



MEMORIA FINAL BECA DE VERANO 2024

INSTITUTO DE ASTROFÍSICA DE CANARIAS

TÍTULO DEL PROYECTO / PROJECT TITLE

Observing the formation of the Milky Way with the James Webb Space Telescope.

INVESTIGADORES PRINCIPALES / PRINCIPAL INVESTIGATORS

Ignacio Martín Navarro, Adriana de Lorenzo-Cáceres Rodríguez, Begoña García Lorenzo

AUTOR / AUTHOR

Fernando Valenciano Ruano

ABSTRACT

Understanding the Milky Way means understanding galaxy evolution, as our models of the physical processes shaping the Universe are calibrated to reproduce the observed properties of our own galaxy. We have access to an unprecedented range of measurements, from the mass of the central supermassive black hole to the detailed chemical composition of individual stars. This project seeks to address a simple yet fundamental question: how would the Milky Way appear if viewed from a different galaxy?

The methodology followed relies on a combination of the publicly available data on Milky Way-like galaxies from the TNG50 simulation with the MILES stellar population models. By associating the age and metallicity values of TNG50 stellar particles with those of MILES, we can generate Integral Field Units (IFU) datacubes of simulated galaxies. We then simulate observations using current-generation Integral Field Spectrographs (IFS), such as NIRSpec onboard the James Webb Space Telescope (JWST) and HARMONI on the future ESO's Extremely Large Telescope (ELT), providing an unique perspective on our galaxy from an external vantage point. The main objective of this work is to develop a comprehensive and consistent Python pipeline to generate these simulated IFU datacubes, given a TNG50 galaxy, at a specified redshift, field of view (FOV), and line of sight (LOS). The results of this work are made publicly available via the `TNG_mockIFU` package on [GitHub](#). This package is intended for use in future scientific studies and to derive technical parameters for proposed observations with these advanced observational instruments.

RESUMEN

Entender la Vía Láctea es fundamental para comprender la evolución de las galaxias, ya que los modelos de los procesos físicos que configuran el Universo están calibrados para reproducir las propiedades observadas de nuestra propia galaxia. Disponemos de un rango sin precedentes de medidas, que van desde la masa del agujero negro supermasivo central hasta la composición química detallada de estrellas individuales. Este proyecto pretende abordar una pregunta simple pero fundamental: ¿cómo aparecería la Vía Láctea si se viera desde otra galaxia?

La metodología empleada se basa en una combinación de datos públicos sobre galaxias similares a la Vía Láctea obtenidos de la simulación TNG50 y los modelos de poblaciones estelares MILES. Al asociar los valores de edad y metalicidad de las partículas estelares de TNG50 con los de MILES, podemos generar cubos de datos de Integral Field Units (IFU) para galaxias simuladas. Posteriormente, simulamos observaciones utilizando espejorógrafos de campo integral de última generación, como NIRSpec a bordo del James Webb Space Telescope (JWST) y HARMONI en el futuro Extremely Large Telescope (ELT) del ESO. El objetivo principal de este trabajo es desarrollar un procedimiento completo y consistente en Python para producir dichos cubos de datos IFU simulados, dado una galaxia de TNG50, un redshift especificado, un campo de visión (FOV) o una línea de visión (LOS). Los resultados de este trabajo están disponibles públicamente a través del paquete `TNG_mockIFU` en [GitHub](#). Este paquete está destinado a ser utilizado en futuros estudios científicos y para proveer a la comunidad científica los parámetros técnicos requeridos en futuras propuestas de observación con dichos instrumentos.

Contents

1	Introduction and current status of the research	1
1.1	Galaxy evolution and formation	1
1.2	Milky Way-like galaxies	1
1.3	James Webb Space Telescope	2
2	Hypothesis and objectives	3
3	Methodology and materials	4
3.1	Illustris TNG50 simulation	4
3.2	MILES stellar population models	8
3.3	Integral spectroscopic datacubes	12
3.4	Simulate JWST observations	15
3.5	Simulate HARMONI observations	17
4	Results and discussion	18
5	Conclusions	19

1 Introduction and current status of the research

1.1 Galaxy evolution and formation

Understanding the formation and evolution of structures such as our galaxy, the Milky Way, is one of the main challenges of galaxy studies. This evolution is governed by a dynamical balance between processes that trigger their formation and other that prevent it, such as gas expulsion or heating. The formation of galaxies generally follows a pathway influenced by the complex interplay between the gravitational potential of the dark matter halos which initiates the collapse of dense molecular clouds where star formation occurs [1], and the thermodynamic properties of the circumgalactic medium (CGM), alongside mechanisms known as feedback processes.

In a general structure formation scenario, dark matter halos act as the seeds for the matter overdensities, where smaller halos merge to form larger and more massive halos [2]. Concurrently, there is a continuous accretion of the intergalactic medium, which supplies gas to clouds that undergo gravitational collapse, leading to star formation and the assembly of a galaxy comprising gas, dust and a central supermassive black hole (SMBH). Dark matter halos experience a process known as *hierarchical clustering*, in the *Cold Dark Matter* (CDM) paradigm we have a *bottom-up* hierarchical scenario: the smaller structures form first in the Universe, assembling virialised halos, and by accretion and merging of these, large and massive structures arise [3] (see [4] for a comprehensive review). These merging events critically influence the evolution of galaxies over cosmic time by affecting the resulting stellar populations.

Star formation is also influenced by the mass of the galaxy. For galaxies with masses below $10^{10.5} M_{\odot}$ supernova feedback is expected to be the dominant mechanism suppressing star formation. In contrast, for more massive galaxies, feedback from active galactic nucleus (AGN) plays a more significant role. Within the mass range of $10^{10} - 10^{11} M_{\odot}$, a phenomenon known as galaxy bimodality is observed. The red sequence consists of galaxies with higher luminosities, older stellar populations, and higher metallicities, typically elliptical and lenticular galaxies, while the blue sequence includes galaxies with lower luminosities, younger stellar populations, and lower metallicities, such as spiral galaxies. The age of a galaxy is associated with the epoch of star formation, which occurs during the collapse of gas clouds, whereas metallicity is linked to energetic processes, such as supernova events, in which massive stars expel their material into the interstellar medium. Massive stars ($> 8 M_{\odot}$) enrich the interstellar medium through stellar winds or core-collapse as Type II supernovae in events that last approximately $\sim 10^8$ years [5]. In contrast, Type Ia supernovae, originating from binary systems, occur over longer timescale of about 1 Gyr. Therefore, it is generally more effective to consider that a galaxy's enrichment is predominantly driven by Type II supernovae.

1.2 Milky Way-like galaxies

The Milky Way is a spiral barred galaxy in which the majority of its baryonic content is located within the thin disk, which has a mass of approximately $5 \cdot 10^{10} M_{\odot}$, a radius of ~ 3.5 kpc, and a vertical thickness of ~ 0.3 kpc [6]. The Sun is situated at a distance of approximately 8 kpc from the galactic center, with a rotational velocity of ~ 220 , km s $^{-1}$. In addition, the Milky Way's thin disk hosts a thick disk that contains about 10 – 20% of the thin disk's mass, a bulge at the center with a mass of $\sim 10^{10} M_{\odot}$, and a stellar halo that contains approximately 3% of the disk's mass. A SMBH is located at the center with a mass of $\sim 2 \cdot 10^6 M_{\odot}$.

In less than a decade, our understanding of the Milky Way has been deeply revolutionized thanks to the success of the Gaia mission [7]. The synergies with deep spectroscopic follow-ups

[8] have enabled an unprecedented characterization of the internal structure [9], dynamical state [10], and chemical composition of our galaxy [11].

These results have revealed that the formation processes of the Milky Way are characterized by three continuous yet relatively distinct phases. The initial *protogalactic stage* is characterized by the formation of metal-poor stars in a stochastic and turbulent environment, resulting in a pressure-supported system that marks the beginning of the Milky Way. This is followed by a *spin-up* phase, a rapid transition towards a fast-rotating disk occurring approximately 2 Gyr after the Big Bang [12]. Finally, the *settle-down* phase results in the late formation of the thin disk of our galaxy, likely associated with the major disruptive event of the Gaia-Sausage Enceladus merger.

Open questions

While the Gaia mission has significantly enhanced our understanding of the Milky Way's formation history, several limitations remain that require further advancements and surveys. Firstly, precise age measurements required to probe the critical early stages of disk formation are currently unfeasible. This presents a significant challenge for studies of ancient stellar populations, as age primarily affects the spectra of long-lived, low-mass stars. Secondly, our understanding of the gas accretion and cooling processes that led to the formation of the Milky Way's disk(s) remains limited due to the current state of knowledge about its CGM. These challenges give rise to several open questions: *Did Milky Way-like disks develop a central metal-poor component? How have star formation, internal structure, and chemical composition evolved over time? Is disk formation driven by cold or hot accretion into dark matter halos? How do Milky Way-like disks form?*

In order to shed light onto the previous questions, different surveys have been proposed, including large-scale observational campaigns and theoretical modeling efforts. Among them, deep spectroscopic observations of Milky Way-like galaxies can offer valuable insights into their stellar populations. In particular, at certain redshifts in which the Milky Way suffered a significant change, such as at the time of the Gaia-Sausage Enceladus merger ($z \sim 1.5$). For this kind of observations, the James Webb Space Telescope provides the community data with unprecedent depth and spatial resolution, therefore, is the most suitable instrument. In the following subsection we proceed to provide a brief description of this state-of-the-art instrument.

1.3 James Webb Space Telescope

The James Webb Space Telescope (JWST) was launched in December 2021 with four principal scientific objectives [13]: to study the Dark Ages of the universe and the epoch of reionization; to investigate the mechanisms of galaxy formation; to observe star formation and protoplanetary systems; and to explore exoplanets and the origins of life. The telescope features a primary mirror with a diameter of 6.6 meters, composed of 18 hexagonal segments and equipped with a solar shield, which is essential for conducting infrared observations at low temperatures. JWST is equipped with four main instruments: MIRI (Mid-Infrared Instrument), NIRCam (Near-Infrared Camera), NIRISS (Near-Infrared Imager and Slitless Spectrograph), and NIRSpec (Near-Infrared Spectrograph). Each instrument plays a crucial role in fulfilling the telescope's scientific goals, providing a broad range of observational capabilities from mid- to near-infrared wavelengths.

