [pc]	175 – 360	900 – 1300
	Thick disk scale height [pc]	625 – 1450	2200 – 3400
	Bar half-length [kpc]	4.3 – 5.2	2.6 – 5.1
SMBH	Mass [$M_\odot$]	4.1 – 4.4×10^{6}	1.1 – 2.3×10^{8}

2 Cosmological simulations as a tool to understand the Universe

We described in Section 1.1 the basic laws and processes that determine the formation of structures in the Universe, from the largest scales, of several hundreds of Mpc, to the intragalactic scales of only some pc. Given the complexity of these equations, a numerical approach is needed to solve them. This is precisely what cosmological simulations of galaxies do, determining the intertwined evolution of dark and baryonic matter (and the latter includes gas, stars and SMBHs), by solving the coupled equations of hydrodynamics and self-gravity, starting from minuscule density perturbations in the almost homogeneous primordial Universe (e.g. around $z = 100$).
The formation and evolution of structure includes physical processes that happen in a wide range of spatial scales and masses: for example, SMBHs, of size $\sim 10^6\,M_\odot$ have effects that extend to several kpc in a galaxy and its surroundings, that is, a spatial dynamic range of $\sim 10^{10}$. In terms of mass, a process like the formation of individual stars in a big galaxy can reach a mass dynamic range of $\sim 10^{12-13}$. Given the computational limitations present today, it is impossible to simulate a big portion of a universe with all the small scale physical processes included. Additionally, there are still physical processes relevant for the formation and evolution of galaxies that are not perfectly known or understood. To overcome these limitations, a subgrid treatment includes some of the physical processes in an effective manner.
Most of the concepts explained in this section are covered in greater detail in a recent review by Vogelsberger et al. (2020).

2.1 How to simulate a universe in a box

Cosmological simulations of galaxies solve the gravitational and hydrodynamical interactions in a comoving volume linked to a homogeneously expanding background that determines the real physical size of the simulated Universe. In practice, this is characterized by the evolution of the scale factor, that obeys the Friedmann equations (see Section 1.1.1). In order to recreate galaxy formation and evolution, it is necessary the modelling of additional astrophysical processes, included in a effective or sub-resolution manner. The basic components of the cosmological galaxy simulations will be DM, gas, stars and SMBHs.

2.1.1 Gravitational forces

The gravitational interaction of DM and stellar particles is described by the collisionless Boltzmann equation coupled to Poisson's equation. Gravity is considered Newtonian, because it is simpler and relativistic corrections would be minimal in these studied regimes. This system of equations can be solved using different techniques: particle-based, mesh-based or hybrid:

The most common particle-based technique is the **tree method** (Barnes & Hut, 1986): to accelerate the calculations of the particle-particle interactions, the contributions to the potential of particles that are very distant can be approximated by a low order multipole expansion.

Poisson's differential equation can be solved in Fourier space, giving the **particle-mesh method** (Hockney & Eastwood, 1988). The potential is calculated on a grid and its values at the particles positions subsequently interpolated. This method can deal directly with periodic boundary conditions.

A hybrid approach, as the **tree-particle-mesh method** (Bagla, 2002), takes the tree approach for shorter range and PM for longer range interactions. This method is implemented in most of current gravity codes for cosmological simulations.

2.1.2 Magneto-hydrodynamics

The hydrodynamic equations can be discretized following different approaches. Most commonly, either a Lagrangian or an Eulerian method are followed.

Among the **Lagrangian** methods, the mesh-free smoothed particle hydrodynamics (SPH) techniques (Lucy, 1977, Gingold & Monaghan, 1977, Monaghan, 1992) are the most common. In the SPH technique, fluid elements are surrounded by a region over which their properties are smoothed by a kernel function. The particles move according to the Newtonian equations, being the forces acting on them due to pressure, gravity or electromagnetic. This scheme has the advantage that it automatically adjusts its resolution according to the particle crowding (i.e., fluid density) and naturally follows the mass flow. Also, it ensures the conservation of energy, linear and angular momentum, mass and entropy. The main limitations of SPH methods come to the handling of shocks, the treatment of contact discontinuities and the artificial suppression of certain fluid instabilities. For a more detailed description of the method and its application in astrophysical problems, see, for example, the review by Springel (2010a). For its application in cosmological simulations, the most widely used implementations are based on the GADGET (Springel et al., 2001a) or GADGET-2 (Springel, 2005) codes or modifications of them.

The most extended **Eulerian** techniques utilize adaptive mesh refinement (AMR, described by Berger & Colella 1989). They are superior in the treatment of shocks and discontinuities, compared to SPH methods. But AMR lacks Galilean invariance and it presents problems when high density regions move at high velocities or around intricate geometries. Two of the most used AMR codes are RAMSES (Teyssier, 2002) and ENZO (Bryan et al., 2014).

Alternatively, there are new methods that try to avoid the limitations of SPH and AMR. A clear example is the unstructured dynamic mesh technique present in the AREPO code (Springel, 2010c). The mesh is a Voronoi tessellation of the space from a finite number of reference points. The points follow the matter movement and therefore the mesh shape will automatically adapt, with the additional advantage of being automatically more refined in densest regions. It is also Galilean invariant and has a precise treatment of shocks and fluid instabilities.

2.1.3 Galaxy formation processes

With the combination of the gravitational and hydrodynamical methods described above it is possible to model a self-gravitating fluid. There will be additional physical processes that need to be included to allow the study of galaxy formation. Many of them have to be implemented in an effective or subgrid manner, due to the resolution limitations (especially in large cosmological boxes).

The involved physical processes are succinctly described in Section 1.1.3. Below, a brief description of the different implementations in the codes is given.

Gas cooling and heating:

Radiative cooling is included. For temperatures $10^5 \lesssim T \lesssim 10^7$ K also metal line cooling is included. The values are tabulated as a function of density, temperature, metallicity, and redshift. If the simulation allows that the ISM cools further, molecular cooling is included below $\sim 10^4$ K.

Modeling ISM:
The modeling of multi-phase ISM is essential but challenging. It is estimated that the cold phase occupies a volume of 1 per cent while containing 90 per cent of the total ISM mass. Due to the very short-time scales and very small space-scales involved in the dense gas, the cold phase is usually described with an effective polytropic equation of state.

The formation of stars and their evolution:
Stars are expected to form in the cold-phase ISM. In typical implementations in cosmological simulations of galaxies, stars can form when the density exceeds a critical value ($n_{\rm H} \approx 0.1\,{\rm cm}^{-3}$). The star formation rate can be calculated from the empirical Kennicutt-Schmidt relation (Schmidt, 1959, Kennicutt, 1989). In this process, star particles will form from the available gas. These particles have masses of several thousands of $M_\odot$ and in the simulations are considered as whole stellar populations with the same formation time and same chemical properties. The population is originally determined according to a certain initial mass function (IMF), for example from Chabrier (2003) or Kroupa (2001).

Feedback from stars and SMBHs:
Stellar feedback, i.e. the injection energy and momentum in the ISM from the stars can occur via supernova or winds. Energy is transferred to the surroundings thermally or kinematically. For the former case, an excessive cooling has to be compensated artificially, via deactivation of the radiative cooling or extra heating of the gas, for a time scale comparable of ≈ 30 Myr.
SMBHs are not resolved in cosmological simulation. They are seeded in the centre of a halo that exceeds a certain mass limit. The seeded particle can grow accreting gas from the surroundings or through mergers. The accretion rate is usually calculated with analytic formulas that take into account the properties of the gas and the BH mass. The assumption for SMBH feedback is that part of the energy from the accreted mass is reinjected in the form of thermal or kinetic energy in the surroundings, in a continuous or a discrete manner. Also, cosmological simulations often distinguish two types of feedback based on the accretion rate, low or high: they are commonly named 'radio' and 'quasar' modes (e.g. Sijacki et al., 2007), respectively, injecting the former thermal energy in a bursty and anisotropic manner, and the latter, energy and momentum in a continuous way.
The role of both types of feedback, stellar and SMBH, is crucial to regulate the star formation in galaxies and avoid the formation of objects with too many stars, being this one of the problems present in earlier simulations, as we will describe in Section 2.2.4.

2.2 Types of simulations of galaxy formation

There are several types of cosmological simulations of galaxy formation:

Depending on the physics involved, we will have DM-only simulations, where only gravitational interactions occur; and simulations with full physics, where the equations of the magnetohydrodynamics of baryonic matter and subgrid models for stellar formation must be solved.

Depending on the type of initial conditions for the simulation, we will have cosmological simulations, when we start from the small density fluctuations present at high redshift and make the universe evolve until the desired structures are obtained; or idealized simulations, when the simulation starts from conditions or configurations that have not been reached through the evolution of primordial fluctuations (for example, arbitrarily placing two galaxies already formed and close to each other to simulate a merger).

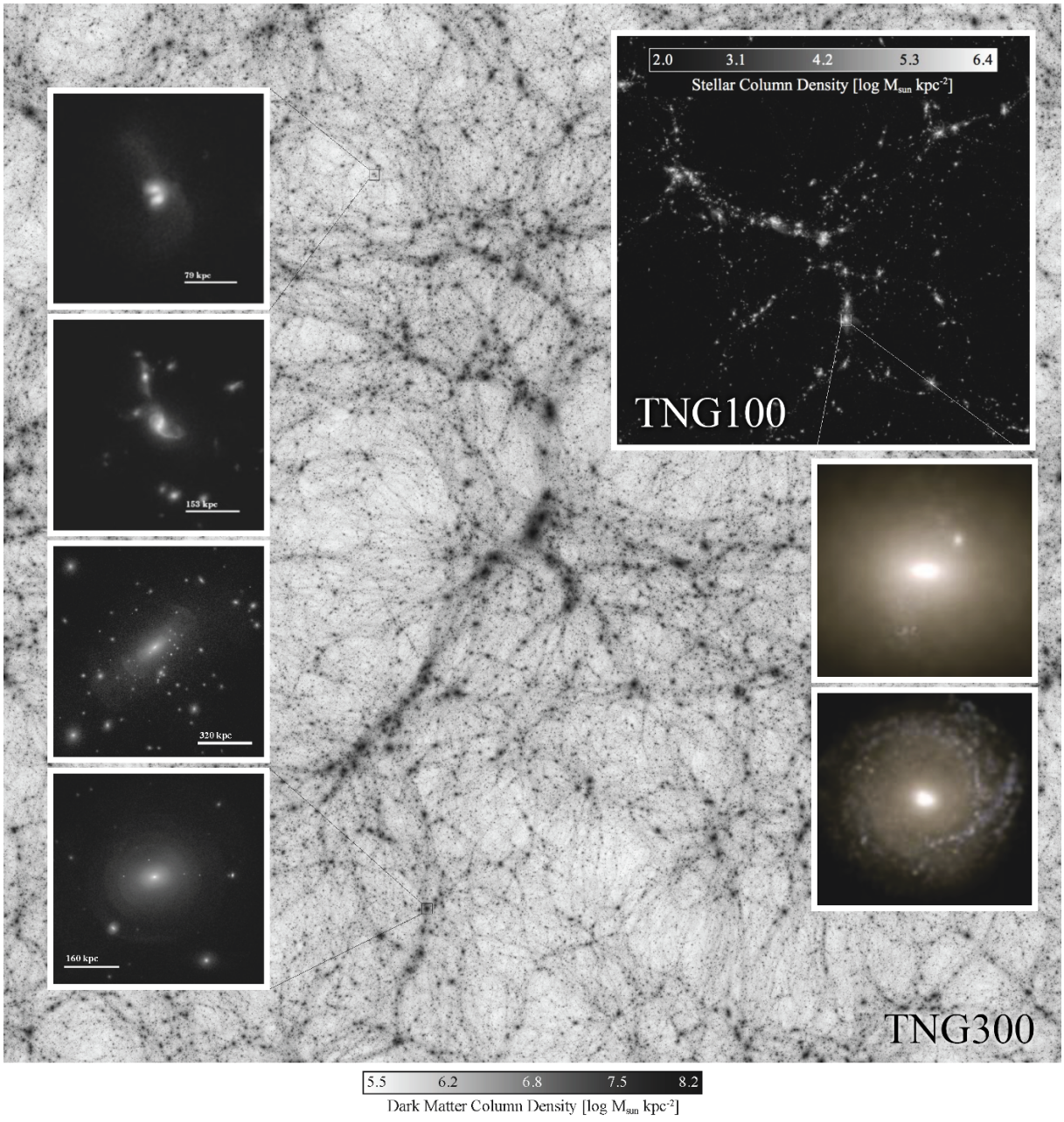

Figure 4: **Cosmological simulation IllustrisTNG.** Composite image with data from the TNG300 and TNG100 simulations showing the high range of structure scales reproduced in the simulations: the background represents the dark matter column density of TNG300. The upper right inset, stellar column density in the TNG100 volume. The left panels show galaxy groups and clusters. The bottom right panels show the stellar distribution of two morphological distinct types of galaxies. (From Pillepich et al. 2018b)

Depending on how the resolution elements are treated, we will have box cosmological simulations, where the resolution of the particles is approximately uniform throughout the simulated volume; and zoom-in simulations, where a higher resolution is applied to certain objects of interest within the simulated volume.

2.2.1 N-body and full-physics

DM dominates the formation of structures at the largest scales, where the effects of the baryonic matter can be completely neglected. As DM interacts only gravitationally, it can be described, in the continuum limit, as a collisionless fluid whose dynamics follow the collisionless Boltzman equation coupled to the Poisson equation. Also a system formed of many stars, like a star cluster or a galaxy, can be described by these equations, if interactions with the surrounding interstellar medium are neglected. It is common to solve the system using the N-body technique, where the continuous fluid is represented as a set of particles of finite mass. The numerical methods employed are described in Section 2.1.1.
The current understanding of the N-body techniques allows for a consensus about the accuracy and importance of the results obtained. To obtain more exact results requires only a higher number of particles and more accurate time integration and gravitational force calculations. Currently there are numerical techniques that can solve this problem very efficiently.
Some examples of the most recent pure DM-only simulations are Cosmo-π (Cheng et al., 2020), Uchuu (Ishiyama et al., 2021) and Last Journey (Heitmann et al., 2021). All of them have more than 10^{12} particles and box side lengths of 2-3 h^{-1} Mpc.
We can take into account only DM to study the structures formation at the largest spatial scales. To study smaller scales regimes and, of course, galaxy formation, the inclusion of baryons is needed. Baryons will be initially included as a collisional inviscid ideal gas whose dynamic is governed by the continuity and the Euler equations. The details of how to solve the magneto-hydrodynamics problem are explained in Section 2.1.2. The techniques to include the additional physics necessary to form stars and galaxies are covered in Section 2.1.3. The outcomes of one of these full-physics cosmological simulations (IllustrisTNG, Naiman et al., 2018, Marinacci et al., 2018, Pillepich et al., 2018b, Nelson et al., 2018, Springel et al., 2018) is shown in Fig. 4: it is possible to appreciate the DM filaments that form the backbone of the cosmic web to smaller structures as galaxy clusters or individual galaxies with its internal structure resolved.

2.2.2 Cosmological boxes and zoom-in simulations

Cosmological simulations try to simulate portions of the Universe starting from high redshift cosmological initial conditions. Both simulations of galaxy formation and of large-scale structure formation are usually performed in large cubic volumes with sides of several Mpc for the former or Gpc for the latter. The mass elements usually have a constant resolution throughout the volume. Examples of the values achieved are, for the TNG50 galaxy formation simulation, with approximately a cubic volume of side 50 Mpc, the stellar particles have masses of $\sim 8.5\times10^4$ $M_\odot$ and the DM, $\sim 4.5 \times 10^5$ $M_\odot$. As an example of simulation of large structures, Uchuu, has DM particles of $\sim 3.27 \times 10^8$ h^{-1} $M_\odot$ and side volume 2 h^{-1} Gpc. These simulations are carried out with periodic boundary conditions to simulate the effects of homogeneity and isotropy at large scales. Some examples of cosmological boxes are EAGLE (Schaye et al., 2015, Crain

et al., 2015), Illustris (Vogelsberger et al., 2014a,b, Genel et al., 2014, Nelson et al., 2015), IllustrisTNG or Simba (Davé et al., 2019).
As we mentioned above, computational power imposes a limit in the dynamic range achieved by the simulations. To increase the dynamical range of cosmological simulations by a few orders of magnitude, a common approach is to use the zoom-in technique. It involves including a high-resolution region of interest (a galaxy, a galaxy cluster, etc) in a larger scale cosmological volume with a coarser resolution. The selection of the object of interest implies in many cases running a DM-only large volume box from which halos with certain characteristics can be identified. For example, zoom-in simulations of MW-like galaxies select haloes of mass $\sim 10^{12}\,M_{\odot}$ as potential hosts of MW analogs. Some recent examples of cosmological zoom-in simulations of MW-like galaxies are the suites Auriga (Grand et al., 2017), APOSTLE (Navarro et al., 2018), Vintergatan (Agertz et al., 2021), NIHAO-UHD (Buck et al., 2020), or Hestia (Local Group analogs, Libeskind et al., 2020).

2.2.3 Idealized simulations: isolated disks and mergers

Compared to cosmological simulations, idealized simulations focus usually on simulating individual galaxy halos in isolation, disregarding the larger cosmological context. These simulations involve a carefully chosen set of components and parameters. The initial conditions are idealized, not cosmological. Consequently, while they are, in principle, easier to interpret, they may lack the level of physical realism found in cosmological simulations. Furthermore, one clear advantage of idealized simulations is that they tend to be computationally less demanding and can therefore be executed with higher resolution compared to cosmological simulations.
If we put the focus on MW/M31-like galaxies, which are also the object of study in this work, idealized simulations have been used to understand the natural evolution of a disky galaxy, disentangled from environmental effects, or to study in detail different merger configurations of the progenitors of this galaxies.

Isolated disks: MW/M31-like spiral galaxies can exist in complex and crowded environments where they interact with neighboring galaxies or larger cosmological structures. Idealized simulations of isolated galaxies help to investigate how a MW-like galaxy would evolve without external disturbances, although it is possible to include certain external effects, like gas inflows. The initial conditions, as they are not cosmological, will assume certain theoretical profiles for the distribution of mass of the DM and stellar components. For example, it is common to sample the DM particles following a NFW profile, or the stellar particles of the disk component following a Miyamoto-Nagai profile. The amount of gas present is often used as a parameter, to study its effect on the stellar component.
Some specific applications of isolated disk simulations are the study of secular radial migration, disk thickening, formation and stability of bars, effects of the CGM in the disk component, effects of AGN-jets in the galaxy, etc.

Idealized merger simulations: Mergers play a crucial role in the in the evolution of MW/M31-like galaxies. There are many parameters that play a role in a binary merger, such as orbital configuration, mass ratio, gas content, impact parameter. In a large enough cosmological volume, many possible realistic merger configurations will appear naturally. Idealized merger simulations allow in this case a not computationally expensive fine sampling of some of the merger parameters. For example, it is possible to make small variations to the mass ration while

all other merger parameters are kept fixed. Some recent examples of the application of idealized merger simulations are shown in Karademir et al. (2019), Petersson et al. (2023)

2.2.4 Historic development: on the difficulty of forming realistic disk galaxies

Although cosmological simulations have long been a fundamental tool for understanding the processes involved in galaxy formation, obtaining realistic disc galaxies, in terms of size, bulge to total fraction, disk thickness (for example, with disks as thin as those observed in the Universe, e.g. Yoachim & Dalcanton, 2006), rotation curve, etc., has been a problem solved only in recent years. They were shown to be too small, or with very bulky and massive bulges.
One of the causes of these problems is the 'overcooling catastrophe' (Balogh et al., 2001) that takes place when gas is allowed to cool radiatively: as this process is unstable, because it increases the gas density, which again increases the cooling rate, gas cooling can occur catastrophically if the star formation and feedback subgrid recipes are not accurately implemented. In turn, it produces big flows of cool gas to the centre of the galaxy and too many and too intense episodes of star formation, leading to galaxies becoming too massive and blue. Other of the causes is the 'angular momentum catastrophe', where most of the angular momentum of galaxies was transferred to the DM halos, producing too compact disks.
Currently, and in the last 10 years, realistic disk galaxies have been produced in simulations. Different authors have advocated for different key elements for this success: improvement in resolution (Governato et al., 2007), the used star formation models (Scannapieco et al., 2008), the stellar feedback (Stinson et al., 2013), or the star formation efficiency (low, Agertz et al. 2011 or high, Sales et al. 2010).
More details about the development of realistic disk galaxies in cosmological simulations can be found in the works by Mayer et al. (2008), Marinacci et al. (2014).

2.3 Rationale and outline of the thesis

In this thesis I study MW/M31-like galaxies using the TNG50 cosmological simulation, the highest resolution run of the IllustrisTNG suite. For this purpose I have a sample of 198MW/M31 analogs, each with a number of stellar particles between approximately 500 thousand and 1 million.
I intend to answer several fundamental questions about the MW and M31, but by doing so I want to gain knowledge about processes that will therefore also be present in similar galaxies and fundamental in driving galaxy evolution:
Firstly, I want to answer whether disky galaxies like the MW and M31 can undergo a recent merger but still be disk galaxies at $z = 0$. This is a widespread debate in the community because it has been assumed that the only possible way for the MW to exhibit such a thin disk as it does, compared to disk galaxies of similar mass, is that it has not had mergers in the last 8-10 Gyr. Large-volume simulations have so far not exhibited galaxies with as much spatial or mass resolution as necessary to carry out the relevant analyses, therefore, most of the zoom-in simulations that have devoted their efforts to reproduce as realistically as possible the MW selected their galaxies having a quiet recent merger history. Other attempts have used idealized simulations, i.e., which do not start from cosmological initial conditions. The advantage of TNG50 is that the galaxy sample is significantly larger (by a factor of at least 10-20, or even more in some cases) compared to the previous studies and all galaxies in the sample have not

been selected a priori with a biased merger history, as many of the previous studies do. In addition, it is interesting to be able to identify some global or structural property of galaxies with recent mergers, for which a sample of a significant size, such as TNG50, is necessary.
Secondly, I want to examine the vertical structure of MW/M31-like galaxy disks, in particular the flaring phenomenon, i.e., an increase in the scale-height of the vertical distribution of stars with galactocentric distance, which has been observed in MW and extragalactic surveys of edge-on galaxies. Moreover, this phenomenom appears more clearly when analyzed for mono-age populations, which is observationally possible only in the MW. This phenomenon has been analyzed from a theoretical point of view and with idealized or zoom-in simulations, with different physical models and different ways of quantifying the flaring. Although an explanation for the physical causes is very complicated, it is possible to study how flaring manifests itself in the TNG50 galaxies, how ubiquitous the phenomenon is, and how it compares quantitatively and formally with the available observations and with previous studies using MW analogs simulations.
Finally, I want to answer a question that has a strong observational counterpart, namely, what low-metallicity stars can tell us about the past of the MW. Specifically, I quantify the presence of very metal-poor ($[\mathrm{Fe/H}] < -2$) stars in the disk of MW-like galaxies in TNG50. These stars are generally very old and a large part of the community holds that they can therefore only be in the Galactic stellar halo. In TNG50 the task of assigning a stellar particle to a morphological component of the galaxy is very straightforward, since their positions and velocities are known in great detail. In this way, it is be possible to make predictions about the number (or proportion) of very metal-poor stars that can be expected to be found in the Galactic disk, specially with upcoming surveys. In addition, it is possible to determine how old these stars are, or whether they formed in the galaxy itself or in satellites that were later accreted.

2.3.1 Thesis outline

In chapter 3 I present the sample of TNG50 galaxies that we consider to be good analogs of the MW and M31. I detail and justify the considered selection criteria and describe some of the properties of the analogous galaxies. The next sections show the main results of my investigations:

In chapter 4 I answer the question of whether it is possible, according to TNG50, that a disky galaxy like the MW or M31 can undergo a recent merger and still show a disky morphology at $z = 0$. This work has been published as Sotillo-Ramos et al. 2022, *The merger and assembly histories of Milky Way- and M31-like galaxies with TNG50: disk survival through mergers*.

In chapter 5 I detail the phenomenon of stellar disk flaring within the MW/M31-like galaxies of TNG50. With the physical model of TNG50 we will see diverse manifestations, which quantitatively and qualitatively encompass all previous studies, starting from the same model of galaxy formation. However, we will see that predicting flaring from other galaxy properties is a complicated task. This work has been published as Sotillo-Ramos et al. 2023a, *Disc flaring with TNG50: diversity across Milky Way and M31 analogues*.

In chapter 6 I discuss the study of very metal-poor ($[\mathrm{Fe/H}] < -2$) stars in MW-like galaxies. I analyze their distribution in the structural elements of the galaxy and take the opportunity to make predictions about how many such stars are expected to be observed in galactic surveys, especially in the galactic disk. This work has been published as (Sotillo-Ramos et al., 2023b),

On the likelihoods of finding very metal-poor (and old) stars in the Milky Way's disc, bulge, and halo.
Finally, I discuss the results of the thesis in chapter 7 and provide and outlook for further and deeper research on the topics of the thesis to improve the results here obtained. I conclude with a summary and final remarks in chapter 8.

II. RESULTS

3 The simulation TNG50 and its MW/M31-like galaxies

The work of this chapter is based on the submitted article
Milky Way and Andromeda analogs from the TNG50 simulation
Annalisa Pillepich, Diego Sotillo-Ramos, Rahul Ramesh, Dylan Nelson, Christoph Engler, Vicente Rodriguez-Gomez, Martin Fournier, Martina Donnari, Volker Springel and Lars Hernquist
Monthly Notices of the Royal Astronomical Society
where I am the second co-author. My specific contributions are: figures 1, 3, 4 (bottom panel), 10, 16; part of Table 1; and the analysis of figure 16.

3.1 IllustrisTNG

This work has been performed using the results of TNG50 (Nelson et al., 2019b, Pillepich et al., 2019), the highest resolution run from the cosmological simulations suite IllustrisTNG (Naiman et al., 2018, Marinacci et al., 2018, Pillepich et al., 2018b, Nelson et al., 2018, Springel et al., 2018).
In Section 2.1 we have described the methodology of cosmological simulations of galaxy formation. Next, we describe the specific implementations that have been used in this simulation. IllustrisTNG is an evolution (literally, 'The Next Generation') of the cosmological simulation of galaxies Illustris. It contains three main runs with different sizes and resolutions, aimed for the study of different aspects of the formation of galaxies, being TNG50 the smallest and highest resolution run. As a periodic cubic volume with side length of ≈ 51.7 comoving Mpc, it contains 2160^3 dark matter (DM) particles and the same initial number of gas cells. DM particles have a mass of $m_{\rm DM} = 4.5 \times 10^5$ M$_\odot$, and gas cells (and stellar particles), $m_{\rm baryon} = 8.5 \times 10^4$ M$_\odot$. Compared to the other two main runs, of box sizes of approximately 100 and 300 Mpc, TNG50 contains a smaller number of galaxies that but at higher resolution, which will be crucial for the study of stellar disks with very short scale heights, comparable to those of the MW.
The coupled dynamics of DM, gas, stars and SMBHs is followed with the quasi-Lagrangian code Arepo (see Section 2.1.2 for details). The initial conditions are established at $z = 127$ and it assumes a LCDM cosmology consistent with Planck results 2015 (Planck Collaboration et al., 2016). It uses an updated and improved version of the Illustris galaxy formation model (Pillepich et al., 2018a, described in great detail in). We briefly summarize here the main characteristics of the model:

- Star formation occurs stochastically in gas cells with density higher than $n_{\rm H} = 0.13\,{\rm cm}^{-3}$, following a Kennicutt-Schmidt law and a Chabrier IMF.
- Stars (stellar populations) evolve and produce chemical enrichment from supernovae Ia, II and winds from AGB stars.
- The production and distribution of nine chemical elements (H, He, C, N, O, Ne, Mg, Si, and Fe) is followed.
- Galactic winds are injected isotropically (Pillepich et al., 2018b).

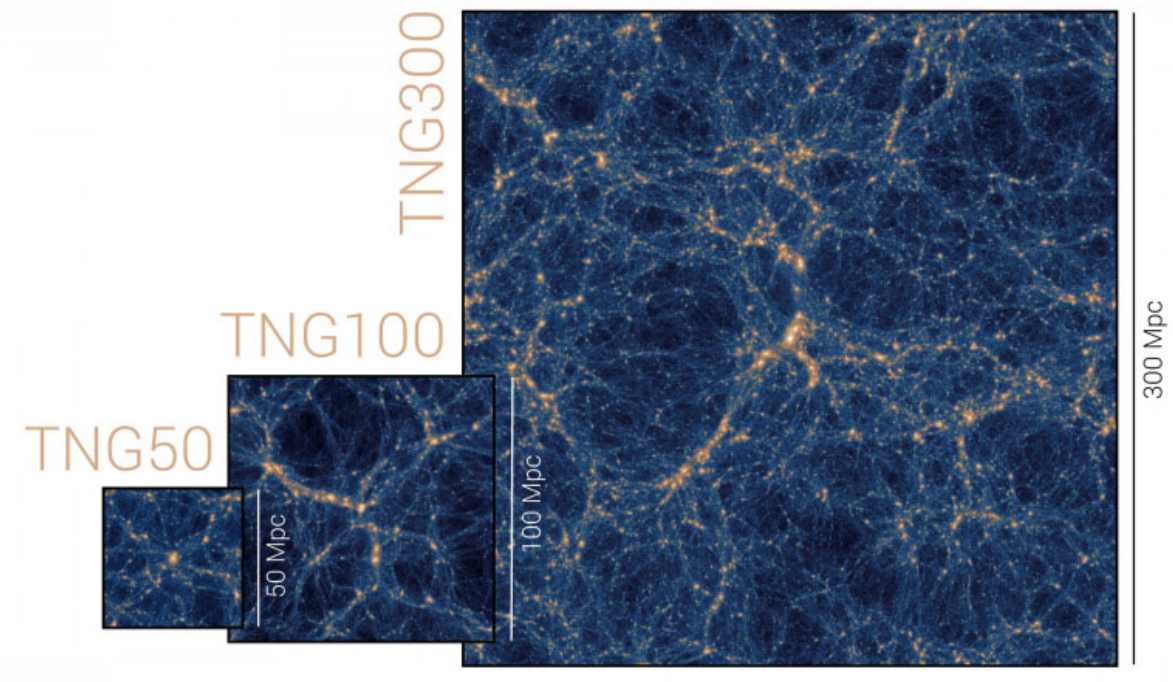

Figure 5: **Cosmological simulation IllustrisTNG.** The three volumes of the IllustrisTNG suite, showing the column density of DM. This work makes use of the higher resolution of TNG50 that enables to resolve and study smaller details (From Nelson et al. 2019b).

- SMBHs are seeded with a mass of 10^5 $M_\odot$ in haloes more massive than 7×10^{10} $M_\odot$.
- The AGN feedback model includes a kinetic mode for low accretion rate and a thermal mode for high accretion rate (Weinberger et al., 2017).
- Cosmic magnetic fields are included. Their evolution starts from a primordial seed field also at $z = 127$.

Scientific results from IllustrisTNG

The IllustrisTNG suite has been used in more than 650 publications for the study of galactic morphology, galaxy formation, galaxy clusters or large-scale structures. Each of these topics can be studied most conveniently in one of the runs, TNG50, TNG100 and TNG300 (we show in Fig. 5 a comparison of the three volumes). Most of the results have been cross-checked with observations, including the full range of scales mentioned above. This has allowed the robustness of the astrophysical models used in IllustrisTNG to be corroborated, giving high reliability to their predictions.
In the following points we highlight some of the results supported by observations:

- Galaxy colors: good quantitative agreement with SDSS, with a sharp transition at stellar mass 10.5 $M_\odot$, that can be explained with the change in the SMBH feedback mode (Nelson et al., 2018)

- Galaxy morphologies: optical morphologies of galaxies are in good agreement with Pan-STARRS observations at $z \approx 0.05$ (Rodriguez-Gomez et al., 2019).
- Galaxy sizes: for redshifts between $z = 0$ and $z = 2$, the stellar galaxy-mass relations, both for SF and quenched galaxies agree with observations within 0.25 dex (Genel et al., 2018).
- Galaxy mergers: for merging galaxies, there is good agreement between the optical morphologies of TNG50 and galaxies from the Kilo-Degree Survey (KiDS) (Guzmán-Ortega et al., 2023).
- Bars in spiral galaxies: the distributions of pattern speeds are similar to those from MaNGA observations at $z = 0$, but the bars in TNG50 are slightly shorter (Frankel et al., 2022).
- Quenched galaxies: for centrals and satellites, similar fractions as in SDSS survey, for redshift $z \leq 3$ (Donnari et al., 2021). If the SF is resolved, massive galaxies exhibit inside-out quenching, in agreement with observations from 3D-HST (Nelson et al., 2021)
- Galaxy clustering: the galaxy 2-point correlation function is in good agreement with SDSS, at z 0.1, both for all galaxies or splitted in colour. (Pillepich et al., 2018b)
- Emissions from galaxy clusters: radio emission from the most massive clusters agree with the observed values for Coma and Perseus. (Marinacci et al., 2018)
- Satellites of MW/M31 analogs: TNG50 satellite abundances are consistent with those of MW, M31 and from similar galaxies from observational surveys like SAGA (Engler et al., 2021b) and exhibit similar quenched fractions, with a minor excess when compared to external galaxies in the same mass range (Engler et al., 2023).

3.2 Comparison to previous simulations of MW/M31 analogs

TNG50 offers a large number of MW/M31-like galaxies with a resolution of about one million stellar particles per galaxy, at z=0. Compared to previous simulations, TNG50 offers clear advantages for an analysis in a cosmological context (see Fig. 6, where the number of analog galaxies in each simulation is plotted against the mass resolution). Compared to zoom-in simulations (shown with diamonds in the Figure), the number of available galaxies increases by a factor of 10-100. Furthermore, in most cases, the galaxies in the zoom-in simulations are selected with biases in their formation history or environment (for example, with a quiet recent merger history or residing in sparse environments). These simulations can employ galaxy formation models that incorporate physics significantly differently from TNG50. For example, as can be seen in the Figure, most zoom simulations do not include AGN feedback, considering it irrelevant for galaxies in the studied mass range. However, the treatment of the interstellar medium or stellar feedback can be modeled in greater detail (e.g., the FIRE model, described by Hopkins et al., 2014). On the other hand, when we compare TNG50 with other cosmological box simulations, with uniform resolution (represented with circles in the Figure), the philosophy followed is the same, i.e., galaxies form and evolve in a cosmological context, but in TNG50 we reach for the first time resolutions that allow us to resolve in sufficient detail the stellar disks

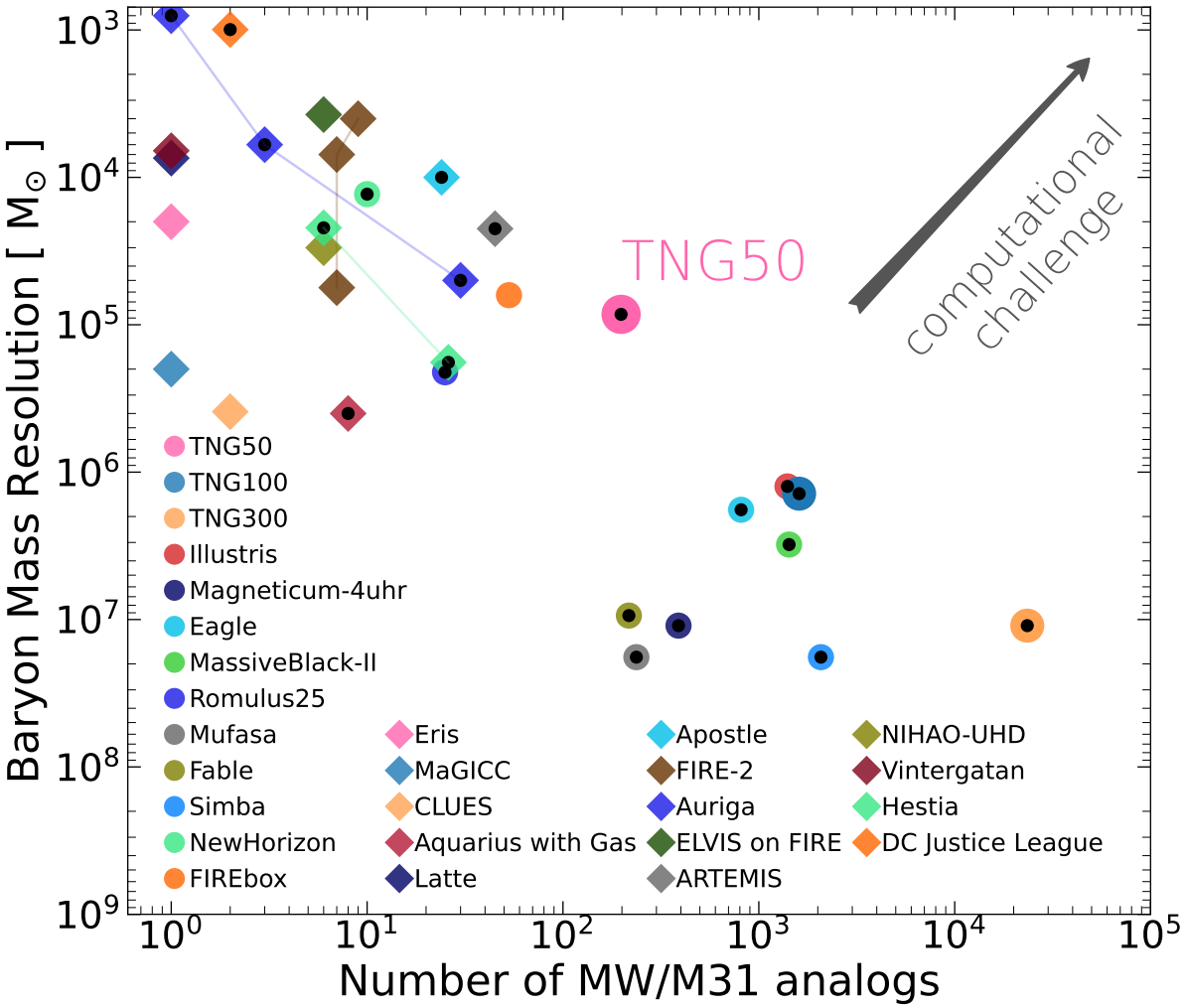

Figure 6: **Cosmological simulations.** TNG50 is compared with another large-volume simulations (circles) and with zoom-in simulations (diamonds). Black dots symbolize the inclusion of SMBH feedback.

of MW-like galaxies, obtaining values compatible with those of the observed galaxies (among them, the MW). The drawback is that to reach this resolution the simulated volume is smaller, so the sample is smaller.

3.3 Analogs of the MW and M31 in TNG50

3.3.1 What is an analog of MW and M31?

TNG50 produces thousands of galaxies with stellar masses in the range $\sim 10^{6-12}\,M_\odot$. It is crucial to define how are galaxies from TNG50 that, at z = 0, can be considered analogues of our Galaxy and Andromeda.
In previous studies, properties such as total or stellar mass in a given range, disky morphology of the stellar component, quiet merger history (since $z \sim 1.5 - 2$; see, e.g. Roca-Fàbrega et al. 2016, Agertz et al. 2021) have been requisites for a galaxy to be considered a good MW analog.

The definition of what constitutes a good analogue of the Milky Way should be in line with the scientific purpose of the study. In this way, in addition to the properties already mentioned, the color, the star formation rate, the total number of satellites, the presence of a LMC and/or SMC companion, etc., could also be included.
The methodology used is also determinant, and the search for analogs will not be the same when it is done within cosmological simulations or astronomical surveys. As an example, only simulations (specially DM-only) allow a selection criterion based on halo mass. This is not possible for observational studies. Similarly occurs when we want to impose a restriction on the assembly history (e.g. a quiet merger history): it will be only possible within the context of cosmological simulations.
Of course, one would think that adding more and more parameters to the selection criteria would yield better MW analogs: one would like not only the correct range of mass and a disky shape, but also, the current SFR, number of satellites, the presence of a bar, etc. It could be expected that applying this conditions to large surveys as SDSS (York et al., 2000) or GAMA (Driver et al., 2011), with millions of galaxies observed, one could obtain a vast amount of analogs. Reality is that the volume of these multi-parameter space increases exponentially, therefore the data becomes more and more sparse extremely quickly [1]. We can find clear examples of this effect in the literature (for example, Fraser-McKelvie et al. (2019) apply a selection criteria based on four parameters to a sample of more than 100000 galaxies in SDSS DR7 (Abazajian et al., 2009), obtaining only ∼100 MW-like galaxies).

3.3.2 The selection in TNG50: observable-based criteria

For this work, we select from the TNG50 box the best analogues of the Milky Way and Andromeda based on their $z = 0$ properties, i.e. without imposing conditions on their evolution or any restrictions on their morphological changes with time.
Because TNG50 starts from general cosmological conditions and is not designed to produce a specific set of galaxies, with these restrictions in the selection we are able to see how representative the MW and M31 are within galaxies with similar properties and at the same time, how diverse the outcomes can be for a relatively narrowly defined sample of galaxies.
The selection is based on observable global properties of the galaxies and pretends to be quite general, meaning that we treat equally the MW and M31, with only a difference in the mass range. The selection has been presented and motivated in Pillepich et al. (2023) and already used by Engler et al. (2021a), Pillepich et al. (2021), Sotillo-Ramos et al. (2022), Sotillo-Ramos et al. (2023a), Chen et al. (2023) and others. To be selected as 'MW- or M31-like', a TNG50 galaxy must meet, at $z = 0$, the following three conditions, based on galaxy stellar mass, disky stellar morphology, and Mpc-scale environment:

A) Galaxy stellar mass: $\log_{10}(M_*/M_\odot)$ within [10.5, 11.2], with stellar mass measured within a 3D circular aperture of 30 kpc.

B) Disky stellar morphology: either of the following conditions:

[1] This phenomenon is commonly referred to as 'the curse of dimensionality' and was first described by Bellman (1961)

(a) Stellar minor-to-major axis ratio (as defined in Pillepich et al., 2019), $c/a < 0.45$, whereby c and a are respectively the minor and major axis of the ellipsoidal distribution of the stellar mass between one and two times the stellar half-mass radius.

(b) Disky by visual inspection of 3-band images, both in face-on and edge-on projections.

C) Environment: no galaxy with stellar mass $\geq 10^{10.5}$ $M_\odot$ within 500 kpc distance and total host mass $M_{200c,\mathrm{Host}} < 10^{13}$ $M_\odot$.

There are 198 galaxies in the TNG50 simulation that match such conditions. In the top panel of Fig. 7 we show how many galaxies are selected in TNG50 according to each one of the three criterion. The stellar mass constraint encompasses the available estimates of the stellar mass of both the MW and M31 (e.g. Bland-Hawthorn & Gerhard, 2016, Boardman et al., 2020). We can then separate galaxies in the sample as MW- and M31-mass, when they stellar mass is in the range $10^{10.5-10.9}$ $M_\odot$ (130 galaxies) or $10^{10.5-10.9}$ $M_\odot$ (68 galaxies), respectively. The purpose of the diskyness criteria is to select all simulated galaxies exhibiting a coherent global stellar disk morphology with or without the presence of spiral arms (see the bottom righ panel of Fig. 7 for a detail comparison of the galaxies included trough the two diskyness criteria considered: c/a ratio and visual inspection). A geometric criterion is popular in extragalactic surveys, where disky galaxies are tipically defined as those with minor-to-major axis ratio < 0.33 (and middle-to-major axis ratio > 0.66; see e.g. fig. 8 of Pillepich et al. (2019) and references therein) and does not require star kinematics, necessary in the case of orbit circularity criteria (e.g. Abadi et al., 2003). Ax expected, the connection of this metric to other measures of "diskyness", e.g. kinematic or photometric D/T ratios, scale length to scale height ratio, etc., is quite complex and depends strongly on where the measurements are performed. We have decided to extend the upper limit on the minor-to-major axis ratio to include the ratio of the stellar thick disk height to the stellar disk length of the Galaxy, according to the mean values from Bland-Hawthorn & Gerhard (2016), with thin and thick disk lengths = 2.6 ± 0.5 and 2 ± 0.2 kpc, respectively, and thin and thick disk heights at the Sun location = 300 ± 50 and 900 ± 180 pc. This gives the aforementioned limit of ~ 0.45. The criterion that ensures isolation guarantees that no galaxy, having a mass equal to or greater than the lower estimate for the Milky Way, exists at a distance closer than 500 kpc. The distance between the Galaxy and Andromeda is ~ 770 kpc (McConnachie et al., 2005, Riess et al., 2012). At the same time it is required that the limit on total host mass excludes galaxies located within a very massive galaxy cluster, which is a much more densely populated environment that the one of the Galaxy and Andromeda. However, it is not necessary that the selected analog is central in its halo, allowing in this way environments and configurations like the Local Group, where it is considered that the Galaxy and Andromeda may be sharing the same DM halo.

Very importantly, in our selection there are no constraints in the galaxies merger histories. This is in clear contrast to previous simulations of Milky Way analogues (e.g. Governato et al. (2004), Okamoto et al. (2005), Eris (Guedes et al., 2011), Agertz & Kravtsov (2015), GARROTXA (Roca-Fàbrega et al., 2016), VINTERGATAN (Renaud et al., 2021), AGORA (Roca-Fàbrega et al., 2021)), that have imposed the last major merger to redshift $z \geq 1$ or at a lookbak time of 8-10 Gyr.

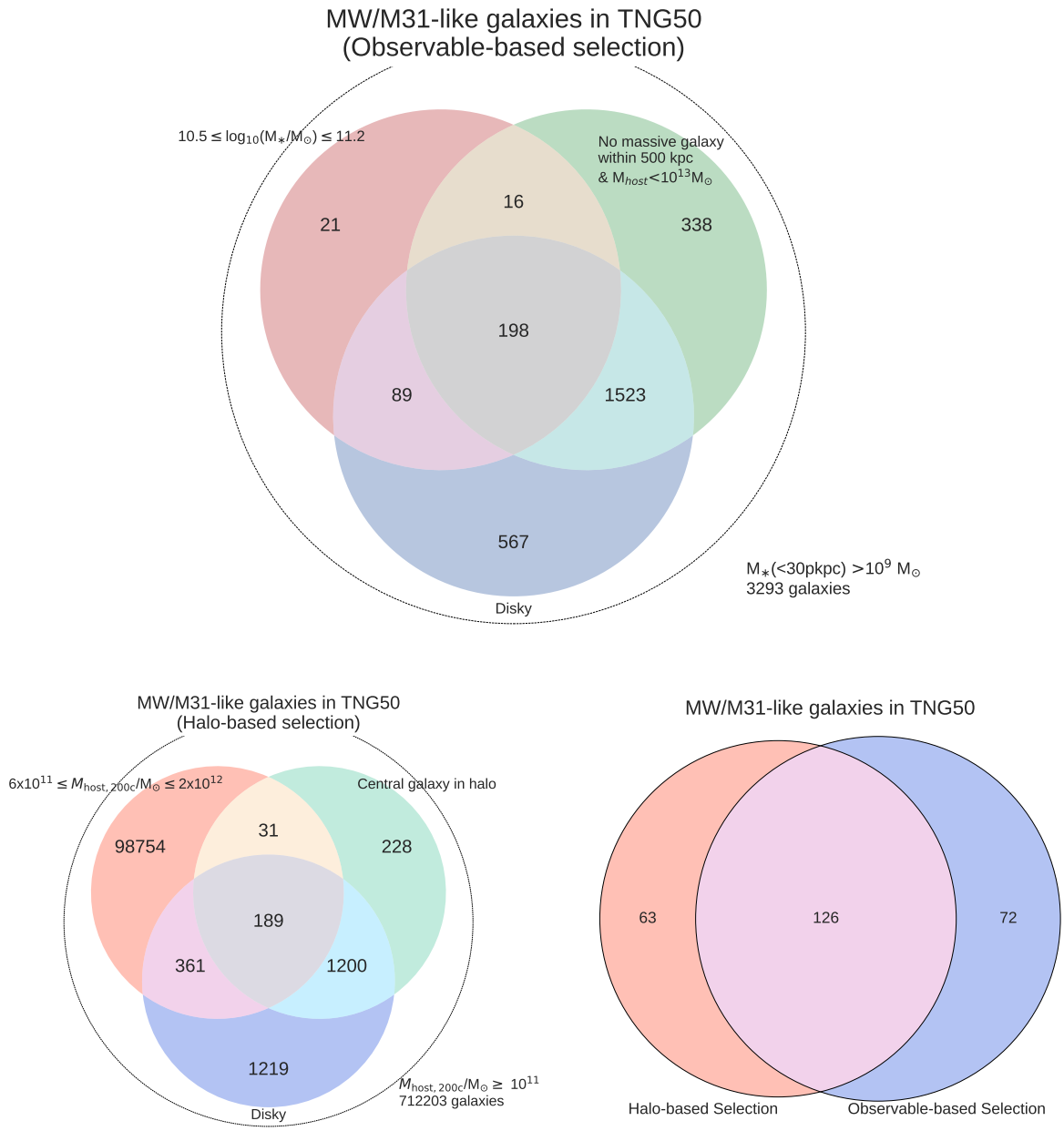

Figure 7: **Selection criteria and number of MW/M31-like galaxies in TNG50** *Top*: Observable-based selection. *Bottom-left*: Halo-based selection, more usual in zoom-in simulations, where the selection is based on the DM halo masses from large DM-only simulations. *Bottom right*: Comparison between the two selection criteria

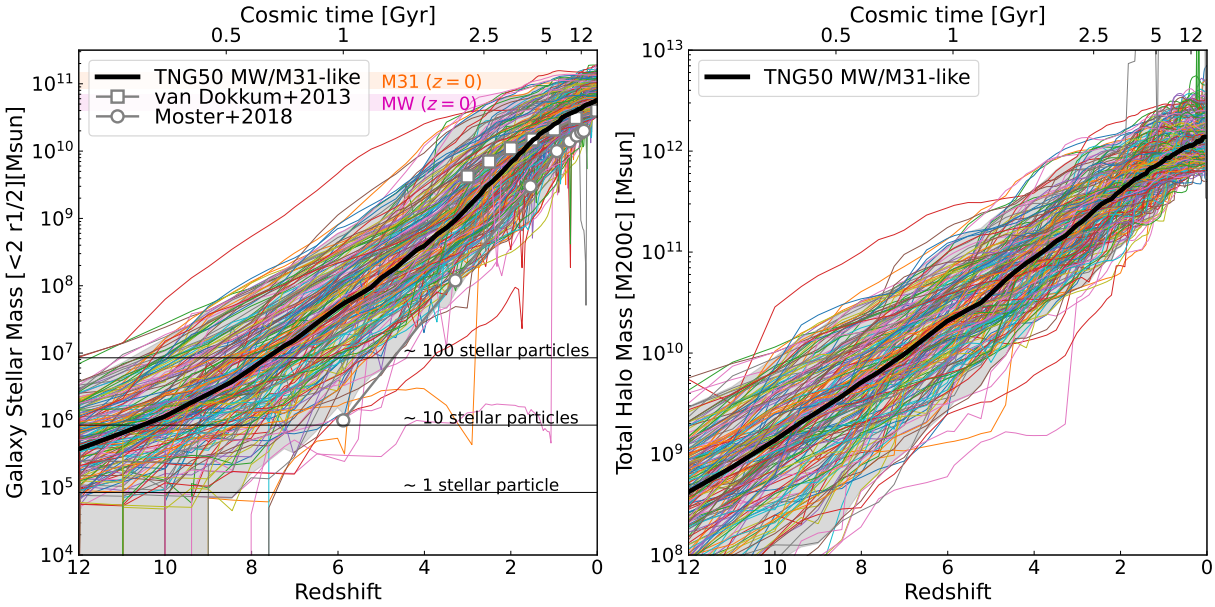

Figure 8: **Mass evolution of MW/M31-like galaxies in TNG50** Stellar mass (left) and halo mass (right) evolution of the 198 MW/M31 analogues. Each thin line represents one galaxy. The median is represented by the thick black line. The grey shaded area shows the 10-90 percentile range. Additional curves represent constraints (grey lines with empty markers).

3.3.3 Basic properties of the MW/M31 analogs

We give a summary of the properties of the MW/M31-like galaxies in TNG50, from their environments to the structure.

Local and large-scale environment: most of the MW/M31-like analogues are located in higher density regions, in the cosmic web, i.e., filaments and sheets. 21 of them are within 10 Mpc of distance to Virgo-mass clusters. 8 of them are not the central galaxy in their haloes.

Local Group analogs: three pairs could be considered LG analogs. Both galaxies in the pair are 500-1000 kpc apart and approach towards each other.

Haloes: haloes of the MW/M31 analogues cover a wide range of masses and sizes. If we consider the region with an average density 200 times the critical of the Universe, the range of mass covers $10^{11.5-12.8}$ $M_\odot$ and the range of radii, 147 – 404 kpc.

Merger histories: TNG50 MW/M31-like exhibit very diverse merger histories. This will be explored in detail in Chapter 4. Most of the galaxies (168/198) had at least one major merger since $z = 5$, and 31 of 198 galaxies had a major merger in the last 5 Gyr, but are still disky at $z = 0$. The progenitors are very varied and, although the $z = 0$ stellar mass range covers only 0.7 dex, at $z = 2$ the stellar mass extends over almost three orders of magnitude, or more than one order of magnitude in its $z = 2$ halo mass (see Fig. 8).

Satellite populations: MW/M31-like analogues in TNG50 host 0-20 satellites with stellar mass $> 5 \times 10^6$ $M_\odot$ for galactocentric distances < 300 kpc, covering the numbers observed in around the MW and M31. A detailed analysis of satellite populations is shown in Engler et al. (2021a).

Star formation activity: MW/M31-like analogues in TNG50 exhibit z=0 SF activities compatible with the estimations for MW (star forming) and M31 (green-valley or quenched galaxy). In

the MW/M31 stellar mass range, observations determine that roughly 50 per cent of the galaxies are star forming. In TNG50, 115 of 198 are either star forming or green valley. SF activity occurs in a smooth or bursty manner. In Chapter 4 we will show that external events, as merger or close encounters with satellites can trigger star formation events.

Stellar population properties: with this we refer to stellar ages and metallicities. The globally-averaged values agree to a good extent with measurements from extragalactic surveys. In Chapter 6 we show how the stellar population properties across morphological components, with an emphasis on the very metal-poor and old stars in the disk.

Disks: TNG50 MW/M31 analogues are strongly disk dominated, with mass fractions of ≈ 0.5 on average. Circular velocity curves are diverse and bracket the measurements for the MW. The extension of the stellar disks exhibit scale lenghts from 2 to 17 kpc, if we describe the radial surface density with an exponential. The vertical distribution can be described with a double functional profile, and the measured scale heights bracket the measurements for the MW or M31. This is strongly dependent on the radial distance of the measurement, and we offer the results at a fixed annulus of 8 kpc or at a variable of $2.7 - 4.7 \times R_{\rm d}$ (see Fig. 9.

