Alma Mater Studiorum - Università di Bologna

School of Science

Department of Physics and Astronomy

Master Degree Programme in Astrophysics and Cosmology

A synthetic view of HI: mock observations of simulated star-forming galaxies

Presentata da:
Yerbolat Saparzhanov

Supervisor:
Prof. Federico Marinacci

Anno Accademico 2024/2025

July 2025

Abstract

This thesis presents an analysis of mock HI observations of Milky Way-like galaxies generated from galaxy formation simulations. In particular, the simulations were performed with the state-of-the-art interstellar medium and stellar feedback SMUGGLE model within the moving mesh code AREPO. Mock 21 cm HI emission cubes based on these simulations were built using the MARTINI software, allowing us to investigate the morphology and kinematics of neutral hydrogen across a range of galaxy inclinations. To analyze these mock data cubes, we applied a series of data processing techniques to extract key physical properties of the gas, including moment maps, velocity channel maps, position–velocity diagrams, column density maps, rotation curves, and total HI mass. From these diagnostics, we inferred the structural and kinematic characteristics of the simulated galaxies.

We also applied a simple and observationally motivated decomposition method to separate the cold disc from gas located outside the disc—commonly referred to as extra-planar gas (EPG)— that shows an anomalous kinematics relative to the gas in the disc. In particular, observational evidence suggests that approximately 10–15% of a galaxy's HI mass resides in such an extra-planar component (also known as HI halo) and that this gas exhibits a vertical velocity gradient of $\sim -10\ \mathrm{km\,s^{-1}\,kpc^{-1}}$ ("lagging halo"). In our analysis, we characterized the kinematics of both the cold disc and the extra-planar gas, identifying their distinct dynamical signatures. We compared our results to observed systems, such as NGC891 or the HALOGAS survey of nearby galaxies, to assess the realism of the simulated galaxy structures and support the presence of lagging halo gas, finding a qualitative agreement with observations. Our findings underscore the value of synthetic observations as a bridge between simulations and real data, particularly in the context of extra-planar gas dynamics in star-forming galaxies. By comparing simulated and observed HI features, we were able to assess the extent to which the simulations reproduce the complex structure and kinematics of neutral hydrogen in real galaxies.

Contents

1

Introduction

Understanding how structures populating our Universe, and in particular galaxies, form and evolve is one of the central goals of modern astrophysics. Galaxies are shaped by a variety of complex physical processes involving interactions between baryonic matter, dark matter (DM), and the intergalactic medium (IGM; Cimatti et al., 2019). Physical processes, such as star formation, stellar and black hole feedback, and large-scale matter accretion, also play a critical role. To investigate these phenomena, researchers integrate observational data with theoretical models that describe the underlying physics of galaxy evolution.

In order to probe the dynamics and gain insights into the mechanisms driving galaxy evolution, astronomers rely on a variety of tracers that reveal the physical and dynamical properties of galaxies. One of the most important of these tracers is neutral atomic hydrogen (HI), which provides crucial information about the cold gas reservoir that fuels star formation. The 21-cm spectral line emitted by HI lies in the radio regime and can penetrate both the Earth's atmosphere and interstellar dust clouds that are opaque at optical wavelengths. This makes HI a powerful tool for probing the structure of galaxies. Indeed, the spatial distribution and kinematics of HI offer crucial insights into galactic structure, particularly in the outskirts of galaxies where starlight is faint or absent. Observations of the 21-cm hyperfine transition enable astronomers to trace cold atomic gas across cosmic time and infer the presence of dark matter through extended rotation curves (e.g., Begeman, 1989).

In Milky Way (MW)-like spiral galaxies, HI is typically found in a disc and in a diffuse, extended gaseous halo. This makes HI an invaluable probe not only for galactic dynamics but also to unveil processes at the interface between galaxies and their environments. Gaseous HI halos have been detected in several galaxies, with roughly 15% of the total HI content residing in these regions. This has been supported by observations of galaxies such as NGC891 (Oosterloo et al., 2007), NGC6946 (Boomsma et al., 2004; Boomsma, 2007), NGC5775 (Lee et al., 2001), NGC2403 (Schaap et al., 2000), and more recently by targeted surveys such as HALOGAS (Marasco et al., 2019). The HI in the circumgalactic medium (CGM) shows signs of coherent rotation, though at velocities lower than those of the galactic disc (Swaters et al., 1997; Fraternali et al., 2002). Gaseous halos are also detectable in Hα emission, displaying similar kinematics to the neutral hydrogen component (Hoopes et al., 1999; Rossa and Dettmar, 2003).

The origin of such gaseous halos remains an open question. One of the leading theories is the galactic fountain model (Shapiro and Field, 1976), which posits that stellar winds and supernovae inject hot gas into the halo. As this gas cools and becomes neutral, it falls back into the galactic disc and may participate in subsequent star formation (Bregman, 1980). Observational support for this model includes Hα emission in NGC6946 (Boomsma, 2007), as well as correlations between the star formation rate (SFR) with X-ray and radio continuum emissions from the halo (Tüllmann et al., 2006; Matthews and Wood, 2003).

However, several tensions with the observations challenge the galactic fountain scenario. For instance, ionized halos appear only in galaxies with SFRs above a certain threshold (Tüllmann et al., 2006), whereas neutral halos have been observed in low-SFR systems like the low surface brightness galaxy UGC7321 (Matthews and Wood, 2003). Another factor that is controversial in the galactic fountain model comes from the studied kinematics of diffuse ionized gas that cannot be described by the simple ballistic models (Collins et al., 2002; Heald et al., 2006). Also, the kinematics of neutral gas does not follow the "pure" ballistic model (Fraternali and Binney, 2006). These discrepancies have led researchers to propose alternative mechanisms for the formation of the HI halos and its interaction with a pre-exisitng CGM that help explain the observed kinematics (Armillotta et al., 2016; Marinacci et al., 2011; Fraternali and Binney, 2008).

Finally, a deeper understanding of CGM is essential for advancing our knowledge of galaxy evolution. The CGM serves as the interface through which galaxies exchange material with the surrounding environment, regulate the inflow and outflow of gas, and mediate the galactic gas cycle. Observing the CGM, particularly their interactions with the gas expelled from galaxies, provides valuable insight into the cosmic baryon cycle.

Interpreting these observations, especially in the era of next-generation high-sensitivity, high-resolution radio surveys such as the Square Kilometre Array (SKA; Weltman et al., 2020), requires robust theoretical models that account for observational limitations, including instrumental resolution, noise, and projection effects that can obscure or distort the intrinsic properties of gas in and around galaxies.

To address these issues, mock observations offer a powerful approach to bridge the gap between theory and observations. By generating synthetic HI data cubes from hydrodynamical simulations, researchers can perform direct comparisons and refine models accordingly. These mock HI observations are constructed by generating synthetic data cubes from hydrodynamical simulations that model galaxy evolution under controlled physical conditions. These mock data cubes are then processed through virtual instruments that replicate the characteristics of real telescopes, including spatial resolution, noise, and spectral sampling. This enables a direct, one-to-one comparison between simulated galaxies and observed data, allowing researchers to test whether theoretical models can reproduce the observable features seen in real systems (Marasco et al., 2025; Roper et al., 2023; Bilimogga et al., 2022).

In this thesis, we focus on generating and analyzing mock HI observations from hydrodynamical simulations of isolated MW-like galaxies. Our overarching goal is to connect theoretical predictions and observational radio data, offering tools and insights for interpreting current and future HI surveys. The motivation behind this research is twofold. The first objective is to improve our understanding of the physical processes governing gas dynamics in disk galaxies. In addition, we would also like to assess how reliably simulation-based models reproduce the observable properties of the HI gas, thereby using these results to potentially constrain the models adopted for describing the complex physics of galaxy formation and evolution in the simulations. In the context of upcoming deep surveys, the ability to generate and analyze mock observations becomes even more critical. As observational capabilities improve, so must our theoretical tools. This thesis contributes to that effort by ensuring that comparisons between simulations and observations are meaningful, self-consistent, and physically interpretable.

To investigate the morphology and kinematics of the HI gas in (star-forming) disc galaxies, we employed a combination of hydrodynamical simulations and synthetic observation techniques. Specifically, we used the SMUGGLE model (Marinacci et al., 2019), a framework for star formation and stellar feedback, implemented within the moving-mesh hydrodynamic code AREPO (Weinberger et al., 2020). SMUGGLE allows for the modeling of a multiphase ISM in a physically motivated way, and the processes that the model takes into account are essential for realistically capturing the formation and evolution of

cold gas structures in and above galactic discs. The simulations analyzed in this thesis produce realistic MW-like galaxies that explicitly model the disc-halo interface, enabling a detailed study of extra-planar gas (EPG) dynamics. We used snapshots of isolated, MW-like disc galaxies at redshift $z \sim 0$ that show prominent gaseous extra-planar components. From these simulations, we extract the gas properties necessary to construct observational analogues.

To convert simulation outputs into observable quantities, we use the MARTINI (Oman, 2024) software package. MARTINI creates synthetic HI data cubes by projecting simulated gas cells along a given line of sight, applying realistic observational effects such as finite angular and spectral resolution, thermal line broadening, instrumental noise, and beam convolution. This process allows us to emulate the output of actual radio telescopes, making the comparison between theory and observation more robust and meaningful. The resulting mock data cubes are analyzed with the same techniques typically used in observational studies. We measure key kinematic quantities, such as rotation curves, vertical velocity gradients, velocity dispersion profiles, and position–velocity (PV) diagrams. We also separate and measure the EPG parts using disc–halo decomposition methods, which involve removing a fitted cold disc model. Furthermore, we investigate how these properties vary with galaxy inclination.

The remainder of this thesis is organized as follows. In Chapter 2, we describe in detail the simulation used as the basis for our study – including the AREPO code and the SMUGGLE galaxy formation model – and outline their numerical setup. Chapter 3 sheds light on the details of the methodology used to generate mock HI observations and the techniques applied for their analysis. In Chapter 4, we present and analyze our results of the kinematics and morphology of the cold disc and EPG. Finally, Chapter 5 summarizes our findings and discusses future directions for research in this area. As an extra, there are Appendices A include comparisons of different modes of observation and describe some of the methods used to derive physical quantities.

2

Galaxy formation simulation framework

Galaxy formation results from a complex and highly non-linear interplay of physical processes. To accurately model these phenomena, modern simulations require advanced numerical techniques capable of solving the governing equations of gravity, gas dynamics, and magnetohydrodynamics (MHD). Gravitational interactions, in particular, are long-range: every mass element within the simulation volume influences any other. Gas dynamics, an inherently non-linear phenomenon, is typically treated using fluid dynamics, where processes such as shocks, turbulence, and radiative cooling and heating are described by Euler or Navier–Stokes equations. Additionally, galaxy formation encompasses a vast range of spatial and temporal scales, requiring simulations to incorporate multiple layers of physics, including star formation, chemical enrichment, feedback from supernovae (SN) and active galactic nuclei (AGN), and the cosmic expansion.

2.1 AREPO

One of the most sophisticated codes used for galaxy formation simulations is AREPO (Weinberger et al., 2020), a moving-mesh code that efficiently solves the equations of a collisionless particle component (the so-called dark matter) and hydrodynamics, within the framework of ΛCDM cosmology, on a uniformly expanding flat Friedmann-Lemaître-Robertson-Walker (FLRW) spacetime:

$$ds^2 = c^2dt^2 - a^2(t)[dr^2 + r^2d\theta^2 + r^2sin^2\theta d\phi^2] \ , \tag{2.1}$$

where $a(t)$ is the scale factor, c is the speed of light, r, θ and ϕ are the comoving spherical coordinates, and t is a proper time for a comoving observer.

An essential component of cosmological simulations is the accurate modeling of gravity in an expanding universe. In the context of galaxy formation, gravitational interactions are the primary drivers of structure formation on large scales, from the collapse of dark matter halos to the aggregation of baryonic matter into galaxies. To account for the effects of cosmic expansion, the gravitational potential is computed by solving the Poisson equation in comoving coordinates:

$$\nabla^2\Phi = 4\pi G(\rho_{\text{total}} - \rho_{\text{mean}}) \ , \tag{2.2}$$

where ρ_{total} is the local total matter density and ρ_{mean} is the mean cosmic matter density and G is the gravitational constant. This formulation ensures that only density fluctuations relative to the cosmic average contribute to the gravitational potential.

The evolution of the scale factor $a(t)$, which determines the rate of cosmic expansion, is governed by the Friedmann equation:

$$H = H_0 \left[\Omega_0 a^{-3} + (1 - \Omega_0 - \Omega_\Lambda)a^{-2} + \Omega_\Lambda\right]^{1/2} , \tag{2.3}$$

where $H = \dot{a}/a$ is the Hubble parameter, H_0 is its current value, Ω_0 is the density parameter of matter, and Ω_Λ is the density parameter of dark energy. The cosmological concepts are essential for understanding galaxy formation and evolution. Although this thesis focuses on simulations of gas dynamics in isolated galaxies, it incorporates key ideas from cosmological modeling, such as the Vlasov equation (see below), to accurately describe the collisionless component, although in a non-expanding background.

2.2 Collisionless component

The collisionless component in AREPO is described by the Vlasov equation:

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \vec{x}}\frac{\vec{v}}{a} + \frac{\partial f}{\partial \vec{v}}\left(\frac{\nabla\Phi}{a^2} + \frac{\dot{a}}{a}\vec{v}\right) = 0 \ . \tag{2.4}$$

From this equation it is possible to derive the equation of motion for the collisionless components of the simulations, that is, dark matter (DM) and stellar particles. These follow simple equations of motion, with their accelerations determined by the gravitational potential:

$$\dot{x} = v/a \ , \tag{2.5}$$

$$\dot{v} = -\frac{\nabla\Phi}{a^2} = \frac{\dot{a}}{a}v \ . \tag{2.6}$$

However, computation of the gravitational acceleration with direct computation for every particle pair is a major bottleneck for the computational effort, since it scales with N^2, where N is the number of particles. For these reasons efficient solver have been developed to reduce this computational cost. We describe them in the next section.

2.3 N-body dynamics solver

To optimize performance, AREPO employs two distinct gravity-solving techniques:

- Hierarchical multipole expansion with an Oct-tree (Barnes-Hut algorithm; Barnes and Hut, 1986) – this method constructs an adaptive tree structure where each tree node stores multipole moments recursively. The tree is walked for each particle, evaluating forces by either accepting the node or further subdividing it based on an accuracy criterion.

- Particle-mesh (PM) algorithm – this method calculates gravitational forces by interpolating the mass onto a Cartesian grid, solving Poisson's equation in Fourier space. It is efficient, particularly for periodic boundary conditions.

In a Tree-PM hybrid approach, AREPO splits the gravitational force into long-range (PM) and short-range (Tree) components, using the following decomposition in Fourier space:

$$\phi_k^{\text{long}} = \phi_k \exp(-k^2 r_s^2), \quad \phi_k^{\text{short}} = \phi_k \left[1 - \exp(-k^2 r_s^2)\right].$$

Long-range forces are computed using Fourier techniques, which are efficient and do not require a high spatial resolution. Short-range forces are calculated via an Oct-tree method that ensures both accuracy in high-density regions and efficiency by approximating distant particle groups using multipole expansions. The scaling of the methods is $N \log N$, which is computationally efficient (see Fig. 2.1). The tree structure is reconstructed at each local time-step, which increases computational cost but mitigates force inaccuracies due to time-step correlations. This reconstruction is particularly important in the centers of halos, where long periods without updates could lead to significant force errors.

For simulations with periodic boundaries, additional considerations are necessary. Normally, the Oct-tree considers only direct gravitational interactions using Newtonian force laws, implying the non-periodic boundary condition, as is the case in our current setup. However, for periodic boundary conditions, AREPO includes mirrored particles to account for the periodic mass distribution.

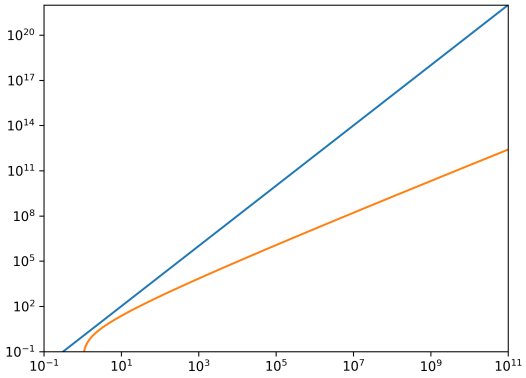


Figure 2.1: Scaling of the computational effort (y-axis) with number of particles (N, x-axis). Blue line: Direct summation ($\propto N^2$). Orange line: Tree-PM hybrid approach ($\propto N \log N$).

2.4 Hydrodynamics solver

In addition to gravity, gas in the simulations is subject to hydrodynamical processes, which must be accurately accounted for to realistically represent the structure formation process. In this section, we briefly outline how the AREPO code evolves gas dynamics.

2.4.1 Finite volume hydrodynamics

One of the defining features of the AREPO code is its use of a second-order accurate finite volume discretization scheme on a fully adaptive, dynamic Voronoi mesh. This mesh is constructed from a set of mesh-generating points, each serving as the center of a Voronoi cell. A Voronoi cell comprises all spatial locations that are closer to its generating point than to any other. The interfaces between neighboring cells are planar surfaces that are perpendicular to the line connecting the respective mesh-generating points and lie equidistant between them.

The spatial relationship between neighboring cells gives rise to a Delaunay triangulation, which is the geometric dual of the Voronoi tessellation. In three dimensions, this triangulation consists of tetrahedra formed such that the circumsphere of each contains no other mesh-generating points. This property ensures that the Delaunay tessellation is unique for a given set of points.

In AREPO, the Delaunay triangulation is constructed iteratively. The process begins with a large enclosing tetrahedron, and mesh-generating points are inserted sequentially. When a new point is added, it is connected to the vertices of the tetrahedron in which it resides, subdividing that region into smaller tetrahedra. If any of the resulting tetra-

hedra violate the empty circumsphere condition, edge-flipping operations are employed to restore the Delaunay property. This iterative refinement continues until all mesh-generating points are incorporated into the triangulation.

Once the final Delaunay triangulation is established, the Voronoi diagram is computed as its topological dual through straightforward geometric calculations. This close coupling between Delaunay and Voronoi structures underpins the mesh adaptivity and accuracy of the AREPO code (for more details Weinberger et al., 2020).

With defined cells, physical quantities such as density (ρ), velocity ($\boldsymbol{v}$), and total internal energy (u) are averaged in those cell volumes and stored at their centers. Gradients are estimated using neighboring cells, allowing piecewise linear reconstructions. To compute fluxes across cell interfaces, AREPO solves a local Riemann problem (an initial value problem for a system of hyperbolic conservation laws, where the initial data is piecewise constant with a single discontinuity at each interface).

In practice, for fluid dynamics the following set of equations is solved:

$$\frac{\partial \mathbf{U}}{\partial t} + \nabla \cdot \mathbf{F}(\mathbf{U}) = 0 \ , \tag{2.7}$$

where $\mathbf{U}$ is a set of conserved quantities:

$$\mathbf{U} = \begin{pmatrix} \rho \\ \rho \boldsymbol{v} \\ \rho e \ , \end{pmatrix}$$

and $\mathbf{F}$ are fluxes:

$$\mathbf{F}(\mathbf{U}) = \begin{pmatrix} \rho \boldsymbol{v} \\ \rho \boldsymbol{v}\boldsymbol{v}^T + P \\ (\rho e + P)\boldsymbol{v} \end{pmatrix} \ ,$$

where $e = u + 1/2\boldsymbol{v}^2$ is the total specific energy (or energy per unit mass) and u is the specific internal energy per unit mass. These are Euler equations, which describe the evolution of a compressible inviscid fluid.

To solve these equations numerically, AREPO uses a finite-volume method and computes the fluxes at cell interfaces by solving the local Riemann problem. This can be done with an exact (iterative) Riemann solver that determines the wave structure and computes intermediate states between left and right states at an interface. Alternatively, AREPO supports approximate solvers, such as the HLLC solver for pure hydrodynamics and the HLLD solver for magnetohydrodynamics (MHD).

2.4.2 Gradient estimation

In hydrodynamic simulations like those run with AREPO, gradient estimation is a fundamental step for reconstructing fluid quantities, such as density, velocity, and pressure, across the computational mesh. These gradients are essential for accurately solving the equations of fluid dynamics, particularly when computing fluxes across cell interfaces using finite-volume methods. The accuracy of these simulations therefore critically depends on how well the gradients are estimated within each cell.

Gradient estimation in Voronoi meshes can be inaccurate in distorted cells, especially when the mesh-generating point deviates significantly from the cell's center of mass. This causes errors because physical variables are defined at the mesh-generating points but are interpreted as values at the cell's center of mass. The resulting discrepancy degrades the accuracy of the gradient estimates. To improve accuracy, AREPO employs a least-squares gradient estimator (Pakmor et al., 2016):

$$\tilde{W}_j = W_i + \boldsymbol{d_{ij}} \nabla W_i \ , \tag{2.8}$$

where the gradient ∇W_i of a primitive variable W_i (ρ, $\boldsymbol{v}$, u) is determined via linear extrapolation from cell i to cell j. Since there is not only one neighbor, the gradient is estimated by minimizing residuals in a weighted least-squares sense:

$$S_{\text{tot}} = \sum_j g_j (W_j - W_i - \boldsymbol{d_{ij}} \nabla W_i)^2, \tag{2.9}$$

where $\boldsymbol{d_{ij}}$ is a position vector of cell j relative to cell i, and $g_j = A_{ij}/|\boldsymbol{d_{ij}}|$ is the adopted weight, A_{ij} is the area of the interface. This is how AREPO accurately recovers the linear gradient even in meshes with randomly distributed mesh-generating points.

The gravitational interactions of the fluid are treated using cell-centered softened gravity forces, with softening lengths tied to the radii of individual cells. These forces act as source terms in the Euler equations applied in a split-step fashion. First, gravity is applied for half a timestep, followed by solving the fluid equations in their conservative form (2.7). Then, another gravity half-step is performed. Although this approach does not conserve the exact total energy, errors decrease with better resolution. Additionally, it avoids unphysical drifts in thermal energy, which can occur when correction factors are included for changes in gravitational softening lengths.

2.4.3 Movement of the mesh

Each mesh-generating point in AREPO moves with the average velocity of the fluid and accelerates under the influence of pressure gradients and gravity (and Lorenz forces in

the MHD case). The velocity of each point is defined at half time steps, which helps to maintain approximately constant mass within each cell over time. As a result, the simulation exhibits a behavior similar to a (quasi-)Lagrangian method.

However, during complex fluid motion, cells can become distorted, often developing high aspect ratios. Such distortions can lead to numerical errors and increased noise, ultimately reducing the accuracy of the simulation (Springel, 2010; Vogelsberger et al., 2012). To mitigate this issue, AREPO applies velocity corrections to regularize the motion of highly distorted cells. A cell is considered highly distorted if the maximum angle under which a face is seen from its mesh-generating point exceeds a certain threshold. This angle is quantified by the parameter α_{face}, defined as:

$$\alpha_{\text{face}} = \left(\frac{A_{\text{face}}}{\pi}\right)^{1/2} / h_{\text{face}}, \tag{2.10}$$

$$\alpha_{\text{max}} = \max\left(\alpha_{\text{face}}\right), \tag{2.11}$$

where A_{face} is the area of the face and h_{face} is the distance from the mesh-generating point to that face.

If α_{max} exceeds the threshold 0.75β, where β is a user-defined free parameter, the velocity of the mesh-generating point is modified. Specifically, the point is moved toward the center of mass of the cell with a fraction f_{shaping} of a characteristic speed v_{char}. The corrected velocity is defined as:

$$\boldsymbol{v}_{\text{corr}} = \begin{cases} 0, & \text{if } \alpha_{\text{max}} \leq 0.75\beta, \\ f_{\text{shaping}} \frac{\alpha_{\text{max}} - 0.75\beta}{0.25\beta} v_{\text{char}} \hat{\boldsymbol{n}}, & \text{if } 0.75\beta < \alpha_{\text{max}} \leq \beta, \\ f_{\text{shaping}}\, v_{\text{char}}\, \hat{\boldsymbol{n}}, & \text{if } \alpha_{\text{max}} > \beta, \end{cases} \tag{2.12}$$

where $\hat{\boldsymbol{n}}$ is the unit vector pointing from the mesh-generating point toward the cell's center of mass. The characteristic speed v_{char} can be chosen as the local sound speed c_s, or alternatively estimated as:

$$v_{\text{char}} = \frac{d}{\Delta t}, \tag{2.13}$$

where d is the characteristic cell size and Δt is the simulation time step. By keeping the offsets between mesh-generating points and the geometric centers of their respective cells small, this regularization scheme effectively minimizes errors in the linear reconstruction step.

2.4.4 Refinement and de-refinement

During the course of a simulation, cells may deviate from their initial mass or size. To address this, AREPO supports local refinement and de-refinement of the mesh to ensure

a nearly constant mass resolution.

Refinement involves splitting a mesh-generating point into two nearby points, separated by a distance of $0.025\, r_{\rm cell}$ in a random direction from the original center. Here, $r_{\rm cell}$ is the effective radius of a sphere with the same volume as the cell, given by:

$$r_{\rm cell} = \left(\frac{3V}{4\pi}\right)^{1/3}, \tag{2.14}$$

where V is the volume of the cell. As a result of this splitting, two new cells are created without significantly disturbing the geometry of the surrounding mesh. The physical quantities (e.g., mass, momentum, energy) originally contained in the parent cell are distributed between the two daughter cells in proportion to their volumes. De-refinement, on the other hand, involves the removal of a mesh-generating point, which eliminates the corresponding Voronoi cell. Its volume is then redistributed among neighboring cells. The conserved quantities of the removed cell are apportioned to these neighboring cells in proportion to the volume they acquire.

The triggering condition for de-refinement is based by default on a mass criterion: a cell may be de-refined if its mass falls significantly below a target resolution mass. Conversely, refinement is triggered when a cell's mass exceeds twice the target mass resolution. This ensures that cell masses remain narrowly distributed around the target resolution, promoting numerical stability and accuracy. Additional refinement criteria can be designed and implemented. For instance, in galaxy formation simulations can be useful to resolve the Jeans length with a specified number of cells to ensures adequate resolution of gravitational collapse and fragmentation processes, thereby reducing artifacts associated with insufficient resolution.

It is also important to note that highly distorted cells are typically excluded from refinement. This precaution prevents frequent successive refinements from introducing severe local irregularities in the mesh. After a refinement event, a few time steps are allowed for the mesh regularization mechanism to restore a locally regular mesh geometry, after which further refinement may proceed.

2.4.5 Time integration

AREPO employs an explicit time integration scheme, which imposes constraints on the time step size to ensure the accuracy and stability of the numerical method. Each particle has its own constraint for the time step. In general, different time step criteria are applied to gravity, hydrodynamics, and source/sink terms (e.g., star formation), if

applicable. The smallest of these constraints is typically selected as the effective time step for each cell/particle.

For gas cells, the Courant–Friedrichs–Lewy (CFL) condition serves as the principal constraint:

$$\Delta t_{\rm CFL} \leq C_{\rm CFL} \frac{r_{\rm cell}}{v_{\rm signal}}, \tag{2.15}$$

where $C_{\rm CFL}$ is a dimensionless Courant factor (free parameter), $r_{\rm cell}$ is the effective cell radius (Eq. 2.14), and $v_{\rm signal}$ is the maximum signal speed in the cell. The signal speed accounts for both the local sound speed and Alfvén speed:

$$v_{\rm signal} = \left(\gamma \frac{p}{\rho} + \frac{\boldsymbol{B}^2}{\rho} \right)^{1/2}, \tag{2.16}$$

where p and ρ are the pressure and the density of cell, and $\boldsymbol{B}$ is magnetic field strength.

Another time step constraint arises from gravitational acceleration:

$$\Delta t_{\rm grav} \geq \sqrt{\frac{2 C_{\rm grav} \epsilon_{\rm soft}}{|\boldsymbol{a}|}}, \tag{2.17}$$

where $C_{\rm grav}$ is a free parameter, ϵ is the gravitational softening length, and $\boldsymbol{a}$ is the local gravitational acceleration.

