Alma Mater Studiorum · Università di Bologna

**Scuola di Scienze**
**Dipartimento di Fisica e Astronomia**
**Corso di Laurea in Fisica**

# MOCK HI OBSERVATIONS OF MILKY WAY-LIKE SIMULATED GALAXIES

**Relatore:**
**Prof. Federico Marinacci**

**Presentata da:**
**Nicolò Landi**

Anno Accademico 2021/2022

**Abstract**

This work briefly illustrates the process behind the creation of mock HI observations of simulated Milky Way-like galaxies and how to properly analyse the generated output by comparing it with real observations. Firstly, we go over the basics of the quantum mechanic theory behind HI line emission, explain its importance and uses in astronomy, and the various methods we can use to obtain information about the properties of the sources emitting such a radiation. Then we focus on the process of mock observing a simulated Milky Way-type galaxy generated with the SMUGGLE model, using the MARTINI software to generate its 21 cm line emission. Finally, we analyse the synthetic data we obtained, comparing them to real data of galaxies with similar properties, to qualitatively assess the physical fidelity of the numerical model employed to evolve the simulated galaxy.

**Sommario**

L'obiettivo di questo lavoro di tesi consiste nel descrivere sia il processo necessario per la creazione di osservazioni sintetiche di galassie simulate simili alla Via Lattea nella riga di emissione a 21 cm dell'idrogeno neutro (HI), sia il lavoro di analisi fondamentale che serve a confrontare in modo efficace l'output generato con delle osservazioni di galassie reali. Come prima cosa è descritta la teoria quantistica che sta alla base dell'emissione a 21 cm di HI, illustrando l'importanza di tale riga di emissione nell'ambito dell'astronomia e come si possano ottenere informazioni fondamentali sulle sorgenti di questa radiazione a partire dai dati osservativi. Il lavoro poi si focalizza sull'utilizzo del software MARTINI per la creazione di osservazioni sintetiche della linea a 21 cm per una galassia simulata con proprietà simili alla Via Lattea generata utilizzando il modello numerico SMUGGLE. Infine, si passa ad una breve descrizione dell'analisi dei dati sintetici creati, e al loro confronto con dei dati provenienti da osservazioni reali di galassie con proprietà simili, per ottenere una valutazione qualitativa della bontà del modello SMUGGLE impiegato nella simulazione numerica.

# Contents

# Introduction

The Universe around us, both locally and billions of light years away, is nowadays studied trough the use of a myriad of instruments and techniques, from optical telescopes to cutting-edge laser interferometers. Nevertheless, one of the best ways to derive a large variety of properties of many astronomical objects is through the investigation of the electromagnetic spectrum. A great deal of information can be inferred from continuous spectra, such as thermal radiation from stars and dust or more specific phenomena like bremsstrahlung. An equally crucial role is played by line spectra. Through quantum mechanics we are able to measure the energy gaps between the discrete levels in atoms and molecules with extreme precision, and thus we also know exactly the energy of photons emitted or absorbed in transitions between said levels. Observing these lines then, gives us the possibility of establishing the chemical composition of the objects emitting such radiation.

Of course these photons do not arrive to the Earth unaltered, but that in itself is something we can exploit to derive more properties about their sources. An important way to do so is measuring how much an emission line is shifted in frequency. This information can give us the line of sight velocity of the source by inverting the Doppler shift relation. Of all the lines that can be observed in the sky, a prominent one is that of the 21 cm line, that corresponds to the transition between states with parallel and anti-parallel spins in the ground state of hydrogen. Even though this spectral line has very small energy and transition rate, it has a central role in astronomical applications. This is because neutral hydrogen is present in large quantities almost everywhere in galactic and extra-galactic environments, even further away than star rich regions, but also because, with its long wavelength, this emission manages to avoids the extinction

that higher energy photons face from dust and other obstacles.

The purpose of this work, though, is not to directly analyze data from HI emission. Instead, we will go over the process of creating and analysing 'mock' observations of neutral hydrogen emission from a simulated Milky Way-type galaxy. This process is of crucial importance because it offers the most realistic possible way to compare real observed galaxies with results form numerical simulations. In fact, one cannot compare directly the output of a simulation with data from real observations because the former gives an output that does not present all of the complexities that have to be taken into account when using actual telescopes, such as interference from other objects or simple measuring errors. With this operational method we can both test the validity of numerical simulations, and gain useful insights about the interpretation of real observations.

To this end, we will start by going over the theory behind the HI emission process, its uses in astronomy, some observational methods and how to analyse the resulting data. Then, we will describe the process of creating mock observations of a Milky Way-type galaxy simulated with the SMUGGLE model (Marinacci et al., 2019), through the use of the MARTINI software (Oman et al., 2019). Finally, we will analyse the mock observations output using different techniques and compare them with data from real observations.

# Chapter 1

# The HI 21 cm line emission

The interstellar medium, or ISM, consists of the matter and radiation that fill the space between star systems in a galaxy. Gas constitutes 99% of the mass of the ISM, and of that mass about 70% is hydrogen (Karttunen et al., 2017). Furthermore the physical state of this medium varies markedly from one region to another because the gas temperature has a dependence on the local energy input and cooling processes. There exist large, cool cloud complexes in which both dust grains and many different molecular species are abundant, and warmer regions where only atomic lines are found (Wilson et al., 2013). Out of those, our interest lies within the part of the ISM that is constituted primarily by neutral atomic hydrogen. One of the most important properties of neutral hydrogen is its capacity to emit in the 21 cm line of the electromagnetic spectrum, the so-called HI line. In this chapter we will first examine the quantum theory behind the atomic transition that generates this particular spectral line, then we will discuss how to detect HI sources and how to extract information and properties about whole galaxies from the observed data.

## 1.1 Hyperfine splitting in the ground state of hydrogen

The electron and the proton in atomic hydrogen constitute tiny magnetic dipoles. Their interaction energy will vary according to the relative orientation of their dipole

moments, thus there will be an energy difference between the two states where the proton and the electron have spins that are parallel (the triplet state) or antiparallel (the singlet state), creating what is known as the hyperfine splitting (Griffiths, 1982). To derive the entity (i.e. the energy gap) of this splitting we can use first-order perturbation theory, where the classical interaction energy is interpreted as the hyperfine structure Hamiltonian.

The energy of a magnetic dipole $\mathbf{m}$, in the presence of a magnetic field $\mathbf{B}$, is given by the following equation:

$$H = -\mathbf{m} \cdot \mathbf{B}\,, \tag{1.1}$$

so to compute the effect of the interaction between the magnetic dipoles of the proton and the electron we need to know the dipole field $\mathbf{B}$ generated by one of those dipoles.

### 1.1.1 Magnetic field generated by a dipole

The vector potential of an ideal magnetic dipole, in an arbitrary spherical coordinate system $\mathbf{r}$, is given by

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{\mathbf{m} \times \hat{r}}{r^2}, \tag{1.2}$$

where $\mathbf{m}$ is the dipole moment and $\mu_0$ is the vacuum permeability. The dipole field can be obtained by the taking the curl of this vector potential:

$$\mathbf{B}(\mathbf{r}) = \left(\frac{\mu_0}{4\pi}\right) \left(\frac{1}{r^3}\right) \left[3\,(\mathbf{m} \cdot \hat{r})\,\hat{r} - \mathbf{m}\right]. \tag{1.3}$$

However, if we were to compute the average magnetic field over a spherical volume of radius R, where we are placing $\mathbf{m}$ at the origin, simply like

$$\mathbf{B}_{av} = \frac{1}{V} \int \mathbf{B}\, dV, \tag{1.4}$$

using (1.3) as $\mathbf{B}$ and $V = \frac{4}{3}\pi R^3$, we notice that the integral diverges. This is not a physically relevant result and the problem lies in the coordinate system choice, since the magnetic field of the dipole is singular at $\mathbf{r} = 0$. To solve this issue we treat the ideal dipole as the point limit of a uniformly magnetized sphere of radius R, magnetization $\mathbf{M}$, and dipole moment $\mathbf{m} = \frac{4}{3}\pi R^3\,\mathbf{M}$. The field outside the sphere (for $r > R$) is given

by (1.3), while the field inside the sphere is uniform:

$$\mathbf{B}_{in}(\mathbf{r}) = \frac{\mu_0}{2\pi}\frac{\mathbf{m}}{R^3} \quad \text{for } r < R. \tag{1.5}$$

In the ideal dipole limit ($R \to 0$) the interior region shrinks to zero, but the field goes to infinity, with their product remaining constant. Therefore, as ($R \to 0$) the field inside the sphere goes to a delta function:

$$\mathbf{B}_{in}(\mathbf{r}) = \frac{2}{3}\,\mu_0\,\mathbf{m}\,\delta^3(\mathbf{r}). \tag{1.6}$$

The magnetic field of an ideal dipole can thus be written as

$$\mathbf{B}(\mathbf{r}) = \left(\frac{\mu_0}{4\pi}\right)\left(\frac{1}{r^3}\right)\left[3\,(\mathbf{m}\cdot\hat{r})\,\hat{r} - \mathbf{m}\right] + \frac{2}{3}\,\mu_0\,\mathbf{m}\,\delta^3(\mathbf{r}), \tag{1.7}$$

with the understanding that the first term applies only to the region outside an infinitesimal sphere at the origin.

### 1.1.2 Quantum energy splitting

By knowing that the magnetic field generated by a magnetic dipole is given by eq. (1.7), and using eq. (1.1), the energy of a magnetic dipole ($\mathbf{m}_1$) in the field of another magnetic dipole ($\mathbf{m}_2$) is

$$H = -\left(\frac{\mu_0}{4\pi}\right)\left(\frac{1}{r^3}\right)\left[3\,(\mathbf{m}_1\cdot\hat{r})\,(\mathbf{m}_2\cdot\hat{r}) - \mathbf{m}_1\cdot\mathbf{m}_2\right] + \frac{2}{3}\,\mu_0\,\mathbf{m}_1\cdot\mathbf{m}_2\,\delta^3(\mathbf{r}), \tag{1.8}$$

where $\mathbf{r}$ is their separation. This represents the interaction energy between the magnetic dipoles generated by the proton ($\mathbf{m}_1$) and the electron ($\mathbf{m}_2$). The formula is symmetric in its treatment of $\mathbf{m}_1$ and $\mathbf{m}_2$, as it should be. In most cases the two dipoles are phyisically separated, and the delta-function term can be ignored; however, it is precisely this part which accounts for hyperfine splitting in the ground state of hydrogen.

First order perturbation theory can now be used to compute the shift in the energy of the ground state, treating the dipole-dipole interaction (1.8) as a perturbation:

$$E' = \int \Psi_0^*\,H\,\Psi_0\,d\tau, \tag{1.9}$$

where $\Psi_0$ is the ground state wave function for atomic hydrogen. The evaluation of the previous integral gives:

$$E' = \frac{2}{3}\,\mu_0\,\langle \mathbf{m}_1 \cdot \mathbf{m}_2 \rangle\,|\Psi_0(0)|^2 = -\frac{2}{3}\frac{\mu_0}{\pi a^3}\langle \mathbf{m}_1 \cdot \mathbf{m}_2 \rangle. \tag{1.10}$$

The magnetic dipole moments are proportional to the spins of their respective particle:

$$\mathbf{m}_1 = \gamma_p\,\mathbf{S}_p, \qquad \mathbf{m}_2 = -\gamma_e 1\,\mathbf{S}_e, \tag{1.11}$$

where the $\gamma$ are the two gyromagnetic ratios form which it follows that

$$E' = \frac{2}{3}\frac{\mu_0}{\pi a^3}\,\gamma_e\,\gamma_p\,\langle \mathbf{S}_e \cdot \mathbf{S}_p \rangle. \tag{1.12}$$

From quantum theory it is known that in a two particles system with total angular momentum $\mathbf{J} = \mathbf{S}_e + \mathbf{S}_p$:

$$\langle \mathbf{S}_e \cdot \mathbf{S}_p \rangle = \frac{1}{2}\left(J^2 - S_e^2 - S_p^2\right). \tag{1.13}$$

The electron and proton carry spin $\frac{1}{2}$, so the eigenvalues of $S_e^2$ and $S_p^2$ are $\frac{3}{4}\hbar^2$. The two spins combine to form a triplet state with $J^2 = 2\hbar^2$ and a singlet state with $J^2 = 0$. Thus there are only two possible values for $\langle \mathbf{S}_e \cdot \mathbf{S}_p \rangle$:

$$E' = \frac{2}{3}\frac{\mu_0\,\hbar^2}{\pi a^3}\,\gamma_e\,\gamma_p \begin{cases} \frac{1}{4} & \text{(triplet)} \\ -\frac{3}{4} & \text{(singlet)}. \end{cases} \tag{1.14}$$

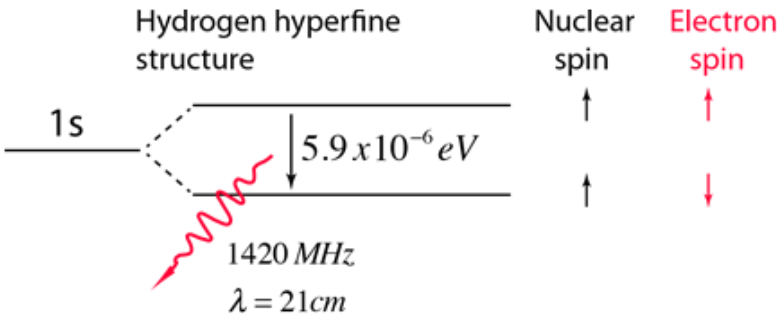


**Figure 1.1:** Hyperfine splitting in the ground state of hydrogen (Rod Nave, 2017).

Evidently the singlet state, in which the spins are antiparallel, carries a somewhat lower energy than the triplet combination (figure 1.1). Knowing that the gyromagnetic ratios are $\gamma = (e/2m)\, g$, where $e$ is the proton charge, $m$ is the mass of the particle, and g its "g factor" (2.0023 for the electron, 5.5857 for the proton) the energy gap between the two states is

$$\Delta E_{hyd} = \frac{\mu_0\, \hbar^2\, e^2}{6\pi a^3} \frac{g_e\, g_p}{m_e\, m_p} = 5.884 \times 10^{-6}\text{eV}. \tag{1.15}$$

The frequency of the photon emitted in a transition from the triplet to the singlet state is then

$$\nu = \frac{\Delta E}{h} = 1422.8 \text{ MHz}, \tag{1.16}$$

and its wavelength is

$$\lambda = \frac{c}{\nu} = 21.07 \text{ cm}. \tag{1.17}$$

## 1.2 Mapping HI emission

Neutral atomic interstellar hydrogen can be found almost everywhere in galactic systems, and its presence is vital to the processes of galactic evolution, dynamics and star formation (see e.g. Marinacci et al., 2019). One of the best way to observe the properties of this gas is through line emission and absorption. Since the exact wavelength of atomic transition is known, a great deal of information about mass, location and the dynamics of the gas can be gathered by observing lines' intensity, broadening and Doppler shift, as will be better discussed later in this section. Out of all atomic hydrogen transitions, observations of the 21 cm line have been especially useful. In principle HI emission should be fairly faint. In fact, the probability of transitioning from the triplet ($s = 1$) to the singlet ($s = 0$) state of the hyperfine splitting can be computed using Einstein coefficients, and is (Wilson et al., 2013):

$$A_{1,0} = \frac{64\pi^4}{3\, h\, c^3}\, \nu_{1,0}^3\, |\mu_{1,0}|^2 = 2.86888 \times 10^{-15}\ \text{s}^{-1}, \tag{1.18}$$

where $|\mu_{1,0}| = (e\hbar)/(2m_e c)$ is the mean magnetic dipole moment for this transition, and $\nu_{1,0}$ is the frequency of the emitted photon from eq. (1.16). This transition probability is

about a factor $10^{23}$ smaller than that of an allowed optical transition, mostly due to the frequency difference, which enters as $\nu^3$, while an additional factor of $5 \times 10^5$ comes from the small magnitude of the magnetic dipole moment. The spontaneous mean half-life time of the triplet state is then

$$t_{1/2} \cong 1/A_{1,0} = 3.49 \times 10^{14} \text{ s} \cong 1.11 \times 10^7 \text{ yr}. \tag{1.19}$$

Despite the low transition probability, the HI emission line is still easily observed due to the sheer amount of neutral hydrogen present in galactic systems; for instance, in Milky Way-type galaxies there is an estimated total HI mass of $M_{\text{HI}} = 6 \times 10^9 \, \text{M}_\odot$ (Rhee et al., 2018). Furthermore, the HI line lies in the radio spectrum, as opposed to other important transitions like for example the Lyman $\alpha$ line, which corresponds to the transition of the electron between states with principal quantum number n = 1 and n = 2, with associated photon energy in the UV spectrum. Since extinction from dust clouds and our atmosphere is very efficient in the visible/UV spectrum, but is negligible for radio waves, HI emission does not suffer from this problem, and radio telescopes on the ground can detect it basically unperturbed. Now that the basic properties of HI emission have been described, we can discuss how we can map the quantity of neutral atomic hydrogen throughout a galaxy from the observation of the 21 cm line.