SMBHs: All analogues host SMBHs. The scatter in mass, at fixed galaxy stellar mass, covers one order of magnitude. In the low mass end, no analogues reach a mass as low as the MW SMBH mass. In the high mass end, M31-SMBH-mass are well represented.

Central bars: Between 128 (Zhao et al., 2020) and 150 bars (Rosas-Guevara et al., 2022) are found among the 198 MW/M31-like galaxies, showing good agreement with extragalactic surveys.

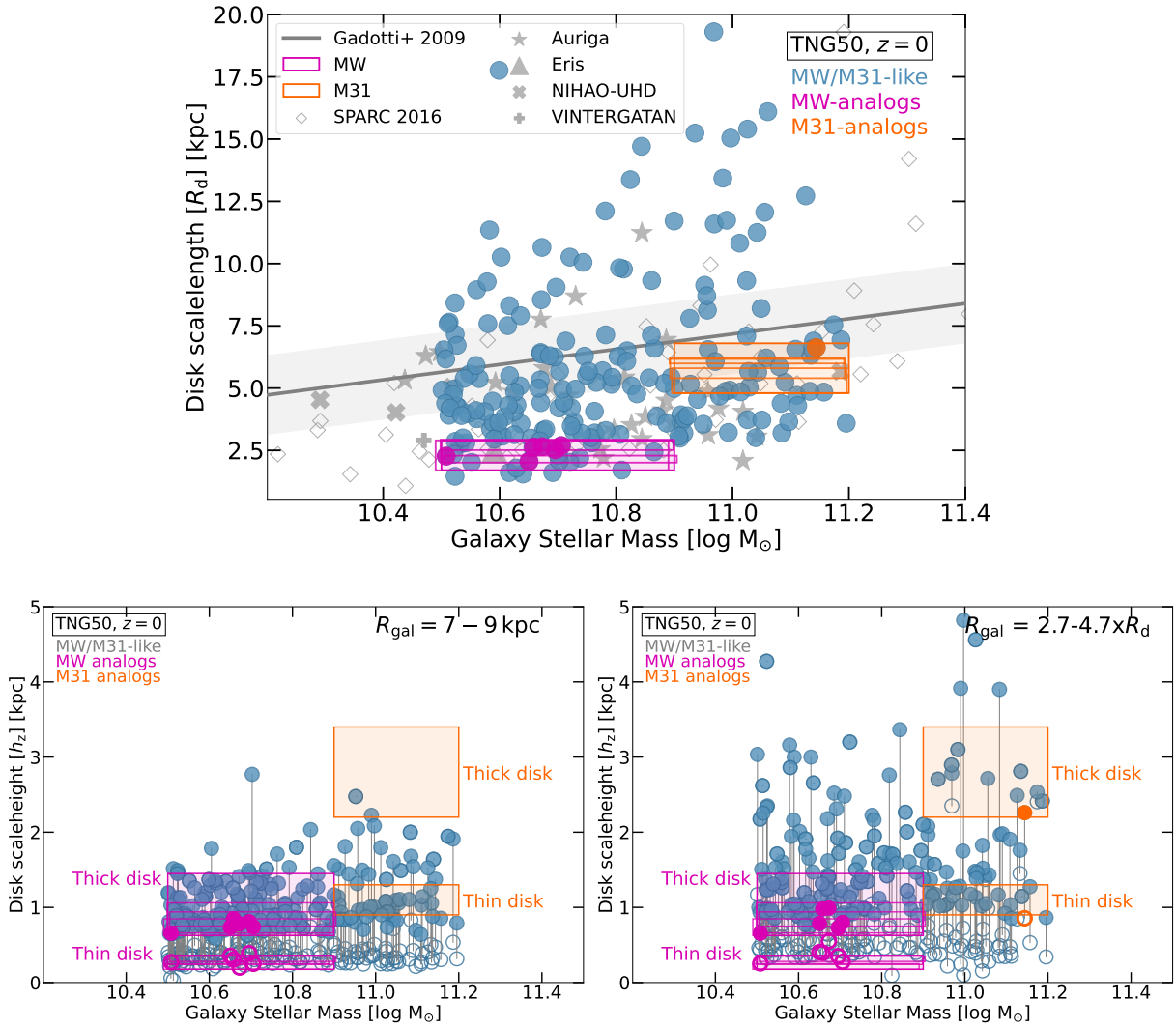

Figure 9: **Structural properties of the disks of MW/M31-like galaxies in TNG50** disk scale-lengths (top) and disk scale-heights (bottom panels, evaluated at cylindrical shells located from the center at 7-9 kpc (left) and at 2.7-4.7×R_d (right)) as a function of the galaxy stellar mass. The magenta and orange rectangles represent observational estimates for the MW and M31. In the top panel, data from other simulations is represented with grey symbols. In the bottom panels, the thin and thick disk of each galaxy are connected with a thin solid line. There are a few galaxies identified as MW and M31 'best' analogs (6 MW, 1 M31) according to their mass, disk length and disk height, represented with magenta and orange dots in the panels.

4 The merger and assembly histories of Milky Way- and M31-like galaxies with TNG50: disk survival through mergers

The work of this chapter is published as:
The merger and assembly histories of Milky Way- and M31-like galaxies with TNG50: disk survival through mergers
Diego Sotillo-Ramos, Annalisa Pillepich, Martina Donnari, Dylan Nelson, Lukas Eisert, Vicente Rodriguez-Gomez, Gandhali Joshi, Mark Vogelsberger and Lars Hernquist
Monthly Notices of the Royal Astronomical Society,
Volume 516, Issue 4, November 2022, Pages 5404–5427

Zusammenfassung

We analyze the merger and assembly histories of Milky Way (MW) and Andromeda (M31)-like galaxies to quantify how, and how often, disk galaxies of this mass can survive recent major mergers (stellar mass ratio $\geq$ 1:4). For this, we use the cosmological magneto-hydrodynamical simulation TNG50 and identify 198 analog galaxies, selected based on their $z = 0$ stellar mass ($10^{10.5-11.2}\ M_{\rm stars}$), disky stellar morphology and local environment. Firstly, major mergers are common: 85 per cent (168) of MW/M31-like galaxies in TNG50 have undergone at least one major merger across their lifetime. In fact, 31 galaxies (16 per cent) have undergone a recent major merger, i.e. in the last 5 Gyr. The gas available during the merger suffices to either induce starbursts at pericentric passages or to sustain prolonged star formation after coalescence: in roughly half of the cases, the pre-existing stellar disk is destroyed because of the merger but reforms thanks to star formation. Moreover, higher merger mass ratios are more likely to destroy the stellar disks. In comparison to those with more ancient massive mergers, MW/M31-like galaxies with recent major mergers have, on average, somewhat thicker stellar disks, more massive and somewhat shallower stellar haloes, larger stellar ex-situ mass fractions, but similarly massive kinematically-defined bulges. All this is qualitatively consistent with the different observed properties of the Galaxy and Andromeda and with the constraints on their most recent major mergers, 8 – 11 and 2 Gyr ago, respectively. According to contemporary cosmological simulations, a recent quiet merger history is not a pre-requisite for obtaining a relatively-thin stellar disk at $z = 0$.

4.1 Introduction

The *Lambda* cold dark matter (ΛCDM) cosmological model successfully explains the evolution of the Universe and the formation of the large-scale cosmic web. Structures grow hierarchically, with more massive haloes growing by smooth accretion and by merging with smaller ones (Genel et al., 2010), in a process dominated by (cold) dark matter, down to galactic scales.
In this context, mergers and interactions between galaxies are expected to be a frequent phenomenon, more so the higher the redshift (Fakhouri & Ma, 2008). In fact, galaxy mergers are known to play a central role in the mass growth of galaxies. Within the hierarchical growth of structure scenario, more massive haloes and galaxies undergo, on average, a larger number of mergers than lower-mass ones (Fakhouri et al., 2010, Rodriguez-Gomez et al., 2015). Moreover, more massive galaxies have been shown, via cosmological hydrodynamical galaxy simulations, to be made of larger fractions of stellar material that is accreted from mergers and orbiting satellites (e.g. Rodriguez-Gomez et al., 2016, Pillepich et al., 2018b).

Following a suggestion originally by Toomre (1977), galaxy mergers have also been shown to be able to drive morphological transformations in galaxies, i.e. to turn disky and rotationally-supported spiral galaxies into elliptical and dispersion-dominated ones. Already a few decades ago, it had been demonstrated via N-body simulations of isolated and idealized systems that the merger of two stellar disks generates descendants with spheroidal morphology (e.g. Barnes, 1988, Hernquist & Barnes, 1991, Barnes, 1992, Hernquist, 1993). Yet, more recently, it has also been found, with full-physics, cosmological hydrodynamical simulations of large samples of objects, that galaxies with similar stellar mass have undergone, on average, similar merger histories irrespective of their stellar morphology upon inspection (e.g. Rodriguez-Gomez et al., 2017, Martin et al., 2018, Jackson et al., 2020, with Illustris and Horizon-AGN).
In the observed Universe, galaxies with disky stellar morphologies are ubiquitous, ranging across wide mass and size scales. In our most immediate vicinity, there are three representative examples: the Milky Way (MW), Andromeda (M31), and Triangulum, all with stellar masses exceeding a few $10^{10}\,M_{\rm stars}$. Beyond the Local Group, at the MW mass scale, stellar disk morphology dominates, i.e. roughly two thirds of the observed galaxies exhibit stellar disk morphology: this is the case not only in the low-redshift Universe (e.g. results with the SDSS and GAMA surveys, Park et al., 2007, Kelvin et al., 2014, respectively) but also up to $z \sim 1.2$ (e.g. results with VIMOS-VLT data by Ilbert et al., 2006).
Recent observations have enabled estimates of the merger histories of both the MW and M31. It is now possible to measure the trajectories of satellite galaxies surrounding the Galaxy and Andromeda, and, for the case of the MW, also the positions and velocities of individual stars (e.g. with RAVE, APOGEE, GAIA, Kunder et al., 2017, Majewski et al., 2017, Lindegren et al., 2016), both in the disk and the stellar halo, and in the satellite galaxies, together with ages and metallicities for many of these stars (e.g. with LAMOST and GALAH, Deng et al., 2012, De Silva et al., 2015). A series of stellar streams have been identified in the stellar halo that are associated with merger events.
Recently, Helmi et al. (2018) and Belokurov et al. (2018) have used data from the *Gaia* mission (Gaia Collaboration et al., 2018) and inferred that the Galaxy has undergone a major merger event sometime at $z = 1 - 2$ (see also Bonaca et al., 2020, Xiang & Rix, 2022): the merging galaxy has been dubbed Gaia-Sausage-Enceladus (GSE) and its mass at the time of the collision has been estimated to be $\sim 5 - 6 \times 10^8\,M_{\rm stars}$ in stars (Naidu et al., 2021). The GSE merger has been associated theoretically with the formation of the inner stellar halo and the thick disk (Grand et al., 2020). Dynamical and chemical evidence of additional events in the MW's past has also been found, i.e. from the stellar remnants of past merging galaxies: these include Sequoia, accreted at a comparable epoch as GSE (Myeong et al., 2019); Thamnos (Koppelman et al., 2019); Helmi Streams (Helmi et al., 1999), that originated from a dwarf galaxy accreted approximately 5-8 Gyr ago (Koppelman et al., 2019), and the long-known Sagittarius Stream (Ivezić et al., 2000). The latter is still clearly identifiable also in configuration space and is being stripped from the Sagittarius dwarf of $\sim 5 \times 10^8\,M_{\rm stars}$ in stars, which is in the process of merging with the MW and whose infall time has been estimated to have occurred ~8 Gyr ago (Dierickx & Loeb, 2017). These, together with the ongoing mergers with the Large and Small Magellanic Clouds (with stellar masses of $\sim 2.7 \times 10^9\,M_{\rm stars}$ and $\sim 3.1 \times 10^8\,M_{\rm stars}$, respectively (van der Marel, 2006), and thus with a current stellar mass ratios of ~1:10 and ~1:100) constitute the six most massive accretion events across the known history of the MW, with the GSE being the largest in estimated mass ratio.

For M31, observational constraints are harder to obtain. However, for example, D'Souza & Bell 2018 recently suggested, by combining the results of cosmological models of galaxy formation and measurements of the ages and metallicities of halo stars, that Andromeda underwent a major merger about 2 Gyr ago (D'Souza & Bell, 2018) and that the observed satellite galaxy M32 is the remnant core of the secondary galaxy involved in the merger: the latter produced the giant stellar stream that is located in Andromeda's stellar halo, generated a recent burst of star formation, but did not destroy the stellar disk.

The merger histories of MW-mass haloes have also been studied from a theoretical perspective, and via cosmological simulations. For example, Stewart et al. (2008) used DM-only or gravity-only N-body ΛCDM simulations and quantified that 95 per cent of MW-mass haloes, i.e. gravitationally-collapsed objects of total mass of $\sim 10^{12} M_{\rm stars}\ {\rm h}^{-1}$ at $z = 0$, have merged with at least one other halo more massive than 1:20 in the last 10 Gyr, this fraction reducing to 70 per cent for mergers with haloes more massive than 1:10. They concluded that halo-halo mergers involving (total) mass ratios of up to 1:5 must not destroy the stellar disks of the hosted galaxies if approximately two thirds of the observed galaxies are disky at $z = 0$.

The reasoning mentioned above encapsulates and exemplifies a debate that has gone on for a few decades: namely, how to reconcile the fact that disk galaxies are found everywhere and at the same time the number of mergers that potentially would transform or destroy them are frequent. In the case of the Galaxy, Toth & Ostriker (1992) claimed via analytical arguments that a thin stellar disk similar to the Milky Way's is only compatible with no minor nor major mergers in the last 10 Gyr, a conclusion similar to that derived via observational data as early as by Wyse (2001) and Hammer et al. (2007).

However, from a theoretical and numerical perspective, it has long been known (e.g. Hernquist, 1989, Barnes & Hernquist, 1996) that the formation and evolution of stellar disks – also and especially in the presence of perturbation mechanisms such as major mergers – cannot be addressed without accounting for the role and availability of gas. Numerical dissipational simulations of idealized mergers by e.g. Springel & Hernquist 2005, Naab et al. 2006, Robertson et al. 2006 showed that the remnants of *gas-rich* mergers can lead to the formation of a spiral and not an early-type galaxy.

In the full cosmological context, i.e. with zoom-in cosmological N-body and hydrodynamical simulations of galaxies, the formation of thin stellar disks in MW-mass galaxies could be achieved only after solving for the "angular momentum problem" or "overcooling catastrophe" and after introducing appropriate (stellar) feedback and star formation recipes (Guedes et al., 2011, Agertz et al., 2011) and refined numerical approaches (Kereš et al., 2012, Sijacki et al., 2012, Vogelsberger et al., 2012). A number of successes followed, with simulations capable of returning large, disk-dominated galaxies that resemble the Milky Way in many respects (e.g. Martig et al., 2012, Stinson et al., 2013, Marinacci et al., 2014). Still, as of just some years ago, such zoom-in simulations of one or a few objects seemed to indicate that disk-dominated galaxies with stellar structures in accord with observations of the Galaxy could emerge, if, and only if, there is no major merger at $z \lesssim 1$ (Rix & Bovy, 2013).

The advent of large-volume cosmological hydrodynamical galaxy simulations (Vogelsberger et al., 2020) such as Illustris (Vogelsberger et al., 2014a,b, Genel et al., 2014, Nelson et al., 2015) and EAGLE (Schaye et al., 2015, Crain et al., 2015) allows us to reconsider the formation of disk galaxies from major mergers in a quantitative manner, thanks to thousand-strong galaxy samples, spanning wide mass ranges and without prior constraints on the past assembly histories

or environments. For example, with Horizon-AGN (Dubois et al., 2014) and Illustris galaxies, Martin et al. (2018) and Peschken et al. (2020) have confirmed that if enough gas is available during a merger event, new stars can form and can reform a stellar disk even if the latter gets destroyed in the collision. A similar scenario has been previously explored using zoom-in simulations by Sparre & Springel (2017), who considered four ~MW-mass (at $z = 0$) galaxies that were disky and underwent one major merger at $z \sim 0.5$. Focusing on very massive disk galaxies from Horizon-AGN ($M_* \geq 10^{11.4}\,M_{\rm stars}$), Jackson et al. 2020 showed that, in the majority of cases, pre-existing stellar disks either survive major mergers or reform shortly after a merger (a pathway in qualitative agreement with observations, e.g. Rothberg & Joseph, 2004, McDermid et al., 2006) and, in fewer cases, stellar disks form after a spheroidal galaxy is established (Hau et al., 2008, Kannappan et al., 2009).
However, so far no quantitative analysis of the merger history and disk survival through mergers has been put forward with a focus on simulated analogues of the Galaxy and Andromeda. On the one hand, zoom-in simulation campaigns have begun to build up impressive samples of Milky Way-like galaxies: however, these are typically not unbiased in formation history/isolation (e.g. Auriga, Grand et al., 2017) or are selected in relatively narrow ranges of total halo mass (e.g. Artemis, Font et al., 2020). On the other hand, large-volume simulation projects have lacked the resolution to capture stellar disks as thin as a few hundred parsecs. Large, unbiased samples at relatively high-resolution are now possible with the cosmological magnetohydrodynamical simulation TNG50 (Nelson et al., 2019b, Pillepich et al., 2019), the highest-resolution run of the IllustrisTNG project (`www.tng-project.org`). There we can identify approximately two hundred galaxies at $z = 0$ with global properties similar to the MW's and M31's and with stellar disk heights as small as $100 - 200$ parsecs. In addition to gravity and hydrodynamics, these simulations account for the effects of stellar and AGN feedback, magnetic field evolution, gas cooling and heating, etc.
In this paper, we hence describe the TNG50 cosmological simulation, the method to create the merger trees, and the selection criteria for the MW/M31 analogues in Section 6.2. We quantify the merger and assembly histories of TNG50 MW/M31-like galaxies in Section 4.3. In Section 4.4 we delve into the evolutionary pathways and physical mechanisms at play in the case of MW/M31 analogues that underwent at least one major merger in recent epochs, e.g. in the last 5 Gyr. We compare the $z = 0$ structural and global properties of these galaxies with the rest of the TNG50 MW/M31-like objects in Section 5.3 and summarize and discuss our findings in Section 6.4.

4.2 Methods: simulation and sample selection

4.2.1 The TNG50 simulation

In this paper, we focus on Milky Way and Andromeda-like (MW/M31-like) galaxies realized within the TNG50 cosmological simulation. TNG50 (Nelson et al., 2019b, Pillepich et al., 2019) is the highest-resolution run of the IllustrisTNG project (Naiman et al., 2018, Marinacci et al., 2018, Pillepich et al., 2018b, Nelson et al., 2018, Springel et al., 2018). It consists of a periodic cubic volume with side length of ≈ 51.7 comoving Mpc, contains 2160^3 dark matter (DM) particles and the same initial number of gas cells. The mass of the DM particles is uniform, $m_{\rm DM} = 4.5 \times 10^5$ M$_\odot$, and the average mass of the gas cells (and stellar particles) is $m_{\rm baryon} = 8.5 \times 10^4$ M$_\odot$. The code Arepo (Springel, 2010c), which makes use of an unstructured

moving mesh that is based on the Voronoi tessellation of the space and that is automatically adaptive in resolution, solves the gravity and magnetohydrodynamical equations in an expanding universe from $z \approx 127$ to $z = 0$. Additionally, a series of physical processes are included to follow galaxy formation: these and their implementation are described in detail in the methods papers by Pillepich et al. 2018a, Weinberger et al. 2017. Crucially and differently from most high-resolution simulations of Milky Way-like galaxies (e.g. zooms without SMBH feedback by Wetzel et al., 2016a, Buck et al., 2020, Renaud et al., 2021), the IllustrisTNG model not only includes star formation and its ensuing expected feedback and metal enrichment, but also magnetic fields and the seeding, growth, and feedback from supermassive black holes (SMBHs). All simulations in the IllustrisTNG series adopt values for the cosmological parameters from Planck Collaboration et al. (2016).

4.2.2 Halo identification and histories of simulated galaxies

Throughout our analysis, the identification of haloes and subhaloes in the simulated domain is performed with the Friends-of-Friends (FoF; Davis et al., 1985) and SUBFIND (Springel et al., 2001b) algorithms, respectively. The latter identifies all particles and resolution elements that are measured to be gravitationally-bound to a self-gravitating structure.
We define as galaxies those subhaloes with non-vanishing stellar mass. The "central" galaxy of a FoF halo is typically the most massive one, whereas the remaining subhaloes within a FoF halo are called satellites. Throughout the analysis, we consider only subhaloes with cosmological origin, i.e. non-spurious subhaloes (Nelson et al., 2019b), and galaxies with at least 50 stellar particles.

Merger histories and merger summary statistics

Information for all resolution elements, subhaloes, and haloes is stored for one hundred snapshots, starting at $z = 20$ (see Nelson et al., 2019b, for details). At recent epochs, the time spacing between consecutive snapshots is approximately 150 million years. The (sub)haloes at different snapshots are linked to describe their merger histories (or merger trees) using the SUBLINK (Rodriguez-Gomez et al., 2015) and LHALOTREE (Springel et al., 2005) algorithms. Both return similar results. In this work, we use SUBLINKGAL, which connects subhaloes across cosmic epochs based on their star particles and star-forming gas cells, instead of DM particles. According to SUBLINKGAL (described in detail Rodriguez-Gomez et al., 2015), the main progenitor of a galaxy is the one with the most massive history behind it: we will refer to its temporal evolution as the "main-progenitor history" (or "main-progenitor branch").
In the following, we describe the merger history of a galaxy by analyzing the time of its mergers and the *stellar-mass* ratio between the two involved galaxies, i.e. between the progenitors. Namely:

- *Time of a merger:* As per Rodriguez-Gomez et al. (2015), the time of a merger ($t_{\rm merger}$) is given by the earliest snapshot when both progenitors are identified by SUBFIND as part of the same unique subhalo. Such a merger time represents the time of coalescence or just after it, i.e. at the next available snapshot.

- *Time of the merger-ratio estimate:* The stellar mass ratio (the smallest of the ratios between the secondary and the primary) is measured at the time when the secondary reached its maximum stellar mass ($t_{\rm max}$). This can be considered as the time when a merger starts.

- *Merger duration:* The duration of a merger is the time elapsed between $t_{\rm max}$ and the time of coalescence, i.e. the time of a merger. This is therefore similar to the infall time, i.e. the time spent by the satellite in orbit within the host halo.

Mergers are classified into three types: major, whereby the stellar-mass ratio is 1:4 or larger; minor, with ratios between 1:10 and 1:4, and mini (below 1:10). In the case of several mergers occurring at the same time, i.e. between the same pair of snapshots for the same galaxy, we account for all of them, separately, and their stellar mass ratios are recorded separately for each binary event.
In comparison to the default functioning of the SublinkGal code, we apply additional measures to avoid counting as mergers the interactions with spurious subhaloes or with subhaloes that are resolved with too few stellar particles, i.e. to account for the selections mentioned above. The former may be objects without a cosmological origin (e.g. fragments of disks or other types of galactic sub structures) that Subfind identifies as subhaloes. They are identified with the *SubhaloFlag* (Nelson et al., 2019a) and are here excluded from the merger statistics. Galaxies whose main progenitor history is very brief (less than three snapshots before the time of coalescence) are also excluded. Finally, we also neglect all mergers between galaxies with fewer than 50 star particles ($\sim 5 \times 10^6 {\rm M}_\odot$), across cosmic epochs. Thinking about the merger history of massive galaxies like MW/M31 analogues at $z = 0$, this allows us to avoid counting as major those mergers whose $t_{\rm max}$ may occur very early on in the cosmic history of a galaxy, when primary and secondary objects might have had comparable masses, but whose actual coalescence occurs several billion years later and at times when the two progenitors actually differ by many orders of magnitude in mass.
The three conditions above allow us to obtain a clean and complete catalog of merger events, particularly at lower redshifts. On the other hand, these conditions also imply that our analysis is neither sensitive nor complete for the merger history of MW/M31-like main progenitors prior to $z \sim 5$.

In-situ vs. ex-situ stars

The cosmological simulations allow us to also follow the evolutionary tracks of individual stellar particles, so that we can identify the galaxy where each formed and the location at subsequent times. Based on the merger trees and for any galaxy at $z = 0$, we categorize its stars as in situ if they formed in a progenitor that belongs to the main-progenitor branch of the galaxy, and ex situ otherwise, according to the definitions and catalogs described in Rodriguez-Gomez et al. (2016). Ex-situ stars are hence stellar particles stripped and consequently accreted into a galaxy from its mergers and orbiting satellites.

4.2.3 Galaxy and stellar-particle properties: circular orbits and diskyness or D/T ratio

Throughout the paper, the following conventions are intended unless otherwise stated.

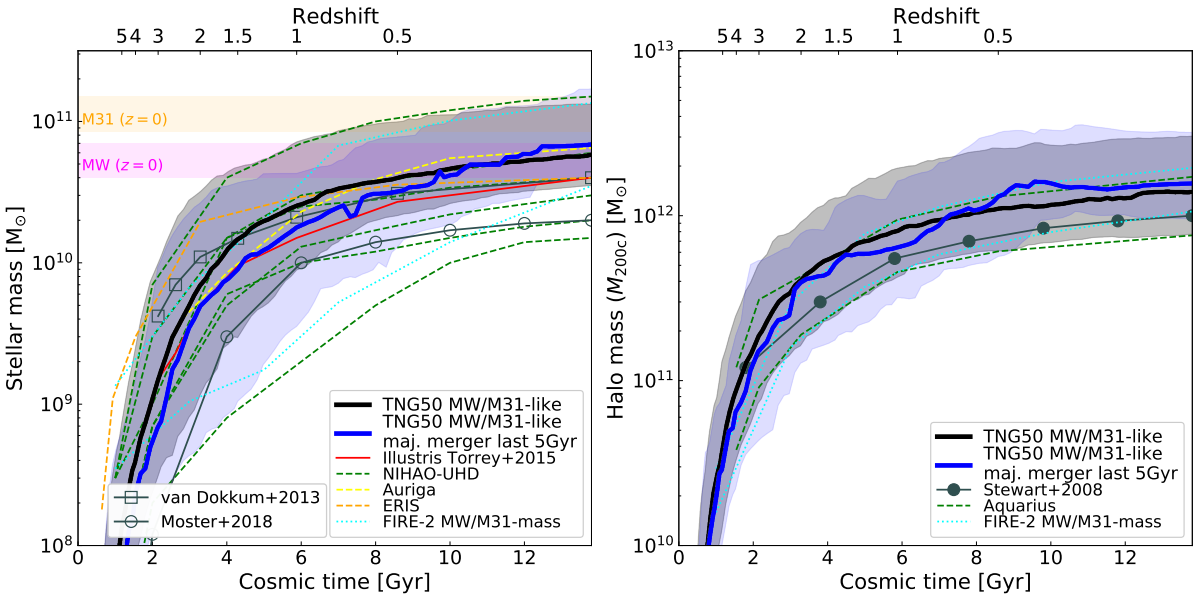

Figure 10: Mass assembly histories of MW/M31 analogues in TNG50, for galaxy stellar mass (left) and total host mass (right). We show the complete TNG50 MW/M31 sample in black and a sub-sample of galaxies with recent major merger in blue (see next Sections): median tracks at fixed cosmic time are marked as thick curves, whereas shaded areas denote interquantiles (10th-90th percentiles). Additional curves represent results from other simulations (colored lines) or observationally-derived constraints (grey lines with empty markers). For TNG50 and the additional simulations, the subhalo's total stellar mass is plotted. For $z = 0$ MW/M31-like galaxies, 10 per cent of their stellar mass is assembled by $z = 2$ and 50 per cent by $z = 1$. At $z = 2$ the scatter in galaxy stellar mass extends for two orders of magnitude ($1.8 \times 10^8 - 2.2 \times 10^{10}$ M$_\odot$, 10th-90th percentiles), in contrast to a $z = 0$ selection ranging 0.7 dex. In terms of total host mass, $z = 0$ MW/M31-like galaxies have assembled 50 per cent of their halo mass by $z = 2$. We show annotations with observational estimates for the stellar mass at $z = 0$ of the MW (in magenta, as an envelope of the values from Flynn et al., 2006, Licquia & Newman, 2015, Bland-Hawthorn & Gerhard, 2016) and M31 (orange, from Geehan et al., 2006, Barmby et al., 2006, Tamm et al., 2012, Sick et al., 2014).

Galaxy stellar mass: sum of the mass of all gravitationally-bound star particles.

Total host mass: the total mass of the FoF group in which a galaxy is found, $M_{200c,\mathrm{Host}}$, measured inside a sphere whose mean density is 200 times the critical density of the Universe (at the considered time).

Stellar half-mass radius of a galaxy ($R_{1/2}$): the radius of the sphere, centered at the particle with the minimum gravitational potential energy in the galaxy, that contains half of the stellar mass.

Star formation rate (total, SFR) is measured by adding the individual instantaneous SFRs of all gas cells that are gravitationally-bound to a galaxy at the time of inspection.

Other targeted galaxy properties are described in the course of the analysis.
For any given galaxy, the orbital properties of its stars are described via the circularity parameter, according to the definition by Scannapieco et al. (2009): $\epsilon = j_z/j_{\mathrm{circ}}$, where j_z is the specific angular momentum of the star in the direction perpendicular to the galactic disk, and j_{circ} is the specific angular momentum of a star at the same radius, but on a perfectly circular orbit. Namely, $j_{\mathrm{circ}} = r v_{\mathrm{circ}}$, with $v_{\mathrm{circ}} = \sqrt{GM(\leq r)/r}$ being the circular velocity of the galaxy at the considered radius. The up vector or vertical axis of a galaxy is chosen as the direction of the total angular momentum of all stars within two half-mass radii. Stellar orbits with $\epsilon \gtrsim 0.7$ are considered circular, i.e. in rotational motion.
We define the diskyness or disk-to-total (D/T) ratio of a galaxy as the fractional mass of stars in circular orbits, i.e. with $\epsilon > 0.7$, minus the fraction of stars with $\epsilon < -0.7$, hence removing the contributions of bulge and stellar halo and assuming that the latter are symmetric around $\epsilon = 0$. The stellar mass in circular orbits and the total stellar mass are evaluated, for the purposes of D/T, within an aperture of five stellar half-mass radii.
The exact choice of separating circular and cold orbits from the rest with the threshold $\epsilon > 0.7$ is not crucial: Aumer et al. (2013a) found that the D/T ratio obtained from our circularity definition is roughly equivalent to the values obtained with another usual definition of the circularities, based on the ratio of the specific angular momentum of the star and the maximum specific angular momentum at the specific binding energy of the star, using in both cases a similar circularity threshold.

4.2.4 Sample selection: MW/M31-like galaxies in TNG50

From the TNG50 box, which at $z = 0$ samples hundreds of massive galaxies, we select the best analogues of the Milky Way and Andromeda based on their $z = 0$ properties, i.e. without imposing conditions on their evolution or any restrictions on their morphological changes with time.
In practice, we use the selection presented and motivated in Pillepich et al. in preparation and already used by Engler et al. (2021a), Pillepich et al. (2021) and others. To be selected as "MW- or M31-like", a TNG50 galaxy must meet, at $z = 0$, the following three conditions, based on galaxy stellar mass, stellar diskyness, and isolation:

A) Galaxy stellar mass: $\log_{10}(\mathrm{M}_*/\mathrm{M}_\odot)$ within [10.5, 11.2], with stellar mass measured within a 3D circular aperture of 30 kpc.

B) Diskyness: either of the following conditions:

   (a) Stellar minor-to-major axis ratio (Pillepich et al., 2019), c/a, < 0.45, whereby c and a are respectively the minor and major axis of the ellipsoidal distribution of the stellar mass between one and two times the stellar half-mass radius.

   (b) Disky by visual inspection of 3-band images, both face-on and edge-on.

C) Isolation: no galaxy with galaxy stellar mass $\geq 10^{10.5}$ $M_\odot$ within 500 kpc distance and total host mass $M_{200c,\mathrm{Host}} < 10^{13}$ $M_\odot$.

There are 198 galaxies in the TNG50 simulation that match such conditions.
The stellar mass constraint is an envelope of the most accurate available estimates of the stellar mass of both the MW and M31 (e.g. Bland-Hawthorn & Gerhard, 2016, Boardman et al., 2020). The diskyness criteria aim at including all simulated galaxies with a global stellar disk morphology and with spiral arms. Whereas typically, disky galaxies are defined as those with minor-to-major axis ratio < 0.33 (and middle-to-major axis ratio > 0.66; see e.g. fig. 8 of Pillepich et al. (2019) and reference therein), the connection of this metric to other measures of "diskyness" (e.g. kinematic or photometric D/T ratios, scale length to scale height ratio, etc.) can be very varied, depending also on where, within a galaxy's body, the measurements are taken. Here we adopt a looser upper limit on the minor-to-major axis ratio to accommodate for the ratio of e.g. the stellar disk height to the stellar disk length of the Galaxy, particularly of its geometrically-thick component. The latter measures ~ 0.45 (in comparison to ~ 0.11 for the geometrically-thin disk) when adopting the best values for the structural properties of the Milky Way disk from Bland-Hawthorn & Gerhard (2016, thin and thick disk lengths = 2.6 ± 0.5 and 2 ± 0.2 kpc, respectively, and thin and thick disk heights at the Sun location = 300 ± 50 and 900 ± 180 pc). The isolation criterion avoids the presence of a galaxy, with mass equal to the lower MW estimate or larger, at a distance closer than 500 kpc – the distance between the Galaxy and Andromeda is ~ 770 kpc (McConnachie et al., 2005, Riess et al., 2012). Additionally, the requirement on total host mass is meant to exclude galaxies located within a very massive galaxy group or cluster, as we know this is not the environment of the Galaxy and Andromeda. However, it is more relaxed than requiring a galaxy to be the central of its halo, while allowing certain galactic environments such as those similar to the Local Group, whereby the Galaxy and Andromeda may be sharing the same dark matter halo.
As the galaxies are selected from a volume-limited sample, a larger number of TNG50 objects have stellar mass closer to the MW's than to the more massive M31's. In fact, among the 198 TNG50 MW/M31 analogues, 130 galaxies have masses more compatible with the MW's ($\leq 10^{10.9}\, M_{\mathrm{stars}}$), and will be dubbed 'MW-mass' and 68 are instead more representative of Andromeda and will be dubbed 'M31-mass'.
No constraints on the past history are imposed, in contrast to what has been typically done with cosmological zoom-in simulations of Milky Way analogues – e.g. those by Governato et al. (2004), Okamoto et al. (2005), Eris (Guedes et al., 2011), Agertz & Kravtsov (2015), GARROTXA (Roca-Fàbrega et al., 2016), VINTERGATAN (Renaud et al., 2021), AGORA (Roca-Fàbrega et al., 2021) –, usually by limiting the last major merger to redshift $z \geq 1$. Furthermore, no a-priori limitations are placed for e.g. the range of current star formation rates, stellar disk structural properties, and SMBH masses see for an overview Pillepich et al. in preparation.

4.3 The past histories and merger statistics of MW/M31-like galaxies, according to TNG50

4.3.1 Mass assembly

We quantify the assembly history of TNG50 MW/M31-like galaxies by plotting their mass across cosmic epochs, i.e. along their individual main progenitor branches. The evolution of the galaxy stellar mass (left) and of the total host mass (within R_{200c}, right) of all MW/M31 analogues are shown in Fig. 10. For TNG50 and the additional simulations the total (i.e. gravitationally-bound) stellar mass is plotted. Thick curves denote the median evolutionary tracks across the galaxy populations at fixed redshifts, whereas the shaded areas indicate the 10th-90th percentiles. It should be noted that, although the medians appear smooth, individual assembly histories may exhibit assembly histories that are very different from the median curves.

The selection by galaxy stellar mass at $z = 0$ – in the range of $3.2 \times 10^{10} - 1.6 \times 10^{11}$ $M_\odot$, within 30 kpc aperture as per Section 4.2.4, left panel – translates into a $z = 0$ host mass range of $M_{200c} = 8.3 \times 10^{11} - 2.5 \times 10^{12}$ $M_\odot$ at $z = 0$, i.e. 0.48 dex, within 10th-90th percentiles (right panel): this is a consequence of the effective stellar-to-halo mass relation resulting in TNG50 (Engler et al., 2021a). On average, the host haloes of MW/M31-like galaxies grow by a factor of ~ 7 between $z = 3$ and today, while the average galaxy stellar mass of their progenitors at $z = 3$ was a factor of ~ 40 lower than today.

The progenitors of MW/M31-like galaxies span a host mass range (within 10th-90th percentiles) at $z = 2$ similar to that at $z = 0$, namely $0.5 - 0.6$ dex: on the other hand, the progenitors of MW/M31-like galaxies at e.g. $z = 2$ span almost three orders of magnitude in stellar mass across all sampled galaxies: $1.8 \times 10^8 - 2.2 \times 10^{10}$ $M_\odot$ within the 10th-90th percentiles, i.e. 2 dex. The pathways leading to MW/M31-like galaxies at $z = 0$ are rich and varied.

These theoretical predictions are of the essence to connect galaxy populations across cosmic time for the purposes of contrasting e.g. MW-like galaxies with their expected progenitors at high redshifts. For comparison, in Fig. 10 we show the corresponding results from previous cosmological simulations of MW-like galaxies, albeit differently selected: solid and dashed thin lines. The thirty zoom-in Auriga MW analogues (yellow dashed curve, for their median growth, Grand et al., 2017), where galaxies are selected as isolated haloes at $z = 0$ with M_{200c} in the range $(1 - 2) \times 10^{12}$ $M_{\rm stars}$, have a very similar median stellar assembly history to those from TNG50. For the Illustris simulation (Torrey et al., 2015a), where MW analogues are all galaxies with stellar mass within the range $(4 - 5) \times 10^{10}$ $M_{\rm stars}$, the median stellar mass is lower at all redshifts (red solid curve). The six NIHAO-UHD MW-like galaxies (green dashed curves, Buck et al., 2020) span a similarly wide range in mass growth across all cosmic epochs as TNG50. On the other hand, the Eris galaxy (orange dashed curve, Guedes et al., 2011) exhibits a much earlier mass growth (Pillepich et al., 2015), similarly as to the FIRE-2 MW-mass zoom-ins by Garrison-Kimmel et al. (2018), of whom we show the growth curves for the least and most massive at $z = 0$ (dotted cyan).

In Fig. 10, assembly histories obtained by assuming that galaxies preserve their number density in time (e.g. van Dokkum et al., 2013) are shown for contrast. As in the original Illustris simulation (Torrey et al., 2015a,b), also according to TNG50 and for MW/M31-like galaxies only, we find that the average stellar mass evolution inferred via a constant comoving number density assumption is systematically shallower than when tracking galaxies via their merger trees, with $z \sim 3$ progenitors characterized by galaxy stellar masses a factor of a few larger than

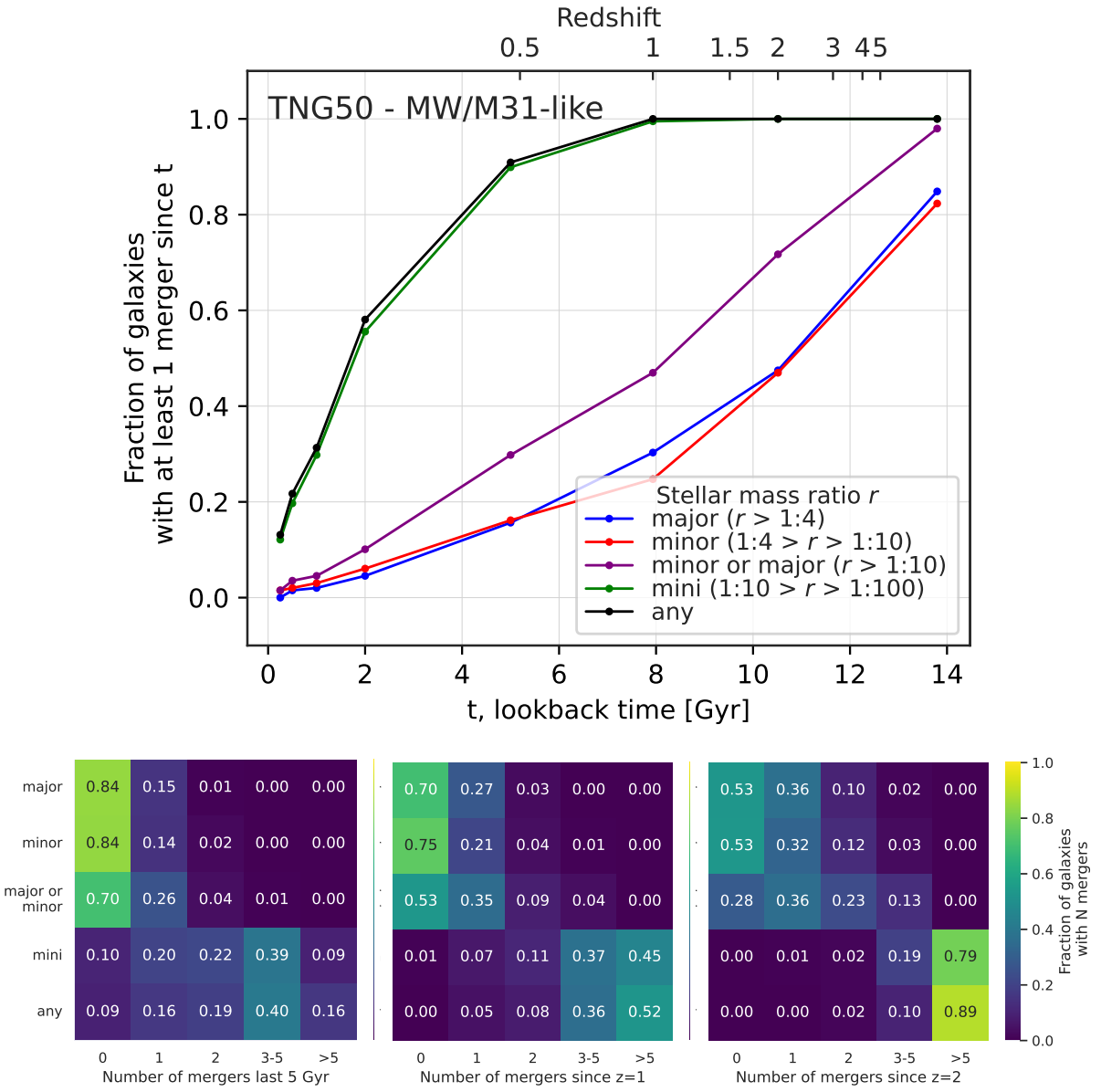

Figure 11: Number of mergers (minor, major, both, mini and any) in different periods of time for MW/M31-like galaxies in TNG50. In the tables, different rows correspond to different stellar mass ratios. *Top panel:* fraction of galaxies that have undergone at least one merger in different time periods (lookback time measured from $z = 0$) and different stellar mass ratios: major (red), minor (blue), minor or major (purple), mini (green) and any (black). Dots represent the exact periods for which the measurements are available. *Bottom panels*: fraction of galaxies undergoing different types of mergers for three defined time periods. *Left*: last 5 Gyr ($z \sim 0.5$). *Middle*: since $z = 1$ (lookback time ~ 7.98 Gyr). *Right*: since $z = 2$ (lookback time ~ 10.51 Gyr). The color denotes the fraction of galaxies undergoing a certain number of mergers. These fractions are not cumulative. 18 per cent of the MW/M31-like galaxies have recent (i.e. over the last 5 Gyr) major mergers. 34 per cent have a major merger since $z = 1$.

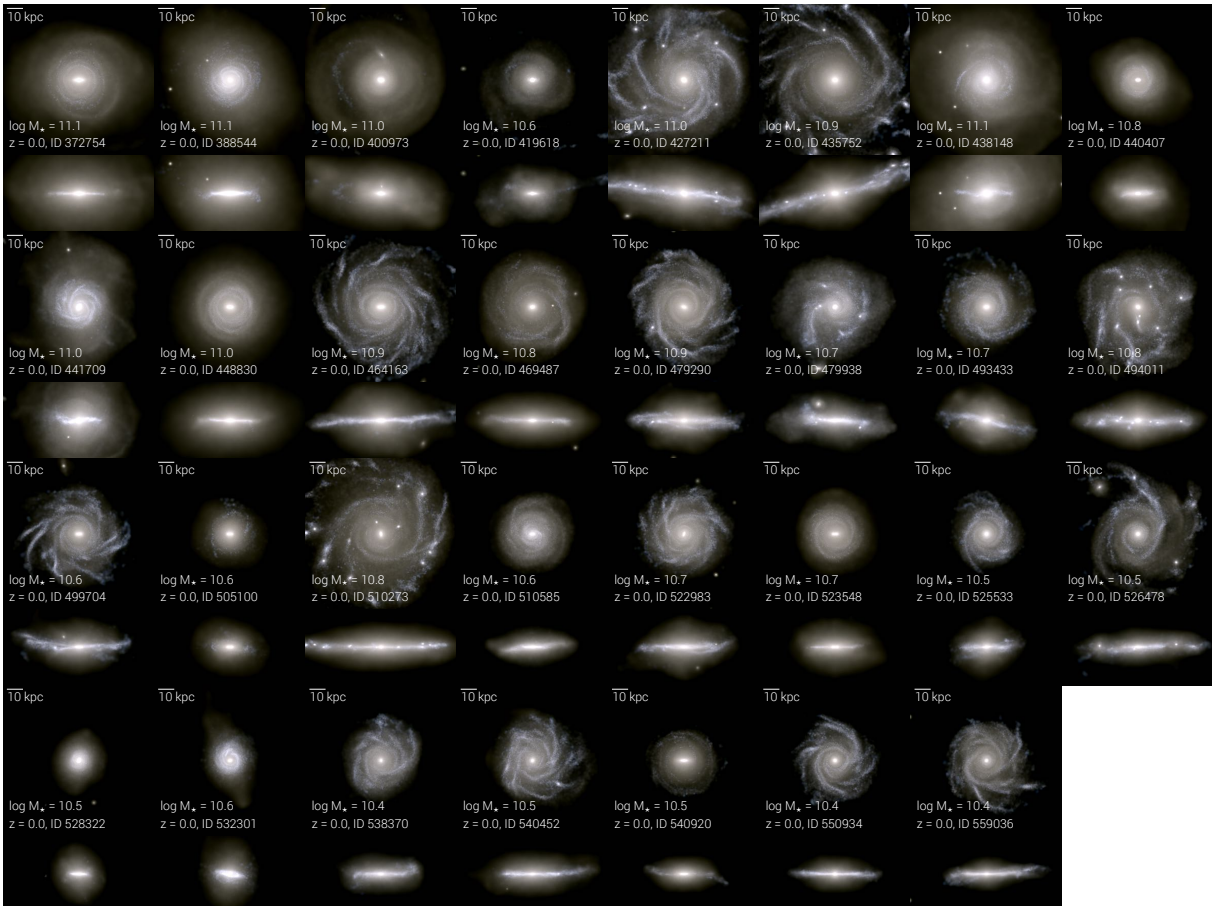

Figure 12: Stellar-light composite images of the 31 MW/M31-like galaxies from TNG50 at $z = 0$ that have undergone a recent major merger (i.e. in the last 5 Gyr). Every disk galaxy is shown in face-on and edge-on projections. Each panel spans horizontally 80 kpc, and each galaxy is identified by its stellar mass and Subhalo ID number in the TNG50 $z = 0$ catalog. These galaxies exhibit a great diversity in stellar sizes and in terms of stellar structures within the disks (see text for comments on galaxies with Subhalo IDs 400973 and 41961): for example, there are cases with marked spiral arms and others with central bars.

those obtained along the main progenitor branches of galaxies. Therefore, the mass evolution from constant comoving number density assumptions cannot be used to validate the results of zoom-in simulations (Buck et al., 2020).

4.3.2 Merger statistics

Fig. 11 shows the merger statistics of MW/M31-like galaxies according to TNG50, selected as described in Section 4.2.4. To our knowledge, this is the first time that such statistics, based on galaxy stellar masses rather than halo masses, are quantified and are obtained from a cosmological hydrodynamical simulation where MW/M31-like galaxies are selected based on observable (rather than host halo) properties.

The upper panel shows what fraction of MW/M31 analogues undergo at least one merger over varying past periods of time (in lookback Gyr from $z = 0$), for different stellar mass ratios: major, minor, major or minor, mini and any ratio. In the lower panels, we give the fractions of galaxies that experienced varying number of mergers, over three time periods: from left to right, in the last 5 Gyr (i.e since z $\sim$ 0.5), since $z = 1$ (lookback time $\sim$ 7.98 Gyr), and since $z = 2$ (lookback time $\sim$ 10.51 Gyr). The color code in the lower panels represents the fraction of MW/M31-like galaxies that underwent a certain number of mergers – these fractions are not cumulative. In all panels, mergers are counted based on the times of coalescence, i.e. the time of the mergers as defined in Section 4.2.2. About thirty per cent of MW/M31-like galaxies have undergone at least one major or minor merger (i.e. at least one merger with stellar mass ratio larger than 0.1) over the last 5 billion years: this fraction increases to 48 (72) per cent since $z = 1$ ($z = 2$).

Interestingly, for MW/M31-like galaxies, on average, the frequency of past major mergers is similar, if not slightly larger, than the frequency of minor mergers. Not only are the fractions of galaxies undergoing at least one minor or one major merger similar (upper panels), but also comparable are the fractions of galaxies undergoing different numbers of minor and major mergers since a fixed period of time (lower panels).

Whereas it has often been assumed, also prior to the newest results with Gaia, that our Galaxy has had a very quiet merger history at least since $z \sim 1$ with no major mergers since then (see Introduction), according to TNG50 and our selection, the fraction of MW/M31-like galaxies that have merged with at least one other similarly-massive galaxy since $z \sim 1$ is about 30 per cent. This fraction is slightly lower, 27 per cent (35 of 130 galaxies), if we only consider TNG50 analogues with stellar mass below $10^{10.9}$M$_\odot$, i.e. if we exclude Andromeda-mass galaxies.

From the lower panels of Fig. 11, it can also be seen that 9 per cent of MW/M31-like galaxies have not merged with any other galaxy, irrespective of its stellar mass, over the last 5 billion years. On the other hand, a handful (12 per cent) of the selected TNG50 galaxies have experienced multiple major mergers since $z = 2$, as many as between 3 and 5. Yet, more than one merger event per galaxy (major or minor) since e.g. $z = 1$ is infrequent[2].

It is important to notice that the numbers summarized in Fig. 11 are based on a full-physics (i.e. not DM only) model for galaxy formation in the full cosmological context, they apply to a specific galaxy selection, and the merger mass ratios are characterized in terms of the stellar mass (and not DM mass) of the merging objects. Fig. 11 can be considered as the stellar-mass

[2]For future references and analyses, the $z = 0$ Subhalo IDs of the TNG50 MW/M31-like galaxies with two or more major mergers since $z = 1$ read: 400973, 435752, 441709, 479938, 526478, 538370; and in the last 5 Gyr: 441709.

based update of e.g. Figs. 5 and 6 of Stewart et al. 2008, which were based on N-body only simulations, and the TNG updates of the merger estimates for MW-mass haloes given for Illustris by Rodriguez-Gomez et al. 2015 and quoted in the Introduction.

4.4 Mergers and disk survival

The findings in the previous Section imply that a selection of TNG50 galaxies at $z = 0$ have global properties – stellar mass, stellar diskyness, and large-scale environment – similar to those of the Galaxy and Andromeda and yet have undergone at least one major merger as recently as over the last 5 Gyr. In particular, 31 of the 198 MW/M31 analogues of TNG50 (16 per cent) are found in this subsample and are hence the focus of the rest of the paper. The choice to focus on galaxies with the time of their last major merger at 5 billion years ago, instead of e.g. 6 or 4 billion years ago, is somewhat arbitrary but is simply meant to qualitatively encapsulate the phenomenology of *relatively recent* major mergers. We will drop the binary classification whenever instructive and possible.