In simulations where star formation is active, an additional time step constraint is introduced based on the star formation rate:

$$\Delta t_{\rm SF} = 0.1 \frac{m}{\dot{m}_{\rm SF}}, \tag{2.18}$$

where $\dot{m}_*$ is the local star formation rate, and m is gas cell mass.

To manage local time steps while maintaining global synchronization, AREPO uses a hierarchical time binning scheme. The total simulation time is divided by 2^N (N is an integer number), creating a hierarchy of discrete time bins. The time step for each particle is then assigned as:

$$\Delta t = \frac{t_{\rm end} - t_{\rm start}}{2^N}, \tag{2.19}$$

where N is the level of the time bin, such that Δt is the smallest value satisfying the cell's accuracy constraints. Each particle is placed into the highest possible time bin, which smaller than the most restrictive time step (i.e., $< \min(\Delta t_{\rm SF}, \Delta t_{\rm grav}, \Delta t_{\rm CFL})$). A particle may move to a shorter time bin at any time. However, transitioning to a longer time step is only permitted if the particle is synchronized with the global target time $t_{\rm end}$.

Once time steps are assigned, AREPO proceeds with hierarchical time integration. This method groups particles according to their time bins and restricts gravitational

calculations within each group. Each particle only computes gravitational interactions with particles in the same time bin, following the approach described in Pelupessy et al. (2012).

For hydrodynamic evolution, AREPO uses a scheme similar to Heun's method, which is second-order accurate in time. This method is computationally efficient, as it requires mesh construction only once per time step. On the moving mesh, both the interface areas A_{ij} and the cell volumes evolve over time. By applying a high-order integration scheme, AREPO accurately accounts for these geometric changes. The volume-integrated conserved variable Q_i of a cell i is defined as:

$$Q_i = \int_{\text{cell } i} \mathbf{U}\, dV, \tag{2.20}$$

where $\mathbf{U}$ is the vector of conserved quantities (mass, momentum, energy, etc.). And updated as:

$$\begin{aligned} Q_i^{n+1} &= Q_i^n - 0.5\Delta t \\ &\times (\sum_j A_{ij} F(W_{ij}^n, W_{ji}^n) + A_{ij}^{n+1} F(W_{ij}', W_{ji}')), \text{ with} \end{aligned} \tag{2.21}$$

$$W_{ij}^n = W_i^n + \boldsymbol{d_{ij}} \cdot \frac{\partial W_i}{\partial x}, \tag{2.22}$$

$$W_{ij}' = W_i^n + \boldsymbol{d_{ij}'} \cdot \frac{\partial W_i}{\partial x} + \Delta t \frac{\partial W_i}{\partial t}\ , \tag{2.23}$$

where $\boldsymbol{d_{ij}}$ is the position vector of the geometric center of the interface between cell i and j relative to center of cell i, W is primitive variable, and $\frac{\partial W}{\partial x}$ and $\frac{\partial W}{\partial t}$ are their rate of changes in space and time. The fluid fluxes are first computed using the initial mesh geometry, and then recalculated on the updated mesh using extrapolated (primed) quantities. Time derivatives are evaluated using spatial gradients via the continuity equation (Pakmor et al., 2016). As for gravitational dynamics, AREPO supports local time stepping for hydrodynamical updates. In this scheme, fluxes between two cells are updated only if at least one of them is in its active phase (Springel, 2010).

By integrating these advanced numerical techniques, AREPO enables high-accuracy simulations of galaxy formation, resolving both large-scale cosmological structures and small-scale processes within individual galaxies.

2.5 SMUGGLE – Stars and MUltiphase Gas in GaLaxiEs

Numerical methods play a crucial role in understanding galaxy formation and evolution. They allow us to explore the complex processes involved in the life cycle of galaxies.

Achieving "realistic" simulations is a challenging task, as these processes span a wide dynamical range, from sub-parsec scales (e.g., supernova explosions) to megaparsec scales (e.g., dark matter halo formation and mergers).

Modern cosmological or galaxy-wide simulations rely on sub-resolution (sub-grid) models: numerical prescriptions that capture unresolved physical phenomena at a given resolution level. For instance, current simulations cannot resolve the evolution of individual stars. Therefore, most relevant physical processes are implemented in a sub-grid fashion to some degree, including radiative cooling, star formation, stellar evolution, SN feedback, active galactic nucleus (AGN) feedback, and more. The resolution level is critical, as it determines the convergence properties of the simulation. Below a certain convergence threshold, the simulation no longer yields new or reliable physical insights.

The physical processes incorporated into numerical methods must be based on solid empirical evidence. These processes include not only hydrodynamics and gravity but also star formation and the interaction of stars with the surrounding interstellar medium (ISM). Observations indicate that only about 20–25% of baryonic matter ends up in stars, and in low-mass dwarf galaxies, this fraction can be as low as 1% (Conroy and Wechsler, 2009; Behroozi et al., 2019). The remaining baryonic matter resides in the ISM, the CGM, or remains outside galaxies altogether (Bregman, 2007; Putman et al., 2012). This suggests the presence of regulatory mechanisms that prevent gas from cooling and forming stars (Benson et al., 2002). It is now widely accepted that *stellar feedback* is the primary factor responsible for the low stellar mass-to-halo mass ratio (Kim and Ostriker, 2018) for galaxies at the MW mass scale and below, but including it in simulations is not a trivial task. This is because stellar feedback is intimately linked to the process of star formation itself and is also affecting the properties of the ISM. The ISM is a complex environment composed of multiple phases, hot, warm, and cold gas, that coexist and interact within galaxies. Numerically modeling this gaseous medium is especially difficult in the cold, dense regions, where critical physical processes occur on short timescales. Therefore, resolving and properly modeling the multiphase structure of the ISM remains one of the central goals in contemporary studies of galaxy formation and evolution.

The SMUGGLE (the Stars and MUltiphase Gas in GaLaxiEs; Marinacci et al., 2019; Barbani et al., 2025) model is a numerical framework designed for modeling ISM gas and stellar feedback in the moving-mesh code AREPO. It aims to resolve the different gas phases of the ISM and implements a detailed treatment of stellar feedback. There are three main physical components considered: gas heating and cooling mechanisms, star formation, and stellar feedback processes. The methodology and implementation details of the SMUGGLE are described in the following subsections.

The inclusion of the SMUGGLE model in AREPO is essential because, while AREPO provides a highly accurate hydrodynamic solver with a moving mesh scheme, it cannot resolve the full complexity of the interstellar medium or the detailed processes of stellar feedback at the resolution scales of cosmological simulations. SMUGGLE thus complements the AREPO code by introducing physically motivated subgrid prescriptions that model the interaction between gas phases and the injection of energy and momentum from stellar sources. Presenting this model is therefore necessary to understand how our simulated HI emission data incorporates realistic ISM dynamics and feedback mechanisms within the framework of a state-of-the-art hydrodynamic code for galaxy formation simulations.

2.5.1 Cooling and heating

The thermal state of the gas is determined by a network of cooling and heating processes that govern the multiphase nature of the ISM. The interplay between heating and cooling mechanisms is essential to maintain the thermal equilibrium of the ISM. Efficient cooling allows gas to condense into dense clouds, thus starting star formation, while heating processes prevent runaway collapse and regulate the distribution of gas phases. This balance influences the morphology and stability of galaxies, shaping their evolution over cosmic time.

The cooling and heating result from two-body interactions in gas, primarily composed of hydrogen and helium: collisional excitation, collisional ionization, recombination, dielectric recombination, and free-free emission, Compton cooling off CMB photons, and photoionization from a UV background. In addition to the mechanisms occurring in the gas of the primordial universe, there are high-temperature metal line cooling ($T \gtrsim 10^4$ K), and low-temperature molecular cooling processes ($T \lesssim 10^4$ K) proceeding in the later Universe. The net metal cooling rate is expressed as a function of temperature, gas density, and redshift based on CLOUDY calculations (Ferland et al., 1998). It was added to primordial cooling and scaled by metallicity. These processes are already present in AREPO (see Vogelsberger et al., 2013).

The molecular cooling is important to reach very low temperatures (~ 10 K) necessary

to reach high densities, and it is implemented by SMUGGLE as:

$$\begin{aligned}\Lambda_{\rm mol} = 2.896 \times 10^{-26} \\ &\left\{\left(\frac{\rm T}{125.215\ {\rm K}}\right)^{-4.9202} + \left(\frac{\rm T}{1349.86\ {\rm K}}\right)^{-1.7288} + \left(\frac{\rm T}{6450.06\ {\rm K}}\right)^{-0.3075}\right\}^{-1} \\ &\times \left(0.001 + \frac{0.1 n_{\rm H}}{1+n_{\rm H}} + \frac{0.09 n_{\rm H}}{1+0.1 n_{\rm H}} + \frac{(Z/Z_\odot)^2}{1+n_{\rm H}}\right) \times \left(\frac{1+(Z/Z_\odot)}{1+0.00143 n_{\rm H}}\right) \\ &\times \exp\left(-\left[\frac{\rm T}{158\ 000\ {\rm K}}\right]^2\right) {\rm erg\ s^{-1} cm^3}\ ,\end{aligned} \tag{2.24}$$

where T is the gas temperature, $n_{\rm H}$ (in cm^{-3}) is the hydrogen number density, Z and $Z_\odot$ are the gas and the solar metalicities respectively.

There is another crucial component taking place in high density gas that regulates its ionization and heating rates, i.e. self-shielding (Vogelsberger et al., 2013). This adjustment typically manifests as a reduction in the effective UV radiation field intensity within dense gas, thereby lowering ionization rates and enhancing cooling from less ionized species. The self-shielding is given as a function of redshift and gas density:

$$f_{\rm ssh} = (1-f)\left[1+\left(\frac{n_{\rm H}}{n_0}\right)^{\beta}\right]^{\alpha_1} + f\left[1+\left(\frac{n_{\rm H}}{n_0}\right)^{\beta}\right]^{\alpha_2}\ , \tag{2.25}$$

where $n_{\rm H}$ is the hydrogen number density and $n_0,\ \alpha_1,\ \alpha_2,\ \beta,\ f$ are redshift dependent parameters, found as the best fit to photoionization function (Rahmati et al., 2013). At redshifts $z > 6$, self-shielding effects are neglected.

The cosmic ray and photoelectric heating, which are important for the thermal balance and stability of cold ($\sim$ 50 K) and warm ($\sim$ 8000 K) gas (Field et al., 1969), are also taken into account in the model. Cosmic ray effects are important for high-density gas, and interact via hadronic losses and Coulomb interactions between cosmic rays and the gas:

$$\Lambda_{\rm CR} = -10^{-16}(0.98+1.65)\tilde{n}_{\rm e} e_{\rm CR} n_{\rm H}^{-1}\ ({\rm erg\ s^{-1}\ cm^3})\ , \tag{2.26}$$

where, $\tilde{n}_{\rm e}$ is the electron number density in units of the hydrogen number density ($n_{\rm H}$) and $e_{\rm CR}$ is the cosmic ray energy density. Photoelectric heating refers to the emission of electrons from dust grains, primarily triggered by the absorption of ultraviolet photons from the interstellar radiation field. This process is especially efficient in small dust grains and polycyclic aromatic hydrocarbons (PAHs), which absorb energetic photons and emit photoelectrons, thereby heating the surrounding interstellar gas (Wolfire et al.,

2003):

$$\Lambda_{\rm phot} = -1.3 \times 10^{-24} \tilde{e}_{\nu}^{\rm pe} n_{\rm H}^{-1} \left(\frac{Z}{Z_{\odot}}\right) \times \left(\frac{0.049}{1+(x_{\rm pe}/1925)^{0.73}} + \frac{0.037(T/10^4 K)^{0.7}}{1+(x_{\rm pe}/5000)}\right) \times {\rm erg\ s^{-1} cm^3}\ , \tag{2.27}$$

$$x_{\rm pe} = \frac{\tilde{e}_{\nu}^{pe} T^{0.5}}{\Phi_{\rm PAH} \tilde{n}_e n_{\rm H}}\ , \tag{2.28}$$

where $\tilde{e}_{\nu}^{pe}$ is the photon energy density, normalized to MW units, $\Phi_{\rm PAH}$ is a factor of uncertainties in the interaction rates between atoms and dust grains. As a default choice SMUGGLE takes the observationally determined value $\Phi_{\rm PAH} = 0.5$ (Jenkins and Tripp, 2001).

2.5.2 Star formation

Star formation is a fundamental process that shapes the evolution of galaxies. It governs the transformation of interstellar gas into luminous stellar populations, influencing both the structure and dynamics of galaxies over cosmic time. The rate at which stars form, the conditions under which this process occurs, and the interplay between stellar feedback and the surrounding medium are crucial in determining the morphology, chemical composition, and overall life-cycle of a galaxy (McKee and Ostriker, 2007). Star formation regulates the baryonic matter cycle by driving the conversion of gas into stars and enriching the ISM with metals through supernovae and stellar winds. The latter feedback mechanisms affect subsequent generations of star formation and contribute to the large-scale properties of galaxies.

In the SMUGGLE model, cold gas is converted into stars using a probabilistic approach (Springel and Hernquist, 2003). Star particles represent a group of stars that formed simultaneously and follow the Chabrier initial mass function (Chabrier, 2001). They are stochastically created based on the probability determined by the star formation rate ($\dot{M}_{\star}$) of each gas cell, given by:

$$\dot{M}_{\star} = \begin{cases} 0, & \rho < \rho_{th} \\ \epsilon \frac{M_{\rm gas}}{t_{\rm dyn}}, & \rho \geq \rho_{th}\ , \end{cases} \tag{2.29}$$

where ϵ is the star formation efficiency factor ($\epsilon = 0.01$), and $t_{\rm dyn}$ is the gravitational dynamical time for a gas cell:

$$t_{\rm dyn} = \sqrt{\frac{3\pi}{32 G \rho_{\rm gas}}}\ . \tag{2.30}$$

A critical condition for star formation is the gas density threshold $\rho_{\rm th}$, set to 100 cm^{-3}, which is within the typical range of molecular cloud densities ($10^2 - 10^3$ cm^{-3}). Additionally, the virial parameter α determines whether a gas cell i is gravitationally bound and able to form stars:

$$\alpha_i = \frac{||\nabla \cdot \boldsymbol{v_i}||^2 + (c_{{\rm s},i}/\Delta x_i)^2}{8\pi G \rho_i} \ , \tag{2.31}$$

where v is a gas velocity, c_s is a speed of sound, Δx is a cell size (2.14), ρ is a gas density and $||\nabla \cdot \boldsymbol{v_i}||^2$ is defined as:

$$||\nabla \cdot \boldsymbol{v_i}||^2 \equiv \sum_{i,j} \left(\frac{\partial v_i}{\partial x_j} \right)^2 \ . \tag{2.32}$$

If $\alpha_i < 1$, the gas remains gravitationally bound, as thermal support and gas motion are insufficient to prevent collapse, and star formation is allowed for cell i.

Given the star formation rate, the stellar mass produced over a time step Δt in a cell i is computed as:

$$M_{\star,i} = M_i \left[1 - \exp\left(-\frac{\dot{M}_\star \Delta t}{M_i} \right) \right] \ , \tag{2.33}$$

where M_i is the mass of the cell i and $\dot{M}_\star$ is defined in Eq. (2.29). The term in square brackets in Eq. (2.33) is the base of probabilistic approach adopted in SMUGGLE. From this equation it is possible to derive the probability $p = 1 - \exp\left(-\frac{\dot{M}_\star \Delta t}{M_i}\right)$ that a stellar population of mass $M_{\star,i}$ forms in a cell i over a time Δt. This probability is compared with a randomly generated value $p_\star$ drawn from a uniform distribution in the range [0,1]. If $p > p_\star$, the gas cell is converted into a star particle with the mass of the original cell.

Alternatively, the model can compute star formation solely based on the molecular gas fraction ($f_{\rm H_2}$), expressed as a function of density and metallicity, by multiplying Eq. (2.29) by f_{H_2} (McKee and Krumholz, 2009), which is:

$$f_{\rm H_2} = 1 - \frac{3s}{4+s} \ , \tag{2.34}$$

with

$$s = \frac{\log[1 + \chi(0.6 + 0.01\chi)]}{0.6\tau} \ , \tag{2.35a}$$

$$\chi = 2.3 \frac{1 + 3.1(Z/Z_\odot)^{0.365}}{3} \ , \tag{2.35b}$$

$$\tau = 0.067 \left(\frac{Z}{Z_\odot} \right) \left(\frac{\Sigma}{\rm M_\odot pc^{-2}} \right) \ , \tag{2.35c}$$

$$\Sigma = \rho \left(\frac{\rho}{||\nabla \rho||} \right) \ , \tag{2.35d}$$

where s is a self-shielding parameter, χ is the dimensionless parameter characterizing UV dissociation field strength, τ is the dust optical depth, Σ is the column density, $\frac{\rho}{||\nabla\rho||}$ is the density scale height, Z and $Z_\odot$ are the gas and the solar metalicities. The molecular gas fraction is forced to be positive ($f_{\mathrm{H}_2} > 0$), therefore, negative values of Eq. (2.34) are set to zero, which ensures that no star formation occurs in those gas cells.

2.5.3 Feedback from supernovae

In addition to star formation, another key physical process implemented in the SMUGGLE model is stellar feedback, with a particular focus on SNe. SN injects a large amount of energy and momentum in the surrounding ISM, and this injection is responsible for driving turbulence in the ISM and (potentially) galactic-scale outflows. However, there are several technical difficulties in implementing effective SN feedback in galaxy formation simulations.

A primary challenge lies in properly capturing the **Sedov-Taylor phase**, which describes the crucial interaction of the over-pressurized supernova remnant with the ambient ISM. During this phase, the expanding remnant sweeps up and accelerates the surrounding material. This energy and momentum injection from supernovae plays a critical role in regulating subsequent star formation by modifying the gas properties and driving large-scale gas flows. In total, the energy injected by SN is parameterized as:

$$E_{\mathrm{SN}} = f_{\mathrm{SN}} E_{51} \ , \tag{2.36}$$

where $E_{51} = 10^{51}$erg and f_{SN} is a model parameter representing SN efficiency (set to 1 as a default choice). The momentum at the time of the explosion is given by:

$$p_{\mathrm{SN}} = M_{\mathrm{SN}} v_{\mathrm{SN}} \ , \tag{2.37}$$

where M_{SN} is the ejecta mass and v_{SN} is blast wave velocity.

In SMUGGLE, each particle represents not only a single star but rather a stellar population. This means that a single particle can account for multiple SN events during a given timestep Δt. To compute amount of energy injected, we must compute the number of SNe occurring in given timestep:

$$E_{\mathrm{SN,tot}} = f_{\mathrm{SN}} E_{51} (N_{\mathrm{SNII}} + N_{\mathrm{SNIa}}) \tag{2.38}$$

and

$$p_{\mathrm{SN,tot}} = p_{\mathrm{SNII,tot}} + p_{\mathrm{SNIa,tot}} \ , \tag{2.39}$$

where N_{SNII} and N_{SNIa} is the number of type II and type Ia SN occurring per time-step Δt. The number of type II supernovae is defined by the integration of the initial mass function of a stellar particle as:

$$N_{\rm SNII,tot} = M_\star \int_{M(t+\Delta t)}^{M(t)} \Phi(m) dm \ . \tag{2.40}$$

$M_\star$ is the mass of the star particle at birth, $M(t)$ is the mass of stars leaving the main sequence at an age t and $\Phi(m)$ is the Chabrier (Chabrier, 2001) initial mass function. There is a physical limit for $M \in$ [8:100] $M_\odot$. For the type Ia SN a different approach is used:

$$N_{SNIa} = \int_t^{t+\Delta t} {\rm DTD}(t') dt' \ , \tag{2.41}$$

where DTD is a delay time distribution that encodes the uncertainties on the progenitors leading to this type of SN events. In the SMUGGLE the following DTD is adopted (see Vogelsberger et al., 2013):

$${\rm DTD}(t) = \Theta(t - \tau_8) N_0 \left(\frac{t}{\tau_8}\right)^{-s} \frac{s-1}{\tau_8} \ . \tag{2.42}$$

Here, $\tau_8 = 40$ Myr, approximates the main sequence life time for star with $M \sim 8\ M_\odot$, $N_0 = 2.6 \times 10^{-3} SN\ M_\odot^{-1}$, s $= 1.12$ and Θ is Heaviside function that sets the delay between birth of stellar population and first SNIa explosion. To determine $N_{\rm SN}$ and $N_{\rm Ia}$ we integrate Eq. (2.40) and Eq. (2.41) over time step $\Delta t_\star$ and take as the expectation value of the Poisson distribution for the number $SN(\lambda)$ to obtain the real discrete number per time step.

The Sedov-Taylor phase lasts until the SN remnant approaches the cooling radius where the gas reaches to the temperature of $\approx 10^6$ K:

$$r_{\rm cool} = 28.4 E_{51}^{2/7} \langle n \rangle^{-3/7} f(Z) \ {\rm pc}, \tag{2.43}$$

and radiative losses becomes important. Here, $\langle n \rangle$ the average gas density within $r_{\rm cool}$, and $f(Z)$ is a function of metallicity defined by Eq. (2.46). In galaxy formation simulations, $r_{\rm cool}$ is not fully resolved so we need to take into account the PdV work done by the hot, post-shock gas:

$$\Delta p_i = \tilde{w}_i {\rm min} \left[p_{\rm SN,tot} \sqrt{1 + \frac{m_i}{\Delta m_i}}, p_{\rm t} \right] \ , \tag{2.44}$$

where, $\tilde{w}_i$ is a weight function, used to spread momentum into the neighboring cells of a given star particle, m_i is the gas cell mass, $\Delta m_i = \tilde{w}_i (M_{\rm SNII,tot} + M_{\rm SNIa,tot})$, and $p_{\rm t}$ is terminal momentum, i.e., the final momentum when SN reaches $r_{\rm cool}$:

$$p_{\rm t} = 4.8 \times 10^5 E_{\rm SN,tot}^{13/14} \left(\frac{\langle n_{\rm H} \rangle}{1\ {\rm cm}^{-3}} \right)^{-1/7} f(Z)^{3/2} M_\odot {\rm km\,s^{-1}}, \tag{2.45}$$

where $\langle n_H \rangle$ is the hydrogen gas density around the star, and:

$$f(Z) = \min\left[\left(\frac{\langle Z \rangle}{Z_\odot}\right)^{-0.14}, 2\right] , \tag{2.46}$$

with $\langle Z \rangle$ being average SPH-weighted gas metallicity around star particle and $Z_\odot = 0.0127$ the adopted value of solar metallicity (Asplund et al., 2009). The terminal momentum is estimated by high-resolution simulations of individual SN blast-waves (Cioffi et al., 1988).

To evaluate how much momentum is imparted to gas cell surrounding the SN event, it is necessary to determine the weighting function $\tilde{w}_i$. To do so, we compute number of affected neighbors:

$$N_{\rm ngb} = \frac{4\pi}{3} h^3 \sum W(|\boldsymbol{r_i} - \boldsymbol{r_s}|, h) , \tag{2.47}$$

where h is a coupling radius, W is the standard cubic spline SPH kernel, $\boldsymbol{r_i}$ and $\boldsymbol{r_s}$ are position vectors of i-th neighbor and star particle. Eq. (2.47) is used to define the scale h iteratively until a predetermined number of neighbors reached (the default choice is $N_{\rm ngb} = 64$).

The weighting function determined such a way that each gas cell within h gets the momentum imparted by a SN proportionally to the fraction of 4π solid angle that lies centered on the star particle:

$$w_i = \frac{1}{2}\left\{1 - \frac{1}{[1 + A_i/(\pi|\boldsymbol{r_i} - \boldsymbol{r}_{\rm s}|^2)]^{1/2}}\right\} , \tag{2.48}$$

where $|\boldsymbol{r_i} - \boldsymbol{r}_{\rm s}|$ distance between cell i and star, A_i is the area of the i-th gas cell. So, the final weighting function given by:

$$\tilde{w}_i \equiv \frac{w_i}{\sum_i w_i} . \tag{2.49}$$

2.5.4 Radiative feedback

Radiation emitted by young and massive stars can significantly influence the ISM. By emitting high-energy photons the star can ionize the gas surrounding them, leading to an increase in the temperature of the ISM. Additionally, radiation pressure can also exert a significant impact on the ISM, an effect that is further enhanced by photon scattering, which becomes particularly substantial at the high densities found in molecular clouds. These processes can shape the environment around a stellar population even before the onset of a supernova explosion. Such processes are often referred to collectively as early stellar feedback.

The SMUGGLE model accounts for these effects of radiative feedback. This is achieved by calculating the ionizing photon rate $N_\star$ emitted by an individual stellar particle as:

$$N_\star = \frac{L_\star}{\langle h\nu \rangle} = \frac{\gamma_\star M_\star}{\langle h\nu \rangle} \ , \tag{2.50}$$

where $L_\star$ is the star particle's luminosity, $M_\star$ is its mass, and $\langle h\nu \rangle$ is the average photon energy above the hydrogen ionization threshold of 17 eV. The parameter $\gamma_\star$ represents a mass-to-light ratio, here set to $\gamma_\star = 10^3 \mathrm{L}_\odot / M_\odot$, which is peak value for blackbody emission at a temperature of 40 000 K (Rybicki and Lightman, 1986). A young massive star can photoionize its surroundings, forming regions of ionized hydrogen (HII regions). The SMUGGLE model incorporates this effect and includes it in its framework.

The mass of gas that can be ionized is usually less than the mass of gas contained within the cell. Therefore, the photoionization process is treated probabilistically, where each cell is assigned an ionization probability:

$$p = \frac{n_\star}{\alpha_{\mathrm{rec}} n_{\mathrm{H}}^2 V} \ , \tag{2.51}$$

where α_{rec} is the recombination rate of hydrogen, n_H is the mean hydrogen number density, V is the cell volume, and $n_\star$ is the injected ionizing photon emission rate and it is equal to $n_\star = \tilde{w}_i N_\star$, where $\tilde{w}$ is calculated via Eq. (2.49). This probabilistic approach is implemented by comparing the expected number of recombinations with the total number of ionizing photons, using a method similar to that employed for star formation. Namely, a random number between 0 and 1 is drawn from a uniform distribution and compared with the computed ionization probability. If the calculated probability exceeds the random number, the cell is ionized. A lower temperature limit of 1.7×10^4 K is imposed, and cooling is disabled for t_{off}, which is set equal to the duration of the time-step of the star particle. To activate the probabilistic mechanism, the cell must satisfy two conditions: (i) the thermal energy per unit mass must be less than 1.2 times the ionization energy per unit mass corresponding to the photoionization temperature T_{phot} ($u_{\mathrm{therm}} < 1.2 \times u_{\mathrm{phot}}$) and (ii) its cooling is not disabled by previous photoionization events ($t_{\mathrm{off}} = 0$). If either condition is not met, the cell is considered already ionized. The ionized medium will experience an increase in pressure, leading to its expansion into neighboring cooler cells. In simulations, this process is represented by the expansion of ionized hydrogen regions around young stellar particles, which transfer momentum and energy to the surrounding medium. It is important to note that, in addition to photoionization, radiation pressure also plays a significant role. Its effects are described in more detail below.

Radiation pressure has an additional effect on ISM. This is included in SMUGGLE as a source of momentum, which is particularly important to precondition the environment before the supernova explosion in dense, optically thick regions. The momentum injection due to this process can be quantified as:

$$\Delta p = \frac{L_\star}{c}(1 + \tau_{\rm IR})\Delta t \ , \tag{2.52}$$

where Δp is the total momentum imparted by stellar photons per timestep, $L_\star$ is the stellar luminosity, c is the speed of light, and $\tau_{\rm IR} = \kappa_{\rm IR}\Sigma_{\rm gas}$ is the optical depth for infrared radiation, $\kappa_{\rm IR}$ is the opacity (set to 10 $[Z/Z_\odot]$ cm^2 g^{-1}; Hopkins et al., 2018), and $\Sigma_{\rm gas}$ is the column density of the gas. Momentum due to radiation pressure is imparted within h, as for SN feedback.