### 1.2.1 Brightness temperature

The brightness temperature $T_B$ of an astronomical source is defined as the temperature of a black body emitting the same spectral radiance $B_\nu$ as the source, based on the Rayleigh-Jeans approximation to Planck's law (Westmeier, 2021), which holds at long wavelengths such as 21 cm emission; hence

$$T_B = \frac{c^2 \, B_\nu}{2 \, k_B \, \nu^2} = \frac{\lambda^2 \, B_\nu}{2 \, k_B} \, , \tag{1.20}$$

where $k_B$ is the Boltzmann constant. Furthermore, the flux density $S_\nu$ of a source is simply defined as the integral of the spectral radiance of the solid angle $\Omega$ of the source in the sky, leading to the following relation between brightness temperature and flux density:

$$S_\nu = \frac{2\,k}{\lambda^2} \iint T_B \, d\Omega, \tag{1.21}$$

which can be simplified if we consider a source of constant brightness temperature:

$$T_B = \frac{\lambda^2 \, S_\nu}{2\,k_B\,\Omega_{bm}}\,, \tag{1.22}$$

where $\Omega_{bm} = \pi\,\theta_a\,\theta_b/\,[4\ln(2)]$ is the solid angle of the telescope, assuming a Gaussian beam, where $\theta_a$ and $\theta_b$ are the beam major and minor axis. If we further assume a circular beam with FWHM (full width at half maximum) $\theta = \theta_a = \theta_b$, we can simplify the equation to

$$T_B = \frac{2\ln(2)\,\lambda^2\,S_\nu}{\pi\,k_B\,\theta^2}\,. \tag{1.23}$$

Since flux density is the quantity measured by radio telescopes, the brightness temperature can be derived for any observed HI source in the sky. We will now derive a relation between brightness temperature and HI column density $\mathcal{N}_H$, defined as the amount of hydrogen atoms per cm$^2$ along a given direction.

### 1.2.2 Determining HI column density

Considering an isothermal cloud of gas and ignoring any background source, the solution of the equation of radiation transfer in terms of brightness temperature will describe a pure emitting process:

$$T_B(v) = T\,\left[1 - e^{-\tau(v)}\right], \tag{1.24}$$

where $T$ is the temperature of the cloud and $\tau(v)$ the optical depth of the cloud at radial velocity $v$ (Wilson et al., 2013). If $\tau(v) \ll 1$, quadratic and higher terms in the Taylor series $e^{-\tau} = 1 - \tau + \tau^2/2 - ...$ can be neglected resulting in

$$T_B(v) = T\,\tau(v) \qquad \text{for } \tau \ll 1. \tag{1.25}$$

Since we already have a relation between $T_B$ and the observed flux density (1.23), we now want to find an explicit relation between optical depth and neutral hydrogen

mass. Optical depth can also be expressed in terms of the absorption coefficient per unit frequency interval $\kappa_v$, as $d\tau(v) = \kappa_v(s)\, ds$, where $s$ is the length of the gas cloud along the line of sight. To compute the absorption coefficient $\kappa_v$, knowledge of the population of the two states relative to the hyperfine splitting in the ground state of hydrogen is required. As seen in (1.19) the half-life time of the excited triplet state is around $10^7$ years. However, a typical interstellar hydrogen atom will change the spin of the electron due to collisions, without the need for photon emission, in about 400 years. Thus in practically all astronomical situations the relative population of the hyperfine structure levels will be determined by collisions, since these radiationless transitions dominate over the photon-emitting ones (Wilson et al., 2013). Let the relative population of the levels be described by an excitation temperature which in this case is usually called the *spin temperature* $T_s$. If we consider a system in full thermodynamic equilibrium, the electrons will be distributed over different states so that the two hyperfine levels are populated according to the Boltzmann distribution

$$\frac{N_1}{N_0} = \frac{g_1}{g_0} \exp\left(-\frac{h\,\nu_{1,0}}{k_b\,T_s}\right), \tag{1.26}$$

where $N_1$ and $N_0$ are the number of electrons per unit volume in the triplet state ($s = 1$) and singlet ($s = 0$) state, respectively, $g_1$ and $g_0$ are the statistical weights of the states, and $\nu_{1,0}$ is frequency of the transition. Since the excited triplet level has a 3 state degeneracy and the singlet level is non degenerate $g_1/g_0 = 3$. Knowing that

$$T_0 = \frac{h\,\nu_{10}}{k_b} = 0.0682 \text{ K}, \tag{1.27}$$

we can safely assume that for every practical temperature $T_s \gg T_0$. In this limiting case equation (1.26) will read

$$\frac{N_1}{N_0} = \frac{g_1}{g_0} = 3. \tag{1.28}$$

Given the following equation for the absorption coefficient (Wilson et al., 2013, page 322):

$$\kappa_v = \frac{c^2}{8\pi}\frac{1}{\nu_{10}^2}\frac{g_1}{g_0} N_0\, A_{10} \left[1 - \exp\left(-\frac{h\,\nu_{10}}{k_b\,T_s}\right)\right] \phi(v), \tag{1.29}$$

where $A_{10}$ is the Einstein emission coefficient and $\phi$ is the spectral line profile, for $T_s \gg T_0$ we get

$$\kappa_v = \frac{3\,c^2}{32\pi}\frac{1}{\nu_{10}}\,A_{10}\,N_H\,\frac{h}{k_b\,T_s}\,\phi(v), \tag{1.30}$$

where the total number of neutral hydrogen atoms per unit volume as been introduced by $N_H = N_0 + N_1 = 4N_0$.

With this absorption coefficient, the differential optical depth can be expressed as

$$d\tau\left(\frac{v}{\mathrm{km/s}}\right) = 5.4873\times10^{-19}\left(\frac{N_H}{\mathrm{cm^{-3}}}\right)\left(\frac{T_s(s)}{\mathrm{K}}\right)^{-1}\left(\frac{\phi(v)}{\mathrm{s/km}}\right)d\left(\frac{s}{\mathrm{cm}}\right). \tag{1.31}$$

Knowing that the integral of $\phi(v)$ over $v$ is normalized to 1, if the spin temperature $T_s$ is constant along the line of sight, integrating (1.31) both over $s$ and over $v$ yields:

$$\int_{-\infty}^{\infty}\tau(v)\,d\left(\frac{v}{\mathrm{km/s}}\right) = 5.4873\times10^{-19}\left(\frac{T_s}{K}\right)^{-1}\int_0^{\infty}\left(\frac{N_H(s)}{\mathrm{cm^{-3}}}\right)d\left(\frac{s}{\mathrm{cm}}\right). \tag{1.32}$$

If we now define column density $\mathcal{N}_H$ as

$$\frac{\mathcal{N}_H}{\mathrm{cm^{-2}}} = \int_0^{\infty}\left(\frac{N_H(s)}{\mathrm{cm^{-3}}}\right)d\left(\frac{s}{\mathrm{cm}}\right). \tag{1.33}$$

then equation (1.32) becomes:

$$\frac{\mathcal{N}_H}{\mathrm{cm^{-2}}} = 1.8224\times10^{18}\left(\frac{T_s}{K}\right)\int_{-\infty}^{\infty}\tau(v)\,d\left(\frac{v}{\mathrm{km/s}}\right). \tag{1.34}$$

Now, using equation (1.25), where $T_s$ is the temperature of the source, we can express column density in terms of the brightness temperature $T_B$ as

$$\frac{\mathcal{N}_H}{\mathrm{cm^{-2}}} = 1.8224\times10^{18}\int_{-\infty}^{\infty}\left(\frac{T_B(v)}{K}\right)d\left(\frac{v}{\mathrm{km/s}}\right). \tag{1.35}$$

Finally, taking the brightness temperature from equation (1.23), and substituting the HI transition wavelength $\lambda = 21.07$ cm, we obtain the following relation between column density and flux density:

$$\frac{\mathcal{N}_H}{\mathrm{cm^{-2}}} = 1.1038\times10^{21}\int_{-\infty}^{\infty}\left(\frac{S_v}{\mathrm{mJy}}\right)\left(\frac{\theta}{\mathrm{arcsec}}\right)^{-2}d\left(\frac{v}{\mathrm{km/s}}\right). \tag{1.36}$$

## 1.3 Rotational velocity of external galaxies

As discussed previously, an important class of observations of the ISM in external galaxies consists of measurements of emission line spectra, both over a one-dimensional or two-dimensional region of the sky covering the galaxy. The Doppler shifts of the observed lines contain information about the line of sight kinematics of the galaxy. These observations can be used to obtain a mean velocity field $V(x, y)$ at a given position $(x, y)$ in the sky, where the Cartesian system $(x, y)$ is defined such that the center of the object is at $(0, 0)$ and the galaxy's major axis lies at $y = 0$. Traditionally, optical measurements are made with a long slit spectrograph, where the observed one-dimensional region is best placed along the major axis, at $y = 0$ (Bovy, 2022). These observations therefore produce a mean velocity field $V(x, 0)$, in the notation above. A typical example are the velocity fields observed by Rubin et al. (1980) (Fig. 1.2). As we can see in the figure, line of sight velocities quickly rise in absolute value moving away from the center, then the trend mostly flattens out in the outer regions. The causes of such a behaviour will be explained later.

Radio 21 cm and CO observations typically provide a two-dimensional velocity field $V(x, y)$ instead. A canonical example are the velocity fields of 22 spiral galaxy observed by Bosma (1978) (Fig. 1.3) These velocity fields display all of the major types of features that can be observed in two-dimensional velocity fields, such as regions of vertical contours of constant velocities near the center, the spider-like shape of contours extending radially from the center, and closed contours along the major axis away from the center. All of these result from different types of galactic rotation, as will be discussed below. These two-dimensional velocity fields contain much more information than one-dimensional slices at $y = 0$, and are essential to determining the best possible representation of a galaxy's rotation curve, which can be defined as the velocity a test particle would have if it was on a circular orbit in an axially-symmetrized version of a galaxy's mass distribution; "axially-symmetrized" stands for the galaxy's mass distribution at radius $\mathbf{R}$ averaged around the axis perpendicular to the disk. The determination of a galaxy's rotation curve from observations of its velocity field will be discussed in more detail in section 1.3.2.

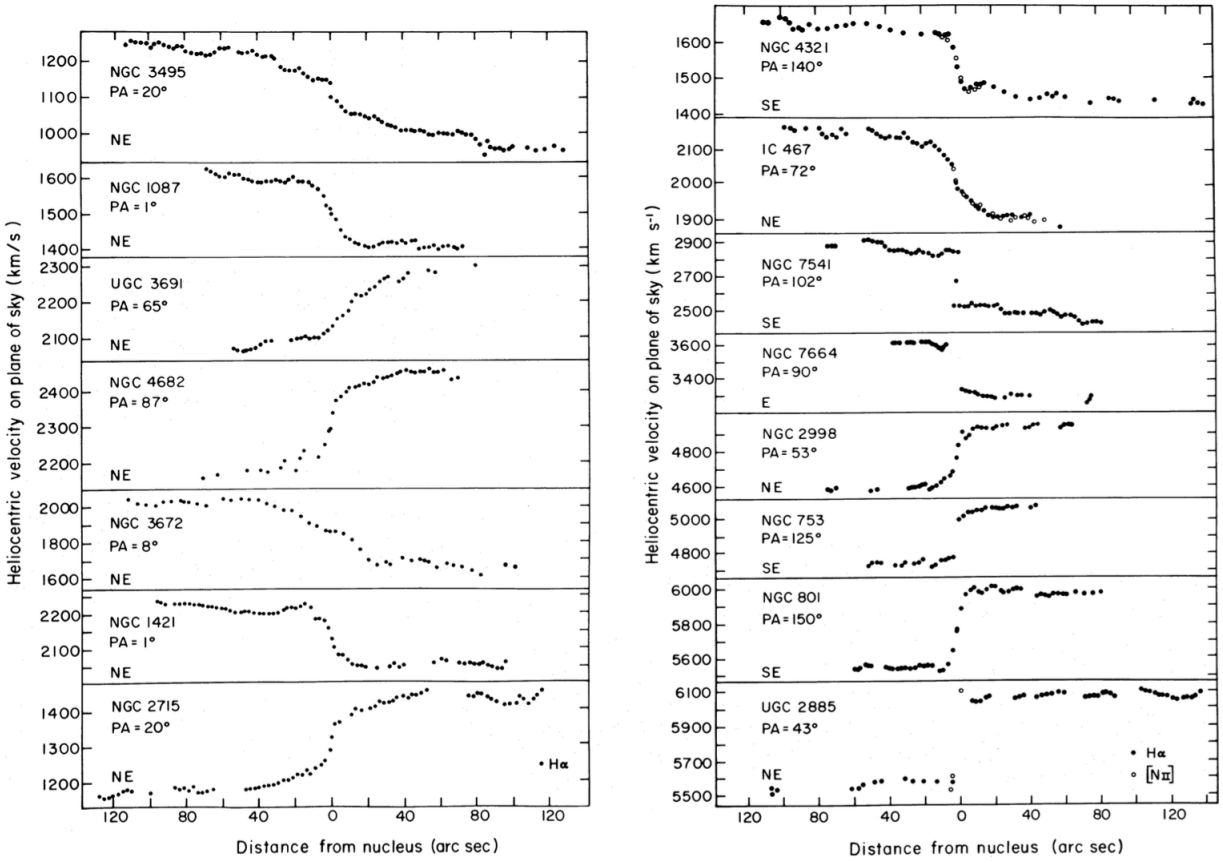


**Figure 1.2:** Heliocentric line of sight velocities on the plane of the sky as a function of angular distance from the nucleus in Sc galaxies observed by Rubin et al. (1980).

### 1.3.1 Modelling velocity fields

To model galactic velocity fields, we are considering a disk in differential rotation around its center with a rotation curve $v_c(R)$, and assuming that the systemic velocity (i.e. the speed of the galaxy's center with respect to the observer) is zero, because it simply adds a constant to all observed velocities. A region at radius $R$ in the galaxy will have a velocity $\mathbf{v} = -v_c(R)\,(\hat{\mathbf{R}} \times \hat{\mathbf{n}})$ where $\hat{\mathbf{R}}$ is the unit vector that points to the location from the galaxy's center and $\hat{\mathbf{n}}$ is a unit vector normal to the disk plane. The observed line of sight velocity $V$ is then the dot product of the velocity $\mathbf{v}$ and the direction of the line of sight $\hat{\mathbf{r}}$ pointing from us to the galaxy:

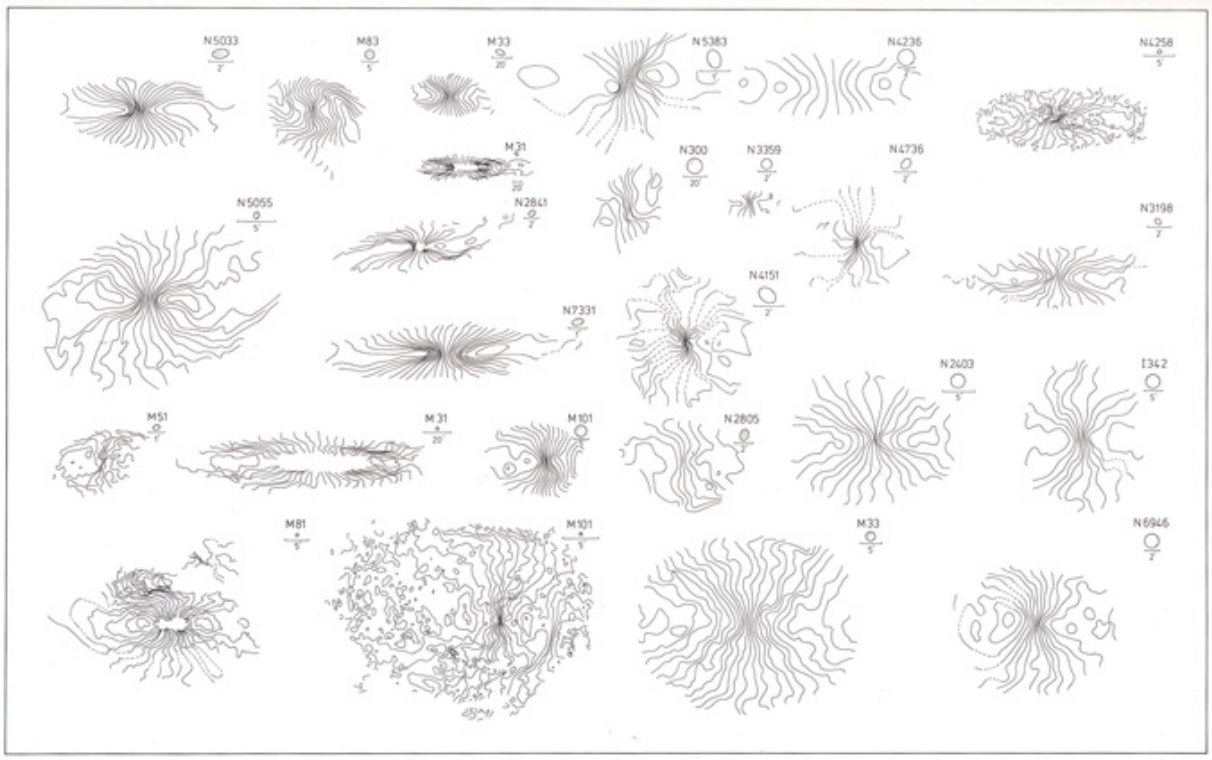


**Figure 1.3:** Velocity field of 22 spiral galaxies observed by Bosma (1978). All pictures have been oriented such that the major axis of observed galaxies is roughly horizontal.