4.4.1 MW/M31-like galaxies as survivors of recent major mergers

In Fig. 32, we show stellar-light composite images at $z = 0$ for the subsample of TNG50 MW/M31 analogues with a major merger within the last 5 billion years, in face-on and edge-on projections. But for a couple of cases – Subhalo IDs 400973 and 419618, for which the identification as MW/M31 analog is in fact borderline – they pass the morphological selection criterion based on the shape of the inner stellar mass distribution but have D/T mass ratio of ~ 0.2 –, these galaxies exhibit clear stellar disk morphologies, although with a great variety in structural properties and extents. Their median D/T mass ratio at $z = 0$ (as defined in Section 4.2.3) is 0.40, but ranges between ~ 0.12 and ~ 0.78 across the whole sample of Fig. 32. For comparison, the D/T ratios of the full MW/M31-like sample range from $\sim$0.10 – 0.90 with a median of 0.55 and six systems below 0.1, all with counter-rotating structures (see also Joshi et al., 2020).
The geometrically-thin stellar disks of MW/M31-like galaxies with recent major mergers have sech^2 heights measured at 4 times the disk scale length spanning between $\sim$50 pc and 2.9 kpc, with median and average of 0.93 and 1.14 kpc (see Section 5.3 for more details). In numerous cases, the stars are organized in very prominent spiral arms and grand-design spiral systems, which are sometimes distributed asymmetrically. There are also central bars with an extension of $\sim 1-4$ kpc (e.g. Subhalo IDs 522983, 523548 and 540920). From the edge-on views, we can appreciate that the stellar haloes are also diverse, even among galaxies with similar stellar disk size: from faint haloes of stars that barely extend beyond the disk, to others that are appreciable in the images for up to a few tens of kpc above and below the disks. A quantification of the mass in the different stellar components (disks, bulges, and haloes) will be given in Section 5.3.
In Fig. 13, we provide additional statistics and properties of the last major mergers experienced by the MW/M31-like galaxies with a recent major merger (blue), in comparison to those of the whole MW/M31-like sample. Firstly, as appreciable also from Fig. 11, 15 per cent, i.e. 30 galaxies among the TNG50 MW/M31 analogues have never undergone a major merger, i.e. never since $z \sim 5$, see Method Section for details. In the top left panel of Fig. 13, these are reported as an orange bar: 6 of those are M31-mass galaxies.
If we look at the times when the last major mergers of MW/M31-like galaxies occurred (time of coalescence, top left panel of Fig. 13), we see that 83 galaxies (42 per cent) experienced their

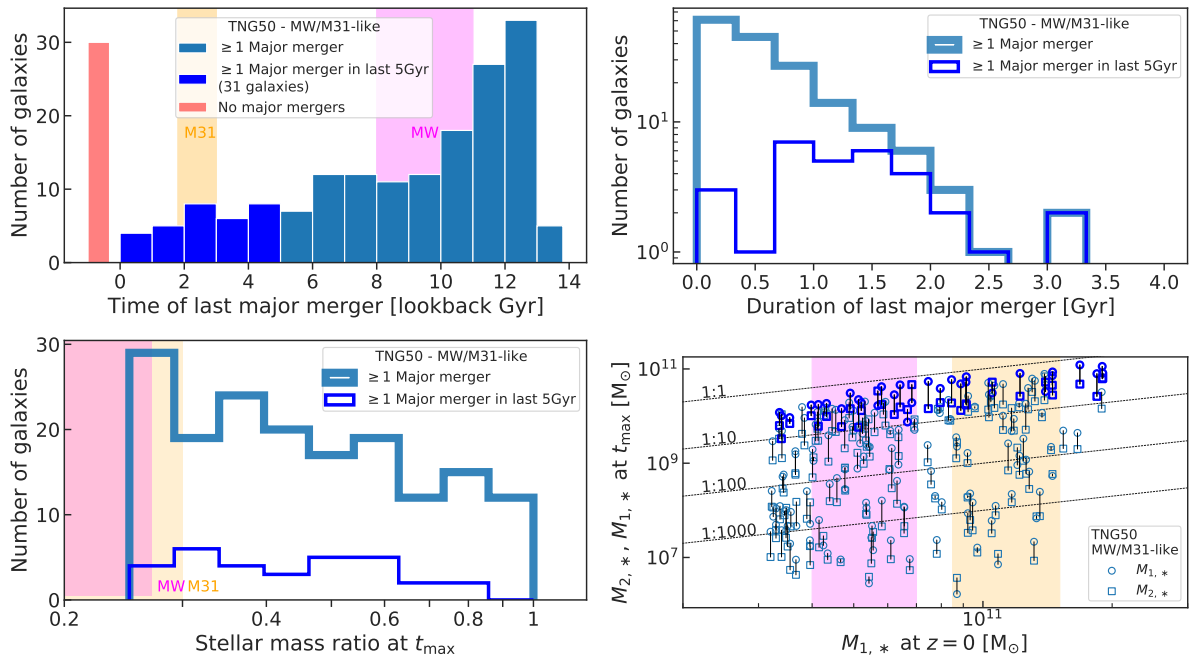

Figure 13: Characteristics of the last major mergers of MW/M31-like galaxies in TNG50 at $z = 0$: major mergers. In all panels, MW/M31-like galaxies that experienced their final major merger in the last 5 Gyr are depicted in blue; MW/M31 analogues with their last major merger at any cosmic time are indicated in light blue. *Top left*: Time of the last major merger (in lookback Gyr) for each of the 198 galaxies, in bins of 1 Gyr: 30 MW/M31-like galaxies did not undergo any major mergers (orange bin); 95 galaxies (approximately 50 per cent) underwent their last major merger more than 9 Gyr ago ($z \gtrsim 1.3$); 31 galaxies have had a late major merger, i.e. as recent as in the last 5 Gyr. The occurrence of the latter is distributed in an approximately uniform way in the considered period of time. *Top right*: Duration of the major mergers, defined as the time elapsed since the secondary reached its maximum in stellar mass (t_{max}) and the moment of coalescence – bins span 330 million years. 77 per cent of the major mergers lasting longer than 1.5 Gyr are recent. *Bottom left*: Stellar-mass ratios between the secondary and the primary, at t_{max}. *Bottom right*: Stellar mass of both progenitors, at t_{max}, versus the stellar mass of the galaxy at $z = 0$. For the Galaxy and Andromeda, we show observational estimates with magenta and orange annotations, respectively: the mass-ratio and the time of the known last major merger are taken for the case of GES (Helmi et al., 2018, Gallart et al., 2019, Naidu et al., 2021) and for M32 (D'Souza & Bell, 2018).

last major merger more than 10 Gyr ago, whereas 95 galaxies (48 per cent) had their last major merger more than 9 Gyr ago – this compares to ∼35 per cent for $M_* \sim 10^{10}\, M_{\rm stars}$ disk galaxies according to Font et al. (2017). Importantly, the MW/M31-like galaxies with a recent major merger are not particularly biased in their merger times and span the entire final 5 billion years of cosmic evolution. Magenta and orange shaded vertical bands denote the current estimates of the last major mergers of the Galaxy, namely of GSE (Helmi et al., 2018, Myeong et al., 2018, Chaplin et al., 2020), and of Andromeda (D'Souza & Bell, 2018), respectively.
In the top right panel, we see that the durations of the last major mergers (see Section 4.2.2) span from ∼ 0.04 to ∼ 3.5 Gyr. The longest mergers are more frequently recent, owing to the shorter dynamical times when the Universe was younger: 77 per cent of the mergers lasting longer than 1.5 Gyr occurred within the last 5 billion years. On average, TNG50 MW/M31-like galaxies experienced major mergers that carried on for 0.7 Gyr (median of 0.5 Gyr), this duration increasing to 1.4 Gyr for those happening since $z \lesssim 0.5$ (median of 1.3 Gyr). Dividing the TNG50 MW/M31-like sample according to stellar mass, and considering only the galaxies that underwent at least one major merger over their history, the last major merger of MW-mass (M31-mass) galaxies occurred on average about 10 (9) billion years ago. The median merger duration is approximately 0.61 and 0.46 Gyr for the two sub-samples.
The stellar mass ratios of the merging galaxies are quantified in the bottom panels of Fig. 13, evaluated at $t_{\rm max}$ (the time of the maximum stellar mass of the secondary, see Section 4.2.2). All ratios in the bottom left panel are above 0.25, by our definition of a major merger. The mass ratios closer to 1 are more frequent for older mergers (see our definitions of merger mass ratios in Section 4.2.2). Only one of the 31 recent major mergers has a mass ratio larger than 0.75. Consistently with the hierarchical model of galaxy assembly, the most recent major mergers have in general lower stellar-mass ratios: galaxies at recent times are more massive but massive galaxies that could merge with them are rarer. The absolute values of the galaxy stellar mass of the merging progenitors are shown in the bottom right panel of Fig. 13. MW/M31-like galaxies with recent major mergers (blue symbols) occupy the highest part of the range, at fixed stellar mass, and are compared to the progenitors of all MW/M31-like galaxies with the last major merger happening at any time (light blue symbols): the former subset of galaxies increases their mass by a small factor in the last few Gyrs, so we do not find galaxies with a recent major merger where both progenitors have low stellar mass and the descendant grows then rapidly enough to be included in our MW/M31-like mass cut. Magenta and orange vertical shaded areas denote the current stellar mass constraints of the Galaxy and Andromeda, respectively: according to TNG50, even galaxies with lower, MW-like mass may experience a recent major merger and still be disky at $z = 0$.
The figures above show that, according to TNG50, MW/M31-like galaxies with recent major mergers have interacted with relatively massive companions for substantial amounts of times, i.e. on average for ∼ 1.4 Gyr with secondaries of $M_* \approx 2 \times 10^{10} \mathrm{M}_\odot$, resulting in mergers of median stellar mass ratio of 0.41 (i.e. 1:2.5).

4.4.2 Gas availability during the mergers

But how is it possible that, despite having undergone major mergers as recently as in the last 5 billion years, galaxies can nevertheless exhibit marked disky stellar morphologies at $z = 0$?
Previous studies (Hopkins et al., 2009, Stewart et al., 2009, Hoffman et al., 2010) had shown with idealized simulations of mergers that the gas content in the merging progenitors is a determinant

factor for the outcome of a merger – see Introduction: it both conditions how destructive a merger is for the stellar component of the involved galaxies and how actively stars can form in the descendant galaxy. In Fig. 14, we hence analyze the gas availability during merger events for all the TNG50 MW/M31 analogues undergoing major mergers across cosmic epochs. Firstly, we estimate the gas content of each galaxy by accounting for all the gravitationally-bound gas at the time of the merger: this choice is supported e.g. by Sparre et al. (2022), who show with simulations that the gas in the circumgalactic medium and even in the outer halo can contribute to the star formation after major merger events between $z = 0.3 - 0.8$ that produce disky, MW-like galaxies. Secondly, we compare the gas mass fractions of each system at coalescence (i.e. the gas mass over the total stellar mass) with that of central galaxies that, at the corresponding major-merger time of TNG50 MW/M31-like galaxies, have the same stellar mass (within a range of ± 0.1 dex) as the galaxy that has resulted from the two progenitors. These galaxies that serve as reference are plotted in gray, with the median gas fractions marked with a gray solid line and the narrower and the broader shaded regions representing, respectively, the 25th-75th and the 5th-95th percentiles. MW/M31 analogues are marked as blue symbols and a median is added to facilitate comparison.
Two facts are noticeable. First, as anticipated (see also e.g. Pillepich et al. 2019 for TNG50 results) and as it is expected from observations of high-redshift galaxies, the average mass fraction of gas in galaxies decreases with cosmic time – for the control sample, the gas mass fraction drops by $\sim$1 dex from $z \sim 3$ (about 2 Gyr after the Big Bang) to $z \sim 0.5$ (about 5 Gyr ago). Generally, the MW/M31 analogues show a gas fraction, on average, comparable to that of the control sample of central galaxies in the similar mass range. Second, and importantly, the difference between samples increases slowly with cosmic time: we think that this is the very manifestation of the fact that galaxies that are selected to be disky at $z = 0$ while having undergone a recent major merger constitute somewhat a biased subset. Namely, MW/M31-like galaxies with recent major mergers exhibit somewhat higher gas fractions ($\sim 0.1 - 0.2$ dex) than the average of other non-merging galaxies with similar mass at the corresponding epoch: 23 of 31 of them lie above the median value of the control sample. Whereas the gas-mass difference is not large, it appears sufficient to trigger star formation in TNG50 galaxies during the merger events, as we show next.

4.4.3 Star formation bursts triggered by gas-rich major mergers

The availability of gas enables star formation during the recent major mergers of TNG50 MW/M31-like galaxies.
We have examined the SFR evolution of the TNG50 MW/M31 analogues with recent major mergers and find that their SFRs can increase substantially in correspondence to their major merger events. Such SF bursts may occur at the time of coalescence or shortly after, and sustained SF may be in place in the descendants through $z = 0$. In fact, SF bursts may happen even before the merging galaxies coalesce, specifically at the close pericentric passages of the secondary progenitor in its approaching orbit. About 24 of the 31 MW/M31 analogues with recent major mergers show appreciable bursts of SF triggered by the merger, seven of which are the depicted in Fig. 15.
There we show the evolution of the instantenous SFR along the main progenitor branch of selected galaxies (magenta) as a function of cosmic time (starting at 8 Gyr after the Big Bang), compared with their distance (in physical kpc) to the secondary progenitor involved in their

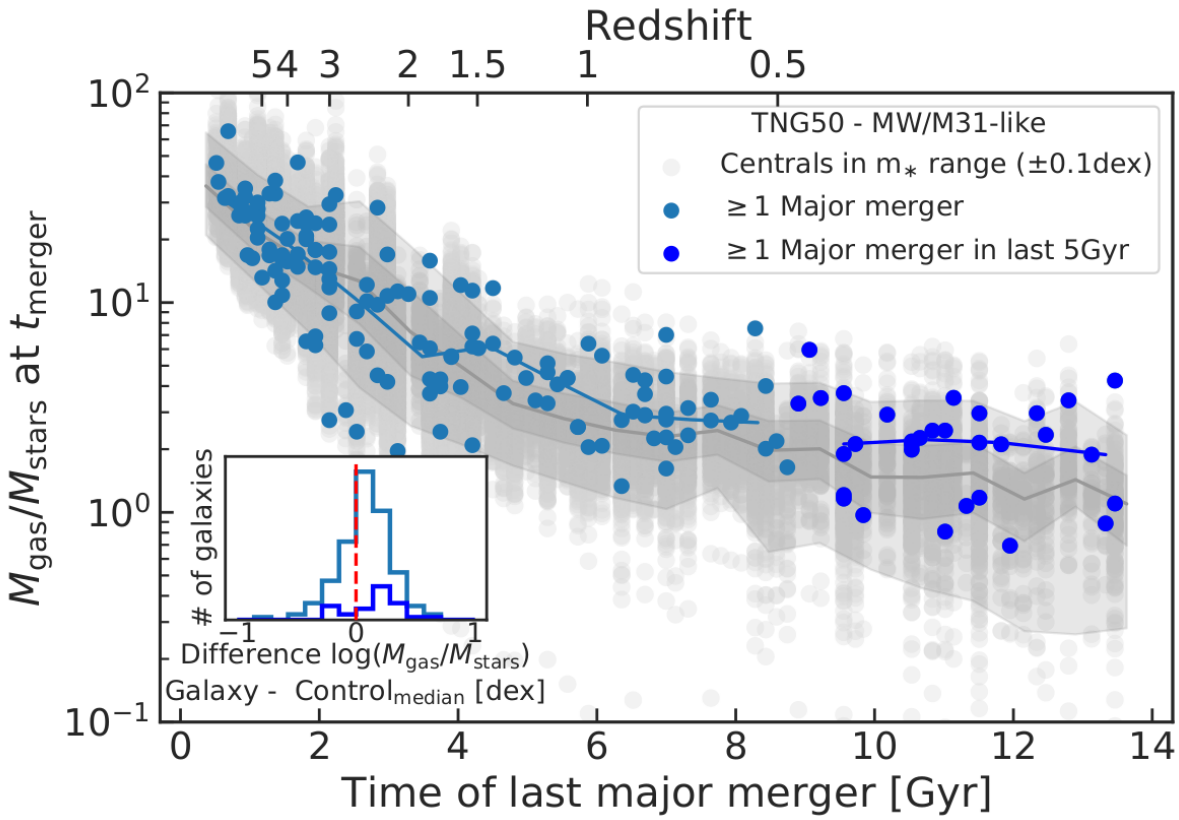

Figure 14: Gas mass fraction (M_{gas}/M_{stars}) of TNG50 MW/M31-like galaxies at the time of their last major merger (time of coalescence). For comparison, in gray, we show the gas fractions of central galaxies in the same stellar mass range (±0.1 dex) at the corresponding merger times: the median is the solid curve. MW/M31-like galaxies with recent major mergers (blue symbols) have lower gas-to-star mass fractions, in absolute terms, than those with more ancient mergers – i.e. the gas mass fraction of galaxies decreases with time. However, they are more gas rich (0.1 – 0.2 dex) than the average central galaxies in TNG50. This gas richness can also be appreciated in the inset histogram, where we show the logarithmic difference of the gas mass fraction, for each of the galaxies, with respect to the median of the control sample in the considered time of the last major merger.

last major merger: green curves. For each galaxy, i.e. row, the time of merger is represented with a vertical black solid line, whereas the last pericentric passages are distinguishable as the minima in the distance curves: dotted black vertical lines. It can be clearly seen that, not only coalescence, but also close galaxy-galaxy interactions prior to the merger time can trigger important events of star formation in the main galaxy and consequently can plausibly alter its structure, morphology, and stellar mass content.
It was previously shown that highly "bursty" SF was suppressed in galaxy major mergers of the original Illustris simulation due to its insufficient numerical resolution: in particular, this was demonstrated via zoom-in simulations of selected merger events with 40 times better mass resolution (Sparre & Springel, 2016). Fig. 15 shows that the numerical resolution of TNG50 is sufficient to capture the compression of gas possibly due to the galaxy-galaxy interactions, the funnelling of gas towards the galaxy centers, higher gas density and hence to reproduce bursts of star formation triggered by mergers and galaxy interactions.

4.4.4 The cases of disks destroyed during the mergers and reformed vs. those surviving during the merger

By inspecting the time evolution of the SFR of the main progenitors of each MW/M31 analogue with recent major mergers and the time evolution and distributions of the orbital properties of their stars (as in Figs. 17 and 18), we find that two main scenarios or pathways are followed by the 31 TNG50 galaxies to be disky at $z = 0$ after undergoing a recent major merger. These are as follows:

1. in 18 cases (58 per cent), the galaxy's stellar disk, being in place prior to the time of the merger or prior to the pericentric passages, is destroyed by the merger; nevertheless, the galaxy by $z = 0$ reforms a stellar disk;

2. in 11 cases (35 per cent), the stellar disk is not completely destroyed by the merger and therefore the galaxy retains its disky morphology down to $z = 0$;

Two remaining galaxies of the sample are difficult to classify. In one case (TNG50 Subhalo ID 419618 at $z = 0$), the D/T mass ratio is actually quite low both prior to and after the merger as well as at $z = 0$: $\sim 0.1 - 0.2$ – we noticed this galaxy in Fig. 32 and possibly this should not have been included in the TNG50 MW/M31-like sample in the first place. In the case of TNG50 Subhalo ID 400973 at $z = 0$, the D/T mass ratio around the time of the merger is biased low by the fact that (possibly another) merger event has triggered the production of new stars in counter-rotating orbits, making the study of its evolution rather specific. In both cases, the galaxies have experienced a major merger as recently as in the last one or two snaphots, i.e. in the last $150 - 300$ million years, so much so that stellar shells are somewhat visible in their stellar light maps.
We have also explicitly checked whether cases exist whereby the recent major merger occurs at a time when the progenitor of the selected MW/M31-like galaxies does not exhibit yet a clear disky stellar morphology (or not anymore due to a previous merger). But for the two galaxies described above, among the 31 TNG50 MW/M31 analogues, this does not occur.
The frequency of major merger events in MW/M31 analogues and the different evolutionary pathways that the TNG50 simulation has allowed us to uncover are summarized in Fig. 16. In

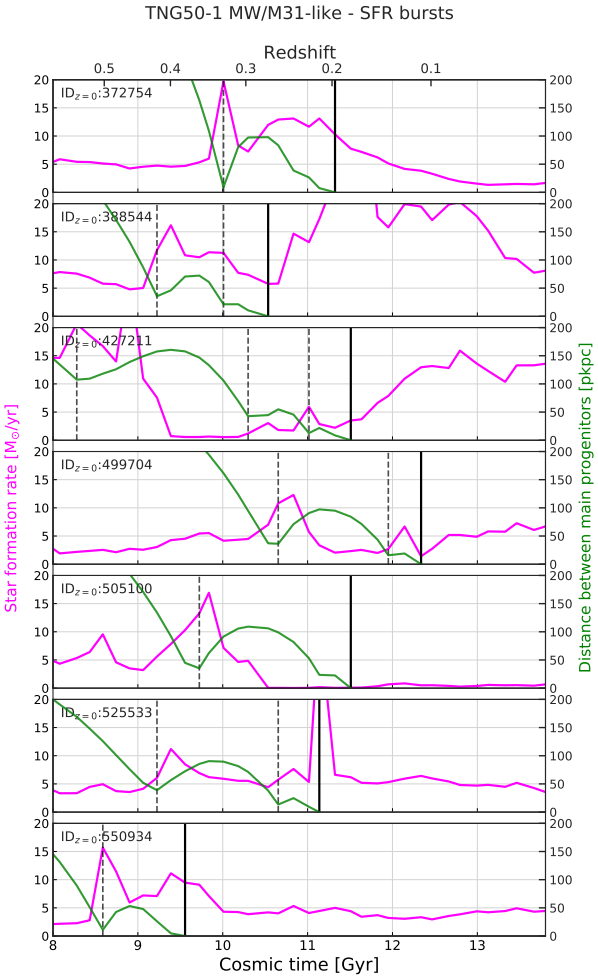

Figure 15: Bursts of star formation in a selection of MW/M31-like galaxies in TNG50 with recent major mergers. SFR in the main progenitor (magenta) and distance between both progenitors (green) are plotted for seven example galaxies, one panel each. The time of the last major merger is marked with a solid vertical line; the last pericentric passages of the secondary progenitor are marked with dotted vertical lines. The galaxy mergers and the close galaxy-galaxy interactions prior to coalescence can trigger substantial bursts of star formation. The SFR after the merger varies depending on the galaxy.

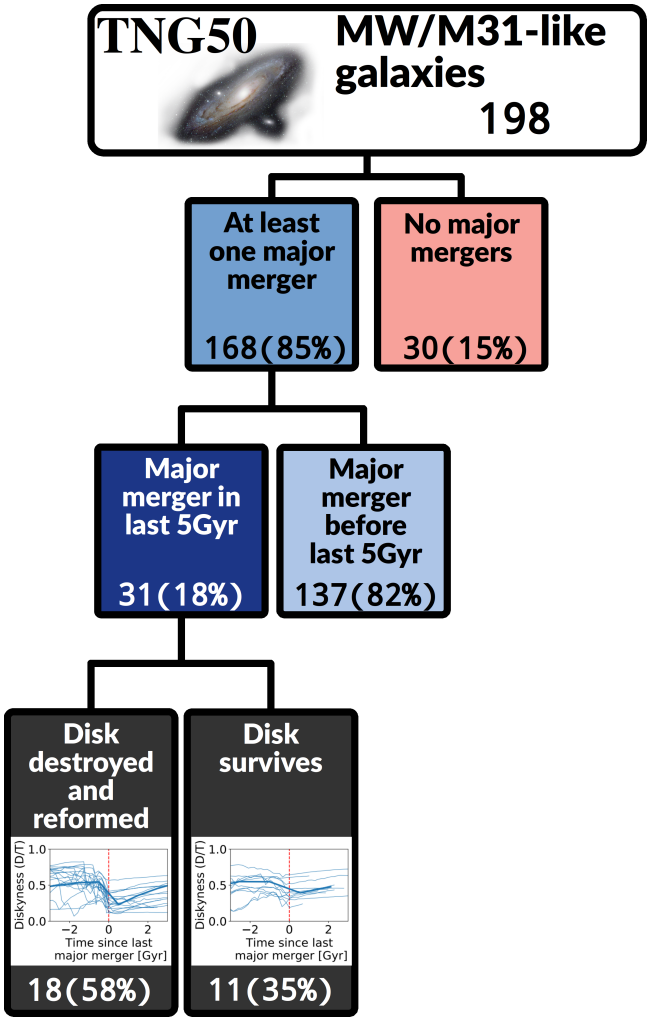

Figure 16: Summary of the evolutionary pathways uncovered via the TNG50 simulation for a sample of galaxies selected at $z = 0$ as MW/M31 analogues. MW/M31-like galaxies may have had at least one major merger throughout their history or not at all: 85 and 15 per cent, respectively – second panels from the top). Among the former, third layer from the top, 31 galaxies, i.e. 18 per cent of the TNG50 MW/M31-like sample with at least one major merger ever, actually underwent one or more major mergers as recently as in the last 5 billion years. For most of them (58 per cent), the previously-existing stellar disk was destroyed but reformed; for another substantial fraction (35 per cent), the previously-existing stellar disk survived through the merger event. Inset diagrams show the evolution of D/T with time, centered at the time of the merger. Thin (thick) curves represent individual galaxies (medians). For the destroyed disk the drop of diskyness is much more noticeable. Two of 198 galaxies cannot be placed in either scenario – see text for details. In each box, we report the number of galaxies in absolute and relative values, being the latter calculated within the downstream sub-samples.

the various boxes from top to bottom we show the number of TNG50 MW/M31-like galaxies undergoing certain conditions and paths.

Somewhat similar scenarios had been found in previous studies, albeit without a specific focus, and hence selection, on the MW/M31 mass scale. For example, Jackson et al. (2020) used the Horizon-AGN cosmological simulation and reported about two main pathways for the existence of massive galaxies with disky stellar morphologies at $z = 0$ ($M_* \geq 10^{11.4}\,M_{\rm stars}$, above our upper end): surviving disks (30 per cent), owing to anomalously quiet merger histories, and spheroidal galaxies that form a young stellar disk (70 per cent) as a consequence of star formation triggered by a recent gas-rich mergers. According to Font et al. (2017) and the GIMIC simulation, on the other hand, stellar disks reform for one third of the $M_* \sim 10^{10}\,M_{\rm stars}$ galaxies undergoing a major mergers in the last 7 Gyr. Also, Sparre & Springel (2017) showed that four $M_* \sim 10^{10.5-11}\,M_{\rm stars}$ disk galaxies that were disky before their last major merger (at $z \sim 0.5$) were able to reform the disk after the merger destroyed it.

Figs. 17 and 18 exemplify the time evolution of one prototypical example galaxy for each of the two scenarios identified for TNG50 MW/M31-like galaxies. Plotted as a function of cosmic time, from top to bottom, are (more details in the caption): stellar and gas mass for the primary and secondary progenitors, SFR of the main progenitor, distance between progenitors, diskyness (D/T ratio) and the distributions of the orbit circularities of the stars formed in situ. Finally, images of the stellar mass density of the main progenitor are shown from face-on and edge-on projections, at the three snapshots immediately prior to the last major merger, at the snapshot after the merger and at $z = 0$ – which helped in the selection of the galaxies in the MW/M31 sample. In each figure, the vertical solid, thick line represents the time of the last major merger, i.e. the time of coalescence. The blue thin vertical lines in the top panels denote the time when the secondary reaches its maximum stellar mass and the stellar mass ratio is evaluated.

In the first pathway, i.e. Fig. 17, a disk of young stars grows in the period of a few billion years between the last major merger and $z = 0$. In these cases, the amount of gas available to form new stars is a key factor. From the Figures, it is also clear that, as already alluded to in Fig. 15, not only does the major merger affect and alter the galaxies, but also the close pericentric passages of the secondary progenitor can have a large impact on the final outcome of the merger and on star formation. In fact, but for periods around coalescence (Fig. 17) and possibly at pericentric passages, new stars are born in circular orbits, i.e. with circularity at birth typically close to unity – this is the case at all times and for all galaxies (Pillepich et al., 2019). On the other hand, the current circularity of the in-situ stars that are in $z = 0$ MW/M31-like galaxies may be very different, i.e. hotter, than that at birth, particularly for stars that formed at early times: top vs. third circularity panels.

For the galaxy of Fig. 17, which is an example of a last major merger that destroys the stellar disk of the main progenitor, the merger produces a drop in the stellar circularities of the galaxy's stars (top and bottom circularity panels). However, stars formed after the merger are born mostly in circular orbits, so that a new stellar disk is present at $z = 0$. This sequence is also visible in the stellar density images. Fig. 18 represents a case of a disky main progenitor whose stellar morphology is not affected by the major merger: compared to the previous case, the secondary progenitor approaches the main progenitor more progressively, with multiple pericentric passages prior to the final merger.

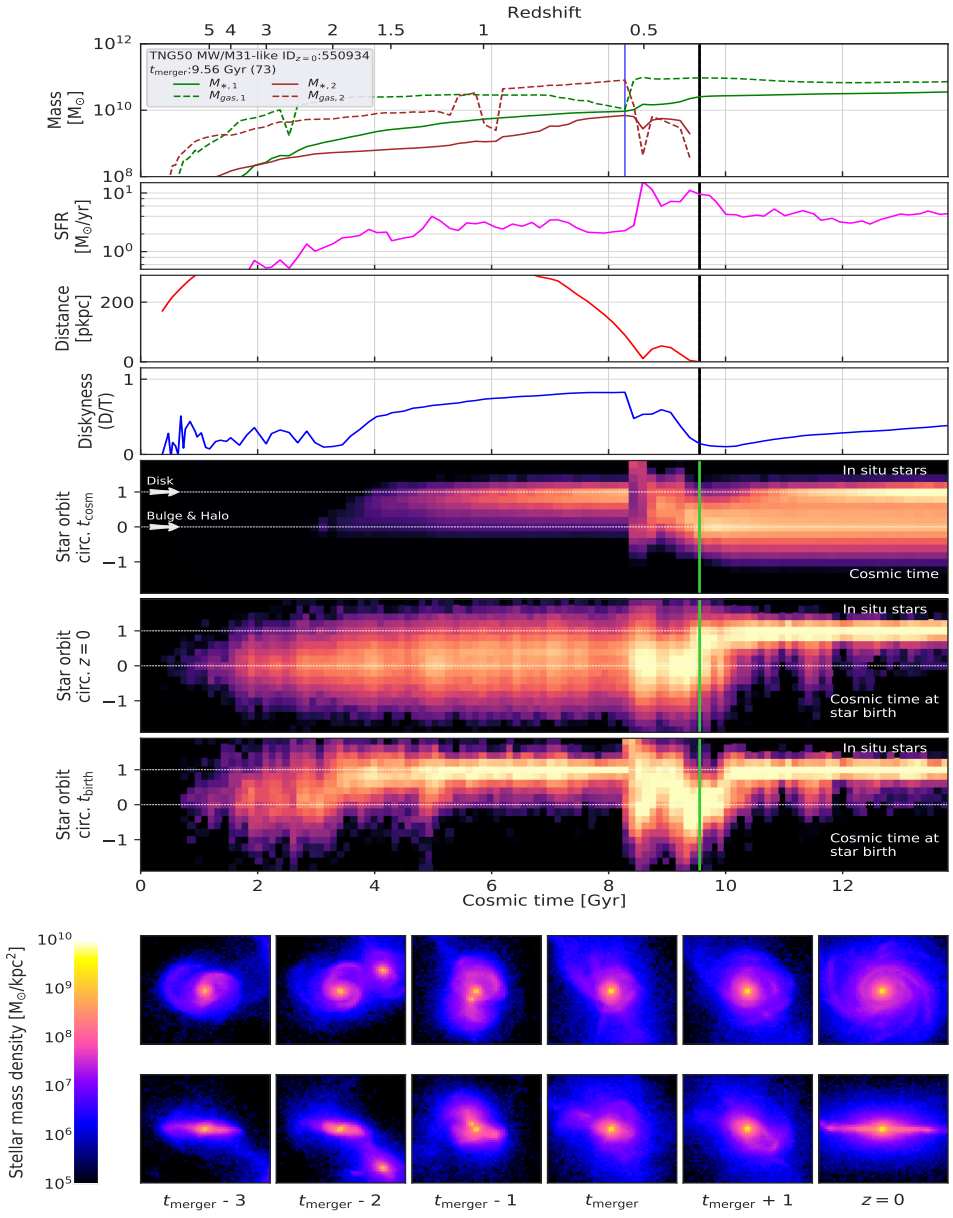

Figure 17: Evolutionary tracks of an example TNG50 MW/M31-like galaxy whose last major merger destroys the stellar disk, but a new one reforms. In all panels, but the images at the bottom, values are shown as a function of cosmic time. The vertical thick solid line represents the time of the last major merger; the thin vertical line in the top panel denotes the time when the stellar mass ratio is evaluated. *Top panel:* Time evolution of the gravitationaly-bound stellar and gas mass (solid and dashed line, respectively) for the primary and secondary progenitors. *Second to fourth panels:* SFR of the main progenitor (fuchsia), distance between progenitors (red), and diskyness i.e. D/T mass ratio (blue – see definition in Section 4.2.3). *Fifth to seventh panels:* Distributions of the orbit circularities of stellar particles. From top to bottom: circularities of the in-situ stars along the main-progenitor branch of the galaxy; in-situ circularities evaluated at $z = 0$ and shown as a function of the birth time of the stars; in-situ circularity evaluated at the time of birth and shown as as a function of birth time. *Bottom panels:* Stellar density images of the main progenitors, centered at the main galaxy, face-on and an edge-on views, across snapshots separated by about 150 Myrs.

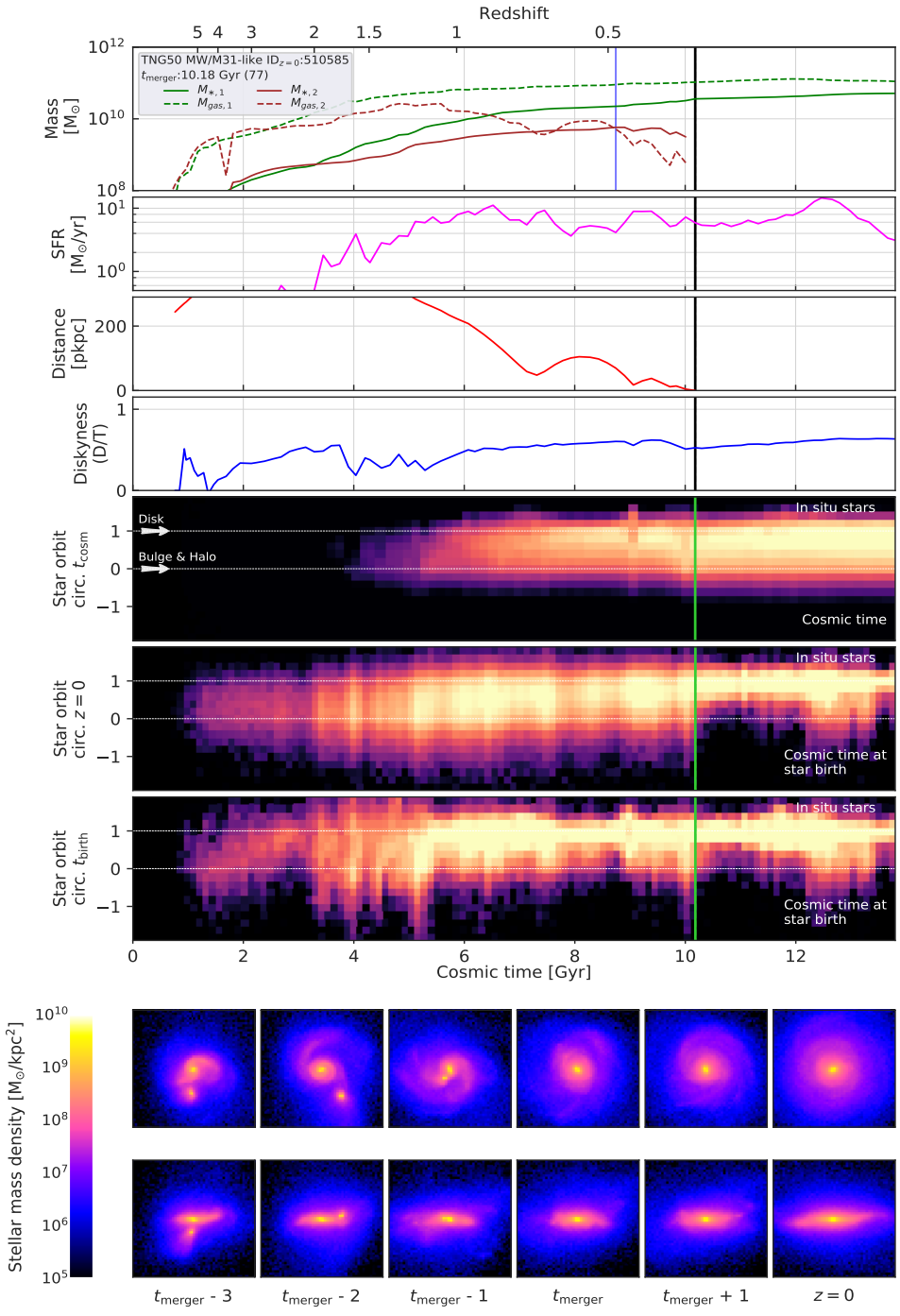

Figure 18: Evolutionary tracks of an example TNG50 MW/M31-like galaxy whose last major merger does not destroy the stellar disk. Panels as in Fig. 17.

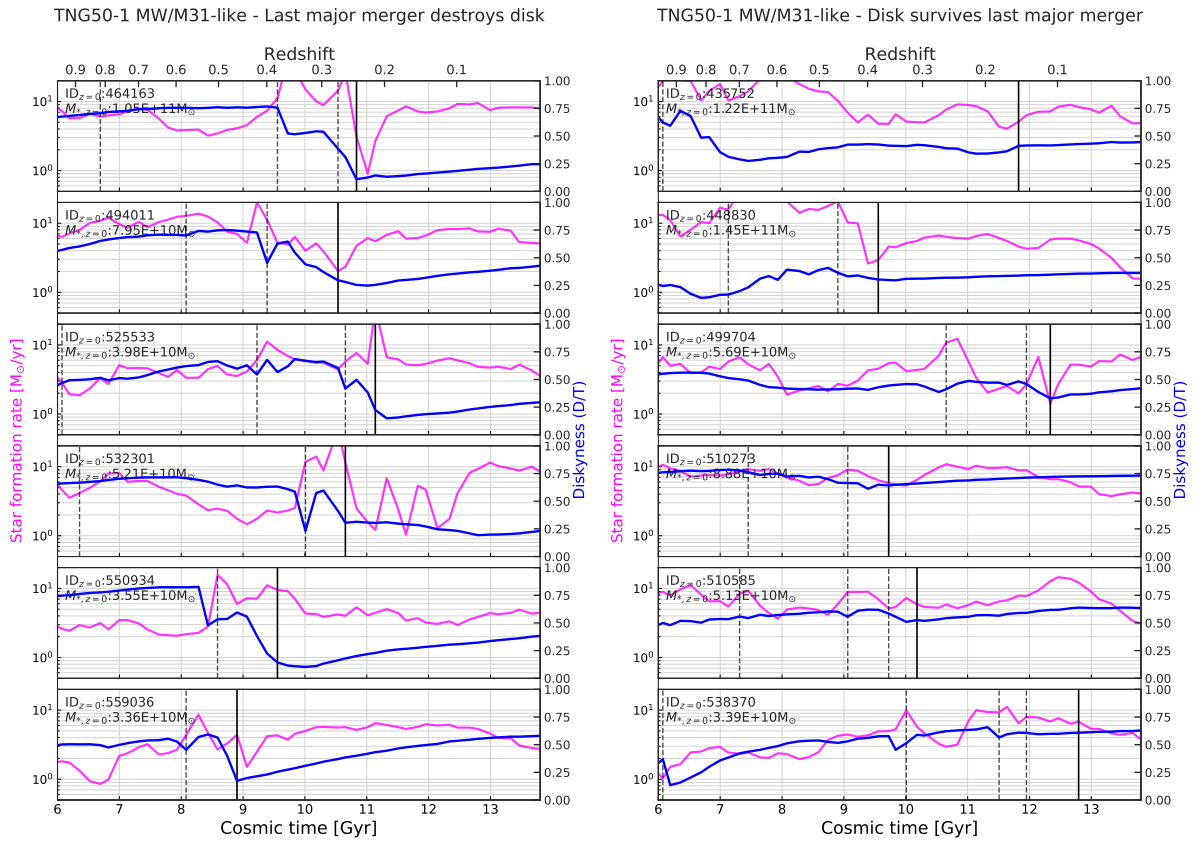

Figure 19: Connection between ongoing star formation and galaxy stellar morphology, i.e. diskyness, for TNG50 MW/M31-like galaxies with recent major mergers. We show the evolution along the main-progenitor branch of individual galaxies of their SFR (fuchsia) and of their D/T ratio (blue), as a function of cosmic time. We show random examples of galaxies whose stellar disk had been destroyed by their last major merger and regenerated (left column: 6 examples of 18 cases) and of galaxies whose stellar disk has not been destroyed (right column: 6 examples of 11 cases). The time of the last major merger is marked with a black solid vertical line; the times of the last pericentric passages are marked with black vertical dashed lines. The destruction of the stellar disks in the left column is identifiable in those cases where a galaxy undergoes a sudden drop in diskyness, and usually this occurs in the period between the last pericentric passage and the time of the merger. Compared to the destroyed disks, the galaxies on the right column exhibit milder and gentler drops in diskyness. In all cases, sustained star formation is in place during and after the merger, often with bursts at pericentric passages (see also Fig. 15) and is temporally coincident with a steady D/T increase after the merger and towards $z = 0$, particularly so in the galaxies where the merger destroyed the disk.

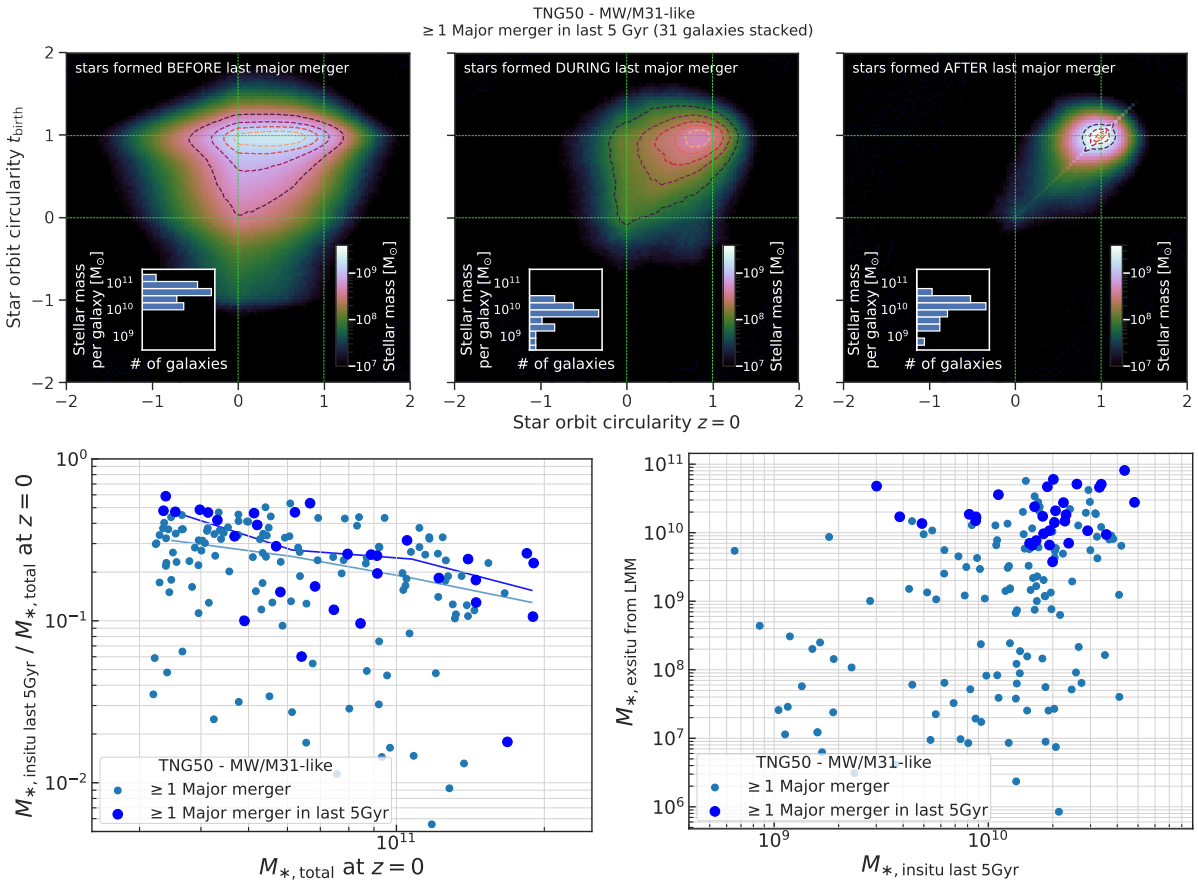

Figure 20: Connection between star formation, orbital circularities of the stars and the amount of in-situ star formation triggered by the merger and of the ex-situ stars brought in by the mergers. *Top*: circularities of the stellar orbits at the time of formation vs. at $z = 0$, stacked for all 31 TNG50 MW/M31-like galaxies with a recent major merger, for stars formed before (before $t_{\rm max}$), during (between $t_{\rm max}$ and $t_{\rm merger}$), and after the last major merger, from left to right. In all panels, stars in perfectly circular orbits have a circularity = 1 (or −1 if they are counter-rotating); stars in radial orbits have a circularity of ∼ 0. Histograms show the total stellar mass per galaxy born in the respective periods of time. *Bottom left*: Fraction of in-situ stellar mass formed in the last 5 Gyr as a function of total stellar mass at $z = 0$. *Bottom right*: Ex-situ stellar mass brought during the last major merger vs. in situ stellar mass in the last 5 Gyr. In both cases, TNG50 MW/M31-like galaxies with a recent major merger (blue) are contrasted with those with at least one major merger across their history but irrespective of when (light blue). Most of the stars were born in circular orbits, this is generally the case both before and after the times of the last major merger and, to a lesser extent, during the last major merger; however, the orbits of the stars born after the last recent major merger remain mostly unaltered, i.e. in circular orbits, all the way to $z = 0$. Furthermore, recent major mergers also seem to trigger more in-situ star formation in the resulting galaxy than in galaxies with more ancient mergers, in addition to bringing large amounts of ex-situ stellar mass.

4.4.5 Connection between in-situ star formation and diskyness during and after the mergers, and on the accreted mass

As pointed out above, within the TNG50 model, the orbits of stars at the time of formation are generally circular – at least at $z \lesssim 2-3$, see 7th panels from the top of Figs. 17 and 18. Namely, if star formation occurs, stars naturally form in circular orbits, because they originate from gas that is in rotationally-supported and disky configurations (see also findings by Pillepich et al., 2019). Here we further expand on the connection between star formation and galaxy stellar morphology.

In Fig. 19, we examine several examples of TNG50 MW/M31-like galaxies with recent major mergers and for each we show the evolution in time of the instantaneous star-formation rate (fuchsia) and of diskyness D/T (blue) along the main progenitor branch. In these plots the time of the merger is marked with a vertical solid black line and the moments of the last pericentric passages with vertical dashed black lines. We show 6 random examples among the cases whose last major merger destroyed the disk but the latter reformed (left panels) and 6 random examples among those galaxies whose last major merger did not destroy the disk.

First, in the left panels, a noticeable feature is the decrease of the diskyness of the galaxy around the time of the merger or some million years before the merger: in the latter cases, this is due to the effects of the proximity of the secondary progenitor at the pericentric passages. The drop of diskyness is roughly at least one third of the D/T value before any interaction with the secondary. This is not the case for the galaxies in the right panel, where the changes in circularity fraction, D/T, are less pronounced.

Second, the evolutionary tracks of Fig. 19 confirm that, after the merger, a disk of young stars can form again: this can be appreciated in the fact that the D/T ratio increases progressively without new drops (the galaxies can undergo additional minor or mini mergers, but no major ones). A progressive increase in the diskyness is here correlated with sustained phases of star formation.

The connection between ongoing and recent star formation and stellar morphology is quantified and extended to the entire galaxy sample in Fig. 20. In the top panels, we compare the circularities of the stellar orbits at birth and at $z = 0$, for the stars formed before, during, and after the last major merger: namely, before $t_{\rm max}$, between $t_{\rm max}$ and $t_{\rm merger}$, and after $t_{\rm merger}$, i.e. after coalescence, respectively. Results are shown by stacking the orbital properties of all the stars in all the TNG50 MW/M31 analogues with a recent major merger – therefore the panel on the top left (right) depicts the properties of stars that are mostly older (younger) than 5 billion years – but see the distribution of the merger times in Fig. 13. Inset histograms show the total stellar mass per galaxy that is formed in-situ in the corresponding periods.

The stars formed before the last major merger (left panel) mostly formed in perfectly circular orbits (circularity ~ 1). On the other hand, the circularity of their orbits changed as time passed, all the way down to $z = 0$, where they span the whole range of values, with circularity $\sim 0-1$, i.e. with also non rotationally-supported orbits. This effect is called orbital heating and is manifested in radial and vertical directions. In fact, at $z = 0$, stars can also exhibit negative circularity values, which corresponds to counter-rotating orbits. The heating quantified in the top left panel of Fig. 20 is thought to be due to secular evolution and to more violent orbital changes induced by mergers. In this plot we cannot identify the exact drivers of the orbital alterations nor to assess the timescales when these old stars were heated up to their $z = 0$ levels. Yet, we notice that a tail of the stars in the left panel of Fig. 20, and mostly formed before the last major

merger, exhibit circularity values at birth also approaching 0, i.e. random motions: these are at least partially stars that are older than 8 – 10 billion years – and this is qualitatively consistent with the picture suggested by observations and also reproduced by simulations whereby galaxies were dynamically hotter at earlier epochs ($z \gtrsim 1.5$, Pillepich et al., 2019, and references therein). Also stars formed *during* the major merger events and at pericentric passages (top, middle panel of Fig. 20) show broad ranges of circularities at birth, with their circularity distribution at $z = 0$ being even more spread out.

Focusing on the two top right panels of Fig. 20, it is manifest (see also inset histograms) that the stars formed after and during the last major merger are less numerous than the stars formed before it – even just because they could form over shorter periods of time. Yet they are sufficient to secure a disky stellar morphology to the $z = 0$ descendants. Stars formed after the merger (top right) were also born in nearly circular orbits, but, in this case, their circularities barely change during the period between the last major merger and $z = 0$: whereas the latter is comparatively short for disk heating to have a large effect, it should be noticed that for a fraction of the stars in this panel the period elapsed after the last major merger can be as long as ~ 5 Gyr. Yet, the phenomenology of the top right panel of Fig. 20 quantifies the idea that those MW/M31-like galaxies that experienced a recent major merger are disky at $z = 0$ because of the sustained star formation and because the newly-formed stars are born in circular orbits and have not yet undergone heating.

These results qualitatively agree with the general picture described by Peschken et al. (2020) for $z = 0$ disk galaxies in the Illustris simulation whose last major merger occurred at $z \lesssim 1.5$: according to that analysis, the stars born before the last major merger form the $z = 0$ spheroidal components and the stars born after the merger constitute a new formed disk.

But how much stellar mass is produced during and after the last recent major mergers? In addition to the insets in the top of Fig. 20, in the bottom left panels we give the fraction of in-situ stellar mass formed in the last 5 Gyr over the total stellar mass (bottom left) and the amount of ex-situ stellar mass brought by the last major merger vs. the in-situ stellar mass in the last 5 Gyr (bottom right). Galaxies with recent major mergers (blue circles) have formed typically a larger fraction of in-situ stellar mass in the last few billion years than their MW/M31 analogues with more ancient last major mergers – this ranges in the median, depending on final mass, between 15 – 30 per cent of the total $z = 0$ mass and the recently formed in-situ stellar masses are, on average, 0.1 – 0.2 dex larger in the recently-merged population. This indicates that a larger amount of in-situ star formation has indeed occurred *because* of the recent major merger, and would have not occurred at the same levels in the absence of a recent major merger. At the same time, recent major mergers bring large amounts of ex-situ stars (bottom right panel), $\gtrsim$ a few $10^9 - 10^{11}\ M_{\rm stars}$. This is somewhat necessary, because recent major mergers are more massive in terms of absolute stellar-mass than ancient ones.

To summarize, recent major mergers bring both larger amounts of ex-situ stellar mass as well as trigger relatively more in-situ star formation in the last few billion years than in the absence of a major merger. This nicely connects to the average stellar mass growth with time of WM/M31-like galaxies – see Fig. 10, left, and compare black thick vs. blue thick curves (i.e. all TNG50 MW/M31-like galaxies vs. those with a major merger in the last 5 billion years): the latter exhibit a suppressed stellar mass growth at $z \gtrsim 0.5$ than the typical MW/M31-like galaxy, and de facto 'managed' to enter in the stellar mass selection thanks to the mass boost (both in-situ and ex-situ) provided by the recent major merger.

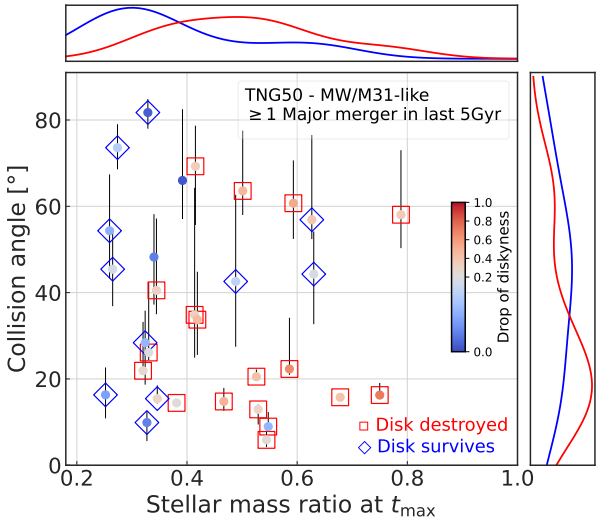

Figure 21: Collision angle of the merger orbit vs. mass ratio at t_{max}, color coded by the absolute drop in diskyness, D/T, during the merger event, for TNG50 MW/M31-like galaxies that experienced a major merger in the last 5 billion years. The points denote the median of the angles across the 5 to 15 snapshots previous to the merger, whereas the errorbars are the 25th-75th percentiles of such angle distribution). The symbols denote whether the last major merger destroyed (squares) or not (diamonds) the previously-existing stellar disk, based on the inspection of the time evolution of the D/T ratios along the main-progenitor branch of each galaxy (see Figs. 17 and 18). The density plots are calculated with a gaussian kernel. Mergers that do not destroy the pre-existing stellar disks tend to populate the parameter space of smaller stellar-mass ratios and larger collisional angles, i.e. spiralling or large impact-parameter orbits (although the stellar mass ratio has a larger impact).

4.4.6 Orbits of the merging galaxies

To understand what determines the disruption or survival of a stellar disk during the merger, we inspect the orbital properties the galaxies follow in their merging process: these can be very diverse. We characterize the orbits of the mergers according to two angles: 1) the angle between the plane of the orbit and the stellar angular momentum of the main progenitor ("orbit plane angle"), and 2) the (acute) angle between the velocity vector of the secondary progenitor with respect to the main galaxy and the position vector between them ("collision angle", as presented in Zeng et al. 2021). These angles are measured at individual snapshots and averaged throughout the 5 last snapshots prior to the merger time (approximately over 800 Myr, to minimize the effect of fluctuations that are sometimes caused by the SUBFIND halo-finder algorithm). The orbit plane angle allows us to determine whether the secondary galaxy orbit is prograde (angle: 0-90°) or retrograde (90-180°) with respect to the rotation of the main galaxy. The collision angle determines whether the merging orbit resembles a smooth spiraling (high values) or a head-on collision (low values).
We find (although do not show) that, throughout all the recent major merger events that involve the progenitors of the TNG50 MW/M31-like galaxies, the majority of the mergers that unfold in retrograde orbits (13 from 31 mergers) end up destroying the stellar disks. Also, within our 31 cases, a disk that survives a major merger is more likely to be the result of a prograde than a retrograde orbit. However, these results are tentative and warrant further investigation.
Instead, we can show that it is the combination of orientation of the collision and of stellar mass ratio of the merger that is a good predictor of whether stellar disks are destroyed or not during a recent major merger. This is shown in Fig. 21, where galaxies with recent major merger are color-coded by the drop (in absolute value) of the D/T mass ratio because of the major merger and the latter is characterized by the collision angle and the stellar mass ratio – all of them are major, but the ratios go from 1:4 to 1:1. Galaxies whose stellar disks are destroyed by the major merger and then reform (red squares) are more frequent towards larger stellar mass ratios and lower collision angles, i.e. towards mergers with small impact parameters. On the other hand, the lower the stellar mass ratio, the higher the probability that the disk survives the merger event (see blue diamonds, in general with smaller drops in diskyness)[3]. To separately quantify the effects of the collision angle and of the stellar mass ratio, we perform Anderson-Darling (AD) and Kolmogorov-Smirnov (KS) tests with the null-hypothesis that the two galaxy samples (destroyed and surviving disks) follow the same distribution. For the angles, we cannot reject the null-hypothesis, namely we cannot exclude that the two distributions are in fact indistinguishable (14 (KS) and 18 (AD) per cent significance levels). On the other hand, for the stellar mass ratios, we can reject the null-hypothesis (as their equality is given with very low significance levels, i.e. 1 (KS) and 0.7 (AD) per cent). We conclude that the stellar mass ratios have more significant effects than the collision angle in the survival of stellar disks.