2.5.5 Effects of stellar winds

Stellar populations influence their surrounding environment through various feedback mechanisms, including stellar winds. This effect is especially pronounced in massive OB stars, which have short lifetimes, and in Asymptotic Giant Branch (AGB) stars. Like photoionization, stellar winds, particularly from OB stars, affect the ISM prior to any SN events, providing a key channel of early feedback.

Stellar feedback operates across multiple timescales and via distinct mechanisms, each contributing differently to galaxy evolution. OB stars begin shaping their environment within the first few million years through intense radiation and powerful stellar winds. These winds can disperse surrounding gas, reducing its density and amplifying the effectiveness of subsequent SN explosions, which typically occur 3–10 Myr after star formation. AGB stars, by contrast, contribute at much later stages, typically after ∼100 Myr, through lower-energy but chemically rich mass loss. While not as energetically disruptive as OB stars or SNe, AGB winds are essential for enriching the ISM and supporting long-term gas recycling. Capturing both the early, energetic input from OB stars and the delayed, chemical feedback from AGB stars is crucial for realistic simulations of star formation, chemical evolution, and ISM dynamics.

In the SMUGGLE model, the effect of OB stellar winds is evaluated first by computing the cumulative mass loss per unit mass of a star particle:

$$m_{\rm closs} = \begin{cases} f(t), & t < 1 \text{ Myr} \\ g(t), & 1 < t < 3.5 \text{ Myr} \\ h(t), & 3.5 < t < 100 \text{ Myr} \end{cases} \ , \tag{2.53}$$

where

$$\begin{cases} f(t) = 4.763 \times 10^{-3}(0.01 + \tilde{Z})t \\ g(t) = 4.763 \times 10^{-3}(0.01 + \tilde{Z})\frac{t^{2.45+0.8\log(\tilde{Z})}-1}{2.45+0.8\log(\tilde{Z})} + f(1) \\ h(t) = g(3.5) - 4.57 \times 10^{-2}\left[\left(\frac{t}{3.5}\right)^{-2.25} - 1\right] + 4.2 \times 10^{-6}(t - 3.5) \end{cases} . \tag{2.54}$$

In the previous equations, $\tilde{Z}$ is the metallicity of the particle in solar units, and time t is given in Myr. For $t > 100$ Myr, the cumulative mass loss per unit mass follows $m_{\rm closs} = h(100)$. Therefore, for a stellar particle of mass $M_\star$ the mass loss within a timestep Δt can be computed as:

$$M_{\rm loss} = M_\star[m_{\rm closs}(t + \Delta t) - m_{\rm closs}(t)] \ . \tag{2.55}$$

The composition of OB stellar winds matches the composition of the stellar particle. Mass loss is computed for OB stars with a minimum stellar mass of $8M_\odot$.

On the other hand, the evolutionary models of AGB stars are incorporated using methods already implemented in the AREPO framework (Vogelsberger et al., 2013). In this case, the mass return of element i from AGB stars is given as:

$$\Delta M_i(t, \Delta t, Z) = \int_{\mathcal{M}(t+\Delta t)}^{\mathcal{M}(t)} (y_i + mZ_i f_{rec}(m, Z))\Phi(m)dm \ , \tag{2.56}$$

where y_i is the yield for each element, $\mathcal{M}(t)$ is a lifetime function, $\Phi(m)$ is the IMF, f_{rec} is a stellar recycling fraction (i.e. the fraction of total mass returned to the ISM), m is a stellar mass and Z_i is the mass fraction of element i. The yields for AGB stars taken from Karakas (2010).

The energy and momentum imparted by stellar winds are then computed using the following formulas (Hopkins et al., 2018):

$$E_{\rm winds} = \Delta t L_{\rm kin} = M_{\rm loss} \times \psi \times 10^{12} {\rm erg\ g^{-1}} \ , \tag{2.57}$$

where $M_{\rm loss}$ is the mass loss defined in Eq. (2.55), and ψ:

$$\psi = \frac{5.94 \times 10^4}{1 + \left(\frac{t_{\rm Myr}}{2.5}\right)^{1.4} + \left(\frac{t_{\rm Myr}}{10}\right)^5} + 4.83 \ , \tag{2.58}$$

$$p_{\rm winds} = \sqrt{2 M_{\rm loss} E_{\rm winds}} \ . \tag{2.59}$$

$t_{\rm Myr}$ is the age of the star particle in Myr. The mass, metallicity, momentum and energy are injected the same way as SN case, with the same maximum coupling radius. The key

important difference is that stellar winds are continuous process, and this is reflected in the SMUGGLE implementation.

This completes the description of the key physical processes that establish the multi-phase ISM nature, including radiative cooling and heating, star formation, and various modes of stellar feedback such as SN, stellar winds, and radiative feedback. Together, these components form the foundation of the SMUGGLE model. With this physical framework in place, we now proceed to discuss the simulations of an isolated MW-like galaxy from which we generated the mock observations analyzed in this work.

2.6 Initial conditions of Milky Way-like galaxy simulations

The initial conditions for the simulations analyzed in this work are generated using the methodology developed by Springel et al. (2005), which enables the construction of multi-component galaxies in approximate dynamical equilibrium. This framework has been further adapted to accommodate the specific requirements of the simulation setup analyzed in this thesis, incorporating the refinements proposed by Barbani et al. (2023) to include hot gas in hydrostatic equilibrium around the galaxy (the so-called corona), which is usually not considered in the hydrodynamical N-body simulations of this kind. To provide a controlled environment for testing feedback mechanisms, a simplified model of a single, isolated galaxy was adopted. This model comprises several structurally distinct components: a dark matter halo, a stellar bulge, a thick stellar disc, a thick gaseous disc, and an extended hot gaseous corona representing the hot phase of the CGM. Below, we briefly describe the properties as included in the setup of the simulations.

The dark matter halo and the stellar bulge are both modeled as spherically symmetric systems following the Hernquist profile (Hernquist, 1990):

$$\rho_*(r) = \frac{M_*}{2\pi} \frac{a_*}{r(r+a_*)^3}, \tag{2.60}$$

where the subscript $*$ denotes either the dark matter halo (**dm**) or the bulge (**b**), a_* is the corresponding scale radius, and M_* is the total mass of the component. The Hernquist profile is chosen for its analytical simplicity and it is well-suited for representing spherical systems in equilibrium.

The stellar and gaseous components follow an exponential distribution in the radial

direction (Freeman, 1970):

$$\Sigma_*(R) = \frac{M_*}{2\pi h_*^2} \exp(-R/h_*) \ , \tag{2.61}$$

where $*$ represents the star ($\star$) and gas ($\mathbf{g}$) components. h_* is the scale length and Σ_* is the surface density of the stellar or gaseous disc. The temperature of the gas disc is set to 10^4 K. The vertical profile of the stellar disc is modeled with a sech^2 profile:

$$\rho_\star(R,z) = \frac{\Sigma_\star(R)}{2z_\star} \mathrm{sech}^2\left(\frac{z}{z_\star}\right) \ , \tag{2.62}$$

where $z_\star$ is the scale height and $\Sigma_\star$ is the surface density of the stellar disc. In contrast, the vertical distribution of the gaseous disc is computed self-consistently to ensure hydrostatic equilibrium in the vertical gravitational potential. This is achieved at the beginning of the simulation by solving iteratively the corresponding equilibrium equations:

$$\frac{1}{\rho_g}\frac{\partial P}{\partial z} = -\frac{\partial \Phi}{\partial z} \ , \tag{2.63}$$

where Φ is the total gravitational potential and P is the gas pressure.

To investigate the interaction between the disc and the surrounding medium, a hot component of the CGM surrounding the galaxy, the so-called galactic corona, was included in the initial conditions. This component is described by a β-model:

$$\rho(r) = \rho_0 \left(1 + \frac{r^2}{r_c^2}\right)^{-3\beta/2} \ , \tag{2.64}$$

where $\beta = \frac{2}{3}$ (Jones and Forman, 1984; Moster et al., 2011), ρ_0 is the central density, and r_c is the core radius. The coronal density profile is truncated at $10R_{200}$, with parameters set to $r_c = 83$ kpc and $\rho_0 = 3.3 \times 10^3\,\mathrm{M_\odot kpc^{-3}}$. R_{200} is the radius where density equal to 200 times the critical density. The motion of the coronal gas is modeled with an azimuthal rotational velocity given by:

$$v_\phi = \alpha v_c, \tag{2.65}$$

where $v_c = \left(R\frac{\partial \Phi}{\partial R}\right)^{1/2}$ is the circular velocity due to the gravitational potential Φ. α is set to 0.4 to achieve a rotational velocity of $v_\phi \sim 90\,\mathrm{km\,s^{-1}}$ near the galactic disc to reproduce observations (Marinacci et al., 2011). The temperature profile of the coronal gas is obtained by solving the equation:

$$\Phi_{\mathrm{eff}} = \Phi(R,z) - \int_\infty^R \frac{v_\phi^2(R')}{R'}\, dR', \tag{2.66}$$

which is similar to the hydrostatic equilibrium equation with the inclusion of a centrifugal term. The metallicity of the coronal gas is assumed to be $0.1Z_{\odot}$, according to observational estimates (Bogdán et al., 2017), while the metallicity of the gaseous disc is modeled with an exponential profile, to maintain the observed metallicity profile in MW (Lemasle et al., 2018). Finally, axisymmetry is enforced for all profiles.

R_{200} (kpc)	$M_{\rm dm}$ ($M_\odot$)	r_s (kpc)	$M_{\rm b}$ ($M_\odot$)	a (kpc)	$M_\star$ ($M_\odot$)	$r_\star$ (kpc)	h (pc)	$M_{\rm g}$ ($M_\odot$)	$r_{\rm g}$ (kpc)	M_{cor} ($M_\odot$)	r_c (kpc)	β
241	1.53×10^{12}	36.46	1.5×10^{10}	1.3	4.73×10^{10}	3.8	380	9×10^{9}	7.6	5×10^{10}	83	2/3

Table 2.1: Structural parameters in simulation. $\boldsymbol{R_{200}}-$ radius where density equal to 200 times of critical density. $\boldsymbol{M}_{\rm dm}-$ dark matter halo mass. $\boldsymbol{r_s}-$ dark matter halo scale length. $\boldsymbol{M}_{\rm b}-$ bulge mass. $\mathbf{a}-$ bulge scale length. $\boldsymbol{M_\star}-$ stellar disc mass. $\boldsymbol{r_\star}-$ stellar disc scale length. $\mathbf{h}-$ stellar disk height. $\boldsymbol{M}_{\rm g}-$ gaseous disc mass. $\boldsymbol{r}_{\rm g}-$ gaseous disc scale length. $\boldsymbol{M}_{\rm cor}-$ mass of corona computed within viral radius (R_{200}). $\boldsymbol{r_c}-$ corona core radius. $\boldsymbol{\beta}-$ beta model parameter.

Table 2.1 summarizes the parameters used to model the galaxy in the simulation. The total mass of the system is 1.6×10^{12} $M_\odot$, with the bulge contributing approximately 1.5×10^{10} $M_\odot$, the stellar disc 4.73×10^{10} $M_\odot$, and the gaseous disc 9×10^{9} $M_\odot$. To improve computational efficiency, the dark matter halo is modeled as a static gravitational potential. The system is enclosed within a cubic volume with a side length of 600 kpc. The gravitational softening length is set to $\epsilon_g = 10$ pc for gas and $\epsilon_\star = 21.4$ pc for stars. Additional resolution parameters, such as the number of particles and the mass per particle, are provided in Table 2.2. The parameters governing the physical processes implemented in SMUGGLE and included in the simulation are listed in Table 2.3.

$\epsilon_\star$, pc	$\epsilon_{\rm gas}$, pc	$m_{\rm gas}$, $M_\odot$	$N_{\rm gas}$	$m_{\rm b}$, $M_\odot$	N_b	$m_{\rm d}$, $M_\odot$	$N_{\rm d}$	$N_{\rm gas,cor}$	$N_{\rm gas,d}$
21.4	10.0	1.1×10^{4}	1.1×10^{7}	2.0×10^{4}	8×10^{5}	1.5×10^{4}	3.2×10^{6}	$1.1{\times}10^{7}$	$3.2{\times}10^{5}$

Table 2.2: Numerical parameters of each component included in the initial conditions. $\epsilon_\star-$softening length of stars, ϵ_g-softening level of gas, $m_{\rm gas}-$mass of gas particles, $N_{\rm gas}-$number of gas particles, m_b-mass of bulge particles, N_b-number of bulge particles, m_d-mass of particles in stellar disc, N_d-number of particles in stellar disc, $N_{\rm gas,cor}-$number of gas particles in corona, $N_{\rm gas,d}-$number of gas particles in disc.

The model used in this work is calibrated to resolve the multiphase ISM of an isolated, non-cosmological MW-like galaxy. It self-consistently implements stellar feedback mechanisms that generate gaseous outflows and regulate star formation. The resulting star formation rate (SFR) remains stable at approximately $3\,M_\odot\,{\rm yr}^{-1}$, in agreement with the Kennicutt–Schmidt relation and observational data. The inclusion of stellar feedback

increases the gas depletion timescale from a few hundred Myr to several Gyr, providing a more realistic star formation history. This model was chosen as the "base" for our mock observations because it captures gas in multiple phases and includes physically motivated feedback processes. These feedback effects inject energy and momentum in regions co-spatial with newly formed stars, allowing for a self-regulated star formation cycle, an essential feature for reproducing observed galaxy behavior. We selected a snapshot from the middle (∼1 Gyr) of the simulation (which evolved for ∼ 2 Gyr), ensuring that the system has reached an approximate equilibrium state while avoiding early-time artifacts (see Fig. 2.2 for a visual impression of the system). This snapshot was used as an input to generate synthetic HI observations.

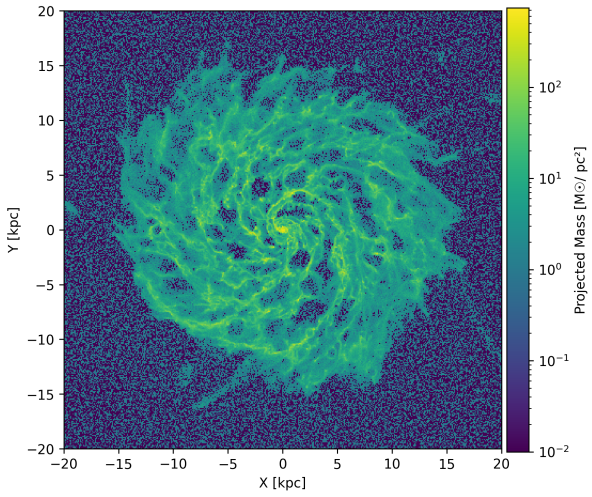


Figure 2.2: Projected total mass along the line of sight for a face-on galaxy. Resolution: 410×410 bins with a size of 97.56 pc.

In summary, this simulation provides a physically motivated environment that includes realistic stellar feedback, multiphase gas dynamics, and a gaseous halo, key ingredients for studying the neutral hydrogen distribution and kinematics in star-forming galaxies. These features make it an ideal base for producing mock observations and investigating the detectability and properties of extra-planar gas. In the next chapter, we detail how synthetic HI observations are generated using this simulation and the tools developed for analyzing them.

Parameter	Fiducial value	Units	Description
		Star formation	
ϵ	0.01	-	Star formation efficiency
ρ	100	cm^{-3}	Star formation density threshold
		Supernova feedback	
f_{SN}	1	-	SN energy relative to fiducial value
E_{SN}	10^{51}	erg	Fiducial energy per SN
N_{ngb}	64	-	Effective neighbor number
ΔN_{ngb}	1	-	Neighbor number tolerance
		Radiative feedback	
κ_{IR}	$10(Z/Z_{\odot})$	$\mathrm{cm}^2\mathrm{g}^{-1}$	Gas infrared opacity
$t_{\star,\mathrm{max}}$	5	Myr	Maximum stellar age for ionizing radiation
$\gamma_{\star}$	10^3	$\mathrm{L}_{\odot}\mathrm{M}_{\odot}^{-1}$	Stellar mass-to-light ratio
T_{phot}	1.7×10^4	K	Temperature of photoionized gas
$\langle h\nu \rangle$	17	eV	Average ionizing photon energy

Table 2.3: Default values of the parameters chosen for the modeling of physical processes included in the simulation with values and descriptions.

3

Generation of mock HI observations

To facilitate comparisons between galaxy formation simulations and observational data, we have developed a methodology to generate synthetic HI emission maps from simulation outputs. This involved modifying the publicly available software package MARTINI (Mock Array Radio Telescope Interferometry of the Neutral ISM; Oman, 2024) to process snapshots from AREPO galaxy formation simulations run with the SMUGGLE model. MARTINI is designed to produce mock observations of HI emission in galaxies, enabling direct comparisons between simulated and observed galaxy properties. We implemented extensions to MARTINI that allow the extraction of realistic HI spectral line data from SMUGGLE simulations, thereby bridging the gap between hydrodynamic simulations and observational radio data.

In this chapter, we present the workflow used to generate and analyze synthetic HI observations for Milky Way-like galaxies. Starting from the simulations, we produced mock 21 cm HI data cubes using the MARTINI software that served as the basis for a detailed investigation of the spatial and kinematic structure of neutral hydrogen in the simulated galaxies. We examined key properties such as HI morphology, velocity fields, and surface brightness profiles and compared them with observations to assess the physical realism of the simulation output.

The chapter is organized as follows. In Section 3.1, we describe the setup and functionality of MARTINI, including its input parameters and the resulting output. Section 3.2 details the data processing pipeline and the techniques used to extract mean-

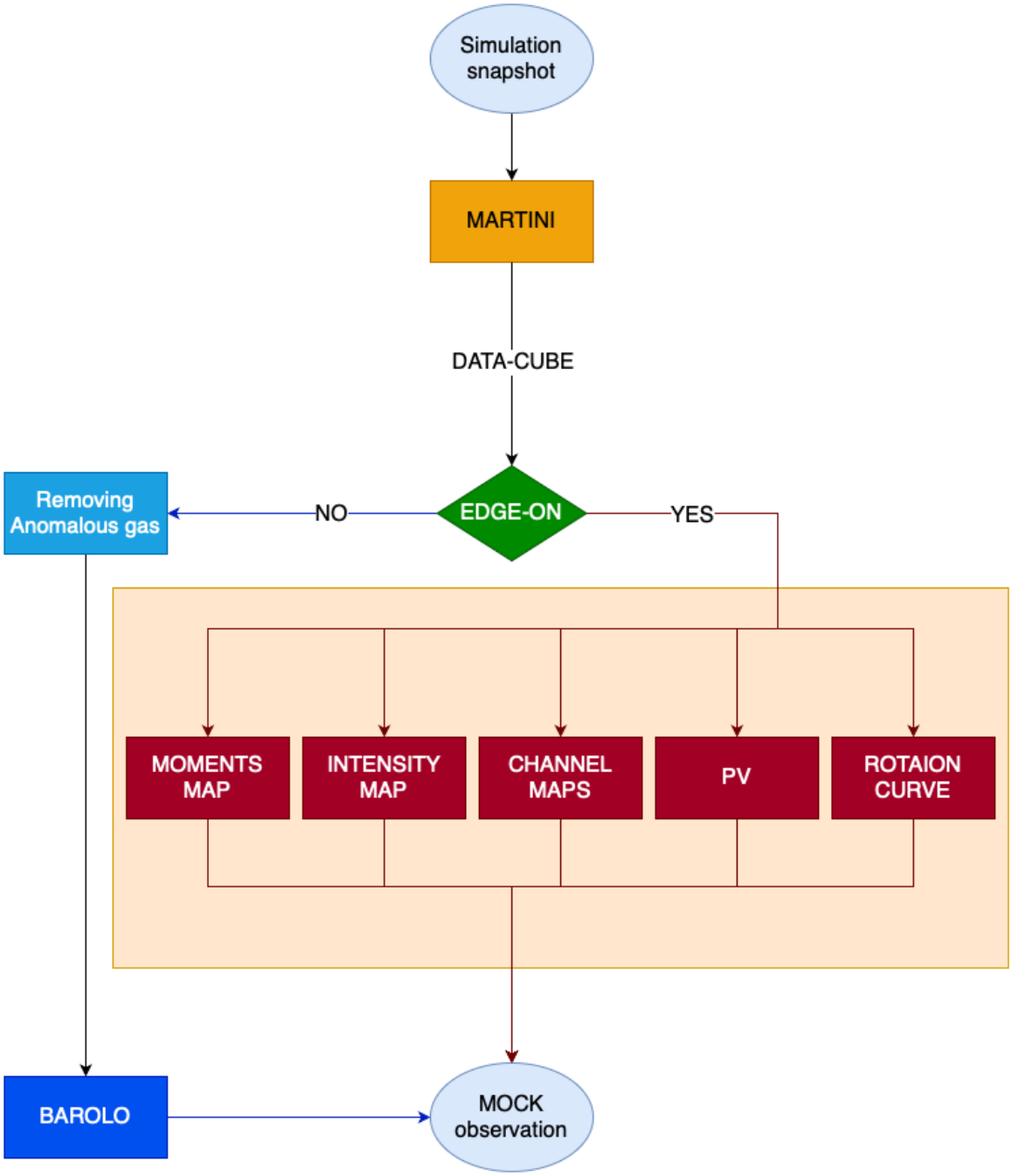


Figure 3.1: The schematic view of obtaining HI mock observations of hydrodynamic galaxy formation simulations with the MARTINI software package.

ingful physical properties from the mock data. Section 3.3 describes the method used to distinguish the EPG from the cold disc. Finally, in Section 3.4, we introduce ^{3D}BAROLO, a software tool used to derive rotation curves, radial velocity dispersion profiles, and surface brightness distributions from the data cubes.

3.1 The MARTINI software package

MARTINI (Oman, 2024) is a software tool developed in the Python programming language. It is designed to generate synthetic 21 cm HI emission line observations from hydrodynamic simulations of galaxy formation. The mock observing process within MARTINI is structured into modular components, each responsible for specific aspects of the generation of the synthetic observation, including data cube creation, source modeling, beam convolution, noise addition, and spectral line modeling. These modular components are implemented as class instances and passed as parameters to the main MARTINI class, allowing the creation of a synthetic observation through the sequential execution of designated functions[1].

The software generates three-dimensional data cubes representing the 21 cm HI emission, comprising two spatial dimensions and a velocity dimension derived from the Doppler shift of the emission line. In addition, the output data contain information that records the observational parameters (such as beam size, velocity channel width, pixel size, etc.) used as input by MARTINI for generating the mock HI observation of the simulation results. MARTINI is typically configured to interface with specific simulation output formats, including large-scale and widely used simulation projects such as EAGLE (Schaye et al., 2015), Illustris-TNG (Nelson et al., 2021), and Simba (Davé et al., 2019). To adapt MARTINI for use with simulation snapshots produced by the SMUGGLE model, modifications to the source code were required. The Illustris-TNG source file, which MARTINI uses in order to read simulation output data and convert it into HI emission, was adopted as a template due to its structural similarities (both types of simulation are run with the AREPO code), but direct compatibility was not possible, necessitating significant code alterations (see Section 3.1.2).

[1]More details on the usage and the implementation of the MARTINI package can be found at https://martini.readthedocs.io/en/latest/index.html

3.1.1 MARTINI parameters

To enable MARTINI to process the data and produce a result in the form of a data cube, it is necessary to specify certain input parameters that define the observational setup. The most important ones are:

- **source** - a description of a HI emitting object that contains the information of the positioning in the sky and distance from the observer.
- **datacube** - a description of the data cube parameters: number of pixels and channels, channel width, and pixel size.
- **beam** - a description of the beam, including its size and model.
- **noise** - a description of the noise, including its root-mean-square (RMS) amplitude.
- **sph_kernel** - a description of the smoothing kernel, used to project the flux density.
- **spectral_model** - a spectral model module that defines the shape of the line emission.

In the **source** module of MARTINI, we specify the required input source, namely the file containing the simulation data. Additionally, we define the spatial parameters, including the distance from the observer to the galaxy, the inclination angle, the position angle, and the rotation around the pole of the galaxy relative to the observer.

The observational setup is configured via the **datacube** module, which requires parameters such as the pixel count, the pixel size, the channel width, the number of spectral channels, and the central velocity of the observation. These parameters determine the spatial and spectral extent of the simulated observation, as well as its resolution.

To account for the finite resolution of real observations, MARTINI includes the **beam** module, which models the telescope's response. For single-dish telescopes, the point spread function (PSF) is approximated using a *GaussianBeam*, parameterized by the beam size (FWHM) and position angle. For interferometric observations, the beam corresponds to the synthesized beam, though a Gaussian approximation may still be used for simplicity in some cases.

Observational noise, arising from the detector electronics or the sky background, is incorporated via the **noise** module. This can be modeled using the *GaussianNoise* function, which requires the RMS noise level expected in the final data cube.

MARTINI supports input data from both particle-based simulations (e.g., SPH) and grid-based simulations (including the moving-mesh schemes like AREPO). In both cases, the **sph_kernel** module in MARTINI ensures consistent mapping of the simulation data to the observational domain by allowing users to specify the interpolation kernel. For SPH simulations, the original kernel is recommended, and MARTINI supports many of the most commonly adopted SPH kernels, such as WendlandC2, cubic spline, quartic spline, and Gaussian kernels. For moving-mesh codes, like AREPO, MARTINI suggests using the Cubic Spline kernel, with a smoothing length of 2.5 size of the cell radii. This is the choice that we have adopted in our analysis.

The emission line modeling in MARTINI is handled by the **spectral_model** module, which offers two approaches: a *DiracDeltaSpectrum*, mostly used as a debugging tool, and a *GaussianSpectrum* for lines with a specified velocity width. The latter is particularly useful for capturing thermal or turbulent broadening effects in the simulated spectra, and it is the adopted default choice for this thesis work.

3.1.2 The SMUGGLE source file

The first step to produce the mock HI observations that are analyzed in this work was to adapt MARTINI to read SMUGGLE simulation snapshots. The original MARTINI code is capable of reading AREPO data in the IllustrisTNG output format only. Therefore, it was necessary to update the MARTINI input routines to handle files produced by the SMUGGLE model. This involved adjusting the way particle data and gas properties are extracted from the snapshot files. For example, TNG source files are categorized into two classes of snapshots, distinguished by the physical quantities, referred to as fields, that they contain (e.g., pressure, chemical abundances, hydrogen fraction). These fields vary depending on the snapshot class. In contrast, SMUGGLE outputs are not divided into such categories: all snapshots consistently include the same set of fields, ensuring uniform data availability across the simulation. Moreover, SMUGGLE provides detailed chemical abundances, including an expanded breakdown of hydrogen into atomic and molecular fractions, which is used to determine the HI emission. This information is not available in the TNG output files, so the molecular fraction needs to be modeled on-the-fly using external libraries instead.

To infer the HI emission, we need to know the neutral hydrogen mass of each gas cell to model the emission source. Since this quantity is not provided directly in the snapshot file, it must be calculated from known fields, such as the total gas mass M, the hydrogen abundance X_{H}, and the fractions of neutral f_{neutral} and molecular hydrogen $f_{\mathrm{molecular}}$ of

each gas cell. After extracting the necessary fields, we can compute the atomic hydrogen fraction and, subsequently, the mass of atomic hydrogen in each cell as follows:

$$f_{\text{atomic}} = f_{\text{neutral}} * (1 - f_{\text{molecular}}) \ , \tag{3.1}$$

$$M_{\text{HI}} = M * X_{\text{H}} * f_{\text{atomic}} \ . \tag{3.2}$$

Figure 3.2 shows the step-by-step separation of the atomic hydrogen mass from the total gas mass. The left panel shows the total gas mass projected along the line of sight. The projection box has a size of 20 kpc. The second panel shows the projection of total hydrogen mass, including the contribution of both ionized and neutral plus molecular gas. The third panel shows after removing the ionized gas mass from the total hydrogen mass. The last panel shows only the atomic hydrogen, after removing the contribution of ionized and molecular hydrogen.