$$V = \hat{\mathbf{r}} \cdot \mathbf{v} = v_c(R)\, \hat{\mathbf{R}} \cdot (\hat{\mathbf{r}} \times \hat{\mathbf{n}}) = v_c(R)\, \hat{\mathbf{R}} \cdot \hat{\mathbf{k}} \sin i \tag{1.37}$$

where $i$ is the galaxy's inclination, defined such that an edge-on galaxy has an inclination of $i = 90°$, and $\hat{\mathbf{k}} = (\hat{\mathbf{r}} \times \hat{\mathbf{n}})/\sin i$ is the unit vector perpendicular to both $\hat{\mathbf{r}}$ and $\hat{\mathbf{n}}$. If we introduce a polar coordinates system $(R, \theta)$ that lies in the plane of the galaxy, then $\hat{\mathbf{k}}$ and $\hat{\mathbf{R}}$ will both fall in this plane and will have Cartesian components $(1, 0)$ and $(\cos\theta, \sin\theta)$ respectively. Under these assumptions equation (1.37) will become

$$V(x, y) = v_c(R) \cos\theta \sin i. \tag{1.38}$$

From this equation we can see why when one is limited to observing a one-dimensional slice of the two-dimensional velocity field, it is advantageous to observe along the major axis, where $\theta = 0$ or $180°$, and therefore the observed velocity $V(x, 0) = \text{sign}(x)\, v_c(R)\ \sin\ i$

is maximum with respect to $\theta$. This produces the S-shape behaviour in the Rubin et al. (1980) curves in Figure 1.2. We now look at a few idealized rotation models to better understand the observed velocity fields.

### Solid body rotation

If we first take a homogeneous density sphere with $\rho(R) = \rho_0 =$ constant for all $R < a$ and zero at $R \geq a$, its gravitational potential for $R < a$ is that of a harmonic oscillator:

$$\Psi(R) = \frac{2\pi\, G\, \rho_0}{3}(R^2 - 3a^2). \tag{1.39}$$

This produces a solid-body rotation with

$$\Omega = \frac{v_c(R)}{R} = \sqrt{\frac{4\pi\, G\, \rho_0}{3}} = \text{constant}. \tag{1.40}$$

If we plug this into equation (1.38) we get

$$V(x, y) = \Omega\, R \cos\theta \sin i. \tag{1.41}$$

Because $x = R\cos\theta$, this becomes simply

$$V(x, y) = \Omega\, x \sin i. \tag{1.42}$$

The velocity therefore does not depend on $y$ and contours of constant velocity will run parallel to the $y$ axis. When vertical contours of constant velocity near the center of a two-dimensional velocity field are observed, see Bosma, 1978 (1.3), this is indicative of solid-body rotation.

### Flat rotation curve

When we instead have a flat rotation curve with $v_c(R) = v_c =$ constant, equation (1.38) becomes

$$V(x, y) = v_c \cos\theta \sin i = v_c \frac{x}{R} \sin i = v_c \frac{\cos i \sin i}{\sqrt{\cos^2 i + \tan^2 \phi}}, \tag{1.43}$$

where $\phi$ is the angular coordinate of the polar coordinate system $(r, \phi)$ that lies in the plane of the sky, so that $\tan\phi = y/x$. Since the observed velocity only depends on $\tan\phi$ (the galaxy inclination $i$ is constant), it only depends on $y/x$. Thus, regions of the two-dimensional velocity field of galaxies in which the contours are straight lines pointing away from the center are indicative of a flat rotation curve.

**Peaked rotation curve**

Finally, we consider what happens when the rotation curve $v_c(R)$ rises to a maximum value and then declines. In this case we should see a closed contour in the two-dimensional velocity field $V(x, y)$ around where the rotation curve reaches its maximum. To see why we need to refer to equation (1.38). Along at the galaxy's major axis at $y = 0$ or $\theta = 0$, $V(x, 0) = v_c(R)\sin i$. In this case there will be two different $R_1$ and $R_2$ at which a certain value of $v_c(R_1) = v_c(R_2)$ is reached, one at a radius before the peak of the rotation curve and one after. Now we trace the contour line at constant $V(x, y)$ for rising values of $\theta$ and starting at $R_1$. Since $\cos\theta$ is falling, then the contour line has to curve towards the x value at which the rotation curve peaks, because $v_c(R)$ needs to rise to keep $V(x, y)$ constant. This will continue until the maximum value of $v_c(R)$ is reached. After that the contour line will turn down towards lower values of $\theta$ until reaching the second point at $R_2$ along the major axis where $v_c(R_1) = v_c(R_2)$. The same happens when continuing down to negative values of $\theta$ until the contour lines closes. Therefore, all contour lines corresponding to $V(x, 0) = v_c(R)\sin i$ for values of $v_c(R)$ that exist both before and after the radius at which the velocity curve peaks must close.

Disk galaxies have light distribution that is observed to be exponential, and if we assign a constant mass-to-light ratio to these galaxies, then the mass distribution should be that of an exponential disk, with a surface density profile that can be written as

$$\Sigma(R) = \Sigma(0)\, e^{-R/R_d}, \tag{1.44}$$

where the scale length $R_d$ is typically of a few kpc. Since the rotation curve for an exponential disk peaks at values of $R \approx 2.15\, R_d$, we would expect the rotation curve of galaxies to peak at $R \approx 10$ kpc if their mass-to-light ratio is constant. Thus we should see that contours in the velocity field start to close around $R \approx 10$ kpc, but this is not

what is observed. In fact, in the velocity fields of Bosma (1978) we see that the contours of the velocity field do not close for many of the observed galaxies.

### 1.3.2 From velocity fields to rotation curves

Now that we have described the basic structure of two-dimensional velocity fields, we will discuss how to determine the one-dimensional rotation curve from them. There are various methods for extracting the one-dimensional rotation curve $v_c(R)$ of a galaxy from a map of its two dimensional velocity field $V(x, y)$. The simplest of these uses the fact that $V(x, 0) = \mathrm{sign}(x)\, v_c(R) \sin i$ and therefore

$$v_c(R) = \frac{V(x, 0)}{\mathrm{sign}(x) \sin i}\,. \tag{1.45}$$

If the two-dimensional velocity field $V(x, y)$ is available, then this method discards too much information. Nevertheless, many observations, especially the ones using optical long-slit spectroscopy, only determine $V(x, 0)$, and using equation (1.45) is then the only option to extract the rotation curve. Notice that in this case the galaxy's inclination needs to be known from photometric observations, otherwise it is degenerate with $v_c(R)$. If we instead have measured the full two-dimensional velocity field $V(x, y)$, using for example radio interferometric observations of HI emission, we can use a more general form of equation (1.38) (see below) and fit for the free parameters. These parameters can include the center of the galaxy on the sky $(x_0, y_0)$, the angle $\phi_0$ of the galaxy's major axis, the systemic velocity of the galaxy $V_0$, the inclination $i$, and parameters describing $v_c(R)$. To account for any deviations from circular motion, the fitting procedure often includes a radial term $v_r(R)$ to model radial motions in the plane. If we include all parameters, then the observed velocity field is given by

$$V(x, y) = V_0 + v_c(R)\cos\theta \sin i + v_r(R) \sin\theta \sin i\,, \tag{1.46}$$

where

$$\cos\theta = \frac{-(x - x_0)\sin\phi_0 + (y - y_0)\cos\phi_0}{R}\,, \tag{1.47}$$

$$\sin\theta = \frac{-(x - x_0)\cos\phi_0 - (y - y_0)\sin\phi_0}{R\cos i}\,. \tag{1.48}$$

A popular method to carry out this fit is to perform it in a set of annuli around the center of the galaxy, where each annulus can in principle have its own values for all of the parameters above. This method is known as the **tilted-ring model** (e.g., Begeman, 1989), because the annuli can potentially be tilted with respect to each other if they have different ($i$, $\phi_0$).

### 1.3.3 Observed rotation curves

Now that we have seen how to extract rotation curves from observed velocity fields, we can discuss some classical results, starting with the rotation curves derived from the one-dimensional velocity field from Rubin et al. (1980) (Fig.1.4). Most of these rotation curves get close to the galaxies' optical radius $D_{25}$, which is defined as the radius corresponding to the isophote where the surface brightness is $25\,\mathrm{mag\,arcsec^{-2}}$ in the B band of the UBV photometric system.

This is a convenient definition of the radius of a galaxy's light, because little light is emitted outside of it. Since the observations extend that far, we should expect to see the rotation curve to peak and then turn down if these disk galaxies were pure (stellar) exponential disks. However, the rotation curves remain mostly flat at the largest observed radius.

Kent (1986) tried to fit these rotation curves with constant mass-to-light ratios, decomposing the galaxies into a bulge and a disk component, and allowed each of them to have a different constant mass-to-light ratio, $(M/L)_b$ and $(M/L)_d$ respectively. In particular, they investigated whether the entire rotation curve can be fitted by adjusting the mass-to-light ratios such that they provide a good fit to the rotation curve in the inner region. Kent (1986) did this for a significant sample of galaxies, but we will look at the results of two particular objects that demonstrate the two different situations that occur (Fig. 1.5).

We can notice that in the case of NGC 801 the fit provides an excellent match to the entire observed rotation curve, as seen on the bottom right panel of Figure 1.5, thus there appears to be no need for additional matter components. Nevertheless, the mass-to-light

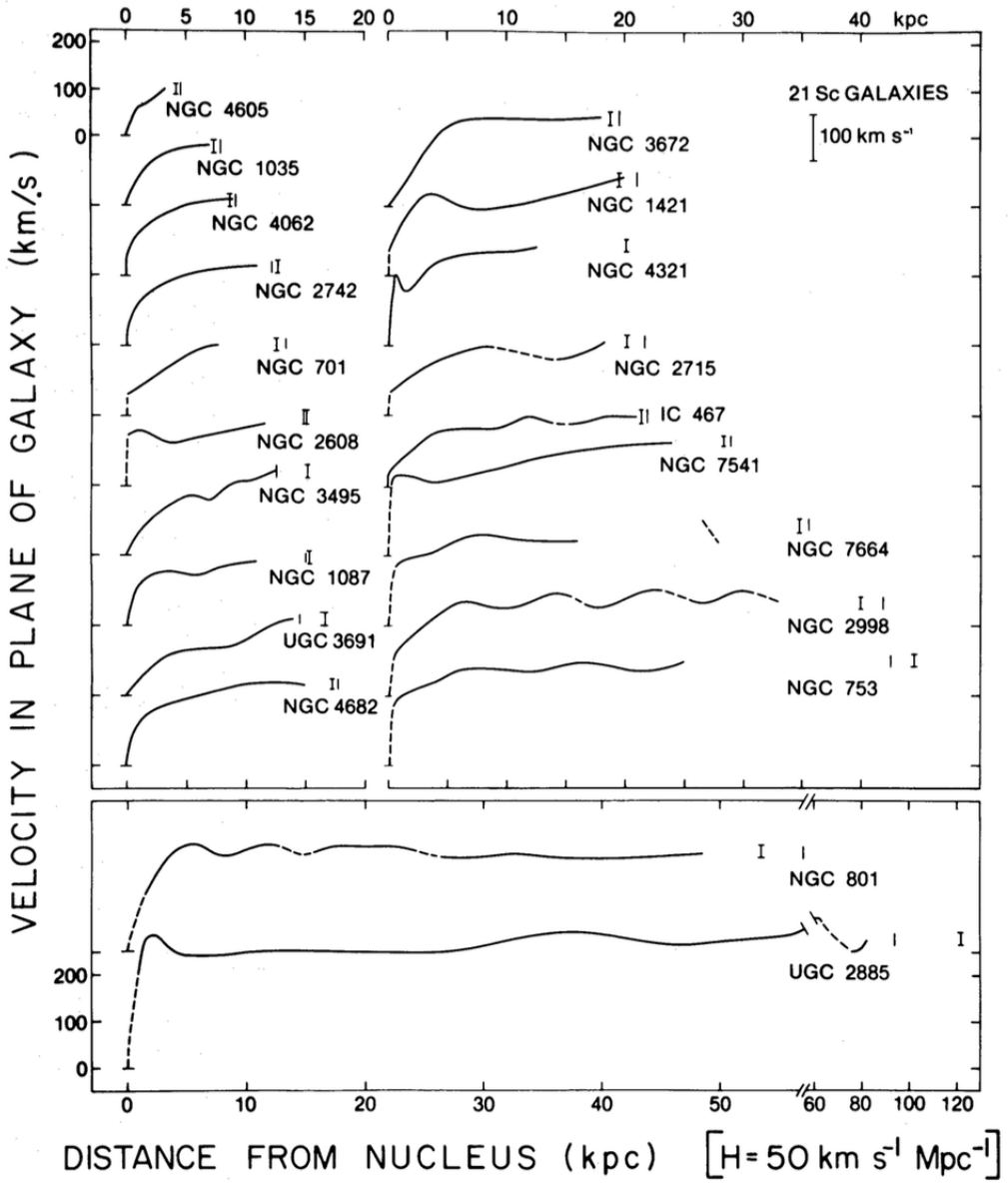


**Figure 1.4:** Rotation curves of Sc galaxies from Rubin et al. (1980). The optical radius $D_{25}$ is displayed by the vertical line segments. Rotation velocity for most galaxy is shown to initially increase, and then flatten when approaching the optical radius.

ratios $(M/L)_b \approx 1.5$ and $(M/L)_d \approx 7$ are somewhat high. NGC 753, however, displays very different behaviour, as can be observed in the bottom left panel of Figure 1.5. The fit successfully reproduces the inner ($R \lesssim 10\,\mathrm{kpc}$) rotation curve, but because the surface

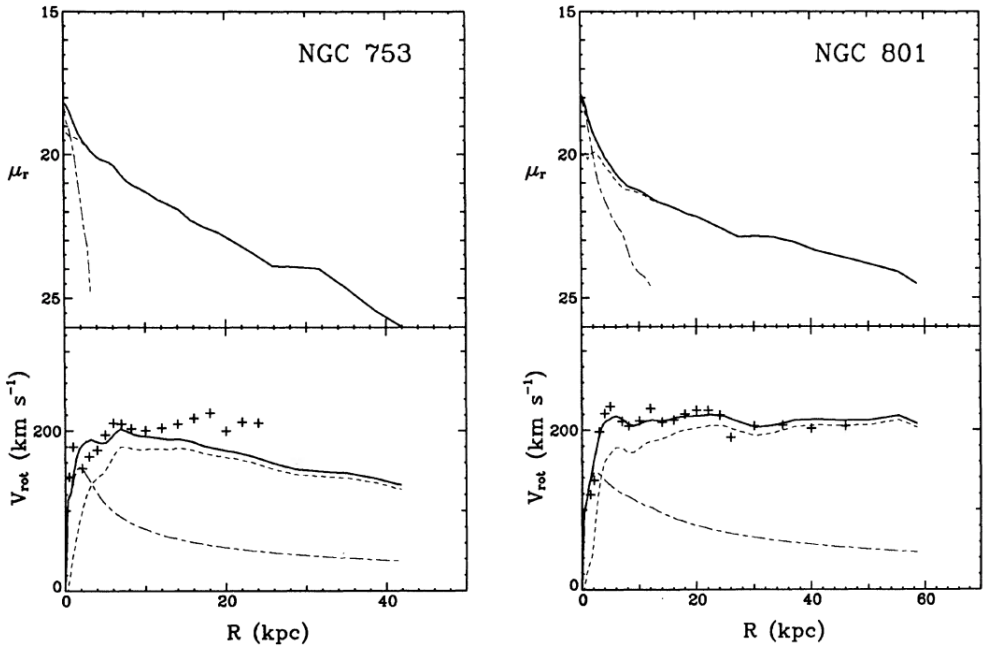


**Figure 1.5:** Top panels: surface brightness profiles of the galaxies NGC 753 (left) and NGC 801 (right) and their decomposition into bulge and disk components. Bottom panels: the rotation curves of those galaxies, with constant mass-to-light ratio fits. The crosses represent the observed data, while the continuous line the fits (from Kent, 1986).

brightness profile steeply declines at large radii, the fit drops at $R \gtrsim 10\,\mathrm{kpc}$, while the observed rotation curve remains approximately constant. Therefore, it is clear that we need an additional component to explain NGC 753's rotation curve. In particular the mass-to-light ratio needs to increase with $R$. Nevertheless, when it comes to NGC 801, and many other galaxies with similar behaviour studied by Kent (1986), there seems to be no such issues and no need for additional matter components, since the fit is a good match for the observed data. However, here the problem lies in the use of optical data. Curves obtained this way typically do not extend far enough to unambiguously show the influence of this additional component (that is identified with a dark matter component).