[3]The reader may notice a galaxy classified as "disk destroyed" but with a very low drop in diskyness. This is Subhalo ID 372754 (top left stamp of Fig. 32): in this case, the evolution of the total D/T does not fully capture what can instead be evinced by inspecting the full distribution of the stellar orbits at the time of birth vs. $z = 0$, as in the corresponding 5th and 7th panels of Figs. 17 and 18 for this galaxy.

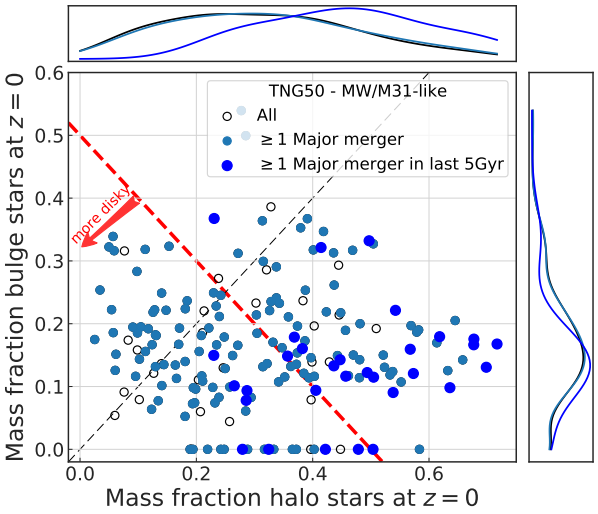

Figure 22: Mass fraction in kinematically-defined bulges vs. mass fraction in kinematically-defined stellar haloes, for TNG50 MW/M31-like galaxies at $z = 0$. We compare three subsamples: all MW/M31 analogues (white circles), analogues that experienced at least one major merger throughout their history (light blue), and analogues with recent major mergers (deep blue). Galaxies below the dashed red line have a total stellar mass fraction in both spheroidal components smaller than 0.5 and are therefore diskier. The black dashed line represents an equal fraction of stars in the bulge and the halo. The density plots are calculated with a gaussian kernel. For most TNG50 MW/M31 analogues the stellar halo is more massive than the bulge. Analogues that underwent recent major mergers have average-mass bulges but more massive stellar haloes, compared to the rest of the MW/M31-like galaxies.

4.5 The properties of MW/M31 analogues with recent major mergers

Are the $z = 0$ properties of recently-merged MW/M31-like galaxies different from the other galaxies in the sample, or from the rest of the galaxies in TNG50 in the same mass range? In the following we inspect a selection of global and stellar structural properties of the TNG50 MW/M31-like galaxies and contrast those with recent major mergers to the whole sample. We also juxtapose the measurements from TNG50 galaxies to analogous constraints from observations of the Galaxy and Andromeda: these are indicated in the following plots as magenta and orange areas and bands, respectively, and encompass the estimates available from the literature (see Table A1 in Pillepich et al in preparation), including systematic differences and measurement errorbars.

4.5.1 Average bulges but more massive and shallower stellar haloes

The origin of stellar bulges in disky galaxies has been debated for at least two decades, with recent observational and theoretical results pointing towards a) a distinction between photometrically- vs. kinematically-defined bulges (e.g. Du et al., 2020, and references therein) and b) an origin of galactic bulges that is not necessarily always connected to mergers, with different pathways for so-called classical and pseudo-bulges (see e.g. **?**, and Gargiulo et al. submitted, for TNG50-based results and references therein).

Here we use the kinematic decomposition proposed by Du et al. (2019), based on a Gaussian mixture separation of the stellar particles in the kinematic phase-space, to separate the stellar structural components of TNG50 galaxies in kinematically-defined stellar bulges and stellar haloes. In Fig. 22, we account for all stellar particles that are gravitationally bound and we give the stellar mass fractions in such bulges and stellar haloes, by denoting TNG50 MW/M31 analogues that underwent recent major mergers with blue circles and others with more ancient or no major mergers with light blue and white circles. Galaxies that fall below the red dashed line have a total star fraction in spheroidal components lower than 0.5 and are therefore diskier (106 over 198 MW/M31 analogues, i.e. 54 per cent).

For the entire sample of TNG50 MW/M31-like galaxies the bulge fraction is always (except for one galaxy) below 0.5, and below 0.25 in most cases (∼85 per cent). The mass fractions in stellar haloes cover a wide range, with values also greater than 0.5 (∼ 12 per cent) and reaching values of ∼0.7. When we consider both spheroidal components, there is a non-negligible fraction of TNG50 MW/M31-like galaxies (44/198, ∼22 per cent) where the bulge is more massive than the stellar halo. But, TNG50 returns also bulgeless galaxies (25 galaxies, ∼13 per cent): among these, there are galaxies with major mergers (including recent ones) and without. These latter numbers seem to reasonably agree with previous observational studies of edge-on galaxies (16 per cent found by Kautsch et al., 2006, albeit differently selected). MW-mass galaxies have lower stellar halo mass fractions than M31-mass analogues (median values of 0.27 vs. 0.37), whereas their bulge mass fractions are similar (median of ∼0.15).

For the galaxies with recent majors mergers, the average bulge fraction is comparable to the average values for the rest of the MW/M31 analogues – in disagreement with e.g. the predictions of Puech et al. (2012) of more bulge dominant galaxies with $z \sim 0.6$ major mergers. A possible interpretative scenario was offered by Hopkins et al. (2010), who pointed to gas richness as a crucial element in limiting bulge growth. However, their stellar halo mass fraction is on average larger, with median halo mass fractions of 0.46 for MW/M31 analogues with recent mergers vs. 0.31 for the complete sample. Again we perform AD and KS to test the null-hypothesis that the two samples (all MW/M31-like galaxies and galaxies having a recent major merger) follow the same distribution, separately for halo and bulge mass fractions. For the halo fraction, the galaxies with recent major mergers have a statistically-distinct stellar halo mass fraction distribution from that of all MW/M31-like galaxies (as the null-hypothesis of equality is given with 0.04 (KS) and 0.1 (AD) per cent significance levels). For the bulge ratios, with 6.9 (KS) and 14.6 (AD) per cent significance levels, we cannot reject the possibility that the two distributions are in fact identical.

The stellar haloes of recently-merged MW/M31-like galaxies are not only more massive, but their stars are also less centrally concentrated, albeit to a somewhat weaker degree. Fig. 23 gives the 3D slopes of the spherically-symmetric stellar profilesfitted with a power law ($\rho(r) = \rho_0 r^{\alpha}$) between twice the stellar half-mass radius and r_{200c} of TNG50 MW/M31-like galaxies as a function of the time of their last major merger (top) and of the stellar mass of the secondary galaxy involved in the merger (bottom). Error bars cover one standard deviation errors of the parameter provided by the non-linear least squares fitting routine (python curve_fit). Galaxies with recent major mergers have slightly shallower slopes, with median ([25th-75th] percentiles) of about $-4.6\,(-[4.3, 5.1])$ vs. $-4.8\,(-[4.4, 5.4])$ for the blue and light blue sub samples, respectively. However, this difference is only mildly significant (the significance levels are of 17 per cent and 22 per cent for the KS and AS tests of consistent distributions). This is a distinction

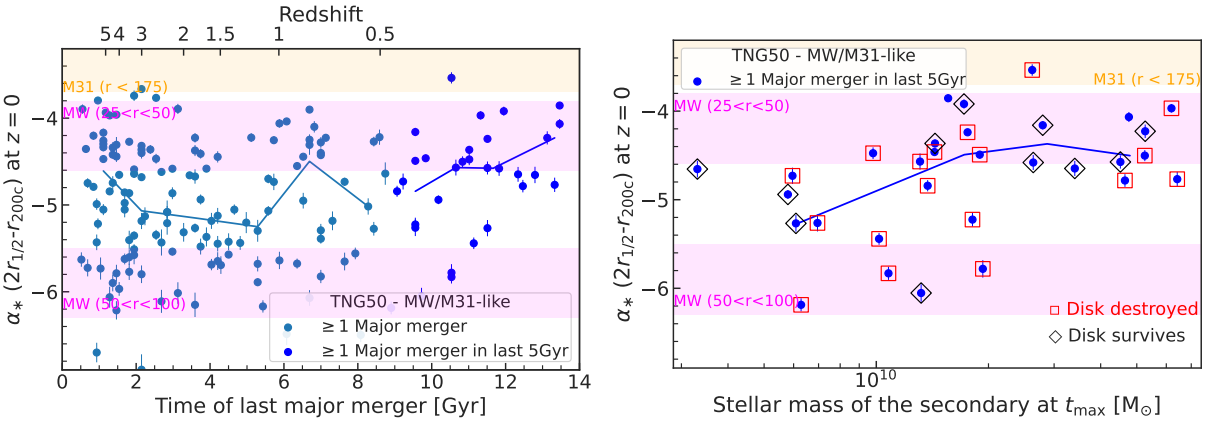

Figure 23: 3D slopes of the stellar mass density profiles at large galactocentric distances, i.e. in the stellar haloes, in TNG50 MW/M31-like galaxies. In the top, we show the stellar halo 3D slopes as a function of the time of the merger, in the bottom as a function of the stellar mass of the secondary progenitor. In both panels, MW/M31-like galaxies with recent major mergers are denoted by blue circles and compared in the top to the rest of the sample (light blue). Error bars are one standard deviation errors of the parameter provided by the non-linear least squares fitting routine (python curve_fit). Magenta and orange shaded area denote current observational constraints in the slopes of the stellar haloes of the MW and M31. In the bottom, we additionally indicate with red squares and black diamonds those recently-merged galaxies whose stellar disk had been destroyed or not by the merger. The stellar haloes of galaxies with recent major mergers have shallower profiles, and among these, the ones whose disk survived the merger but that merged with more massive secondaries are, in general, shallower. We show observational estimates for the MW: for a radius between 25 and 50 kpc, Bell et al. (2008), Watkins et al. (2009), Sesar et al. (2011), Deason et al. (2011), and for a radius between 50 and 100 kpc Deason et al. (2014); for M31: Gilbert et al. (2012), Ibata et al. (2014).

that is qualitatively in line but weaker than the findings based on Illustris galaxies by Pillepich et al. (2014). Similarly as found there, TNG50 also predicts the 3D stellar halo slope to depend on galaxy stellar mass: among the MW/M31 analogues, MW-mass galaxies have steeper stellar haloes (-5.0 ± 0.7) than M31-mass ones (-4.5 ± 0.7), a fact that is qualitatively consistent with the estimates of the stellar halo profiles of the Galaxy and Andromeda – see magenta and shaded areas. Furthermore, the more massive the merging companion (within the major merger mass ratios), the shallower the stellar haloes of the descendant galaxy.

4.5.2 Larger fractions and amounts of ex-situ stellar mass

As in the hierarchical growth of structure scenario stellar haloes of MW-like galaxies are mostly formed by accretion (Zolotov et al., 2011, Font et al., 2011, Pillepich et al., 2015), it is not surprising then that MW/M31-like galaxies with recent major mergers also exhibit larger amounts and fractions of ex-situ, i.e. accreted, stars at $z = 0$.
This is quantified in Fig. 24, in terms of the distributions of the ex-situ stellar mass fraction across different galaxy samples (top) and the total ex-situ stellar mass as a function of galaxy stellar mass at $z = 0$ (bottom). In both panels, we compare all TNG50 galaxies in the depicted

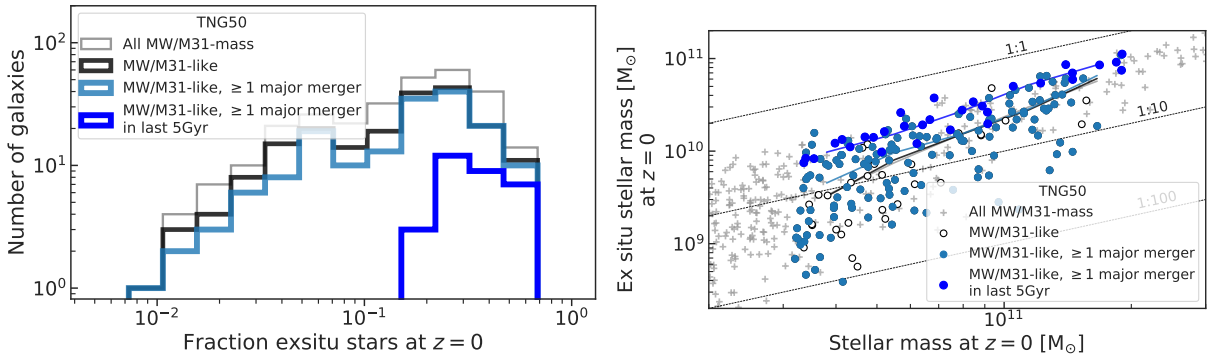

Figure 24: Ex-situ, i.e. accreted, stellar mass and stellar mass fractions of TNG50 MW/M31-like galaxies at $z = 0$. In the top, we show the distributions of the ex-situ fractions, in the bottom the accreted stellar mass vs. galaxy stellar mass, for different sub-samples. Solid curves in the bottom panel denote medians at fixed stellar mass. Galaxies with recent major mergers have on average larger values of ex-situ stellar mass (absolute and relative). At the opposite end, TNG50 MW/M31-like galaxies that never experienced a major merger (white circles) exhibit lower-than average ex-situ stellar masses, given their galaxy mass.

mass range (gray annotations) to TNG50 MW/M31-like galaxies (black and light blue lines in the top and white and light blue circles in the bottom) and MW/M31-like galaxies that underwent recent major mergers (blue lines and circles).
The ex-situ fractions range widely: from a few percents to ~ 65 per cent, with a median value of ~ 0.19 for TNG50 MW/M31-like galaxies. This is in the ball park of the findings by Rodriguez-Gomez et al. (2016) and Pillepich et al. (2018b), with Illustris and TNG100/TNG300 MW-mass galaxies and haloes, respectively. For the sub-sample with a recent major merger, the ex-situ fraction is biased towards the higher end of the distribution, with a median ex-situ mass fraction throughout the galaxy bodies of ~ 33 per cent. Moreover, at a fixed stellar mass, galaxies with a recent major merger have more ex-situ stellar mass than the average, by $0.3 - 0.5$ dex. For the TNG50 MW/M31 analogues with recent major mergers, the total ex-situ mass is always larger than $10^{9.8-10}\,M_{\rm stars}$. Again, the recent major mergers not only trigger star formation (see Figures 19 and 20) but also bring in higher-than average amounts of accreted stars.

4.5.3 Somewhat thicker and hotter stellar disks

Whether $z = 0$ galaxies with thick stellar disks have been born with large stellar velocity dispersions and less flat stellar mass distributions and/or whether their stellar disks puffed up across cosmic epochs, because of secular or externally-triggered processes, is an active field of research (see e.g. Pillepich et al., 2019, Park et al., 2019, for recent works with state-of-the-art cosmological galaxy simulations). Here we do not address this general issue but show that a recent major merger may indeed affect the vertical structure of the stellar disk of a specific selection of galaxies, those with $z = 0$ global properties similar to the Galaxy and Andromeda. In Fig. 25, we plot the stellar disk heights (top two rows) of all TNG50 MW/M31-like galaxies (light blue circles) and of TNG50 MW/M31-like galaxies that underwent a recent major merger (blue circles). These are shown as a function of stellar disk length (top panels) and of the time of the last major merger (middle panels) for two reasons: a) to control for the influence of galaxy

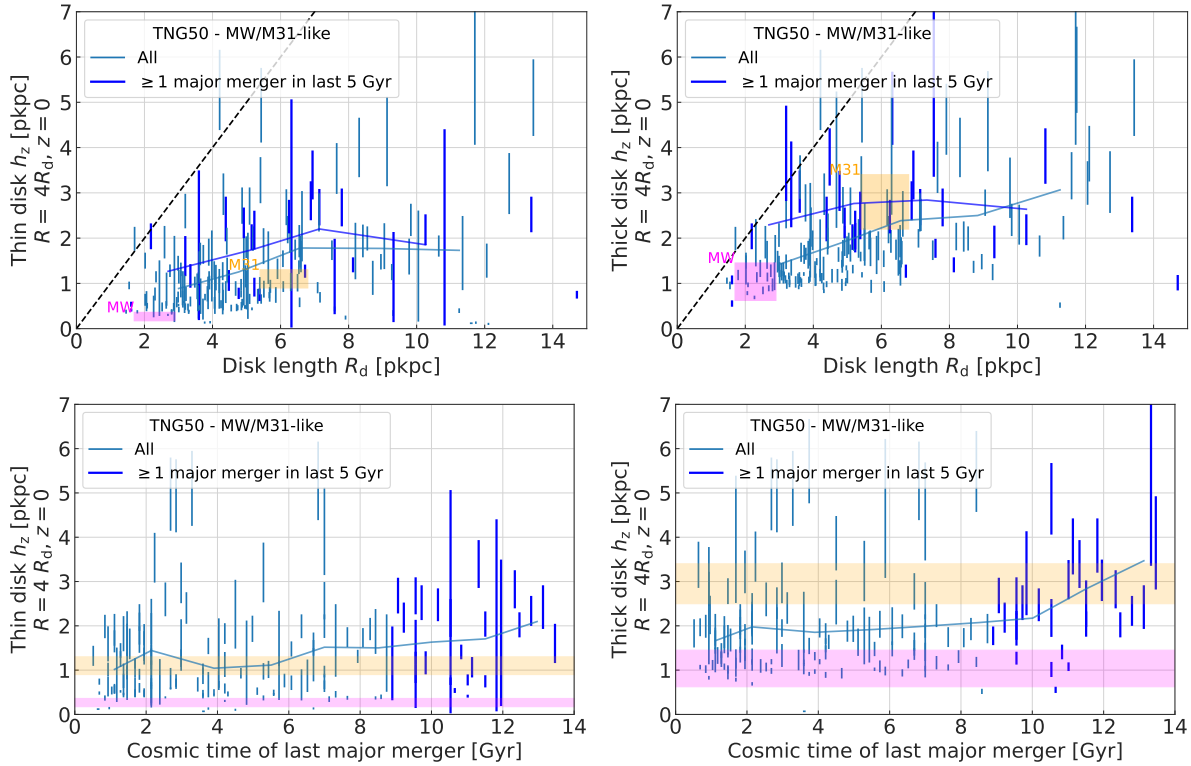

Figure 25: Stellar disk heights of TNG50 MW/M31-like galaxies at $z = 0$. We show thin (left) and thick (right) vertical heights vs. disk length (top row) and time of the last major merger (bottom row). In all panels, vertical colored bars span, for each galaxy, the best-fit values for the scale height obtained from two complementary functional forms: double sech and double sech^2: light blue vertical bars represent TNG50 MW/M31-like galaxies, blue vertical bars denote those that underwent a recent major merger. Solid thick curves are medians in bins of the quantity on the x-axes of the mean between the best-fit values from the two functional forms. At fixed disk length, galaxies with a recent major merger have somewhat thicker thin and thick stellar disks. Observational estimates for the Galaxy and Andromeda are given as magenta and orange boxes, which encompass, and hence marginalize over, diverse measurements including their errorbars: 1.7 – 2.9 kpc, 175 – 360 pc, and 625 – 1450 pc for the disk length, geometrically thin- and thick-disk heights of the MW (Gould et al., 1996, Ojha, 2001, Siegel et al., 2002, Jurić et al., 2008, Rix & Bovy, 2013, Bland-Hawthorn & Gerhard, 2016) and 4.8 – 6.8 kpc, 900 – 1300 pc, and 2200 – 3400 pc for the disk length, geometrically thin- and thick-disk heights of M31 (Worthey et al., 2005, Barmby et al., 2006, Hammer et al., 2007, Collins et al., 2011).

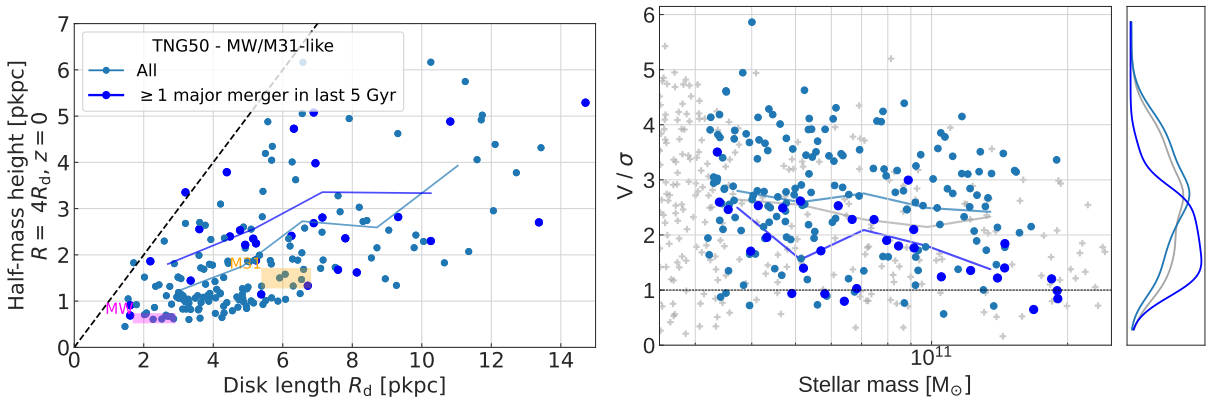

Figure 26: Stellar non-parametric disk heights and stellar vertical kinematics of TNG50 MW/M31-like galaxies at $z = 0$. In the top panel, we measure disk heights in a non-parametric way, namely by measuring the stellar half-mass disk heights at $3.5-4.5$ times the disk length, for comparison to the heights in Fig. 25. Observational constraints on the Galaxy and Andromeda (magenta and yellow areas) are derived from the average relationships in TNG50 between half-mass heights and thin and thick disk heights. In the bottom panel, we show the ratio between the maximum rotational velocity and the vertical velocity dispersion of the stars. Light blue circles represent TNG50 MW/M31-like galaxies, blue circles denote those that underwent a recent major merger, and gray crosses represent the rest of the galaxies in TNG50 in the depicted stellar mass range. Solid curves are medians in bins of the quantity on the x-axes. At fixed stellar mass, galaxies with a recent major merger exhibit thicker stellar disks and lower V/σ ratios, i.e. hotter orbits.

stellar mass and hence disk size; and b) to judge the effects in terms of how long ago a major merger may have perturbed a possibly pre-existing stellar disk. For the disk lengths, for each galaxy projected face on, we fit an exponential profile to the radial distribution of the disk stellar mass surface density in a certain aperture (see Appendix 4.7 for details). For the disk heights, for each galaxy projected edge-on, we select only stars in circular orbits in the cylindrical shell between 3.5 and 4.5 times the disk length and fit a double parametric formula to the vertical stellar mass density profile, to allow for both geometrical thin and thick disks. More details about the fitting technique are given in the Appendix 4.7: for each galaxy, we both fit a double hyperbolic secant and a double squared hyperbolic secant profile and represent in the figure the range spanned by both best-fit values. The averages in bins of disk length or time of last major merger (solid thick curves, top and bottom, respectively) run over the mean of the best-fit values from the two functional fitting functions.

According to TNG50, more extended galaxies exhibit thicker stellar disks, particularly for the geometrically thick component. Yet, the galaxy-to-galaxy variation is very large, and in TNG50 we recover example galaxies whose stellar disk structural properties are very similar to both the Galaxy's (magenta shaded areas) and Andromeda's (orange).

From the top panels of Fig. 25, we uncover that galaxies with a recent major merger have, on average, somewhat larger stellar disk heights, at fixed disk length, but this effect is more pronounced for the thick (right) rather than the thin (left) component and for smaller galaxies. From the bottom panels, we also see an increase in disk thickness that mildly correlates with the time of the last major merger, with thicker disks the more recent the last major merger. Although these relationships have significant scatter, they are in place despite the fact that more recent major mergers impact galactic stellar structures that are expected to be thinner and colder (because they exist at lower redshift) than those impacted by merger at higher redshifts, at least at fixed stellar mass (Pillepich et al., 2019). According to AD and KS tests, we find that, for the geometrically-thin component, we cannot reject the possibility that the distributions of the heights of all MW/M31-like galaxies and of those with recent major mergers are indistinguishable (15 (AD) and 20 (KS) per cent significance levels for the null-hypothesis of equality in the case of the double sech^2; 3 (AD) and 2 (KS) for the double sech). On the other hand, for the geometrically-thick component, the distributions of disk heights are clearly different (0.7 (AD) and 0.2 (KS) per cent significance levels for the double sech^2; 0.3 (AD) and 0.03 (KS) for the double sech).

A similar general picture is in place also when the thickness of the stellar disk is evaluated in a non-parametric way. In Fig. 26, top, we show the stellar half-mass heights of TNG50 MW/M31-like galaxies, measured in the annuli between 3.5 and 4.5 times the disk length: TNG50 MW/M31-like galaxies with recent major mergers have larger stellar half-mass heights than the whole population – and the stellar half-mass height correlates better with the thick-disk height of galaxies than with the thin-disk height. Moreover, when turning to quantify the *kinematics* of the stellar disks (Fig. 26, bottom), we find that the ratio between the maximum rotational velocity and the vertical velocity dispersion of the stars (as in Pillepich et al., 2019) is lower for MW/M31 analogues with recent major mergers, namely, they are kinematically hotter than the rest of the sample, which concurs at least qualitatively with the structural analysis.

We conclude that MW/M31-like galaxies that have undergone recent major mergers exhibit somewhat thicker and clearly hotter stellar disks, but with a significant scatter. TNG50 does return example galaxies that experienced a recent major merger and have thin and thick disk

heights similar to those of Andromeda (orange areas). Interestingly, however, for galaxies with disk length smaller than 3 – 4 kpc and hence smaller than Andromeda (see magenta vs. orange observational constraints), TNG50 seems to imply that it is improbable, but not impossible, for a galaxy to have emerged from a major merger over the last few billion year and exhibit at $z = 0$ thin-disk heights as low as that of the Galaxy. This is qualitatively in line with the ideas of Toth & Ostriker (1992), Wyse (2001), Hammer et al. (2007), who however excluded the possibility. In fact, such statements strongly depend on how and where the vertical structure of the stellar disks is assessed. For example, in terms of thick-disk or stellar half-mass heights, a few TNG50 galaxies appear as compact and similarly thick as the Galaxy. Moreover, there are two TNG50 galaxies that underwent a recent major merger and whose thin disks has a similar disk length and height as the Galaxy's (Subhalo IDs 528322 and 532301). Finally, when we measure disk heights of TNG50 galaxies at a fixed galactocentric distance of e.g. 8 kpc (irrespective of disk size), we do find a few galaxies that underwent recent major mergers and have a stellar disk as thin as $\lesssim 350 - 400$ pc, i.e. similar to the MWs.
These last numbers and considerations indicate that the limited numerical resolution of TNG50 should not affect our scientific conclusions, in that simulated stellar disks can be thinner than the simulation's softening length (288 pc for DM and stellar particles, ≥72 pc for the gas cells – see Section 6.2). An extended study of the resolution effects on galaxy sizes and heights in TNG50 is presented in Pillepich et al. (2019), where it was shown that, in terms of stellar half-mass heights, stellar disk thickness can be considered to be converged in TNG50 to better than 20–40 per cent.

4.5.4 Hints of more massive SMBHs, larger gas reservoirs and star formation rates

We conclude this analysis by enunciating (without showing) the differences for a few additional global properties at $z = 0$ of TNG50 MW/M31-like galaxies with and without a recent major merger.
Whereas the average stellar-to-halo-mass relation of TNG50 MW/M31-like galaxies with recent major mergers is comparable to the entire sample of MW/M31 analogues, the former seem to be somewhat biased towards larger SMBH masses, at fixed stellar mass. Two effects can explain this: the SMBHs present in the two progenitors were on average larger, even just because they merged at more recent epochs than the others, or the SMBH of the resulting galaxy has accreted more gas and grown more quickly in the post-merger epoch than the SMBHs in the control-sample galaxies. As the effect is not particularly large, we do not investigate further on this here and we refer the interested reader to Pillepich et al. in preparation for a discussion on the masses of TNG50 SMBHs vs. that of Sgr A* in the Galaxy.
On the other hand, at fixed galaxy stellar mass, the gas mass fraction and SFR (within $2R_{1/2,*}$ at $z = 0$) are higher for the galaxies with recent major mergers. This is overall consistent with the picture described above, whereby the recent major merger brought in gas, which in turn cooled down, becoming eligible for star formation. Therefore, not only the gas fractions were already larger at the time of coalescence (Fig. 14), but we also find that MW/M31-like galaxies with recent major mergers are biased high in terms of gas availability with respect to the rest of MW/M31 analogues still at $z = 0$.

4.6 Summary and conclusions

We have used the cosmological magneto-hydrodynamical galaxy simulation TNG50 to quantify the mass assembly and merger histories of 198 MW/M31-like galaxies selected at $z = 0$. We have placed a special focus in studying how frequently, and how, MW/M31-like galaxies may undergo a recent major merger (e.g. occurring within the last 5 billion years and with stellar mass ratio > 1:4) and still exhibit a disky stellar morphology at $z = 0$.

The main findings that emerge from our analysis are:

- The progenitors of TNG50 MW/M31-like galaxies at $z \sim 3$ were, on average, ~40 times less massive in stellar mass than today, but with a large past scatter: $\gtrsim 2$ dex between the 10th-90th percentiles (7 times less massive in total halo mass, with scatter 0.6 dex; Fig. 10).

- Major mergers are common: 168 of the 198 MW/M31-like galaxies in TNG50 (85 per cent) have undergone at least one major merger throughout their history (or specifically, at $z \lesssim 5$; Fig. 11).

- TNG50 returns galaxies with stellar mass compatible with the Galaxy and Andromeda and with overall disky stellar morphology at $z = 0$ (Fig. 32) even in the cases when these have undergone a recent major merger: 31 MW/M31-like galaxies in TNG50 (16 per cent) have experienced at least one major merger in the last 5 Gyr (Fig. 13).

- Galaxies with recent major mergers have interacted with relatively massive companions for significant amounts of times, i.e. on average for ~1.4 billion years. The companions, i.e. secondary progenitors, are massive objects, with median stellar mass of $\sim 2 \times 10^{10}\, M_{\rm stars}$ (i.e. average stellar mass ratio of 1:2.5; Fig. 13).

- According to TNG50, there are two main pathways that can lead to a disky MW/M31-mass galaxy at $z = 0$ after a recent major merger: i) the pre-existing stellar disk is destroyed during the interactions and merger with the companion, but reforms (Fig. 17); and ii) somewhat less frequently, the latter does not disrupt the pre-exisisting stellar disk (Fig. 18). Whether the one or the other occurs depends, for example, on the merger configuration, with mergers with larger stellar mass ratios and smaller impact parameters following more frequently the first scenario (Fig. 21).

- In both cases, gas was sufficiently available to either trigger star-formation bursts at pericentric passages or at coalescence or both, as well as to sustain prolonged star formation until $z = 0$, with the ensuing (re)formation of a disk of young(er) stars (Fig. 14).

By comparing the $z = 0$ structural and global properties of TNG50 MW/M31 analogues that underwent a recent major merger with those with more ancient last major mergers, we find that the former have, on average:

- larger amounts of in-situ stellar mass produced over the last few billion years, namely, as a consequence of the recent merger (Fig. 20);

- similarly massive kinematically-defined bulges (Fig. 22);

- more massive and somewhat shallower stellar haloes (Fig. 22);

- larger amounts and relative fractions of ex-situ, i.e. accreted, stellar mass also at fixed $z = 0$ galaxy mass (Fig. 24);

- thicker and hotter stellar disks, but only for the geometrically thick-disk components and for galaxies with smaller disks (Fig. 25);

- somewhat more massive SMBHs at fixed $z = 0$ galaxy mass, and larger gas mass reservoirs and higher star formation rates even at $z = 0$.

These results suggest that it may be possible to associate a probability for an observed galaxy to have experienced a recent major merger on the basis of some key structural and global observable properties (Zhu et al., 2022a, **?**).
Importantly, the results quantified in this paper align well with the current observational constraints on the properties and on the recent past assembly histories of the Galaxy and Andromeda. As anticipated in the Introduction, whereas the last major merger of the Milky Way may have occurred as early as 8 – 11 billion years ago, Andromeda's last major merger may have happened only a few billion years ago. As indicated throughout the paper by reporting known observational properties from the literature, the Milky Way exhibits a steeper stellar halo profile, a thinner and colder stellar disk, and a higher star formation rate – given its mass – than Andromeda: all this is consistent with the average phenomenology of TNG50 MW/M31-like galaxies with ancient and recent last major mergers.
With this paper we clearly signal that, with current state-of-the-art cosmological galaxy simulations that encompass wide ranges of merger histories within the ΛCDM hierarchical growth of structure, there is no need to simulate galaxies with quiet recent merger histories to obtain galactic stellar disks at $z = 0$. This approach had been frequently adopted over the last decade with zoom-in cosmological simulations of *MW- or M31-mass* galaxies, whose parent DM haloes had to be chosen from lower-resolution, DM-only volumes for re-simulation. Whereas this choice has almost always been dictated by the difficulties of the overcooling problem and by practical strategies (quieter merger histories imply faster computing times), it has the drawback that, in practice, no past and recent zoom-in simulations could really address the question of whether the absence of a major merger since $z \lesssim 1$ is a necessary condition for a galaxy to have a stellar disk as thin as the Galaxy's (i.e. 175 – 360 pc and 625 – 1450 pc for geometrically thin and thick components, respectively). Additionally, high-resolution galaxy simulations that are suitable for understanding the formation and evolution of Andromeda have been, by imposition or necessity, exceedingly rare, at least until recently and until TNG50. In fact, TNG50 produces

a few example galaxies whose more detailed stellar disk structures are also compatible with the Galaxy's and Andromeda's. Even more interesting, TNG50 produces example galaxies that did experience a recent major merger and have thin and thick disk heights similar to those of Andromeda (0.9 – 1.3 and 2.2 – 3.4 kpc, respectively). Moreover, according to TNG50, it does seem improbable, but not impossible, for a galaxy as small (in disk size) as the Galaxy to have emerged from a major merger over the last 5 billion years with thin-disk height as small as the Galaxy's. However, these statements do depend on where within the disk and how the vertical structure is assessed.
We conclude this discussion by highlighting that important structural changes and star formation episodes can be triggered in the progenitors of MW/M31-like galaxies also prior to coalescence with a major companion, i.e. during pericentric passages (see Figs. 15 and 19). It is hence to be expected, and to be searched for, that the Milky Way's disk and stellar halo may exhibit perturbations (such as waves and vertical spirals) that are linked to the passage of the Large Magellanic Cloud, as recently pointed out with observational data by e.g. Vasiliev et al. (2021) and Conroy et al. (2021). Our findings also offer a qualitative glimpse into what our Galaxy may experience in just a a couple of billion years in the future (Cautun et al., 2019), when the Large Magellanic Cloud will eventually merge with it.

Acknowledgements

DS, AP, and MD acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 138713538 – SFB 881 ("The Milky Way System", subprojects A01 and A06). The TNG50 simulation was realized with compute time granted by the Gauss Centre for Super- computing (GCS), under the GCS Large-Scale Project GCS-DWAR (2016; PIs Nelson/Pillepich) on the GCS share of the supercomputer Hazel Hen at the High Performance Computing Center Stuttgart (HLRS). DN acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) through an Emmy Noether Research Group (grant number NE 2441/1-1).

4.7 Fit of radial and vertical stellar density profiles

In this appendix, we explain how the measurements of the stellar disk lengths and the scale heights are performed in this paper.
For the disk length, for any given galaxy, we select all stellar particles with almost circular orbits (i.e. disk stars, ϵ >0.7) between one and four times the half-mass radius, so excluding the bulge region. We fit an exponential profile to the radial stellar surface density distribution in face-on projection, in bins of 2 kpc:

$$\Sigma(R) = \Sigma_{\rm d} \exp\left(-\frac{R}{\rm R_d}\right), \tag{4}$$

where Σ_d is the central stellar mass surface density of the disk and R_d is the disk scalelength. For each galaxy, 100 fits starting with random initial values (around the values taken at the limits of the cylindrical shell) are performed. We take the mode of the distribution of the 100 best-fit values as the best measure of the scale length, i.e. the one we quote for each galaxy; the error is obtained by the interquartile range.

For the disk heights, for each galaxy projected edge-on, we select stars in circular orbits in the cylindrical shell between 3.5 and 4.5 times the disk length and we fit a double function profile (either linear or square hyperbolic secant) to the vertical stellar mass distribution, between 50 pc and 5 kpc in bins of 200 pc:

$$\rho(z) = \rho_{\rm thin}\ \mathrm{sech}\left(\frac{z}{\mathrm{h_{thin}}}\right) + \rho_{\rm thick}\ \mathrm{sech}\left(\frac{z}{\mathrm{h_{thick}}}\right)\ ; \tag{5}$$

$$\rho(z) = \rho_{\rm thin}\ \mathrm{sech}^2\left(\frac{z}{2\mathrm{h_{thin}}}\right) + \rho_{\rm thick}\ \mathrm{sech}^2\left(\frac{z}{2\mathrm{h_{thick}}}\right)\ . \tag{6}$$

This gives us the scale heights of both a "thin" and a "thick" disk component (geometrical). The factor 2 for the squared case allows that the scale heights of the linear and squared cases are comparable in magnitude. Both profiles are similar at high heights but significantly different close to the galactic midplane. We justify the choice of adopting two functional forms because both have been extensively used to describe vertical galaxy density profiles and yet may give in certain instances rather different results. The sech^2 profile has been employed by e.g. **?**Bizyaev et al. (2014) for observations of edge-on spiral galaxies, and Stinson et al. (2013), Ma et al. (2017), Park et al. (2021) for simulated galaxies. The sech profile, justified theoretically by van der Kruit (1988) and Banerjee & Jog (2007), has been used by Schwarzkopf & Dettmar (2000) in galaxy surveys or by Qu et al. (2011) for simulations. On average, the scale heights obtained from the double sech^2 profile are ∼10-20% smaller than those obtained from the double sech profiles (in agreement with Qu et al. (2011).

As for the lengths measurements, for each galaxy and choice of functional form, we perform 100 fits starting with random initial values of the scale heights (in the ranges 20 to 1000 pc and 800 to 7000 pc for the thin and thick disks, respectively), for each functional form choice. We take the mode of the distribution of of the 100 best-fit values as the best measure of the scale height for a given parametric function. If needed, an error can be represented by the interquartile range of such distribution. In the main part of the manuscript, we opt instead to provide an estimate of the (systematic) error by quoting, for each galaxy, the best scale-height measures from both fitting functions of eqs. 10 and 6.

5 Disk flaring with TNG50: diversity across Milky Way and M31 analogs

The work of this chapter is published as:
Disc flaring with TNG50: diversity across Milky Way and M31 analogues
Diego Sotillo-Ramos, Martina Donnari, Annalisa Pillepich, Neige Frankel, Dylan Nelson, Volker Springel and Lars Hernquist
Monthly Notices of the Royal Astronomical Society,
Volume 523, Issue 3, August 2023, Pages 3915–3938

Zusammenfassung

We use the sample of 198 Milky Way (MW) and Andromeda (M31) analogs from TNG50 to quantify the level of disk flaring predicted by a modern, high-resolution cosmological hydrodynamical simulation. Disk flaring refers to the increase of vertical stellar disk height with galactocentric distance. The TNG50 galaxies are selected to have stellar disky morphology, a stellar mass in the range of $M_* = 10^{10.5-11.2}$ $M_\odot$, and a MW-like Mpc-scale environment at $z = 0$. The stellar disks of such TNG50 MW/M31 analogs exhibit a wide diversity of structural properties, including a number of galaxies with disk scalelength and thin and thick disk scaleheights that are comparable to those measured or inferred for the Galaxy and Andromeda. With one set of physical ingredients, TNG50 returns a large variety of flaring flavours and amounts, also for mono-age stellar populations. With this paper, we hence propose a non-parametric characterization of flaring. The typical MW/M31 analogs exhibit disk scaleheights that are $1.5 - 2$ times larger in the outer than in the inner regions of the disk for both old and young stellar populations, but with a large galaxy-to-galaxy variation. Which stellar population flares more, and by how much, also varies from galaxy to galaxy. TNG50 de facto brackets existing observational constraints for the Galaxy and all previous numerical findings. A link between the amount of flaring and the $z = 0$ global galaxy structural properties or merger history is complex. However, a connection between the scaleheights and the local stellar vertical kinematics and gravitational potential is clearly in place.

5.1 Introduction

Understanding the formation and evolution of our Galaxy, of Andromeda, and of other disk galaxies is one of the main quests of modern astrophysics. Over the last decade, large spectroscopic surveys have constrained quantities such as the ages, element abundances and phase-space properties of the stars in the Milky Way, mostly in the proximity of the Sun but also at several kpc distance, throughout the disk, bulge and stellar halo. These include LAMOST (Deng et al., 2012), RAVE (Steinmetz et al., 2020, and references therein), SEGUE/SDSS (Blanton et al., 2017), APOGEE (Majewski et al., 2017), GALAH (Martell et al., 2017), H3 Survey (Conroy et al., 2019a) and finally *Gaia* (Gaia Collaboration et al., 2016), with the delivery of positions and proper motions for more than 1.4 billion stars in the third data release (Gaia Collaboration et al., 2022). Similarly, albeit from a distance of about 750 kpc, photometric and spectroscopic surveys like PHAT (Dalcanton et al., 2012) and SPLASH (Gilbert et al., 2009) have mapped large portions of the disk of Andromeda and its disk-halo interface.

5.1.1 The stellar disk of the Galaxy

A remarkable feature uncovered over the past few years about the stellar disk of our Galaxy is the existence of two different stellar populations in the solar neighbourhood: on the one hand, alpha-rich and metal-poor stars seem to associate well with the *geometrical* or *morphological* "thick" disk, with scaleheight of $\sim$ 600-1400 pc and with old and kinematically-hotter stars; on the other hand, metal-rich stars with lower, i.e. solar [α/Fe] abundances are thought to populate the *geometrical* or *morphological* "thin" disk, with scaleheight of $\sim$ 150-350 pc and characterised by young and kinematically-colder stellar populations (e.g. Gilmore & Reid, 1983a, Jurić et al., 2008, Adibekyan et al., 2012, Haywood et al., 2013, Bland-Hawthorn & Gerhard, 2016, the latter for a compilation of several measurements). Interestingly, whether this correspondence in the solar neighbourhood extends also away from it is not obvious, with theoretical analyses suggesting that it does not (Minchev et al., 2015, Martig et al., 2016) and that the geometrical and the chemical definitions of the thin and thick disk do, in general, describe different stellar populations.
Recent large scale surveys have allowed to collect 6D phase space data, ages and abundances beyond the solar neighbourhood. From these, a global picture has emerged, whereby the low-alpha disk is young and the high-alpha disk formed early (see, e.g., Bensby et al., 2014, Miglio et al., 2021, Xiang & Rix, 2022). This is consistent with the so-called "inside-out" and "upside-down" scenario, whereby at early times stars were born in a radially-compact but vertically-thick disk and, later on, a thin and more extended disk developed (e.g. Robin et al., 2014, Bovy et al., 2016, for observational inferences) and (Bird et al., 2013, Stinson et al., 2013, Minchev et al., 2014a, Buck et al., 2020, Agertz et al., 2021, Nelson et al., 2021, Bird et al., 2021, Yu et al., 2022, for numerical modeling).
However, observations of our Galaxy have suggested that two simultaneous facts are in place at larger galactocentric distances than 8 kpc: on the one hand, the scaleheight of the Galactic stellar disk increases at larger galactocentric distances, a phenomenon called *flaring* (see below; first observed in the HI disk); on the other hand, the Milky Way's geometric thick disk, here denoted as stars at large heights over the disk plane ($\gtrsim$ 2 kpc), also contains young stars (Ness et al., 2016, Xiang et al., 2017, 2018, Feuillet et al., 2019).

5.1.2 Disk flaring in the Galaxy

The flaring of the stellar disk of our Galaxy has been studied quantitatively, namely by inferring from observations the changes in stellar disk height with galactocentric distance. Studies have found that, considering the whole stellar population, the stellar disk of the Galaxy is flared in the outskirts (López-Corredoira & Molgó, 2014, Kalberla et al., 2014, Wang et al., 2018), but this phenomenon is unlikely to be present in the inner disk (Mateu & Vivas, 2018). Flaring may be more appreciable when the disk stars are dissected into mono-age (e.g. Lian et al., 2022) and/or mono-abundance populations (e.g. Minchev et al., 2017), populations, which in general are not equivalent. Bovy et al. (2016) binned APOGEE stars in mono-abundance populations and quantified the changes in stellar disk height between 4 and 15 kpc from the Galactic Center. They found that the high-[α/Fe] population – mainly associated with old stars – does not show any evidence of flaring, whereas low-[α/Fe] stars – associated with young populations – present clear evidence of a flaring (also described by Ness et al., 2019), with scaleheights exponentially increasing as a function of galactocentric distance. Mackereth et al.

(2017) binned APOGEE RGB stars between 3 and 15 kpc from the Galactic Center in mono-age & mono-[Fe/H] populations and reached similar albeit not identical conclusions as above (and ones shared by Minchev et al. 2017): all mono-age populations flare although to different levels: in the Galaxy, younger populations flare more. The results of these two studies have different implications for the early stages of the MW: while a low-[α/Fe] population with no flaring could be associated with a more quiet early merger history, a picture where the older stars flare less is compatible with an early active phase of accretion that suppresed the flaring (e.g. Minchev et al., 2014b). Finally, Ting & Rix (2019) studied the vertical motion of low-alpha disk stars via their vertical actions, and published an analytical function for the mean vertical action of stars at given age and radius. Their findings imply a manifest flaring of the young stellar population (3 Gyr) with scaleheights of 120 and 500 pc at 4 and 14 kpc, respectively (see §5.4).
These recent results point to a consistent picture for the Galaxy, at least, and only when mono-age or mono-abundance stellar populations are analyzed separately: young stars in the Galaxy exhibit some level of flaring, at least at radii $\gtrsim 10-11$ kpc (see otherwise Mateu & Vivas, 2018). However, as of today, some confusion and uncertainties remain as to how different levels of flaring may map into the distributions of stellar ages in the height vs. radial distance plane and as to how the flaring of different mono-age stellar populations translates into the flaring of the morphological thin and thick disks. Finally, it remains unclear whether the phenomenology in the Milky Way is representative of most spiral galaxies or not.

5.1.3 Thin and thick Galactic disks

Whether the morphological or geometrical thin and thick disks of the Galaxy are two distinct components, or just the manifestation of a single variable structure, is also a matter of debate. The former is the classical view, as described for example in Jurić et al. 2008, whereas **other** analyses (e.g. Bovy et al., 2012, Rix & Bovy, 2013) argue for the latter, with the vertical structure of the Galactic disk being a continuum of stellar populations. The claim is that, even if the vertical stellar mass density profile is well described by a double exponential fit or similar, this does not necessarily imply a physically-originated decomposition. It is because of these arguments that it is now customary to characterize the vertical stellar disk structure in terms of mono-age or mono-abundance stellar populations, which should promise clearer physical insight.

5.1.4 Flaring of Andromeda and other spiral galaxies

Also the stellar disk of Andromeda seems to be well described by a double vertical component, with thin and thick disks separating in both kinematics and metallicity (Collins et al., 2011), but with scaleheights approximately two to three times larger than those observed in the Milky Way: about 0.9-1.3 and 2.2-3.4 kpc, respectively. On the other hand, the level of disk flaring in Andromeda remains unclear and, de facto, unaccessible. On the other hand, disk flaring has been suggested by observations in a number of edge-on spiral galaxies (de Grijs & Peletier, 1997, Narayan & Jog, 2002, Kasparova et al., 2016, Rich et al., 2019, Sarkar & Jog, 2019), although other studies were not able to find any significant amount of flaring (Comerón et al., 2011b, for the thick disks of external edge-on galaxies).

5.1.5 Disk flaring with theoretical and numerical models

The formation of thick disks and disk flaring are two closely linked phenomena. Several processes have been suggested to be responsible for the thickening of the stellar disk in general or for the disk flaring in particular: radial migration, accretion of stars from satellites, heating of a thinner pre-existing disk through mergers, and in-situ star formation from gas-rich mergers. The thickening of the stellar disk with galactic radius has been suggested to be a natural consequence of radial orbit migration by Sales et al. (2009), Schönrich & Binney (2009), Loebman et al. (2011), Roškar et al. (2013). On the other hand, other studies support that radial migration does not significantly contribute to the thickening of stellar disks (Minchev et al., 2012, Martig et al., 2014, Vera-Ciro et al., 2014, 2016, Minchev et al., 2014a, Grand et al., 2016) but it can affect flaring: in a quiescent disk, radial migration may enhance flaring (Minchev et al., 2012), but in the presence of mergers, radial migration may function as a suppressor of flaring, bringing kinematically colder stars to the galactic outskirts heated by mergers (Minchev et al., 2014a, Grand et al., 2016, Ma et al., 2017).
In fact, flaring could also be a consequence of heating caused by external triggers, i.e. not because of secular processes but rather events such as the infall of satellites or the interaction with other flying-by galaxies (e.g. Kazantzidis et al., 2009, with N-body only models of disks bombarded by cosmologically-consistent subhaloes).
Cosmological hydrodynamical simulations of well-resolved Milky Way (MW)-like galaxies, which have become increasingly realistic over the last decade (see e.g. Guedes et al., 2011, Wetzel et al., 2016b, Grand et al., 2017, Agertz et al., 2021, for Eris, LATTE, the Auriga sample, and VINTERGATAN, respectively), have also allowed to address the question of disk flaring.
Minchev et al. (2015), analysing two simulated galactic disks formed in a cosmological context – one from Martig et al. 2012 and one from Aumer et al. 2013b –, demonstrated that a non-flaring thick disk can actually be in place even if several mono-age populations with different levels of flaring are superposed – a statistical phenomenon commonly known as *Simpson's paradox*: a trend can appear in several groups of data but disappear or reverse when the groups are combined (studied in more detail by Minchev et al. 2019). By using a cosmological zoom-in simulation from the FIRE project of a MW-mass galaxy ($M_{\rm stars} \simeq 6 \times 10^{10} {\rm M}_{\odot}$ at $z = 0$), Ma et al. (2017) found that the scaleheight of mono-age stellar populations shows an outward and somewhat linear flaring, being higher at larger galactocentric distances. However, differently from Minchev et al. (2015), in the FIRE galaxy the scaleheights of both the thin and thick disks are found to be flared, with nearly the same slope of the mono-age populations. Also in all the 30 Auriga MW-analogs (Grand et al., 2017), an exponential flaring is a common feature: the flaring is in place for both young stars (< 3 Gyr) and the whole stellar populations, even though by different amounts. Indeed, by fitting the flaring with an exponential trend, Grand et al. (2017) found that, in the majority of the Auriga galaxies, young stars show a higher degree of flaring with respect to the global ones.
With one of the APOSTLE cosmological hydrodynamical simulations, Navarro et al. (2018) demonstrated that the stellar-disk flaring reflects the flaring of the gaseous disk – as stars inherit the properties of the gas at their birth – and argued that the age and metallicity gradients are settled at birth and are not the result of radial migration or disk instabilities.
The flaring of mono-age populations is non negligible in all the five NIHAO-UHD MW-like galaxies (Buck et al., 2020) and in the MW-mass disk galaxy VINTERGATAN (Agertz et al., 2021). However, some of them flare linearly, others flare with an exponential radial trend of

the heights. Moreover, the increasing of the scaleheight in the NIHAO-UHD sample is found to be much stronger for the old stellar populations, unlike the case of the Galaxy, and mild-to-no flaring is appreciable when all stellar populations are combined, similarly to the cases by Minchev et al. (2015).
Finally, more recently, García de la Cruz et al. (2021) expanded the work of Minchev et al. (2015) by showing the vertical structures of 27 MW-like galaxies with $M_{\rm stars} \simeq 10^{10} - 2 \times 10^{11} {\rm M}_{\odot}$ at $z = 0$: they found that in 44 per cent of their galaxies, the morphological thick disk does not flare and this typically occurs in galaxies with $M_{\rm stars} < 5 \times 10^{10} {\rm M}_{\odot}$, with a thin disk (< 1kpc) and a rather quiescent merger history. On the other hand, the remaining 15 galaxies show a flared thick disk and they are more massive, have a thicker disk and have undergone a major merger with respect to their non-flaring counterparts.
Despite the many recent results on the topic put forward by the simulation community, the scientific and general interpretation of the findings above, and their applicability to the cases of the Galaxy or Andromeda, are impeded by a number of limitations. Firstly, most of the analyses based on state-of-the-art cosmological models remain qualitative and refer to one or just a few galaxies formed within a specific galaxy-formation model: namely, they are often reduced to the plotting of the stellar scaleheights (and/or vertical stellar velocity dispersion) as a function of radius for stars in e.g. different age bins, and are associated to only one or just a few specific realizations of galaxies that span a limited range (if any) of mass, merger history and stellar disk structure. Secondly, when the study of more than a handful of objects is possible, the quantification of the flaring is not consistently derived across the analyses, making the comparison of the predicted outcome problematic.

5.1.6 TNG50 and the scope of this paper

In this paper, we use the most recent and highest-resolution simulation of the IllustrisTNG project (Pillepich et al., 2018b, Nelson et al., 2018, Marinacci et al., 2018, Naiman et al., 2018, Springel et al., 2018), TNG50 (Pillepich et al., 2019, Nelson et al., 2019b), and quantify the stellar disk flaring of 198 MW- and M31-like galaxies, thereby tripling the number of cosmologically-simulated galaxies analyzed to this end. This is possible thanks to the mass and spatial resolution of the simulation, which returns galaxies with disks as thin as 100 – 200 pc (Pillepich et al., 2019, Sotillo-Ramos et al., 2022), and to the encompassed volume, with realistic galaxy properties and galaxy populations across a wide range of masses, types, and environments i.e. not only for the case of disk, star-forming galaxies that form in $10^{12} {\rm M}_{\odot}$ haloes.
We hereby focus on the vertical distribution of the stellar mass in disks at $z = 0$ and on its connection to the vertical stellar velocity dispersion. We assess the flaring both for the morphological thin and thick disks (i.e. when single-component vertical fits are not appropriate to obtain scaleheights) and especially by separately studying mono-age stellar populations. We again postpone to future work the study of how the latter connect to mono-abundance populations in the context of the IllustrisTNG model and enrichment, but we give particular emphasis to whether and how often –i.e. across the selected galaxy sample –, young disk stars flare more or less than old disk stars, and we explore the relationship between the degree of the flaring and $z = 0$ galaxy and disk properties.
In Section 6.2, we hence summarize the salient aspects of the TNG50 simulation, describe the adopted selection of MW/M31-like galaxies, and define the ways we characterize the simulated stellar disks. In Section 5.3 we show the range of stellar-disk structures encompassed by the

TNG50 MW/M31 analogs, including their scaleheights as a function of radius and the cases of warped and disturbed stellar disks. We quantify the vertical disk structure and flaring predicted by TNG50 for MW/M31-like galaxies in §5.4. There we also argue for, and propose, a non-parametric and more-generally applicable and comparable method to quantify the amount of the disk flaring and compare the flaring of stars of different ages and to the inferences for our Galaxy. We connect stellar disk heights to the underlying stellar kinematics and potential in §5.5. In Section 5.6, we quantitatively compare the TNG50 results to those from previous simulations, by casting them all under the same general and non-parametric flaring quantification; we discuss our results, limitations, and the possible origin of the diversity predicted by TNG50, and connect to observations of the distributions of the stellar ages as a function of galactocentric radius and height. Summary and conclusions are given in Section 6.4.