It is worth noting that several methods can be used to remove the contribution of molecular hydrogen from the HI emission. To test the robustness of the approach presented in Fig. 3.2, we compared it with an empirical method derived from observations of nearby galaxies (Leroy et al., 2013), as discussed in Appendix B.1. Since the differences between the methods were negligible, we chose to rely on the primary data, namely, the properties contained in the snapshot file, to estimate the molecular hydrogen fraction in each gas cell.

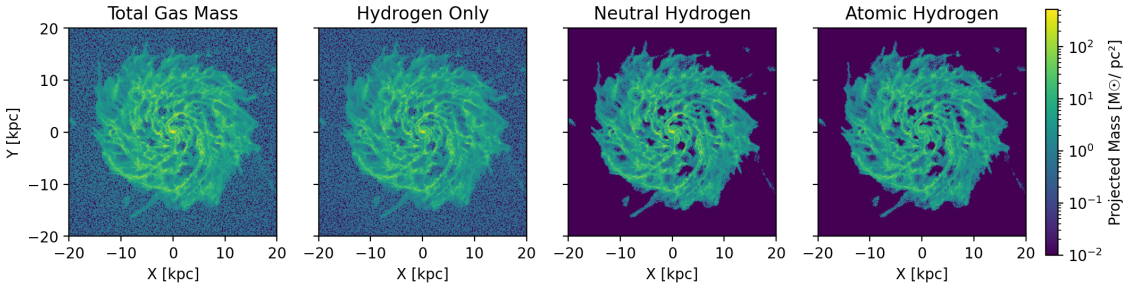


Figure 3.2: The figure shows four stages of inferring atomic hydrogen emission. Each panel presents binned (256×256) data of mass distribution integrated along the line of sight on each bin with a size of 156.25 pc. Left: the total mass of all gas. Mid-left: total hydrogen mass. Mid-right: mass of hydrogen after removing the contribution of ionized hydrogen. Right: mass after removing molecular hydrogen.

Finally, MARTINI requires the galaxy to be positioned at the center of the grid. While galaxy centers are directly provided in the Illustris-TNG outputs, this is not the case for the SMUGGLE simulations analyzed here. To address this, we centered the galaxy by subtracting half the size of the simulation box from the coordinates of each gas particle.

3.1.3 Creation of the HI data cube

Now that the simulation data can be processed, it is essential to define the appropriate observational parameters to obtain the HI data cube. To ensure consistent results and maximize similarity to real observations, we adopt parameters from actual observational campaigns. To this end, we establish the following key observational parameters:

- The minimum detectable flux density, which defines the sensitivity limit.
- The spatial resolution of the observations.
- The spectral resolution of the observations.
- The galactic parameters, such as distance and inclination.

For our study, we adopted the observational parameters of the edge-on spiral galaxy NGC891 (as reported in Oosterloo et al., 2007), located at a distance of 9.5 Mpc. This galaxy is considered a typical spiral galaxy, structurally and dynamically similar to the MW, making it an ideal reference for our mock HI emission analysis. Table 3.1 summarizes the observational parameters adopted in this work.

Inclination angle ($degree$)	90, 60, 30, 0
Distance (Mpc)	9.5
Beam size ($arcsec^2$)	33.2×23.9
Noise per channel ($mJy\ beam^{-1}$)	
no smoothing	0.09
with Hanning smoothing	0.06
Channel width ($\mathrm{km\,s^{-1}}$)	
no smoothing	8.2
with Hanning smoothing	16.4

Table 3.1: Observational parameters used to generate the synthetic data cube in MARTINI. The beam size, noise level, and channel width are based on observational data of NGC891, with the galaxy inclination (90° in the original observation) being the only parameter varied in our study. Both smoothed and non-smoothed data cubes are presented, with corresponding values for channel width and noise.

The first step in creating the data cube involves determining its spatial and spectral dimensions. This is achieved through preliminary analysis provided by MARTINI (see Fig. 3.3 and Fig. 3.4), which identifies the spatial extent and spectral channels containing

emission associated with the galaxy under examination. Spectral analysis clearly reveals emission in the velocity range from ∼350 to ∼1000 $\mathrm{km\,s^{-1}}$, centered on the systematic velocity $V_{sys} = H_0 D = 665\ \mathrm{km\,s^{-1}}$ obtained by MARTINI, where $H_0 = 70\ \mathrm{km\,s^{-1}\ Mpc^{-1}}$ is the Hubble constant and D is the distance to the galaxy from the observer. The velocity field map (left panel of the Fig. 3.4) clearly shows signal-emitting particles distributed within a 25 kpc radius from the center.

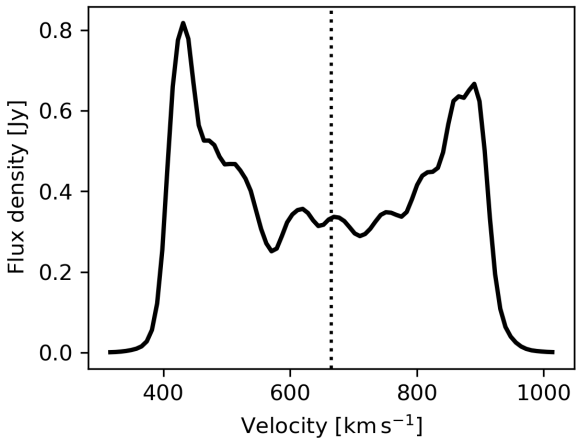


Figure 3.3: Global line profile showing the galaxy extends in spectral profile. Dotted line showing systemic velocity (665 $\mathrm{km\,s^{-1}}$).

After defining the data cube's spatial and spectral extent, we must establish the appropriate sampling resolution. Given the channel width, beam size, and spectral resolution (see Table 3.1), we apply the Nyquist-Shannon sampling theorem (Cárcamo et al., 2018) to determine the maximum allowable pixel size:

$$\Delta x \leq \frac{\theta_{\mathrm{beam}}}{2}\ , \tag{3.3}$$

where θ_{beam} is the telescope beam width and Δx is the spatial pixel size. This condition ensures that the synthesized image is sufficiently sampled to prevent aliasing effects. However, decreasing the pixel size increases the total number of pixels in the data cube, thereby raising computational costs. Since the chosen kernel (Cubic Spline) scales computationally as $\mathcal{O}(N)$, with $N \propto N_x \times N_y \times N_v$, it is crucial to balance resolution with performance. In practice, we selected a pixel size (2 arcsec) that satisfies the Nyquist condition while keeping the data volume and processing time computationally manageable.

Another essential component in simulating realistic observations is the addition of noise, which includes both instrumental noise and background radiation detected by the telescope. This step is crucial, as real galaxies are not observed in isolation. They

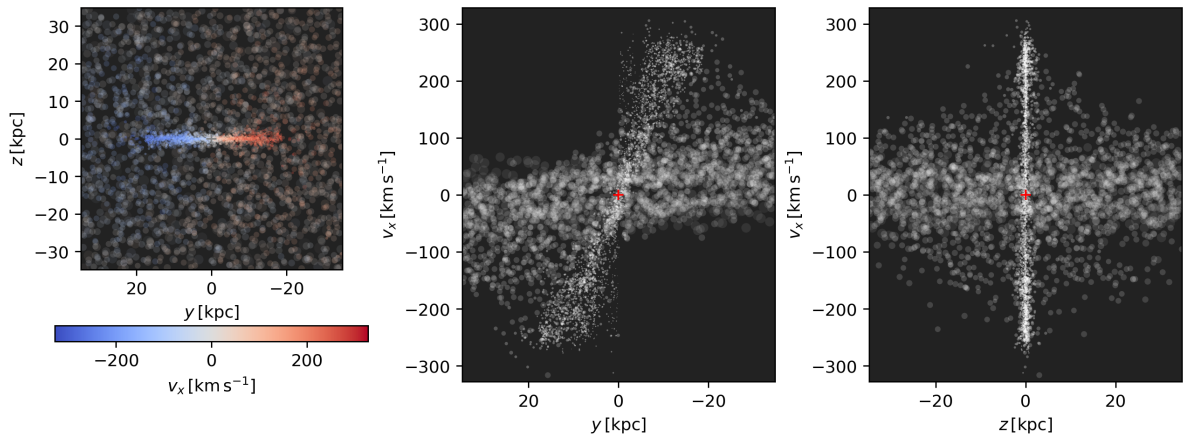


Figure 3.4: Preview image of the galaxy derived from the MARTINI HI data cube, showing the spatial and spectral extent of the galaxy. Left: Velocity field map. Mid: Velocity distribution along the major axis. Right: Velocity distribution along the minor axis.

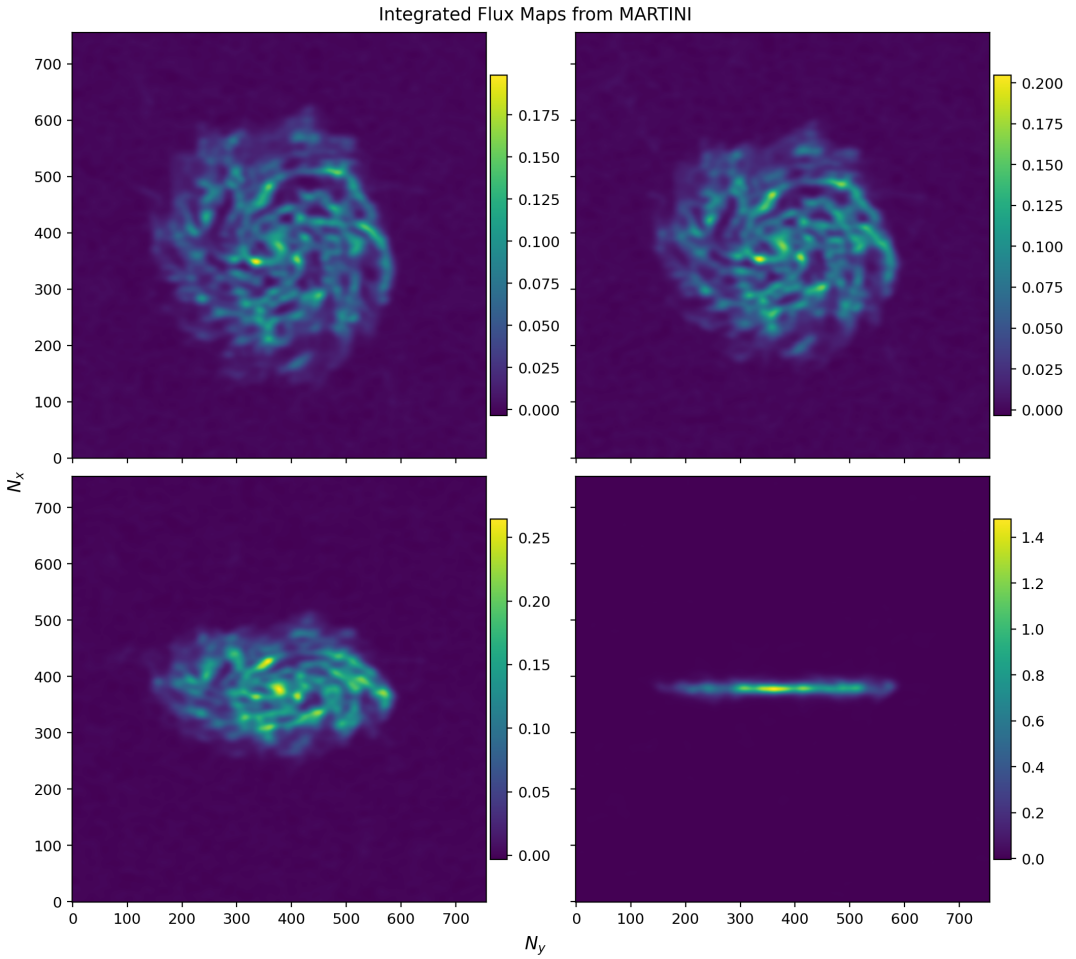


Figure 3.5: Data cubes produced with MARTINI for different galaxy inclinations. Top-left: 0°. Top-right: 30°. Bottom-left: 60°. Bottom-right: 90°. Color bars represent integrated flux level (Jy beam^{-1}).

exist within complex environments, and the instruments used for observation introduce intrinsic uncertainties. To replicate these effects, we added Gaussian noise with standard deviation values specified in Table 3.1, offering a statistically valid representation of the observational conditions.

Incorporating noise not only enhances the realism of the mock data but also ensures that the analysis techniques applied to simulated data are directly comparable to those used in real observations. This is particularly important when assessing the detectability of low-surface-brightness features, such as EPG, and in evaluating the robustness of data reduction and interpretation methods under realistic observational constraints.

The final data cube (see Fig. 3.5) has dimensions of 756×756 spatial pixels and 86 spectral channels, incorporating both noise convolution and beam-smearing effects appropriate for our observational setup. All data cube parameters are listed in Table 3.2.

N_x	N_y	N_v	Δx	V_{sys}	σ_{ISM}
756	756	86	$2''$	665 $\mathrm{km\,s^{-1}}$	8 $\mathrm{km\,s^{-1}}$

Table 3.2: Data cube parameters. N_x-number of pixels along the x-axis, N_y-number of pixels along the y-axis, N_v-number of channels, $\Delta x-$pixel size, $V_{sys}-$systemic velocity of the galaxy or central spectral channel, $\sigma_{ISM}-$standard deviation for Gaussian spectral model.

3.2 Data processing methods

Before starting the analysis of the mock data cube, it is necessary to process the data to extract scientific features. As a first step, it is useful to apply some smoothing along the velocity axis to reduce the noise level and suppress fluctuations in each channel of the data cube. To this end, we apply the Hanning smoothing (Springob et al., 2005), which is a standard procedure and is implemented as a weighted moving average defined as:

$$y_i = 0.25 \cdot x_{i-1} + 0.5 \cdot x_i + 0.25 \cdot x_{i+1} \ , \tag{3.4}$$

where x_i is the original value of the flux in a given pixel of the $i-$th velocity channel, x_{i-1} and x_{i+1} are the flux values in the neighboring velocity channels.

Moreover, to suppress spurious detection, masking with $3 \times \sigma_{rms}$ level has been applied:

$$\mathrm{Mask}(x, y, v) = \begin{cases} 1 & \text{if } I(x, y, v) \geq 3 \times \sigma_{\mathrm{rms}} \\ 0 & \text{otherwise} \end{cases} \ . \tag{3.5}$$

By applying this masking procedure, we exclude the flux values below a defined threshold (typically set relative to the noise level in each channel, $\sigma_{\rm rms}$), thereby removing low-significance emission and preventing noise from contaminating the analysis.

Now, it is possible to operate with these masked and smoothed data to derive some physical quantities related to the HI emission. Examples of such quantities are:

1. Moments map;
2. HI column density and total mass;
3. Velocity channel maps;
4. Position – Velocity diagram;
5. Rotation curve.

We describe the procedures to obtain and analyze them in more detail in the following subsections.

3.2.1 Moments map

One of the tools used to investigate the kinematics and dynamics of galaxy are moment maps. Their production is a way to simplify data to some key quantities (e.g., mean velocity, standard deviation). They are produced by processing a 3D data cube, and as a result a 2D maps is obtained (Thompson et al., 2017).

A moment-0 is a map showing integrated flux revealing the spatial distribution and is defined as:

$$M_0(x,y) = \int_{v_{\rm min}}^{v_{\rm max}} I(x,y,v)\, dv \ , \tag{3.6}$$

where M_0 is the integrated intensity along the line of sight (los), $I(x,y,v)$ is the intensity of each 3D data cube pixel, and dv is the velocity channel width. This type of map gives information on the spatial extent of the source and can be used to identify filaments, arms or outflows from the regions with noticeable flux.

The moment-1 map is an intensity-weighted velocity map, obtained as:

$$M_1(x,y) = \frac{\int v \cdot I(x,y,v)\, dv}{\int I(x,y,v)\, dv} \ , \tag{3.7}$$

where M_1 is the velocity field along the los, v is the velocity of the incoming flux. The moment-1 map reveals the bulk motion of gas as an averaged velocity field. From this we can derive kinematic characteristics that allow us to identify rotation by detecting a

coherent velocity pattern, trace inflows or outflows, and highlight non-circular motions through deviations in the velocity structure. In edge-on galaxies, symmetric velocity gradients are also evident.

A moment-2 map is an intensity-weighted velocity dispersion map defined as:

$$M_2(x,y) = \sqrt{\frac{\int (v - M_1)^2 \cdot I(x,y,v)\, dv}{\int I(x,y,v)\, dv}} \ , \tag{3.8}$$

where M_2 is the dispersion along the los, and M_1 is the velocity field along the same los. By obtaining it, we could measure turbulence or thermal broadening and trace feedback effects. For galaxies with an inclination of $\sim 90°$, we expect a high dispersion near mid-plane due to beam smearing and/or unresolved rotation.

3.2.2 HI column density map

From the mock HI data cubes, it is possible to create HI column density maps by integrating the emission along the velocity axis as a moment-0 map and converting the resulting flux into column density. Assuming optically thin emission, the HI column density $\mathcal{N}$ is calculated by integrating the surface brightness along over the velocity channels as (Chengalur et al., 2013):

$$\frac{\mathcal{N}(x,y)}{cm^{-2}} = 1.823 \times 10^{18} \int_{-\infty}^{\infty} \left(\frac{T_B(v)}{K}\right)\ d\left(\frac{v}{km/s}\right)\ , \tag{3.9}$$

where T_B is the brightness temperature of the source.

The brightness temperature of an astronomical source is defined under the Rayleigh–Jeans approximation to Planck's law as the temperature of a blackbody reproducing a specific intensity B_ν at frequency ν. In this limit, the brightness temperature is related to the specific intensity by

$$T_B = \frac{c^2 B_\nu}{2k_B \nu^2} = \frac{\lambda^2 B_\nu}{2k_B}, \tag{3.10}$$

where c is the speed of light, k_B is Boltzmann's constant, and $\lambda = c/\nu$ is the wavelength corresponding to the frequency ν.

The *flux density* S_ν received from the source is obtained by integrating the specific intensity over the source's apparent solid angle:

$$S_\nu = \iint B_\nu \, \mathrm{d}\Omega \ . \tag{3.11}$$

Substituting $B_\nu = 2k_B T_B/\lambda^2$ gives

$$S_\nu = \frac{2k_B}{\lambda^2} \iint_{\mathrm{src}} T_B \, \mathrm{d}\Omega \ . \tag{3.12}$$

If the source uniformly fills the telescope beam of solid angle $\Omega_{\rm bm}$, this simplifies to

$$S_\nu = \frac{2k_B T_B \Omega_{\rm bm}}{\lambda^2} \ , \tag{3.13}$$

solving for T_B yields

$$T_B = \frac{\lambda^2 S_\nu}{2k_B \Omega_{\rm bm}} \ . \tag{3.14}$$

For a telescope beam with a Gaussian profile and full width at half maximum (FWHM) axes θ_a and θ_b (measured in arcseconds), the beam solid angle is given by

$$\Omega_{\rm bm} = \frac{\pi \theta_a \theta_b}{4 \ln 2}. \tag{3.15}$$

Expressed in practical units, the brightness temperature T_B in K can be approximated as

$$\frac{T_B}{\rm K} \approx 1.360 \left(\frac{\lambda}{\rm cm}\right)^2 \frac{S_\nu}{\rm mJy} \left(\frac{\theta_a \theta_b}{\rm arcsec^2}\right)^{-1} . \tag{3.16}$$

Finally, the column density could be found from the flux per beam by substituting brightness temperature (3.16) in (3.9):

$$\frac{\mathcal{N}(x,y)}{cm^{-2}} = 1.104 \times 10^{21} \int_{-\infty}^{\infty} \frac{I(x,y,v)}{mJy} \left(\frac{\theta}{arcsec}\right)^{-2} d\left(\frac{v}{km/s}\right) \ , \tag{3.17}$$

where $I(x, y, v)$ is the intensity in each data cube cell, θ^2 is the area of the beam, and dv is the velocity channel width.

3.2.3 Total mass

Following the derivation of the column density map, the total HI mass can be calculated by integrating the column density over the observed area. This provides a direct link between synthetic observation and one of the key global properties of the galaxy.

The column density $N_{H\rm I}$ is typically obtained from the integrated flux in each pixel of the data cube using Eq. (3.9). Once the column density map is available, the total HI mass is calculated as:

$$M_{H\rm I} = m_{\rm H} \sum_i N_{H\rm I,i} \cdot A_{{\rm pix},i} \ , \tag{3.18}$$

where m_H is the mass of a hydrogen atom, $N_{H\rm I}$ is the column density in the i-th pixel, and $A_{{\rm pix},i}$ is the physical area of that pixel, in cm^2. In practice, the area per pixel can be computed from the pixel size (in arcseconds) and the distance to the galaxy, using the small-angle approximation. For a uniform grid, the equation can be simplified to:

$$M_{H\rm I} = m_{\rm H} \cdot \Delta A \sum_i N_{H\rm I,i} \ , \tag{3.19}$$

where ΔA is the constant physical area of one pixel.

In the data cube, if integrated flux densities are presented in units of Jy km s^{-1} per beam, the total HI mass can be inferred using the standard formula:

$$\frac{M_{\rm HI}}{M_\odot} \approx 2.36 \times 10^5 \left(\frac{\rm S_{int}}{\rm Jy\ km/s}\right) \left(\frac{D}{\rm Mpc}\right)^2 , \tag{3.20}$$

where D is the distance to the object and $\rm S_{int}$ is the integrated flux along the spectral channels. This equation assumes optically thin emission at the 21 cm line and is commonly used for both real and synthetic observations. In this work, the equation (3.20) was applied to ensure consistency with observational practices.

The inferred HI mass is a critical diagnostic for evaluating the realism of galaxy simulations. It provides a basis for validating the simulated gas content, star formation efficiency, and the effectiveness of feedback mechanisms. In this study, we isolated the anomalous HI component, often associated with extra-planar or halo gas, from the total HI distribution (see Section 3.3). We then quantified its relative mass fraction and compared it with observational estimates from real galaxies. This comparison helps assess whether the simulation accurately reproduces the complex gas dynamics and multi-phase structure observed in galactic halos.

3.2.4 Velocity channel maps

Velocity channel maps are used to visualize the spatial distribution of HI emission at specific line-of-sight velocities. These maps are generated directly from the mock data cube by selecting individual spectral channels, each corresponding to a narrow velocity interval.

We construct the velocity channel maps by extracting the intensity distribution at selected velocity channels across the cube. These channels are sampled evenly across the full velocity range to provide representative "snapshots" of the HI emission at different systemic velocities. Thus, each map shows how HI is spatially distributed at a specific velocity, making it possible to trace the motion of gas components across the galaxy.

Channel maps are particularly useful for identifying coherent kinematic structures, such as rotating discs, spiral arms, or anomalous gas flows. They also allow for qualitative assessment of the HI morphology at various velocities and serve as a complementary tool to position–velocity diagrams in the study of galaxy kinematics.

3.2.5 Position—velocity (PV) diagram

To analyze the kinematic structure of the HI gas in galaxies similar to the MW, we construct PV diagrams from mock HI data cubes.

To generate a PV diagram with correct spatial scaling, we first process the data cube by converting its units to physical values and applying appropriate normalization. As previously noted, for our standard choices in building the data cube (see Table 3.2), each spatial pixel spans 2 arcseconds, while the synthesized beam size along the minor axis is approximately 23 arcseconds. To match the spatial resolution of the mock observations, we spatially averaged the emission over groups of neighboring pixels within a single beam. This involves summing the emission from 11 adjacent pixels (since $23\,\mathrm{arcsec}/2\,\mathrm{arcsec} \approx 11$), effectively smoothing the data to the beam resolution. The integration is expressed as:

$$PV(x, v) = \int_{\delta S_{beam}} I(x, y, v)\; dS_{beam}\;, \tag{3.21}$$

where $I(x, y, v)$ is the intensity at position (x, y) and velocity v.

The PV diagram is constructed by extracting a slice along a chosen spatial axis (e.g., the galaxy's major or minor axis), where each point represents the integrated HI intensity as a function of position and line-of-sight velocity. This yields a 2D representation that encodes both the spatial distribution and the kinematic signatures of the gas.

This method enables us to trace the galaxy's velocity structure, identify rotational features, and detect non-circular motions or disturbances in the gas. PV diagrams also provide a diagnostic tool for comparing kinematic properties across different models or observational configurations. Moreover, they serve as a starting point to determine the galaxy rotation curve (i.e., the gas rotational velocity as a function of galactocentric radius), which yields key information on the mass distribution of such objects.

3.3 Cold gas separation

A galaxy simulated with the SMUGGLE model generally consists of several components, including a CGM or gas halo surrounding the galactic disc. The cold phase of the latter manifests as extended high-velocity dispersion features in the HI moment maps. For studying the kinematics of galaxies with inclination angles below 90°, it is beneficial to remove this gas component or at least minimize its influence in subsequent analysis. A spectral fitting method was implemented to achieve this objective (Fig. 3.6).

As illustrated in the figure, the method analyzes the spectral shape of each spatial pixel by fitting it to a Gaussian profile (Eq. 3.22), centered on the emission peak within

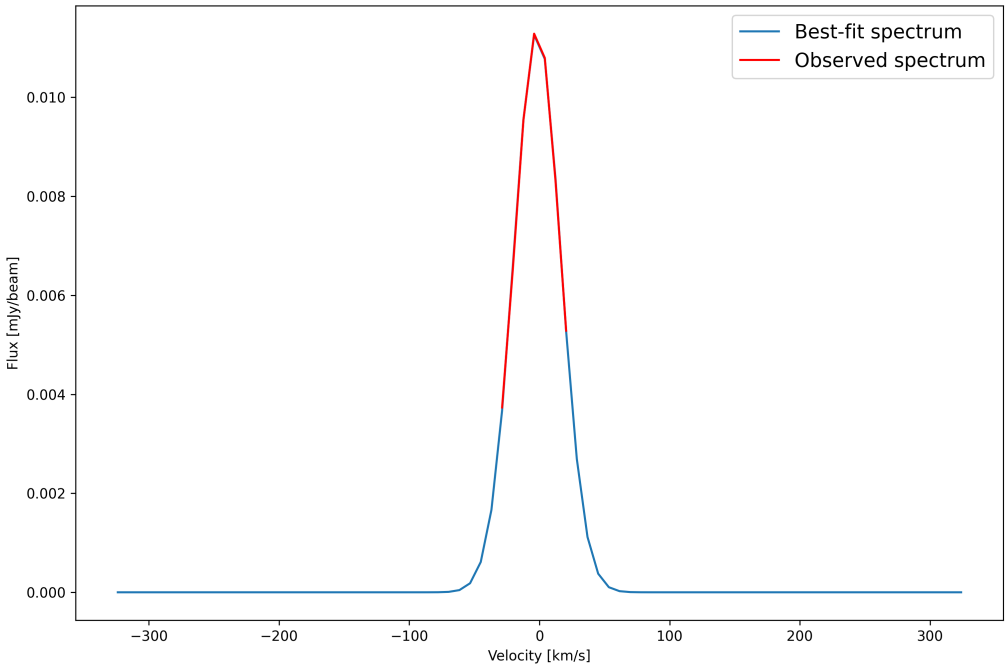


Figure 3.6: Spectral fit of arbitrary pixel in galaxy (x=378,y=410). Red line: observed spectral shape above 30% of the peak flux. Blue line: Gaussian best fit to the observed data. In this example, the spectral fit has been taken from the mock observation of the galaxy with 0° inclination.

that pixel. The fitting procedure considers only flux values exceeding 30% of the peak intensity and adopts the following functional form:

$$G(v) = A \exp\left[-\frac{(v - v_0)^2}{2\sigma^2}\right], \tag{3.22}$$

where A represents the amplitude of the Gaussian, v_0 denotes the central velocity, and σ corresponds to the velocity dispersion of the line profile. To maintain physical validity, fits producing unrealistically large velocity dispersion values were rejected. The Gaussian model was identified as representing the cold disc component, while the residuals were classified as anomalous gas. This method effectively separates the high-dispersion CGM features from the disc kinematics, facilitating more accurate analysis of cold gas dynamics while preserving the extra-planar component for independent investigation. The efficacy of this approach is particularly demonstrated by the cleaner velocity dispersion maps of the cold disc obtained after the removal procedure (see Fig. 3.7).