The NIRSpec (Near-Infrared Spectrograph) is the primary instrument for conducting spectroscopic observations in the near-infrared range, covering wavelengths from 0.6 to 5.3 μm [14]. It offers spectral resolutions of approximately 100, 1000, and 2700, depending on the observational setup. NIRSpec is particularly well-suited for spectroscopic surveys focused on galaxy formation and evolution, enabling detailed characterization of stellar populations

2 Hypothesis and objectives

The primary objective of this study is to develop integral field spectroscopic datacubes of simulated Milky Way-like galaxies at $z \sim 1.5$ using the Illustris TNG50 cosmological simulation, and to simulate how these galaxies would appear through the JWST NIRSpec spectrograph. This work aims to support ongoing observational efforts by the host research group at the Instituto de Astrofísica de Canarias (IAC) to study a sample of Milky Way-like galaxies with the JWST. This approach serves as a direct test of JWST's capabilities to constrain the formation history of Milky Way-like galaxies, by combining insights from both cosmological simulations and observational data. To achieve this goal, the study is structured around the following short-term objectives:

- (i) Utilize the Illustris TNG50 simulation to generate two-dimensional maps of the stellar population properties of Milky Way-like galaxies at $z \sim 1.5$.
- (ii) Employ the MILES stellar library [15] to generate synthetic spectra for the different stellar populations within the simulated Milky Way-like galaxy.
- (iii) Investigate different methodologies for assigning MILES age and metallicity values to the mock galaxy and analyze the most suitable method for our objectives.
- (iv) Develop a computational pipeline for generating spatial-spectral datacubes with integral field units (IFUs) for TNG galaxies. The framework should incorporate input parameters such as the subhalo ID, snapshot (corresponding to redshift), galaxy radius (which affects the field of view, FOV), and spaxel size (or, alternatively, the number of spaxels).
- (v) Translate these physical properties into datacubes as they would be observed through the JWST [NIRSpec](#) spectrograph, taking into account factors such as spatial and spectral resolution, field of view (FOV), and spaxel size relevant to the observational parameters.
- (vi) As an additional objective, utilize the synthetic datacubes to simulate observations of the galaxy using the [HARMONI](#) spectrograph for the future Extremely Large Telescope (ELT) [16].

With this work, we aim to obtain a robust and accurate methodology to translate Milky Way-like galaxies from hydrodynamical simulations, to the spectroscopic three-dimensional datacubes that the current generation of telescopes such as the JWST will observe. This results will provide advanced technical parameters to enhance the description for future observational proposals for the JWST.

The code developed through the internship, `TNG_mockIFU` is publicly available on [GitHub](#), so everyone can use it to develop their own IFU datacubes from all TNG50 galaxies.

3 Methodology and materials

Once the galaxy of TNG50 has been selected, we aim to generate a mock JWST-like datacube. We describe the steps followed below, and the methodology is summarized in Fig. 1.

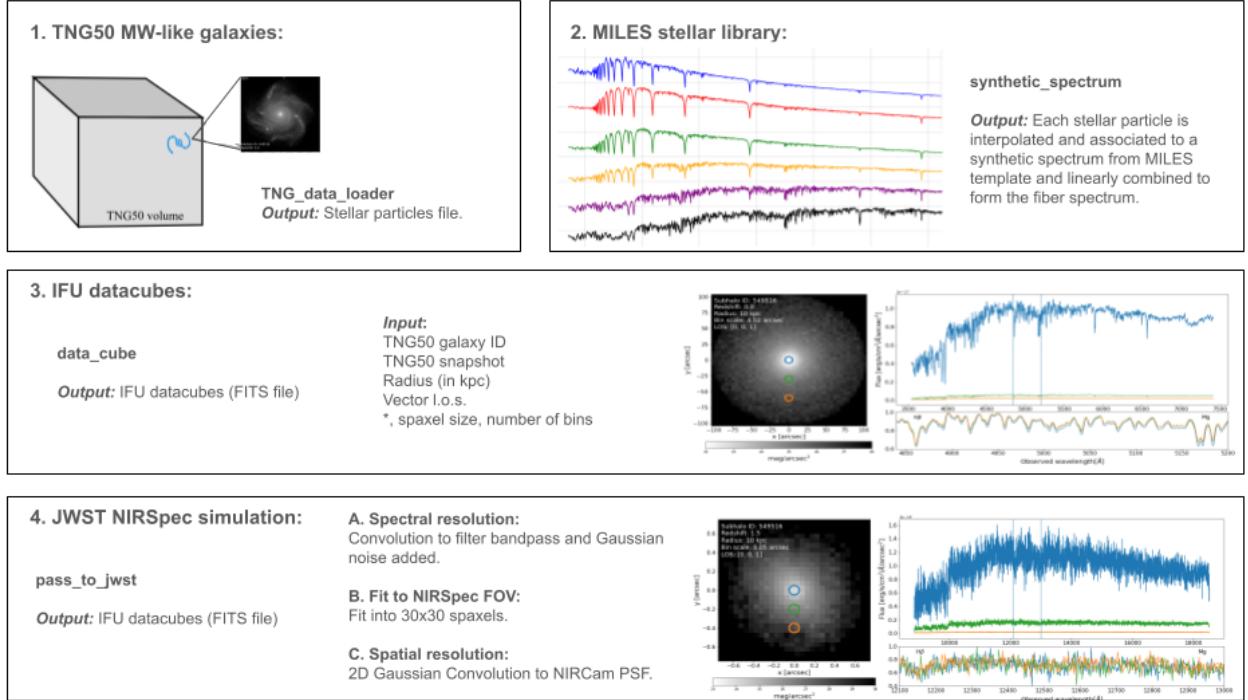


Figure 1: Scheme followed to build a JWST-like datacube from a TNG50 simulated galaxy.

3.1 Illustris TNG50 simulation

TNG50 [17, 18] is a cosmological magneto-hydrodynamical (MHD) simulation that models the formation and evolution of galaxies within a Λ CDM universe [19]. It employs the moving-mesh code AREPO [20] and incorporates the fiducial TNG galaxy formation model [21]. TNG50 evolves components including cold dark matter (CDM), gas, stars, SMBHs, and magnetic fields within a cubic volume of 51.7 comoving Mpc on each side, spanning from $z = 127$ to the present epoch and encompassing thousands of galaxies. The simulation achieves a resolution with a target gas mass cell and stellar particle mass of $8.5 \cdot 10^4 M_{\odot}$, and a dark matter particle mass of $4.6 \cdot 10^5 M_{\odot}$. This places TNG50 at the intersection of large-volume and zoom-in hydrodynamical galaxy simulations.

The simulation introduce astrophysical processes including primordial and metal-line cooling; heating from an homogeneous UV/X-ray background and X-ray radiation from SMBHs; a density threshold based star formation; stellar-driven outflows; seeding, growth and feedback of SMBHs; evolution of cosmic magnetic fields via MHD; stellar evolution, stellar mass loss, and enrichment from SNII and SNIa tracked via nine elements (see [21] for a complete description).

A. TNG50 MW/M31 analogs:

One of the significant achievements of current MHD simulations is their ability to reproduce galaxies with properties and structures analogous to those of the Milky Way. For instance, IllustrisTNG Collaboration presents a selection of Milky Way/M31-like galaxies derived from cosmological TNG50 simulations [22]. Specifically, they establish several criteria that must be simultaneously fulfilled for galaxies in the TNG50 simulation at $z = 0$ to be included in the catalog of Milky Way/M31-like galaxies:

- (i) Stellar mass: the galaxy stellar mass must be in the following range $M_{\odot}(< 30\text{kpc}) = 10^{10.5-11.2} M_{\odot}$
- (ii) Stellar morphology: galaxies with a stellar disk-like morphology, including the presence of spiral arms. This is achieved by establishing a relationship between the minor-to-major axis ratio, or by observing the presence of arms in three-band images.
- (iii) Environment: no other galaxy with stellar mass $\leq 10^{10.5} M_{\odot}$ is within 500 kpc distance.

With this definition, they are left with 198 MW/M31-like galaxies. Therefore, this simulated galaxies can be used as mock galaxies for deeper studies, as we are able to obtain all the information regarding its gas, dark matter, stars and SMBH in different redshift snapshots.

The results from IllustrisTNG are publicly available and can be accessed through the [Public Data Access](#) provided by the IllustrisTNG Collaboration. For our specific interest in Milky Way/M31-like galaxies, data can be obtained from the [TNG50 MW/M31 Sample and Data Release](#), where we can retrieve the datacubes for each galaxy based on their subhalo ID and corresponding snapshot.

The initial step involves filtering the data by specifying a given radius and inclination. This is accomplished using the `TNG_data_loader` function, where the radius in kpc and the vector corresponding to the desired line of sight (with the galaxy's minor axis oriented along the z-axis) must be defined. This filtering process also excludes wind particles that are included in the same dataset as the stellar particles by removing particles with negative values in the ("GFM_StellarFormationTime") field.

Following this, a preliminary analysis is conducted to familiarize ourselves with the type of data available. For particle positions tracked by the simulation, we utilize the ("RotatedCoordinates") dataset. Figures 2 and 3 illustrate the distribution of different components—gas, dark matter, and stars—at a fixed radius of 50 kpc, shown for face-on and edge-on views of a Milky Way-like galaxy at redshifts $z = 1.5$ and $z = 0$, respectively. This visualization is generated using the `mk_hist_plots` function from the package developed during the summer internship.

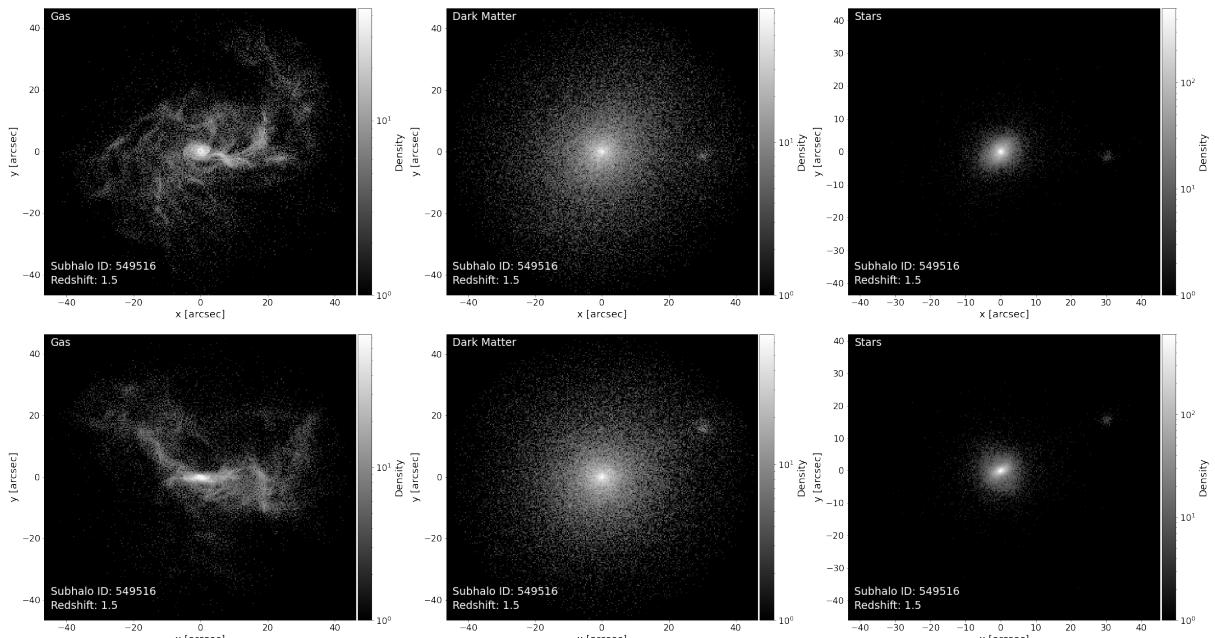


Figure 2: Mass density of gas (*left*), dark matter (*middle*) and stars (*right*) for the galaxy ID 549516 at the snapshot 40 ($z = 1.5$) and a FOV of 100kpc. Upper (lower) panels correspond to the face-(edge-) on visualization.