5.2 Methods

5.2.1 The TNG50 simulation

The TNG50 simulation is, among the flagship runs of the IllustrisTNG project (Nelson et al., 2019b), the smallest in volume but best in resolution: it evolves a cubic box of ~ 50 comoving Mpc a side, sampled by 2160^3 dark-matter particles and 2160^3 initial gas cells (Nelson et al., 2019b, Pillepich et al., 2019), with a resulting gas-cell and stellar-particle mass resolution of about 8.5×10^4 $M_\odot$ and a dark-matter mass resolution of about 4.5×10^5 $M_\odot$.
TNG50 uses the code Arepo (Springel, 2010b) and includes the IllustrisTNG galaxy-formation model introduced and described in the method papers by Weinberger et al. 2018, Pillepich et al. 2018a: in practice, it solved for the coupled equations of gravity and magneto-hydrodynamics in an expanding Universe, in addition to prescribing the cooling and heating of the cosmic gas, star formation, stellar evolution and enrichment, as well as phenomena such as stellar feedback and the seed, growth and feedback from supermassive black holes (SMBHs). The initial conditions of TNG50 have been initialized at redshift $z = 127$ and assume a cosmology compatible with the Planck 2015 results (Planck Collaboration et al., 2016).
As in previous large-scale and zoom-in cosmological simulations of MW-mass galaxies, also in TNG50 stellar particles do not represent individual stars but rather simple, mono-age stellar populations of thousands of stars characterized by an initial stellar mass function (Chabrier, 2003, for TNG50). On the other hand, a few modeling elements set apart TNG50 from the great majority of cosmological simulations that have been used so far to study the vertical structure and flaring of galactic disks: chiefly, the inclusion of magnetic fields and the effects of SMBH feedback (both also in Auriga, Grand et al., 2017). Importantly, TNG50 is a relatively large uniform-volume simulation and so, differently than in many of the aforementioned zoom-in simulations, it returns a large number of massive galaxies ($\simeq 800$ at $z = 0$ above 10^{10} $M_\odot$) and hence, among them, also many MW and M31-mass objects spanning a wide range of merger histories, i.e. *without* any a-priori choice about the number and time of their past major mergers (Sotillo-Ramos et al., 2022).
TNG50 is suitable for studying disk flaring thanks to its mass and spatial resolution (see Nelson et al., 2019b, Pillepich et al., 2019, 2021, for more details). The smallest gas cell in TNG50 at $z = 0$ measures 9 pc across, whereas the average gas cells within the star-forming regions of massive galaxies at $z = 0$ are typically of the order $50 - 200$ pc: this means that processes such as star formation and feedback are implemented below such spatial scales. The gravitational

potential, on the other hand, is softened on different scales for different types of resolution elements: the softening length of the stellar and DM particles read 288 pc, the smallest softening length of the gas cells is 72 pc. These are sufficient to capture half-light disk heights of 200 – 400 pc for the typical massive star-forming galaxy at low redshift, but also thinner ones (i.e. thinner than the softening length for stellar particles; see Pillepich et al. (2019) and next Sections). Pillepich et al. (2019) presents also a study of the resolution effects on galaxy sizes and heights in TNG50: stellar disk thickness (as stellar half-mass height) can be considered to be converged in TNG50 to better than 20–40 per cent.
Finally, the choice and functioning of the IllustrisTNG model as implemented in TNG50 have been validated against observations not only of the Galaxy or Andromeda, but of large galaxy populations (see Pillepich et al., 2021, for a summary).

5.2.2 Galaxy selection: choosing MW and M31 analogs

In the following, we identify (sub)halos within the TNG50 volume by using the Friends-of-Friends (FoF) and SubFind algorithms (Davis et al., 1985, Springel et al., 2001b). Also, we define a "virial" halo boundary, R_{200c}, as the radius within which the mean enclosed mass density is 200 times the critical density of the Universe. We refer to the total mass enclosed within this radius as the virial mass, M_{200c}, of the host halo. Additionally, all galaxies residing within one virial radius of the host center are dubbed as "satellite" or "galaxy", whereas the galaxy settled at the deepest potential within a FoF is named "central", and it is typically, but not always, the most massive one. The galaxy stellar mass adopted in this work ($M_{\rm stars}$) is the sum of all stellar particles within a fixed aperture of 30 physical kpc, unless otherwise stated.
With these definitions in mind, we select MW/M31-like galaxies from TNG50 at $z = 0$ by means of the following three criteria, all in turn based on *observable* rather than *halo-based* properties. Extended motivations and characterizations for this selection are given in Pillepich et al. (2023) and the resulting sample has been already used in Engler et al. (2021b, 2023), Pillepich et al. (2021), Sotillo-Ramos et al. (2022), Chen et al. (2023), Ramesh et al. (2023).
Namely, at $z = 0$ we select galaxies:

- with $M_{\rm stars}$ (<30 kpc) = $10^{10.5-11.2}$M$_\odot$;
- with a disky stellar shape (see Section 5.2.2 for more details)
- in isolation: no other galaxies with $M_{\rm stars} > 10^{10.5}M_\odot$ within 500 kpc and host halo mass $M_{200c} < 10^{13}$M$_\odot$.

This leads to a sample of 198 TNG50 MW/M31-like galaxies, and since it is not required for a galaxy to be the central of its halo, our sample also includes pairs of a few Local Group-like systems. Additionally, we note that, differently from the majority of zoom-in simulations of MW-mass haloes (see for example Guedes et al., 2011, Roca-Fàbrega et al., 2016, Agertz et al., 2021, Renaud et al., 2021), our sample is not a-priori *biased* in its history, namely, we have not imposed a recent quiescent merger history for the MW/M31-like galaxies to be part of our sample (see Sotillo-Ramos et al. (2022) for an in-depth analysis of the merger history of all TNG50 MW/M31 analogs).
Throughout the text and in some selected Figures, we at times label *MW-mass* and *M31-mass* those galaxies within the MW/M31-like sample with stellar mass in the ranges $10^{10.5-10.9}$M$_\odot$

and $10^{10.9-11.2}$M$_\odot$, respectively. MW-mass (M31-mass) galaxies are sampled, on average, by $\simeq 5.5 \times 10^5$ (1.3×10^6) gravitationally-bound star particles.

Stellar morphology selection

As stated above, each galaxy in the TNG50 MW/M31 sample has been selected to be *disky*, i.e. they either satisfy a stellar-morphology constraint based on the minor-to-major axis ratio (c/a) of the stellar mass distribution or are disky by visual inspection. The first criterion is satisfied if $c/a \geq 0.45$ (see Chua et al., 2019, Pillepich et al., 2019, for more details), being c and a the minor and major axis of the ellipsoidal distribution of stellar mass between 1 to 2 times the stellar half-mass radius ($R_{\rm stars,1/2}$). Additionally, in the TNG50 MW/M31-like sample are also present galaxies that, even if with $c/a > 0.45$, clearly appear disky and with well-defined spiral arms by visual inspection, based on 3-band images in face-on and edge-on projections. Of 198 TNG50 MW/M31-like galaxies, 25 have been included via the visual-inspection step. See Pillepich et al. (2023) for more details.

5.2.3 Measurement of stellar disk properties

Definition of disk stars

Throughout this paper, we quantify the structures of the simulated galactic stellar disks based on mass, i.e. based on the spatial location of stellar particles and the stellar mass densities they sample.
We call disk stars all the gravitationally-bound (according to SubFind) stellar particles that are in circular orbit upon inspection, i.e. with circularity $\epsilon = L_z/L_{z,\rm circ} > 0.7$. Here L_z is the z-component of the angular momentum of a given star particle and $L_{z,\rm circ}$ is the angular momentum of a star located at the same radius but following a perfectly circular orbit. The z direction for each galaxy (its "up vector") is chosen to be the direction of the total angular momentum of all stars within $2\times R_{\rm stars,1/2}$. The galactic plane is hence the plane perpendicular to this up vector. The center of a galaxy is chosen as the location of its most gravitationally-bound element, typically the location of its SMBH.

Stellar disk lengths

We measure the disk length, for any given galaxy, selecting all stellar particles with circular orbits (i.e. disk stars, $\epsilon >0.7$) between one and four times the half-mass radius (i.e. excluding the bulge region). We fit an exponential profile to the radial stellar surface density distribution in face-on projection, in bins of 2 kpc:

$$\Sigma(R) = \Sigma_{\rm d} \exp\left(-\frac{R}{R_{\rm d}}\right), \tag{7}$$

where Σ_d is the stellar mass surface density of the disk at $R = 0$ and $R_{\rm d}$ is the disk scalelength, a characteristic scale that is commonly used as a proxy for the extension or the size of the stellar disk.
We perform 100 fits for each galaxy, by starting with random initial values around the values taken at the limits of the cylindrical shell. The fitting routine (python curve_fit) uses a non-linear

least squares method to fit our defined function to the data. The best measure of the scalelength of a given galaxy is then the mode of the distribution; an error can be obtained as the interquartile range.

Stellar disk heights

Given the up-vector for each galaxy, we analyze its vertical stellar disk structure, by rotating its stars in edge-on projections and by extracting and fitting the vertical stellar mass density distribution of the disk stars. We determine the latter at various different radii, centered at integer multiples, from 1 to 5 in steps of 0.5, of the the scalelength of the galaxy, $R_{\rm d}$, and dividing the galactic disk into radial annuli (cylindrical shells) of width one integer. The latter are also centered at multiple physical radii, in cylindrical shells of 2 kpc.
We use either a single or a double parametric formula to fit the vertical mass profiles at fixed gactocentric distance. In the literature, there are a variety of formulas to describe these profiles: the most common are exponential, hyperbolic secant and squared hyperbolic secant. All three can be seen as special cases of the general formula (see, e.g., van der Kruit, 1988):

$$f(z) \propto \mathrm{sech}^{2/n}(nz/2h_z), \tag{8}$$

for the cases $n \to \infty$, $n = 2$ and $n = 1$, respectively. All three tend to the exponential profile as z increases, being the shape of the profile for low values of z the main difference among the three. The single squared hyperbolic secant profile reads

$$\rho(z) = \rho_0 \, \mathrm{sech}^2(z/2h_z) \, , \tag{9}$$

where ρ_0 is the normalization and h_z is the disk scaleheight. The factor 2 in the denominator allows that the scaleheights of the exponential, and of the linear and squared hyperbolic secant cases, are comparable in magnitude. As shown in the upcoming sections, this provides a good description of the vertical mass distribution of mono-age stellar populations, for all selected galaxies. So the fit of Eq. 9 is the one we adopt to quantify the stellar disk scaleheight of mono-age stellar populations.
However, when all disk stars in a galaxy are considered, a two-component vertical formula returns a better fit for the majority of TNG50 MW/M31-like galaxies:

$$\rho(z) = \rho_{\rm thin} \, \mathrm{sech}^2(z/2h_{\rm thin}) + \rho_{\rm thick} \, \mathrm{sech}^2(z/2h_{\rm thick}) \, . \tag{10}$$

This gives us the scaleheights of both a "thin" (h_{z1}) and a "thick" (h_{z2}) disk component (Gilmore & Reid, 1983a, Yoachim & Dalcanton, 2006, Comerón et al., 2011b, 2012, Ma et al., 2017, Buck et al., 2020, Agertz et al., 2021, Navarro et al., 2018). We stress here that the division into a thin and thick disk of our TNG50 MW/M31-like galaxies is purely geometrical, i.e. morphological, and does not necessarily imply a meaningful physical decomposition into two separate orbital structures or into two distinctly-formed structures (see Bovy et al., 2012).
We choose to proceed with the single and double sech^2 formula because it has been extensively used: e.g. by Yoachim & Dalcanton (2006), Bizyaev et al. (2014) with observed edge-on spiral galaxies and Villalobos & Helmi (2008), Stinson et al. (2013), Ma et al. (2017), Park et al. (2021) with simulated galaxies. The main justification is the physical motivation: it represents the vertical density variation of a self-gravitating iso-thermal population (Spitzer, 1942, van

der Kruit & Searle, 1981). However de Grijs & Peletier (1997), Hammersley et al. (1999) claim that it never reproduces well the densities of the Galactic midplane. In fact, other works have preferred the exponential formula: it has been used as fit of the vertical stellar density profile of the Galaxy (Pritchet, 1983, Siegel et al., 2002, Jurić et al., 2008, Bovy et al., 2016, Mackereth et al., 2017) and of other galaxies (Comerón et al., 2011a); it has also been used for simulations in the works by Roca-Fàbrega et al. (2016), Buck et al. (2020), Agertz et al. (2021). van der Kruit (1988) proposed that an intermediate solution, i.e. a sech profile, would reproduce better the stellar vertical densities in the MW. It is adopted in the works by de Grijs & Peletier (1997), Matthews (2000).
Similarly for what is done for the scalelengths and in Sotillo-Ramos et al. 2022, we perform 100 fits for each galaxy, galactocentric distance, age bins, etc. The initial guess fit values are in the ranges 20 to 1000 pc and 800 to 7000 pc for the thin and thick disks when using the double function, and 100 to 7000 for the single function. We choose the mode of the distribution as the best measure of the scaleheight and quote errors as one standard deviations of the estimated parameters, provided by the fitting function.

5.3 The structural and age properties of the stellar disks in TNG50 MW/M31-like galaxies

Before quantifying the stellar disk flaring according to TNG50, we first comment on the structural properties of the stellar disks of the 198 TNG50 MW/M31-like galaxies – additional global and structural properties can be found in Pillepich et al. (2023) and references therein. An extensive analysis of the merger history of each MW/M31-like galaxy, and on how stellar disks can survive major mergers, is instead given in Sotillo-Ramos et al. 2022.

5.3.1 Diversity of stellar disk lengths and heights

An in-depth analysis of the global structural properties of the stellar disks of the 198 TNG50 MW/M31-like galaxies can be found in Pillepich et al. (2023). We refer the reader to that work for details, whereas here we report the most relevant facts.
TNG50 predicts a wide range of stellar disk sizes, also at fixed stellar mass. Within the TNG50 MW/M31 sample, the stellar disk scalelengths vary between $\sim$ 1.5 and $\sim$17 kpc, denoting a remarkable variety of disk extents in such a narrow range of stellar mass (Pillepich et al. (2023), their fig. 16, top). These sizes are consistent with previous zoom-in simulations of $\sim 10^{12}\,M_{\odot}$ haloes, e.g. Auriga (Grand et al., 2017), Eris (Guedes et al., 2011), NIHAO-UHD (Buck et al., 2020) and VINTERGATAN (Agertz et al., 2021). Also, TNG50 disk sizes are compatible with those measured for local disky and spiral galaxies (based on stellar light rather than stellar mass Gadotti, 2009, Lelli et al., 2016) and for the Galaxy (Hammer et al., 2007, Jurić et al., 2008, Bovy & Rix, 2013) and Andromeda (Worthey et al., 2005, Barmby et al., 2006, Hammer et al., 2007). We note that a number of TNG50 MW/M31-like galaxies fall within the observed values for the scalelength and stellar mass of the Galaxy and Andromeda, whereas the rest have more or less extended stellar disks for their mass: compared to the total TNG50 sample of MW/M31 analogs, the Milky Way has a rather compact stellar disk given its mass, as it settles at the lower end of the TNG50 distribution, while for Andromeda the value is rather average.
The scaleheights of TNG50 galaxies, evaluated at galactocentric distances of a few times the disk length, can be as small as $\simeq$ 200 pc (lowest 10th percentiles). Yet, stellar disks of TNG50

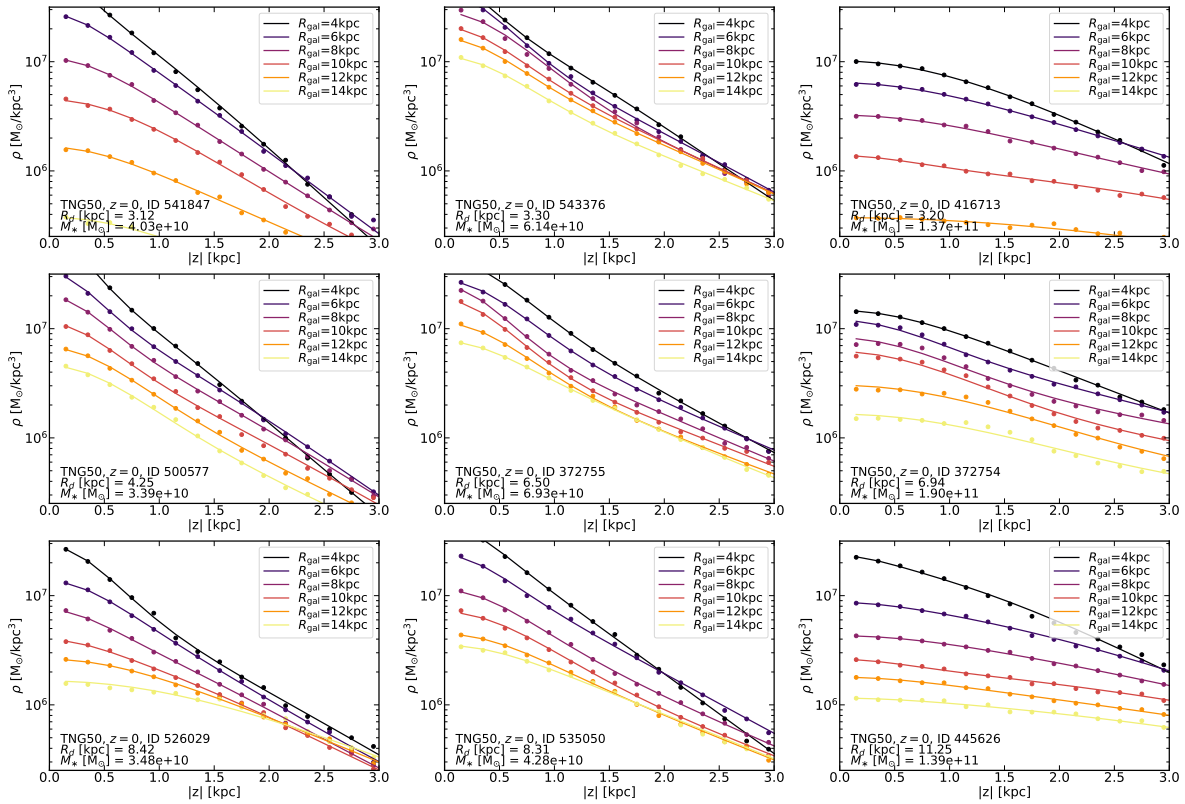

Figure 27: **Vertical surface density profiles of example TNG50 MW/M31-like galaxies.** In each panel, we show the vertical stellar mass density profiles at different radii for one of nine TNG50 MW/M31 analogs. Solid curves represent the two-component fit, as described in the text (see Eq. 10). These galaxies are selected to be representative of the whole sample, below the 25^{th} percentile, in between 25^{th} and 75^{th} percentile and above the 75^{th} percentile, of both galaxy stellar mass (left to right columns) and scalelength (top to bottom rows).

MW/M31-like galaxies (as selected in Section 5.2.2) can be as thick as a few kpc (Pillepich et al. (2023), their fig. 16, bottom panels). As is the case for the disk extent, TNG50 MW/M31-like galaxies have typically thicker thin disks than the Galaxy but not necessarily than Andromeda. TNG50 disk heights are consistent with those of zoom-in simulations (e.g. Guedes et al., 2011, Grand et al., 2017, Ma et al., 2017, Buck et al., 2020). There is even a number of TNG50 MW/M31-like galaxies that exhibit thin and thick disks with similar heights as the observational estimates of both the Galaxy and Andromeda.
We in particular highlight and take note of six galaxies (Subhalo IDs 516101, 535774, 538905, 550149, 552581, 536365), which we refer to as *MW-analogs* and whose stellar disk properties are within the observational estimates for the Galaxy. These MW-analogs are chosen among the TNG50 MW/M31-like galaxies that have thin and thick disk heights consistent with those of the Galaxy (approximately in the range 175 – 360 pc and 625 – 1450 pc, respectively), measured at either 7-9 kpc or $2.7 - 4.7 \times R_{\rm d}$; and with disk scalelength and stellar mass in the ranges 1.7 – 2.9 kpc and $10^{10.5-10.9}$ $M_{\odot}$, encompassing available literature contraints. There is also one galaxy (Subhalo ID 432106) that could be considered a *M31-analog*, based on its stellar mass, disk scalelengths and thickness (Pillepich et al., 2023).

5.3.2 Vertical stellar mass profiles

In Fig. 27 we show the vertical surface mass density profiles at different radii of disk stars for nine TNG50 MW/M31 analogs. These are selected to be representative of the whole sample, namely below, in between and above the 25^{th} and 75^{th} percentiles, in either galaxy stellar mass (from left to right) or in stellar disk scalelength (top to bottom). The scalelength and stellar mass are labeled in each panel. Dots represent the measured values of the stellar density, solid curves represent the resulting function (as per Eq. 10), whereas different colors denote different radii.
Amid differences in the details, all galaxies in Fig. 27 (and all the others not shown here) exhibit vertical stellar mass disk profiles that, in general, are steeper towards higher altitudes (or shallower closer to the galactic midplane, i.e. better described with a $\rho(z) \propto \mathrm{sech}^2$ profile than with a pure exponential) and whose normalization decreases with galactocentric radius. More massive galaxies (right-hand columns) typically exhibit shallower profiles as a function of physical vertical distance from the mid plane, consistently with the mild trend with mass of the scaleheights at fixed galactocentric distance present in the whole sample.
By selecting disk stars according to their ages, we find that the vertical mass distribution of each stellar population can be better fitted with a single-function profile, in agreement with observations (e.g. Martig et al., 2014, Xiang et al., 2018) and with previous simulations (Ma et al., 2017, Buck et al., 2020) – this further justifies our choice of using a single hyperbolic secant squared functional form (Eq. 9). In Fig. 28, we show an example of the vertical surface mass density profiles of mono-age disk stars (in bins of stellar ages of $\Delta t_{\rm age}$=2 Gyr and a final bin for all stars older than 10 Gyr): this is the TNG50 object with Subhalo ID 538905, selected to have a stellar mass and a disk scalelength similar to our Galaxy. Here we show the vertical stellar mass profiles of disk stars at three different radii, centered at ~ 4, 8, and 12 kpc, and covering a radial range of ±1 kpc (from top to bottom, respectively). We see that the scaleheight at all radii is smallest for the youngest stars or, in other words, that the vertical profile is flatter, i.e. thicker or more extended, with increasing stellar age. Also, in the case of this galaxy, at fixed age, the vertical profiles are shallower at increasing radius, indicating a clear presence of

a flaring.
A qualitatively similar vertical stellar mass profile is found for our Galaxy by Xiang et al. 2018 in the solar neighborhood with data from the LAMOST Galactic Survey, who however speculate that the sharp density profiles recovered for the old stellar populations may have suffered contamination from young main sequence stars, which in turn may increase the density near the disk mid-plane (see their figures 19 and 20).

5.3.3 Age distributions of disk stars

The great diversity in the disk structures of TNG50 MW/M31 analogs is also reflected in the ages of the stars that populate the disks. As we aim at characterizing the vertical stellar disk structure and disk flaring for mono-age stellar populations, here we give an overview of the age distributions of disk stars in TNG50 MW/M31-like galaxies.
Within the sample, TNG50 returns both disks characterized by a sizeable young component, as well as disks that are mostly old – consistently with the fact that not all TNG50 MW/M31-like galaxies are indeed star forming (Pillepich et al., 2021, 2023). In Fig. 29, we show the age distribution of disk stars for a selection of MW/M31-like galaxies in TNG50, divided in *young* disks, i.e. with a mean stellar age younger than 4 Gyr, and *old* disks, i.e. with a mean stellar age older than 9 Gyr. We further divide the MW/M31 sample into two bins of stellar mass, to separate them in MW-mass and M31-mass galaxies.
Now, young stellar disks are found only in MW-mass galaxies (7 galaxies, top panel), whereas old stellar disks are returned in both low and high stellar-mass galaxies (9 galaxies, middle panel, and 8 galaxies, bottom panel, respectively). All the other TNG50 MW/M31-like galaxies (174 of 198, not shown) have intermediate-age stellar disks, i.e. with broad distributions across all stellar ages.
Interestingly, most of the TNG50 MW/M31-like analogs exhibit many short periods of intense star formation. In Sotillo-Ramos et al. 2022 we have shown that these can be caused by merger events and close pericentric passages of gas-rich companions and in Boecker et al. 2023 we have shown that the latter can induce bursts of star formation also in the innermost regions of the disks.

5.3.4 The cases of warped and disturbed stellar disks

Vertical perturbations within or in the outer regions of galactic disks may affect our statistical quantification of their structural properties. They may also lead to an incorrect impression and quantification of the flaring (see next Sections). These can appear under the form of the common *S-shaped* warps, as well as asymmetries, and have been generally ascribed to e.g. the tidal distortions imparted by an external (merging) satellite in fly-by (Ostriker & Binney, 1989, Kazantzidis et al., 2009, Gómez et al., 2013, D'Onghia et al., 2016, Gómez et al., 2017, Semczuk et al., 2020) or to a misaligned accretion of high angular momentum cold gas (see for example Roškar et al., 2010, Aumer et al., 2013b).
In this paper we do not attempt an accurate and in-depth analysis of warped stellar disks, which we leave to future works, but we at least signal those TNG50 MW/M31-like galaxies that may exhibit some warps or disturbed stellar disks based on a visual inspection of their edge-on stellar maps. Among the TNG50 MW/M31 sample, we identify 21 galaxies that have a well-defined S-shaped warp and 17 galaxies with a generally disturbed or distorted stellar disk, which we

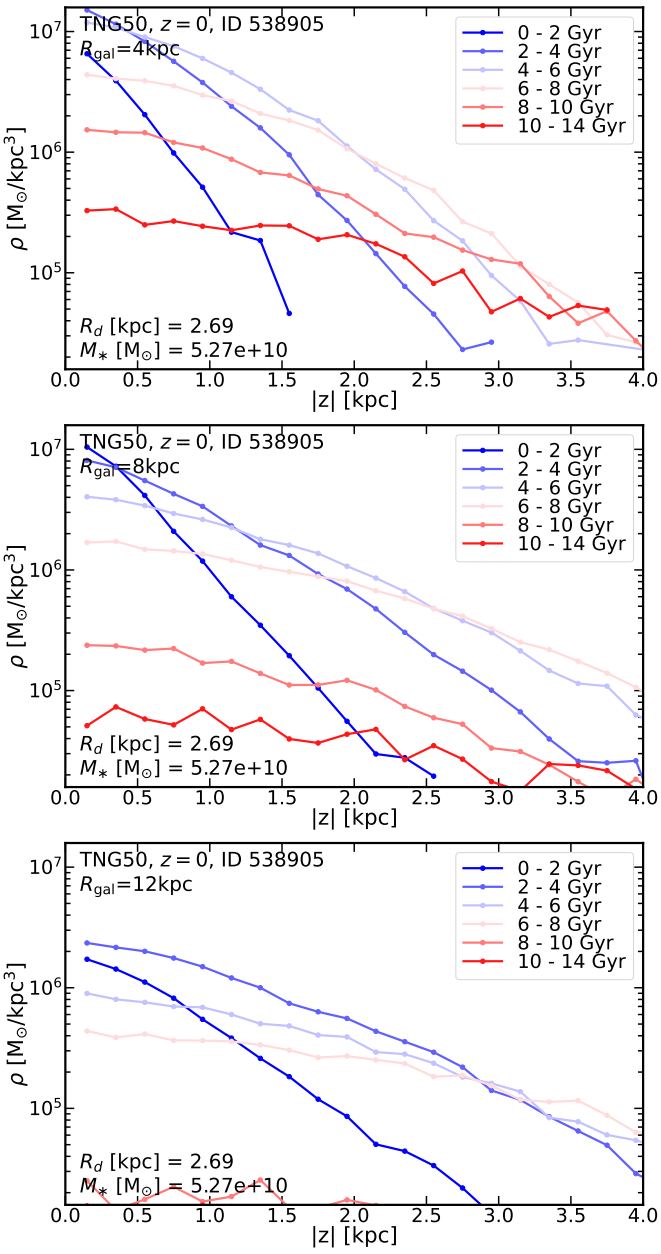

Figure 28: **Vertical surface density profiles of mono-age stellar populations in an example TNG50 MW/M31-like galaxy**. We show the vertical stellar surface mass density profile of mono-age stellar populations at three different radii: 3-5 kpc (top), 7-9 kpc (middle), and 11-13 kpc (bottom). Different colors in each panel denote different stellar ages, as labeled in the legend.

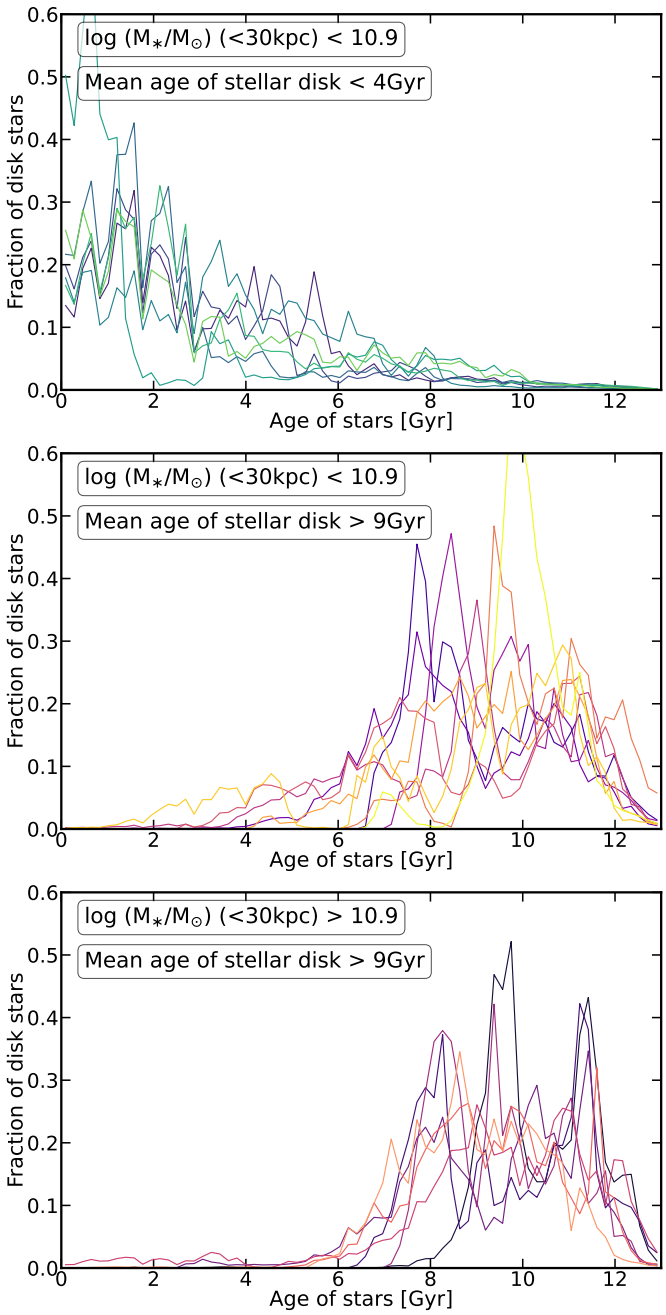

Figure 29: **Age distributions of disk stars in selected TNG50 MW/M31-like galaxies.** Each line represents one galaxy. In the top, we show example galaxies with a *young* disk, i.e. with a mean stellar age younger than 4 Gyr. In the middle and bottom panels, we show example galaxies with an *old* disk, i.e. with a mean stellar age older than 9 Gyr. Additionally, we split this sample into two bins of stellar mass, below (middle) and above (bottom) $10^{10.9}$M$_\odot$. Only low-mass galaxies show a relatively young stellar disk (7 galaxies, top panel) while old stellar disks are present in both low and high-stellar mass MW/M31-like galaxies (9 and 8 galaxies in the middle and bottom panel, respectively). All the other MW/M31-like galaxies (174 of 198, not shown) have an intermediate-age stellar disk.

refer to as *disturbed* disk. When needed, we will properly identify such galaxies in plots and discuss them in the text (e.g. Section 5.4.3). A catalog with corresponding flags is released with this paper.

5.4 Disk flaring with TNG50

Equipped with the vertical stellar mass distributions of stars and of mono-age stellar distributions throughout the simulated stellar disks, we can quantify how, and by how much, if at all, the disk scaleheights of all, young and old stellar populations increase with galactocentric distance. Namely, in the following, we provide results from TNG50 about the the vertical structure of stellar disks in MW/M31 analogs across galactocentric distances. The questions we would like to answer are the following: how often does flaring occur in MW/M31-like galaxies according to TNG50? And across all MW/M31-mass disk galaxies, how often do young and old stellar populations display the same or different amount of flaring?

5.4.1 Diversity in disk flaring across TNG50 MW/M31 analogs

Fig. 30 shows the disk scaleheights as a function of galactocentric distance of mono-age stellar populations in six MW/M31-like galaxies from the TNG50 simulation. Curves of different colors represent different bins of stellar ages, as labeled in the legend (blue to red from young to old), with scaleheights obtained via a single-component fit (Eq. 9). The dashed, dotted curves accounts for all disk stars in the simulated galaxies, with scaleheights obtained via a double-component fit (Eq. 10): black for geometrical thick disk and grey for the geometrical thin disk. These galaxies are chosen to highlight the variety of disk flaring that we find across the TNG50 MW/M31-like sample:

1. *Top left panel*: both mono-age and total stellar populations exhibit a substantial flaring, with the older stars flaring, upon visual impression, a bit more than younger ones. In this galaxy, the flaring of most mono-age populations is almost exponential (in analogy with Minchev et al. 2015, Grand et al. 2017 and with some galaxies of Buck et al. 2020).

2. *Top right panel*: both mono-age and total stellar populations exhibit a flaring, which appears in this case linear rather than exponential (in analogy with Ma et al. 2017, Agertz et al. 2021, García de la Cruz et al. 2021 and some galaxies of Buck et al. 2020).

3. *Middle left panel*: flaring is more evident when considering mono-age populations than when considering all disk stars at once (colored curves vs. black and grey ones; consistently with the results of Minchev et al., 2015).

4. *Middle right panel*: in this galaxy, young and old stellar populations follow two somewhat different trends, that can be identified, respectively, with the geometrical thin and thick disks.

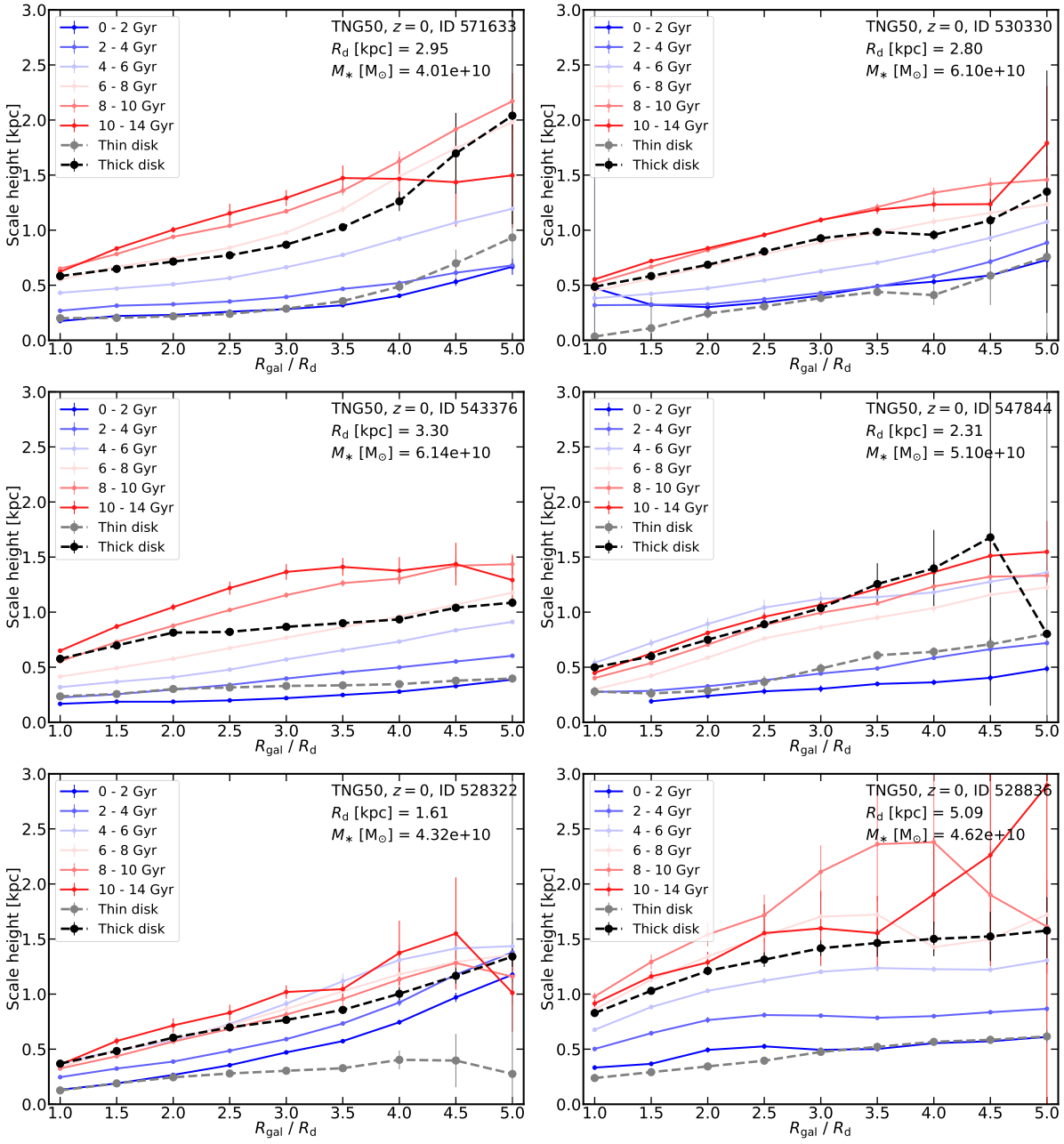

Figure 30: **Visualization of disk flaring in a few example TNG50 MW/M31-like galaxies**. We show disk scaleheight as a function of galactocentric distance normalized by the scalelength for six TNG50 MW/M31 analogs, chosen to highlight the diverse ways in which the flaring is manifested across the whole galaxy sample (see the text for details). Curves of different colors represent different stellar ages, as labeled in the legend. The black (grey) curves with circles denote the geometrical thick (thin) disk, i.e. all disk stars without splitting in stellar ages.

5. *Bottom left panel*: in this galaxy, all mono-age stellar populations and the thick disk are flared, but not the thin disk (or at least not out to 4.5×R_d). The latter does not follow any of the mono-age populations: clearly a double functional profile was necessary to unravel it.

6. *Bottom right panel*: a linear flaring of both mono-age and total populations is manifest in the inner part of the galaxy. The flaring disappears in the outer parts, as the scaleheights remains either constant or non monotononic.

The significance of the results uncovered here is remarkable: with a single set of physical-model ingredients, TNG50 returns all the manifestations of flaring that have been found so far in individual simulations and simulated galaxies. This demonstrates that the diversity of disk flaring outcomes can arise naturally from the diversity of the galaxy population.
It is also manifest from Fig. 30 that the change of scaleheights with galactocentric distance is not always smooth nor monotonic. This can be explained by the fact that realistic disk galaxies, even if simulated, can have complex structures, including over- and underdense regions across their disks. Namely, TNG50 galaxies are certainly not akin to idealized smooth exponential disks: this is reflected in their more complex radial and vertical stellar density distributions, which in turn may simply be not well described by parametric functions (see Sections 5.2.3 and 5.2.3). We notice that these effects are more accentuated towards the disks' outskirts, whereby, together with sparser populations of stars, the error bars on the scaleheight estimates get typically larger.

5.4.2 A new non-parametric and generally-accessible quantification of disk flaring

As TNG50 returns a wide variety of disk flaring, a unique prescription to quantify it would be inadequate. We have seen, for example, that the radial trends of scaleheights can be both exponential and linear. Past works have fitted the flaring with an exponential formula only (López-Corredoira et al., 2002, Grand et al., 2017) or with a linear function only (Evans et al., 1998, Alard, 2000, García de la Cruz et al., 2021, Lian et al., 2022). The TNG50 phenomenology suggests that a non-parametric quantification of the flaring is of the essence.
We propose a quantification of the flaring that is independent of the shape of the flaring (linear, exponential or otherwise) and that can be applied in the most general manner to any data, so long as the latter allows the estimate of the disk scaleheight at two different galactocentric distances. Namely, we advocate for a quantification of the flaring based simply on the *relative enhancement* of disk heights at two locations (inner and outer disk), i.e. the difference between the scaleheights at two fixed galactocentric distances divided by the height in the innermost location. Now, as discussed in §5.3.1, stellar disks can span a wide diversity of extents even in a relatively narrow range of galaxy mass: $\sim 1.5 - 17$ kpc. This hence requires to quantify the amount of flaring upon normalizing the galactocentric distance by the scalelength of each MW/M31-like galaxies. We then propose and evaluate the amount of flaring in between (1-5)×$R_{\rm d}$ (see also Minchev et al., 2014a, for a similar approach):

$$\tau_{\rm flare} = \frac{h_{z,5\times R_{\rm d}} - h_{z,1\times R_{\rm d}}}{h_{z,1\times R_{\rm d}}}\,, \tag{11}$$

where $h_{z,r}$ denotes the vertical scaleheight in a narrow radial annulus at distance r, in our case according to §6.2. A $\tau_{\rm flare}$ equal to 1 (3) means that the disk height is 2 (4) times larger at $5 \times R_{\rm d}$ than at the inner radius.

5.4.3 What flares more? Young or old stars?

We proceed our analysis by quantifying the degree of flaring of the stellar disk populations as per Eq. 11, separately for old and young stars, i.e. with stellar ages of $8-10$ and $2-4$ Gyr, respectively. As noted before, the vertical distributions of stars in relatively narrow age intervals are well described by a single-component formula (Eq. 9), which hence gives the corresponding scaleheight, in agreement with previous simulations (Martig et al., 2014, Minchev et al., 2015, Ma et al., 2017) and observations (e.g. Bovy et al., 2016, assuming the chemistry as a good proxy for ages).

The main results of this paper are shown in Fig. 31. In the top panel, we show the comparison between the flaring of young and old stellar populations across 159 TNG50 MW/M31-like galaxies[4].

The black dashed line separates galaxies in which old stars flare more than the young ones (above the bisector) or vice versa, young stars flare more than the old ones (below the bisector): according to TNG50, a slight majority of MW/M31-like galaxies (89) galaxies, i.e. 56 per cent) have stellar disks whereby the young stars flare more than the old ones. However, a non-negligible fraction of TNG50 galaxies display a similar amount of flaring between young and old stars, settling on top of (or very close to) the dashed line. For the average or typical TNG50 MW/M31-like galaxy, young and old stellar populations exhibit scaleheights in the outer disk that are $\sim 1.5-2$ times larger than those in the inner part of the disk (median values of 1.44 and 1.66 for young and old stars, respectively; see inset of the main panel of Fig. 31).

A few galaxies populate the bottom right and top left corners, where the flaring of the young stars is considerably more pronounced than of the old stars, or vice versa. Still, across the studied TNG50 MW/M31 sample, slightly more frequent are galaxies where the young, rather than the old, stars reach high levels of flaring, e.g. $\tau_{\rm flare} \gtrsim 4$, corresponding to scaleheights at large distances that are $\gtrsim 5$ times larger than in the inner disk regions. Young stars show a somewhat broader diversity of flaring (with a weak peak at $\tau_{\rm flare} = 1-4$), whereas the old stellar populations exhibit a narrower distribution concentrated around $\tau_{\rm flare} = 0-3$.

To visualize how the flaring quantification corresponds to diverse vertical structures, we show the change of scaleheight as a function of galactocentric distance for three galaxies, each representing different regions of the $\tau^{\rm young}_{\rm flare} - \tau^{\rm old}_{\rm flare}$ plane. A series of stellar light images of TNG50 MW/M31-like galaxies with substantial flaring are shown in Fig. 32.

Importantly, a few galaxies from TNG50 seem to reproduce the flaring phenomenology inferred observationally for our Galaxy. The vertical magenta areas in the panels of Fig. 31 represent the flaring (evaluated using Eq. 11) of the "young" stellar population of our Milky Way, with the scaleheights inferred from the vertical-action values estimated in Ting & Rix (2019). In the iso-thermal regime, a population of stars with a mean vertical action $\langle J_z \rangle (R,t)$ has vertical distribution $p(z) \sim {\rm sech}^2(\frac{z}{2h_z})$ with a scaleheight $h_z(R,t) = \sqrt{\frac{\langle J_z \rangle (R,t)}{2\nu_z(R)}}$, where $\nu_z(R)$ is the local vertical frequency which we determine with Galpy (Bovy, 2015) in the same way as Ting &

[4]In fact, in 39 of the 198 TNG50 MW/M31, one or both the stellar age bins are not sufficiently populated to ensure a good vertical profiling and fitting of the stellar vertical mass distribution.

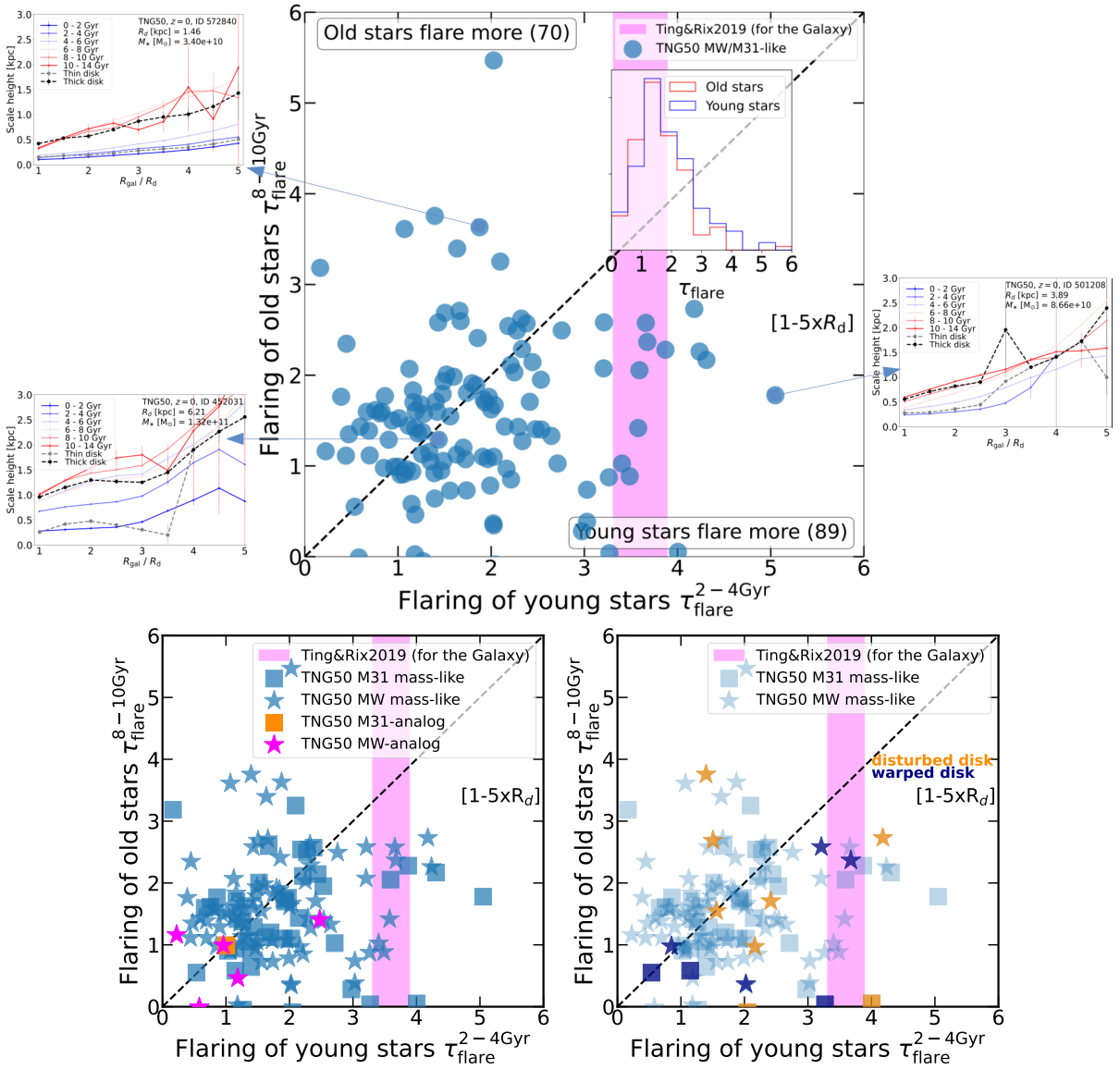

Figure 31: **Flaring of young vs. old stellar populations in TNG50 MW/M31-like galaxies.** We compare between the flaring of young (2-4 Gyr) and old (8-10 Gyr) stellar populations. The black solid line separates galaxies in which old stars flare more than the young ones (above the bisector) or vice versa (below the bisector). The vertical magenta areas indicate the flaring of young stellar population of our Milky Way, inferred from Ting & Rix (2019) – see the text for more details. We have selected three galaxies from three different regions of the plane and show in the small panels their scaleheights vs. radius. In the bottom left panel, TNG50 MW/M31-like galaxies are separated according to their stellar mass, above (squares, M31-mass like) and below (stars, MW-mass like) log $(M_{\rm stars}/{\rm M}_\odot) = 10.9$. Moreover, galaxies dubbed as MW-analogs (i.e. having stellar mass, scalelength, thick and thin scaleheight similar to the Milky Way) are depicted with pink stars. In the bottom right panel, we highlight the cases in which galaxies show a warped (in navy) or disturbed (in orange) stellar disk, showing that these visually-identified features do not seem to systematically bias our quantification of flaring.

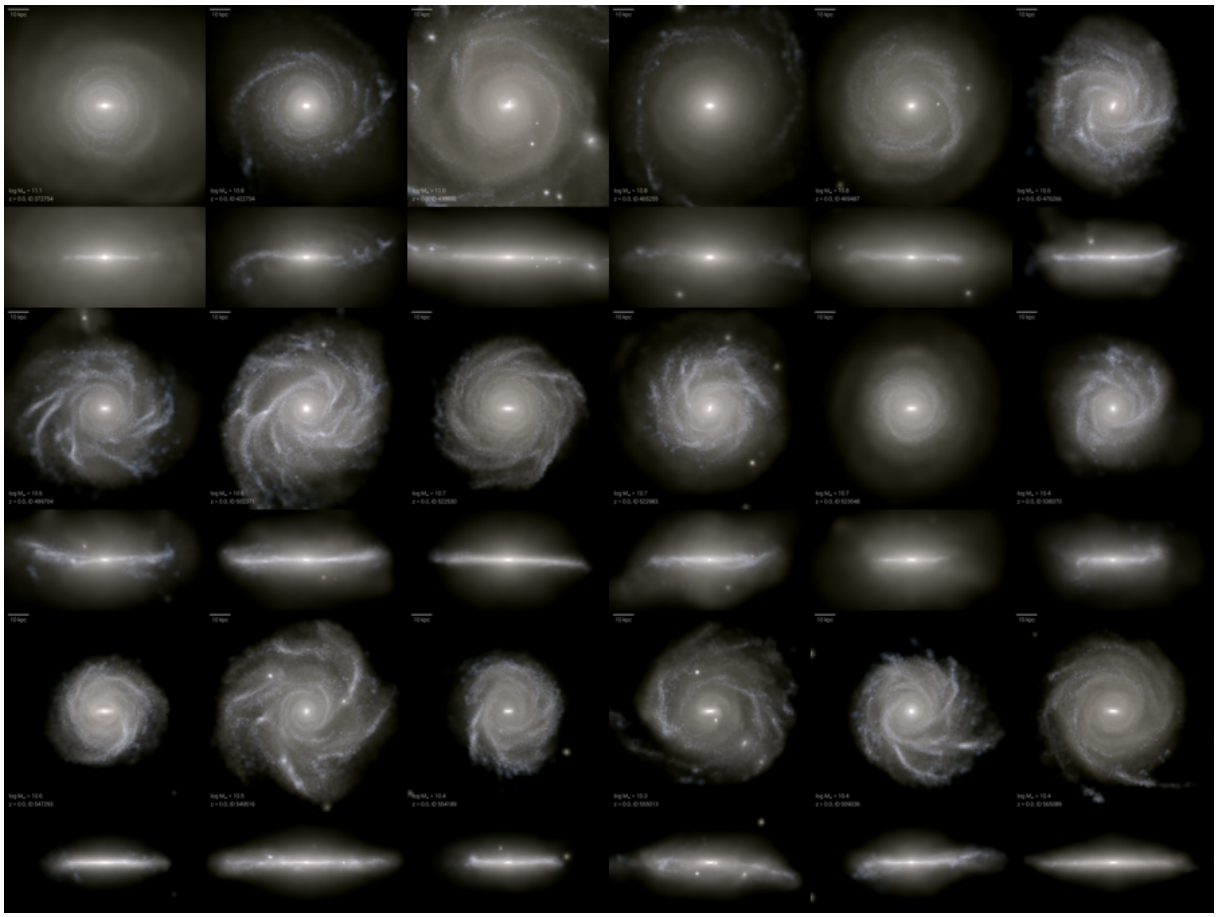

Figure 32: **Stellar-light composite images of selected TNG50 MW/M31-like galaxies**. We show the face-on and edge-on projections of 18 MW/M31 analogs from TNG50 at z = 0 that exhibit substantial flaring (i.e. $\tau > 3$). As it may be a characteristic of subsets of the stellar populations, flaring may not be visually manifest in stellar maps (whether light-based, as here, or mass ones).