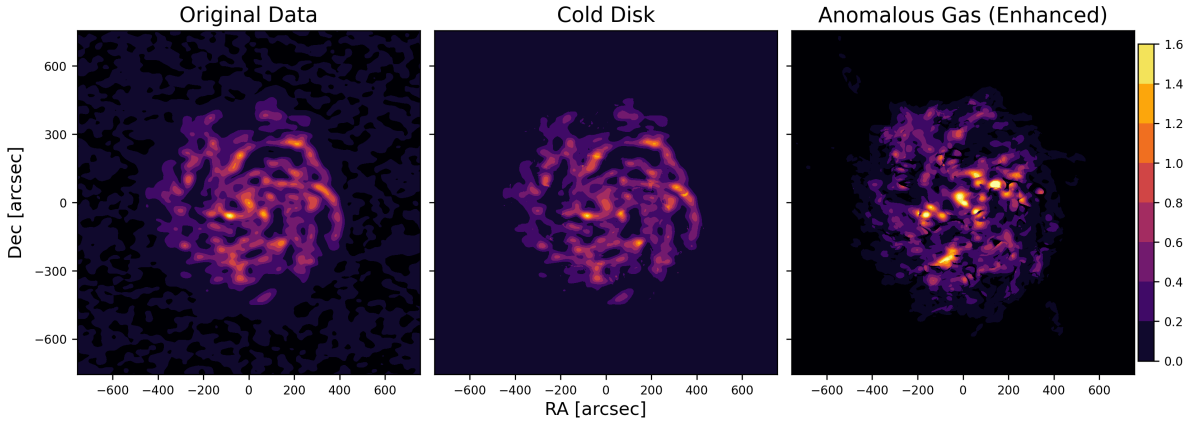


Figure 3.7: The figures show integrated intensity maps of a galaxy with an inclination of 0°. Left: Original data cube. Mid: Cold gas disk, obtained from data cube. Right: The CGM, gas surrounding the galactic disc. The cold gas and original data have the same normalization, while the one of the anomalous gas has been enhanced by a factor of 7 for better visibility. The color bar represents the integrated flux (Jy beam^{-1}).

3.4 ^{3D}BAROLO

A more advanced way to analyze data cubes to derive gas kinematics is by fitting a tilted-ring model (Rogstad et al., 1974) to the observations. This method is used in the ^{3D}BAROLO software (Di Teodoro and Fraternali, 2015). ^{3D}BAROLO breaks the galaxy into rings and fits rotation and other properties for each ring. This is especially useful for galaxies that are not perfectly edge-on (i.e., with inclination angles different from 90°), where projection effects can make the velocity field harder to interpret. Using this approach, we can get a more accurate estimate of the galaxy's rotation curve and other kinematic properties.

The model used in ^{3D}BAROLO describes the galaxy as a set of concentric rings with non-zero thickness, forming a three-dimensional disc structure. Each ring is characterized in a six-dimensional domain: three dimensions for spatial position and three for velocity components. The resulting emission is a three-dimensional projection, consisting of two spatial dimensions and one spectral dimension. The comparison between the model and the observed data is then performed ring by ring. The best-fit parameters for each disc ring are determined using the Nelder–Mead optimization method (Nelder and Mead, 1965), which minimizes the difference between the model and the data through iterative steps (see Appendix C).

^{3D}BAROLO software accepts some input parameters (for example: inclination angle, position angle, height of the disc, etc.) as initial conditions for the model for the fit-

ting procedure. If no initial parameters are provided, 3DBAROLO can estimate them automatically. The fitting process proceeds through the following main steps:

1. Disc model;
2. Convolution;
3. Normalization;
4. Residuals.

Each step is described in the subsections below.

Disc model

3DBAROLO uses a 3D tilted-ring model (Rogstad et al., 1974) to fit the observational data. The model space is filled randomly with emitting gas using a stochastic function. Each ring, with radius R and width W, is described by the following parameters:

- Coordinates of the center (x_0, y_0)
- Systemic velocity $V_{\rm sys}$
- Inclination angle i with respect to the observer
- Position angle ϕ of the major axis on the receding side of the galaxy, measured counterclockwise from north in the sky
- Rotation velocity $V_{\rm rot}$
- Velocity dispersion $\sigma_{\rm gas}$
- Face-on gas column density Σ
- Scale height of the gas layer z_0

Each ring is populated with gas clouds distributed in cylindrical coordinates, defined by a radial distance R_c, an azimuthal angle θ_c, and a height z_c above the disc plane. Radial and azimuthal coordinates are assigned from a uniform distribution, while height z_c is drawn from a specified vertical distribution. Available vertical distributions include Gaussian, sech^2, exponential, Lorentzian, and box profiles. The distributed gas is then projected according to the specified position angle and inclination.

Once the positions of the clouds are determined, their velocities are calculated by combining systemic, rotational, and random motions. The velocity profile of each cloud

is modeled as a Gaussian centered on the mean velocity, with a total dispersion $\sigma^2 = \sigma^2_{\rm gas} + \sigma^2_{\rm instrumental}$, where $\sigma_{\rm gas}$ is the intrinsic gas dispersion and $\sigma_{\rm instrumental}$ accounts for instrumental broadening.

The velocities are then discretized and recorded in a data cube with the same size as the observed data cube. By default, 3DBAROLO sets instrumental broadening $\sigma_{\rm instrumental} = W_{\rm ch}/\sqrt{2\log 2}$, where $W_{\rm ch}$ is the width of a spectral channel in the data cube. However, users can provide their own value for the spectral resolution if desired.

Convolution

Since the "raw" model data cube is generated at nominal resolution, it must be convolved to match the resolution of the observed data cube. This is done using the observational point spread function (PSF), modeled as a two-dimensional Gaussian function applied to each spectral channel.

The Gaussian PSF is defined by three parameters: the beam sizes along the two spatial axes (a, b) and the position angle ϕ. The Gaussian kernel is given by:

$$k(x,y) = \frac{1}{2\pi\lambda_\kappa\lambda_\eta}\exp\left[-\frac{1}{2}\left(\frac{\kappa^2(x,y)}{\lambda^2_\kappa} - \frac{\eta^2(x,y)}{\lambda^2_\eta}\right)\right], \tag{3.23}$$

where (x, y) are the spatial offsets from the beam center, and (κ, η) are coordinates in the rotated frame aligned with the beam axes:

$$\kappa = x\sin\phi - y\cos\phi,$$
$$\eta = x\cos\phi + y\sin\phi.$$

The full width at half maximum (FWHM) values a and b along the principal axes are related to the standard deviations by:

$$a = \sqrt{8\log 2}\,\lambda_\kappa,$$
$$b = \sqrt{8\log 2}\,\lambda_\eta,$$

and the values of a, b, and ϕ are taken from the observed data cube. To improve computational efficiency, 3DBAROLO performs the convolution by using fast Fourier transforms (FFTs), which significantly reduce the processing time.

Normalization

Normalization allows the code to remove one parameter from the fit: the surface density Σ of the gas. There are two types of normalization that can be performed by 3DBAROLO: (i) pixel-by-pixel and (ii) azimuthally averaged.

In the pixel-by-pixel normalization approach, the model is adjusted such that the column density maps of the model and the observations match at each spatial pixel. This means that the integral over the spectral axis (i.e., the total flux) at each pixel in the model cube is forced to equal that of the corresponding pixel in the observed cube, yielding identical column densities. This technique allows the model to capture non-axisymmetric features in the surface brightness distribution, thereby minimizing the influence of localized structures such as clumps or holes on the global fit (see, e.g., Lelli et al., 2012).

Alternatively, in azimuthally averaged normalization (the default method in 3DBAROLO), the model is scaled to match the average flux in each radial ring. While Di Teodoro and Fraternali (2015) recommend using pixel-by-pixel normalization in most cases due to its robustness against asymmetries, we adopt the azimuthally averaged method in this study. This choice is justified by the fact that both approaches yield consistent results for our data, with no significant impact observed on the derived kinematic parameters (see Appendix C).

Residuals

3DBAROLO calculates residuals between the model fit and data on a pixel-by-pixel basis. To do so, it computes a quantity F that measures how well a model fits the data and is defined as the weighted sum of the residuals over all considered pixels:

$$F = \frac{1}{n} \sum_{i=1}^{n} \Delta r_i \, w(\theta_i), \tag{3.24}$$

where n is the number of considered pixels, Δr_i are the residuals, and $w(\theta_i)$ is a weighting function. A pixel is considered only if its flux is larger than zero or if it belongs to an identified region of the galaxy. The weighting function is defined as $w(\theta) = |\cos\theta|^m$, where θ is the azimuthal angle and $m = 0, 1, 2$. In the previous formula, $m = 0$ represents uniform weighting and $m \neq 0$ weighting gives priority to the emission along the major axis ($\theta = 0$).

The residuals Δr can be computed in different ways:

$$\Delta r = \frac{(M - D)^2}{\sqrt{D}} \ , \tag{3.25a}$$

$$\Delta r = |M - D| \ , \tag{3.25b}$$

$$\Delta r = \frac{|M - D|}{M + D} \ , \tag{3.25c}$$

where M and D are the flux values of the model and the data, respectively. Eq. (3.25a) corresponds to a χ^2-like residual. If $D = 0$, it is replaced by the root mean square (RMS) value of the data cube. Eq. (3.25b) is appropriate for regions with intermediate flux, while equation (3.25c) gives more weight to regions with faint and diffuse emission.

Additional features

There are also other features available in ^{3D}BAROLO. In cases where the position of the galaxy (or galaxies) is unknown, a *Source detection* algorithm is implemented (Whiting, 2012), which can identify all the sources in the data cube and automatically begin fitting them.

To obtain a good fit, ^{3D}BAROLO builds a mask that identifies the regions attributable to the galaxy. By default, this mask is generated from the regions identified by the source finder algorithm. Alternatively, the data cube can be smoothed by a factor n (default is 2), and only the regions with emission higher than m times (default is 3) the $\sigma_{\rm rms}$ of the smoothed data cube are considered.

Another important feature is the automatic estimation of initial guesses for fit parameters. For instance, the galaxy center is determined as the flux-weighted average position of the source. The systemic velocity is chosen as the midpoint between the two velocities where the flux is greater than 20% of the peak of the global line profile. The position angle is estimated as the line that maximizes the velocity gradient along the line of sight. The inclination angle is derived by fitting the model column density map—smoothed to the same spatial resolution as the data—to the observed map. The rotation velocity is estimated from the inclination-corrected half-width of the global line profile at 20% of its peak value. The velocity dispersion and the disc thickness are set to default values of 8 $\mathrm{km\,s^{-1}}$ and 150 pc, respectively. In most cases, these initial guesses yield a good fit. However, the inclination angle is generally the most uncertain parameter.

Unphysical discontinuities can occur when kinematic and geometrical parameters are fitted simultaneously. To mitigate this, a two-step fitting procedure is implemented. First, all parameters are fitted together. In the second step, the geometrical parameters are fixed to a chosen functional form, and only the circular velocity is fitted. ^{3D}BAROLO performs this *parameter regularization* automatically: in the second fit, only the rotation velocity and velocity dispersion are left free.

Finally, the estimation of errors is performed using a Monte Carlo method, since no direct error computation is possible. Random Gaussian samples are drawn around the minimum found by the minimization procedure. The errors are then defined as the

range over which the residuals, approximated as a quadratic function, increase by n percent (default is 5%) from the minimum. This method yields error estimates that are in good agreement with those derived from more traditional techniques (Di Teodoro and Fraternali, 2015).

We used 3DBAROLO to derive radial profiles of rotation velocity (VROT), velocity dispersion (VDISP), surface brightness (DENS), and scale height (Z0) of the cold disc component for galaxies with inclinations of 0°, 30°, and 60°. These kinematic parameters were set as free variables during the fitting procedure. The remaining input parameters were chosen to avoid biasing the fits toward any extreme values. The radial spacing (RADSEP) of the rings was selected to ensure that the resulting measurements were statistically independent, with each ring covering a region larger than the beam size. The number of rings (NRADII) was set to fully sample the galaxy without leaving significant regions unfitted. Geometrical parameters such as inclination (INC), position angle (PA), and distance (DISTANCE) were specified based on the known properties of the model galaxy. We adopted a Gaussian vertical profile (LTYPE = 1) for the disc emission and minimized residuals using the absolute difference between the model and data (FTYPE = 2). The weighting function applied a $\cos\theta$ dependence (WFUNC = 1), giving greater weight to pixels near the major axis. To improve model convergence, we employed the default two-stage fitting procedure mentioned above: first determining the geometric parameters and then fitting the velocity dispersion and rotation velocity in a second pass. Additionally, masking was used to exclude pixels not associated with the galaxy based on a defined flux threshold. The following configuration summarizes the key parameters used in out 3DBAROLO fits:

```
NRADII      16
RADSEP      30
XPOS        377.5
YPOS        377.5
PA          270
INC         60    # or 0 or 30
Z0          5.0
VRAD        0
VDISP       15
FREE        VROT,Z0,VDISP,DENS
LTYPE       1      # Gaussian vertical profile
FTYPE       2      # Residuals: |model - observation|
```

```
DISTANCE    9.5
MASK        SEARCH
THRESHOLD   2e-4
LOGFILE     true
WFUNC       1       # Weighting: |cos({$\theta$})|
TWOSTAGE    true
```

4

Results

In this chapter, we present the results of our mock observation process and analyze them using the same techniques applied to real galaxy data. However, a key distinction is that, in this case, we have complete knowledge of all the intrinsic properties of the simulated galaxy. This methodology provides the most realistic framework for comparing real observational data with numerical simulations.

Direct comparisons between simulation outputs and telescope observations can be misleading because of fundamental differences between the two. Simulation data lack many of the complexities related to astronomical observations, such as spectral line shifting, contamination of foreground and background objects, instrumental noise, and measurement uncertainties. By simulating the observational process, we can bridge this gap, testing the validity of our numerical models and allowing us for a more accurate interpretation of real observations.

The goal of this chapter is to examine how observable signatures, such as column density distributions, velocity structures, and flux patterns, vary with galaxy inclination and how accurately they reflect the underlying physical conditions of the simulated galaxy. This analysis allows us to validate the simulation results via the methods commonly used in radio astronomy and enhances our ability to interpret real observational data using insights derived from numerical simulations.

The chapter is structured as follows. We begin with the analysis of the edge-on galaxy (90° inclination). This includes a detailed examination of moment maps (moment-

0, moment-1, and moment-2) to visualize the distribution of flux, the velocity field, and the velocity dispersion across the disc. We then present the column density map, highlighting key morphological features such as holes, filaments, and asymmetries. Next, we investigate velocity channel maps and position–velocity (PV) diagrams to explore the galaxy's kinematics and search for evidence of non-circular motions.

The same set of analyses is applied to galaxies at different inclinations: intermediate angles (60° and 30°) and a face-on configuration (0°). Finally, we introduce the use of ^{3D}BAROLO to construct tilted-ring models from the mock data. This modeling yields radial profiles of rotation velocity, velocity dispersion, and surface brightness, which we compare across the different inclinations.

Through this analysis, we aim to assess the fidelity of synthetic observations in reproducing observational features and to evaluate the performance of widely adopted analysis tools such as ^{3D}BAROLO under idealized yet realistic conditions.

As a reminder, the synthetic data cubes were generated by applying the MARTINI software package (see Section 3.1) to a SMUGGLE simulation snapshot, using the observational parameters listed in Table 3.1. The galaxy was placed at a distance of 9.5 Mpc and rendered at the four inclinations mentioned above (90°, 60°, 30°, and 0°). The root-mean-square (rms) noise level is 0.09 mJy per channel in the unsmoothed data and decreases to 0.06 mJy per channel after Hanning smoothing. The spectral resolution corresponds to a channel width of 8.2 km s^{-1}, increasing to an effective resolution of 16.4 km s^{-1} post-smoothing.

4.1 Edge-on galaxy

In this section, we present and analyze the moment maps (Fig. 4.1) derived from mock observations of HI emission in the edge-on (90°inclination) MW-like galaxy. The maps reveal significant details about the gas distribution and kinematics in both the galactic disc and the halo.

The left panel of Fig. 4.1 shows the integrated HI emission distribution. The contours are defined logarithmically, with the lowest contour set to three times the rms noise level, $\sigma_{\rm rms} \times \Delta v = 6.88$ mJy beam^{-1} km s^{-1}, where Δv is the velocity channel width. The highest contour corresponds to 95% of the peak emission detected on the map. To obtain surface brightness units, the flux values were multiplied by the channel width.

The HI emission peaks strongly near the galactic disc, consistent with the high concentration of neutral gas within the disc plane, which extends radially to approximately $\sim$ 25 kpc (500″). In addition to the disc emission, we also detect substantial vertical

extensions, with emission reaching up to ∼ 3.5–4 kpc (75″–85″) above and below the mid-plane. Numerous small clumps are detected outside the main disc structure. These morphological features will be described in more detail below in Section 4.1.1.

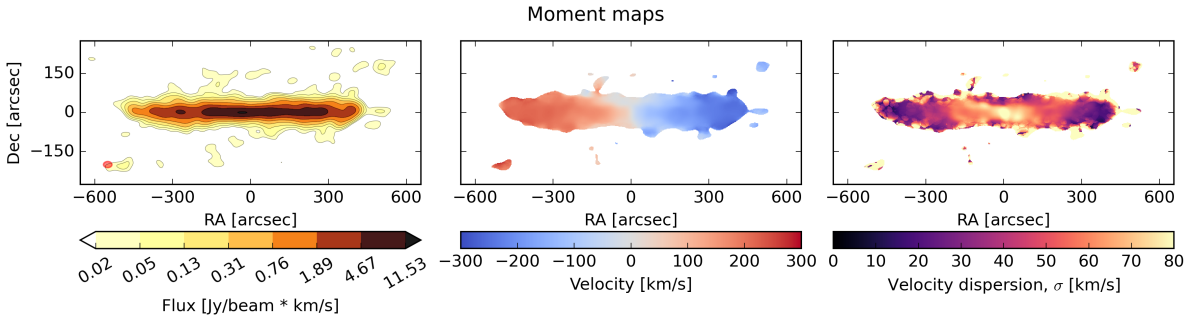


Figure 4.1: Moments maps of the edge-on galaxy (90° inclination). Left: Integrated intensity map. Middle: Velocity map. Right: Velocity dispersion map. Red circle: beam size (33.2 × 23.9 arcsec^2).

The middle panel of Fig. 4.1 displays the velocity field (moment-1 map), which shows the characteristic symmetric distribution expected from rotating disc galaxies. The velocity field reaches maximum values of ±250 km s^{-1}, with the zero-velocity point coinciding with the galactic center. This symmetric pattern confirms the regular rotation of the HI disc, with no significant signs of kinematic disturbances or asymmetries.

The right panel presents the velocity dispersion map, revealing distinct spatial variations. The galactic outskirts ("wings") show low velocity dispersion, typically below 20 km s^{-1}, indicative of quiescent gas in the outer disc and halo regions. In contrast, the central 8 kpc region exhibits significantly higher dispersion values of 50–80 km s^{-1}, which is related to the projection effects of the edge-on galaxy.

4.1.1 Column density distribution

Figure 4.2 displays the column density, derived by converting the observed surface brightness (Eq. 3.9). The contour levels have been chosen logarithmically, with the lowest contour equal to $3 \times \sigma_{rms,HI}$ where $\sigma_{rms,HI} \approx 0.86 \times 10^{18}$ cm^{-2} and the highest is 95% of the maximum detectable value.

The maximum column density reaches $\sim 1.4 \times 10^{22}$ cm^{-2}, and the inferred total mass from the data cube (Eq. 3.20) is $4.95 \times 10^{10} M_\odot$, which is consistent with the HI mass in the simulation (see Appendix A). The black line shows the concentration of 90% of the total mass from the data cube, around $4.89 \times 10^{10} M_\odot$. It is concentrated in a thin disc with a vertical extent of ∼1.1 kpc or ∼ 25″.

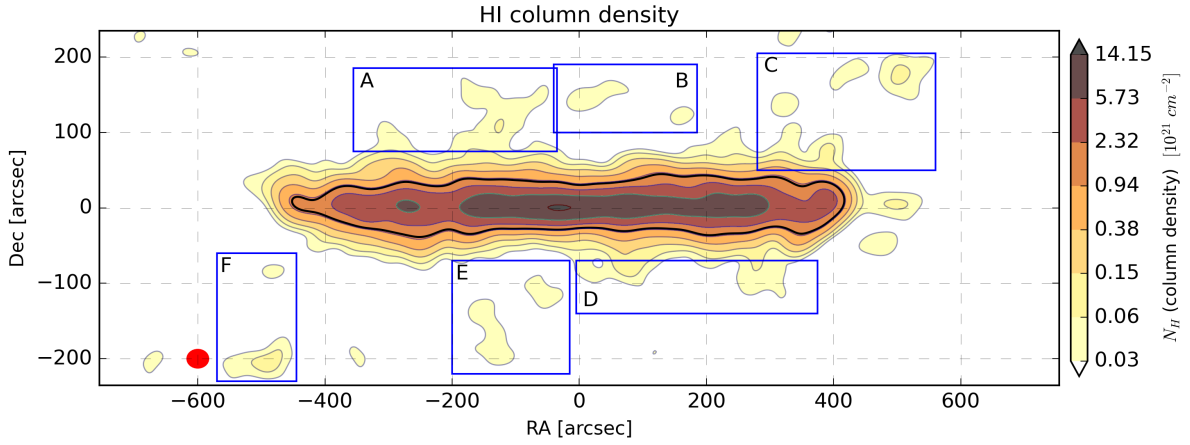


Figure 4.2: Column density of HI emission of the edge-on galaxy. Black solid line: level at which data cube contains 90% of total mass. Red circle: beam size (33.2×23.9 arcsec^2).

There are morphologically irregular features connected to the disc: region A, located above the disc on the receding side, and region D, found below the disc on the approaching side. Regions C and F may be indicative of rotational outflows. In contrast, the features labeled A and C are likely associated with energetic processes within the disc itself. Specifically, region A may represent a manifestation of the galactic fountain mechanism (Shapiro and Field, 1976), where hot gas, expelled by supernova activity, cools and falls back onto the disc in a fountain-like cycle.

This vertical distribution is further quantified in Fig. 4.3, where we analyze the gas distribution across four quadrants (see Appendix B.2). The average vertical gas distribution extends to approximately 3.5 – 4 kpc ($\sim 1.5'$) and has a symmetrical distribution in all directions.

4.1.2 Position – velocity diagrams

Figure 4.4 shows us the velocity distribution along the major axis obtained by integrating the central pixels within the beam size (Eq. 3.21). The contour levels are obtained logarithmically and lie in the range between three times $\sigma_{rms,PV} \sim 0.6$ mJy beam^{-1} and 95% of the peak emission detected on the PV diagram (~ 1.4 Jy beam^{-1}).

In the central part, we can see the velocity distribution from -300 km s^{-1} to +300 km s^{-1} which is followed by flattening of the velocity at larger radii, which is the typical behavior for galactic rotation, influenced by dark matter halos. Also, there are some inhomogeneities in the diagram, emission from the region with low velocity (∼140–190 km s^{-1}) at the approaching side (430″–540″), and a very small disturbance at the receding side with velocity ∼50–100 km s^{-1}(-360″ – -320″).

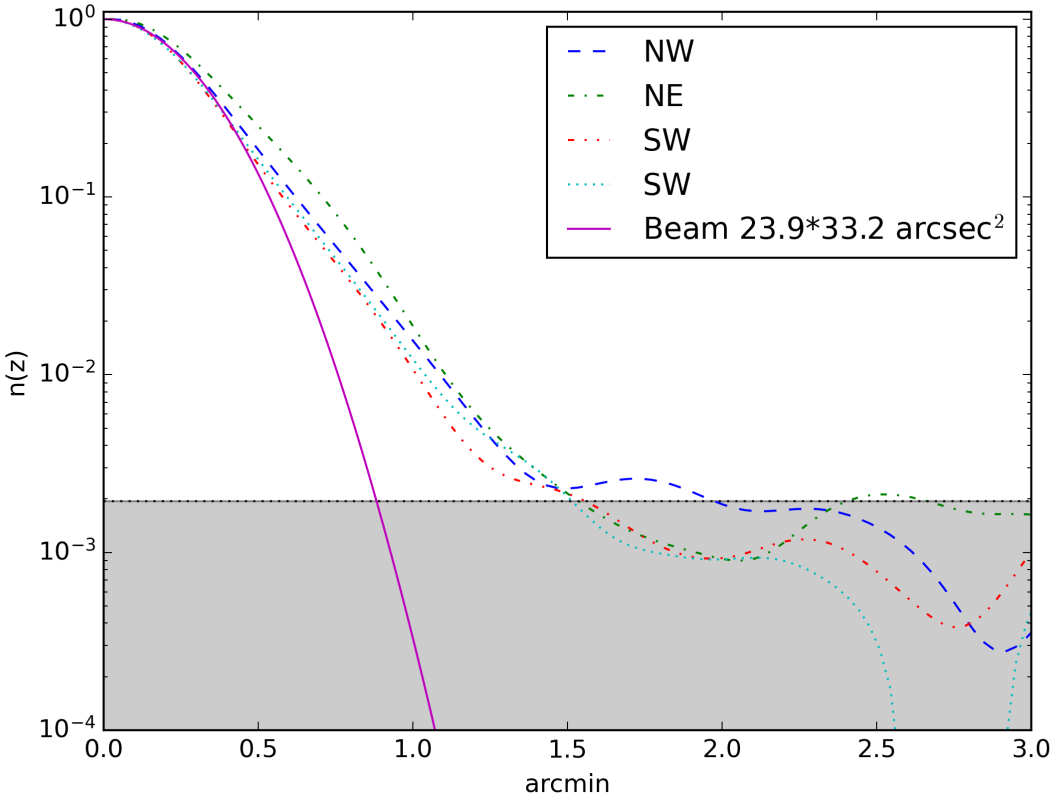


Figure 4.3: Vertical distribution of HI. The distribution is taken as the mean of HI column density in 4 quadrants, and normalized to the peak value. The purple line is the beam size. The grey box is the region below the $\sigma_{rms,HI}$ level.

In the PV diagram along the minor axis (Fig. 4.5), the velocity is distributed in the range between [-300 km s^{-1} – +300 km s^{-1}]. The diagram overall looks symmetrical, except for irregular velocity distribution below the disc (-100″– -70″), corresponding to some disc extension along the minor axis.

In Fig. 4.6 we can see the PV diagrams along the major axis at different distances from the disc (1.6 and 3 kpc above and below). These diagrams reveal an irregular velocity distribution, indicative of extra-planar gas components. A detectable signal is present even at $\sim$ 3 kpc from the disc, although it is significantly weaker than that along the disc. It reveals the presence of gas at such an offset from the disc, with velocities lower than observed on the disc, showing some "lagging" component outside the disc.