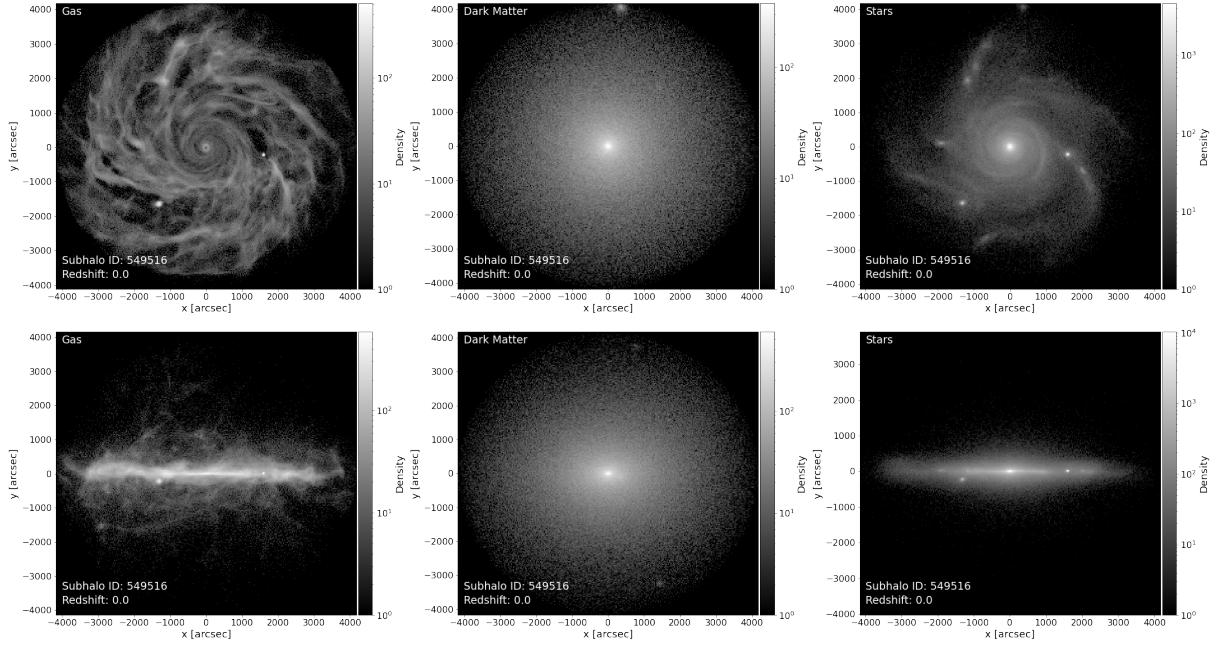


Figure 3: Mass density of gas (*left*), dark matter (*middle*) and stars (*right*) for the galaxy ID 549516 at the snapshot 99 ($z = 0$) and a FOV of 100kpc. Upper (lower) panels correspond to the face- (edge-) on visualization.

B. Stellar population properties:

As outlined in the objectives, our focus is on the stellar particles from the TNG50 simulation. We will extract the following stellar datasets, which contain the necessary information for constructing the IFU datacubes:

- (i) (“GFM_Metallicity”): To convert the metallicity into relative abundance units [Fe/H], one must divide by the solar metallicity value of $Z_{\odot} = 0.0127$ and then take the logarithm.
- (ii) (“GFM_StellarFormationTime”): The age is computed changing the stellar formation time (given in terms of the scale factor) into redshift with the following equation:

$$a(z) = \frac{1}{1+z} \quad (1)$$

Assuming a Planck16 cosmology [19], the redshift can be converted into age in gigayears (Gyr) using the *astropy.cosmology.Planck16* package.

- (iii) (“GFM_InitialMass”): Stellar mass of each stellar particle, in units of $10^{10} M_{\odot}$.
- (iv) (“RotatedCoordinates”): Position of each stellar particle in units of comoving kpc/ h .
- (v) (“RotatedVelocities”): Tangential velocity will be useful in the implementation of the effect of stellar kinematics to their spectrum. They are referenced to the center of mass of the galaxy and in units of km/s.

With this information, we can identify the essential components required for developing the datacubes. Notably, for a galaxy with a radius of 10 kpc, the number of stellar particles exceeds 700,000 at redshift $z = 0$. Consequently, minimizing the number of datasets extracted will help to reduce RAM usage. This process is managed by the `TNG_data_loader` function. Figure 4 and Fig. 5 illustrates the distribution of metallicity, age, and tangential velocity for a given galaxy at $z = 1.5$ and $z = 0$, respectively.

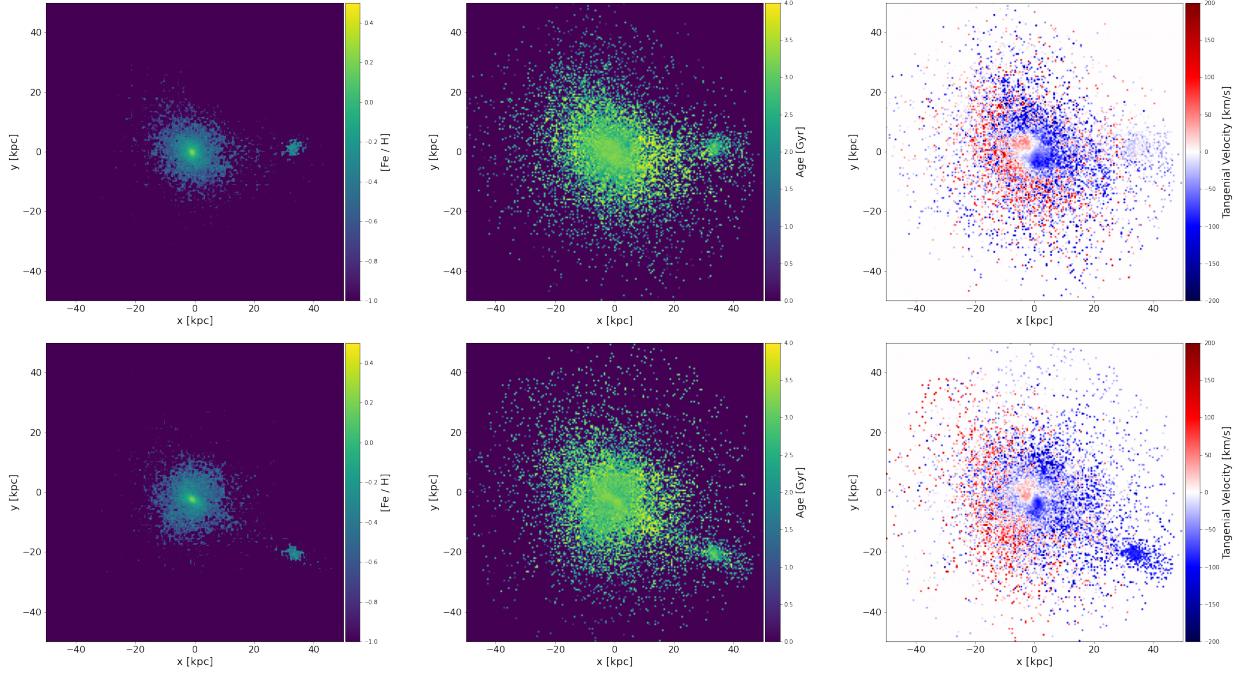


Figure 4: Stellar particles histograms of metallicity (*left*) age (*middle*) and tangential velocity (*right*) for the galaxy ID 549516 at the snapshot 40 ($z = 1.5$) and a FOV of 100kpc (radius of 50kpc). Upper (lower) panels correspond to the face- (edge-)on visualization.

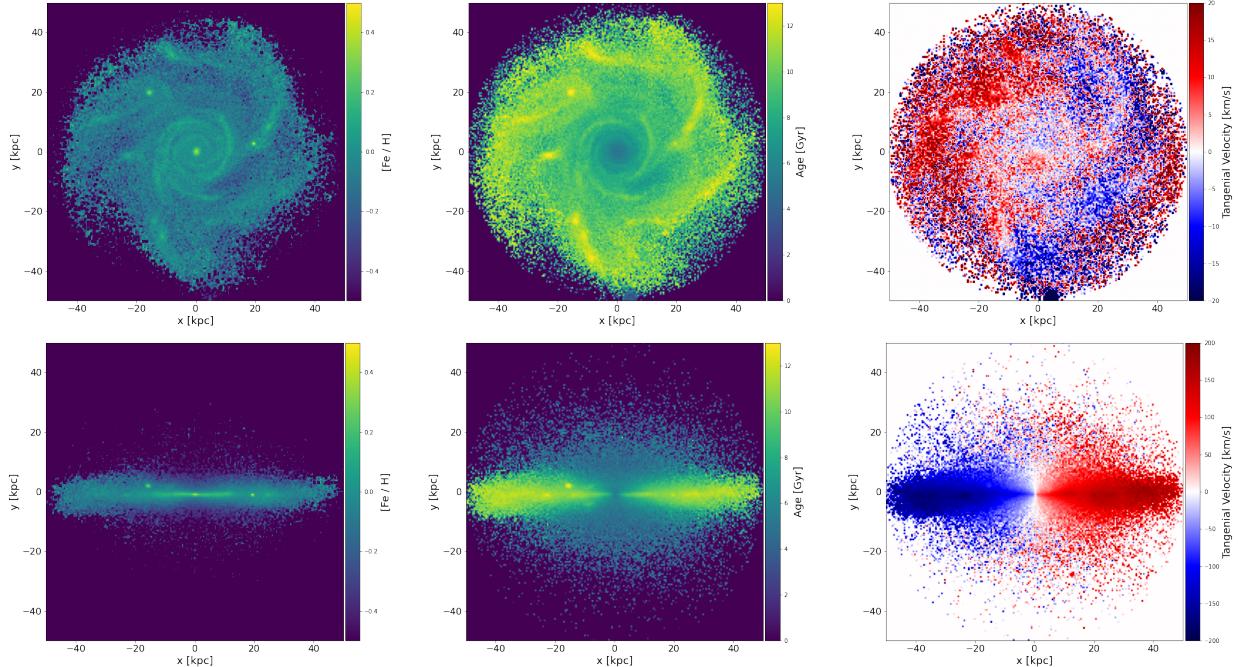


Figure 5: Stellar particles histograms of metallicity (*left*) age (*middle*) and tangential velocity (*right*) for the galaxy ID 549516 at the snapshot 99 ($z = 0$) and a FOV of 100kpc (radius of 50kpc). Upper (lower) panels correspond to the face- (edge-)on visualization.