Radio observations of the 21 cm line of neutral hydrogen allow the velocity field to be mapped at larger distances than the optical rotation curves that we have discussed so far. In Figure 1.3, we showed the velocity maps from Bosma (1978). Many of these maps extend beyond the optical radius of the galaxy, typically to between the optical

radius and twice that radius. Therefore, we expect to see a drop-off in the rotation curves in the outer observed regions, and we should see clear closed contours in the two-dimensional velocity maps $V(x, y)$. However, looking at the velocity maps, we see that closed contours are rare and that most of the galaxies in Bosma's sample have the "spider" shape indicative of flat rotation curves. Bosma used a model similar to that described for equation (1.46) to extract the one-dimensional rotation curve from the two-dimensional velocity fields (Fig. 1.6). The majority of these rotation curves remain flat to the furthest observed point, typically well beyond the optical radius of the galaxy. Bosma used an inversion of the observed velocity curves $v_c(R)$ to obtain the galaxies' surface densities $\Sigma(R)$ and mass profiles $M(< R)$ (Fig. 1.7). We can clearly see that a flat rotation curve implies $M(< R) \propto R$. These galaxies have large masses enclosed within the radius of the last point where $v_c(R)$ was measured and the mass should be expected to keep rising, because the increase barely flattens at large radii. These masses are larger by a factor of at least a few than the mass observed within the optical radius. Thus, we see that the evidence from HI rotation curves of external galaxies points toward a clear indication: galaxies must be surrounded by an extended distribution of non-luminous, dark matter that causes the total mass-to-light ratio to increase with radius.

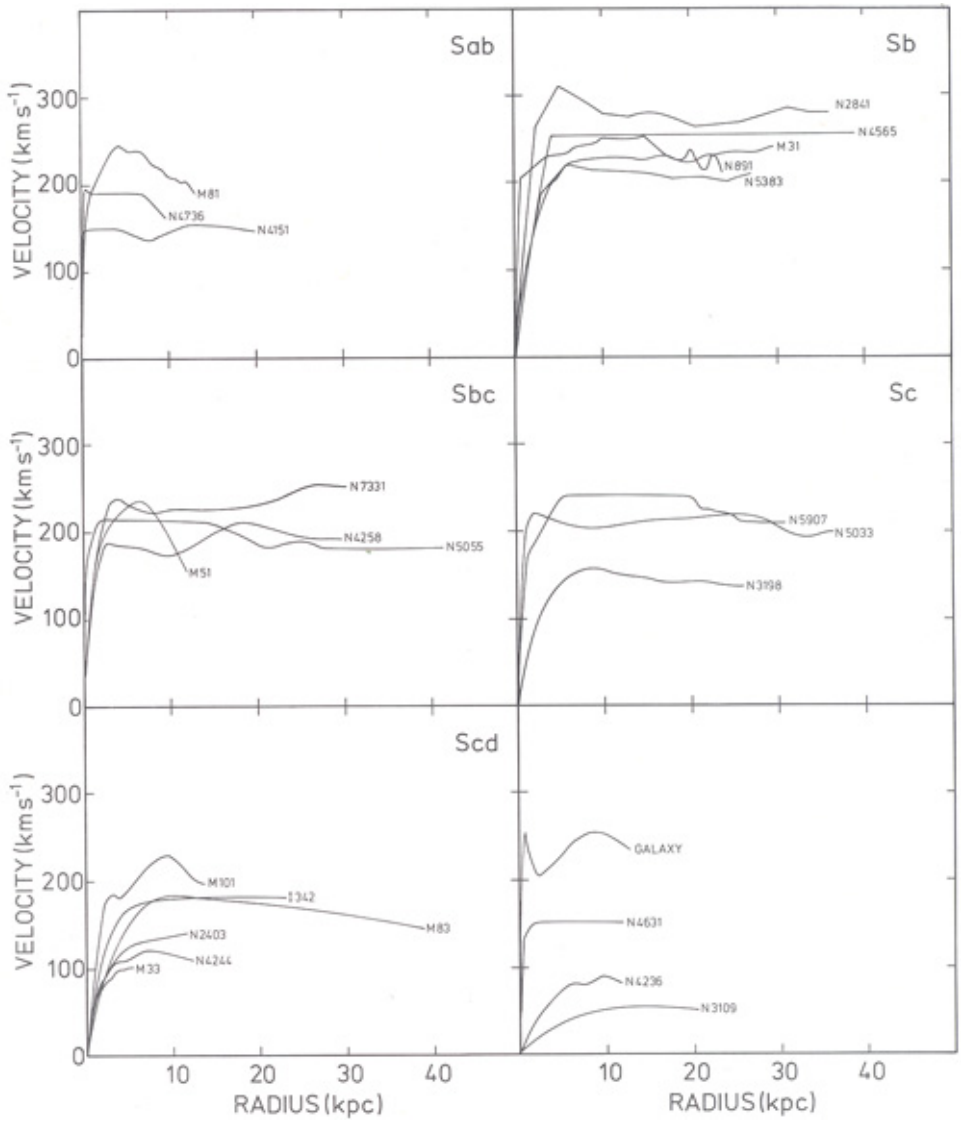


**Figure 1.6:** Rotation curves of various galaxies observed by Bosma (1978), arranged by Hubble type. The "Galaxy" rotation curve is that of the Milky Way.

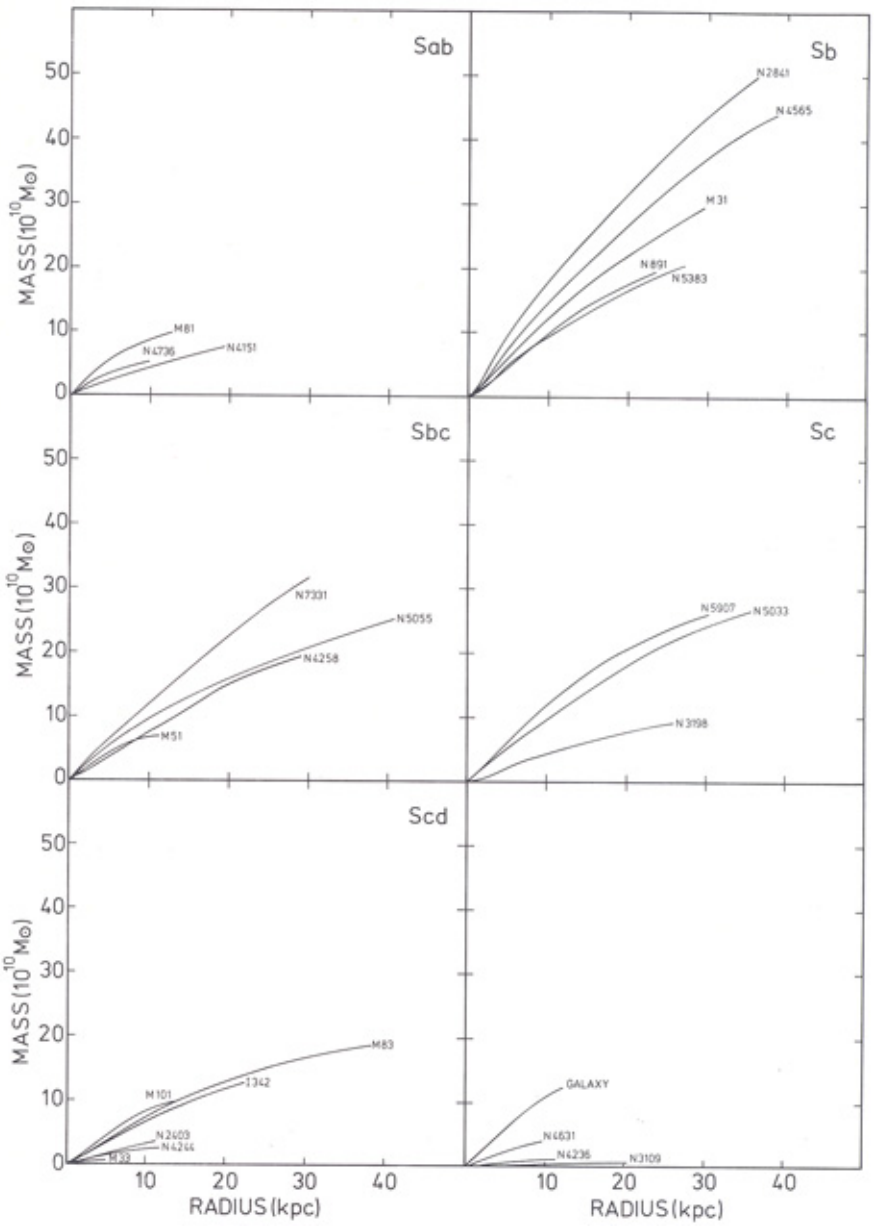


**Figure 1.7:** Mass profiles of the galaxies in Figure 1.6

# Chapter 2

# Generating mock HI observations

Now that we have discussed the theory behind the HI 21 cm emission and the ways we can exploit such an emission to study the properties of HI sources and in particular inferring the dynamical properties and mass distribution of an entire galaxy, we will use and test those methods on a simulated galaxy and compare the results with data from real observations. This way we are able to test the accuracy of current galaxy formation simulations and gain insight into how we can best interpret these HI observations of real galaxies. In this analysis, we are taking a simulated, isolated Milky Way-like disc galaxy, generated by the moving-mesh code AREPO (Weinberger et al., 2020) using the ***SMUGGLE*** model (Marinacci et al., 2019), and we will briefly describe its functioning below. Afterwards, we will introduce the ***MARTINI*** (Mock Array Radio Telescope Interferometry of the Neutral ISM) package (Oman, 2019), the software that was used to create a mock observation of HI line emission from the simulated galaxy. We will describe its operational methods and the interface that has been developed to allow MARTINI to operate on the SMUGGLE simulation, and then discuss the setup we used to create the mock observations.

## 2.1 The SMUGGLE model for galaxy formation simulations

Let us briefly introduce the SMUGGLE model that was used to create the simulated galaxy (Marinacci et al., 2019). Understanding and modeling the formation of galaxies is a daunting task owing to the large dynamical range in space and time as well as the wide range of physical mechanisms involved. To get a feel for the wide range of spatial scales needed in principle to carry out such a task, notice for instance that the physics of a single supernova explosion occurs on sub-parsec scales, but megaparsec scales are required to capture the gravitational collapse and assembly of a halo such as that of the Milky Way. So, modern simulations require both the modelling of large volumes, while still resolving small scales in order to capture the relevant galaxy formation physics. To that end, sub-grid models are used. They are models that numerically describe physical processes that occur on scales at or below the resolution of the simulation.

Nearly all physical processes involved in galaxy formation are implemented to some degree as sub-grid models including radiative cooling, star formation, stellar evolution, SN feedback, Active Galactic Nuclei (AGN) feedback, etc. The most important physical processes beyond gravity and hydrodynamics are the formation of stars and their interaction with the surrounding interstellar medium. However, the ISM structure in galaxies is complex, with hot, warm and cold phases of gas coexisting and interacting, and following this complex structure becomes computationally more expensive when the numerical resolution increase. One common way to overcome this obstacle is to impose an effective equation of state for the dense gas instead of directly resolving the individual gas phases and processes within the ISM, but the overall structure of galaxies in such models is overly-smooth on scales of a few hundred parsecs, the scale of molecular clouds. This severely limits the predictive power of these simulations on small scales. Current state-of-the-art zoom-in simulations have reached a resolution where it becomes desirable to remove the effective ISM approach. SMUGGLE represents a new ISM model that captures the multi-phase gas structure alongside a more explicit local stellar feedback model, and yields converged results at the resolution of current cosmological simulations, resolving smaller scales withing the ISM. Figure 2.1 provides a high level schematic overview

of the physical processes that are considered in the model. Figure 2.2 is an image of the stellar light of the galaxy simulated with this model, that will be further studied through HI line emission later.

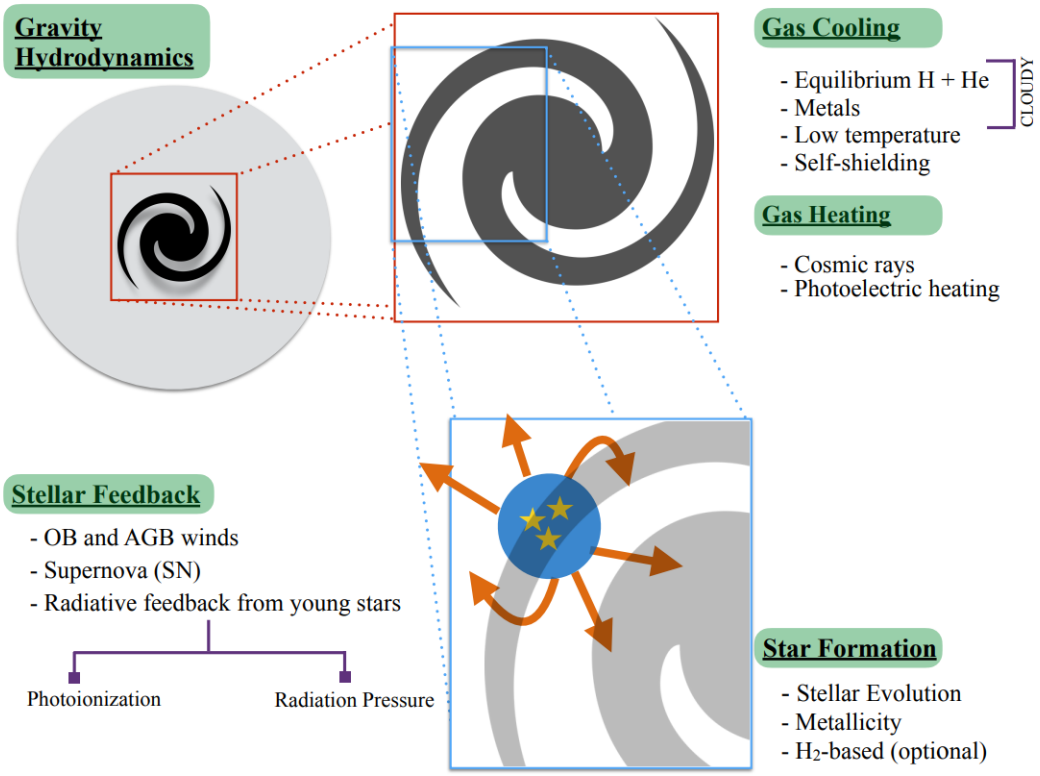


**Figure 2.1:** Schematic overview of the main physical processes included in the SMUGGLE model. These are: gravity and gas dynamics, radiative cooling of the gas, including metal line cooling and low-temperature atomic and molecular cooling with a prescription for self-shielding from the UV radiation, gas heating caused by the interaction with cosmic rays and photoelectric heating, a stochastic prescription for star formation, stellar evolution with the associated mass and metal return, and three main channels for stellar feedback, namely stellar winds, supernova feedback and radiation feedback from massive stars. See Marinacci et al. (2019) for a more detailed description of these processes.