Additional minor axis PV slices were obtained at projected distances of $\pm 2'$ and $\pm 4'$ (5.6 and 11.2 kpc, respectively) from the galactic center, in both directions (Fig. 4.7). These slices show asymmetric extension depending on the side of the galaxy, with the

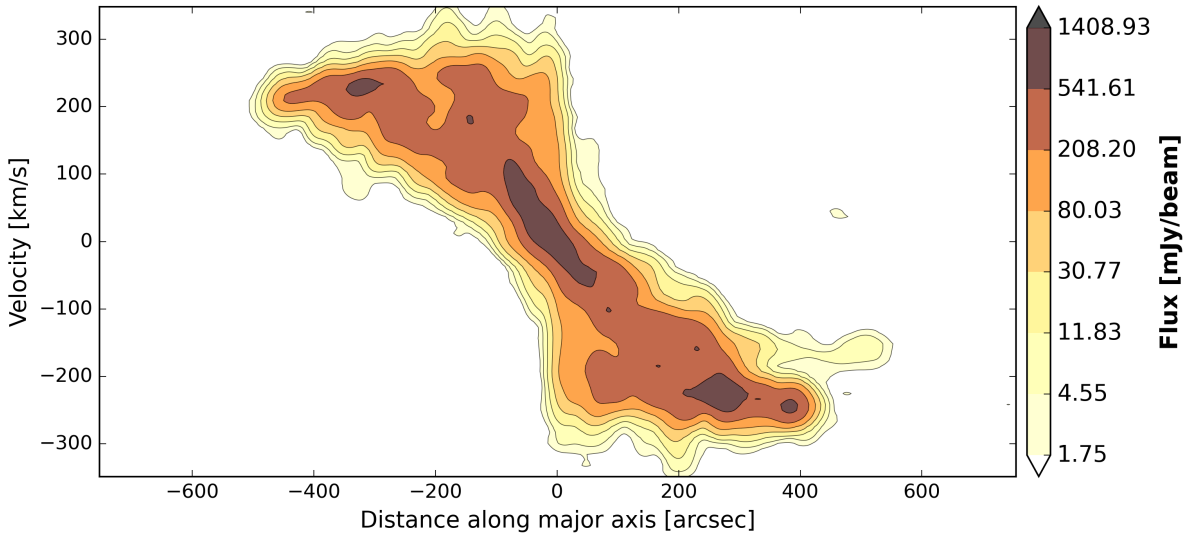


Figure 4.4: Position–velocity diagram along the major axis for edge-on galaxy.

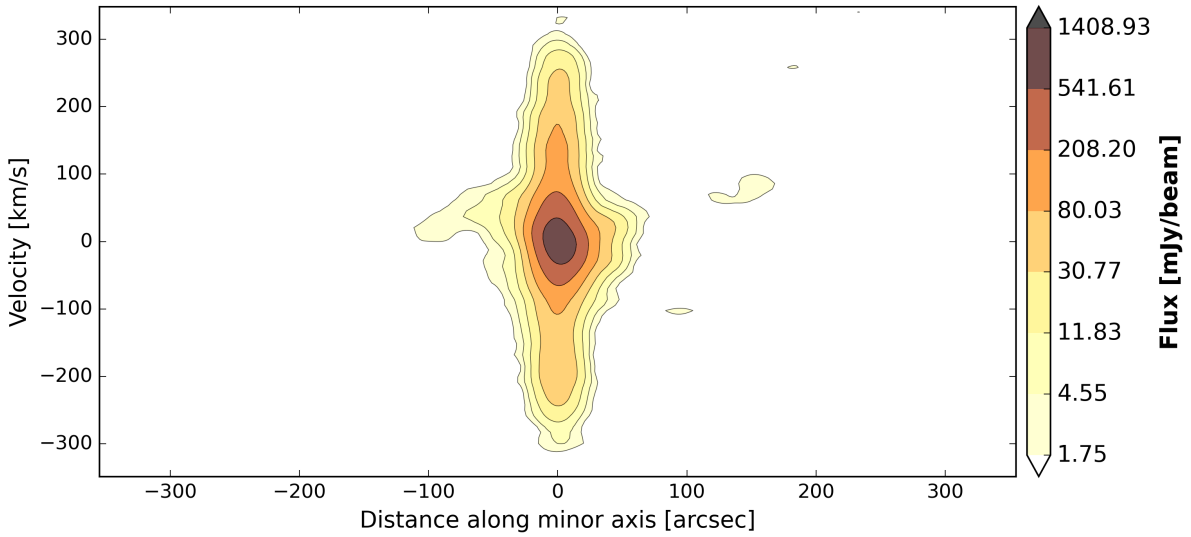


Figure 4.5: Position–velocity diagram along the minor axis for edge-on galaxy.

outermost slice displaying a narrower velocity range but a slightly increased vertical extent in the spatial coordinate. For instance, on the map at the distance of 5.6 kpc from the center (top-middle), we could observe a velocity distribution from 0 $\mathrm{km\,s^{-1}}$ to 300 $\mathrm{km\,s^{-1}}$, and it shows a triangular shape with a wide side close to the central velocity (0 $\mathrm{km\,s^{-1}}$). The same pattern was noted by Sancisi and Allen (1979) and Oosterloo et al. (2007) for NGC891 and reproduced by the lagging-halo model. Also, in the same panel, we could observe some inhomogeneities above and below the disc, which are also seen on the integrated flux map (bottom-right).

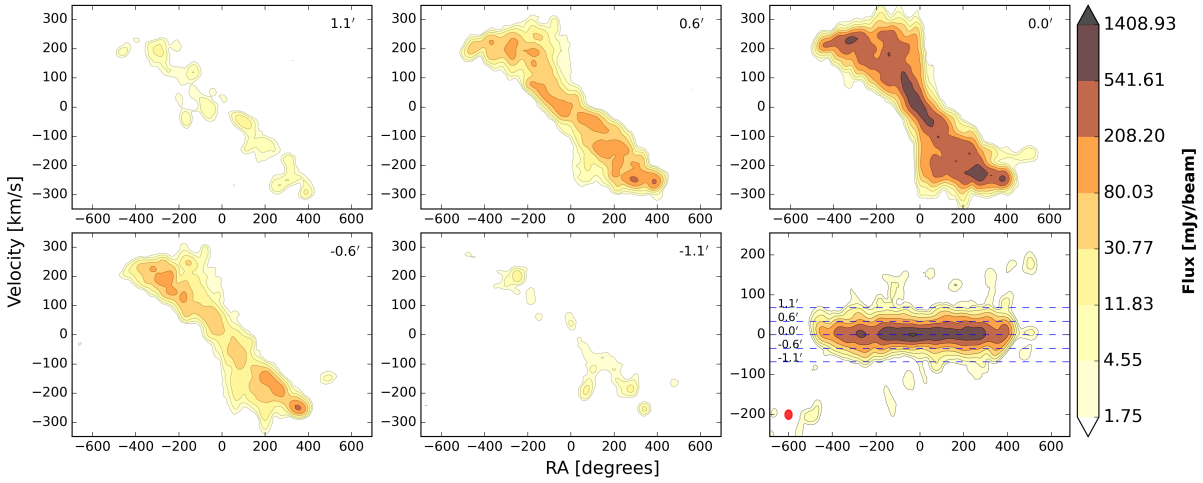


Figure 4.6: Position-velocity diagrams along major axis at different heights from disc: $4'$ (3 kpc), $2'$ (1.6 kpc), $0'$ (0 kpc), $-2'$ (-1.6 kpc), $-4'$ (-3 kpc). The bottom-right panel shows the integrated flux map. The dashed lines represent the axis along which PV diagrams were obtained. Red circle: beam size (33.2×23.9 arcsec^2).

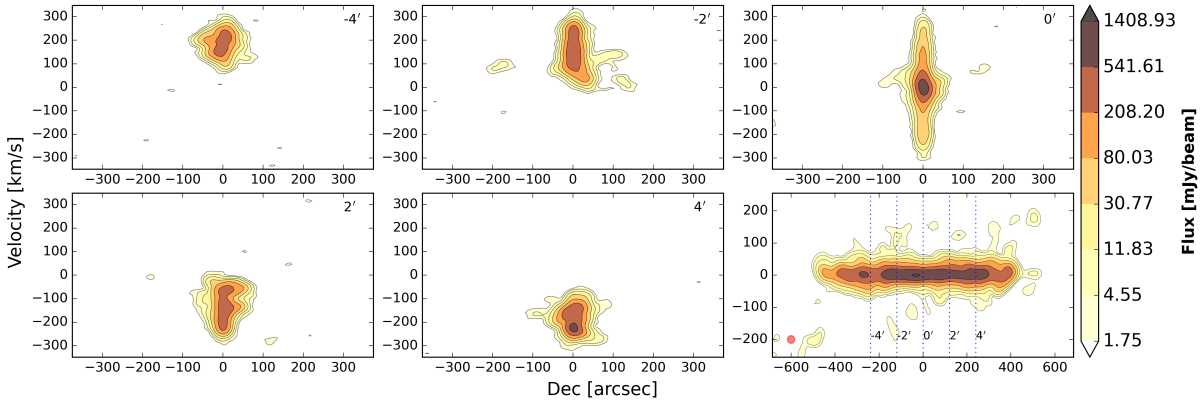
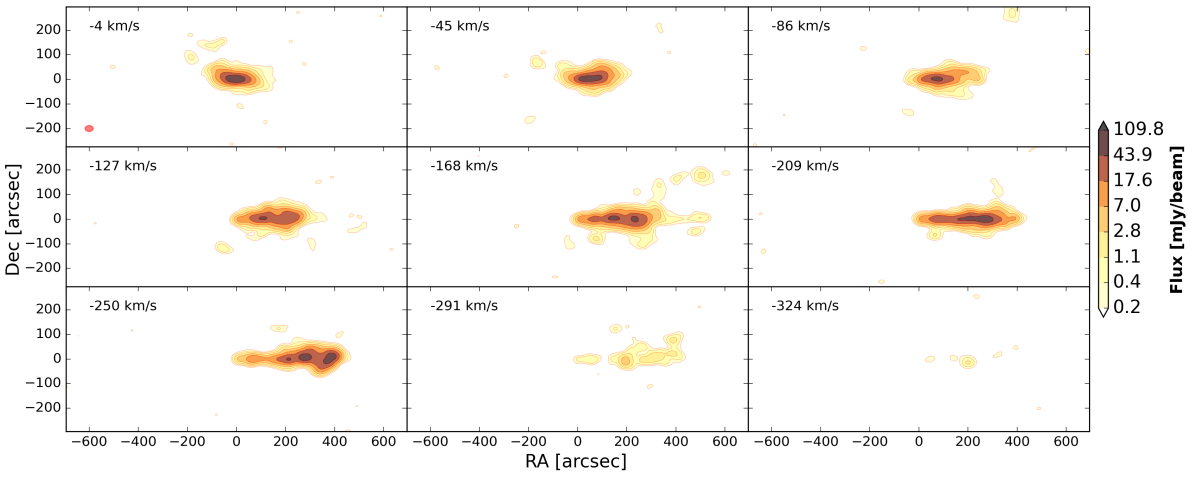


Figure 4.7: Position-velocity diagrams along the minor axis for the edge-on galaxy at different radial distances from the center: $4'$ (11.2 kpc), $2'$(5.6 kpc), $0'$ (0 kpc), $-2'$ (-5.6 kpc), $-4'$ (-11.2 kpc). The bottom-right panel shows the integrated flux map. The dashed lines represent the axis along which PV diagrams were obtained. Red circle: beam size (33.2×23.9 arcsec^2).

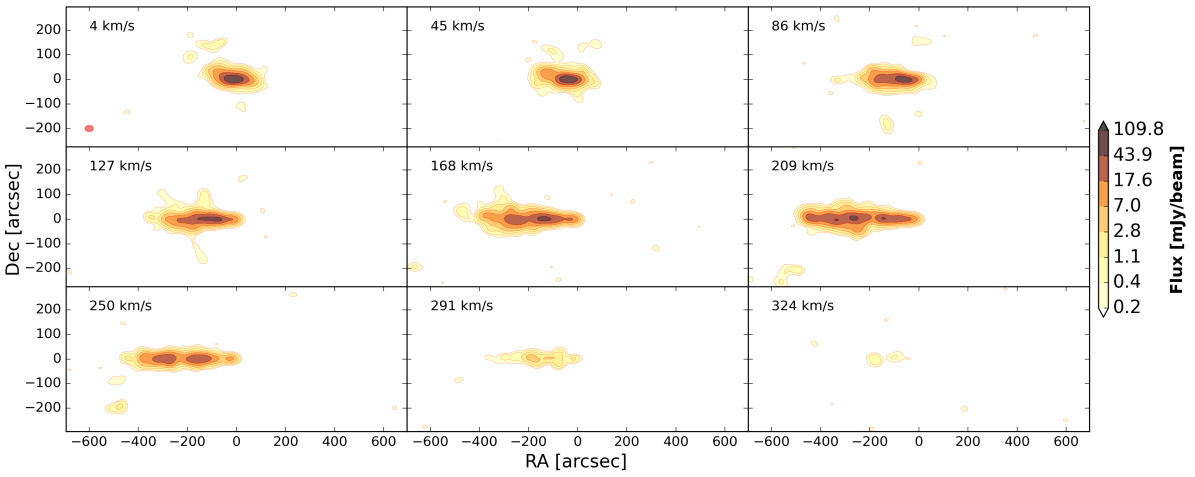
4.1.3 Velocity channel maps

In the channel maps (Fig. 4.8), we present the spatial distribution of the HI flux for each velocity channel. These maps show how the emission varies across the galaxy at different line-of-sight velocities. The systemic velocity of the galaxy (originally $\sim$665 km s^{-1}) was subtracted from the data cube so that the central velocity now lies around 0 km s^{-1}. The channels have been chosen to have a flux higher than the $3 \times \sigma_{rms}$ level and evenly distributed across the velocity interval of [-324 km s^{-1} – +324 km s^{-1}]. The contour levels

correspond to the minimum detectable level and 95% of the maximum flux of the data cube. The emission is clearly detected over a velocity range extending up to $\pm 324\ \mathrm{km\,s^{-1}}$



(a) Velocity channel maps of approaching side: $[4\ \mathrm{km\,s^{-1}} - 324\ \mathrm{km\,s^{-1}}]$



(b) Velocity channel maps of receding side $[4\ \mathrm{km\,s^{-1}} - 324\ \mathrm{km\,s^{-1}}]$

Figure 4.8: Emission of the edge-on galaxy from velocity channels in the range: -332 – +324 $\mathrm{km\,s^{-1}}$. Each panel represents one velocity channel. Red circle: beam size (33.2×23.9 arcsec^2).

and distributed approximately symmetrically in velocity space around 0 $\mathrm{km\,s^{-1}}$. This symmetric velocity distribution is characteristic of a rotating disk. Since the galaxy is observed nearly edge-on, the Doppler shifts due to rotation are maximally projected along the line of sight, allowing a clear detection of both the approaching and receding sides. This configuration is particularly well-suited for analyzing the vertical structure of the HI disk, as the velocity information across the channels corresponds to different positions along the galaxy's rotation curve and height above the mid-plane.

The majority of the flux originates from the galactic disk, as expected for a rotating edge-on system. The emission is concentrated in the disc of the galaxy, and the spatial

extent, both radial and vertical, increases with increasing velocity offset from the center. In the central channels (e.g., ± 4 km s^{-1}), the emission spans approximately 144″ (∼6.63 kpc) along the radial direction and about 54″ (∼2.5 kpc) vertically. The radial extent increases progressively with the velocity up to $\sim \pm 250$ km s^{-1}, beyond which the emission becomes fainter.

The vertical thickness of the disk appears to be relatively constant across the channels, indicating a geometrically thin disk component. However, in several intermediate-velocity channels (notably at ± 127 km s^{-1} and ± 168 km s^{-1}), additional off-plane emission features are observed. These may suggest the presence of EPG or potential outflow components, possibly linked to feedback processes or galactic fountains.

4.1.4 Rotation curve

The rotation curve of the galaxy has been obtained with the envelope tracing method, which is based on the principle that the rotational velocity of the gas in the edge-on case is related to the so-called terminal velocity. This quantity can be defined at a specific intensity level in the line profile, referred to as the terminal intensity, as:

$$I_t = [(\eta I_{max})^2 + I_c^2]^{1/2} \ , \tag{4.1}$$

where I_{max} is the maximum intensity level, I_{LC} is the intensity corresponding to the lowest contour level, and η represents the critical intensity level of the line profile. η generally ranges from 0.2 to 0.5 (Sofue, 2016), and we adopt $\eta = 0.3$ for our analysis. The rotational velocity $V_{\rm rot}$ is then given by:

$$V_{\rm rot} = |V_t - V_{\rm sys}| - \sqrt{\sigma_{\rm ISM}^2 + \sigma_{\rm obs}^2} \ , \tag{4.2}$$

where V_t is the terminal velocity, $V_{\rm sys}$ ($= 665$ km s^{-1}) is the systemic velocity of the galaxy, $\sigma_{\rm ISM}$ is the velocity dispersion of the interstellar medium, which typically ranges from 7 to 10 km s^{-1} (we fixed this to 8 km s^{-1}), and $\sigma_{\rm obs}$ is the velocity resolution of the (mock) observations (see Table 3.1).

In the case of the edge-on galaxy, most of the emission comes from the envelope of the disc, the region where rotation is ordered, faster, and less affected by beam smearing. This method also minimizes the contribution of EPG and non-circular motions. Additionally, since we are observing the galaxy edge-on, by adopting the envelope-tracing method, we reduce errors caused by projection effects and can directly measure the actual rotational velocities.

The resulting curve (see Fig. 4.9) exhibits a rapid increase in rotational velocity, reaching approximately 240 km s^{-1} within the inner 5 kpc. Beyond this radius, the

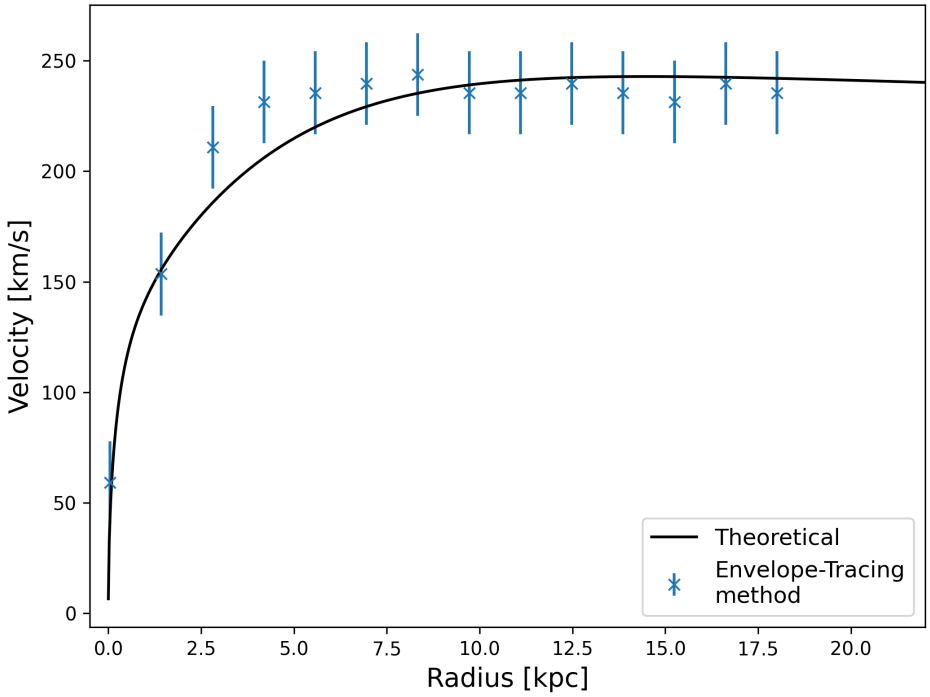


Figure 4.9: Observed rotation curve compared to the one sampled in the initial condition of the simulation.

velocity remains roughly constant, forming a flat rotation curve out to at least 20 kpc. The observationally obtained rotation curve aligns closely with the theoretical trends (see Appendix B.4) expected from the initial parameters of our simulated galaxy, discussed earlier in Section 2.1.

Furthermore, in Fig. 4.10, we show the derived rotation curve overlaid on the position-velocity (PV) diagram. In the image, we clearly observe that the rotation curve generally follows the overall structure of the galaxy.

4.2 Mildly inclined galaxies (30°, 60°)

In this section, we present the results for galaxies with intermediate (60°) and low (30°) inclinations. By analyzing these systems, we gain deeper insight into how observable quantities vary with inclination, particularly due to the effects of projection. Unlike the edge-on case, lower inclinations allow us to disentangle certain morphological features (holes and filaments) of the galaxy. At the same time, however, projection effects can degrade the visibility of other parameters, such as velocity gradients. This analysis

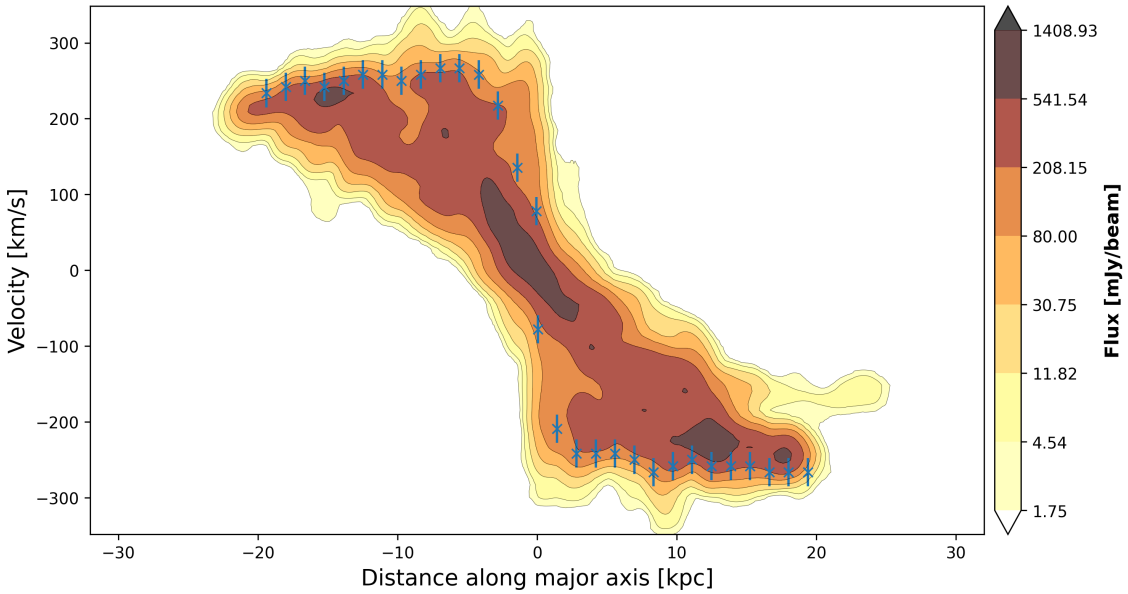


Figure 4.10: Observed rotation curve overlayed on the position velocity diagram along the major axis of the galaxy projected in the edge-on configuration.

enables a more accurate assessment of observational limitations across different viewing angles.

Figure 4.11 presents the moment maps of the galaxies at low and intermediate inclinations. These maps were generated after removing the anomalous gas component (see Section 3.3) from the data cubes and applying a mask with $3 \times \sigma_{\rm rms}$. Anomalous gas refers to HI emission that deviates from the expected rotational motion of the galactic disk. Such components may result from gas inflows, outflows, tidal interactions, or other non-circular motions unrelated to the regular disk dynamics.

This removal step is crucial, as the presence of anomalous gas can distort the interpretation of both the velocity field and velocity dispersion maps. For instance, halo gas, typically rotating more slowly than the disk, appears as a lagging component in the velocity field and contributes to artificially elevated values in the velocity dispersion (Marasco and Fraternali, 2010). By excluding this component, we ensure a clearer characterization of the primary disk structure and avoid biases in the derived morphological and kinematic properties.

The lowest contour in the moment-0 maps corresponds to $3 \times \sigma_{\rm rms}$ of the integrated flux, while the highest contour reaches 95% of the peak emission for each galaxy. Surface brightness units were obtained by scaling the flux densities by the spectral channel width. The moment-1 maps are shown in the middle panels and use the same velocity range (± 260 km s^{-1}). These velocity fields are presented without correction for inclination to

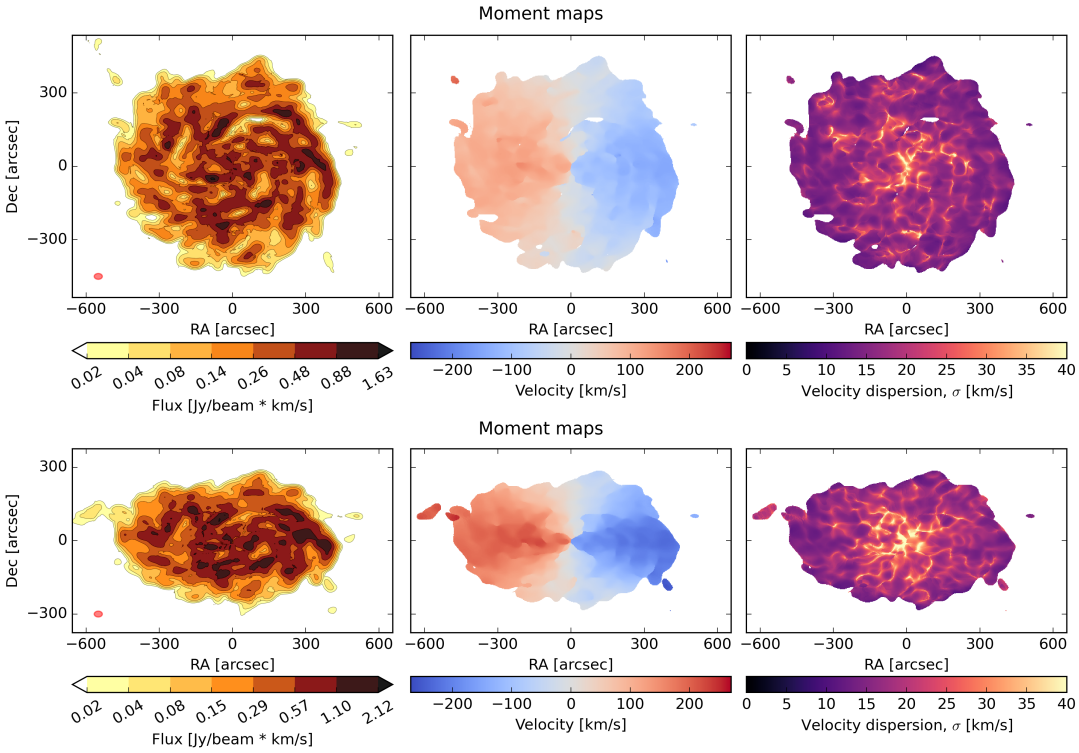


Figure 4.11: Moment maps of galaxies with inclination of 30°(*top*) and 60°(*bottom*)

highlight its effects on observed velocity values. The moment-2 maps, displayed in the right panels, represent the velocity dispersion and are also plotted using the same scale for both galaxies.

The moment-0 maps show how flux distribution is affected by projection effects. The galaxy with a higher inclination angle exhibits a significantly higher observed flux level due to the increased column density of HI along the line of sight. In fact, the peak flux increases by a factor of ~ 1.6. In the lower inclination galaxy, filamentary structures are visible, whereas in the more inclined case, holes and filaments become less distinguishable due to line-of-sight integration.

The range of velocities differs noticeably between the two inclinations. In the intermediate inclination case, the velocity field shows a stronger gradient along the major axis, with clear separation between the approaching and receding sides. For example, the galaxy inclined at 30° shows a velocity range of approximately $-130\ \mathrm{km\,s^{-1}}$ to $+130\ \mathrm{km\,s^{-1}}$, while the 60° case shows a gradient extending from $-220\ \mathrm{km\,s^{-1}}$ to $+220\ \mathrm{km\,s^{-1}}$.

The moment-2 maps, presented on the right side of Fig. 4.11, reveal a consistent pattern. Most of the disc area exhibits a low value of velocity dispersion around 15 $\mathrm{km\,s^{-1}}$, while higher dispersions of about 35–40 $\mathrm{km\,s^{-1}}$ are concentrated near the galaxy

centers. The galaxy with higher inclination shows slightly broader dispersion overall, once again as a result of projection effects.

4.2.1 Column density distribution

Figure 4.12 and Figure 4.13 show the column density maps for the galaxies for the low and intermediate inclination cases, respectively, after the removal of the anomalous gas components from the data cubes. The lowest contour level corresponds to $3 \times \sigma_{\mathrm{rms,HI}}$, while the highest contour represents 95% of the peak detected column density.