3.2 MILES stellar population models

At this stage, we introduce the tool for assigning a synthetic spectrum to a given stellar population using the MILES models [15]. These models consist of single stellar population (SSP) spectra characterized by their age and metallicity values, covering a spectral range from $\lambda\lambda$ 3535 to 7500 Å. The construction of this library involves an empirical stellar library, theoretical isochrones, and modeling of the initial mass function (IMF). The empirical stellar library includes spectra of 985 stars from the solar neighborhood, spanning a broad range of atmospheric parameters. These spectra were obtained using the 2.5m Isaac Newton Telescope and cover the range $\lambda\lambda$ 3525 to 7500 Å with a spectral resolution of 2.3 Å (full width at half maximum, FWHM) [23]. The isochrones utilized are those derived by Basti [24], which incorporate stellar evolution in terms of effective temperature, luminosity, and metallicity. The IMF is adopted from Kroupa [25] and follows a power-law distribution similar to the Salpeter IMF [26] with a typical slope of 1.35 for stars with masses above $0.5M_{\odot}$, but with a flatter distribution for lower mass stars.

It is important to note that the shape of the synthetic spectrum of a stellar population is fully determined by its age and metallicity values. In Fig. 6 we illustrate the spectra for different stellar populations from the MILES models, in which we can distinguish a relevant variation depending on their age and metallicity values.

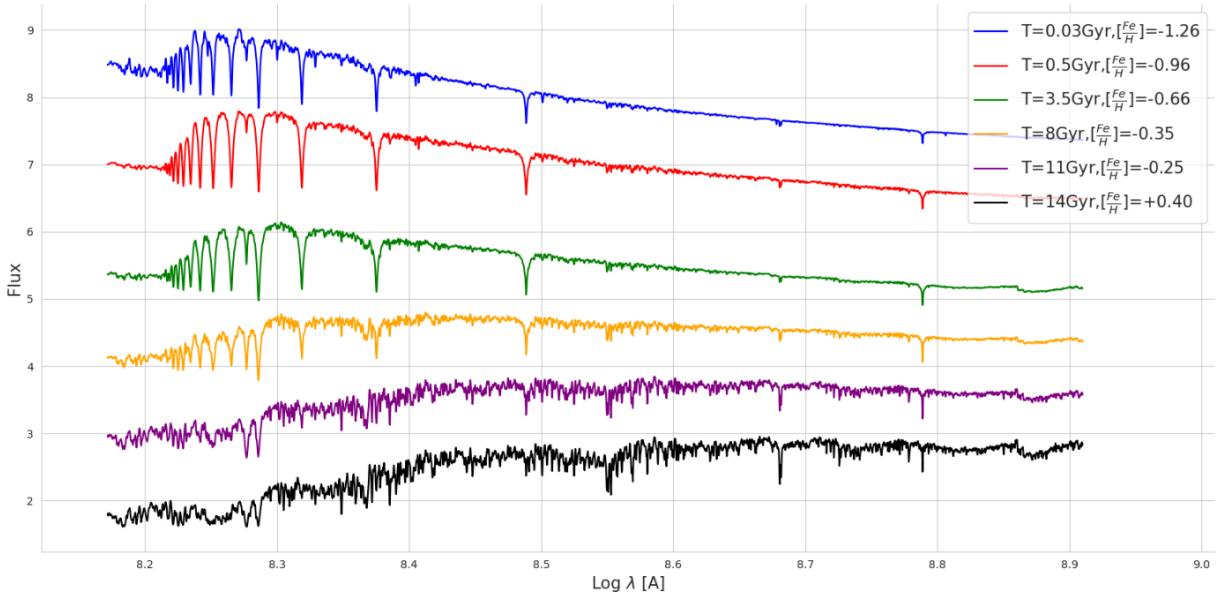


Figure 6: Different synthetic spectra obtained from the MILES stellar library [15] for different age and metallicity values.

A. Association between TNG50 and MILES age and metallicity values

The MILES library consists of discrete age and metallicity values, with 12 metallicity values spanning the range $[-2.27, 0.4]$ and 53 age values ranging from 0.03 to 14 Gyr. In contrast, the TNG50 simulation provides a continuous distribution of age and metallicity values, as illustrated in Figure 7 for the same galaxy at $z = 1.5$ (upper panels) and $z = 0$ (lower panels). Notably, the metallicity distribution for the galaxy at higher redshift appears shifted towards lower values. This observation can be attributed to cosmic time; earlier epochs generally have fewer supernovae, resulting in less enrichment of the circumstellar medium with metals.

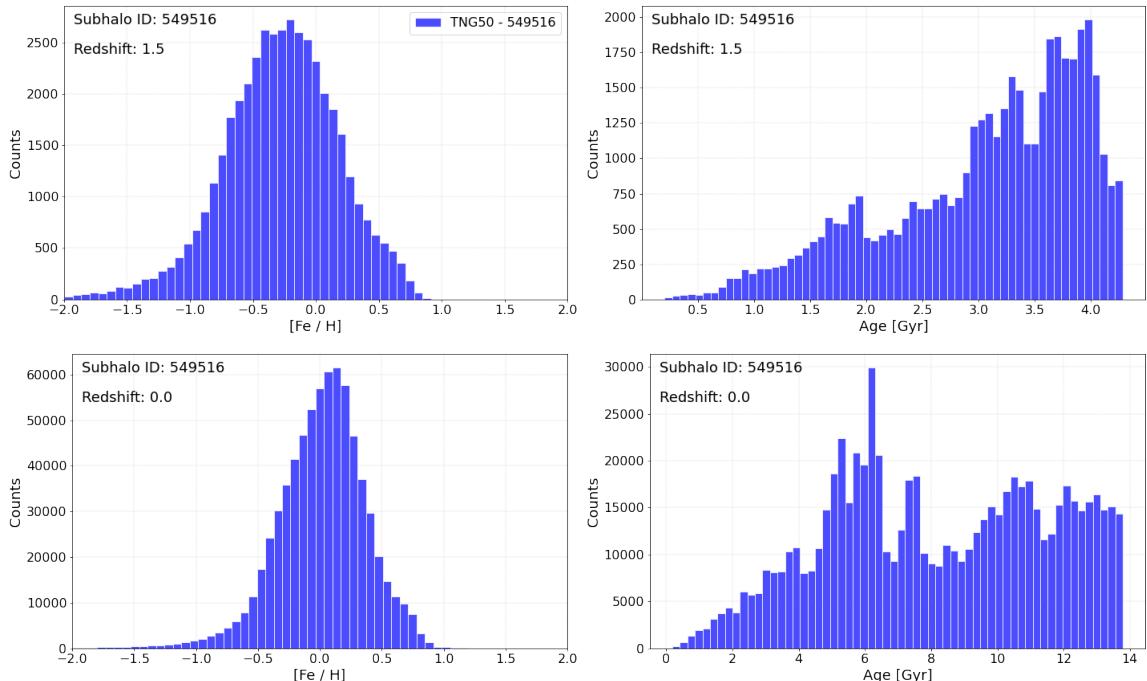


Figure 7: Histogram of the metallicity (*left*) and age (*right*) distribution of the stellar particles of the galaxy ID 549516 with a radius of 10kpc for the snapshots at $z = 1.5$ (upper panels) and $z = 0$ (lower panels).

At this stage, it is necessary to discuss how to associate the MILES age and metallicity values with each TNG stellar particle to assign a synthetic spectrum to each particle. As shown in Figure 7, TNG does not include a significant number of stellar particles with metallicities below $[Fe/H] = -1.5$. Consequently, we can exclude the MILES metallicity values of -2.27 and -1.79 to optimize computational resources. The association of MILES spectra to TNG star particles can be performed using three primary methods:

- (i) **Nearest Value:** This approach involves assigning to each TNG stellar particle the MILES age and metallicity value closest to that of the particle. The integrated synthetic spectrum is then computed by multiplying each assigned spectrum by the corresponding stellar mass and summing the results.
- (ii) **Interpolation:** This widely-used technique interpolates between MILES spectra for the permitted age and metallicity values. For metallicity values that exceed the maximum provided by MILES, the highest available value is assigned. The interpolation process utilizes the *scipy.interpn* function.
- (iii) **Extrapolation:** This method extends the interpolation approach by linearly extrapolating for metallicities beyond the maximum MILES values.

Figure 8 displays an integrated synthetic spectrum for the three methods described. Initially, there is no significant difference between them, even upon closer examination (see lower panel). However, the extrapolation method shows more deviation to the nearest method compared to the interpolation method. The choice of methodology — followed through the internship evolution — involves the creation of the IFU datacubes, which will be detailed in the following subsection.

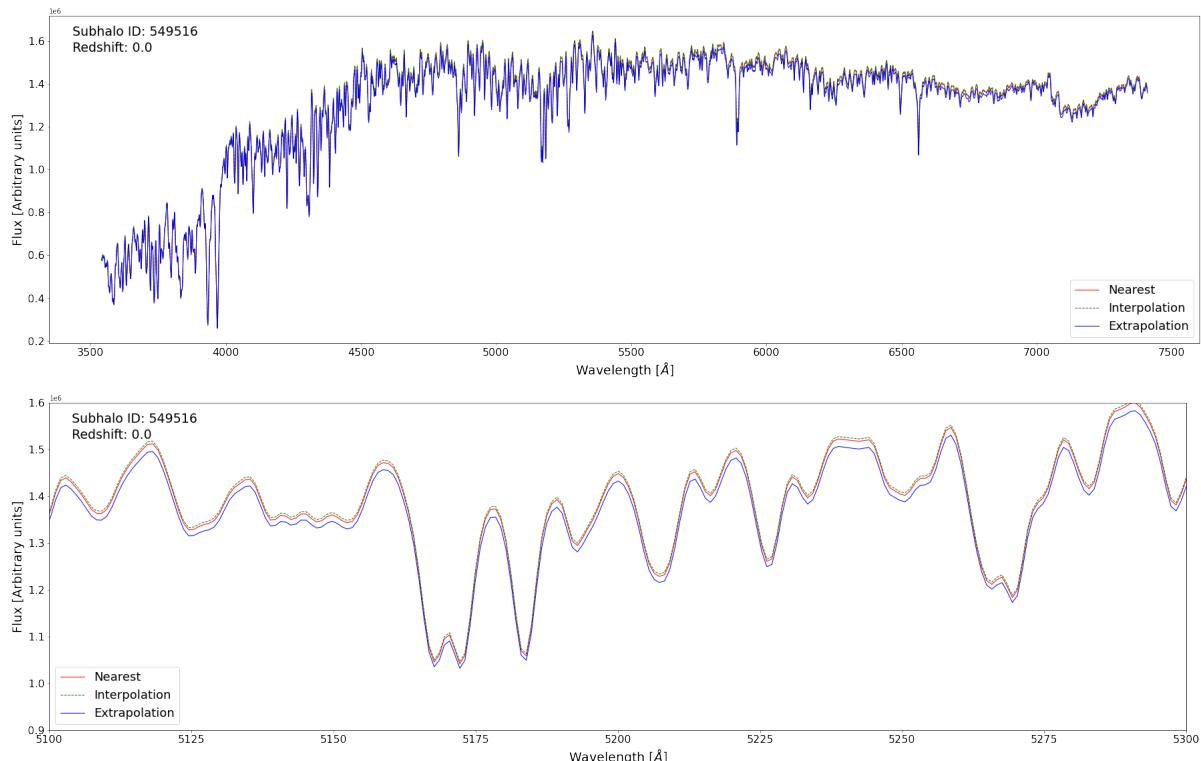


Figure 8: Integrated synthetic spectrum for TNG galaxy ID 549516 at $z = 0$ compared with the three techniques presented for the MILES wavelength range (upper panel) and in a thinner range (lower panel).