## 2.2 The MARTINI Package

We will now take a look at MARTINI, the software used for creating mock observations of the 21 cm line emitted by the simulated galaxy analysed in this work (e.g. used in Oman et al., 2019). MARTINI is a modular package for the creation of synthetic resolved HI line observations (data cubes) of (smoothed-particle) hydrodynamics simulations of galaxies. MARTINI is object oriented: the various aspects of the mock-observing processes are divided logically into sub-modules, each providing a class which can be configured as desired.

### 2.2.1 MARTINI sub-modules

The MARTINI software package is responsible for creating the synthetic HI data cubes from simulated data. To use this code one must first create instances of classes from each of the required and optional sub-modules, then create a 'martini' object with these instances as arguments. The object can then be used to create synthetic observations in the form of a data cube, i.e. a $3^{\text{rd}}$-rank tensor-like object, of which each pixel has data on the flux density of the 21 cm line depending on location in the sky (the first two dimensions), and on line of sight velocity (the third dimension) accounting for the Doppler shift in the emission line. The following is a list of the needed parameters (for a more extensive description of each module see Oman, 2019):

- **source** - A description of the HI emitting object, including position in the sky, geometry, distance from observer, and an interface to the simulation data.
- **datacube** - A description of the datacube to create, including pixels, channels, sky position.
- **beam** - A description of the beam for the simulated telescope, including beam model (such as a Gaussian beam) and axis angular size.
- **noise** - A description of the simulated noise, including the root mean square amplitude of the noise field (rms).

- **sph_kernel** - A description of the SPH (Smoothed-Particle Hydrodynamics) smoothing kernel, used for projecting flux density in the pixels of the data cube.
- **spectral_model** - A description of the HI line produced by a particle of given properties. Examples are a Dirac-delta spectrum model, or a temperature-dependent Gaussian line model.

In theory, once the values for these sub-module instances have been chosen, one would just need to create a martini class object passing all of the above as parameters and the mock observation process should be complete. However, in our case we still needed to develop a custom source, so that MARTINI would be able to operate on our simulated data (see below).

### 2.2.2 Arepo custom source

In the generation of our mock HI observation we used almost all of the base classes provided by MARTINI. The only problem resided within the 'source' sub-module, since our simulated galaxy came from a simulation performed with te moving-mesh code AREPO and MARTINI does not provide a class interface to deal with for this particular type of simulation data. Fortunately, it does provide a class for the 'IllustrisTNG' simulations (Nelson et al., 2019), the output of which is very similar to that of AREPO. So a custom source class was implemented by modifying the 'TNGSource' class as needed. For instance, the various points where distinctions between galaxy groups and clusters (Halos in TNG) and individual galaxies (Sub-Halos) were removed, since we are only interested in observing one specific galaxy. Then other parameters needed to be changed such as how to set the galaxy's center position and peculiar velocity. Lastly, MARTINI used a method to compute the neutral atomic hydrogen fraction for IllustrisTNG that was specifically developed for that simulation model, but in our case produced poor resolution, very smoothed data cubes, which glossed over most fine ISM features. That method has been replaced with a more direct and suitable approach. The AREPO output already provides, for every position in the simulated galaxy, both the individual element abundances of nine tracked species (H, He, C, N, O, Ne, Mg, Si, Fe) and the neutral atomic fraction over all the possible states for hydrogen. So, to get the HI mass

needed for the mock observing process, we simply took the total gas mass of each cell and multiplied it by the hydrogen mass fraction, and then by the neutral atomic fraction of that element. With these modifications MARTINI runs smoothly with the AREPO simulation data analysed in this work.

### 2.2.3 Mock observation parameters setup

Finally let us look at the setup of MARTINI that we used to generate our mock observations. In our setup the galaxy was placed, for simplicity, at the center of the observed region in the sky, so its center lies at 0 degrees in both right ascension and declination (this is irrelevant for the analysis that it is going to be performed on the resulting datacube). Its distance from the observe has been set at an arbitrary 10 Mpc, that is a typical distance for a local galaxy that is outside the Local Group (see, e.g. Oosterloo et al., 2007). Furthermore, the systemic line of sight velocity of the galaxy as been set to 0. So, the center of the galaxy should moving away from us with a speed dictated only by the expansion of the Universe and, as per Hubble's law, that amounts to $v = H_0\, D = 700$ km/s, by assuming $H_0 = 70$km/s/Mpc. By default, the galaxy is face-on, with an inclination $i = 0°$, but we have also generated models rotated around the galaxy's major axis to also observe it at inclinations $i = 30$, 60, and 90°. The data cube has been set up to contain $768 \times 768$ pixels for the spatial dimensions, with each pixel's size being 2.5 arcseconds (corresponding to 0.12 kpc at a distance of 10 Mpc), and 64 velocity channels, with a channel width of 16 km/s, which are sufficient to span the entire velocity range of the galaxy. For the mock telescope beam, the default MARTINI model of a Gaussian beam was used, with a value for both its major and minor axis of 20 arcseconds, making it a circular beam. A temperature-dependant Gaussian model was used for the broadening of the HI emission line. Finally, regarding the noise affecting the mock telescope, the default Gaussian noise model from MARTINI was used, with a rms = 0.09 Jy/beam. We chose these values for parameters of the mock observations to mimic Oosterloo et al. (2007) full resolutions observations of a similar galaxy (NGC 891). We adopt this approach so that we may better compare our results with real galaxy observations and analyse similarities and deviations that they may help us to notice possible issues regarding the simulation or our mock observation process.

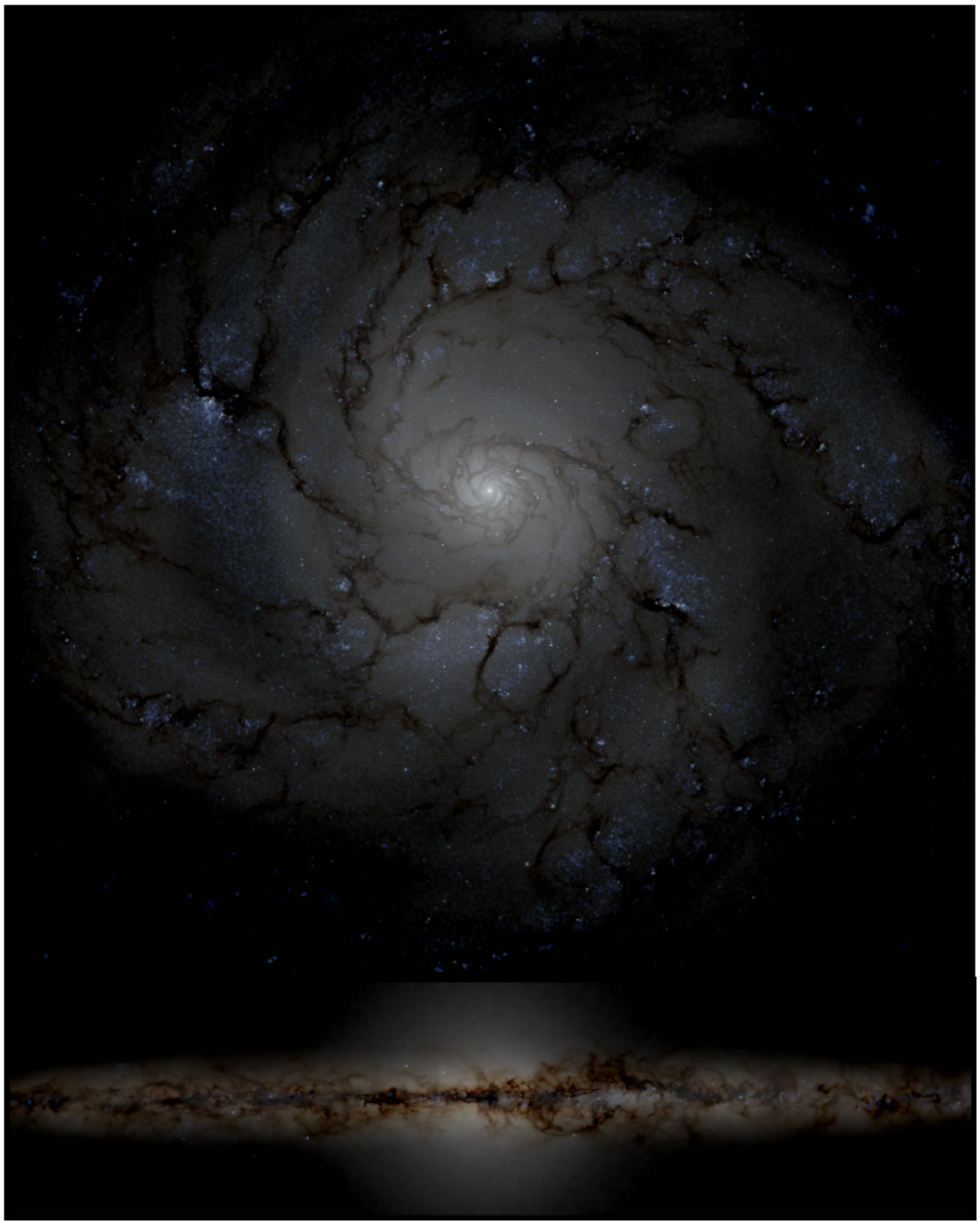

**Figure 2.2:** Stellar light (ugr-band) images of the simulated galaxy, both face on (top) and edge on (bottom).

# Chapter 3

# HI mock observation analysis

In this chapter we will present the results of our mock observation process, and will proceed to analyse them as we would do with data from real galaxies, with one key difference: we have access to all the properties of the galaxy we are "observing". We would like to stress that this is the most realistic possible way to compare real observed galaxies with results form numerical simulations. Indeed, the output of a simulation is very different from the actual data obtained from real telescopes, and comparing them directly might be very misleading. In fact, simulation output is missing all the complexities faced in astronomical observations, ranging from line shifting, to interference from other objects – both in the background and in front of the source, to simple measurement errors and noise that real equipment is subject to. This approach then, in addition to testing the validity of our current simulation models, may also help us to gain insight into the interpretation of real observations.

Specifically, this chapter will focus on different approaches to analyse the synthetic observations of our isolated galaxy. We will look at the maps we produced for HI column density, position velocity plots, moment maps, and finally we will briefly examine the output of the 3D-Barolo software (Di Teodoro et al., 2015), used to create a tilted-ring model fit for our galaxy.

## 3.1 HI column density maps

Figure 3.1 shows maps of HI column density for various inclinations of our simulated galaxy. These maps were obtained by integrating the flux of the datacube generated by MARTINI along the third axis, that of the velocity channels, using eq. (1.36). We set the minimum detectable column density at a value of $\mathcal{N}_H = 3\,\sigma_{\text{rms}} \approx 1.2 \times 10^{19}\,\text{cm}^{-2}$, where $\sigma_{\text{rms}} = 3.975 \times 10^{18}\,\text{cm}^{-2}$ is the HI column density corresponding to the noise level of the datacube, adopted from the NGC 891 observations by Oosterloo et al. (2007). From a visual analysis, we can clearly recognise the spiral nature of this galaxy, with various arms protruding from the center. In particular a dense disk can be observed with a radius of around 20 kpc, while a more rarefied HI gas component extends up to around 30 kpc. These dimensions are slightly inflated at higher inclinations, since having more matter along the line of sight will result in larger column densities, and so we are able to detect HI emission in regions with lower amount of gas further away from the galactic center.

### 3.1.1 Inferred total HI mass

From these observations we can estimate the total mass of neutral atomic hydrogen present in the galaxy. Under the usual assumptions – also used the derivations of section 1.2 – of optically thin gas, we can calculate the HI mass of the galaxy from the integrated HI flux using

$$\frac{M_{HI}}{\text{M}_\odot} \approx 2.36 \times 10^5 \left(\frac{S_{\text{int}}}{\text{Jy km/s}}\right) \left(\frac{d}{\text{Mpc}}\right)^2, \tag{3.1}$$

where $S_{\text{int}}$ is the integrated HI flux (i.e. the sum of the fluxes of all the pixels in the datacube) and $d$ is the distance to the source. We find that at higher inclinations we observe slightly more mass, since, as stated above, by having more matter along the line of sight, more pixels will reach the minimum detectable column density and will thus be counted towards the total mass value. By computing the mean value of total neutral atomic hydrogen mass for the different inclinations we get $M_{HI} \approx 3.3 \times 10^9\,\text{M}_\odot$. This value is compatible with observations of total HI mass in Milky Way-type galaxies, like for instance from Rhee et al. (2018). The exact total mass for neutral hydrogen

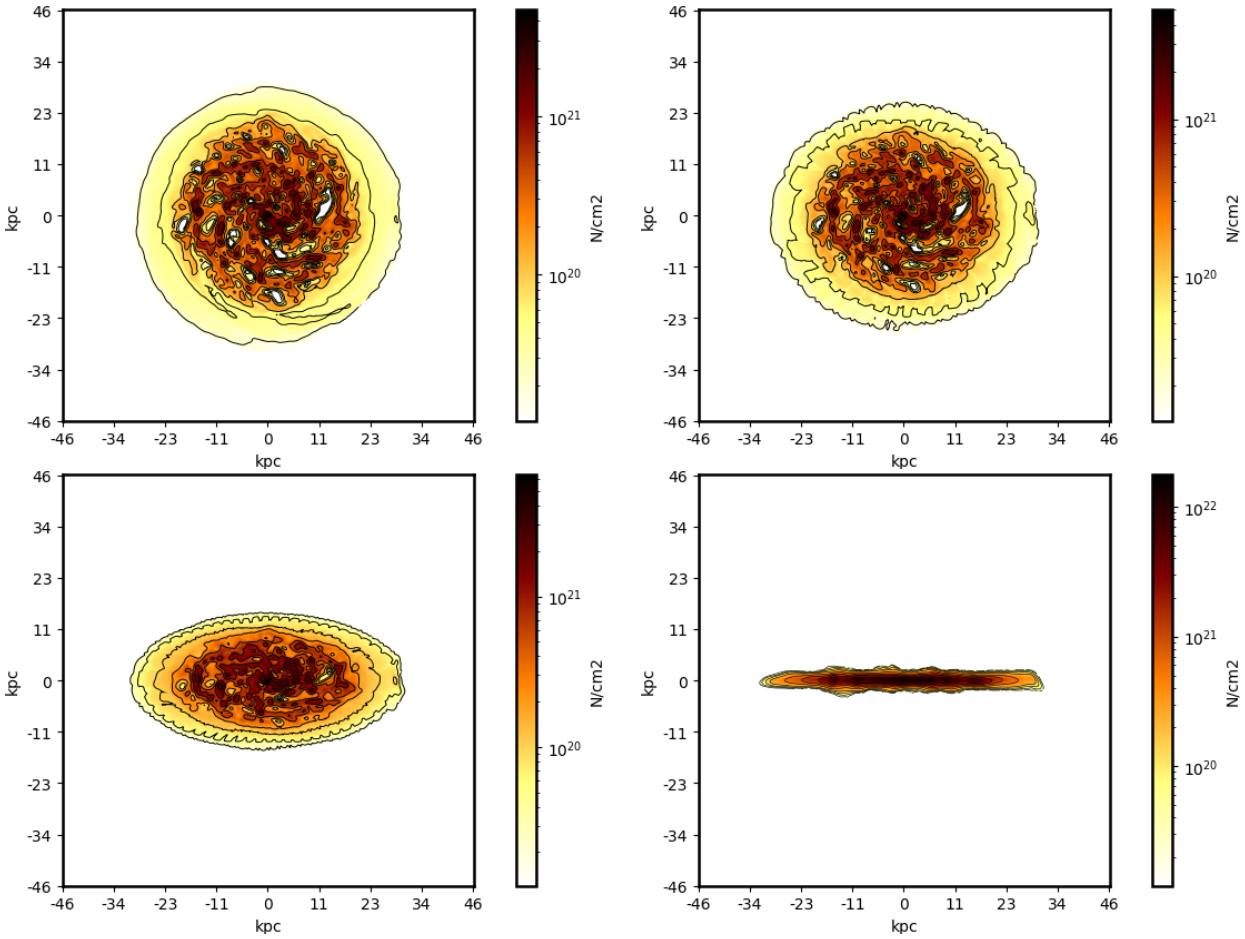


**Figure 3.1:** Column density maps, for four different inclinations: $i = 0°$ (top left), $i = 30°$ (top right) $i = 60°$ (bottom left), and $i = 90°$ (bottom right). The contour levels are 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1.0, 2.0, and 5.0 $\times 10^{21}$ cm$^{-2}$.

determined from the simulation output is $M_{HI} = 3.4 \times 10^9\,\mathrm{M}_\odot$, which means our mock-observation is self consistent. The slightly lower value may be attributed to zones with lower column density than the minimum detectable value (stated above), or may just be due to the effect of noise. Furthermore, since the hydrogen gas mass of most galaxies of this type represent roughly $5-10\,\%$ of total stellar mass, we can say that this galaxy will have a total stellar mass of around some $10^{10}\,\mathrm{M}_\odot$, even without directly observing stellar emission in the optical spectrum. This is also consistent with the value for total stellar mass that was fixed in the initial conditions of the simulation to $M_S = 4.7 \times 10^{10}\,\mathrm{M}_\odot$.