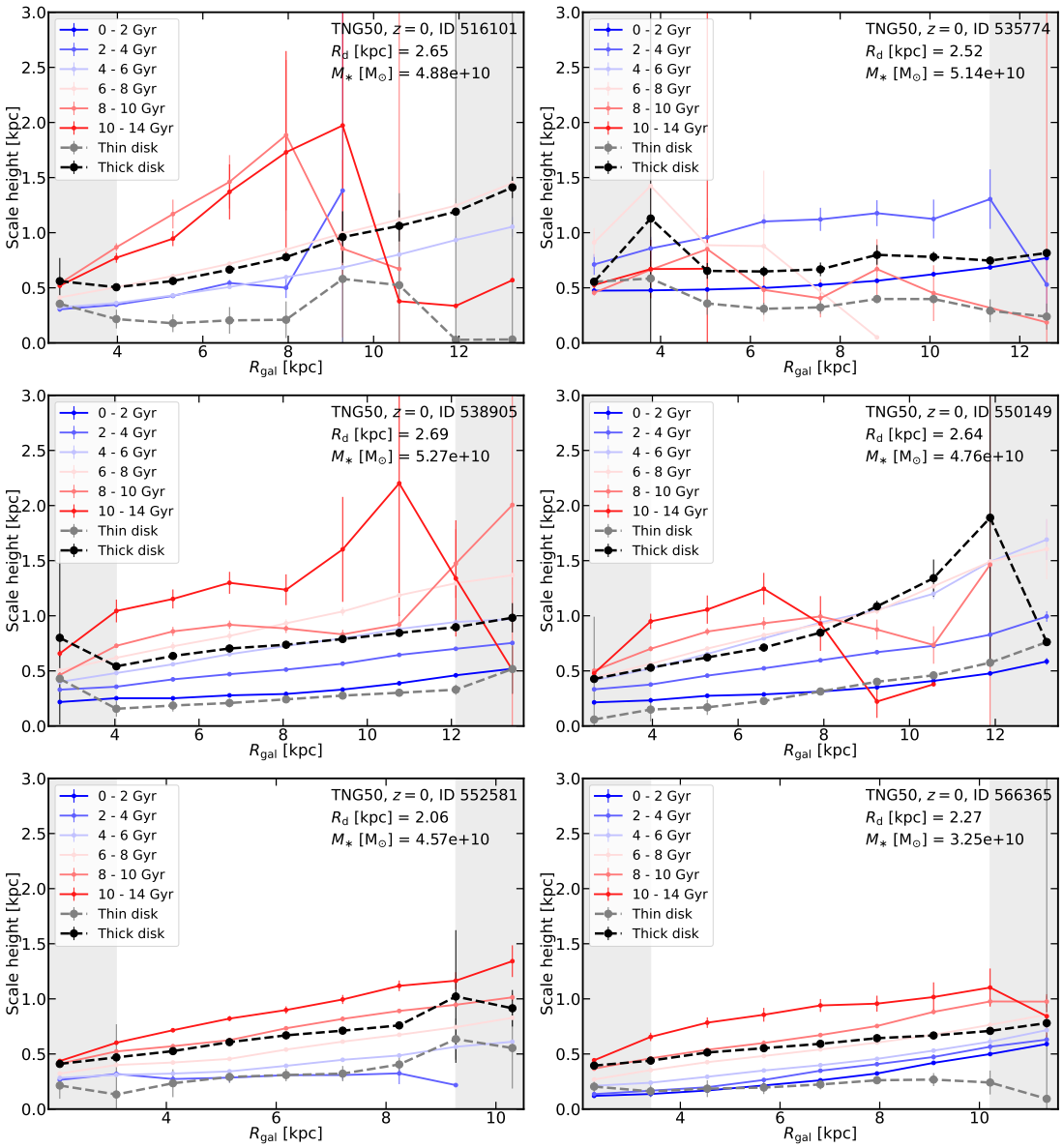

Figure 33: **Vertical disk structure and flaring of six TNG50 MW-like galaxies.** The scaleheight as a function of the galactocentric distance (normalized by the scalelength) is shown for the six MW analogs with stellar mass and disk properties most similar to the Galaxy. Error bars represent one standard deviation errors of the parameters, provided by the fitting function.

Rix (2019). Therefore, assuming a vertical stellar density distribution proportional to sech^2 and a Milky Way gravitational potential from Bovy (2015), we can infer a level of flaring for our Galaxy. About seven galaxies in the TNG50 sample show a flaring similar to the one inferred observationally for the Galaxy.
In particular, a good fraction of TNG50 galaxies with flaring similar to the Galaxy also have similar galaxy stellar mass. In the bottom left panel of Fig. 31, we show the same plot as in the top but with different symbols denoting different subsamples of the TNG50 MW/M31-like galaxies: squares indicate M31-mass objects (53 in total, $\geq 10^{10.9}$M$_\odot$), stars indicate the 106 MW-mass galaxies. The magenta star (orange square) symbols represent the MW analogs (M31 analogs) identified in §5.3, i.e. the six (one) galaxies that have detailed stellar disk structural properties consistent with the Galaxy (Andromeda). Within TNG50, there is no simulated galaxy with the same stellar disk structure, including extent, thickness, and flaring, of the Galaxy. However, our Galaxy represents one among many realizations of disky galaxies and, in terms of flaring of the young stars per se, according to TNG50, it appears rather common.
Finally, the right bottom panel of Fig. 31 is meant to convince us that the general picture depicted so far is not systematically affected or biased by cases of warps or disturbed disks (Section 5.3.4). These are highlighted in blue or orange, based on the visual inspection presented above. Although our measurements are all azimuthally-averaged, distorted disks could potentially imply an under or over estimation of the amount of flaring. There is no manifest bias toward large or small degrees of disk flaring when we focus on warped and disturbed stellar disks, namely they populate all regions of the depicted space as the rest of the population. However, the few identified warped disks seem to tend to have a young stellar population that flares more than the old one.

5.4.4 The cases of TNG50 galaxies with stellar disk properties compatible with the Galaxy's

As already mentioned in Section 5.3.1, in the TNG50 MW/M31 sample there are six galaxies with stellar mass and disk sizes similar to the Milky Way and one galaxy with disk properties similar to Andromeda. For the MW-analogs, and to connect more directly with the observational opportunities in our Galaxy, we show in Fig. 33 the change of the scaleheight as a function of galactocentric distance.
As already pointed out throughout this work, even for galaxies sharing the main disk properties, the flaring can be qualitatively very diverse. Indeed, for two MW-analogs, Subhalo IDs 538905 and 566365, the flaring is quite linear, whether we consider mono-age stellar populations (colored curves) or all disk stars (black and grey curves); on the other hand, the galaxy with Subhalo ID 552581 shows an exponential flaring. Additionally, the MW-analogs 516101 and 535774 show more irregular trends where mono-age population and the thin and thick disks flare quite differently. Once we evaluate the degree of flaring in the way proposed in this work (i.e. by using Eq. 11), the relative enhancement between 1 and 5R_d turns out to be diverse (see bottom left panel of Fig. 31, magenta stars with black contours): for two of them, young stars flare more than old ones, with different level of flaring. Additionally, we have an analog where young and old stars flare equally, and an analog where old stars flare much more than young stars.

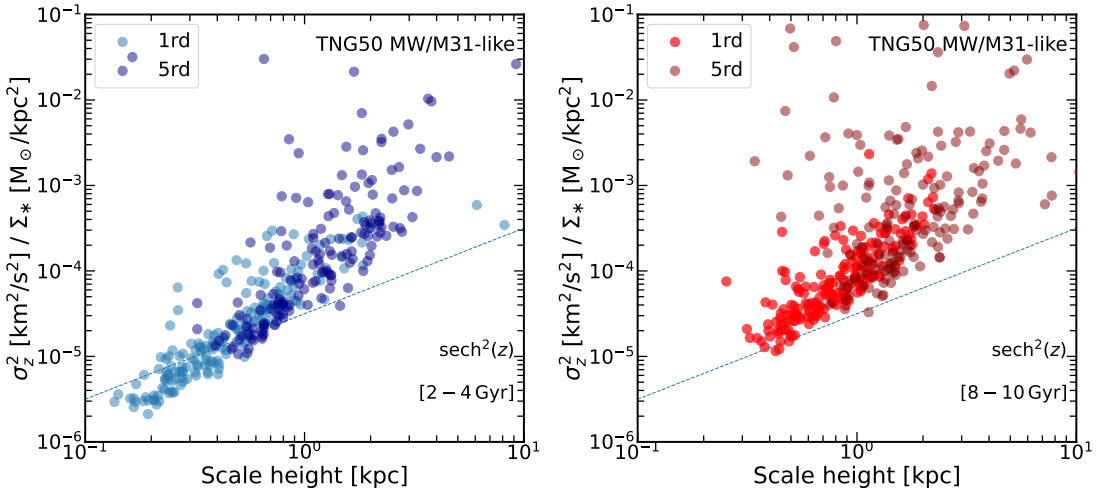

Figure 34: **Relationship between stellar heights and stellar kinematics for TNG50 MW/M31-like galaxies**. We plot the ratio between the squared local stellar vertical velocity dispersion and the stellar surface density vs. stellar disk scaleheight, for young (left) and old (right) stellar populations. We do so at different locations within the stellar disk. The dashed line defines the linear correspondence that is expected for an idealized self-gravitating disk.

5.5 Disk flaring and kinematics in TNG50

Our fiducial quantification and definition of disk flaring is based on a geometrical estimation of the stellar disk scaleheights. However, the structural properties of galaxies are expected to be the global manifestation of the underlying orbital configuration and interaction of all present matter components. In Pillepich et al. 2019, we showed that the vertical structure of both the stellar and gaseous components of TNG50 star-forming galaxies across epochs indeed is the resolved outcome of an ensemble of physical ingredients – chiefly, the shape and depth of the overall gravitational potential, which in turn is the result of the interaction and orbital mixture of both collisional and collisionless material. Here we expand upon that analysis by focusing on MW/M31-like galaxies, and link the flaring with the kinematics of disk stars: we hence examine their vertical velocity dispersion.

From a theoretical point of view, the scaleheight for an isothermal sheet is related, above and below a certain position, to the local vertical velocity dispersion (σ_z^2) and the local stellar surface density (Σ_*), according to the relation (Spitzer, 1942): $h_z \propto \sigma_z^2/\Sigma_*$.

We show in Fig. 34 this connection for the disk stars of the TNG50 MW/M31 analogs, separating between star in the inner vs. outer disk (different shades of colors) and young vs. old stellar populations (left vs. right panels, respectively). Here we adopt our fiducial choice for the heights measurement as in the previous Sections.

A number of interesting considerations can be made. Firstly, despite the complexity of the realistic disks realized by TNG50 and even though stars are not the only matter components in the disk regions, the stellar disk heights of both young and old stellar populations relate to the underlying vertical velocity dispersion and average mass surface density: the higher the σ_z^2/Σ_* ratio, the thicker the disk. However, and secondly, the relation may be not perfectly linear and is steeper than $h_z \propto \sigma_z^2/\Sigma_*$ (dotted lines): this is probably due to the gas and dark matter contributing to the disk potential and also to the morphological complexity of a real galactic

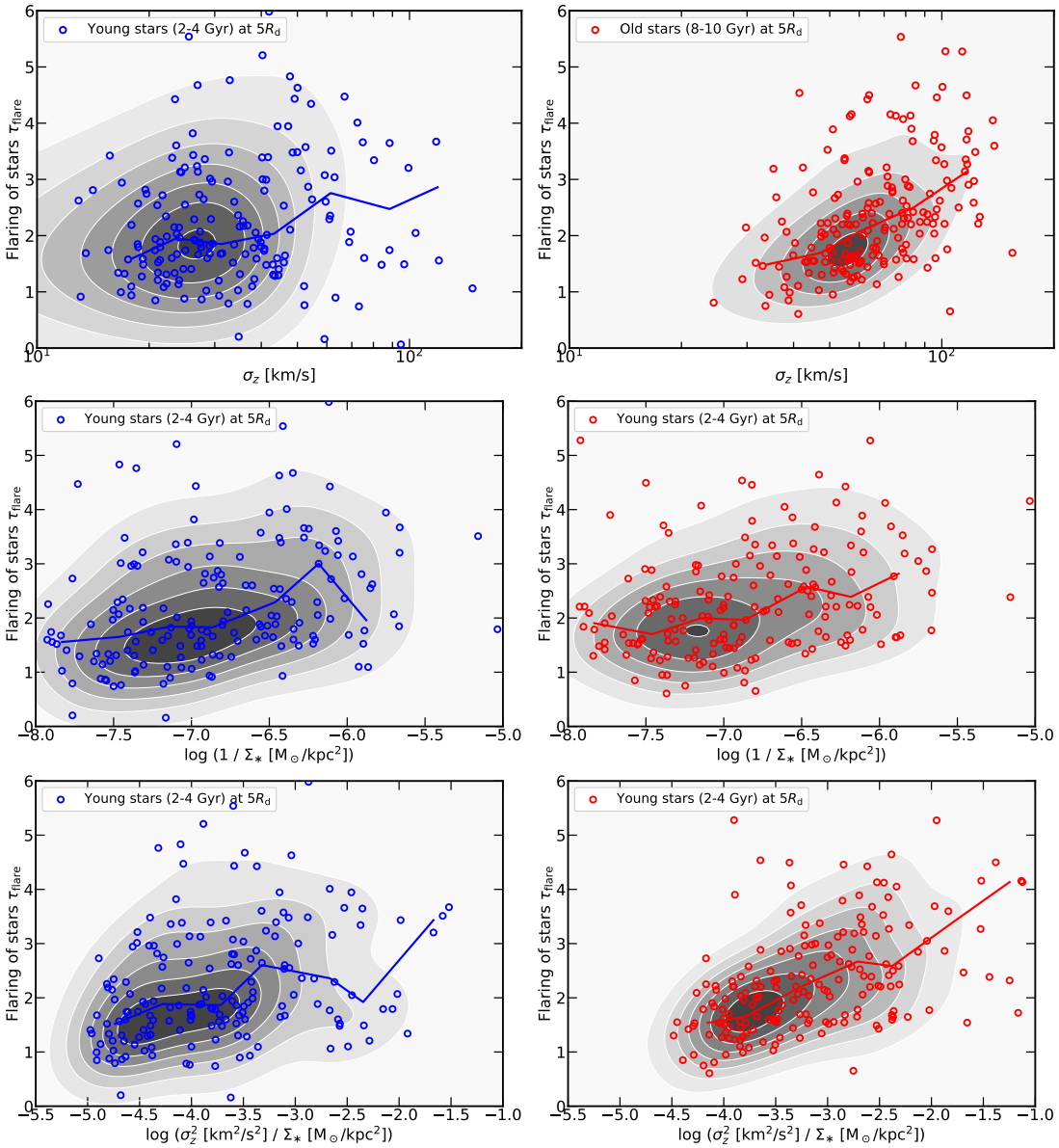

Figure 35: **Disk flaring vs. kinematic properties for TNG50 MW/M31-like galaxies**. We plot the amount of flaring of young (left column, in blue) and old (right column, in red) stellar populations as a function of the stellar vertical velocity dispersion at $5R_{\rm d}$ (top) and the inverse of the stellar surface density at $5R_{\rm d}$ (middle), and finally their ratio (bottom). In each panel, the galaxy number density is estimated with a Gaussian kernel and represented with the shaded contour areas. Solid curves are medians in bins of the quantity on the x-axes. Galaxies with hotter or less dense outer stellar disk flare more.

stellar disk compared to the idealized isothermal sheet.
We have quantified the same relation by using the half-mass heights of the disks instead of the best fit to a squared hyperbolic secant function, as a more robust estimate of the scaleheight against disk internal structure and inhomogeneities: we notice, although we do not show, that the scatter in the relationships of Fig. 34 at fixed galactocentric distance is considerably reduced in such a case. This suggests that the galaxy-to-galaxy scatter in Fig. 34 is not all due to physical effects and the most severe cases of galaxy outliers are those where the parametric functional forms are not a very good description of the vertical stellar mass distribution in the disk.
Fig. 34 shows that, also according to TNG50 and as argued in previous works, at fixed galactocentric distance in the disk, older stars are not only thicker but also hotter than younger ones, i.e. exhibit overall larger velocity dispersions (left vs. right panel). The same phenomenology has been measured in the Milky way (Nordström et al., 2004, Dorman et al., 2015) and also in other cosmological simulations (Buck et al., 2020).
We also find, although do not show, that at fixed galactocentric distance, the stellar disk heights exhibit a correlation with both the local stellar velocity dispersion and the local stellar mass surface density, individually: namely, stellar disks are thicker when they are hotter or less dense. But so, what does determine the flaring of the stellar heights?
In Fig. 35, we show how the level of disk-height flaring of TNG50 MW/M31-like galaxies depends on stellar velocity dispersion (top), inverse of the stellar mass surface density (middle) and the ratio of the two, i.e. σ_z^2/Σ_* (bottom), whereby these quantities are evaluated in the outer disk regions. To avoid any issue with fitting the vertical mass profiles with parametric functions, here we quantify the flaring by measuring the stellar half-mass heights (see also §5.6). We also quantify the flaring and its relationship with stellar kinematics and potential for young and old stellar populations separately: left vs. right panels of Fig. 35.
From this analysis we can see that, according to TNG50, galaxies with hotter or less dense outer stellar disks flare more strongly, in both young and old stellar populations. By comparing the top to the middle panels, we can also conclude that the diversity in flaring is mostly driven by the diversity in outer-disk temperatures, i.e. in stellar kinematics, rather than by a diversity in stellar surface density.

5.6 Discussion

5.6.1 On other methods to quantify the disk flaring

One of the main messages of this paper is that, not only across differently-simulated galaxies, but also within one given simulation, the nature and amount of disk flaring can be very diverse, even for disk galaxies within a narrow range of galaxy stellar mass and environments. We hence advocate for a non parametric quantification of flaring, as provided in §5.4.2.
One possible drawback of this proposition is that, at least in our fiducial implementation, it still relies on a parametric measure of the stellar disk heights. In §5.2.3, we have described different methods that are commonly used to quantify disk heights and all results in the previous sections, unless otherwise stated, are based on our fiducial choice: profile fitting with single and double squared hyperbolic secant, for mono-age stellar populations and all-ages disk stars, respectively.
We show now in Fig. 36 how the flaring values for young and old stellar populations depend on the method for height measurements. We show again the main panel of Fig. 31 in the top left, unchanged, and repeat it for the additional parametric cases in the top right (hyperbolic secant)

and bottom left (exponential). The bottom right panel shows the flaring measurement based on non-parametric half-mass height measurements, as in Fig. 35.
All four cases exhibit roughly similar flaring ranges and three of them also have similar distributions on the depicted plane (squared and single hyperbolic secant and half-mass heights), whereas for the exponential profiles the distribution appears more sparse and dissimilar. The fraction of galaxies where the young stars flare more than the old stars (and vice-versa) also changes from method to method, although not in qualitatively significant manners, barring the case of the exponential fits. The latter is the method that returns the most different qualitative and quantitative quantification of flaring. The quantification based on the non-parametric measurements of the stellar half-mass disk heights provides confidence that, in the population, fit-based assessments are not dramatically affected by possible numerical/fitting issues or by the possibility that parametric functional forms do not describe the stellar mass distributions well. On the other hand, individual galaxies may exhibit somewhat different amounts of flaring depending on the adopted method to measure their disk heights. Comparisons can hence only be done once the same operational definitions are adopted on all sides.
Considering that some of these galaxies may have a bar at a galactocentric distance of $1R_{\rm d}$, we have repeated the measurements of the flaring parameter in the range $(2-5)\times R_{\rm d}$ instead of $(1-5)\times R_{\rm d}$: the results are shown in the top left panel of Fig. 36, in grey. The values are smaller (as we can expect from the flaring phenomenon) but they are distributed similarly to the original definition on both sides of the identity line.

5.6.2 A note to observers: on the "flaring" based on the spatial distribution of stellar ages

As introduced in Section 6.1, it has become customary to inspect the vertical structure of the Galactic stellar disk by looking at the mean or median stellar ages as a function of galactocentric radius $R_{\rm gal}$ and as a function of vertical distance from the midplane $|z|$. This is typically done in observations of the Galaxy (Ness et al., 2016, Xiang et al., 2017, Feuillet et al., 2019) as well as with simulation data (Ma et al., 2017, Buck et al., 2020, Agertz et al., 2021).
The common emergent picture is that of a *funnel* shape, in which at each radius $R_{\rm gal}$, the mean age of disk stars increases with $|z|$, whereas at fixed $|z|$ the mean or median stellar age decreases with galactocentric radius. This can be appreciated in Fig. 37, where we have compiled examples from the literature, for the Galaxy (left column Feuillet et al., 2019, Xiang et al., 2017, Ness et al., 2016) and from MW-like zoom-in cosmological simulations (central column Agertz et al., 2021, Ma et al., 2017, Navarro et al., 2018), and where it can be seen that young stars at large galactocentric distances can indeed be found at large altitudes.
For comparison, on the right column of Fig. 37, we show five TNG50 MW/M31 analogs that are representative of the whole sample and that qualitatively reproduce previously-quantified observational and theoretical scenarios, for $R_{\rm gal}=3-14$ kpc and $|z|=0-4$ kpc. Here we divide all stars in 100×100 bins on the $|z|-R_{\rm gal}$ plane, and the colors denote the mean stellar ages in each pixel. Also in TNG50 MW/M31-like galaxies, the mean stellar age increases with Z at fixed $R_{\rm gal}$, and decreases with $R_{\rm gal}$, at fixed Z, but with different detailed patterns depending on the galaxy.
However, it is important to realize that, even if these funnel features have been often connected to the flaring of the disk, they in fact do not imply it. Flaring is the increase of the stellar disk thickness at increasing radii. The change with radius and height of the typical ages of the disk

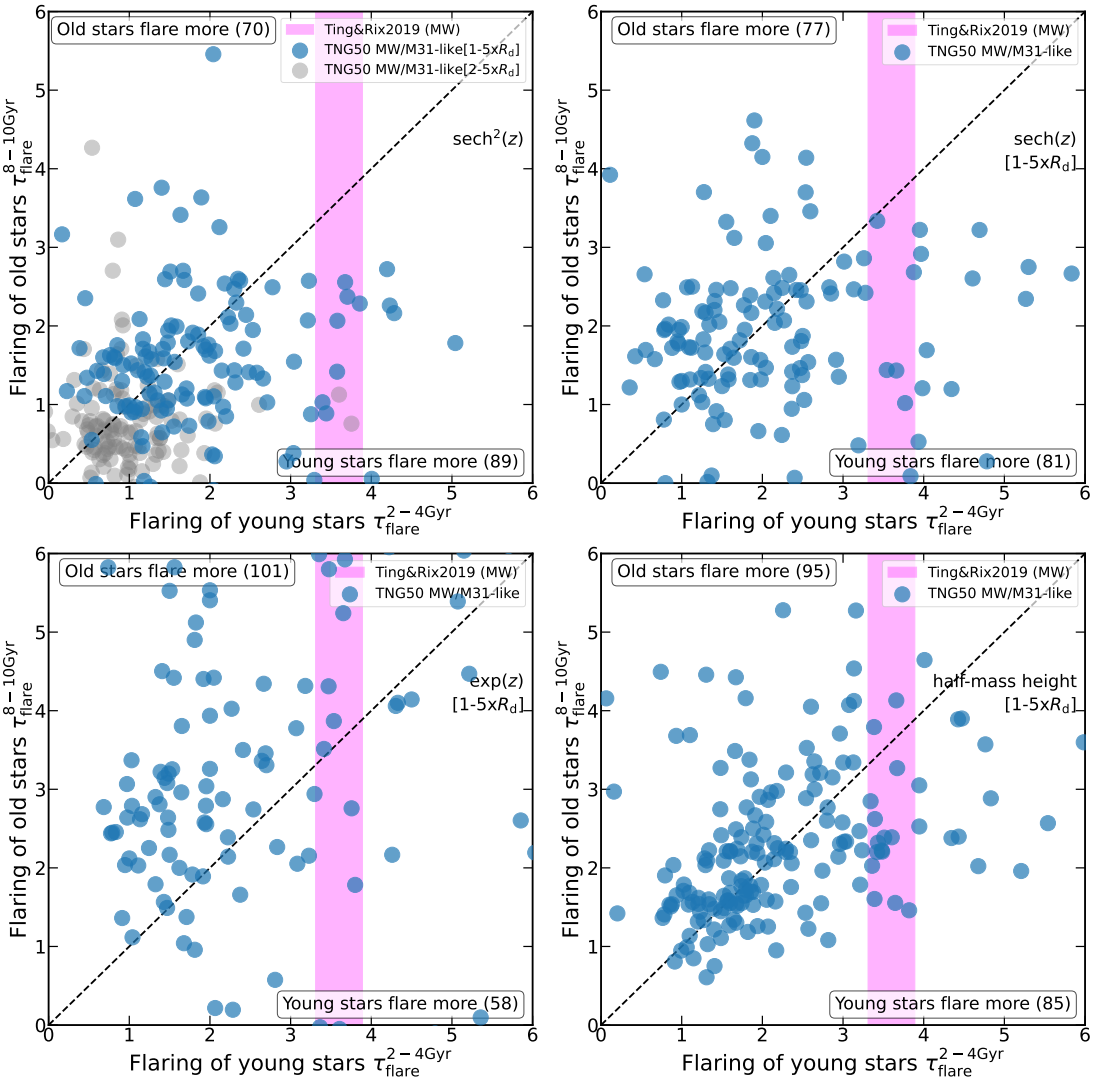

Figure 36: **Flaring of young vs. old stellar populations in TNG50 MW/M31-like galaxies for different methods to measure disk heights.** Top left: quantification of the flaring with the fiducial height measurements based on fitting a squared hyperbolic secant profile (as in all figures so far): blue data points refer to the flaring measured between 1 and 5×R_d (same as main panel of Fig. 31), whereas in grey the flaring is evaluated at 2 vs. 5×R_d. Top right: hyperbolic secant profiles. Bottom left: exponential profiles. Bottom right: half-mass heights. Estimating disk heights with the exponential fit returns, among the others, the most different results, both quantitatively and qualitatively.

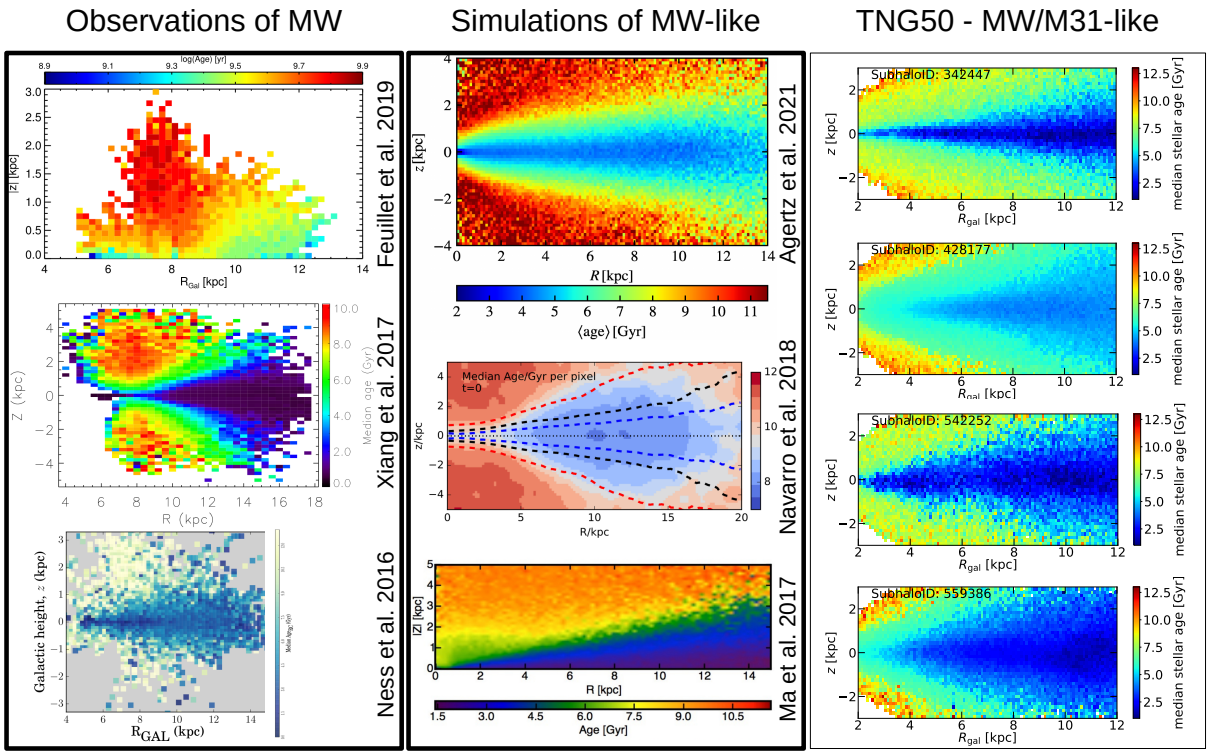

Figure 37: **Stellar age distributions in the MW/M31 midplane**: vertical distance from the midplane as a function of galactocentric distance with the color code representing the mean or median stellar age. Here we compare observations from Ness et al. (2016), Xiang et al. (2017), Feuillet et al. (2019) (left), simulations from Ma et al. (2017), Navarro et al. (2018), Agertz et al. (2021) (center), and a sample of five TNG50 MW/M31 analogs (right). Barring those in the right column, the plots are replicated from the respective research studies without modification.

stars *do not* imply flaring: an un-flared inside-out model could still produce the age maps with a funnel shape.
To demonstrate this we produce two MW disk mocks from the model described in Frankel et al. (2020) and inspect for both cases the stellar density and mean age as a function of radius and height (we do not show the plots here). Without flaring and inside-out growth, no funnel shape is present. However, when we activate inside-out growth, we can recover the same pictures as in Fig. 37.
To illustrate this further, we show in Fig. 38 the $R - z$ plane color-coded by mean age that result from two toy models, one with flaring and one without flaring. Despite the absence of flaring, the model without flaring exhibits the typical funnel structure purely as a result from inside-out formation and vertical heating. The model with flaring is adapted from the model described in Frankel et al. (2020), which builds on the best fit vertical distribution of Ting & Rix (2019). In summary, the model describes a star formation history of the disk from the inside-out, forming on an exponential radial profile, and the subsequent orbit evolution in the plane (radial migration and radial heating as diffusion in angular momentum and radial action) and out of the plane as a scale-height that increases with radius and time. The toy model without flaring has exactly the same formation scenario and parameters and the same in-plane orbit evolution, but the vertical distribution is now modelled as a scale-height that increases with time as $h_z(t) = 0.15 + 0.1(t/1\mathrm{Gyr})$ independent of radius. As it can be appreciated, a funnel-like distribution of stellar ages can be in place with or without actual flaring, i.e. with or without orbital changes of the stars.

5.6.3 Comparison to previous simulations

As we have already decribed in Section 5.1.5, even if the flaring of the stellar disk has been studied and investigated in a number of previous theoretical works in the literature, a comparison among them is non-trivial. This is due to the different ways to quantify the flaring and the diverse ways in which the flaring is manifest in the various simulated galaxies and simulation models. In this section we attempt a closer comparison between the TNG50 results and those from selected zoom-in simulations of MW-mass disk galaxies, such as NIHAO-UHD (Buck et al., 2020), Latte (Ma et al., 2017), VINTERGATAN (Agertz et al., 2021), simulations by García de la Cruz et al. (2021) and Auriga (Grand et al., 2017).
We note that in all the simulations used for this comparison, the scaleheight of the mono-age stellar populations that constitute the stellar disk is evaluated from a single exponential fit, which is not our fiducial choice – this is the case for all but for Latte and García de la Cruz, where a single sech^2 profile is used instead. We hence proceed as follows. From the available literature, we extrapolate the scaleheight of young and old stellar population at 4 and 12 kpc from the galactic center (to be consistent with all the selected models). Unless otherwise stated we try to stick to our definition of *young* (2-4 Gyr) and *old* (8-10 Gyr) age bins. Then, using Eq. 11 (but with the heights at the physical distances of 4 and 12 kpc instead of $1 \times R_\mathrm{d}$ and $5 \times R_\mathrm{d}$), we evaluate the amount of flaring τ_flare of each model and we plot them in concert with the TNG50 results in Fig. 39. In order to make an apples-to-apples comparison, we measure the scaleheights of also the TNG50 sample by fitting single exponential profiles to the stellar vertical density distribution.
In the left panel of Fig. 39, TNG50 is compared with NIHAO-UHD (black plus symbols), Latte (black diamond), García de la Cruz (pentagons) and VINTERGATAN (black crosses). We note

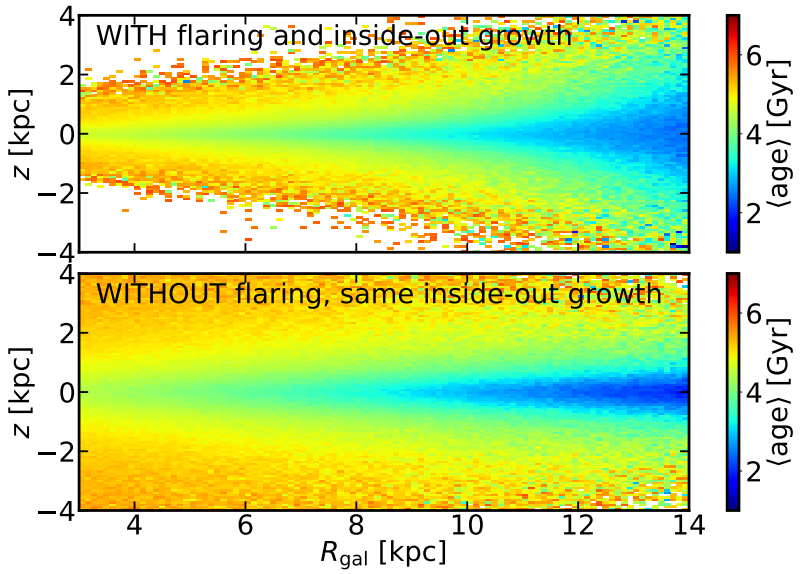

Figure 38: **Stellar mean age distributions for two MW-like disk mocks.** We illustrate the $R-z$ plane of two model galaxies color-coded by mean age (as in Fig. 37 from a toy disk model that includes flaring (top) and one that does not (bottom). The toy model is adapted from the best-fit of Frankel et al. (2020), Ting & Rix (2019). In the top panel, the vertical distribution of stars is a sech^2 function with an age- and radius-dependent scale-height as described in Frankel et al. (2020) that captures the flaring of the Galactic disk as described and fit by Ting & Rix (2019). In the bottom panel, the model is the same, but the vertical distribution of stars is replaced by a non-flaring toy distribution, where the scaleheight is essentially only a function of age. In particular, we took $h_z(t) = 0.15 + 0.1(t/1\text{Gyr})$ kpc a (not physically motivated, but simple) example. The funnel shape of iso-age contours in the bottom panel arises naturally from the combination of inside-out formation of the disk and subsequent vertical heating, not from disk flaring.

that in Latte, the "old" stellar population is composed by all stars older than 8 Gyr while in VINTERGATAN the bins of stellar ages are different from the one used in this work, being Δage = 1 Gyr. To be consistent, we then plot both of them, a thick cross denoting stellar populations of 3-4 Gyr and a thin cross representing stars of 2-3 Gyr.
Because of the diverse treatment of the flaring of the mono-age stellar populations adopted for the Auriga galaxies, we choose to separate this comparison from the others, and show it in the right panel of Fig. 39. In this case, we apply to our TNG50 sample the same choices as in the Auriga's paper. Indeed, here the young stars are defined to be all disk stars below 3 Gyr and their flaring is shown against all disk stars. For the 30 MW-like galaxies in the Auriga sample, the vertical density distribution of the disk is fitted at a series of different radii with a single exponential profile. However, as already mentioned in §5.3.2, the vertical profiles of the TNG50 MW/M31-like galaxies are more often better described with a double functional than a single functional profile when all disk stars are considered. Nevertheless, also in this case we measure the heights for the TNG50 galaxies by fitting a single exponential profile. As in the left panel, the flaring is evaluated between 4 and 12 kpc.
As it is clear from Fig. 39, the TNG50 MW/M31-like sample returns and brackets all the other theoretical findings, including the most extreme cases: very small flaring of the old population, as in Latte, or cases where the old stars flare much more than the young ones, as in one of the NIHAO-UHD galaxies. We see also that in the TNG50 sample we have galaxies where the flaring values are larger than in previous simulations. The consistency of the outcomes in Fig. 39 is indeed a remarkable result. Until now it was not possible to say whether the diversity of flaring manifestations predicted by simulations was a genuine manifestation of galaxy-to-galaxy variation or was due to different numerical codes, galaxy-formation models, galaxy/halo selection or even different numerical resolution. The results with TNG50 demonstrate that MW/M31-like galaxies can exhibit very diverse levels and flavours of flaring simply due to galaxy-to-galaxy diversity.

5.6.4 Disk flaring vs. $z = 0$ structural and global properties of MW/M31-like galaxies

In §5.5, we have shown that the flaring of the stellar disk thickness is a direct manifestation of larger stellar velocity dispersions and lower stellar surface densities in the outer disk regions. Are there other global and/or structural properties of galaxies that correlate with the disk flaring? In this section we proceed to see whether or not the flaring of the mono-age stellar populations is connected with the global or structural $z = 0$ properties of the TNG50 MW/M31-like galaxies. For the sake of clarity, we divide the $z = 0$ properties in two subgroups: disk structural properties, including gas mass fraction in the disk region (Fig. 40) and mass properties (Fig. 42). In each figure, the amount of flaring of young (left columns) and old stellar populations (right columns) are shown separately, with lines representing the running medians.
From Fig. 40, we see that there is no clear trend or correlation between the degree of flaring and the scalelength of the disk. On the other hand, a clearer positive trend appears when we plot the dependence of flaring with the scaleheight of the stellar disk, both for the young and for the old stars, albeit with a significant scatter: namely, galaxies with thicker stellar disks appear to host a larger amount of disk flaring. As Fig. 42 shows, the current galaxy stellar mass, stellar disk mass and disk-to-total mass ratio of a MW/M31-like galaxy are not predictors for disk flaring (even though this statement may differ across a larger range of stellar or halo masses). However, interestingly and related to the discussions of Section 5.5, larger gas mass fractions in the disk

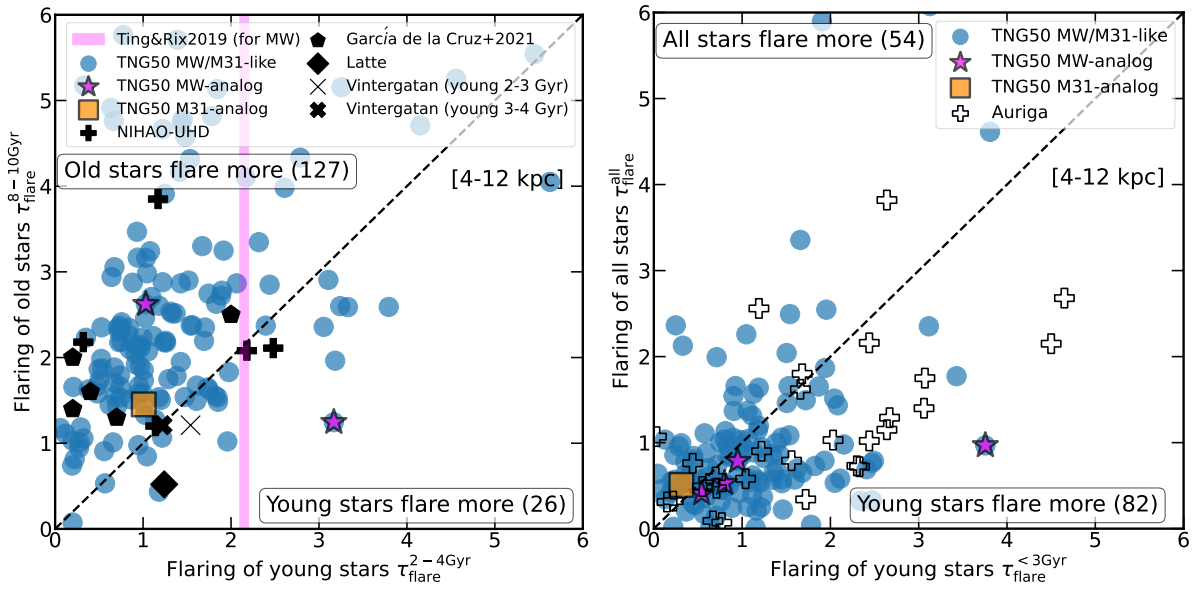

Figure 39: **Flaring of TNG50 MW/M31 analogs in comparison to the results of other cosmological MW-like galaxy simulations**. We show again the flaring of old and young stars for TNG50 MW/M31 analogs in comparison to, on the left, simulations from Agertz et al. (2021, VINTERGATAN: thin and thick crosses, with young stellar populations of 2-3 Gyr and 3-4 Gyr, respectively), Ma et al. (2017, Latte: diamonds), García de la Cruz et al. (2021, pentagons), and Buck et al. (2020, NIHAO-UHD: plus symbols); on the right, simulations from the Auriga project (Grand et al., 2017). The magenta area denotes the stellar flaring of young stars of the Milky Way extrapolated from Ting & Rix (2019), as in Fig. 31. We note that in all the simulation models used for this comparison (except for Latte that uses sech^2), the scaleheight of the mono-age stellar populations in the disk is evaluated from a single exponential fit. In these plots, the heights are hence measured using exponential profiles also for TNG50.

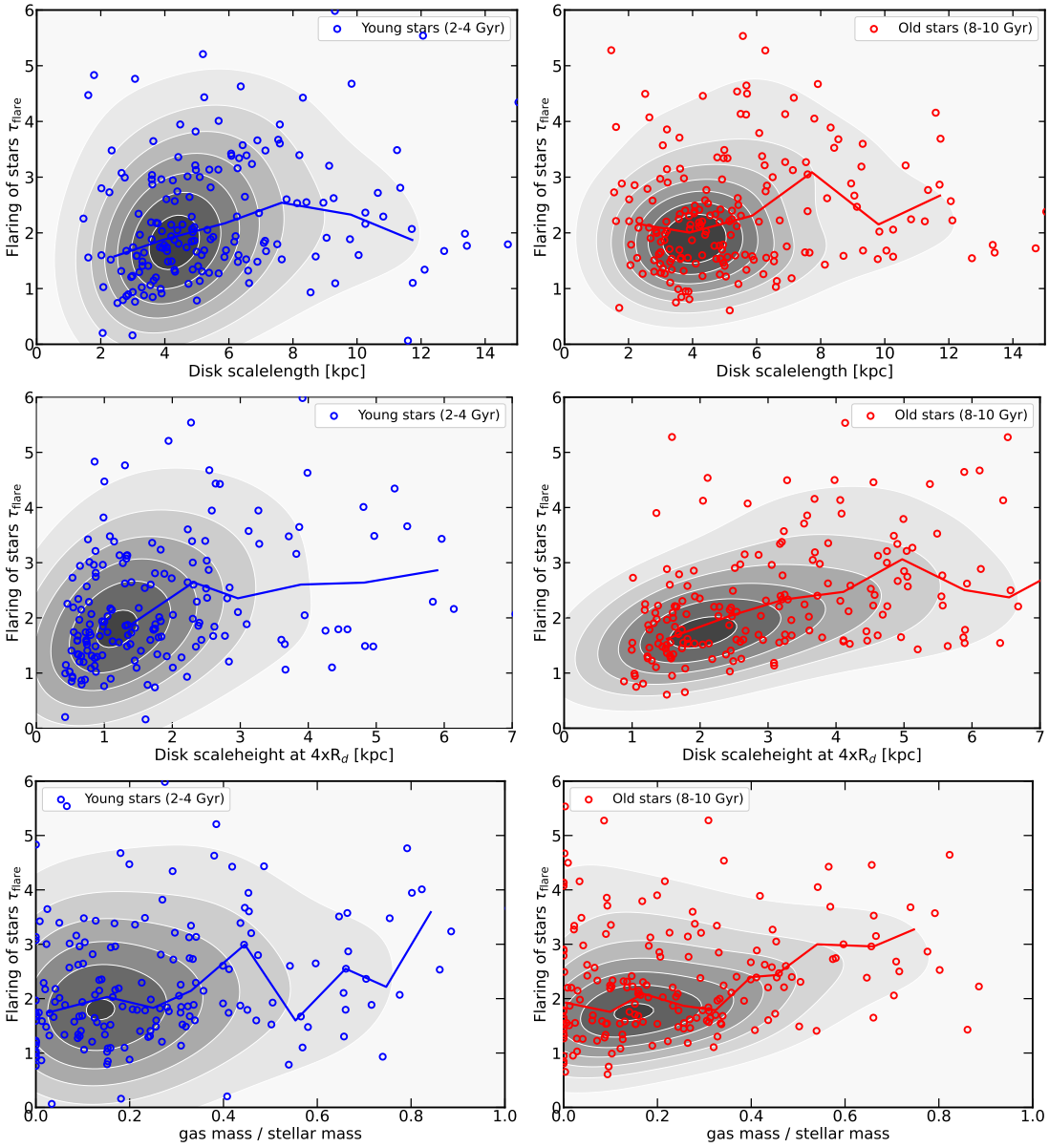

Figure 40: **Disk flaring vs. disk properties for TNG50 MW/M31-like galaxies.** We plot the amount of flaring of young (left column, in blue) and old (right column, in red) stellar populations as a function of stellar disk scalelength (top), disk scaleheight at $4\times R_{\rm d}$ (middle) and gas fraction within the disk (bottom). In each panel, the galaxy number density is estimated with a Gaussian kernel and represented with the shaded contour areas. Solid curves are medians in bins of the quantity on the x-axes. Galaxies with thicker stellar disks and larger gas disk fractions seem to be characterized by somewhat larger degrees of disk flaring.

of these galaxies imply larger degrees of flaring.
We confirm, although do not show, that the findings above hold irrespective of whether the flaring is evaluated based on stellar half-mass heights or scaleheights from parametric fits of the vertical stellar mass distribution.
Finally, we have compared the amount of flaring for young and old stellar populations with additional galactic and environmental properties (we omit the corresponding plots for brevity). The proportion of barred galaxies is not significantly different in the two subsets of systems where the young population flares more than the old stellar populations or viceversa. The presence itself of a bar does not seem to correlate either with the amount of flaring. We have examined whether the presence of other (satellite) galaxies in the proximity of our MW/M31 analogs at $z = 0$ may be a predictor of, or may be associated to, larger degrees of flaring – we cannot discern any statistically-robust trend.

5.6.5 Disk flaring vs. merger histories of MW/M31-like galaxies

Is the amount of disk flaring determined by the past merger history of a galaxy?
As already mentioned in the Introduction, our TNG50 MW/M31-like sample is built with no constraint on past history, so that the $z = 0$ MW/M31-like galaxies from TNG50 are the result of very diverse merger histories. In Fig. 41, we investigate whether the amount of flaring is correlated with aspects of the past merger history. To this aim, we highlight those galaxies that have undergone at least one major merger in the last 2 Gyr (red), 5 Gyr (orange) and since $z = 1$ (navy). There seems to be a trend whereby MW/M31-like galaxies that underwent recent major mergers are more likely to have young stars flaring more with respect to the old population. To test the statistical significance of this statement (for the case of the major merger since $z = 1$, where the counts are higher), we perform two sample tests of proportions, obtaining a p-value of 0.12 for both the Z-test and the chi-square test: this means that the trend that we see Fig. 41 is only weak.
Now, we can speculate that even if stars younger than 2 Gyr were not born yet at the time of the major merger, they were born later in a "perturbed" stellar disk likely heated-up by the merger event. Alternatively, as discussed in Ma et al. (2017), the young stars flare more because they inherit the flaring of the gas from which they formed, being this possibly perturbed by the recent mergers. However, the occurrence of recent or less recent major mergers does not seem to imply overall systematically stronger disk flaring. In fact, we have not identified trends between disk flaring and a number of merger-history statistics, including the number of all major and minor mergers, the ex situ total mass and the ex situ fraction of the galaxy.

5.6.6 On possible resolution effects

Before closing, we offer a few remarks on the possible effects of numerical resolution. In this paper we have characterized the stellar disk structures and flaring predicted for MW/M31-like galaxies by one simulation only, the highest-resolution run of the IllustrisTNG project: TNG50. As shown in Pillepich et al. (2019), stellar disk heights are affected by the underlying resolution: TNG50 galaxies are on average thinner than those in the 8-times lower mass resolution counterpart TNG50-2: see their figures B2 and B3. However, by how much galaxies are thinner in the higher-resolution simulation in comparison to lower-resolution runs depends on redshift and stellar mass range. With this in mind, we notice that all the results shown in this paper are

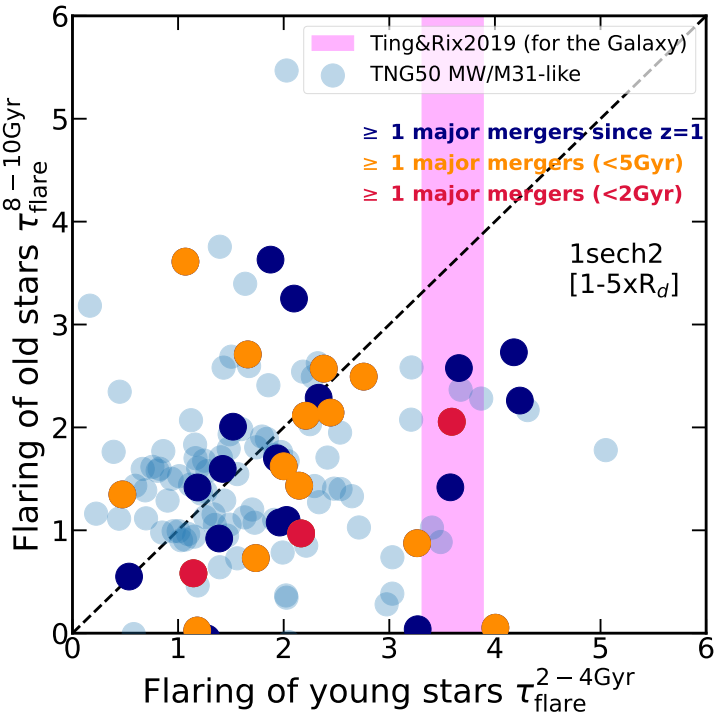

Figure 41: **Flaring of young vs. old stellar populations in TNG50 MW/M31-like galaxies in connection to their past merger history.** Here we show the same results, layout and method of measuring heights as in Fig. 31, but we additionally highlight the different merger histories: i.e. galaxies that have undergone at least one major merger in the last 2 Gyr (red), in the last 5 Gyr (orange) and since $z = 1$ (navy). Whereas a recent or ancient major merger does not seem to systematically influence the degree of flaring, at the galaxy population level, it would appear that galaxies with recent major mergers are more likely to host young stellar populations that flare more than their older counterparts.

based on the relative enhancement, $\tau_{\rm flare}$, of scaleheights across the simulated stellar disks. It is therefore plausible that any effect of numerical resolution on the values of the disk scaleheights should be mitigated by this definition. Furthermore, the levels of disk flaring predicted by the TNG50 simulation are consistent with those returned in MW analogs simulated at much better numerical resolution, such as Latte and Vintergatan. This gives further confidence on the quantitative soundness of the results provided herein.

5.7 Summary and conclusions

In this paper we have presented a comprehensive study of the stellar disk structure of a large, unbiased sample of Milky Way (MW) and Andromeda (M31)-like galaxies from TNG50, the highest-resolution cosmological magneto-hydrodynamical large-volume simulation of the IllustrisTNG project (§5.2.1).
Our TNG50 MW/M31-like sample includes objects with stellar disky morphology, with a stellar mass in the range of $M_* = 10^{10.5-11.2}$ $M_\odot$, and within a MW-like Mpc-scale environment at $z = 0$ (§5.2.2). We have focused on the vertical structure of the TNG50 MW/M31-like galaxies and, in particular, on the flaring of their stellar disks, distinguishing across mono-age stellar populations. Throughout this paper, by disk flaring we mean the quantification of by how much the scaleheight of the stellar disk changes (i.e. increases) with galactocentric distance, once stars are selected according to their ages (or other properties).
Our analysis and results rely on the resolution, sample size, and realism of the TNG50 simulated galaxies. In fact, thanks to its high numerical resolution approaching the one typical of "zoom-in" simulations, TNG50 returns at $z = 0$ 198 different realizations of MW/M31-like galaxies. The stellar disk scalelength of TNG50 MW/M31-like galaxies ranges across $\sim 1.5 - 17$ kpc, with a good qualitative and quantitative agreement when compared to other models and available observational findings of local disky and spiral galaxies (Pillepich et al., 2023). Moreover, the vertical stellar distribution in most TNG50 MW/M31-like galaxies can be well described with a double (squared hyperbolic secant) profile, allowing us to distinguish between a thin and a thick disk *geometric* component. For some galaxies, the stellar thin (thick) disk is found to be as thin (thick) as the observed one for the Milky Way, i.e. with a scaleheight of about 175 – 360 (900 – 1300) pc.

Our main results on disk flaring can be summarized as follows:

- By fitting the vertical stellar density distribution of each simulated galaxy with a double or single functional profile, we have estimated the stellar disk scaleheights at a series of different galactocentric distances and for different mono-age stellar populations (Figs. 27 and 28). TNG50 predicts, in general, systematically higher values of the stellar scaleheight moving outward from the galactic center, i.e. it predicts "disk flaring". In fact, we show that, with one unique set of physical ingredients, TNG50 is able to reproduce diverse levels and kinds of flared stellar disks (Fig. 30).

- Because, according to TNG50, the increase of stellar scaleheight with galactocentric distance can be linear, exponential or other, depending on the galaxy and stellar population, and because disk galaxies in a narrow range of stellar mass can exhibit disk scalelengths varying by up to factors of 8, with this paper we propose and advocate for an easy, non-parametric, fit-independent measurement of the degree of flaring. Namely, we propose to

quantify flaring simply based on the relative enhancement of the scaleheight between 1 and 5 times the scalelength of each MW/M31 galaxy (Eq. 11).

- We have compared the amount of flaring displayed by the young (2-4 Gyr) and old stellar populations (8-10 Gyr) finding that which stars flare more, and by how much, changes from galaxy to galaxy, with both typically exhibiting 1.5 – 2 thicker disk heights in the outskirts than towards the center (Fig. 31). The young stellar populations in about eleven MW/M31-like galaxies exhibit a similar degree of flaring of our Milky Way, for which we have extrapolated the scaleheight as a function of radius and ages from Ting & Rix (2019).