Fig. 9 demonstrates the relative difference associated with the interpolation (left) and extrapolation (right) methods, compared with the nearest method, which serves as the baseline approach. The interpolation method yields a relatively uniform difference across spaxels, approximately $\sim 4\%$. Conversely, the extrapolation method displays spatial dependence, reflecting the galaxy's morphological features. Specifically, Fig. 9 (right) reveals the spiral arm structure of the galaxy and its central region, highlighting that the extrapolation method's difference is more pronounced in regions of higher metallicity. Moreover, the accuracy of the extrapolation method is contingent upon the extrapolation technique employed (e.g., linear, cubic spline). Therefore, the interpolation method is deemed more precise for our objectives. Consequently, all subsequent analyses will utilize the interpolation method for the association between MILES and TNG age and metallicity values.

B. Including stellar kinematics:

The next major enhancement to the calculation of our synthetic spectra, aiming for increased realism, is to include the effect of the Doppler shift resulting from stellar kinematics. This shift is influenced by the tangential velocities and can be quantified by the following expression:

$$\frac{\Delta\lambda}{\lambda} = \frac{v}{c} \quad (2)$$

where λ represents the wavelength, v denotes the tangential velocity, and c is the speed of light. Fig. 10 demonstrates the impact on the spectrum for two spaxels at opposite location in the galactic plane, where the Doppler shift plays a major effect. It is evident that in the synthetic spectrum, which sums contributions from all stellar particles, the net result resembles a velocity dispersion effect, characterized by broader line widths than the spectrum without including stellar kinematics, due to the dispersion in velocities of the stellar particles.

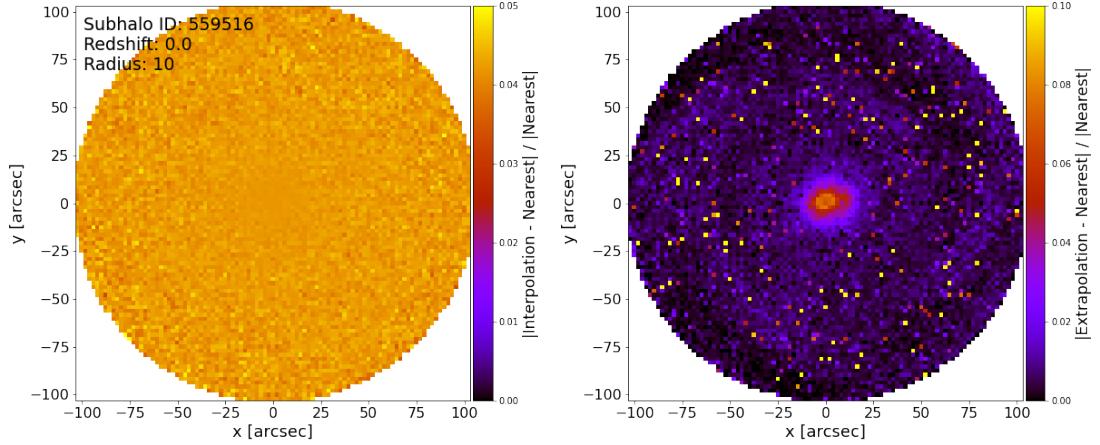


Figure 9: Comparison of the IFU datacubes for TNG galaxy ID 549516 at $z = 0$ and a radius of 10kpc, obtained using the three methods outlined in subsection 3.2 *B*. *Left*: relative difference between the interpolation and nearest methods. *Right*: relative difference between the extrapolation and nearest methods, revealing a clear correlation with higher difference in the spiral arms. This suggests that the extrapolation method may be less accurate in regions with higher metallicities.

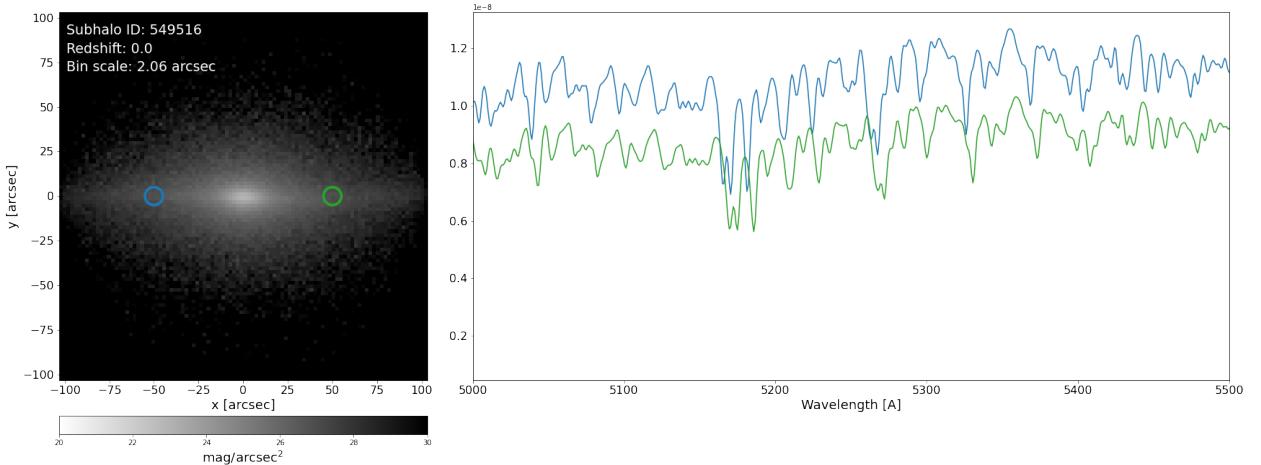


Figure 10: TNG galaxy ID 549516 at $z = 0$ and edge-on visualization. *Left*: photometric magnitude in SDSS g' band. *Right*: synthetic spectra of two spaxels along the galactic plane where the Doppler shift by tangential velocity is more significant (as can be see in Fig. 3 lower right panel).

3.3 Integral spectroscopic datacubes

Integral field spectrographs integrate imaging and spectroscopic capabilities in the infrared or optical wavelength ranges to obtain spatially resolved spectra across a two-dimensional region from a single exposure. The resulting measurements are derived by integrating the light from multiple sub-regions within the field, effectively condensing photometric and spectroscopic information of the same target into a three-dimensional datacube. Current and upcoming generations of telescopes are equipped with at least one IFS instrument; for example, the JWST features the NIRSpec IFU, while the future ELT will be equipped with HARMONI. As outlined in the objectives, we will generate mock IFU datacubes from simulations and subsequently incorporate the relevant resolution and technical parameters to simulate observations through these two instruments.

In the previous subsection, we briefly outlined the methodology employed to generate the IFU datacubes. A field of view (FOV) is defined based on the redshift, radius (in kpc), and spaxel size, or alternatively, the number of spaxels (which correspond to the angular distance assuming a Planck16 cosmology). The stellar particle information for a given galaxy is saved in an HDF5 file and subsequently projected onto the sky according to the specified line-of-sight vector. Each stellar particle is then assigned a synthetic spectrum from the MILES library, which are linearly combined to construct the fiber spectrum. Prior to this, however, it is necessary to convert the flux from arbitrary units to physical units.

A. Translate to photometric units:

The approach adopted to convert the flux into physical units was based on a calibration relative to a typical magnitude of a galaxy. The following steps outline the procedure for this calibration:

1. Correct by redshift: This initial step involves accounting for the redshift of the galaxy, particularly important for spectra at higher redshifts. The observed wavelength is shifted according to:

$$\lambda_{\text{obs}} = \lambda_0 \cdot (1 + z) \quad (3)$$

where λ_0 is the rest-frame wavelength. This correction shifts the spectrum to the observed wavelength values, which will be used in the following steps.

2. Calibration region: We select a central region of 1 kpc radius within the galaxy for calibration, defined by an integrated calibration magnitude m_{cal} . All stellar particles within this central region are identified, and their corresponding synthetic spectra are computed and integrated.
3. Calibration magnitude: The integrated spectrum is then translated into magnitude units using a specific filter. For galaxies at $z = 0$, the [SLOAN/SDSS g'](#) filter is employed, whereas for galaxies at $z = 1.5$, the [JWST NIRSpec F162M](#) filter is used. The filter response functions for these photometric bandpasses are obtained from the [SVO Filter Profile Service](#). To derive the integrated stellar luminosity within the selected bandpass, the stellar synthetic spectrum is convolved with the filter response function. We use the AB magnitude system, defined as follows [27]:

$$m_{\text{AB}} = -2.5 \log_{10} f_{\nu} - 48.6 \quad (4)$$

where the average f_{ν} in $\text{erg s}^{-1}\text{cm}^{-2}\text{Hz}^{-1}$ is given by

$$f_{\nu} = \frac{1}{c} \frac{\int ST\lambda d\lambda}{\int T/\lambda d\lambda} \quad (5)$$

where S is the spectrum with units $\text{erg s}^{-1}\text{cm}^{-2}\text{\AA}^{-1}$, T the filter curve, λ the wavelength in Angstrom, and c the speed of light in \AA s^{-1} . From this calculation we can obtain the corresponding magnitude in the AB magnitude system of our stellar particles within the 1kpc central m_{TNG} .

- Flux scale factor: Finally, we can multiply each of our spectra to a constant flux scale factor that is defined as the difference in fluxes:

$$\alpha = 10^{-|m_{\text{TNG}} - m_{\text{cal}}|/2.5} \quad (6)$$

This is the factor by which we will multiply each spectrum to convert it into physical units. In the pipeline developed it is obtained through the `flux_scale_factor` function.

By convolving the spectrum of each spaxel with the corresponding filter response function, we can generate a map of the integrated magnitude for each spaxel. To convert these magnitudes into surface brightness (surface magnitude), we first transform the magnitudes into flux units, divide by the spaxel area, and then revert to photometric magnitudes using Eq. (4). Figure 11 (12) illustrates the resulting IFU datacubes for face- (edge-) on visualization at different snapshots. This is performed with the `plot_IFU` function, after running `data_cube` function.