### 3.1.2 HI holes in emission maps

By looking at the column density maps of Figure 3.1, particularly at the face-on one, we can clearly identify holes in the HI distribution with different spatial extent. Similar holes in HI maps were studied by Boomsma et al. (2008), and a possible explanation of them was found by comparing these maps with observation of other parts of the electromagnetic spectrum, like from HII (ionized hydrogen) regions and X-ray emission. They found a correlation between some small ($< 200-300\,\mathrm{pc}$) HI holes and regions with strong HII and X-Ray emission. The larger ones ($> 800\,\mathrm{pc}$) instead show at their interior no bright optical/UV emission, no radio continuum and no far infrared emission. This suggest that these holes are really devoid of gas, dust and star formation. Boomsma et al. (2008) concluded that in young, small holes, at the beginning of their formation, we should find HII regions, created by a high concentration of short-lived blue stars, that during their life cycles emit large amount of UV light capable to ionize the surrounding gas. This means that these would be regions with a high level of star formation rate. These large stars are very short-lived, and so will quickly explode as supernovae. Older, larger holes then represent regions where star formation has already stopped, and where multiple large supernova explosions already occurred, resulting in strong stellar winds that wipe out almost all the dust and gas in these regions. Furthermore, these expanding shells are expected to compress the surrounding ISM, enhancing gravitational collapse and triggering new star formation around the edges of these regions. It would be interesting to also study our simulated galaxy in other regions of the electromagnetic spectrum to check if these scenario is supported by the numerical model. However, since

this thesis focuses on the observation of HI emission, we defer this analysis to future work.

## 3.2 Position-velocity plots

In this section we will discuss position-velocity plots of the observed galaxy. These plots show HI emission flux density as a function of a spatial coordinate $x$, in our case the one parallel to the galaxy's major axis, and of line-of-sight velocity. This means that we must use a fixed value for the $y$ coordinate. This is equivalent to the result of optical long-slit spectrographs, that register flux density and relative line of sight velocity along a one-dimensional line. In figure 3.2 we can see some examples of these plots. As was already mentioned in section 1.3.2, by using this method we are technically throwing away information, since we have access to the full two-dimensional data set. Nevertheless, these plots can still prove useful in showing some interesting features, some of which will be discussed in the following sections.

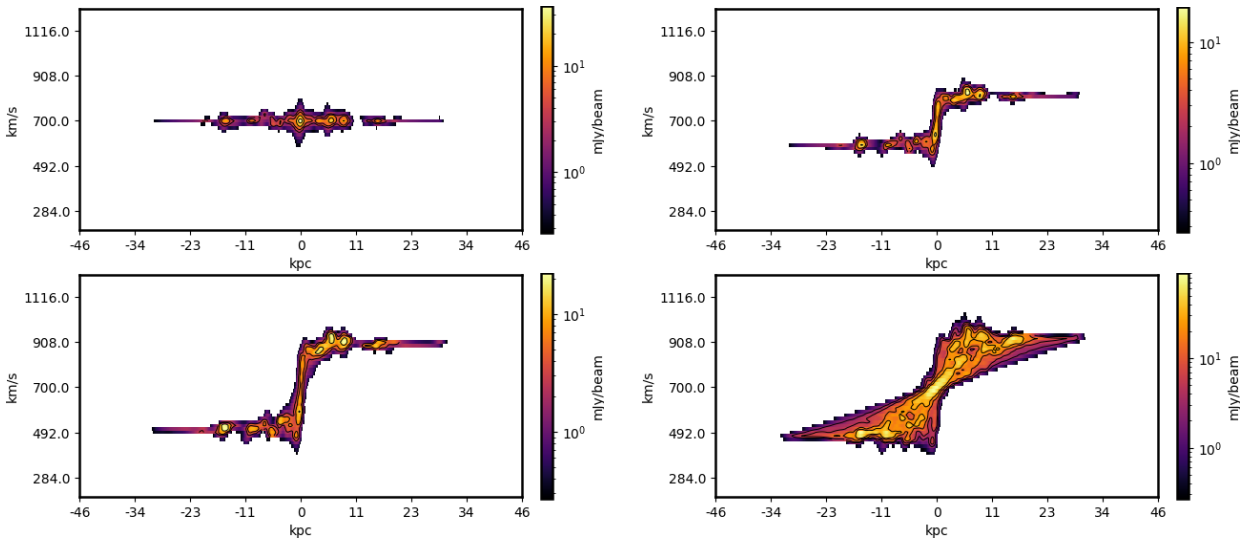


**Figure 3.2:** Mock position-velocity plots, for inclinations $i = 0°$ (top left), $i = 30°$ (top right) $i = 60°$ (bottom left), and $i = 90°$ (bottom right). The $y$-axis shows heliocentric velocity. The center of the galaxy moves away from the observer at 700 km/s due to Universe expansion as per Hubble's law.

### 3.2.1 Galactic fountain flow

Since we are only observing line of sight velocities through the use of Doppler shift, we note that, on average, the observed speed of the gas decreases at lower inclinations $i$ of the galaxy. This is because in these projections more and more of the rotation velocity of the gas will lie in the plane of the sky, and thus will not cause any Doppler shift of the 21 cm line. Because of this effect, we should expect the face-on plot at $i = 0°$ to be a straight line lying at the base velocity value of 700 km/s. At zeroth-order approximation this may appear true, but we can clearly notice some non-negligible deviations from this trend, with some gas near the center of the galaxy displaying velocities of up to 100 km/s. This is something totally different from the rotation velocity around the galactic center, as it does not lie in the plane of the disc, but is instead perpendicular to it. One possibility is that we are witnessing the effect of a galactic fountain flow (Bregman, 1980). In their article Marinacci et al. (2019) show that our simulated galaxy establishes a gas circulation with its surroundings. In this process a non-negligible quantity of gas is ejected from the galactic plane by stellar winds and supernova explosions, like the ones that form the holes in the HI map discussed in section 3.1.2. The ejected gas can later fall back onto the galactic disk to fuel late star formation, and this has major implications for the evolution of the galaxy. In our case this phenomenon is not observed as clearly as seen in the Marinacci et al. (2019) analysis, but that is to be expected since the stellar feedback that eject material from the disk also heat up the involved gas (up to $\sim 10^6$ K). Therefore, most of the hydrogen will be in an ionized state and will not be emitting in the 21 cm line. We can still speculate that the velocities observed in our HI map come from the gas involved in this galactic fountain flow that is both not hot enough to be found in an ionized state (perhaps because it is farther away from the supernova-rich regions), and not dense enough to be in a molecular state.

### 3.2.2 Edge-on position-velocity plots at different heights

Focusing now on the velocity plot of the edge-on projection we expect so see a figure that closely resembles that of the velocity curves discussed in section 1.3.3. Indeed that is the leading trend, with an inner disk dominated by differential rotation with velocity

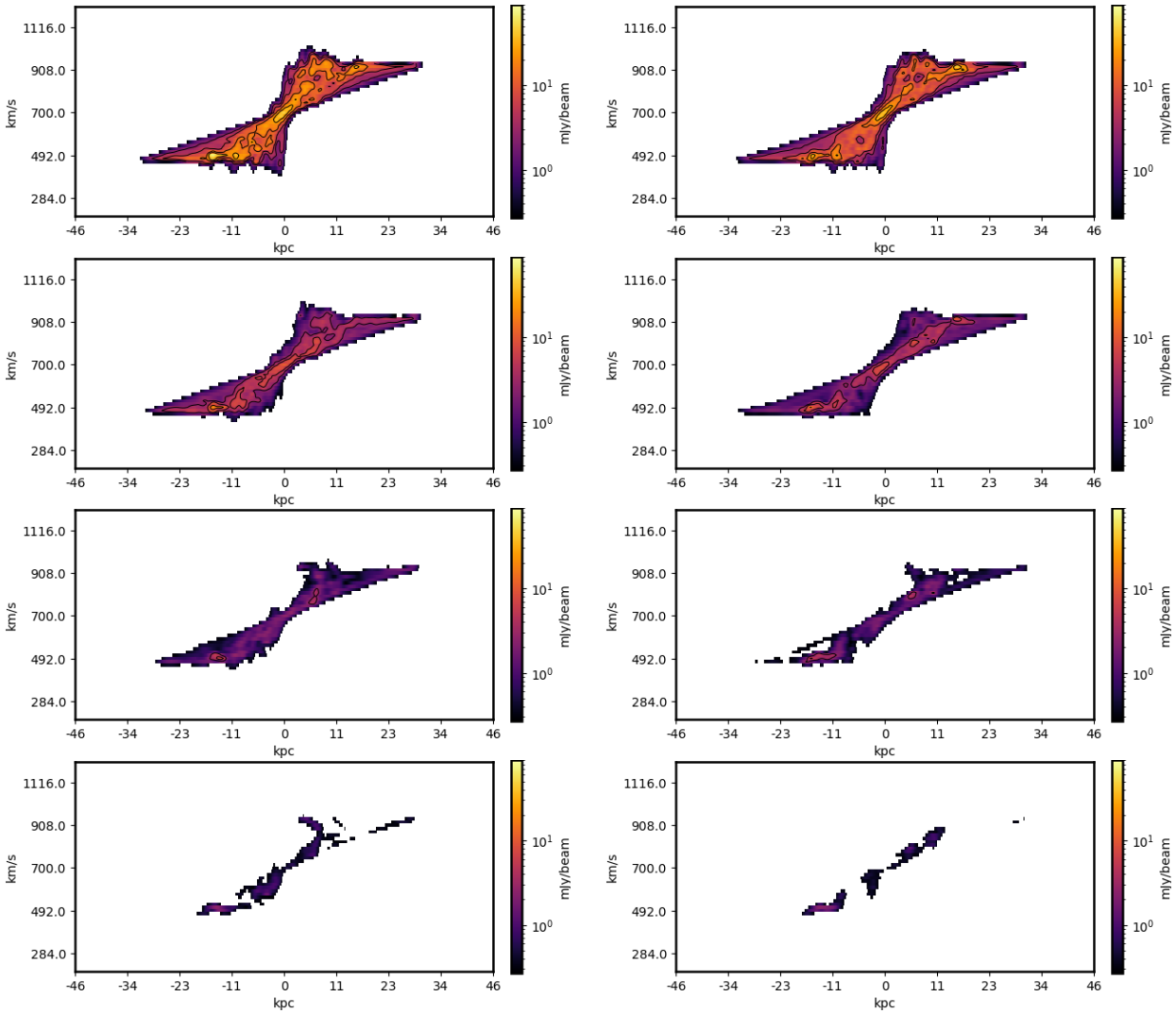


**Figure 3.3:** Position-velocity plots for edge-on galaxy projection at various distances from the galactic plane. From the top, respectively at 0.6, 1.2, 1.8 and 2.4 kpc. On the left column maps are shown for the region above the galactic plane, on the right column for the region below.

increasing with distance from the center, followed by a mostly flat curve in the external region. Beyond this though, we can see that rotation velocity is characterized by an inner peak, as also seen in other HI observations (see, e.g. Oosterloo et al., 2007). This feature is not very symmetric with respect to the center and is in fact more pronounced for positive values of the x coordinate, meaning it might interest a side of the galaxy more than the other. Interestingly, if we take a look at the other plots in Figure 3.3 we can see that above and below the plane of the disk the shape of the diagrams changes from that of a typical rotating disk to that of solid body rotation. In fact, we can clearly see, especially when moving further away from the mid-plane of the disk, that the curve seems to follow a more linear trend, instead of the usual differential to flat trend. This indicates a slower rotation of the gas present in the extra-planar region with respect to that in the disk.

## 3.3 Moment maps

We now briefly discuss two-dimensional mean velocity fields, also known as 1$^{st}$ moment map, and the respective velocity dispersion maps, or 2$^{nd}$ moment maps (Fig. 3.4 and 3.5). These moment maps are generated by using as a weigh for the line of sight velocity/line of sight velocity dispersion of every location the observed HI flux density.

In the face-on projection ($i = 0°$), we can clearly see that mean velocity is consistent with to zero across the galactic plane. Still, the dispersion value (and the mean velocity as well) is somewhat higher in the central regions and along active star formation zones, and that is indicative of gas motion perpendicular to the plane, as we discussed in section 3.2.1.

In the edge-on projection ($i = 90°$), the 1$^{st}$ moment map shows, as expected, rising velocity from the center and then a plateau. A closed contour signifying a velocity peak is observed only in the positive horizontal semi-axis, around x = 5 kpc (the same can be said for $i = 30°$ and 60°). The 2$^{nd}$ moment map also has a peak at the same distance. This location corresponds to the region where we have observed the high velocity, non-symmetrical structure of section 3.2.1.

Lastly, looking at the maps with inclinations $i = 30°$ and 60°, we can clearly see the

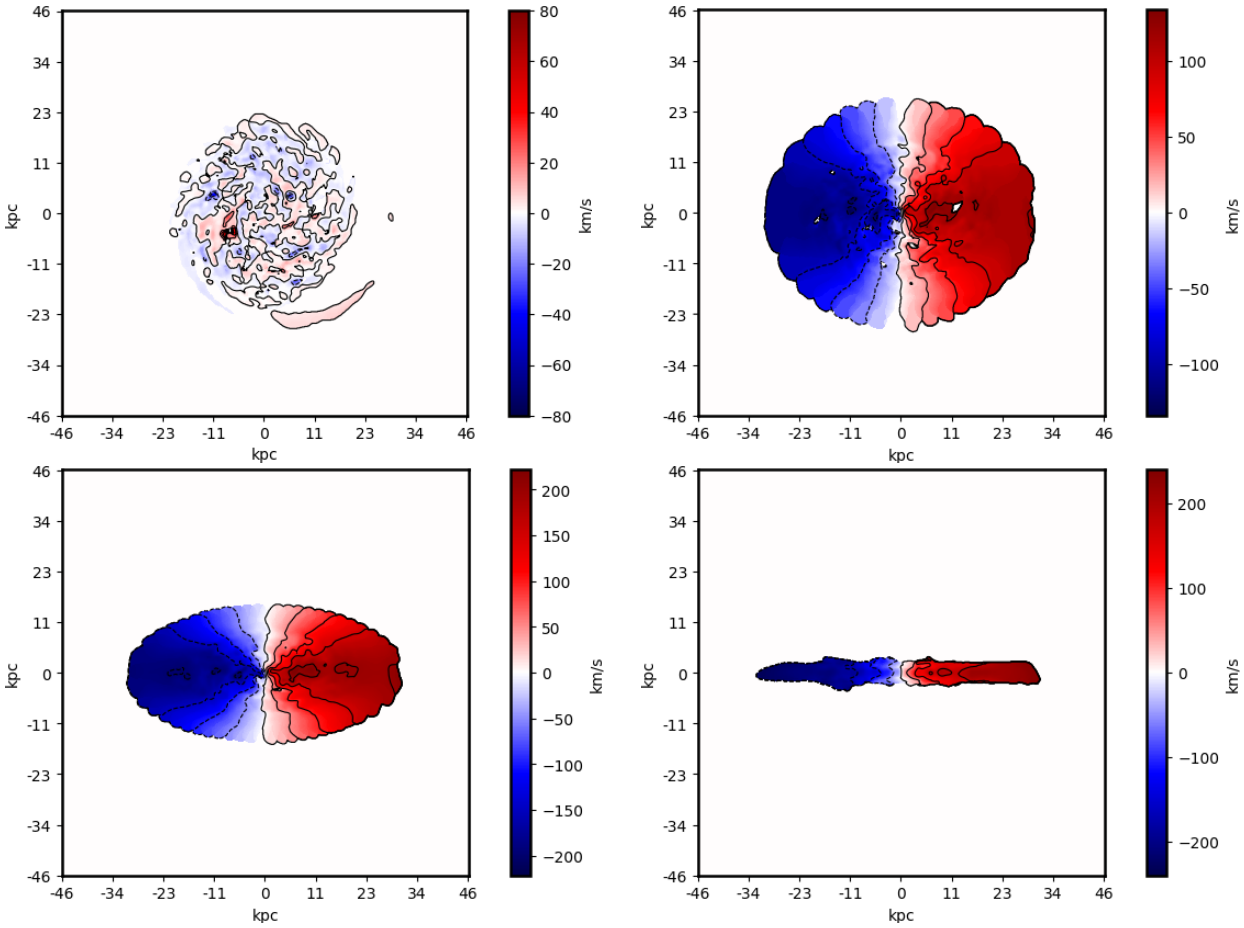


**Figure 3.4:** 1st moment (mean velocities) maps for various galaxy inclinations. The inclination angles are $i = 0°$ (top left), 30° (top right), 60° (bottom left) and 90° (bottom right).