- We have applied our method to the data available in the literature for selected zoom-in simulations of MW-like galaxies: namely, Latte, VINTERGATAN, NIHAO-UHD and Auriga. The amount of flaring of old and young stellar populations found in TNG50 encompasses qualitatively and quantitatively all the aforementioned simulations, implying that the stellar-disk flaring of the unbiased sample of TNG50 MW/M31 analogs, returned by a fixed physical model, covers all the previous findings of zoom-in simulations performed with different codes, in some cases with better resolution, different galaxy formation models and varying assumptions on the past assembly history of the simulated galaxies (Fig. 39).

- The scaleheights we measure in the TNG50 simulated galaxies are a manifestation of the underlying orbits and overall potential, with their values exhibiting a clear correlation with the local stellar vertical velocity dispersion and the local stellar surface density (Fig. 34), as predicted by theoretical models. Namely, stellar disks are thicker if hotter and with lower surface densities.

- According to our analysis, galaxies with hotter or less dense outer stellar disks flare more strongly, in both young and old stellar populations, and the diversity in flaring seems to be mostly driven by a diversity in vertical stellar velocity dispersion in the outer disk regions (Fig. 35).

- On the other hand, the flaring of the mono-age populations does not manifestly depend on global $z = 0$ structural properties of the MW/M31 sample. However, two key albeit mild trends can be highlighted: old stellar populations flare more in galaxies with larger disk scaleheight and with larger gas mass fraction in the disks, the former relationship being in place also for young stars (Figs. 40, 42).

- Finally, the stellar population properties in the mid plane of each TNG50 MW/M31-like galaxy exhibit a qualitative agreement with what typically observed: the average stellar age *decreases* at increasing galactocentric distance (and at fixed vertical height), and *increases* at increasing vertical height (and at fixed radius – Fig. 37, right panels compared

to left and middle ones). However, we argue that these observed trends do not necessarily imply flaring, and shall not be used in fact as a proxy for flaring, as a change in average stellar age with galactocentric distance can be realized with a varying trend with radius of stellar age distributions and not necessarily by a change in stellar orbits (Fig. 38).

Overall, with the analysis of this paper we have demonstrated that a sample of 198 $z = 0$ MW/M31-like galaxies – selected to be disky, isolated (but allowing Local Groups), and within a narrow range of stellar mass (covering the MW and M31 stellar masses) –, exhibit a great diversity in their vertical disk structures. The heterogeneity of the stellar disk properties and the diverse flavors and amounts of stellar disk flaring are a key result. Indeed, TNG50 is a cosmological magneto-hydrodynamical simulation that – with a given set of physical-model ingredients and with a resolution that bridges the gap between the volume and zoom-in simulations – is able to reproduce, replicate, and expand upon the features of the stellar disk flaring already investigated and shown in the literature, based on zoom-in cosmological simulations.
The fact that there is no trivial correlation between the amount of flaring and the $z = 0$ structural *global* properties of galaxies is intriguing. We have also tentatively investigated whether flaring may be correlated with the number of the more or less recent major/minor/total mergers, finding no obvious dependence (Fig. 41). These considerations, together with the findings on kinematics, lead us to speculate that the increasing of the stellar disk scaleheight with the galactocentric distance could be ascribed to a possible heating of the stellar disk due to external perturbations such as flybys (without the need for coalescence). In fact, keeping in mind that we have not imposed constraints on past history, the pathways leading to $z = 0$ MW/M31-like galaxies could be very diverse across the whole sample (see also Sotillo-Ramos et al., 2022). Future and further analyses are required to connect the disk flaring to the past of MW/M31-like galaxies and to the dynamical evolution of their stars.

Data availability statement

The entire data of the IllustrisTNG simulations, including TNG50, are publicly available and accessible at `www.tng-project.org/data` (Nelson et al., 2019a). Additional and easier-to-use data products and particle data cutouts related to the 198 MNW/M31-like galaxies from TNG50 used in this paper are also publicly available (Pillepich et al., 2023). With this paper, we make public also a series of catalogs for the various measures of disk flaring and for the flags of warped and disturbed TNG50 MW/M31-like galaxies.

Acknowledgements

DS, MD, and AP acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 138713538 – SFB 881 ("The Milky Way System", subprojects A01 and A06). DN acknowledges funding from the Deutsche Forschungsgemeinschaft (DFG) through an Emmy Noether Research Group (grant number NE 2441/1-1). The TNG50 simulation used in this work has been run on the HazelHen Cray XC40-system at the High Performance Computing Center Stuttgart under the Gauss centers for Super-computing (GCS) Large-Scale Project GCS-DWAR (2016; PIs Nelson/Pillepich).

5.8 Disk flaring vs. additional structural properties of MW/M31-like analogs

We also examine the level of flaring, observed in both young and old stars, as a function of selected galaxy mass properties: stellar mass, disk stellar mass and disk-to-total stellar mass ratio (Fig. 42).

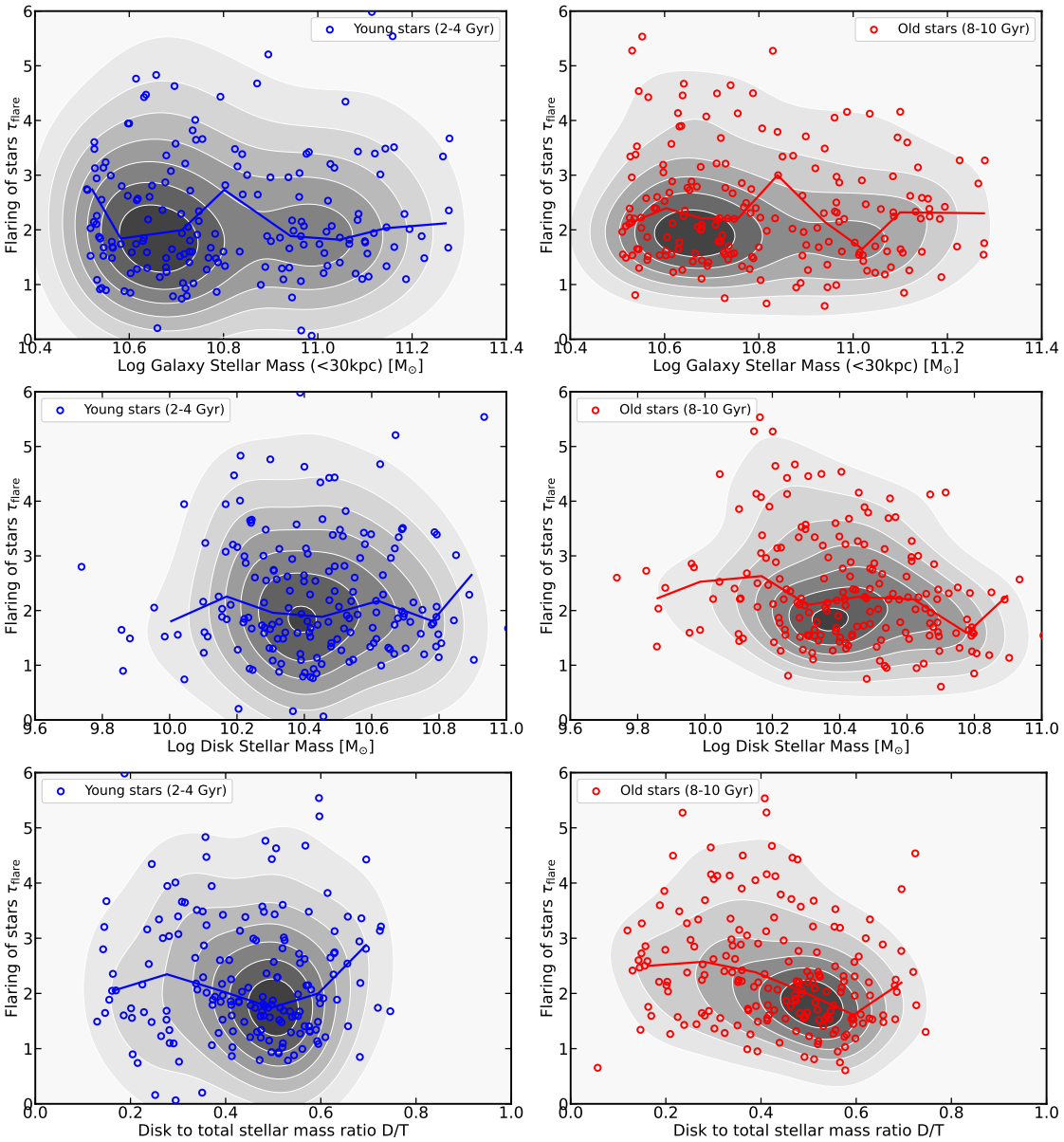

Figure 42: **Disk flaring vs. galaxy masses for TNG50 MW/M31-like galaxies.** We show the amount of flaring of young (left column, in blue) and old (right column, in red) stellar populations as a function of three selected global galaxy properties: stellar mass (top), disk stellar mass (middle), and disk-to-total stellar mass ratio (bottom). In each panel, the galaxy number density is represented with the shaded contour areas. Solid curves are medians in bins of the quantity on the x-axes. No discernible trend seem to be in place.

6 The composition of the stellar disk: metal-poor, old and ex situ stars

The work of this chapter is published as:
On the likelihoods of finding very metal-poor (and old) stars in the Milky Way's disc, bulge, and halo
Diego Sotillo-Ramos, Maria Bergemann, Jennifer K.S. Friske and Annalisa Pillepich
Monthly Notices of the Royal Astronomical Society: Letters, Volume 525, Issue 1, October 2023, Pages L105–L111

Zusammenfassung

Recent observational studies have uncovered a small number of very metal-poor stars with cold kinematics in the Galactic disc and bulge. However, their origins remain enigmatic. We select a total of 138 Milky Way (MW) analogs from the TNG50 cosmological simulation based on their $z = 0$ properties: disky morphology, stellar mass, and local environment. In order to make more predictive statements for the MW, we further limit the spatial volume coverage of stellar populations in galaxies to that targeted by the upcoming 4MOST high-resolution survey of the Galactic disc and bulge. We find that across all galaxies, ~20 per cent of very metal-poor ($[Fe/H] < -2$) stars belong to the disk, with some analogs reaching 30 per cent. About 50±10 per cent of the VMP disc stars are, on average, older than 12.5 Gyr and ~70±10 per cent come from accreted satellites. A large fraction of the VMP stars belong to the halo (~70) and have a median age of 12 Gyr. Our results with the TNG50 cosmological simulation confirm earlier findings with simulations of fewer individual galaxies, and suggest that the stellar disc of the Milky Way is very likely to host significant amounts of very- and extremely-metal-poor stars that, although mostly of ex situ origin, can also form in situ, reinforcing the idea of the existence of a primordial Galactic disc.

6.1 Introduction

One of the most exciting questions in modern observational astrophysics is the existence of primordial, so-called Population III stars (e.g. Beers et al., 1985, Beers, 2000, Christlieb et al., 2002, Sneden et al., 2003, Frebel et al., 2005, Frebel & Norris, 2015). Despite decades of theoretical and observational research, no such objects have been discovered yet, although their successors have been identified and it has become possible to link their properties to enrichment in individual stellar explosions (e.g. Keller et al., 2014, Howes et al., 2015, Takahashi et al., 2018, Ji et al., 2020, Yong, 2020, Hansen et al., 2020, Skúladóttir et al., 2021, Lagae et al., 2023).

In this work, we explore the possibility that the Milky Way galaxy may host a primordial disc, that is, qualitatively, an old in situ disc formed at $z \gtrsim 2$ out of stars born from the rotationally-supported pristine gas. In other words, the question is whether very metal-poor stars[5] could be

[5]Following Beers & Christlieb (2005), stars are defined as *metal-poor* ($[Fe/H] < -1$), *very metal-poor* ($[Fe/H] < -2$), *extremely metal-poor* ($[Fe/H] < -3$), *ultra metal-poor* ($[Fe/H] < -4$), *hyper metal-poor* ($[Fe/H] < -5$), and currently these categories are being extended to $[Fe/H] < -10$.

hiding in the Galactic disc, in addition to their established association with the Galactic halo (e.g. Schörck et al., 2009, Youakim et al., 2020, Bonifacio et al., 2021) and the bulge (e.g. Schlaufman & Casey, 2014, Howes et al., 2016, Reggiani et al., 2020). Owing to the overall progression of cosmic chemical enrichment, one expects that more metal-poor stars form in larger systems at earlier times or in smaller systems at later times and their presence would be intricately linked to the hierarchical growth of galaxies (White & Springel, 2000). From a theoretical point of view (e.g. Searle & Zinn, 1978, Salvadori et al., 2010) and also confirmed by recent cosmological simulations, such as FIRE (Hopkins et al., 2014), APOSTLE (Sawala et al., 2016, Fattahi et al., 2016), and TNG50 (Nelson et al., 2019b, Pillepich et al., 2019), it is expected that most metal-poor (MP) stellar populations follow isotropic distributions of orbits and are therefore preferentially confined to the spheroidal components, bulge and halo (e.g. Chen et al., 2023, for TNG50). This is expected, because they form either ex situ and have been accreted by progressively stripping smaller satellites to form mostly the halo; or they form in situ, at the early stages, when there was no rotationally supported component or in the primordial disc on orbits then heated by mergers (Starkenburg et al., 2017, El-Badry et al., 2018, Chen et al., 2023). Observational searches have yielded over 10^5 stars with metallicity [Fe/H] $\lesssim -2$ (Bonifacio et al., 2000, Li et al., 2018, Chiti et al., 2021, Huang et al., 2022, Andrae et al., 2023), representing all morphological components of the Galaxy, including the bulge (Howes et al., 2015, Koch et al., 2016, Reggiani et al., 2020), the halo (Hayes et al., 2018, Limberg et al., 2021), and the disc (Di Matteo et al., 2020, Sestito et al., 2020, Carter et al., 2021, Fernández-Alvar et al., 2021). With estimated fractions of $25-30$ per cent (Sestito et al., 2019, 2020), very metal-poor (VMP) stars with disky orbits are, perhaps surprisingly, not rare and are confined close to the Galaxy's midplane (see also Venn et al., 2020, Di Matteo et al., 2020). There is evidence for these systems being preferentially on prograde orbits (Carter et al., 2021, Carollo et al., 2023). However, the origin of the kinematic asymmetry is currently debated. Santistevan et al. (2021), using FIRE-2 simulations (Hopkins et al., 2018), confirm the preference for prograde orbits for the UMP disky stars and a prograde-to-retrograde ratio of $\sim 2:1$, associating the rotational bias with a single major merger event. Sestito et al. (2021) use 5 Milky Way (MW) analogs from the NIHAO-UHD project (Buck et al., 2020) to show that $[Fe/H] < -2.5$ stars in retrograde disc orbits were accreted in the first billion years of the galaxy formation, whereas the prograde subpopulation was mostly accreted at later stages.

In this work, we use the TNG50 cosmological simulation, which is the highest resolution run of the IllustrisTNG project (Naiman et al., 2018, Marinacci et al., 2018, Pillepich et al., 2018b, Nelson et al., 2018, Springel et al., 2018), to assess the fraction of metal-poor stars expected in the different Galactic morphological components. We follow up on the analysis by Chen et al. (2023), who performed the analysis of extremely metal-poor stars in TNG50 MW and M31-like galaxies. Differently from the latter study, here we aim to a) quantify the presence and origin of stars across a wider range of metallicity levels; b) put a special focus on the Galactic disc and c) estimate the statistics and properties of VMP, EMP and UMP stars by including their [Mg/Fe] ratios, ages and origin (in or ex situ). Crucially, we show where these stars are distributed in MW simulated analogues by using the nominal spatial selection informed by the upcoming 4MOST high-resolution disc and bulge survey (Bensby et al., 2019), in order to provide predictions for the detectability of VMP stars in next-generation observational programs. Compared to previous works based on zoom-in simulations, we increase the MW-analogs sample size by a factor of ~10-20.

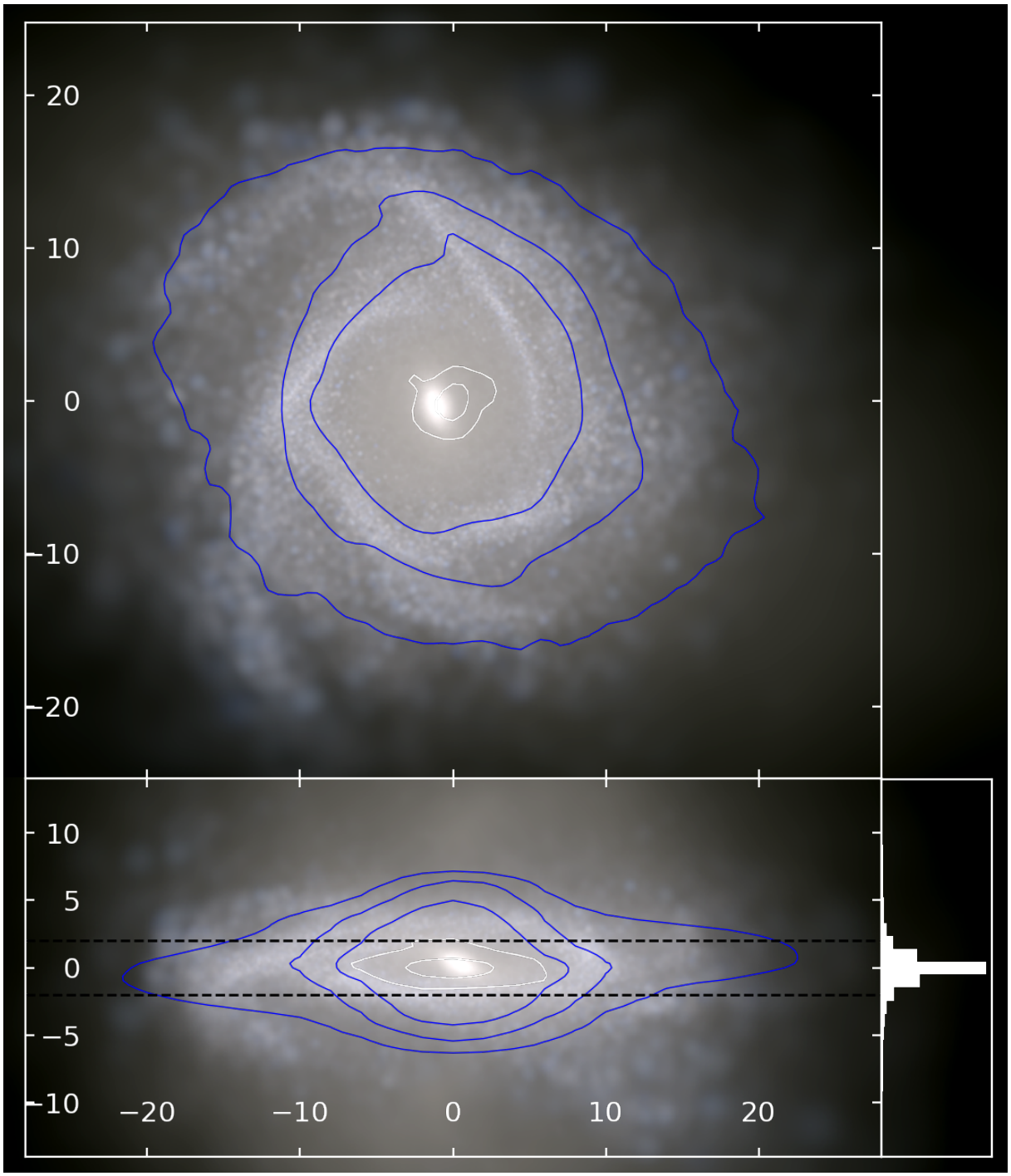

Figure 43: Stellar-light composite image of one MW-like galaxy from the TNG50 simulation in face-on and edge-on projections, among 138 MW analogs. Spatial scales are given in the units of kpc. For this example (Subhalo IDs at $z = 0$: 535774) disc scale-length, and think and thick disc scale-height are compatible with the current estimations for the Galaxy (e.g. Bland-Hawthorn & Gerhard, 2016). Blue contours trace the stellar surface density. White contours trace the surface density of VMP stars. The side histogram shows the vertical distribution of VMP stars. In all galaxies in the sample, most of the VMP stars are concentrated very close to the midplane.

The paper is organised as follows: in Sec. 6.2 we describe the cosmological simulation TNG50 and how the MW analogs are selected. We describe in Sec. 6.3 the populations of metal-poor stars across morphological components, exploring additional properties such as their ages, [Mg/Fe] abundances and origin. We discuss possible implications for the current understanding of the origin and composition of the Galaxy's disc in Sec. 6.4.

6.2 Methods

In this paper, we focus on MW-like galaxies realized within the cosmological simulation TNG50 (Nelson et al., 2019b, Pillepich et al., 2019). For details on the simulation, we refer to the latter papers, and here provide a brief account of the main properties of the galaxies.
The simulation comprises a periodic cubic volume with a side length of 51.7 comoving Mpc, contains 2160^3 dark matter (DM) particles and equal initial number of gas cells. The DM particles have an uniform mass of $m_{\rm DM} = 4.5 \times 10^5$ M$_\odot$, while the gas cells (and stellar particles) have an average (initial) mass of is $m_{\rm baryon} = 8.5 \times 10^4$ M$_\odot$. The star formation follows Springel & Hernquist (2003): the gas is transformed into star particles stochastically when the density exceeds $n_{\rm H} = 0.1 {\rm cm}^{-3}$ on time scales to reproduce the Kennicutt-Schmidt relation (Kennicutt, 1989). Stellar particles represent stellar populations that are born at the same time and have an initial mass distribution by Chabrier (2003). Detailed information about all the particles, subhaloes, and haloes, is stored in 100 snapshots. The (sub)haloes at different snapshots, i.e. across cosmic times, are linked via the SUBLINK (Rodriguez-Gomez et al., 2015) and LHALOTREE (Springel et al., 2005) algorithms, so that the assembly histories of galaxies is available. In this paper, we use the baryonic version of SUBLINK and the main progenitor of a galaxy is the one with the most massive history. We will also identify *in situ* stars as the ones formed in the main progenitor; accreted stars will be referred to as *ex situ* (as per Rodriguez-Gomez et al., 2016).
From the TNG50 simulation box, which returns hundreds of massive galaxies at $z = 0$, Pillepich et al. (2023) identify the 198 most suitable counterparts to the MW and M31 based on their properties at $z = 0$ according to galaxy stellar mass, stellar diskyness, and environment. This galaxy sample has been previously used and extensively detailed in terms of its stellar content also by Engler et al. (2021a), Sotillo-Ramos et al. (2022), Engler et al. (2023), Chen et al. (2023). With an additional cut in galaxy stellar mass ($10^{10.5-10.9}$ M$_\odot$), in this paper we identify the 138 best MW analogs from TNG50 at $z = 0$. We note that this selection does not impose any constraints on the evolutionary paths of galaxies nor a-priori on the detailed structural and chemical properties of the stellar disc and bulge.
In Fig. 43 we show the stellar-light composite face-on and edge-on images of one MW-like galaxy from the TNG50 sample, at $z = 0$. This simulated analog has disc scale-length and thin and thick disc scale-heights compatible with the current estimations for the Galaxy (see Sotillo-Ramos et al., 2023a, for more details on the calculations and the MW reference values). We overlay the positions of VMP stellar particles with white contours and shows that a high fraction lies within a few kpc of the midplane, as is the case for the Galaxy.

6.2.1 Morphological decomposition of MW analogues

There are many methods to decompose a galaxy in its stellar morphological components. Recent methodologies applied to simulated galaxies have been described by Du et al. (2019), Gargiulo

et al. (2022), Zhu et al. (2022b), are based on earlier works by e.g. Abadi et al. 2003, Doménech-Moral et al. 2012, Obreja et al. 2018 and combine structural and kinematical information. In this paper, we choose the approach by Zhu et al. (2022b), which is based on the orbit circularity ϵ_z[6] (as defined by Sotillo-Ramos et al., 2022) and galactocentric distance r_* of the stars. In brief, the four main stellar components are defined as follows (see also Chen et al., 2023):

- Cold disk: $\epsilon_z > 0.7$
- Warm disk: $0.5 < \epsilon_z < 0.7$ and $r_{cut} < r_* < r_{disk}$
- Bulge: $\epsilon_z < 0.7$ and $r_* < r_{cut}$
- Stellar halo: $\epsilon_z > 0.5$ and $r_{disk} < r_* < r_{halo}$ or $\epsilon_z < 0.5$ and $r_{cut} < r_* < r_{halo}$,

with $r_{cut} = 3.5$ kpc, $r_{disk} = 6 \times r_d$, the exponential scale-length of the stellar disk, as measured by Sotillo-Ramos et al. (2022), $r_{halo} = 300$ kpc the maximum galactocentric distance to which we consider that the stellar halo extends: see also fig. A1 in Chen et al. (2023) for a visual depiction of the components in the $\epsilon_z - r_*$ plane. 'Cold' and 'warm' discs are similar, but are *not* apriori equivalent, to the geometrically defined 'thin' and 'thick' discs based on fitting the vertical stellar density profiles (e.g. Gilmore & Reid, 1983b).

In the majority of the TNG50 MW analogues (96), the cold disc is the most massive component (by galaxy selection), with median values of $\sim 1 - 3 \times 10^{10}$ $M_\odot$, and increasing with galaxy stellar mass. There is also a significant, more than half an order of magnitude, galaxy-to-galaxy variation in the mass of all morphological components. The warm disc is the least massive component and it is one order of magnitude less massive than the cold disk. The bulge is in most cases the second most massive component, except for some galaxies at the high-mass end, where the stellar halo is more massive.

In relative terms, the cold disc represents ~60 per cent of the total stellar mass for the less massive MW analogs and decreases to ~40 per cent in the case of the most massive ones in the sample. The relative contribution of the other components does not change significantly with galaxy stellar mass. This dominance of the disc is related to the selection of the galaxies: it is in good agreement with the analysis of tens of edge-on spiral galaxies in the local universe by Comerón et al. (2014) and, considering the D/T fraction, with the properties of MW analogs from the NIHAO simulations by Obreja et al. (2018).

In this paper, we focus on the qualitative comparison with the expected properties of stars to be observed within the 4MIDABLE-HR survey (Bensby et al., 2019) that will be carried out at the 4MOST facility (de Jong et al., 2019). This survey will provide coherent homogeneous characterisation of a very large number (over 3 million) of stars in the Galactic disc and bulge, including their detailed metallicities, abundances, ages, and kinematics. To mimic the spatial coverage of 4MIDABLE-HR, we apply a cut on the volume occupied by the simulated stellar particles of 5.5 kpc in heliocentric distance centered at a random point positioned at 8 kpc from the galaxy center. We note that the 4MIDABLE-HR survey selects the targets based on apparent magnitude but, owing to instrumental limits and the complex observational strategy, in practice most stars to be observed (with the exception of selected fields) will be confined within the given

[6] $\epsilon_z = j_z / j_{circ}$, with j_z the specific angular momentum of the star in the direction perpendicular to the galactic disk, and j_{circ} the specific angular momentum of a star at the same radius, on a circular orbit, i.e., $j_{circ} = r v_{circ}$, with $v_{circ} = \sqrt{GM(\leq r)/r}$ the circular velocity of the galaxy at the considered radius.

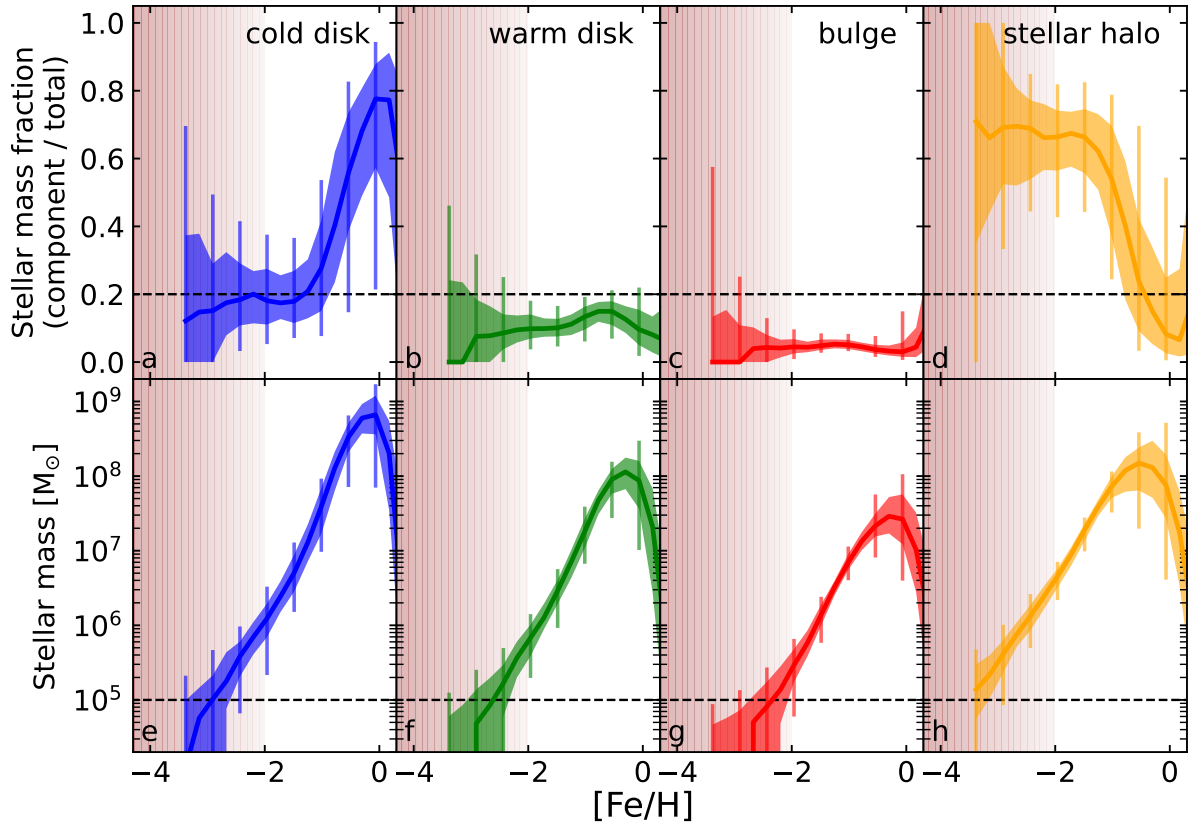

Figure 44: Metallicity distributions of stars in TNG50 MW-like galaxies, grouped by their respective morphological component: cold disc, warm disc, bulge and stellar halo, from left to right. The top panels quantify the stellar mass fractions in each component to the total stellar mass; the bottom ones are the metallicity distributions in each morphological component, in stellar mass. The solid lines represent the medians and the shaded areas and error-bars represent inter-per centile ranges across the galaxy sample: 16 to 84 and 2 to 98, respectively. The vertical red shaded bands highlight metallicities [Fe/H]≤ -2, i.e., VMP stars. We remind the reader that the distributions are shown for the characteristic observable volume fraction of the Galaxy, as it will be 'seen' by the 4MIDABLE-HR survey on 4MOST (see Section 6.2.1). Specifically, we apply a volume cut of 5.5 kpc in heliocentric distance, where the fiducial 'Sun' is placed at 8 kpc in the simulated galaxy.

spatial volume: this lends credibility to our procedure, despite its simplicity. It should also be noted that, with this spatial cut, the bulge component is represented only by its most external stars. Finally, we have checked that the results of the paper are qualitatively the same than if we had placed the fiducial 'Sun' not at a fixed 8 kpc distance but at 4 times the disc length of each galaxy, to account for the diversity in galaxy sizes (see e.g. Figure 13 of Pillepich et al., 2023). In the next section, we explore in detail the temporal, chemical, and evolutionary properties of these four main Galactic components, and analyse their distributions by focusing on the metal-poor and old populations.

6.3 Results

We begin with the analysis of the temporal (i.e. of the stellar ages) and chemical properties of the components in the simulated galaxies, and then proceed with the assessment of the statistical properties of the distributions in the volume that will be accessible to next-generation spectroscopic surveys of the Galaxy, such as with 4MOST (Bensby et al., 2019, Chiappini et al., 2019, Christlieb et al., 2019), WEAVE (Dalton et al., 2012, 2016), and MOONS (Gonzalez et al., 2020). As described in the previous section, the focus here is on the 4MOST 4MIDABLE-HR survey and we refer the reader to the science case for more details on its scope and strategy (Bensby et al., 2019).

6.3.1 Trends with metallicity

Fig. 44 shows the metallicity distribution functions (MDFs) for all four Galactic components of all TNG50 MW-like galaxies described in Section 6.2. The distributions include a cut on the volume to mimic the spatial coverage of 4MIDABLE-HR, as detailed in the previous Section. The solid lines represent the medians across galaxies. Shaded areas and error-bars represent inter-per centile ranges also across the studied galaxy sample: 16 to 84 and 2 to 98 per cent, respectively. Cold disk, warm disk, bulge and halo are represented, respectively, with blue, green, red and orange lines, from left to right. We also show the stellar mass fraction per component, that is stars in the component relative to the total number of stars: top row.
In most TNG50 MW analogues, the majority of Sun-like, [Fe/H]$\sim$ 0 stars reside in the kinematically cold 'thin' disk, but with some scatter that encompasses $\sim$ 20 per cent of the total number of stars in the disc. The bulge is the second most populated component when no cut is applied, whereby we note that the apparent difference in the total stellar mass in the bulge and in the thin disc is caused by the application of the fiducial volume cut to account for the survey selection of 4MIDABLE-HR. In general, just a small, although non-negligible (at the level of 20 per cent) fraction of the solar-metallicity stars can be found in the halo or the warm disk, although we note a significant galaxy-to-galaxy variation.
For stars with [Fe/H] $\lesssim -1$ the trend reverses. As expected, in the vast majority of galaxies, most of these low-metallicity stars reside in the stellar halo, with a median fraction of $\sim$ 60 per cent. However, and this is one of most interesting results, the fraction of VMP stars in the cold disc component still reaches up to $\sim$ 20 per cent, in the typical galaxy. In fact, in some MW-like galaxies, as many as 40 per cent of the stars with [Fe/H]~ -2 follow cold disky orbits. For progressively lower metallicity values, the trends change slope. Most of the stars with metallicity values [Fe/H] $\lesssim -2$ across all galaxies can be found in the stellar halo, with the median fraction of $\sim$ 80 per cent. For the cold (thin) disk, the median values are around $\sim$ 15 per

cent, but we consistently find a significant (∼25 per cent) number of galaxies where the fraction of VMP stars is ≥ 25 per cent. These results are qualitatively and quantitatively very similar if we place the 'Sun' at four times the disc length: the median values change only minimally across metallicity and component, although the scatter is in all cases is larger.
In summary, we find that large fractions of ultra, extremely and very metal-poor stars are present in all morphological components of the simulated Milky Way analogues. Most intriguingly, their MDFs suggest that such stars should be abundantly present in the cold disk, typically referred to as the thin disc. This finding confirms recent results of kinematically cold VMP targets in the literature (Sestito et al., 2019, Di Matteo et al., 2020), with our results indicating that the fraction of such stars in the Galactic disc could be even higher than the current observational evidence suggests.

6.3.2 Trends with stellar age

In order to understand the temporal history of metal-poor populations, in Fig. 45 we show the age distribution functions of the halo, bulge and disc components in the simulated galaxies. In the top row, we normalise the fraction of stars per [Fe/H] bin in each population to the total number of stars in all populations. In the second row, as in Fig. 44, we also provide the stellar mass in each component per metallicity bin, and in the bottom row we show the corresponding mass of each stellar population for metal-poor stars with [Fe/H] < −1.
The bulge and stellar halo appear as the oldest components, followed by the warm disc and the cold disc as the youngest component, although each of these populations show a significant temporal extent spanning the entire range of ages up to ∼ 13 Gyr, with only a mild dependence on the galaxy. Even in the cold disc, a non-negligible fraction of stars of ∼ 20 per cent, have ages greater than 10 Gyr, and the cold discs of some MW analogues stand out with fractions as large as 50 per cent. There is no galaxy in our TNG50 sample that does not host an old cold disc. On the one hand, properties of these distributions are consistent with what observers would usually describe as the canonical formation picture of the MW (Freeman & Bland-Hawthorn, 2002). On the other hand, the extended star formation histories of all components, especially that of the disc and (to a lesser extent) of the bulge, are striking and indicate that the Milky Way may host a primordial disc that we explore in more detail in Sec. 6.3.3.
In the bottom row of Fig. 45, we show the age distributions for the metal-poor populations ([Fe/H] < −1) of the discs, the bulge, and the stellar halo. Here we do not apply the heliocentric cut, in order to be able to find a significant enough number of EMP and UMP stars. It is clear, and expected, that the age distribution of each component exhibits a trend toward older ages with decreasing metallicity: middle vs. bottom panels of Fig. 45 (Bergemann et al., 2014). Specifically, the median age of all four components is now skewed towards ages ≥ 8 Gyr, and the mode values peak at ≳ 12 Gyr for all galaxies and components. Cold disc stars with [Fe/H]≈ −1 have an age distribution that closely resembles that of the warm disc and the halo (with the median ages of ∼10 Gyr to 11 Gyr), whereby the bulge appears to be made of the oldest population, as consistently seen in all MW analogues. We find very similar distributions for VMP, EMP and UMP stars. We also note, and this will be discussed in more detail in the next section, that generally, halo stars of all metallicities are on average younger than the stars in the bulge (Fig. 46).
Finally, by exploring the median [Mg/Fe] abundance ratios (Fig. 46), we also find a strong dependence of the distributions on metallicity. In line with observational evidence (e.g. Bensby

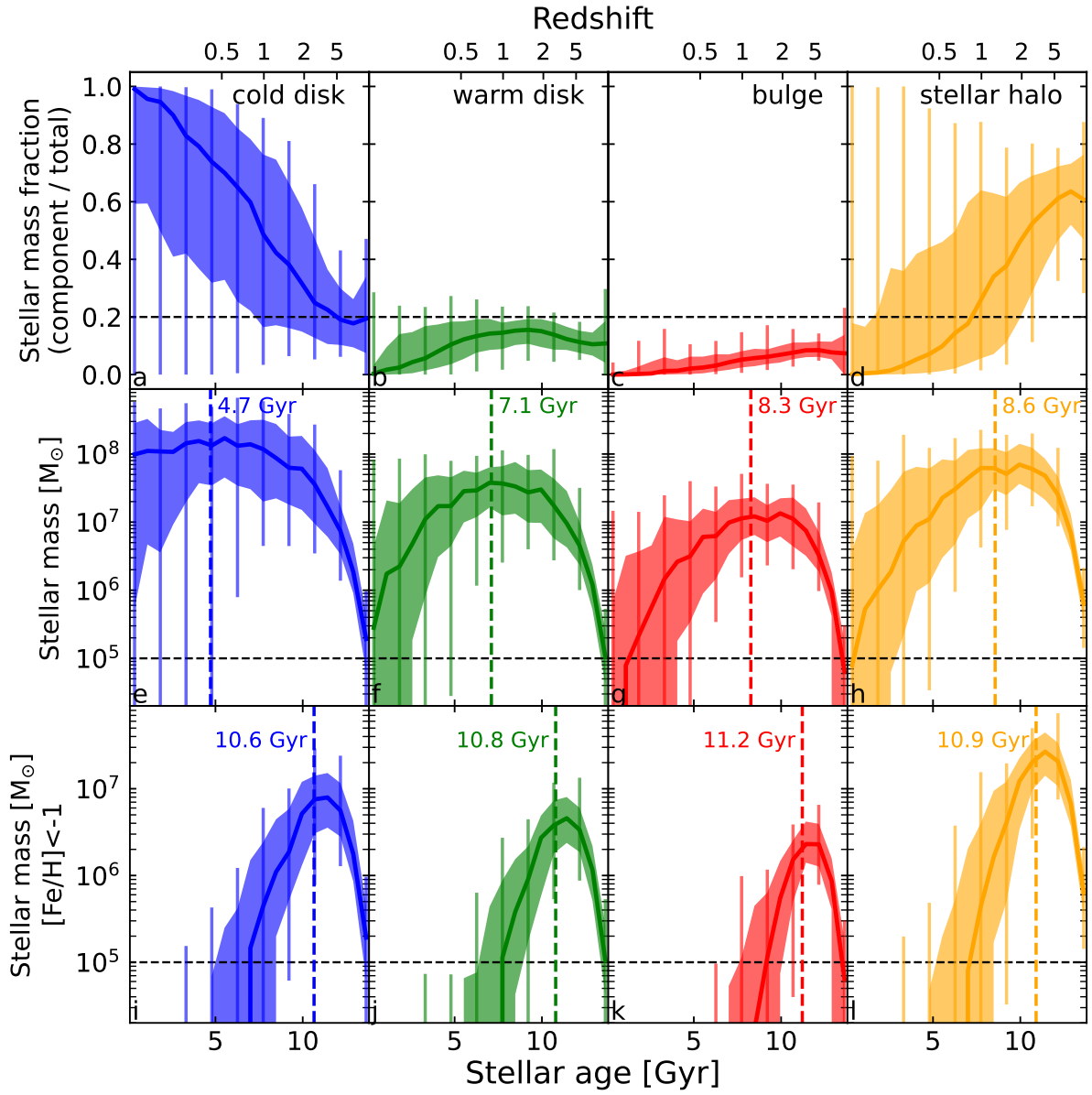

Figure 45: Same as Fig. 2, but for the distribution of stellar ages in the morphological components. *Top*: Stellar mass fraction per component to total stellar mass. *Middle*: Stellar mass per component. *Bottom*: Mass of MP stars per component. The vertical dashed lines represent the median age of the stars in the component (within the volumetric cut), across all galaxies.

et al., 2014, Bergemann et al., 2017, Nissen & Gustafsson, 2018), [Mg/Fe] increases as the [Fe/H] values decrease (from the top to the bottom panels). At [Fe/H]$\sim$ 0, the cold disc has the lowest [Mg/Fe]. For metal-poor stars the bulge exhibits, on average, slightly higher (by $\sim$ 0.05 dex) values of [Mg/Fe]. Such a systematic difference for the bulge is indeed observed in the Milky Way (Rich & Origlia, 2005, Cunha & Smith, 2006, Ryde et al., 2010, Rich et al., 2012), but see also Jönsson et al. (2017) and Griffith et al. (2021), who advocate smaller chemical differences between the disc and the bulge (but with the caveat here that the different works use different definitions for the morphological components). Interestingly, we also find an increasingly large scatter of [Mg/Fe] ratios for $-4 \lesssim$ [Fe/H] $\lesssim -3$, which is consistent with the observational distributions of EMP stars by Howes et al. (2016), although their data suggest a higher scatter also for stars with [Fe/H] $\lesssim -2.5$. At higher metallicities, [Fe/H] $\gtrsim -2$, the stellar halo and both disc components show similar trends of progressively declining [Mg/Fe] values. However, we emphasize that owing to very large differences in the stellar mass in each metallicity bin, it is unlikely that one would observe many low-α halo stars or high-α cold disc stars, respectively.

6.3.3 Origins of metal-poor stars

The detailed properties of the stellar particles in the TNG50 simulation allow the identification of their origin, in particular, whether they formed in the main galaxy or whether they were accreted. Fig. 46 shows that as with the age and [Mg/Fe] abundance, there are clear trends that depend on the stellar metallicity. More metal-poor stellar populations have, on average, a higher ex situ fraction. This applies to all Galactic components, although the halo, on average, has a comparatively higher ex situ fraction for all metallicities. These distributions are qualitatively consistent with observations, e.g. Conroy et al. (2019b) have shown that the accreted fraction increases with decreasing metallicity for the Milky Way halo stars.
Kinematically cold and warm discs are dominated by the ex situ component for metallicities below [Fe/H] $\lesssim -2$. Yet, interestingly, even at the lowest metallicities a non-negligible fraction of stars in these components have an in situ origin. The median ex situ fraction values for the discs are, in these cases, $\approx$ 65, $\approx$ 75 and $\approx$ 100 per cent, for metallicities [Fe/H] of -2, -3, and -4, respectively. However, importantly, we find also some MW analogs (Fig. 46 bottom row, right) where 10 to 30 per cent of the UMP disk stars were formed in situ, and given the age of these stars, most likely in a primordial disk. This finding is not in contradiction with Ruchti et al. (2015). They report no strong evidence for an accreted disc component, however, their observed sample is dominated by targets with [Fe/H] $\gtrsim -1$, and only a very small fraction of their stars are more metal-poor than [Fe/H] ~ -2, which, as we see in Fig. 46, represents a transition from the in situ to the ex situ dominated regime. Indeed, TNG50 suggests that most of the stars with [Fe/H]$\gtrsim -1$ have a strongly in situ dominated origin, regardless of their parent Galactic component. However, for some of the analogs ($\sim$10 per cent), the accreted disc fraction can be also significant ($\gtrsim$15 per cent). The properties of early galactic discs, such as their alignment with respect to the present-day orientation, are complex and will be a subject of another paper. This has also been addressed, e.g. in Belokurov & Kravtsov (2022). However, here we note that from a quick inspection of the angular momenta, we find that at ancient times, galaxies show a large spread of angular momentum vectors close to a uniform distribution. This could be expected if stars formed in a chaotic way, with multiple gas inflows from many random directions and mergers, that potentially destroy and heat the primordial discs and bring stars in randomly distributed orbits. We cannot detect any preferred angle for the orientation of the

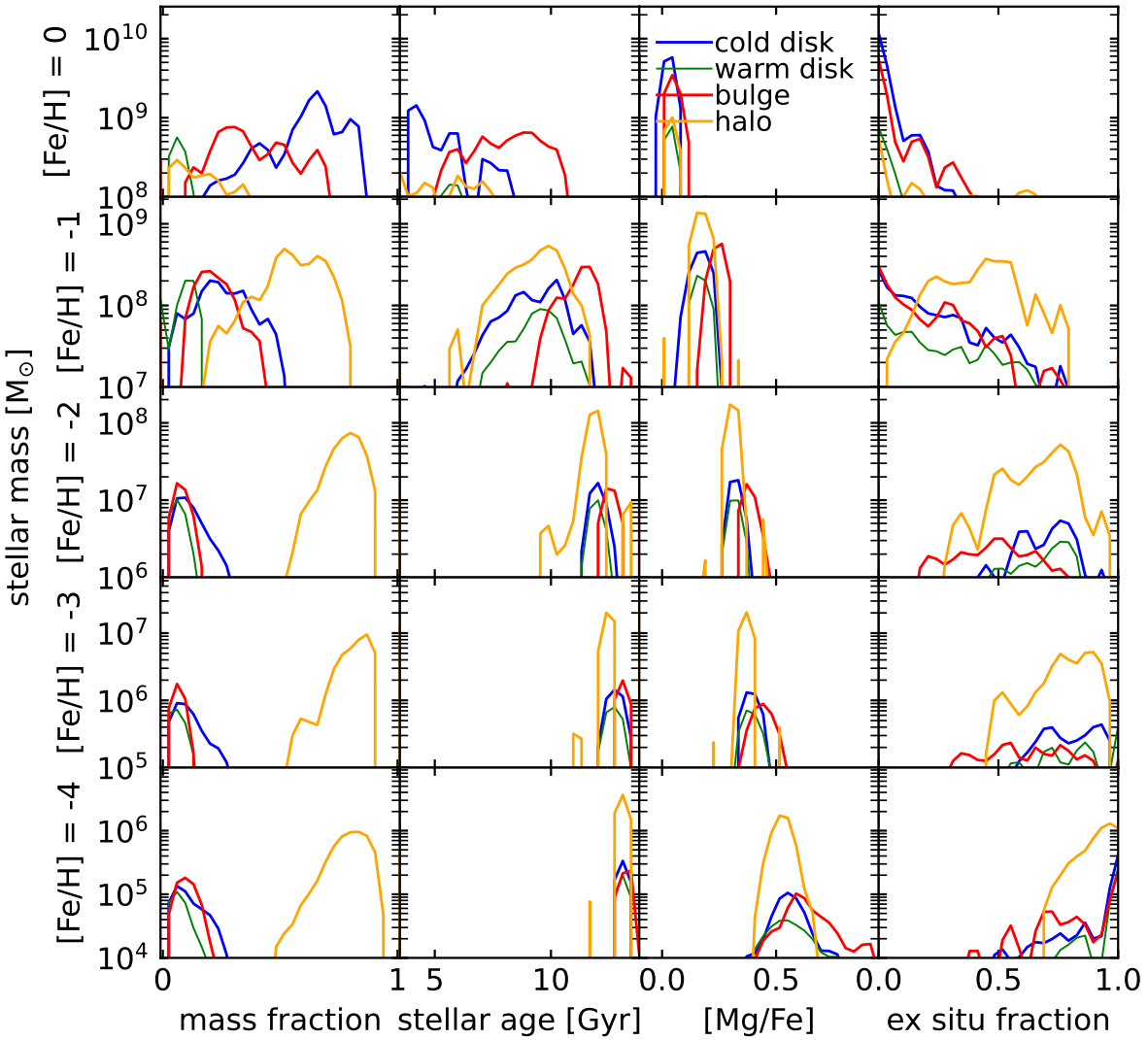

Figure 46: Mass fraction per component, stellar age, [Mg/Fe] and ex situ fraction distributions, for stellar samples with different values of [Fe/H]. We show the distributions of the median values across all TNG50 138 MW-like galaxies weighted by stellar mass and do not apply a heliocentric cut.

primordial disk, but the alignment steadily increases as the galaxies evolve and approach the present-day age, $z = 0$.

6.4 Summary and conclusions

We have used the cosmological magneto-hydrodynamic galaxy simulation TNG50 to explore the fraction of very metal-poor stars, [Fe/H] $\lesssim -2$, in Milky Way like galaxies. The selection of galaxies follows our detailed previous work presented in Pillepich et al. (2023). We furthermore apply observationally motivated limits to the theoretical distributions, aiming to understand which fraction of the metal-poor stars would be observable in the Galactic disc, bulge, and halo with next-generation facilities, such as the 4MIDABLE-HR survey on 4MOST (Bensby et al., 2019).

Through the statistical analysis of the stellar populations of simulated galaxies, specifically their metallicities, ages, and Mg/Fe ratios, we find that metal-poor stars are common in all morphological components of the MW analogues. As expected, the stellar halo is the component primarily hosting VMP stars (see also Chen et al., 2023). However, we find that in the cold 'thin' discs of TNG50 MW-like galaxies, the fraction of VMP, EMP, and UMP stars is typically ≈ 20 per cent of the total number of stars, and in some MW-like galaxies stars with [Fe/H] ≈ -3 reach up to 50 per cent. Most of these low-metallicity stars are formed ex situ. The temporal properties of these populations suggest that all galaxy components, i.e. the cold (thin) disc, the warm (thick) disc, the halo, and the bulge, have very extended evolutionary histories with ages reaching $\gtrsim 13$ Gyr. This suggests that the Galaxy could host a primordial cold disc even though a significant fraction of metal-poor stars in disc orbits originated from small satellites and got subsequently accreted.

The large unbiased sample of MW analogs of the TNG50 simulation confirms that metal-poor stars can help unveil the first steps of the formation of the Galaxy. Contrary to what has been largely expected thus far, it is very likely that many of these stars follow cold disc orbits. We hence recommend that current and future Galactic surveys should target not only the stellar halo, but also the disc for the search of the most metal-poor old stars, which will be equivalent to exploring the regime of redshifts $z \sim 2$ to > 5.

Acknowledgements

We are grateful to the anonymous referee for the insightful suggestions that helped to improve the work. We acknowledge Timothy C. Beers and Anirudh Chiti for valuable discussions. DSR, AP, and MB acknowledge support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 138713538 – SFB 881 ("The Milky Way System", subprojects A01, A05, A06, A10). MB is supported through the Lise Meitner grant from the Max Planck Society. This project has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (Grant agreement No. 949173). JF acknowledges support from University College London's Graduate Research Scholarships and the MPIA visitor programme. The TNG50 simulation was realized with compute time granted by the Gauss Centre for Super- computing (GCS), under the GCS Large-Scale Project GCS-DWAR (2016; PIs Nelson/Pillepich) on the GCS share of the supercomputer Hazel Hen at the High Performance Computing Center Stuttgart (HLRS). This work benefited

from a workshop supported by the National Science Foundation under Grant No. OISE-1927130 (IReNA), the Kavli Institute for Cosmological Physics, and the University of Chicago Data Science Institute.

Data Availability

Data directly related to this publication and its figures are available on request from the corresponding author. The IllustrisTNG simulations, including TNG50, are publicly available and accessible at `www.tng-project.org/data` (Nelson et al., 2019a). A special data release is also available for the TNG50 Milky Way and Andromeda like galaxies, as per Pillepich et al. (2023).

III. DISCUSSION, OUTLOOK AND SUMMARY

7 Discussion

The results presented here demonstrate the significant power of TNG50 and the galaxy formation model of TNG in faithfully reproducing many observed properties of MW/M31-like galaxies, thus highlighting as well its strong predictive capability. Additionally, I will now summarize some further findings obtained through the exploration of the same sample of MW/M31-like galaxies. However, it is essential to acknowledge a number of inherent limitations in TNG50 (and other cosmological simulations), which I specify below. The obtained results also have direct implications for future galaxy surveys or studies conducted with existing surveys, and I provide some examples.

7.1 Other results on the TNG50 MW/M31-like

We have already shown in section 3.1 a selection of scientific results of the IllustrisTNG suite, although only a small sample, since more than 650 papers have been published to date. We will now list some additional results on MW/M31-like galaxies:

- X-ray emitting bubbles, shells, and cavities emerge within the circumgalactic gas enveloping TNG50 MW/M31-like galaxies. These structures bear a resemblance to the eROSITA and Fermi bubbles found within our own Galaxy. At $z = 0$, two-thirds of the analogs exhibit zones of overpressurized gas penetrating the gaseous halo (Pillepich et al., 2021).
- The CGM of MW-like galaxies exhibit a great structural complexity up to R_{200c}, with multimodal distributions of temperature, density and entropy, as shown by Ramesh et al. (2023). The CGM properties are connected with the SF rate of the galaxy. Also, a radially decreasing magnetic field is predicted.
- Extremely metal-poor stars in the MW/M31-like galaxies of TNG50 can be found in all components, but predominantly in the stellar haloes (Chen et al., 2023).
- An analysis of the bulges shows that $\approx$ 20 per cent of the analogs have a high Sérsic index, while $\approx$ 80 per cent have lower values (Gargiulo et al., 2022). No correlation between the bulge properties and the environment is found. On the contrary, a dependency is predicted by TNG50 between the presence of bars and the type of bulge, as well as higher Sérsic indices among galaxies that have a higher fraction of ex situ stars.
- A diversity in integrated radial metallicity profiles is realized by TNG50 in MW/M31 analogs: a broken profile, as the one observed in the MW, is present only in a small fraction of the galaxies (Lian et al., 2023).
- The analysis of the satellites surrounding the MW/M31 analogs shows numbers similar to MW, M31 and similar observed galaxies: there is no missing satellites problem according to IllustrisTNG, as shown by Engler et al. (2021b), with their quenched fractions also in good agreement with observations (Engler et al., 2023).