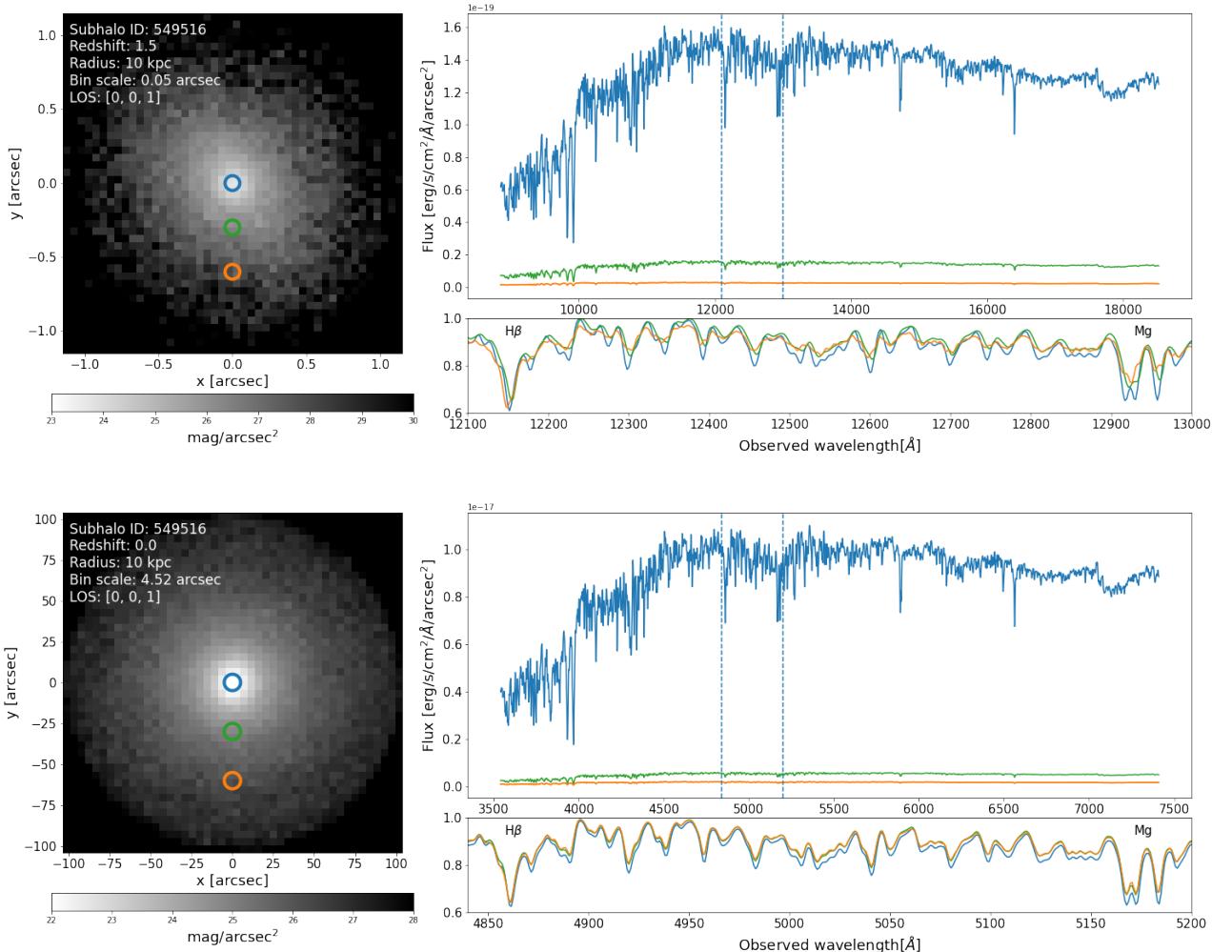


Figure 11: Face-on visualization of TNG50 galaxy ID 549516 at snapshot $z = 1.5$ (*upper panel*) and $z = 0$ (*lower panel*). The *left column* presents the surface brightness (in magnitudes) in the SDSS g' filter for $z = 0$ and in the NIRCam F162W filter for $z = 1.5$. The *upper right panel* shows spectra at various radii in flux units, highlighting the differences in spectral features across the galaxy. The *bottom right panel* provides a zoom-in of the spectral region between the dashed blue lines indicated in the upper plot, where the $\text{H}\beta$ line and the Mg triplet are distinctly visible.

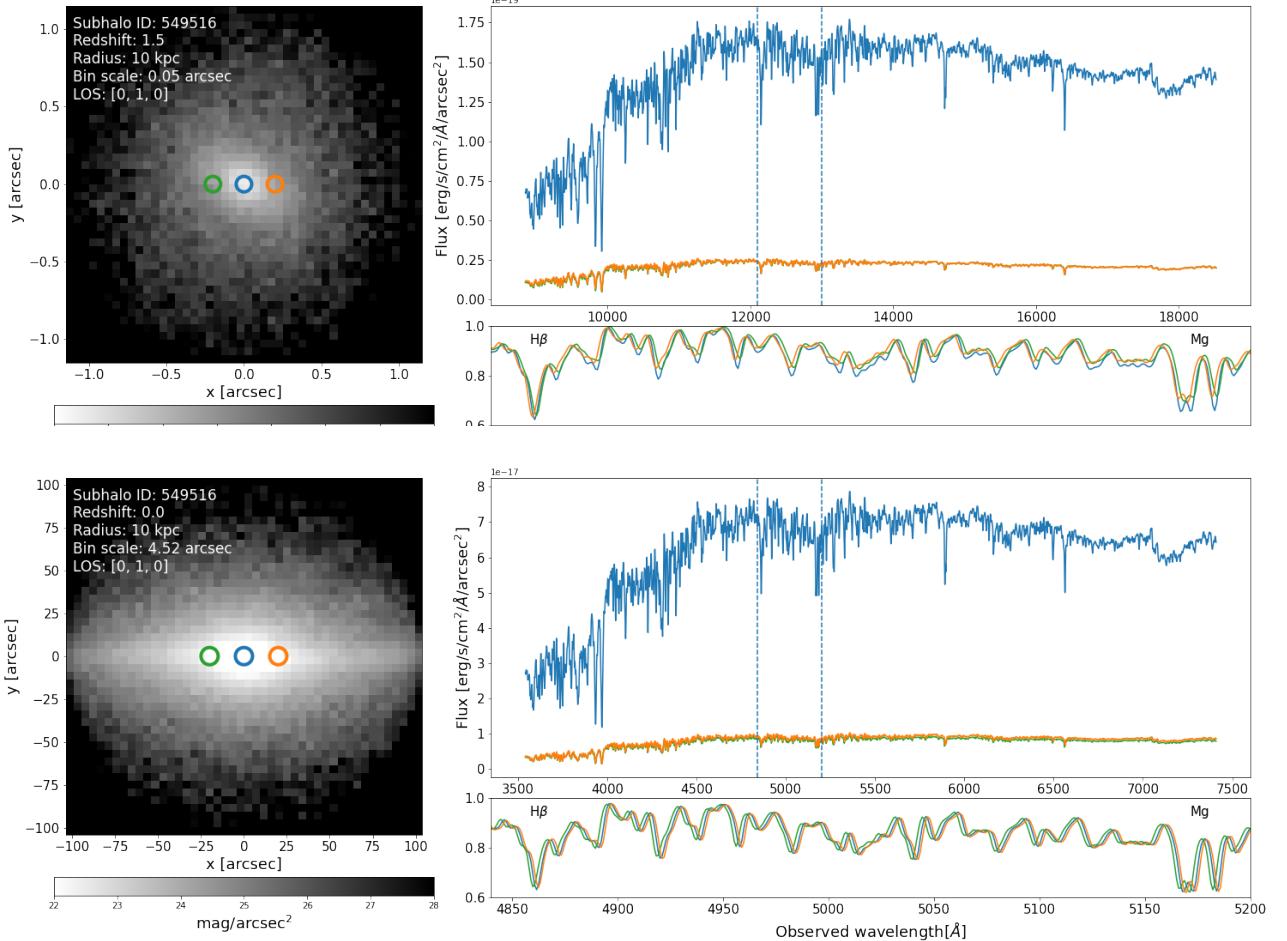


Figure 12: Edge-on visualization of TNG50 galaxy ID 549516 at snapshot $z = 1.5$ (*upper panel*) and $z = 0$ (*lower panel*). The *left column* presents the surface brightness (in magnitudes) in the SDSS g' filter for $z = 0$ and in the NIRCam F162W filter for $z = 1.5$. The *upper right panel* shows spectra at various radii in flux units, highlighting the differences in spectral features across the galaxy. The *bottom right panel* provides a zoom-in of the spectral region between the dashed blue lines indicated in the upper plot, where the $H\beta$ line and the Mg triplet are distinctly visible.

From Figure 11, we observe the anticipated exponential decrease in flux with increasing distance from the galaxy center, and a relative higher abundance of young, blue stars at $z = 1.5$. In the lower right panels, the $H\beta$ line exhibits nearly identical depths for the green and orange spaxels, serving as an indicator of the stellar population's age. Meanwhile, the middle absorption features, corresponding to Fe lines, are more pronounced in the central region, reflecting the higher contribution of Type Ia supernovae to the metallicity there. This pattern is similarly evident in the Mg line, further emphasizing the influence of the central region's enriched stellar environment.

B. Transport to $z = 0$:

During the internship we found interesting to perform a comparison between the galaxy at the snapshot corresponding to $z = 1.5$ and the galaxy at $z = 0$ displaced at $z = 1.5$ taking into account the redshift dimming and the angular distance. This was performed by shifting the wavelength according to the Eq. (3) and the spaxel size scaled to the difference of angular distance diameter $D_{A,1.5}/D_{A,0}$. In our pipeline this is done with the `displace_galaxy` function.

3.4 Simulate JWST observations

We now proceed to simulate the observation of our IFU datacubes using the NIRSpec instrument on the JWST. For this simulation, several key factors must be considered:

A. Spectral resolution:

NIRSpec offers various grisms, each with different spectral resolutions [28]. Although higher resolution grisms are available, we will utilize the PRISM grism, which provides an adequate resolution for observations at $z = 1.5$. The spectral degradation to match NIRSpec’s resolution is implemented through a Gaussian convolution applied to the spectrum of each spaxel, using a sigma value derived from the FWHM of the instrument. The main contributors to the resolution degradation include the galaxy’s intrinsic velocity dispersion, the spectral resolution of NIRSpec, and the resolution of the MILES library spectra used in the simulation. These factors are combined to accurately replicate the observational conditions, ensuring that our simulated datacubes reflect the realistic performance of NIRSpec under the chosen observational settings. The corresponding FWHM is:

$$\text{FWHM}_i^2 = \text{FWHM}_{\text{JWST}}^2 - \text{FWHM}_{\text{MILES}}^2 \quad (7)$$

where $\text{FWHM}_{\text{MILES}}^2 = 2.51\text{\AA}$ [29] and the pixel size of MILES is $0.9\text{\AA}/\text{pix}$. The spectral resolution R as a function of the wavelength for NIRSpec is obtained from [NIRSpec Dispersers and Filters webpage](#) (see Figure 4). To determine the FWHM specific to JWST, denoted as $\text{FWHM}_{\text{JWST}}^2$, we must also consider its dependence on redshift, which is characterized by the relation:

$$\text{FWHM}_{\text{JWST}} = \frac{\lambda}{R \cdot (1 + z)} \quad (8)$$

It is important to note that the FWHM varies with each wavelength step due to the dispersive nature of the PRISM grism. To calculate the corresponding sigma for the Gaussian convolution at each wavelength, we use the relation:

$$\sigma = \frac{\text{FWHM}_i}{2\sqrt{2\log(2)}} \quad (9)$$

Finally, to complete the spectral resolution adjustment, we divide the calculated sigma value by the MILES pixel size of $0.9\text{\AA}/\text{pixel}$ [29]. The convolution is then performed using the `scipy.gaussian.filter1d` function, applied to the spectrum of each spaxel in our grid.