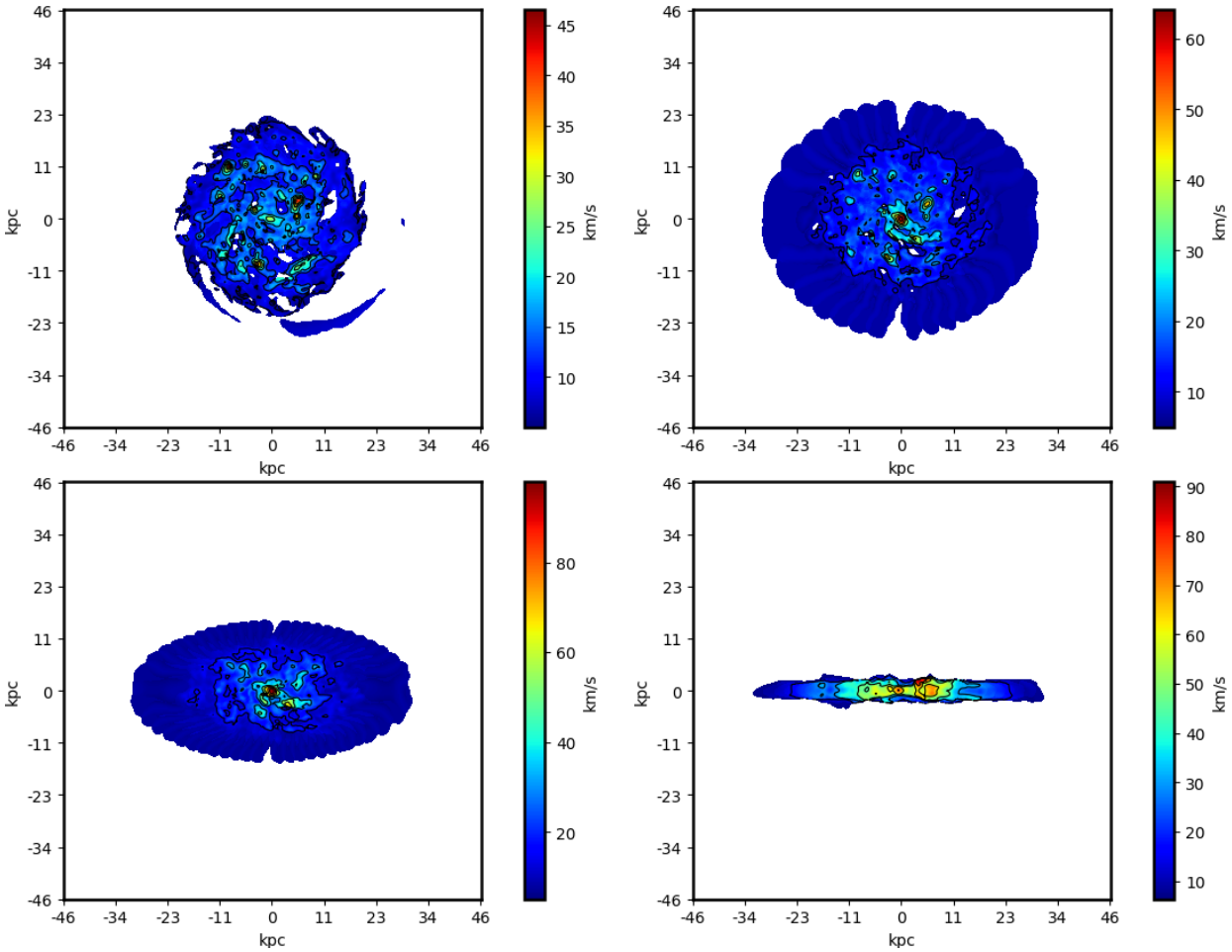


**Figure 3.5:** 2nd moment (velocity dispersion) maps for various galaxy inclinations. The inclination angles are $i = 0°$ (top left), 30° (top right), 60° (bottom left) and 90° (bottom right).

same patterns observed by Bosma (1978) and discussed in chapter 1.3. Near the center, parallel and almost vertical lines can be observed, indicative of solid body rotation, but then the 'spider' pattern emerges, indicative instead of a flat rotation curve. A few closed contours are present where peak velocity values are reached. However, a single marked peak followed by a steep decrease, which would be clearly seen if the mass distribution of the galaxy followed an exponential disc model, is never observed. As previously discussed, these features can only be explained through the presence of large quantities of non-luminous dark matter in the global mass distribution of the galaxy. Of course, this behavior of the simulation is to be expected since the initial conditions that were used to evolve the simulated galaxy do indeed account for the presence of a fairly massive ($\sim 10^{12}\,M_\odot$) dark matter halo.

## 3.4 Tilted-rings model fits

Finally, to show the full potential of the mock observation approach, we will briefly go over a more advanced method used to analyse observation data:the fitting of our synthetic datacube with a tilted-ring model (see section 1.3.2). To do so we make use of the 3D-Barolo software (Di Teodoro et al., 2015). The main feature of Barolo is to fit data cube observations, starting from the model of a rotating gaseous disc, and to compare it with real data (which, in our case, is the data from our mock observation with MARTINI). The model disc is made up by a number of concentric rings with non-zero thickness. The emission from gas in each ring is generated in a 6D domain (three dimensions for the spatial location and three for the components of velocity) and these rings are then projected into the 3D data cube (i.e. two spatial and one spectral dimensions). At each step, if the model is good enough, the algorithm moves to the following ring, otherwise it updates the disc parameters until the best match between model and observations is found.

Each ring is described by the following parameters:

- Spatial coordinates of the center $(x_0, y_0)$.
- Systemic velocity $V_{sys}$.

- Inclination angle $i$ with respect to the observer (90° for edge-on).
- Position angle $\phi$ of the major axis on the receding half of the galaxy, taken anti-clockwise from the north direction on the sky.
- Rotational velocity $V_{rot}$.
- Radial velocity $V_{rad}$.
- Velocity dispersion $\sigma_{gas}$.
- Face-on gas column density $\Sigma$.

A schematic flowchart of the main fitting algorithm is shown in Figure 3.6.

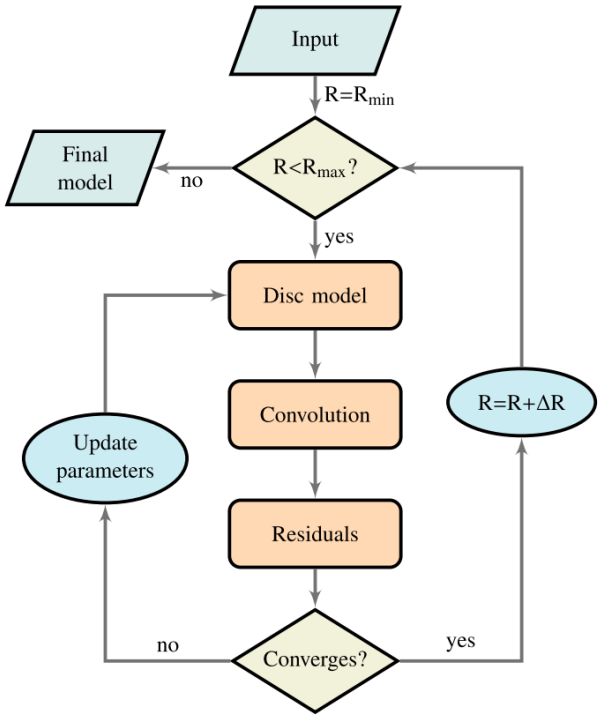


**Figure 3.6:** Flowchart of the 3D-Barolo main algorithm. For each ring R, the algorithm builds a 3D model, convolves it with the observational beam and compares it with the data. If no convergence has been achieved, Barolo updates the parameters and starts over. When the algorithm converges, it moves to the next ring (Di Teodoro et al., 2015).

All of these quantities are allowed to vary from ring to ring. In our case, we made use of only a fraction of Barolo potential, but we will still present its output about the data cube regarding inclination $i = 60°$, reserving a more thorough analysis for possible future developments of our work. For our fitting process, we fixed the values of some of the parameters:

- Coordinates of center of the galaxy $(x_0, y_0)$ were set to (383.5, 383.5) pixels, the known center of our data cube.
- Systemic $V_{sys}$ was set to 700 km/s.
- Radial velocity $V_{rad}$ was set to 0.

Then we allowed the algorithm to run, fitting the remaining free parameters for each ring. The software divided the galaxy into 35 rings, each spaced 20 arcseconds from one another (corresponding to approximately 1 kpc).

In figures 3.7 and 3.8 we can see the comparison between position-velocity plots, flux density maps, velocity fields and dispersion fields regarding our data cube, and the ones generated by the tilted ring model. We notice that this model provides a good fit of our synthetic data. This type of maps and their most important features were already described in sections 3.1, 3.2 and 3.3, except for the position velocity plot along the minor axis in figure 3.7. This plot shows velocities for a vertical cut of the galaxy along the y axis, while keeping $x = 0$. The distribution of the HI emission is rather concentrated around the galaxy's systemic velocity, except for a peak in the central region, extending to a large range of line-of-sight velocities velocities, corresponding to the high velocity dispersion we find in the center of the galaxy, as can bee seen in 2$^{nd}$ moment maps (e.g. in Figures 3.5 and 3.8).

Figure 3.9, instead, shows a channel-by-channel comparison of flux density between the datacube and the model cube. These panels give us similar, but more detailed, information to that of 1$^{st}$ moment maps (e.g. 3.4), by decomposing it along some representative velocity channels. The figure makes it readily apparent that the motion of the gas in the galactic disk is dominated by rotation.

Finally, Figure 3.10 shows the tilted-ring model fit parameters for each ring. We can notice that the model correctly estimated the galaxy's inclination to be around 60°

(center-left panel). Furthermore in the top-right panel we see that face-on column density follows an exponential mass distribution, as discussed in section 1.3.1. The rotation curve (top-left panel) does not show a peaked trend, as should be the case if the only mass components were those described by the exponential distribution, but is instead flat, out to the most external region occupied by the gas disc. Of course this type of rotation curve is the expected, since our simulated galaxy was initialized with an additional matter component representing dark matter. As real disc galaxies also have this type of rotation curve (see e.g. Figure 1.6), its flatness is indicative that they too must include such an additional non-luminous matter component in their mass distribution.

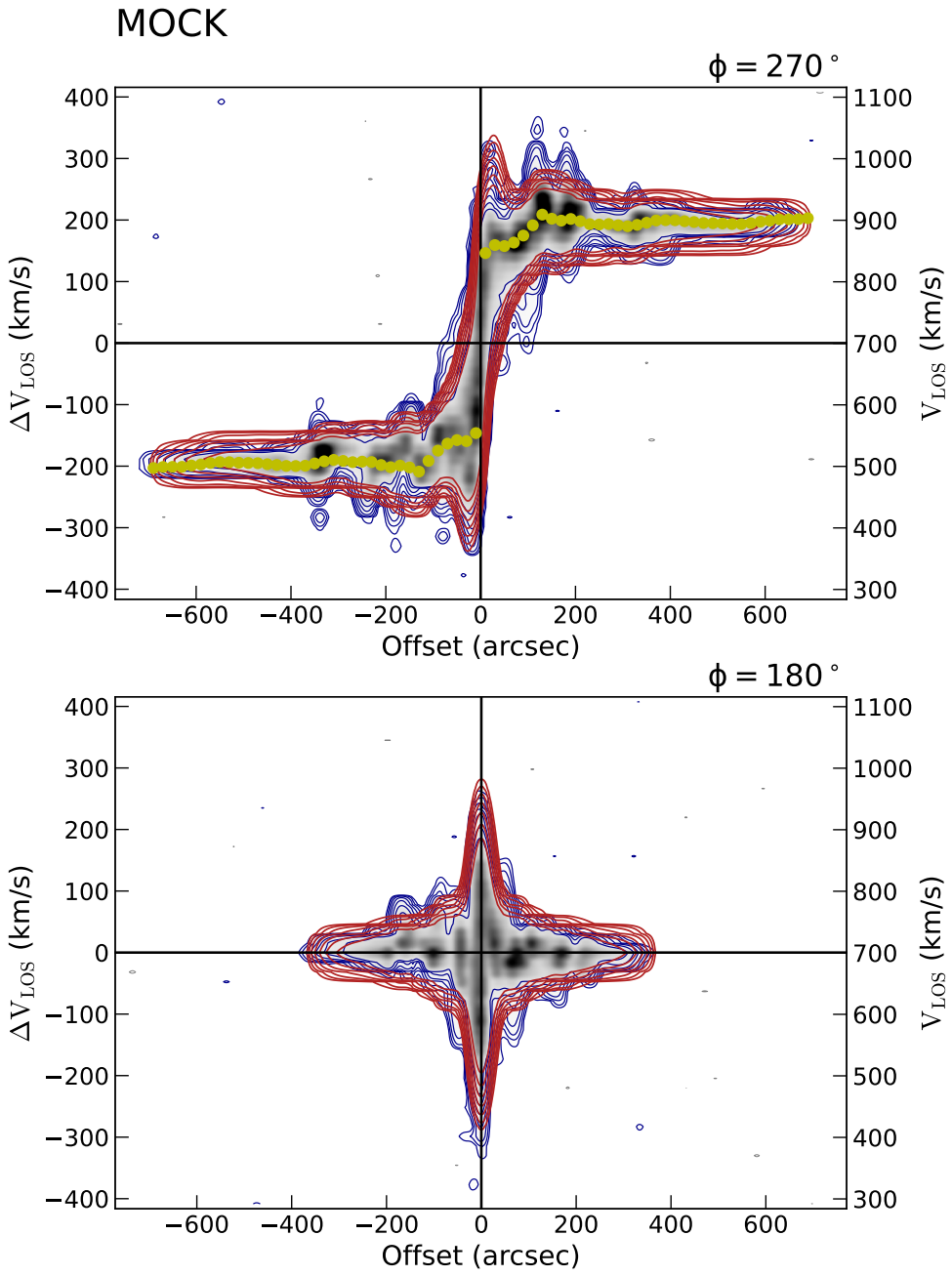


**Figure 3.7:** Position-velocity plots along major (top) and minor (bottom) axes. Data from our data cube is represented in blue, the results from the tilted-ring model fit in red, and the derived rotation curve in yellow.