7.2 Limitations of TNG50

TNG50 allows an unprecedented number of MW/M31-like galaxies to be analyzed at high resolution, comparable to many of the recent cosmological zoom-in simulations of MW analogs, which instead only simulate a handfull of systems. At $z = 0$, TNG50 galaxies have a number of stellar particles between approximately 500 thousand and 1 million. The number of DM particles in its halo ranges approximately between 1 and 10 millions. This number of particles allows a very detailed reproduction of their morphology and structure. However, the progenitors of these galaxies will contain a smaller number of particles, which will impose limitations. For example, if we consider a minimum number of 100 stellar particles to be able to determine the galactic morphology and structure, this will not be possible for epochs before $z \approx 8$.
For a fixed physical model, changes in the resolution may alter the properties of the galaxies: this effect is called resolution convergence. In TNG50, galaxy sizes and heights are converging at all mass scales ant times, among the different resolution levels (TNG50, and lower resolution variants TNG50-2, -3 and -4). Above a certain mass limit that depends on redshift, the values across the different resolution levels are in agreement (see, e.g. Pillepich et al., 2019, for more details). A further
The volume extension of the simulation also limits the sample. We already mentioned that by imposing a larger number of conditions to find analogs (e.g., for a MW, the presence of a bar, correct number of satellites, presence of a M31 in the surroundings, a quiet recent merger history, etc.), the sample would be reduced to a minimum. As an example, only three pairs of the 198 galaxies are found in Local Group analogs. As another example, none of the analogues has a SMBH with a mass similar to that of the MW central BH. However, analyzing in TNG300 galaxies in haloes of a similar mass to those considered in our study (giving a sample of ~40000 MW/M31-analogs instead of 200), the width of the SMBH mass distribution extends to much lower masses, giving values compatible with those of the MW's SMBH.

7.3 Implications for surveys i.e. connection to observations

All the studies conducted in this thesis harness the power of TNG50 to uncover properties of galaxies and their stellar components that are challenging to estimate from observations.
The merger history of an observed galaxy is impossible to determine a priori. Through cosmological simulations of galaxies, it is possible to identify certain properties that may also be determinable through observations, and that can be associated with specific events in the history of galaxies. This conclusion was reached in the first study, in Chapter 4, where MW/M31-like galaxies with recent major mergers exhibit some $z = 0$ global and structural properties that distinguish them (thicker stellar disks, more massive and somewhat shallower stellar haloes, etc), on average, from the overall sample's average properties. This means that based on these measured properties, a probability can be assigned for an observed galaxy to have experienced a recent major merger.
We have also found in that study that the progenitors of MW/M31-like galaxies exhibit a significant diversity. For instance, at $z \approx 2$, their stellar mass span two orders of magnitude. This should be taken into account in studies that use galaxy surveys to investigate what the MW must have been like in ancient times, searching for galaxies that may resemble the MW's progenitors.
Flaring is a phenomenon that can be directly observed and quantified in the MW but also

external galaxies. In the latter, however, flaring measurement for monoage stellar populations is challenging. In the MW, these measurements have been possible with the aid of data from different surveys. With the results of the study presented in Chapter 5, I have shown that some of the flaring properties appear more frequently in galaxies with certain structural characteristics, which can be used in galaxy surveys to estimate characteristics related to the flaring of some of the stellar populations in these galaxies. For example, according to TNG50, MW/M31-like galaxies where old stellar populations show stronger flaring have, on average, larger stellar disk scale-height and higher gas fractions in the disks.
The study of very metal-poor stars in MW-like galaxies, presented in Chapter 6, aims to make reasonable predictions about the number of stars of this metallicity range that can be expected in the Milky Way. Through current and future surveys, it is possible to measure the metallicity of the observed stars. Based on the simulation results, it can be determined how likely it is for a very metal-poor star to belong to the disk, halo, or bulge. On average, we have quantified that around 20 percent of these stars belong to the disk, across the whole sample of MW-like galaxies in TNG50. This implies that future surveys aiming to find VMP stars should also target the Galactic disk.

7.4 Outlook

With the results obtained in this thesis, we gain a deeper understanding of the Milky Way, M31, and similar galaxies. These findings open the door to address additional scientific questions that were not tackled in these studies. Moreover, the use of cosmological simulations of galaxies have also implicit limitations; therefore, improvements in them are suggested to enhance or extend the obtained results. Below, I summarize some of these aspects.

In the case of the disk survival study, this work would clearly benefit from a larger number of galaxies. The total sample number of 198 MW/M31 analogs is quickly reduced if we add aditional criteria: for example, for the subsample of galaxies with a recent major merger, the number is reduced to 31 galaxies. Moreover, taking into account the two different scenarios (stellar disk that is destroyed but reforms and disk that survives), it is further reduced. A larger sample would help to reinforce the statistical significance of some of the statements we have made above. For this, one computationally inexpensive possibility would be to perform re-simulations of the TNG100 galaxies that inhabit halos such as those in MW and M31 at the resolution of TNG50. This would enlarge the sample by a factor of ≈ 8. Another scientific extension would be to deepen the analysis of the orbital parameters of the major mergers: relationships between the spins of the colliding galaxies, duration of the interaction, etc. Even extending the analysis presented in the study to other mergers, for example not recent or with other mass ratios, could bring more significance to the results. Finally, one element that this first study lacked was the influence of the environment (small, medium or large scale) on the probability of occurrence of the mergers or on the two merger scenarios.

The study of the flaring of the stellar disk helped us to better understand this phenomenon in MW/M31-like galaxies, although it is difficult to predict how flaring would manifest itself in a particular galaxy. One of the basic questions in the theory of galactic disk formation is whether stars that at $z = 0$ show a high vertical velocity dispersion were born with this kinematic property, probably inherited from a hotter and more turbulent gas, or whether it is an acquired property resulting from galactic evolution and possible interactions and perturbations. Analyzing these

phenomena from the flaring perspective can help to understand the causes. As an extension of the analysis carried out in the study, the effect of radial migration of stars on the flaring phenomenon can also be investigated, as it has been done in previous studies using cosmological or idealized simulations of one or a few galaxies.

In the study of very metal-poor stars in the galactic disks, we have provided some estimates of the number of stars that we expect the next surveys to find in the disk and the other morphological components of the galaxy. Some simplifications were made here that distort the results. Firstly, the volumetric spherical cut around the equivalent solar position does not represent the selection function of the surveys: they are bounded by the observable limiting magnitudes as well as by a certain observable fraction of the celestial sphere, limited by several factors depending on the survey. Secondly, the stellar particles of the simulation were considered without calculating the equivalent stellar populations that they intent to represent: this is determinant for calculating the magnitude and visibility of all stars from an equivalent heliocenter. In addition, it is also necessary to calculate the effects of extinctions on the apparent magnitude. On the other hand, the selection of a helioequivalent position was azimuthally random, simply respecting the values of the distance to the galactic center, but not its relative position with respect to possible structures such as arms or the central bar. By taking these improvements into account, more accurate results can be given about the number of VMP stars expected to be observed in the Galactic disk.

8 Summary and final remarks

In this thesis I have used state-of-the-art cosmological simulations of galaxies to study fundamental aspects of the formation and evolution of MW/M31-like galaxies: how it is possible that they can undergo a recent merger but be morphologically disky at $z = 0$, how the stellar disk flaring phenomenon manifests itself and how diverse it is in this type of galaxies, and how frequently very metal-poor stars can be found in the galactic disk and the other morphological stellar components, as well as the clues this provides us about the origin and evolution of these galaxies (with the direct implications for our own Galaxy). The TNG50 cosmological simulation of galaxies, which is the primary tool used to conduct the studies contained in this thesis, is currently almost unique in its class because it combines the variety and multiplicity of objects that is characteristic of a large-volume simulation with the high numerical resolution previously only available in zoom-in simulations, which usually only focus on one or a few galaxies. TNG50 contains 198 MW/M31-like galaxies, thus adding statistical significance to the results obtained.
In the first part of the Introduction (§1), I described the most important theoretical concepts of cosmology and large-scale structure formation in the Universe, as well as the physical processes involved in galaxy formation, reviewing some of the observations that support our theories. I then described the diversity of galaxies, with a special emphasis on disk galaxies, which are the main focus of this thesis. Finally, I described in great detail some of the properties that we know about the MW (and M31 to a lesser extent), how they compare with respect to other galaxies of its type that we know, and what kind of observations and surveys have allowed us to accumulate this knowledge, and which ones are planned in the near future to deepen it.
In the second part of the Introduction (§2), I described the main type of tools I used in this thesis, namely cosmological simulations of galaxies, and in particular TNG50. I first gave an overview, describing what is the practical implementation of the large and small scale physical processes

needed to form elements with properties comparable to the observed galaxies. I highlighted the progress made in recent years and the main difficulties encountered in simulating disky galaxies. Then I also described the most common types of cosmological simulations, especially those used to analyze MW-like galaxies, as well as the pros and cons of each of them.
In §3, I detailed the cosmological simulation IllustrisTNG, including the implementations of the physical processes used in it and that constitute the TNG model of galaxy formation. I described the highest resolution run of the suite, TNG50, enumerating the virtues that make it the optimal cosmological simulation for the investigations carried out in this thesis. I highlighted some previous astrophysical results that have been reviewed and contrasted, as well as predictions that have been obtained using the IllustrisTNG model. I described the selection criteria of the sample of MW/M31-like galaxies in TNG50, which were the basis for the rest of the analyses carried out in this thesis. I also detailed some of the general properties of the 198 galaxies that were obtained in the sample, which show that the adopted selection criteria provide a sample whose properties cover non-narrow ranges of values (e.g., for star formation rates, disk scale-length, number of satellites, etc.) that, in general, contain the values observed for MW and M31 in them.
In the next section I present the results obtained in my analysis. In §4, I studied to what extent major mergers affect the stellar disks of this type of galaxies, in particular by answering the question whether a galaxy like the MW can undergo a major merger in a recent period (in the last 5 Gyr) and yet remain a disk galaxy at $z = 0$. I found that according to TNG50, major mergers are very frequent, with only 30/198 galaxies having none of this type since $z = 5$. 31/198, moreover, underwent at least one recent major merger. We find here two main scenarios: before the recent major merger, the galaxy already had disk morphology (except for 2 galaxies), and for 2/3 of the galaxies, the merger or the preceding galaxy-galaxy interaction destroys the stellar disk, but a new disk is able to reform until $z = 0$. In all other cases, the stellar disk survives the merger. I showed that the progenitors of these galaxies have more gas than similar galaxies at the same epoch, and that the orbital configuration and the exact ratio of stellar masses between progenitors is, to some extent, a factor that will determine which of the two merger scenarios the resulting galaxy will be located in. In terms of their properties, these galaxies show some characteristics that make them different when compared to the total sample. For example, on average, more massive and shallower stellar halos, higher fractions of ex situ stars, and, very importantly, wider and kinematically hotter stellar disks, although we will also find some galaxies whose disk are as thin as that of the MW, contradicting some of the paradigms that were assumed about the formation of our Galaxy.
In §5, I quantified and analyzed the flaring phenomenon in the MW/M31-like galaxies of TNG50, that is, the increase of the scale-height of the stellar disk as a function of the increase of the Galactocentric distance. Among the galaxies in the sample are some with disk dimensions (scale length and scale height) comparable to those of MW and M31. The observed flaring has many possible manifestations, and I analyzed it with an emphasis on the monoage populations, where it is more easily quantifiable and comparable with previous studies based on observations and simulations. To this end, I introduced a nonparametric measure of flaring based on two regions, in the inner and outer parts of the stellar disk. Typically, the scale-heights increase by a factor of 1.5 to 2 between these two points. I was able to find that the values and types of flaring found cover a range of values observed in a disaggregated manner in previous simulations using various physical models and galaxy samples. However, it is not straightforward to predict how

flaring will manifest itself in the stellar populations of a given galaxy. Nevertheless, I was able to show that locally, scale-heights correlate with surface density and vertical velocity dispersion, and to demonstrate that galaxies whose outer disk is less dense or kinematically hotter exhibit a higher level of flaring.

Finally, in §6, using the sample of 138 MW-like galaxies in TNG50, I analyzed how likely very metal-poor stars are to be found in the disk at $z = 0$. To make the predictions more compatible with future surveys such as 4MOST, I limited the analysis to a heliocentric sphere at an azimuthally random position 8 kpc from the galaxy center. I found that, on average, 20 percent of the VMP stars are found in the disk, with some galaxies reaching values of 30 percent. The stellar halo will contain most of the VMP stars, 70 percent on average. The VMP stars are generally very old, and I was able to show that among those in the disk, more than half are 12.5 Gyr or older. Their origin is, in most cases (on average 70 percent), ex situ, i.e., in satellites that at some time coalesced with the main galaxy. However, this leaves an open door to the possibility that in most of these galaxies of TNG50 (and therefore, in the MW), stars that belonged to a primordial disk, formed already at the dawn of the creation of these objects, can be found. The implications of these results therefore rule out the traditional idea that the Milky Way stellar halo must be the oldest component and suggest that future surveys will also look for VMP stars in the Galactic disk.

In summary and conclusion, the work presented here has served to confirm or refute some of the hypotheses regarding the history and evolution of our Milky Way and similar galaxies, relying on the descriptive and predictive power of a state-of-the-art cosmological simulation of galaxies, such as TNG50, while also ensuring the validation of results against available observations. Furthermore, it lays the foundation for future studies on the topics here presented, through more advanced simulations or by employing a more sophisticated treatment of the data within the current simulation. Lastly, these findings also provide valuable guidance for future galaxy surveys, including those focused on the Milky Way itself.

References

Abadi M. G., Navarro J. F., Steinmetz M., Eke V. R., 2003, ApJ, 591, 499

Abazajian K. N., et al., 2009, ApJS, 182, 543

Adibekyan V. Z., Delgado Mena E., Sousa S. G., Santos N. C., Israelian G., González Hernández J. I., Mayor M., Hakobyan A. A., 2012, A&A, 547, A36

Agertz O., Kravtsov A. V., 2015, ApJ, 804, 18

Agertz O., Teyssier R., Moore B., 2011, MNRAS, 410, 1391

Agertz O., et al., 2021, MNRAS, 503, 5826

Alard C., 2000, arXiv e-prints, pp astro–ph/0007013

Alpher R. A., Bethe H., Gamow G., 1948, Phys. Rev., 73, 803

Andrae R., Rix H.-W., Chandra V., 2023, arXiv e-prints, p. arXiv:2302.02611

Aumer M., White S. D. M., Naab T., Scannapieco C., 2013a, MNRAS, 434, 3142

Aumer M., White S. D. M., Naab T., Scannapieco C., 2013b, MNRAS, 434, 3142

Bagla J. S., 2002, Journal of Astrophysics and Astronomy, 23, 185

Balogh M. L., Pearce F. R., Bower R. G., Kay S. T., 2001, MNRAS, 326, 1228

Bamford S. P., et al., 2009, MNRAS, 393, 1324

Banerjee A., Jog C. J., 2007, ApJ, 662, 335

Barmby P., et al., 2006, ApJ, 650, L45

Barnes J. E., 1988, ApJ, 331, 699

Barnes J. E., 1992, ApJ, 393, 484

Barnes J. E., Hernquist L., 1996, ApJ, 471, 115

Barnes J., Hut P., 1986, Nature, 324, 446

Baumann D., 2022, Cosmology. Cambridge University Press, doi:10.1017/9781108937092

Beaton R. L., et al., 2007, ApJ, 658, L91

Beers T. C., 2000, in Weiss A., Abel T. G., Hill V., eds, The First Stars. p. 3 (arXiv:astro-ph/9911171), doi:10.1007/10719504_1

Beers T. C., Christlieb N., 2005, ARA&A, 43, 531

Beers T. C., Preston G. W., Shectman S. A., 1985, AJ, 90, 2089

Behroozi P. S., Wechsler R. H., Conroy C., 2013, ApJ, 770, 57

Bell E. F., et al., 2008, ApJ, 680, 295

Bellman R., 1961, Adaptive Control Processes. Princeton University Press

Belokurov V., Kravtsov A., 2022, MNRAS, 514, 689

Belokurov V., et al., 2009, MNRAS, 397, 1748

Belokurov V., Erkal D., Evans N. W., Koposov S. E., Deason A. J., 2018, MNRAS, 478, 611

Bender R., et al., 2005, ApJ, 631, 280

Bensby T., Feltzing S., Oey M. S., 2014, A&A, 562, A71

Bensby T., et al., 2019, The Messenger, 175, 35

Bergemann M., et al., 2014, A&A, 565, A89

Bergemann M., Collet R., Schönrich R., Andrae R., Kovalev M., Ruchti G., Hansen C. J., Magic Z., 2017, ApJ, 847, 16

Berger M. J., Colella P., 1989, Journal of Computational Physics, 82, 64

Bird J. C., Kazantzidis S., Weinberg D. H., Guedes J., Callegari S., Mayer L., Madau P., 2013, The Astrophysical Journal, 773, 43

Bird J. C., Loebman S. R., Weinberg D. H., Brooks A. M., Quinn T. R., Christensen C. R., 2021, MNRAS, 503, 1815

Bizyaev D. V., Kautsch S. J., Mosenkov A. V., Reshetnikov V. P., Sotnikova N. Y., Yablokova N. V., Hillyer R. W., 2014, ApJ, 787, 24

Blaña Díaz M., Wegg C., Gerhard O., Erwin P., Portail M., Opitsch M., Saglia R., Bender R., 2017, MNRAS, 466, 4279

Bland-Hawthorn J., Gerhard O., 2016, ARA&A, 54, 529

Blanton M. R., et al., 2017, AJ, 154, 28

Boardman N., et al., 2020, MNRAS, 498, 4943

Boecker A., Neumayer N., Pillepich A., Frankel N., Ramesh R., Leaman R., Hernquist L., 2023, MNRAS, 519, 5202

Bonaca A., et al., 2020, ApJ, 897, L18

Bonifacio P., Monai S., Beers T. C., 2000, AJ, 120, 2065

Bonifacio P., et al., 2021, A&A, 651, A79

Bovy J., 2015, ApJS, 216, 29

Bovy J., Rix H.-W., 2013, ApJ, 779, 115

Bovy J., Rix H.-W., Hogg D. W., 2012, ApJ, 751, 131

Bovy J., Rix H.-W., Schlafly E. F., Nidever D. L., Holtzman J. A., Shetrone M., Beers T. C., 2016, ApJ, 823, 30

Bryan G. L., et al., 2014, ApJS, 211, 19

Buck T., Obreja A., Macciò A. V., Minchev I., Dutton A. A., Ostriker J. P., 2020, MNRAS, 491, 3461

Bullock J. S., Kravtsov A. V., Weinberg D. H., 2001, ApJ, 548, 33

Carollo D., Christlieb N., Tissera P. B., Sillero E., 2023, ApJ, 946, 99

Carter C., et al., 2021, ApJ, 908, 208

Cautun M., Deason A. J., Frenk C. S., McAlpine S., 2019, MNRAS, 483, 2185

Chabrier G., 2003, PASP, 115, 763

Chaplin W. J., et al., 2020, Nature Astronomy, 4, 382

Chatzopoulos S., Fritz T. K., Gerhard O., Gillessen S., Wegg C., Genzel R., Pfuhl O., 2015, MNRAS, 447, 948

Chen B., et al., 2001, ApJ, 553, 184

Chen L.-H., Pillepich A., Glover S. C. O., Klessen R. S., 2023, MNRAS, 519, 483

Cheng S., Ting Y.-S., Ménard B., Bruna J., 2020, MNRAS, 499, 5902

Chiappini C., et al., 2019, The Messenger, 175, 30

Chiti A., Frebel A., Mardini M. K., Daniel T. W., Ou X., Uvarova A. V., 2021, ApJS, 254, 31

Christlieb N., et al., 2002, Nature, 419, 904

Christlieb N., et al., 2019, The Messenger, 175, 26

Chua K. T. E., Pillepich A., Vogelsberger M., Hernquist L., 2019, MNRAS, 484, 476

Colless M., et al., 2003, arXiv e-prints, pp astro–ph/0306581

Collins M. L. M., et al., 2011, MNRAS, 413, 1548

Comerón S., et al., 2011a, ApJ, 729, 18

Comerón S., et al., 2011b, ApJ, 741, 28

Comerón S., et al., 2012, ApJ, 759, 98

Comerón S., Elmegreen B. G., Salo H., Laurikainen E., Holwerda B. W., Knapen J. H., 2014, A&A, 571, A58

Conroy C., et al., 2019a, ApJ, 883, 107

Conroy C., Naidu R. P., Zaritsky D., Bonaca A., Cargile P., Johnson B. D., Caldwell N., 2019b, ApJ, 887, 237

Conroy C., Naidu R. P., Garavito-Camargo N., Besla G., Zaritsky D., Bonaca A., Johnson B. D., 2021, Nature, 592, 534

Crain R. A., et al., 2015, MNRAS, 450, 1937

Cunha K., Smith V. V., 2006, ApJ, 651, 491

DESI Collaboration et al., 2016, arXiv e-prints, p. arXiv:1611.00036

D'Souza R., Bell E. F., 2018, Nature Astronomy, 2, 737

Dalcanton J. J., et al., 2012, ApJS, 200, 18

Dalton G., et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. p. 84460P, doi:10.1117/12.925950

Dalton G., et al., 2016, in Evans C. J., Simard L., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 9908, Ground-based and Airborne Instrumentation for Astronomy VI. p. 99081G, doi:10.1117/12.2231078

Dame T. M., 1993, in Holt S. S., Verter F., eds, American Institute of Physics Conference Series Vol. 278, Back to the Galaxy. pp 267–278, doi:10.1063/1.43985

Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieferantsoa M. H., Appleby S., 2019, MNRAS, 486, 2827

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, ApJ, 292, 371

Dawson K. S., et al., 2013, AJ, 145, 10

De Silva G. M., et al., 2015, MNRAS, 449, 2604

Deason A. J., Belokurov V., Evans N. W., 2011, MNRAS, 416, 2903

Deason A. J., Belokurov V., Koposov S. E., Rockosi C. M., 2014, ApJ, 787, 30

Deng L.-C., et al., 2012, Research in Astronomy and Astrophysics, 12, 735

Di Matteo P., Spite M., Haywood M., Bonifacio P., Gómez A., Spite F., Caffau E., 2020, A&A, 636, A115

Dierickx M. I. P., Loeb A., 2017, ApJ, 836, 92

van Dokkum P. G., et al., 2013, ApJ, 771, L35

Doménech-Moral M., Martínez-Serrano F. J., Domínguez-Tenreiro R., Serna A., 2012, MNRAS, 421, 2510

Donnari M., Pillepich A., Nelson D., Marinacci F., Vogelsberger M., Hernquist L., 2021, MNRAS, 506, 4760

Dorman C. E., et al., 2015, ApJ, 803, 24

Dressler A., 1980, ApJ, 236, 351

Driver S. P., et al., 2011, MNRAS, 413, 971

Du M., Ho L. C., Zhao D., Shi J., Debattista V. P., Hernquist L., Nelson D., 2019, ApJ, 884, 129

Du M., Ho L. C., Debattista V. P., Pillepich A., Nelson D., Zhao D., Hernquist L., 2020, ApJ, 895, 139

Dubois Y., et al., 2014, MNRAS, 444, 1453

Dullo B. T., Graham A. W., Knapen J. H., 2017, MNRAS, 471, 2321

D'Onghia E., Madau P., Vera-Ciro C., Quillen A., Hernquist L., 2016, The Astrophysical Journal, 823, 4

Eggen O. J., Lynden-Bell D., Sandage A. R., 1962, ApJ, 136, 748

Einstein A., 1917, Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften, pp 142–152

El-Badry K., et al., 2018, MNRAS, 480, 652

Engler C., et al., 2021a, MNRAS, 500, 3957

Engler C., et al., 2021b, MNRAS, 507, 4211

Engler C., Pillepich A., Joshi G. D., Pasquali A., Nelson D., Grebel E. K., 2023, MNRAS, 522, 5946

Euclid Collaboration et al., 2022, A&A, 662, A112

Evans N. W., Gyuk G., Turner M. S., Binney J., 1998, ApJ, 501, L45

Fakhouri O., Ma C.-P., 2008, MNRAS, 386, 577

Fakhouri O., Ma C.-P., Boylan-Kolchin M., 2010, MNRAS, 406, 2267

Fall S. M., Efstathiou G., 1980, MNRAS, 193, 189

Fattahi A., et al., 2016, MNRAS, 457, 844

Fernández-Alvar E., et al., 2021, MNRAS, 508, 1509

Feuillet D. K., Frankel N., Lind K., Frinchaboy P. M., García-Hernández D. A., Lane R. R., Nitschelm C., Roman-Lopes A. r., 2019, MNRAS, 489, 1742

Flynn C., Holmberg J., Portinari L., Fuchs B., Jahreiß H., 2006, MNRAS, 372, 1149

Font A. S., McCarthy I. G., Crain R. A., Theuns T., Schaye J., Wiersma R. P. C., Dalla Vecchia C., 2011, MNRAS, 416, 2802

Font A. S., McCarthy I. G., Le Brun A. M. C., Crain R. A., Kelvin L. S., 2017, Publ. Astron. Soc. Australia, 34, e050

Font A. S., et al., 2020, MNRAS, 498, 1765

Frankel N., Sanders J., Ting Y.-S., Rix H.-W., 2020, ApJ, 896, 15

Frankel N., et al., 2022, ApJ, 940, 61

Fraser-McKelvie A., Merrifield M., Aragón-Salamanca A., 2019, MNRAS, 489, 5030

Frebel A., Norris J. E., 2015, ARA&A, 53, 631

Frebel A., et al., 2005, Nature, 434, 871

Freeman K., Bland-Hawthorn J., 2002, ARA&A, 40, 487

Friedmann A., 1922, Zeitschrift fur Physik, 10, 377

Gadotti D. A., 2009, MNRAS, 393, 1531

Gaia Collaboration et al., 2016, A&A, 595, A1

Gaia Collaboration et al., 2018, A&A, 616, A1

Gaia Collaboration et al., 2022, arXiv e-prints, p. arXiv:2208.00211

Gallart C., Bernard E. J., Brook C. B., Ruiz-Lara T., Cassisi S., Hill V., Monelli M., 2019, Nature Astronomy, 3, 932

Gamow G., 1946, Physical Review, 70, 572

García Pérez A. E., et al., 2013, ApJ, 767, L9

García de la Cruz J., Martig M., Minchev I., James P., 2021, MNRAS, 501, 5105

Gargiulo I. D., et al., 2022, MNRAS, 512, 2537

Garrison-Kimmel S., et al., 2018, MNRAS, 481, 4133

Geehan J. J., Fardal M. A., Babul A., Guhathakurta P., 2006, MNRAS, 366, 996

Genel S., Bouché N., Naab T., Sternberg A., Genzel R., 2010, ApJ, 719, 229

Genel S., et al., 2014, MNRAS, 445, 175

Genel S., et al., 2018, MNRAS, 474, 3976

Ghez A. M., et al., 2008, ApJ, 689, 1044

Gilbert K. M., et al., 2009, ApJ, 705, 1275

Gilbert K. M., et al., 2012, ApJ, 760, 76

Gilmore G., Reid N., 1983b, MNRAS, 202, 1025

Gilmore G., Reid N., 1983a, MNRAS, 202, 1025

Gingold R. A., Monaghan J. J., 1977, MNRAS, 181, 375

Gómez F. A., Minchev I., O'Shea B. W., Beers T. C., Bullock J. S., Purcell C. W., 2013, MNRAS, 429, 159

Gómez F. A., White S. D. M., Grand R. J. J., Marinacci F., Springel V., Pakmor R., 2017, MNRAS, 465, 3446

Gonzalez O. A., et al., 2020, The Messenger, 180, 18

Goto T., Yamauchi C., Fujita Y., Okamura S., Sekiguchi M., Smail I., Bernardi M., Gomez P. L., 2003, MNRAS, 346, 601

Gould A., Bahcall J. N., Flynn C., 1996, ApJ, 465, 759

Governato F., et al., 2004, ApJ, 607, 688

Governato F., Willman B., Mayer L., Brooks A., Stinson G., Valenzuela O., Wadsley J., Quinn T., 2007, Monthly Notices of the Royal Astronomical Society, 374, 1479

Grand R. J. J., Springel V., Gómez F. A., Marinacci F., Pakmor R., Campbell D. J. R., Jenkins A., 2016, MNRAS, 459, 199

Grand R. J. J., et al., 2017, MNRAS, 467, 179

Grand R. J. J., et al., 2020, MNRAS, 497, 1603

Griffith E., et al., 2021, ApJ, 909, 77

de Grijs R., Peletier R. F., 1997, A&A, 320, L21

Guedes J., Callegari S., Madau P., Mayer L., 2011, ApJ, 742, 76

Guth A. H., 1981, Phys. Rev. D, 23, 347

Guzmán-Ortega A., Rodriguez-Gomez V., Snyder G. F., Chamberlain K., Hernquist L., 2023, MNRAS, 519, 4920

Hammer F., Puech M., Chemin L., Flores H., Lehnert M. D., 2007, ApJ, 662, 322

Hammersley P. L., Cohen M., Garzón F., Mahoney T., López-Corredoira M., 1999, MNRAS, 308, 333

Hansen T. T., et al., 2020, ApJ, 897, 183

Hau G. K. T., Bower R. G., Kilborn V., Forbes D. A., Balogh M. L., Oosterloo T., 2008, MNRAS, 385, 1965

Hayes C. R., et al., 2018, ApJ, 852, 49

Haywood M., Di Matteo P., Lehnert M. D., Katz D., Gómez A., 2013, A&A, 560, A109

Heitmann K., et al., 2021, ApJS, 252, 19

Helmi A., White S. D. M., de Zeeuw P. T., Zhao H., 1999, Nature, 402, 53

Helmi A., Babusiaux C., Koppelman H. H., Massari D., Veljanoski J., Brown A. G. A., 2018, Nature, 563, 85

Hernquist L., 1989, Nature, 340, 687

Hernquist L., 1993, ApJ, 409, 548

Hernquist L., Barnes J. E., 1991, Nature, 354, 210

Hockney R. W., Eastwood J. W., 1988, Computer simulation using particles

Hoffman L., Cox T. J., Dutta S., Hernquist L., 2010, ApJ, 723, 818

Homma D., et al., 2019, PASJ, 71, 94

Hopkins P. F., Cox T. J., Younger J. D., Hernquist L., 2009, ApJ, 691, 1168

Hopkins P. F., et al., 2010, ApJ, 715, 202

Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, MNRAS, 445, 581

Hopkins P. F., et al., 2018, MNRAS, 480, 800

Howes L. M., et al., 2015, Nature, 527, 484

Howes L. M., et al., 2016, MNRAS, 460, 884

Huang Y., et al., 2022, ApJ, 925, 164

Hubble E. P., 1926, ApJ, 64, 321

Ibata R. A., et al., 2014, ApJ, 780, 128

Ilbert O., et al., 2006, A&A, 453, 809

Ishiyama T., et al., 2021, MNRAS, 506, 4210

Ivezić Ž., et al., 2000, AJ, 120, 963

Jackson R. A., Martin G., Kaviraj S., Laigle C., Devriendt J. E. G., Dubois Y., Pichon C., 2020, MNRAS, 494, 5568

Jeans J. H., 1961, Astronomy and cosmogony. Dover

Ji A. P., et al., 2020, ApJ, 889, 27

de Jong R. S., et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 8446, Ground-based and Airborne Instrumentation for Astronomy IV. p. 84460T (arXiv:1206.6885), doi:10.1117/12.926239

de Jong R. S., et al., 2019, The Messenger, 175, 3

Jönsson H., Ryde N., Schultheis M., Zoccali M., 2017, A&A, 598, A101

Joshi G. D., Pillepich A., Nelson D., Marinacci F., Springel V., Rodriguez-Gomez V., Vogelsberger M., Hernquist L., 2020, MNRAS, 496, 2673

Jurić M., et al., 2008, ApJ, 673, 864

Kalberla P. M. W., Kerp J., Dedes L., Haud U., 2014, ApJ, 794, 90

Kannappan S. J., Guie J. M., Baker A. J., 2009, AJ, 138, 579

Karademir G. S., Remus R.-S., Burkert A., Dolag K., Hoffmann T. L., Moster B. P., Steinwandel U. P., Zhang J., 2019, MNRAS, 487, 318

Karim T., Mamajek E. E., 2017, MNRAS, 465, 472

Kasparova A. V., Katkov I. Y., Chilingarian I. V., Silchenko O. K., Moiseev A. V., Borisov S. B., 2016, MNRAS, 460, L89

Kautsch S. J., Grebel E. K., Barazza F. D., Gallagher J. S. I., 2006, A&A, 451, 1171

Kazantzidis S., Zentner A. R., Kravtsov A. V., Bullock J. S., Debattista V. P., 2009, The Astrophysical Journal, 700, 1896–1920

Keller S. C., et al., 2014, Nature, 506, 463

Kelvin L. S., et al., 2014, MNRAS, 444, 1647

Kennicutt Robert C. J., 1989, ApJ, 344, 685

Kereš D., Vogelsberger M., Sijacki D., Springel V., Hernquist L., 2012, MNRAS, 425, 2027

Koch A., McWilliam A., Preston G. W., Thompson I. B., 2016, A&A, 587, A124

Koposov S. E., Belokurov V., Torrealba G., Evans N. W., 2015, ApJ, 805, 130

Koppelman H. H., Helmi A., Massari D., Price-Whelan A. M., Starkenburg T. K., 2019, A&A, 631, L9

Kordopatis G., et al., 2015, MNRAS, 447, 3526

Kroupa P., 2001, MNRAS, 322, 231

van der Kruit P. C., 1988, A&A, 192, 117

van der Kruit P. C., Searle L., 1981, A&A, 95, 105

Kunder A., et al., 2017, AJ, 153, 75

Laevens B. P. M., et al., 2015, ApJ, 813, 44

Lagae C., Amarsi A. M., Rodríguez Díaz L. F., Lind K., Nordlander T., Hansen T. T., Heger A., 2023, A&A, 672, A90

Lamers H. J. G. L. M., Cassinelli J. P., 1999, Introduction to Stellar Winds

Lelli F., McGaugh S. S., Schombert J. M., 2016, AJ, 152, 157

Li C., Zhao G., 2017, ApJ, 850, 25

Li H., Tan K., Zhao G., 2018, ApJS, 238, 16

Lian J., et al., 2022, MNRAS, 513, 4130

Lian J., Bergemann M., Pillepich A., Zasowski G., Lane R. R., 2023, Nature Astronomy,

Libeskind N. I., et al., 2020, MNRAS, 498, 2968

Licquia T. C., Newman J. A., 2015, ApJ, 806, 96

Limberg G., et al., 2021, ApJ, 907, 10

Lindegren L., et al., 2016, A&A, 595, A4

Liu C., Hao L., Wang H., Yang X., 2019, ApJ, 878, 69

Loebman S. R., Roškar R., Debattista V. P., Ivezić Ž., Quinn T. R., Wadsley J., 2011, ApJ, 737, 8

López-Corredoira M., Molgó J., 2014, A&A, 567, A106

López-Corredoira M., Cabrera-Lavers A., Garzón F., Hammersley P. L., 2002, A&A, 394, 883

Lucy L. B., 1977, AJ, 82, 1013

Ma X., Hopkins P. F., Wetzel A. R., Kirby E. N., Anglés-Alcázar D., Faucher-Giguère C.-A., Kereš D., Quataert E., 2017, MNRAS, 467, 2430

Mackereth J. T., et al., 2017, MNRAS, 471, 3057

Majewski S. R., et al., 2017, AJ, 154, 94

Marinacci F., Pakmor R., Springel V., 2014, MNRAS, 437, 1750

Marinacci F., et al., 2018, MNRAS, 480, 5113

Martell S. L., et al., 2017, MNRAS, 465, 3203

Martig M., Bournaud F., Croton D. J., Dekel A., Teyssier R., 2012, ApJ, 756, 26

Martig M., Minchev I., Flynn C., 2014, MNRAS, 442, 2474

Martig M., Minchev I., Ness M., Fouesneau M., Rix H.-W., 2016, ApJ, 831, 139

Martín-Navarro I., Pillepich A., Nelson D., Rodriguez-Gomez V., Donnari M., Hernquist L., Springel V., 2021, Nature, 594, 187

Martin G., Kaviraj S., Devriendt J. E. G., Dubois Y., Pichon C., 2018, MNRAS, 480, 2266

Masters K. L., et al., 2011, MNRAS, 411, 2026

Mateu C., Vivas A. K., 2018, MNRAS, 479, 211

Matthews L. D., 2000, AJ, 120, 1764

Mau S., et al., 2020, ApJ, 890, 136

Mayer L., Governato F., Kaufmann T., 2008, Advanced Science Letters, 1, 7

McConnachie A. W., Irwin M. J., Ferguson A. M. N., Ibata R. A., Lewis G. F., Tanvir N., 2005, MNRAS, 356, 979

McDermid R. M., et al., 2006, MNRAS, 373, 906

McMillan P. J., 2011, MNRAS, 414, 2446

van der Marel R. P., 2006, The Large Magellanic Cloud: Structure and kinematics. Cambridge University Press, p. 47–71, doi:10.1017/CBO9780511734908.005

Miglio A., et al., 2021, A&A, 645, A85

Minchev I., Famaey B., Quillen A. C., Dehnen W., Martig M., Siebert A., 2012, A&A, 548, A127

Minchev I., Chiappini C., Martig M., 2014a, A&A, 572, A92

Minchev I., et al., 2014b, ApJ, 781, L20

Minchev I., Martig M., Streich D., Scannapieco C., de Jong R. S., Steinmetz M., 2015, ApJ, 804, L9

Minchev I., Steinmetz M., Chiappini C., Martig M., Anders F., Matijevic G., de Jong R. S., 2017, ApJ, 834, 27

Minchev I., et al., 2019, Monthly Notices of the Royal Astronomical Society

Monaghan J. J., 1992, ARA&A, 30, 543

Mukhanov V., 2005, Physical Foundations of Cosmology, doi:10.2277/0521563984.

Mulchaey J. S., Regan M. W., 1997, The Astrophysical Journal, 482, L135

Myeong G. C., Evans N. W., Belokurov V., Sanders J. L., Koposov S. E., 2018, ApJ, 863, L28

Myeong G. C., Vasiliev E., Iorio G., Evans N. W., Belokurov V., 2019, MNRAS, 488, 1235

Naab T., Ostriker J. P., 2017, ARA&A, 55, 59

Naab T., Jesseit R., Burkert A., 2006, MNRAS, 372, 839

Naidu R. P., et al., 2021, ApJ, 923, 92

Naiman J. P., et al., 2018, MNRAS, 477, 1206

Narayan C. A., Jog C. J., 2002, A&A, 390, L35

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Navarro J. F., et al., 2018, MNRAS, 476, 3648

Nelson D., et al., 2015, Astronomy and Computing, 13, 12

Nelson D., et al., 2018, MNRAS, 475, 624

Nelson D., et al., 2019a, Computational Astrophysics and Cosmology, 6, 2

Nelson D., et al., 2019b, MNRAS, 490, 3234

Nelson E. J., et al., 2021, MNRAS, 508, 219

Ness M., Hogg D. W., Rix H. W., Martig M., Pinsonneault M. H., Ho A. Y. Q., 2016, ApJ, 823, 114

Ness M. K., Johnston K. V., Blancato K., Rix H. W., Beane A., Bird J. C., Hawkins K., 2019, ApJ, 883, 177

Nieten C., Neininger N., Guélin M., Ungerechts H., Lucas R., Berkhuijsen E. M., Beck R., Wielebinski R., 2006, A&A, 453, 459

Nissen P. E., Gustafsson B., 2018, A&ARv, 26, 6

Nordström B., et al., 2004, A&A, 418, 989

Obreja A., Macciò A. V., Moster B., Dutton A. A., Buck T., Stinson G. S., Wang L., 2018, MNRAS, 477, 4915

Ojha D. K., 2001, MNRAS, 322, 426

Okamoto T., Eke V. R., Frenk C. S., Jenkins A., 2005, MNRAS, 363, 1299

Ostriker E. C., Binney J. J., 1989, MNRAS, 237, 785

Park C., Choi Y.-Y., Vogeley M. S., Gott J. Richard I., Blanton M. R., SDSS Collaboration 2007, ApJ, 658, 898

Park M.-J., et al., 2019, ApJ, 883, 25

Park M. J., et al., 2021, ApJS, 254, 2

Peñarrubia J., Ma Y.-Z., Walker M. G., McConnachie A., 2014, MNRAS, 443, 2204

Peschken N., Łokas E. L., Athanassoula E., 2020, MNRAS, 493, 1375

Petersson J., Renaud F., Agertz O., Dekel A., Duc P.-A., 2023, MNRAS, 518, 3261

Pillepich A., et al., 2014, MNRAS, 444, 237

Pillepich A., Madau P., Mayer L., 2015, ApJ, 799, 184

Pillepich A., et al., 2018a, MNRAS, 473, 4077

Pillepich A., et al., 2018b, MNRAS, 475, 648

Pillepich A., et al., 2019, MNRAS, 490, 3196

Pillepich A., Nelson D., Truong N., Weinberger R., Martin-Navarro I., Springel V., Faber S. M., Hernquist L., 2021, MNRAS, 508, 4667

Pillepich A., et al., 2023, arXiv e-prints, p. arXiv:2303.16217

Planck Collab. et al., 2020, A&A, 641, A1

Planck Collaboration et al., 2016, A&A, 594, A13

Planck Collaboration et al., 2020, A&A, 641, A6

Pritchet C., 1983, AJ, 88, 1476

Puech M., Hammer F., Hopkins P. F., Athanassoula E., Flores H., Rodrigues M., Wang J. L., Yang Y. B., 2012, ApJ, 753, 128

Qu Y., Di Matteo P., Lehnert M. D., van Driel W., 2011, A&A, 530, A10

Ramesh R., Nelson D., Pillepich A., 2023, MNRAS, 518, 5754

Rees M. J., Ostriker J. P., 1977, MNRAS, 179, 541

Reggiani H., Schlaufman K. C., Casey A. R., Ji A. P., 2020, AJ, 160, 173

Renaud F., Agertz O., Read J. I., Ryde N., Andersson E. P., Bensby T., Rey M. P., Feuillet D. K., 2021, MNRAS, 503, 5846

Rich R. M., Origlia L., 2005, ApJ, 634, 1293

Rich R. M., Origlia L., Valenti E., 2012, ApJ, 746, 59

Rich R. M., et al., 2019, MNRAS, 490, 1539

Riess A. G., et al., 1998, AJ, 116, 1009

Riess A. G., Fliri J., Valls-Gabaud D., 2012, ApJ, 745, 156

Rix H.-W., Bovy J., 2013, A&ARv, 21, 61

Robertson B., Bullock J. S., Cox T. J., Di Matteo T., Hernquist L., Springel V., Yoshida N., 2006, ApJ, 645, 986

Robin A. C., Reylé C., Fliri J., Czekaj M., Robert C. P., Martins A. M. M., 2014, A&A, 569, A13

Roca-Fàbrega S., Valenzuela O., Colín P., Figueras F., Krongold Y., Velázquez H., Avila-Reese V., Ibarra-Medel H., 2016, ApJ, 824, 94

Roca-Fàbrega S., et al., 2021, ApJ, 917, 64

Rodriguez-Gomez V., et al., 2015, MNRAS, 449, 49

Rodriguez-Gomez V., et al., 2016, MNRAS, 458, 2371

Rodriguez-Gomez V., et al., 2017, MNRAS, 467, 3083

Rodriguez-Gomez V., et al., 2019, MNRAS, 483, 4140

Rosas-Guevara Y., et al., 2022, MNRAS, 512, 5339

Rothberg B., Joseph R. D., 2004, AJ, 128, 2098

Roškar R., Debattista V. P., Brooks A. M., Quinn T. R., Brook C. B., Governato F., Dalcanton J. J., Wadsley J., 2010, MNRAS, 408, 783

Roškar R., Debattista V. P., Loebman S. R., 2013, MNRAS, 433, 976

Ruchti G. R., et al., 2015, MNRAS, 450, 2874

Russeil D., Zavagno A., Mège P., Poulin Y., Molinari S., Cambresy L., 2017, A&A, 601, L5

Ryde N., et al., 2010, A&A, 509, A20

Sales L. V., et al., 2009, MNRAS, 400, L61

Sales L. V., Navarro J. F., Schaye J., Dalla Vecchia C., Springel V., Booth C. M., 2010, MNRAS, 409, 1541

Salvadori S., Ferrara A., Schneider R., Scannapieco E., Kawata D., 2010, MNRAS, 401, L5

Sandage A., Visvanathan N., 1978, ApJ, 225, 742

Santistevan I. B., Wetzel A., Sanderson R. E., El-Badry K., Samuel J., Faucher-Giguère C.-A., 2021, MNRAS, 505, 921

Sarkar S., Jog C. J., 2019, A&A, 628, A58

Sawala T., et al., 2016, MNRAS, 457, 1931

Scannapieco C., Tissera P. B., White S. D. M., Springel V., 2008, MNRAS, 389, 1137

Scannapieco C., White S. D. M., Springel V., Tissera P. B., 2009, MNRAS, 396, 696

Schaye J., et al., 2015, MNRAS, 446, 521

Schlaufman K. C., Casey A. R., 2014, ApJ, 797, 13

Schmidt M., 1959, ApJ, 129, 243

Schönrich R., Binney J., 2009, MNRAS, 396, 203

Schörck T., et al., 2009, A&A, 507, 817

Schwarzkopf U., Dettmar R. J., 2000, A&A, 361, 451

Searle L., Zinn R., 1978, ApJ, 225, 357

Semczuk M., Łokas E. L., D'Onghia E., Athanassoula E., Debattista V. P., Hernquist L., 2020, MNRAS, 498, 3535

Sérsic J. L., 1963, Boletin de la Asociacion Argentina de Astronomia La Plata Argentina, 6, 41

Sesar B., Jurić M., Ivezić Ž., 2011, ApJ, 731, 4

Sestito F., et al., 2019, MNRAS, 484, 2166

Sestito F., et al., 2020, MNRAS, 497, L7

Sestito F., et al., 2021, MNRAS, 500, 3750

Sick J., Courteau S., Cuillandre J.-C., Dalcanton J., de Jong R., McDonald M., Simard D., Brent Tully R., 2014, Proceedings of the International Astronomical Union, 10, 82–85

Siegel M. H., Majewski S. R., Reid I. N., Thompson I. B., 2002, ApJ, 578, 151

Sijacki D., Springel V., Di Matteo T., Hernquist L., 2007, MNRAS, 380, 877

Sijacki D., Vogelsberger M., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 424, 2999

Silk J., 1977, ApJ, 211, 638

Silk J., Mamon G. A., 2012, Research in Astronomy and Astrophysics, 12, 917

Silk J., Rees M. J., 1998, A&A, 331, L1

Skúladóttir Á., et al., 2021, ApJ, 915, L30

Smith M. C., et al., 2007, MNRAS, 379, 755

Sneden C., et al., 2003, ApJ, 591, 936

Somerville R. S., Davé R., 2015, ARA&A, 53, 51

Sotillo-Ramos D., et al., 2022, MNRAS, 516, 5404

Sotillo-Ramos D., Donnari M., Pillepich A., Frankel N., Nelson D., Springel V., Hernquist L., 2023a, Monthly Notices of the Royal Astronomical Society

Sotillo-Ramos D., Bergemann M., Friske J. K. S., Pillepich A., 2023b, MNRAS, 525, L105

Sparre M., Springel V., 2016, MNRAS, 462, 2418

Sparre M., Springel V., 2017, MNRAS, 470, 3946

Sparre M., Whittingham J., Damle M., Hani M. H., Richter P., Ellison S. L., Pfrommer C., Vogelsberger M., 2022, MNRAS, 509, 2720

Spina L., et al., 2021, MNRAS, 503, 3279

Spitzer Lyman J., 1942, ApJ, 95, 329

Springel V., 2005, MNRAS, 364, 1105

Springel V., 2010a, ARA&A, 48, 391

Springel V., 2010b, MNRAS, 401, 791

Springel V., 2010c, MNRAS, 401, 791

Springel V., Hernquist L., 2003, MNRAS, 339, 289

Springel V., Hernquist L., 2005, ApJ, 622, L9

Springel V., Yoshida N., White S. D. M., 2001a, New Astron., 6, 79

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001b, MNRAS, 328, 726

Springel V., Di Matteo T., Hernquist L., 2005, MNRAS, 361, 776

Springel V., et al., 2018, MNRAS, 475, 676

Starkenburg E., et al., 2017, MNRAS, 471, 2587

Starobinsky A. A., 1979, ZhETF Pisma Redaktsiiu, 30, 719

Steinmetz M., et al., 2006, AJ, 132, 1645

Steinmetz M., et al., 2020, American Astronomical Society, 160, 82

Stewart K. R., Bullock J. S., Wechsler R. H., Maller A. H., Zentner A. R., 2008, ApJ, 683, 597

Stewart K. R., Bullock J. S., Wechsler R. H., Maller A. H., 2009, ApJ, 702, 307

Stinson G. S., Brook C., Macciò A. V., Wadsley J., Quinn T. R., Couchman H. M. P., 2013, MNRAS, 428, 129

Takahashi K., Yoshida T., Umeda H., 2018, ApJ, 857, 111

Tamm A., Tempel E., Tenjes P., Tihhonova O., Tuvikene T., 2012, A&A, 546, A4

Tegmark M., et al., 2004, Phys. Rev. D, 69, 103501

Tempel E., Saar E., Liivamägi L. J., Tamm A., Einasto J., Einasto M., Müller V., 2011, A&A, 529, A53

Teyssier R., 2002, A&A, 385, 337

Ting Y.-S., Rix H.-W., 2019, ApJ, 878, 21

Toomre A., 1977, in Tinsley B. M., Larson Richard B. Gehret D. C., eds, Evolution of Galaxies and Stellar Populations. p. 401

Torrey P., et al., 2015a, MNRAS, 447, 2753

Torrey P., et al., 2015b, MNRAS, 454, 2770

Toth G., Ostriker J. P., 1992, ApJ, 389, 5

Tsukui T., Iguchi S., 2021, Science, 372, 1201

VERA Collaboration et al., 2020, PASJ, 72, 50

Vasiliev E., Belokurov V., Erkal D., 2021, MNRAS, 501, 2279

Venn K. A., et al., 2020, MNRAS, 492, 3241

Vera-Ciro C., D'Onghia E., Navarro J., Abadi M., 2014, ApJ, 794, 173

Vera-Ciro C., D'Onghia E., Navarro J. F., 2016, ApJ, 833, 42

Villalobos Á., Helmi A., 2008, MNRAS, 391, 1806

Vogelsberger M., Sijacki D., Kereš D., Springel V., Hernquist L., 2012, MNRAS, 425, 3024

Vogelsberger M., et al., 2014a, MNRAS, 444, 1518

Vogelsberger M., et al., 2014b, Nature, 509, 177

Vogelsberger M., Marinacci F., Torrey P., Puchwein E., 2020, Nature Reviews Physics, 2, 42

Wang H.-F., Liu C., Xu Y., Wan J.-C., Deng L., 2018, MNRAS, 478, 3367

Watkins L. L., et al., 2009, MNRAS, 398, 1757

Wegg C., Gerhard O., Portail M., 2015, MNRAS, 450, 4050

Weinberger R., et al., 2017, MNRAS, 465, 3291

Weinberger R., et al., 2018, MNRAS, 479, 4056

Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016a, ApJ, 827, L23

Wetzel A. R., Hopkins P. F., Kim J.-h., Faucher-Giguère C.-A., Kereš D., Quataert E., 2016b, ApJ, 827, L23

White S. D. M., Rees M. J., 1978, MNRAS, 183, 341

White S. D. M., Springel V., 2000, in Weiss A., Abel T. G., Hill V., eds, The First Stars. p. 327 (arXiv:astro-ph/9911378), doi:10.1007/10719504_62

Worthey G., España A., MacArthur L. A., Courteau S., 2005, ApJ, 631, 820

Wyse R. F. G., 2001, in Funes J. G., Corsini E. M., eds, Astronomical Society of the Pacific Conference Series Vol. 230, Galaxy Disks and Disk Galaxies. pp 71–80 (arXiv:astro-ph/0012270)

Xiang M., Rix H.-W., 2022, Nature, 603, 599

Xiang M., et al., 2017, ApJS, 232, 2

Xiang M., et al., 2018, ApJS, 237, 33

Yoachim P., Dalcanton J. J., 2006, AJ, 131, 226

Yong D., 2020, in Bragaglia A., Davies M., Sills A., Vesperini E., eds, A Vol. 351, Star Clusters: From the Milky Way to the Early Universe. pp 19–23, doi:10.1017/S1743921319007361

York D. G., et al., 2000, AJ, 120, 1579

Youakim K., et al., 2020, MNRAS, 492, 4986

Yu S., et al., 2022, arXiv e-prints, p. arXiv:2210.03845

Zeng G., Wang L., Gao L., 2021, arXiv e-prints, p. arXiv:2105.09722

Zhao G., Zhao Y.-H., Chu Y.-Q., Jing Y.-P., Deng L.-C., 2012, Research in Astronomy and Astrophysics, 12, 723

Zhao D., Du M., Ho L. C., Debattista V. P., Shi J., 2020, ApJ, 904, 170

Zhu L., et al., 2022a, A&A, 660, A20

Zhu L., et al., 2022b, A&A, 664, A115

Zibetti S., White S. D. M., Brinkmann J., 2004, MNRAS, 347, 556

Zolotov A., Hogg D. W., Willman B., 2011, ApJ, 727, L14

Zucker D. B., de Silva G., Freeman K., Bland-Hawthorn J., Hermes Team 2012, in Aoki W., Ishigaki M., Suda T., Tsujimoto T., Arimoto N., eds, Astronomical Society of the Pacific Conference Series Vol. 458, Galactic Archaeology: Near-Field Cosmology and the Formation of the Milky Way. p. 421

Declaration of authorship

I hereby declare that the report submitted is my own unaided work. All direct or indirect sources used are acknowledged as references. I am aware that the Thesis in digital form can be examined for the use of unauthorized aid and in order to determine whether the report as a whole or parts incorporated in it may be deemed as plagiarism. For the comparison of my work with existing sources I agree that it shall be entered in a database where it shall also remain after examination, to enable comparison with future Theses submitted. Further rights of reproduction and usage, however, are not granted here. This paper was not previously presented to another examination board and has not been published.

Heidelberg, 11th September 2023

Diego Sotillo Ramos