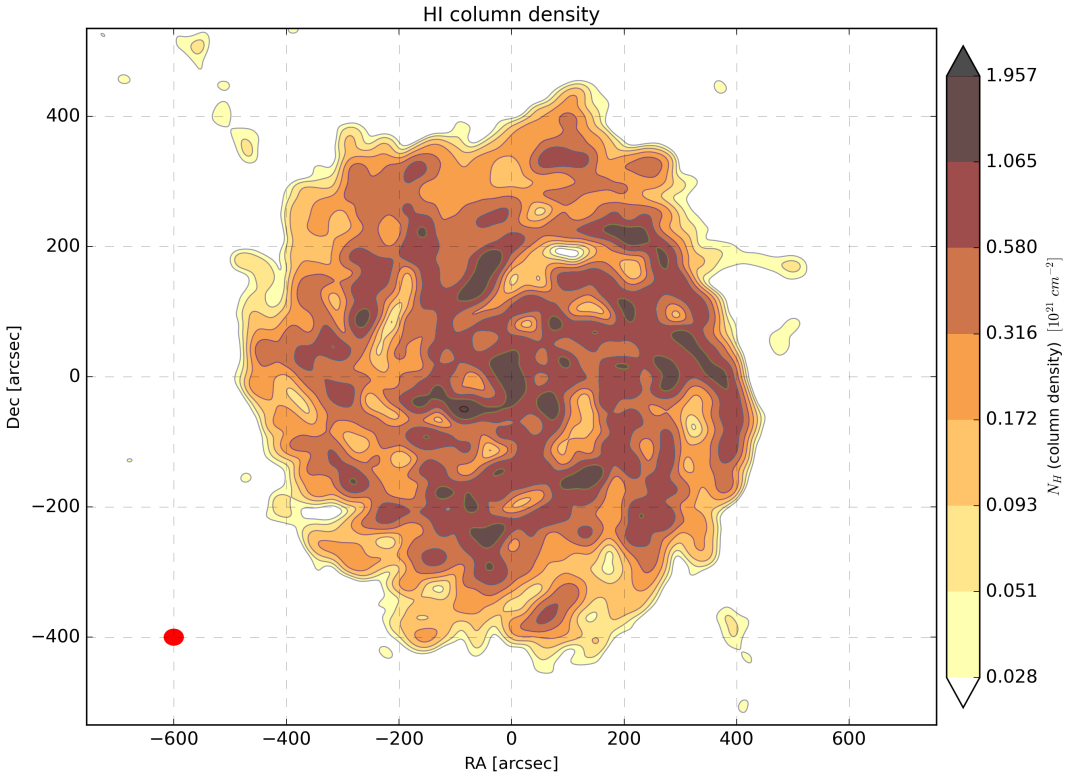


Figure 4.12: HI column density of the galaxy with an inclination of 30°. Red circle: beam size (33.2×23.9 arcsec^2).

At a first glance, the galaxy with the higher inclination (60°) exhibits overall higher column density values, approximately 1.3 times larger than those observed in the 30° case. This enhancement is a projection effect due to the increased line-of-sight path length through the disc. In addition, the 60° map displays broader regions of high column density, and different structures (filaments and holes) are less resolved.

In contrast, the 30° galaxy reveals more detailed morphological features, such as filamentary structures and low-density cavities (holes), which are less discernible in the more inclined case. Interestingly, a region that appears detached from the main body in

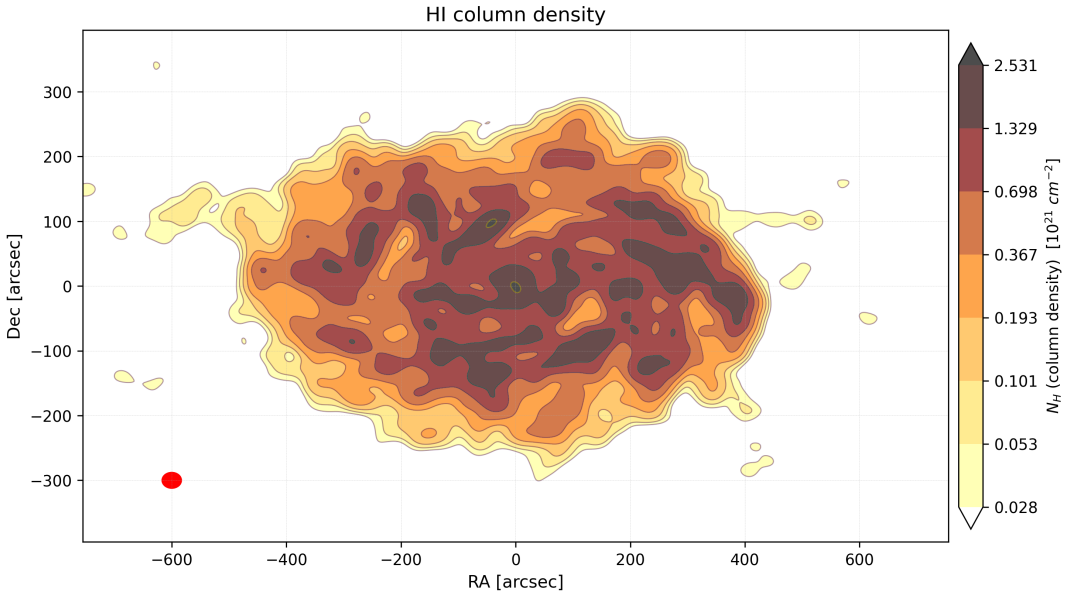


Figure 4.13: HI column density of the galaxy with an inclination of 60°. Red circle: beam size (33.2×23.9 arcsec^2).

the 30° map becomes merged with the disc in the 60° map, likely due to projection and blending effects.

Despite these apparent differences, the total HI masses derived from the cubes are very similar: $4.93 \times 10^9\,M_\odot$ for the 30° galaxy and $4.95 \times 10^9\,M_\odot$ for the 60° galaxy. This consistency confirms that the inclination primarily affects the observed spatial and column density distribution, rather than the total mass measurement.

4.2.2 Position – velocity diagram

The PV diagrams along the major and minor axes for galaxies with inclination angles of 60° are shown in Fig. 4.14, while the corresponding plots for the 30°case are presented in Fig. 4.15. The lowest contour level is defined as $3 \times \sigma_{\mathrm{rms,PV}} = 1.72$ mJy beam^{-1} for 60° and $= 1.55$ mJy beam^{-1} for 30°, while the highest contour corresponds to 95% of the peak flux. The flux and velocity values are derived from the original data cubes, without inclination correction and removal of anomalous gas components.

In the major axis diagrams, clear velocity gradients are visible in the central regions, becoming more pronounced in the galaxy with higher inclination. At larger radii, the rotation curve flattens with only minor deviations. On the receding side (negative distances), the velocity decreases slightly with radius, while on the approaching side (positive distances), a slight increase is observed with respect to central velocities of approximately 5 km s^{-1}. This asymmetry is related to the non-symmetric gas distribution

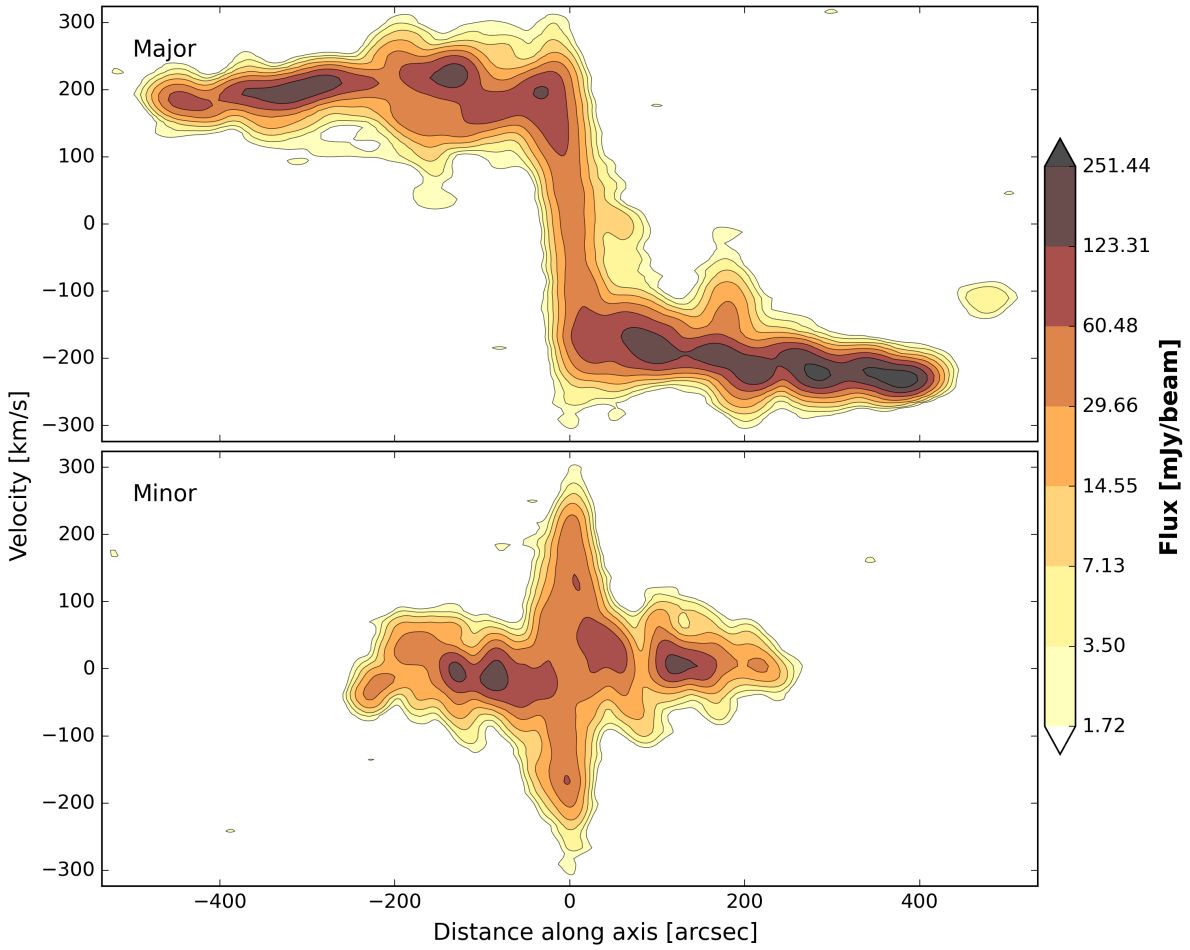


Figure 4.14: Position–velocity diagram along the major (top) and minor (bottom) axis for the galaxy with an inclination of 60°.

on galaxy, which are observed on other inclinations.

Significant flux is detected in regions with velocities of approximately ±240 – ±250 $\mathrm{km\,s^{-1}}$ along the major axis, and around ±30 $\mathrm{km\,s^{-1}}$ along the minor axis. The receding side consistently shows lower flux than the approaching side, the same pattern that observed in edge-on system.

There are deviation from the main flux distribution on PV diagrams, which are related to EPG. In the major axis diagrams, the features are detected at offsets of 180″ on the receding side and at approximately 20″ − 50″ and 180″ on the approaching side. These features appear in both inclination cases, but are more prominent at higher inclination, where line-of-sight integration enhances their visibility.

The minor axis PV diagrams exhibit a typical pattern: a steep velocity distribution near the center, transitioning to flat, low-velocity emission around 0 $\mathrm{km\,s^{-1}}$ at larger radii. The higher inclination case shows a slightly higher central difference, which is

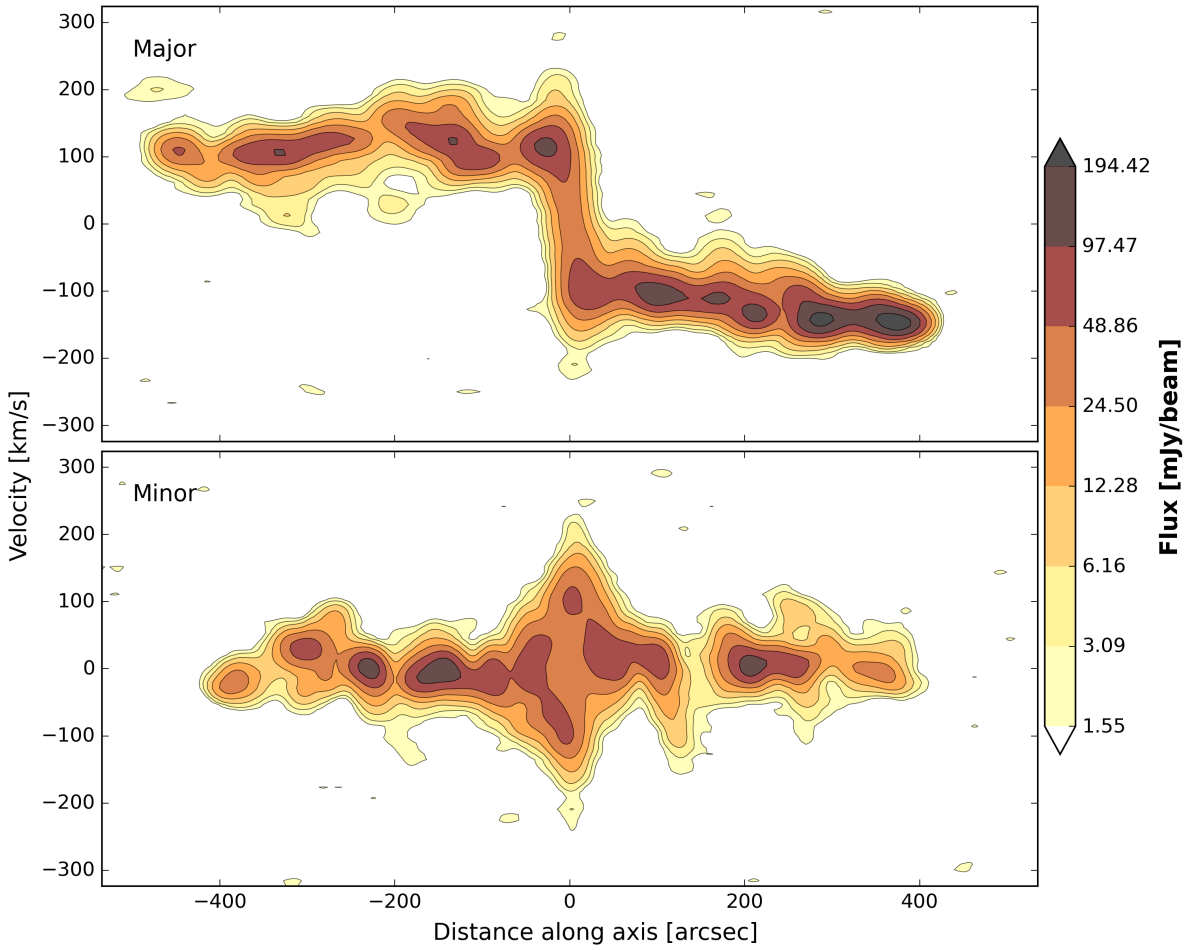


Figure 4.15: Position–velocity diagram along the major (top) and minor (bottom) axis for the galaxy with an inclination of 30°.

expected due to the increased projection of line-of-sight velocities. A minor broadening of the emission is observed near an offset of 120″ on the positive RA side in both minor axis diagrams. This feature may also be associated with EPG.

4.2.3 Velocity channel maps

The velocity channel maps for the low and intermediate inclination cases are shown in Fig. 4.16 and Fig. 4.17.

The maps were generated after separating anomalous gas from the data cubes, as discussed above (see Section 3.3). To accommodate the broader spectral extent of the more inclined galaxy, 24 channels were used for the 60° case, compared to 18 channels for the 30° case. All channels are centered around 0 $\mathrm{km\,s^{-1}}$. The lowest contour level corresponds to $3 \times \sigma_{\mathrm{rms}}$ per channel, while the highest contour is set to 95% of the peak flux in the 60° galaxy in order to clearly highlight the projection effects due to the longer

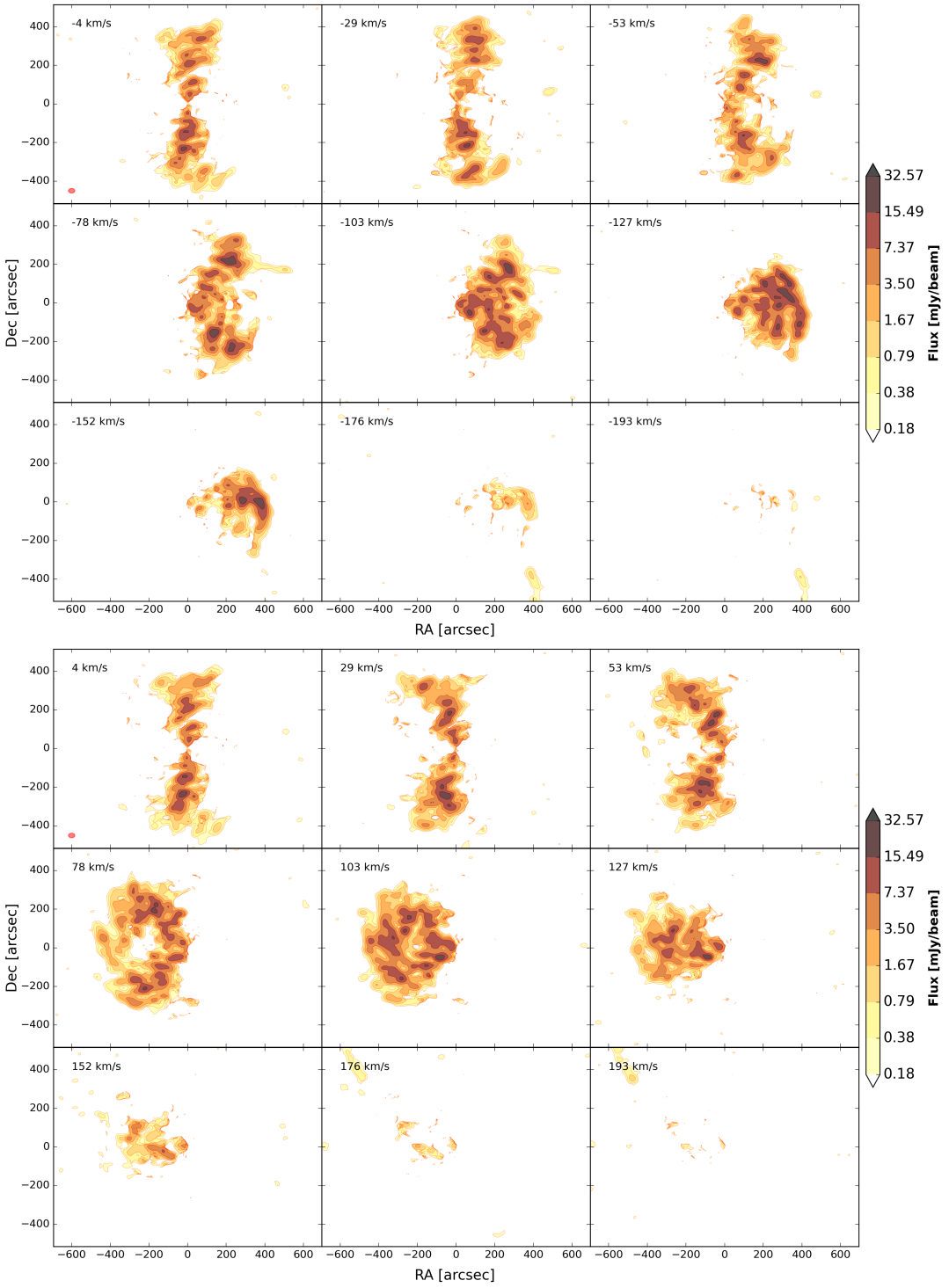


Figure 4.16: Velocity channel maps for the galaxy with an inclination of 30°. Top: Velocity channel maps range: [-193 $\mathrm{km\,s^{-1}}$ to -4 $\mathrm{km\,s^{-1}}$]. Bottom: Velocity channel maps range: [4 $\mathrm{km\,s^{-1}}$ to 193 $\mathrm{km\,s^{-1}}$]. Red circle indicates the beam size (33.2 × 23.9 $\mathrm{arcsec^2}$).

line-of-sight path through the gas.

Both sets of channel maps exhibit a typical behavior: flux is distributed symmetrically

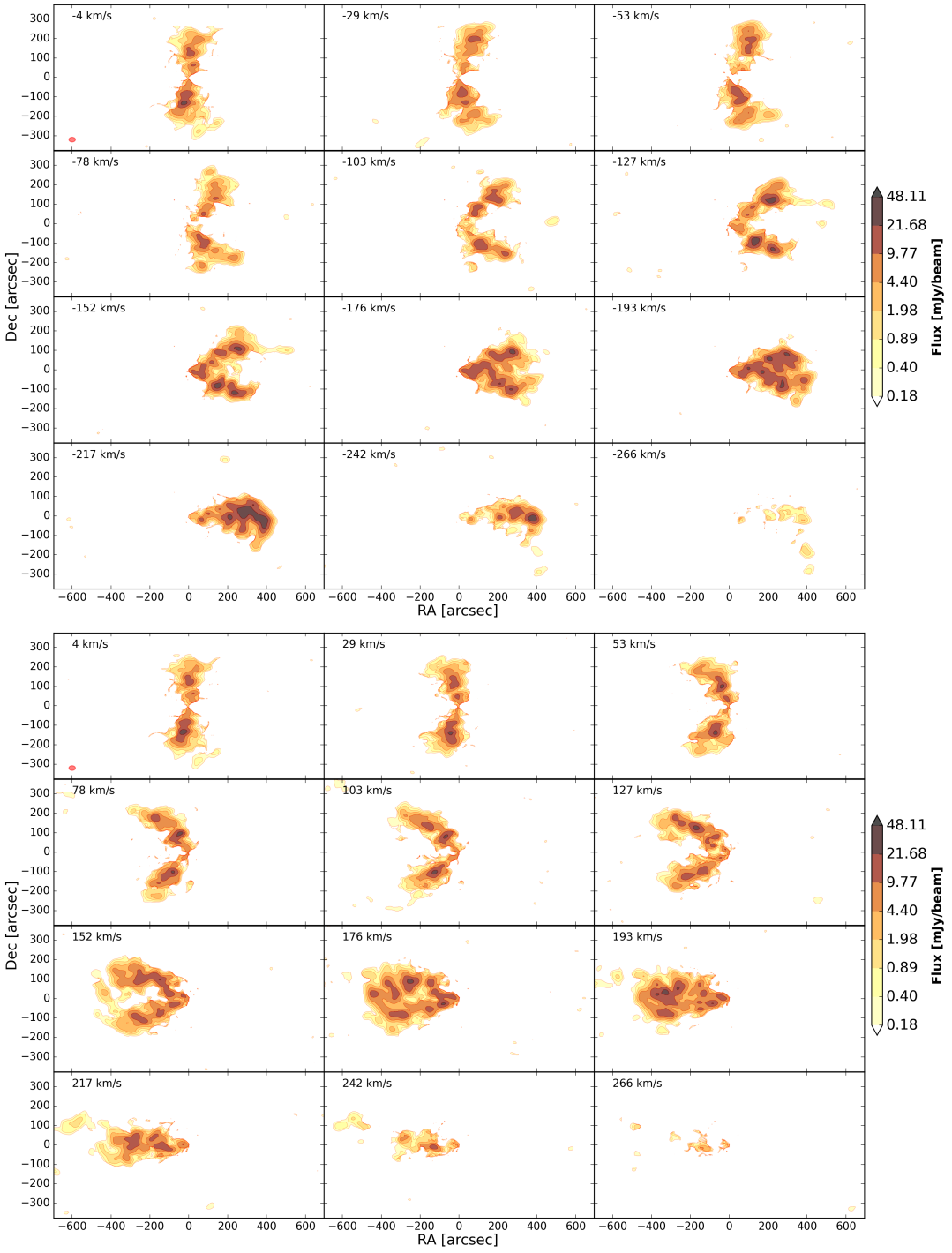


Figure 4.17: Velocity channel maps for the galaxy with an inclination of 60°. Top: Velocity range: $[-266\ \mathrm{km\,s^{-1}}$ to $-4\ \mathrm{km\,s^{-1}}]$. Bottom: Velocity range: $[4\ \mathrm{km\,s^{-1}}$ to $266\ \mathrm{km\,s^{-1}}]$. Red circle indicates the beam size (33.2×23.9 arcsec^2).

around the central velocity ($0\ \mathrm{km\,s^{-1}}$). In the galaxy with 60° inclination, the emission is more extended in the velocity space, reaching up to approximately $\pm 242\ \mathrm{km\,s^{-1}}$. In contrast, the 30° galaxy spans a narrower range of $\pm 152\ \mathrm{km\,s^{-1}}$. As expected, the more

inclined galaxy also displays smoother and more spatially extended flux distribution in each channel, due to projection effects along the line of sight.

For both galaxies, slightly stronger emission is observed in channels corresponding to the approaching side, which is consistent with the asymmetric gas distribution detected for edge-on galaxy observation (Section 4.1).

4.3 Face-on galaxy (0°)

Figure 4.18 presents the moment maps of our galaxy viewed face-on, i.e., with an inclination of 0°. The maps were derived after separating anomalous and cold HI gas components and filtering the flux above a threshold of $5 \times \sigma_{\rm rms}$ for the moment-1 (velocity field) and moment-2 (velocity dispersion) maps. The contours in the moment-0 (integrated intensity) map were calculated using a logarithmic scale, starting from a lower threshold of $3 \times \sigma_{\rm rms}$, up to 95% of the peak emission. The flux values were multiplied by the channel width to yield units of surface brightness.

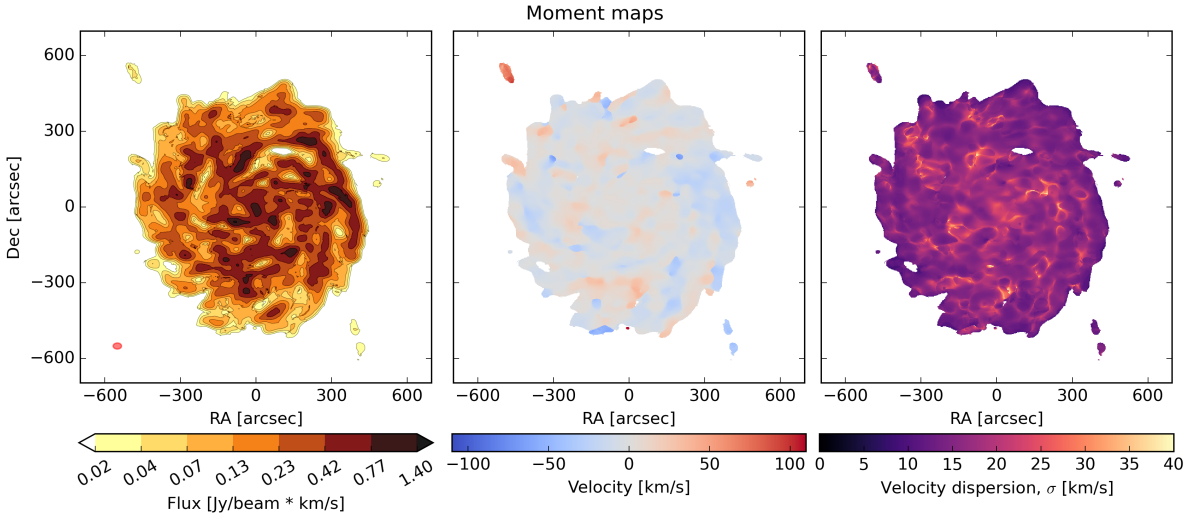


Figure 4.18: Moments maps of the face-on galaxy (0° inclination). Left: Integrated intensity map. Mid: Velocity map. Right: Velocity dispersion map. Red circle: beam size (33.2×23.9 arcsec^2).

The moment-0 map (left panel) displays a patchy structure composed of alternating low-density cavities and high-density filaments, characteristic of HI emission (Puche and Westpfahl, 1994). In the outer regions, the HI emission weakens, and signs of gas outflow into the CGM become apparent. This is inferred from the detection of filamentary structures with substantial surface brightness (> 0.07 Jy beam^{-1} km s^{-1}) and high mean velocities (~ 100 km s^{-1}), which appear spatially detached from the main galactic disc.