Additionally, in order to reproduce accurately the reality we must include a random Gaussian noise of mean $\mu = 0$ and sigma $\sigma = 1$ scaled to the unity. Then, we normalize the noise to each spaxel median spectrum and divide it into 5 in order to get a signal to noise ratio (SNR) of 5, as is derived from previous studies for a 6 hour exposure for an element resolution of 24 mag/arcsec² for Milky Way-like galaxies at $z = 1.5$ with the JWST. Then we add this random Gaussian noise to our observed spectrum after the spectral convolution.

B. Number of spaxels:

The JWST NIRSpec IFU has a $3'' \times 3''$ field of view (FOV) that is divided into 30 slices, resulting in 900 spatial elements within the FOV (see [NIRSpec IFU webpage](#), Figure 2). Therefore, the next step is to configure our IFU datacube to match the 30×30 spaxel layout of the NIRSpec spectrograph. While it might be tempting to set the number of bins directly to 30 in the `data_cube` function to reduce computational load, constructing the datacubes with smaller spaxel sizes (and thus a higher number of spaxels) initially enhances spatial resolution. This allows for finer spatial detail, which can be preserved even if the datacube is later cropped to 30×30 .

C. Spatial resolution:

The final piece is to introduce the spatial resolution of the JWST, which is also performed by a Gaussian convolution in 2D. In this case, the corresponding spatial FWHM depends linearly with the wavelength [NIRCam PSF webpage](#) (see Figure 5), so we can perform a linear regression to those values to obtain the spatial FWHM. And then, similar to the spectral resolution the sigma is obtained and the convolution performed with `scipy.gaussian_filter`.

After the introduction of this elements we are able to obtain with the `pass_to_jwst` a realistic simulation of a IFU datacube of a MW-like galaxy at $z = 1.5$ through the NIRSPec of the JWST. Fig. 13 illustrates the procedure described above.

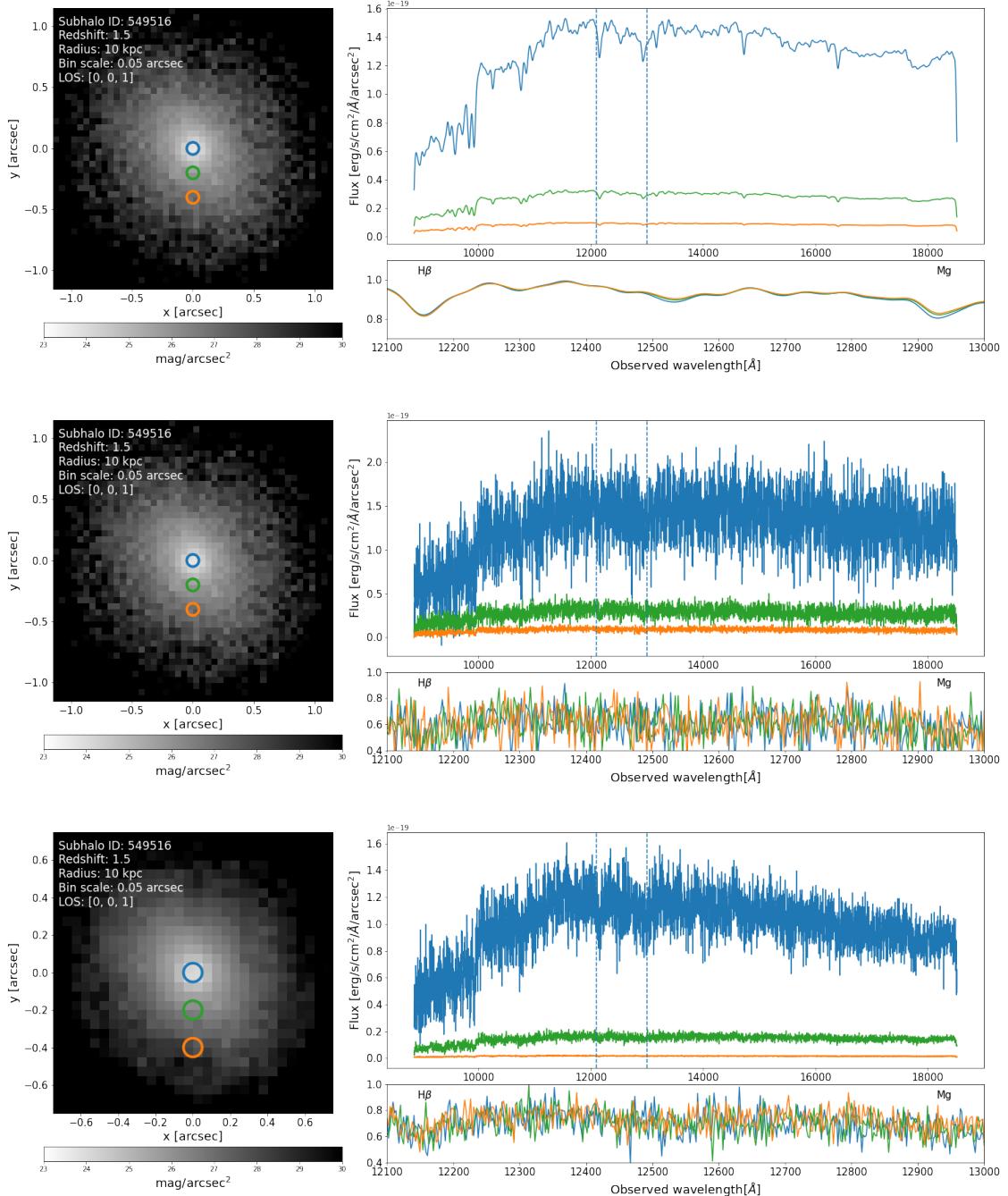


Figure 13: Face-on visualization of TNG50 galaxy ID 549516 at snapshot $z = 1.5$. Upper panel illustrates the spectral convolution to the grism PRISM. Middle panel shows the addition of the Gaussian noise to get a SNR= 5. Lower panel shows the fit to the 30×30 IFU datacube spaxel and the spatial convolution to the NIRCam PSF.

3.5 Simulate HARMONI observations

The High Angular Resolution Monolithic Optical and Near-Infrared Integral field spectrograph (HARMONI) is the visible and near-infrared adaptive-optics-assisted, integral field spectrograph for ESO’s ELT. It has four choices of spaxel scale and 11 grating choices with resolution powers from $R \sim 3000$ to $R \sim 17000$. The conjunction of ELT+HARMONI will transform the landscape of observational astronomy by providing a big leap in sensitivity and resolution, with a spatial-resolution of ~ 10 milliarcseconds (mas) over most of the sky. Fig. 14 shows the various spaxel scales and gratings available for use with the HARMONI instrument. For our study, we will use the 20×20 mas spaxel scale, as it effectively covers the 1 kpc central region of a Milky Way-like galaxy at $z = 1.5$, and the Iz grating, which encompasses the spectral range extrapolated from the MILES library for these redshifts. For a more realistic future work, the procedure will involve to extend our spectral coverage to the J and H filter.

To simulate HARMONI observations, we will employ the publicly available HARMONI simulator **HSIM3**. HSIM3 simulations incorporate detailed models of the sky, telescope, instrument, and detectors, allowing for the generation of realistic mock datacubes [30]. Fig. 15 presents an example of the simulation output of an IFU datacubes in units of electrons.

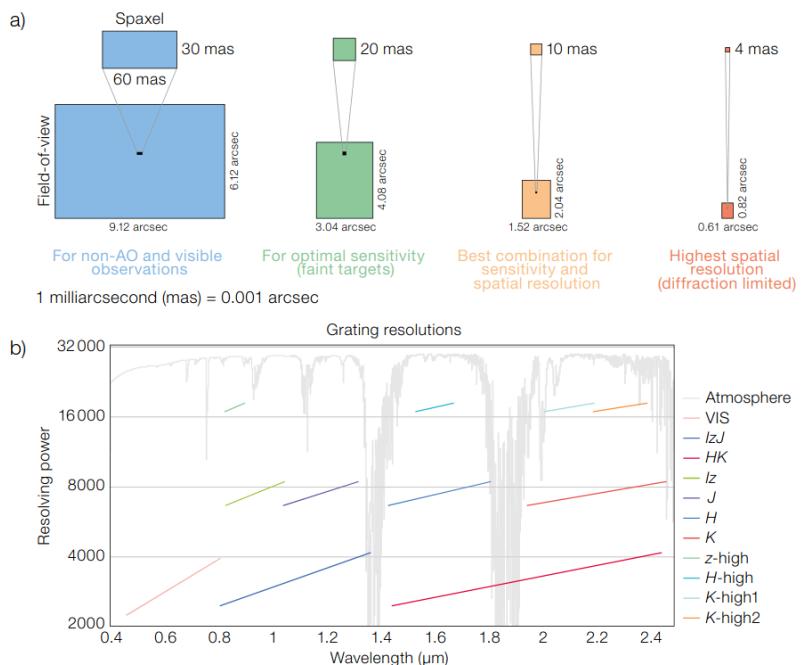


Figure 14: a. Spatial layout of the HARMONI science field, showing the spaxel sizes and FOV at the four different spaxel scales. b. Spectral coverage and resolving power ranges for each of the 11 HARMONI grating choices.

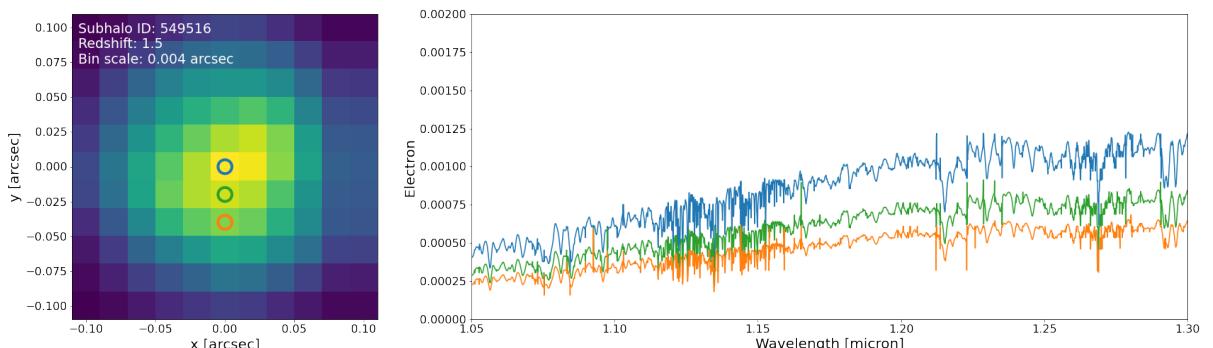


Figure 15: HSIM3 simulation output for TNG galaxy ID 549516 at $z = 1.5$ with 20×20 spaxel scale and Iz grating.