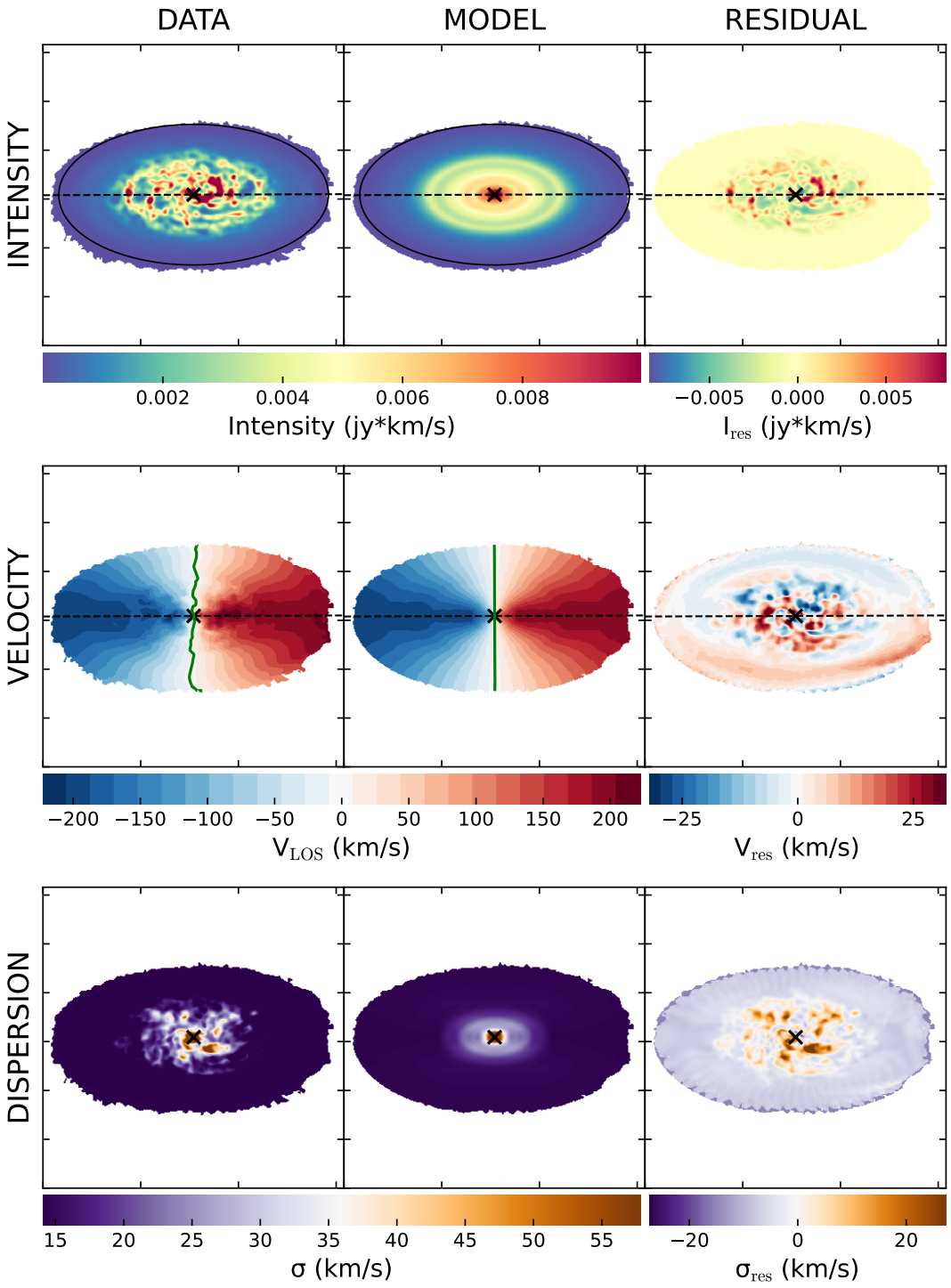


**Figure 3.8:** Flux density maps ($0^{\text{th}}$ moment), velocity fields ($1^{\text{st}}$ moment) and velocity dispersion fields ($2^{\text{nd}}$ moment), of the synthetic datacube ($1^{\text{st}}$ column), the model cube ($2^{\text{nd}}$ column) and the residual function to be minimized in the fit ($3^{\text{rd}}$ column), which in our case is the absolute value of the difference between the values of the model cube and the data cube: $|model - data|$.

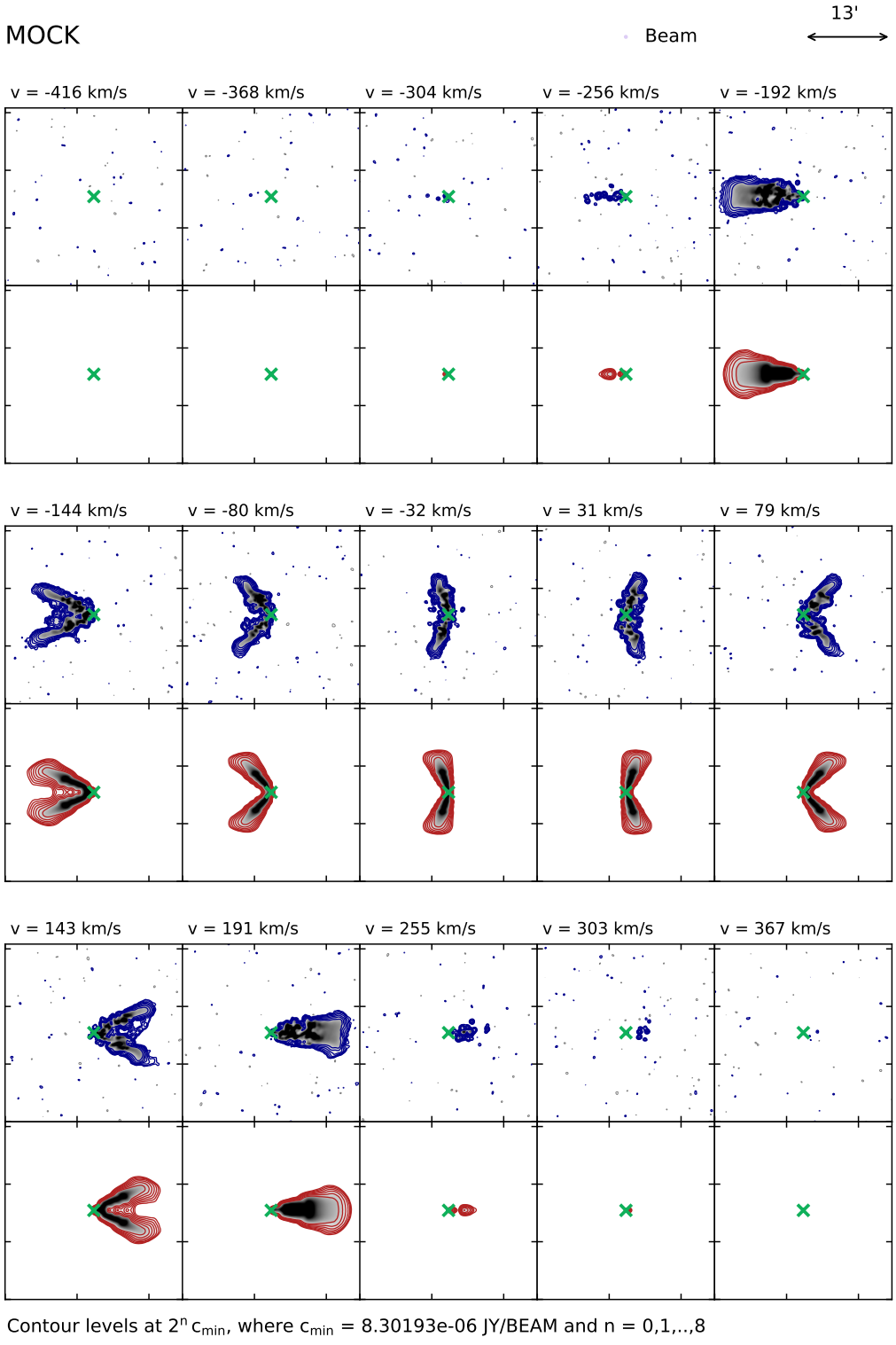


**Figure 3.9:** Comparison between data (blue) and model (red) cubes of flux density over some representative velocity channels. The velocity difference between consecutive panels is 48 km/s, corresponding to 3 velocity channels of the synthetic datacube.

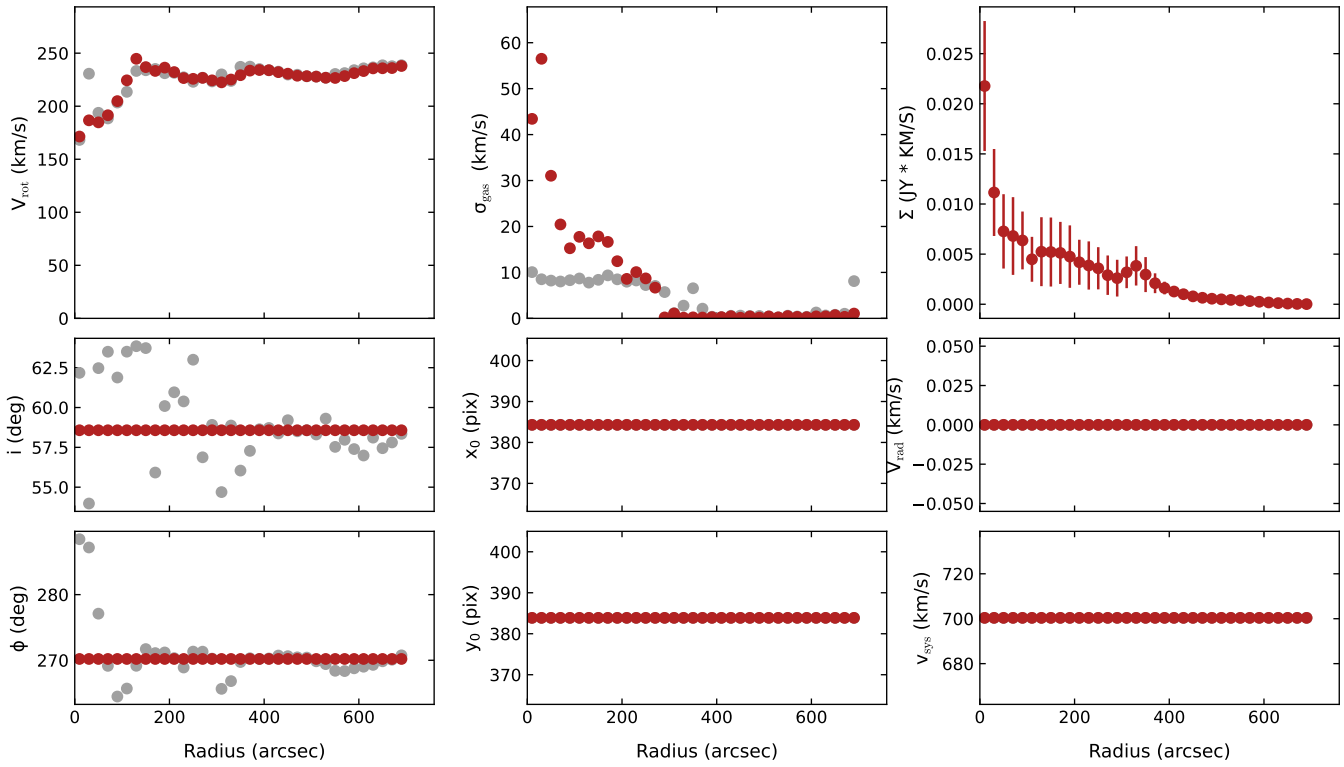


**Figure 3.10:** Tilted ring fit parameters derived from Barolo for a galaxy inclination of $i = 60°$. From left to right, and then from top to bottom the graphs plot: (i) rotational velocity $V_{rot}$, (ii) velocity dispersion $\sigma_{gas}$, (iii) face-on gas column density $\Sigma$, (iv) inclination $i$, (v) horizontal coordinate of the center $x_0$, (vi) radial velocity $V_{rad}$, (vii) position angle $\phi$ of the major axis, (viii) vertical coordinate of the center $y_0$, (ix) systemic velocity $V_{sys}$. We can see that $(x_0, y_0)$, $V_{rad}$ and $V_{sys}$ are fixed in value. For the others, the grey dots represent the first fit with all parameters kept free. The red dots or lines represent the second fit of $V_{rot}$ and $\sigma_{gas}$ after the regularization of the other parameters.

# Summary and conclusions

The purpose of this thesis work is to show the potential of using 21 cm line HI mock observations of simulated galaxies and the crucial role that they can play in testing and understanding both simulation models and real observations. In this work we provide practical methods for the creation of such observations and discuss some key results that can be obtained from their analysis. First, we started by giving a brief overview of the quantum theory behind this fine energy splitting in the hydrogen atom, and then we went on to discuss its use in modern astrophysics. Specifically, we discussed the properties of 21 cm line emission sources and how we can analyse such observations to obtain an estimate of crucial properties of such gas component. In particular, we derived the most important expressions relating the properties of the HI emission and that of the emitting gas. Finally, we discussed how we can get crucial information on the mass distribution on galactic scales and on galactic kinematics through the analysis of their HI emission. After giving an overview of the most important classical results on the topic by discussing findings from the seminal works of Rubin et al. (1980) and Bosma (1978), we applied the same observational methods to a simulated galaxy, for which the baseline properties are known. This can be in principle used both to help interpreting observations and to test to what degree current galaxy formation simulations are an accurate model of real galaxies. We managed to do so through to use of the MARTINI package, which allowed us to create synthetic resolved HI line observations of our simulated galaxy. We then analysed the data obtained through our mock observation as if we were looking at measurements from a real galaxy, mimicking actual observations. Our results were in line with what we would expect for a star-forming galaxy similar to the Milky Way confirming the validity of this approach as an important tool to analyse numerical simulations of galaxy

formation, to meaningfully compare their results to real observations and to advance our knowledge on the formation and evolution of galaxies in general.

# Bibliography


Begeman, K. G. (1989). HI rotation curves of spiral galaxies. I. NGC 3198. *Astronomy and Astrophysics*, *223*, 47–60.

Boomsma, R., Oosterloo, T. A., Fraternali, F., Hulst, J. M. v. d., & Sancisi, R. (2008). HI holes and high-velocity clouds in the spiral galaxy NGC 6946. *Astronomy & Astrophysics*, *490*(2), 555–570. https://doi.org/10.1051/0004-6361:200810120

Bosma, A. (1978). *The distribution and kinematics of neutral hydrogen in spiral galaxies of various morphological types* (Doctoral dissertation). University of Groningen, Netherlands.

Bovy, J. (2022). *Dynamics and Astrophysics of Galaxies*. Princeton University Press, Princeton, NJ (in preparation). https://galaxiesbook.org/

Bregman, J. N. (1980). The galactic fountain of high-velocity clouds. *The Astrophysical Journal*, *236*, 577–591. https://doi.org/10.1086/157776

Di Teodoro, E. M., & Fraternali, F. (2015). 3D barolo: A new 3D algorithm to derive rotation curves of galaxies. *Monthly Notices of the Royal Astronomical Society*, *451*(3), 3021–3033. https://doi.org/10.1093/mnras/stv1213

Griffiths, D. J. (1982). Hyperfine splitting in the ground state of hydrogen. *American Journal of Physics*, *50*(8), 698–703. https://doi.org/10.1119/1.12733

Karttunen, H., Kröger, P., Oja, H., Poutanen, M., & Donner, K. J. (Eds.). (2017). *Fundamental Astronomy*. Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-53045-0

Kent, S. M. (1986). Dark matter in spiral galaxies. I - Galaxies with optical rotation curves. *The Astronomical Journal*, *91*, 1301. https://doi.org/10.1086/114106

Marinacci, F., Sales, L. V., Vogelsberger, M., Torrey, P., & Springel, V. (2019). Simulating the interstellar medium and stellar feedback on a moving mesh: Implementation and isolated galaxies. *Monthly Notices of the Royal Astronomical Society*, *489*(3), 4233–4260. https://doi.org/10.1093/mnras/stz2391

Nelson, D., Springel, V., Pillepich, A., Rodriguez-Gomez, V., Torrey, P., Genel, S., Vogelsberger, M., Pakmor, R., Marinacci, F., Weinberger, R., Kelley, L., Lovell, M., Diemer, B., & Hernquist, L. (2019). The IllustrisTNG simulations: public data release. *Computational Astrophysics and Cosmology*, *6*(1), 2. https://doi.org/10.1186/s40668-019-0028-x

Oman, K. A. (2019). MARTINI: Mock spatially resolved spectral line observations of simulated galaxies. https://ui.adsabs.harvard.edu/abs/2019ascl.soft11005O

Oman, K. A., Marasco, A., Navarro, J. F., Frenk, C. S., Schaye, J., & Benıtez-Llambay, A. (2019). Non-circular motions and the diversity of dwarf galaxy rotation curves. *Monthly Notices of the Royal Astronomical Society*, *482*(1), 821–847. https://doi.org/10.1093/mnras/sty2687

Oosterloo, T., Fraternali, F., & Sancisi, R. (2007). The cold gaseous halo of NGC 891. *The Astronomical Journal*, *134*(3), 1019–1036. https://doi.org/10.1086/520332

Rhee, J., Lah, P., Briggs, F. H., Chengalur, J. N., Colless, M., Willner, S. P., Ashby, M. L. N., & Le Fèvre, O. (2018). Neutral hydrogen (H I) gas content of galaxies at z  0.32. *Monthly Notices of the Royal Astronomical Society*, *473*, 1879–1894. https://doi.org/10.1093/mnras/stx2461

Rod Nave, C. (2017). Hyperphysics: The Hydrogen 21-cm Line. http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/h21.html

Rubin, V. C., Thonnard, N., & Ford, W. K., Jr. (1980). Rotational properties of 21 SC galaxies with a large range of luminosities and radii, from NGC 4605 /R = 4kpc/ to UGC 2885 /R = 122 kpc/. *The Astrophysical Journal*, *238*, 471. https://doi.org/10.1086/158003

Weinberger, R., Springel, V., & Pakmor, R. (2020). The AREPO Public Code Release. *The Astrophysical Journal Supplement*, *248*(2), 32. https://doi.org/10.3847/1538-4365/ab908c

Westmeier, T. (2021). Tools - Useful equations for radio astronomy. https://www.atnf.csiro.au/people/Tobias.Westmeier/tools_hihelpers.php

Wilson, T. L., Rohlfs, K., & Hüttemeister, S. (2013). *Tools of Radio Astronomy*. Springer Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39950-3