4 Results and discussion

At this stage, we have achieved our objective of developing a Python pipeline capable of generating JWST-like IFU datacubes for Milky Way-like galaxies at any desired redshift. For the purposes of this study, we are particularly interested in the redshift value of $z = 1.5$, as it corresponds to the epoch of the Gaia-Sausage Enceladus merger. In this section, we present the results for various galaxies for a random l.o.s., focusing exclusively on the JWST-like IFU datacubes and adopting the $3'' \times 3''$ FOV of the NIRSpec IFU Spectroscopy.

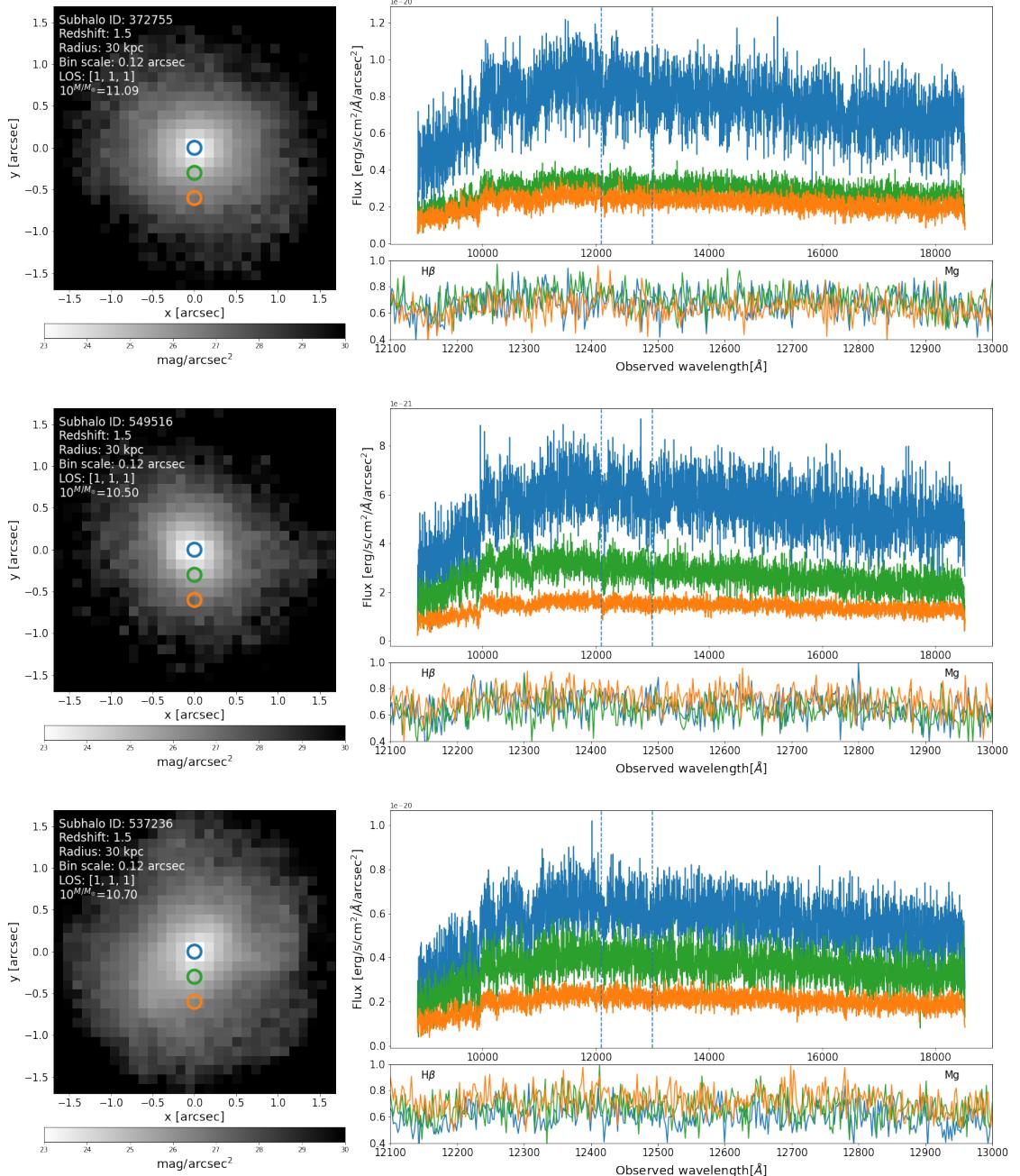


Figure 16: Three TNG50 simulated galaxies illustrate the simulated JWST observations. *Left:* surface brightness (in magnitudes per arcsecond) in the NIRSpec F162W filter at $z = 1.5$, convolved with the NIRCam PSF and adjusted to the $3'' \times 3''$ FOV of NIRSpec. *Upper Right:* spectra at various radii, convolved with the NIRSpec PRISM grism and superimposed with Gaussian noise to achieve a SNR=5. *Lower Right:* the same spectra, but zoomed into the region highlighting the H β and Mg absorption lines.

5 Conclusions

In this work we have investigated the longstanding challenge of observing high-redshift galaxies with the current instrumental capabilities by simulating JWST and HARMONI observations of mock IFU datacubes from TNG50 Milky Way-like galaxies. Throughout this analysis, we have utilized state-of-the-art hydrodynamical simulations and stellar population models, gaining substantial insights into the association between simulated stellar particles with the stellar population models. Then, we have developed a consistent Python code to create IFU mock datacubes from simulated Milky Way-like galaxies. The input required is the TNG50 galaxy ID, snapshot, radius, line of sight and the spaxel scale. The output are a IFU datacube in physical units of $\text{erg}/\text{s}/\text{cm}^2/\text{\AA}/\text{arcsec}^2$ and a photometric superficial magnitude per arcsec, both in FITS file. Then, as detailed we can simulate their observation through the JWST by including the spectral resolution of the NIRSpec F162W filter, spatial resolution of the NIRCam PSF, and by adding a random Gaussian noise to obtain a SNR=5.

From the IFU datacubes developed we can infer several expected features of the evolution of Milky Way-like galaxies. Firstly, at $z = 1.5$ there is no clear evidence of a disk structure among the analyzed sample, suggesting that disk formation may occur later in the galaxy's evolution. Secondly, the spectra reveal a notable presence of young, blue stellar populations within the galaxy at $z = 1.5$, as indicated by the increased intensity of blue components in the spectra. However, a more in-depth analysis and interpretation of these results were beyond the scope and temporal constraints of the summer internship.

Regarding future perspective of this work, the primary objective of this work was to simulate JWST observations to derive technical parameters for future proposals. Future simulations could incorporate calculations of the exposure time required to achieve a sufficient SNR for observations. Additionally, significant effort will be required to reverse-engineer the process: starting from real JWST datacubes, and using inverse methods to reconstruct the stellar population properties of Milky Way-like galaxies.

Acknowledgments:

This work has been supported with the funding from the competitive grant “Becas de verano para iniciación en investigación Astrofísica” from Instituto de Astrofísica de Canarias (IAC), with reference PS-2024-018.

I personally thank to my supervisors Nacho, Adri and Bego for welcoming me with open arms during this summer experience and always being ready to explain and assist me in navigating this unfamiliar yet fascinating field of astrophysics. To my colleagues during this summer internship, Ángel, Pere, Óscar, Salva and Pablo. To our fruitful and eternal discussions and thanks to become a great company during this two months.

References

- [1] S. Kurtz *et al.* *Protostars & Planets IV*, Ed. V, 299-326, (2000).
- [2] C. Lacey & S. Cole. *MNRAS*, **262**, 627-649, (1993).
- [3] C. S. Frenk & S. D. M. White. *Ann. Phys.*, **524**, 9-10, 507-534, (2012).
- [4] J. Zavala & C. S. Frenk. *Galaxies*, **2019**, 7(4), 81, (2019).
- [5] G. Worthey, S. M. Faber & J. J. González. *Astrophys. J.*, **398**, 69, (1992).
- [6] H. Mo *et al.* *Galaxy Formation and Evolution*, Cambridge University Press, (2010).
- [7] T. Prusti *et al.* (Gaia Collaboration). *Astron. & Astrophys.*, **595**, 36, (2016).
- [8] A. Bonaca *et al.* *Astrophys. J. Lett.*, **897**, L18, (2020).
- [9] A. B. A. Queiroz *et al.* *Astron. & Astrophys.*, **656**, A156, (2021).
- [10] E. Vasiliev. *Proceedings IAU*, **2019**, 536-539, (2019).
- [11] H. Rix *et al.* *Astrophys. J.*, **941**, 1, 15, (2022).
- [12] V. Belokurov & A. Kravtsov. *MNRAS*, **514**, 1, 689-714, (2022).
- [13] J. P. Gardner *et al.* *Space Science Reviews*, **123**, 4, 485-606, (2006).
- [14] T. Böker *et al.* *PASP*, **135**, 038001, (2023).
- [15] A. Vazdekis *et al.* *MNRAS*, **404**, 4, 1639-1671, (2010).
- [16] N. A. Thatte *et al.* *The Messenger*, **182**, 7-12, (2021).
- [17] D. Nelson *et al.* *MNRAS*, **490**, 3, 3234-3261, (2019).
- [18] A. Pillepich *et al.* *MNRAS*, **490**, 3, 3196-3233, (2019).
- [19] P. A. R. Ade (Planck Collaboration). *Astron. & Astrophys.*, **594**, 63, (2016).
- [20] V. Springel. *MNRAS*, **401**, 2, 791-851, (2010).
- [21] R. Weinberger *et al.* *MNRAS*, **465**, 3, 3291-3308, (2017).
- [22] A. Pillepich *et al.* *arxiv:2303.16217*, (2023).
- [23] P. Sánchez-Blázquez *et al.* *MNRAS*, **371**, 2, 703-718, (2006).
- [24] A. Pietrinferni *et al.* *Astrophys. J.*, **612**, 1, 168-190, (2004).
- [25] P. Kroupa. *MNRAS*, **322**, 2, 231-246, (2001).
- [26] E. E. Salpeter. *Astrophys. J.*, **121**, 161, (1955).
- [27] J. B. Oke & J. E. Gunn. *Astrophys. J.*, **266**, 713-717, (1983).
- [28] P. Jakobsen *et al.* *Astron. & Astrophys.*, **661**, A80, (2022).
- [29] J. Falcón-Barroso *et al.* *Astron. & Astrophys.*, **532**, A95, (2011).
- [30] S. Zieleniewski *et al.* *MNRAS*, **453**, 4, 3754-3765, (2015).