The cavities are likely filled with hot gas likely produced by SN explosions (Barbani et al., 2025). These feedback events heat and expel the surrounding ISM, leading to compression and the formation of dense, star-forming filaments.

In the middle panel of Fig. 4.18, the moment-1 map reveals a lack of any large-scale velocity range distribution, as expected for a face-on orientation. However, localized velocity structures are visible along some filaments, likely associated with vertical gas flows, such as outflows and inflows. Most of the HI emission comes from the velocity fields around the central velocity.

The right panel shows the moment-2 map, representing the velocity dispersion. Most of the galaxy exhibits low dispersion values, indicating minimal line-of-sight broadening. Exceptions include a few narrow filaments with elevated velocity dispersions of approximately 35–40 $\mathrm{km\,s^{-1}}$, where stellar feedback is more active.

4.3.1 Column density distribution

The column density map (Fig. 4.19) reveals the spatial distribution and extent of the galaxy. The lowest contour level corresponds to $3 \times \sigma_{\mathrm{rms},N_H} \approx 2.7 \times 10^{19}$ $\mathrm{cm^{-2}}$, while the highest contour represents 95% of the peak detected column density, approximately $\sim 1.9 \times 10^{21}$ $\mathrm{cm^{-2}}$.

As mentioned previously, the galaxy exhibits a porous structure characterized by low-density cavities (holes) and high-density filaments.

High column densities are detected in the central region and along filamentary structures, predominantly located on the western (positive RA, between 0″–400″) side of the galaxy, which is consistent with previous observations of asymmetries (Sections 4.1 and 4.2). Some of these filaments extend to the galaxy's outer edge. These dense filaments serve as reservoirs of cold gas, potentially fueling future star formation.

4.3.2 Position – Velocity diagram

Figure 4.20 presents the position–velocity (PV) diagrams extracted along both the major (x-axis) and the minor (y-axis) of the galaxy, intersecting at the central coordinate. These diagrams were generated prior to the removal of the anomalous gas component in order to visualize its contribution to the kinematic structure.

As expected for a galaxy with an inclination angle of 0°, no clear velocity gradient is observed in either diagram. The spatial extent of the galaxy appears to be similar along both axes, approximately 420″. Most of the HI emission is concentrated within a velocity

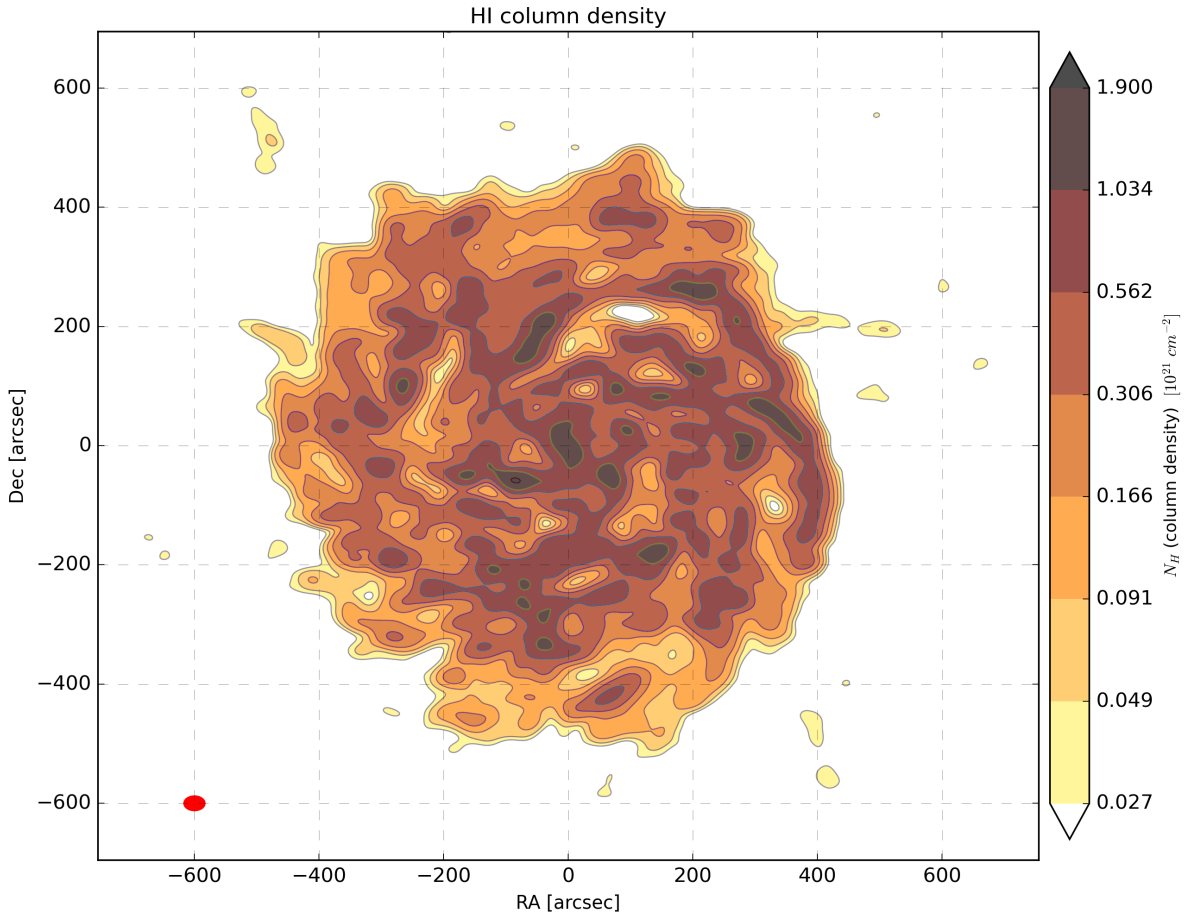


Figure 4.19: HI column density of the face-on galaxy. Red circle: beam size (33.2×23.9 arcsec^2).

range of ± 30 km s^{-1}. In the PV diagram along the major axis, a noticeable asymmetry is present, with stronger emission originating from the positive distances along the axis.

Both PV diagrams show deviations from the typical ± 30 km s^{-1} range, indicating the presence of kinematically distinct features. Notably, high-velocity emission is detected in the central region, reaching up to ± 150 km s^{-1}. Along the major axis, additional features are observed with velocities of approximately ± 100 km s^{-1} at projected offsets of 200″–250″ from the center. Similarly, along the minor axis, filamentary structures are detected with velocities around -100 km s^{-1}, located at angular offsets of approximately $+150''$ and $-220''$ from the center.

These high-velocity features likely trace an extra-planar or outflowing gas, and their distribution highlights the complex kinematics associated with feedback processes in the interstellar medium.

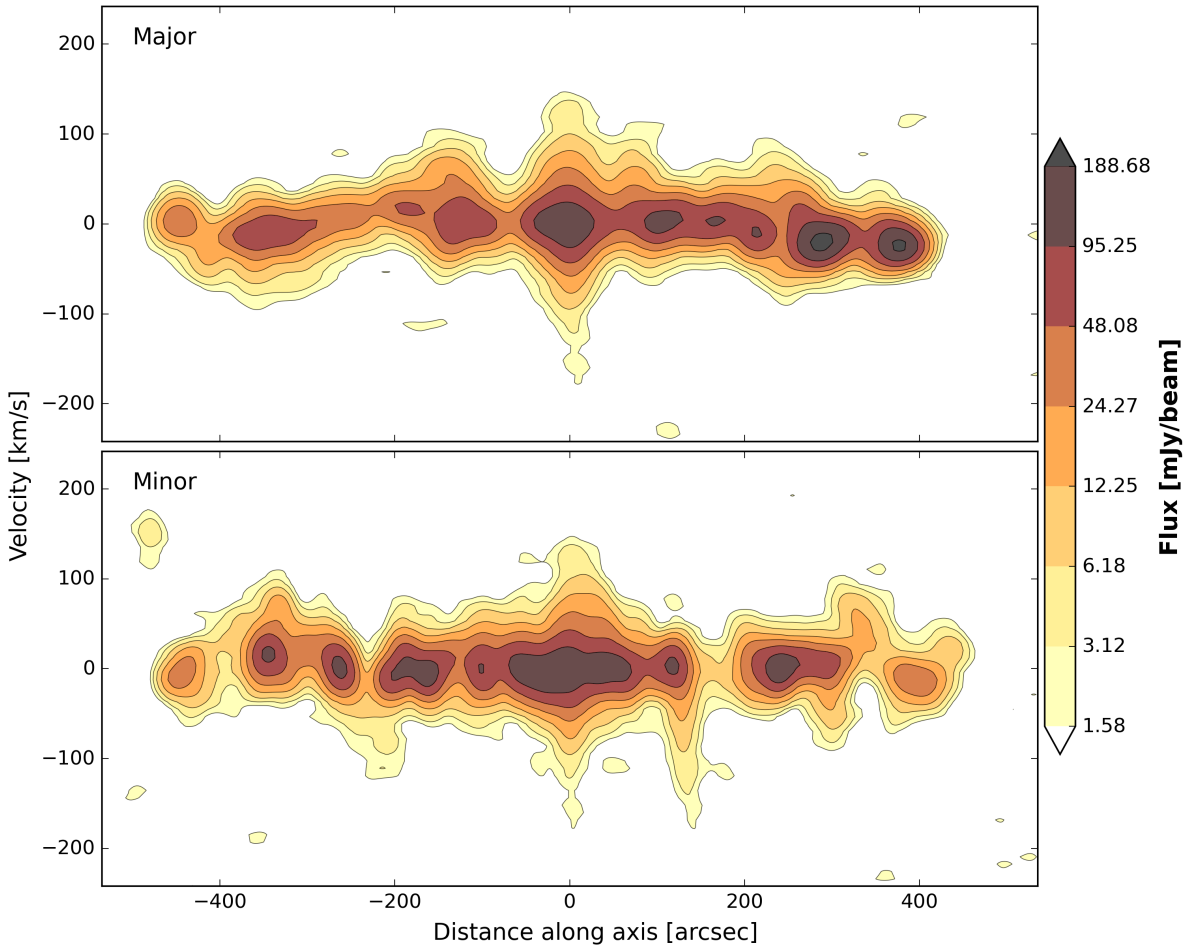


Figure 4.20: Position–velocity diagrams of major (top) and minor (bottom) axis of the face-on galaxy.

4.3.3 Velocity Channel Maps

The velocity channel maps (Fig. 4.21) display emission across 16 selected velocity channels, centered around 0 km s^{-1} ($V_{\rm obs} - V_{\rm sys}$) and covering a total velocity range from −62 to +62 km s^{-1}. Contour levels are plotted on a logarithmic scale, with the lowest contour corresponding to $3 \times \sigma_{\rm rms}$, and the highest representing 95% of the peak emission in the data cube.

The maps show that within the central velocity range of approximately ±37 km s^{-1}. The bulk of the HI flux is concentrated in the central channels, particularly within ±12 km s^{-1}, beyond which the emission gradually fades. At velocities of ±53 km s^{-1}, only about 10% of the peak emission remains detectable, and even less is visible in the outermost channel maps.

Some of this low-level emission at higher absolute velocities may originate from gas

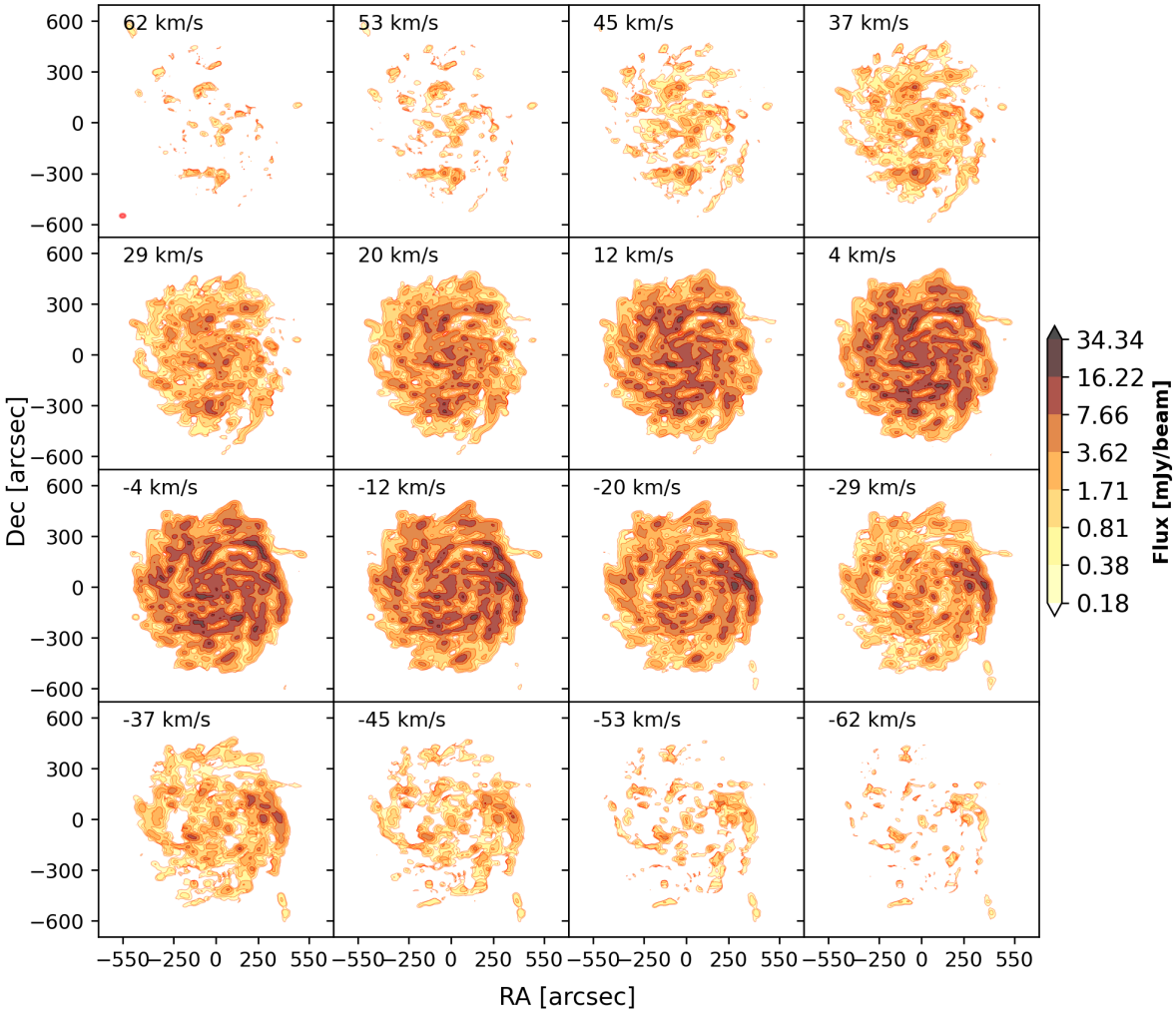


Figure 4.21: Velocity channel maps of face-on galaxy in the velocity range: [-62 km s^{-1} − 62 km s^{-1}]. Red circle: beam size (33.2 × 23.9 $arcsec^2$)

transitioning into or out of the galactic disc, contributing to the expected spectral profile of a face-on system. However, it is also possible that a fraction of these features results from uncertainties or limitations in the method used to separate the anomalous gas component. Despite this, the channel maps display a notable degree of symmetry with respect to the systemic velocity (0 $\mathrm{km\,s^{-1}}$), which is consistent with expectations for a galaxy observed at an inclination of 0°.

4.4 Radial profiles

In this section, we compare the radial profiles of the rotation velocity, velocity dispersion, and surface brightness inferred from the data cubes. The radial profiles of galaxies with inclinations of 0°, 30°, and 60° were obtained using ^{3D}BAROLOby fitting the data cubes with a tilted-ring model (Rogstad et al., 1974). Prior to fitting, the EPG component was subtracted from the data cubes (see Section 3.3), and Hanning smoothing was applied to reduce noise (see Section 3.2).

For the edge-on galaxy (inclination 90°), we derived, instead, the rotation curve using the envelope-tracing method (as described in Section 4.1.4) and inferred the surface brightness profile via an Abel transformation (Appendix B.5; Hickstein et al., 2019) to de-project the observed data. The Abel de-projection assumes axisymmetry and allows the recovery of the intrinsic radial distribution of surface brightness from the line-of-sight integrated emission observed in the edge-on view. Since no reliable method exists for de-projecting velocity dispersion in such systems, we adopted a constant value, as assumed by default in the envelope-tracing method.

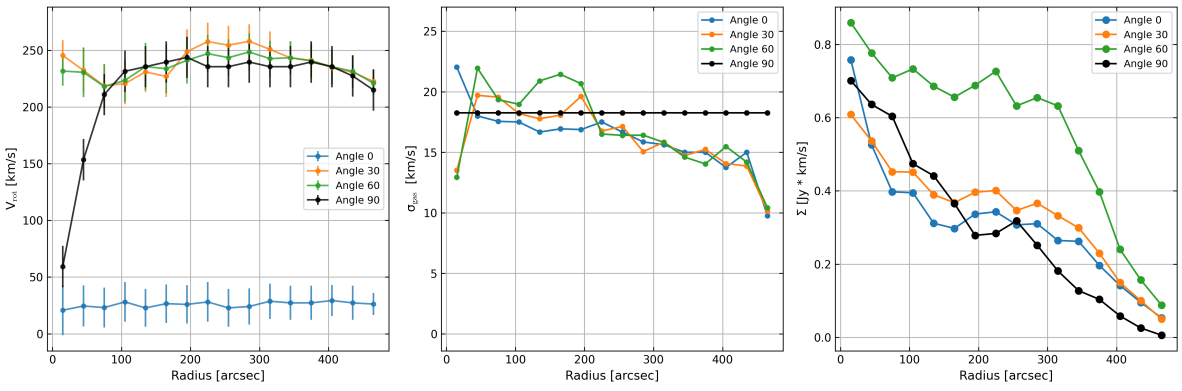


Figure 4.22: Radial profile of parameters. Left: Rotational velocity. Mid: Velocity dispersion. Right: Column density profile. Black: 90°, green: 60°, orange: 30°, blue: 0°.

In the left panel of Fig. 4.22, we present the rotation curves (V_{rot}) of the galaxy at different inclinations. With the exception of the galaxy at 0° inclination, the curves show

good agreement beyond a radius of 100″ (4.6 kpc) and show average values in a range of 236–242 km s^{-1}. In the range between 200″ and 350″, projection effects become evident, as the rotational velocity gradually decreases with increasing inclination.

In the radial profiles of velocity dispersion, we observe general agreement across inclinations. At larger radii (>200″), the velocity dispersion consistently declines with increasing radius. However, in the inner regions (<200″), galaxies with higher inclinations exhibit significantly elevated velocity dispersion values. Notably, at the very center, the galaxy with 0° inclination shows a slightly larger dispersion with respect to less inclined galaxies.

The right panel of Fig. 4.22 shows the surface brightness profile as a function of radius. An anomalous behavior is observed for the galaxy with 60° inclination, which exhibits significantly higher surface brightness (which would imply a higher HI column density) in the radial range of 100″–400″. In contrast, the other three inclinations show relatively consistent profiles. It is important to recall that the profile of the edge-on galaxy was derived using a de-projection method (Abel transformation) to account for projection effects.

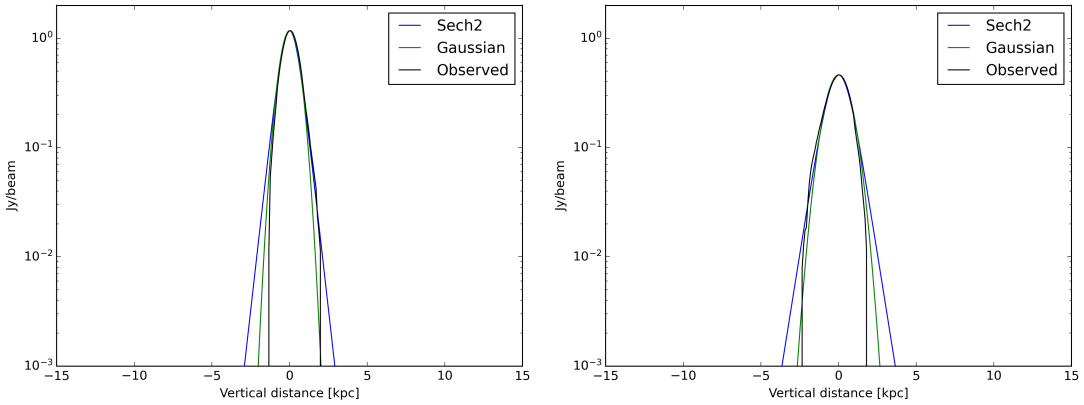
4.5 Extra-planar gas

In this section, we present an analysis of the extra-planar gas component for the four different inclinations. The analysis starts by defining the cold disc component using the method described in Section 3.3. This cold disc component is then subtracted from the original cube, and the residual emission represents the EPG for galaxies with inclinations of 0°, 30°, and 60°. For the edge-on galaxy, the separation between the cold disc and the EPG requires additional assumptions based on prior knowledge. Specifically, we determined an emission threshold above which 90% of the total HI mass is contained. We then applied Gaussian fitting to the vertical flux distribution (showed less residuals) in regions along the major axis with significant emission (see Fig. 4.23a and Fig. 4.23b).

This approach allowed us to characterize the cold disc and isolate the anomalous gas as the residual emission after subtraction. Additionally, we estimated the disc scale height by correcting the observed vertical extent $\sigma_{\rm obs}$ of the cold disc for beam smearing effects using the relation (Wilson et al., 2013):

$$h_{\rm disc} = \sqrt{\sigma^2_{\rm obs} - \sigma^2_{\rm beam}} \tag{4.3}$$

The edge-on configuration was also used for a more detailed analysis of the kinematics and vertical structure of the EPG component.



(a) 0.69 kpc from center. **(b)** 10.36 kpc from center.

Figure 4.23: Vertical profile fitting of HI emission at different radial distances from the galaxy center. Two models were applied to the data (black): a sech^2 profile (blue) and a Gaussian profile (green). The comparison illustrates the differences in how each model captures the shape of the emission, with implications for estimating the scale height and structure of the gas disc.

We computed the total and extra-planar HI masses for all inclinations. The total masses of the data cubes were calculated using Eq. (3.20), and are approximately (4.93–4.95) $\times\ 10^9\, M_\odot$ for all inclinations. The mass of the anomalous EPG was estimated to be in the range of (0.49–0.55) $\times\ 10^9\, M_\odot$, corresponding to approximately 11% of the total mass of the data cube. The derived masses and other relevant parameters are summarized in Table 4.1.

The observed fraction of EPG is consistent with results of the HALOGAS survey (Marasco et al., 2019). The main finding of that survey was that, mostly, EPG made up around 10-15% of total mass, reaching up to 25% in a few cases. This agreement supports the realism of the simulated gas distribution and the reliability of our methodology for identifying and quantifying anomalous gas components.

In Fig. 4.24 and Fig. 4.25, we present the position–velocity (PV) diagrams for galaxies inclined at 30° and 60°, respectively. The left panels show the total HI emission along the major axis. The middle panels isolate the cold gas emission, highlighting the contribution from the thin HI disc, while the right panels display the residuals, which we interpret as the extra-planar or anomalous gas component.

These figures are included to demonstrate how the anomalous gas can be separated from the regular disc component based on its distinct kinematics. By varying the inclination, we evaluate the robustness of our method and investigate how projection effects influence the identification and morphology of extra-planar features in PV space.

The PV diagrams of the EPG reveal broader velocity structures and asymmetries,

Inclination (°)	M_t ($10^9 M_\odot$)	M_{cd} ($10^9 M_\odot$)	M_{ag} ($10^9 M_\odot$)	f_{halo}	h_z (pc)	$V_{\rm rot}$ (km s^{-1})
0	4.93	4.37	0.54	0.109	257.8	—
30	4.93	4.40	0.51	0.103	248.7	241.7
60	4.95	4.38	0.55	0.111	240.7	238.9
90	4.96	4.42	0.53	0.107	241.2	236.2

Table 4.1: Inferred parameters of the galaxies with different inclinations. M_t is the total HI mass detected in the data cube, M_{cd} is the mass of cold disc gas, M_{ag} is the mass of anomalous gas, f_{halo} is the fraction of mass of extra-planar HI gas, h_z is the scale height of the gaseous disc, $V_{\rm rot}$ is a mean rotation velocity of the "flat" part of the rotation curve ($R > 4.6$ kpc).

indicative of anomalous kinematics relative to the rotating disc. These features suggest the presence of non-circular motions and potential vertical outflows or inflows. This decomposition of the HI emission enables a clearer distinction between the regular disc kinematics and disturbed gas components, which may be associated with galactic fountains, accretion flows, or/and feedback-driven processes.

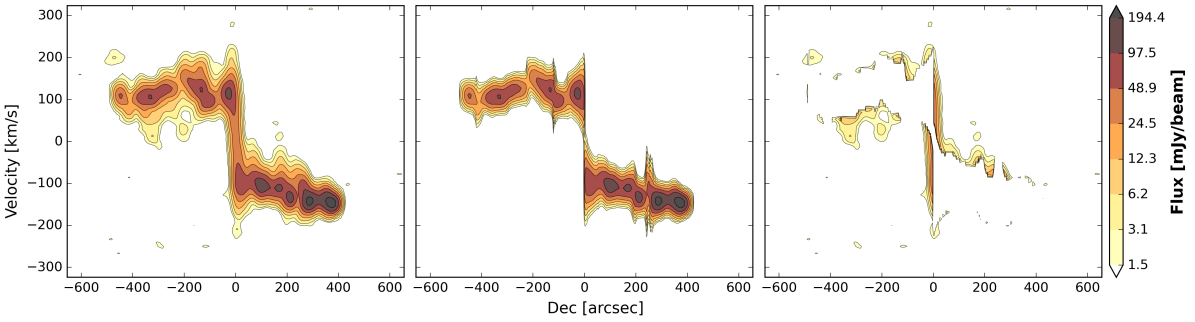


Figure 4.24: Position velocity diagrams of galaxy with 30° inclination. Left: Total HI emission. Mid: Emission of the HI disc. Right: Extra-planar gas emission.

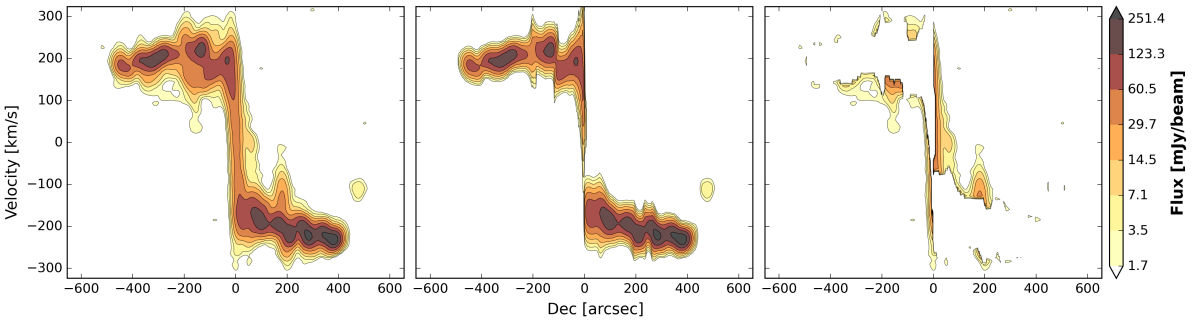


Figure 4.25: The same as Fig. 4.24, but for 60° inclination.

To characterize the kinematics of the EPG, we derived the rotation curves of the

edge-on galaxy at three representative heights, 0.51, 1.46, 2.35 kpc above and below the galaxy mid-plane (Fig. 4.26). The curves were obtained using the envelope-tracing method with parameter $\eta = 0.25$ to capture faint emission, as described in Section 4.1.4, by applying it independently in four quadrants and computing the mean rotation curve at any given height (see Fig. 4.28).

The rotation curves exhibit a declining trend with increasing vertical distance from the plane. While the rotation velocity rises steeply near the disc plane, it increases more gradually at larger heights, indicating a vertical gradient in rotational velocity $\partial v/\partial z < 0$. The existence of this gradient provides key information about the structure

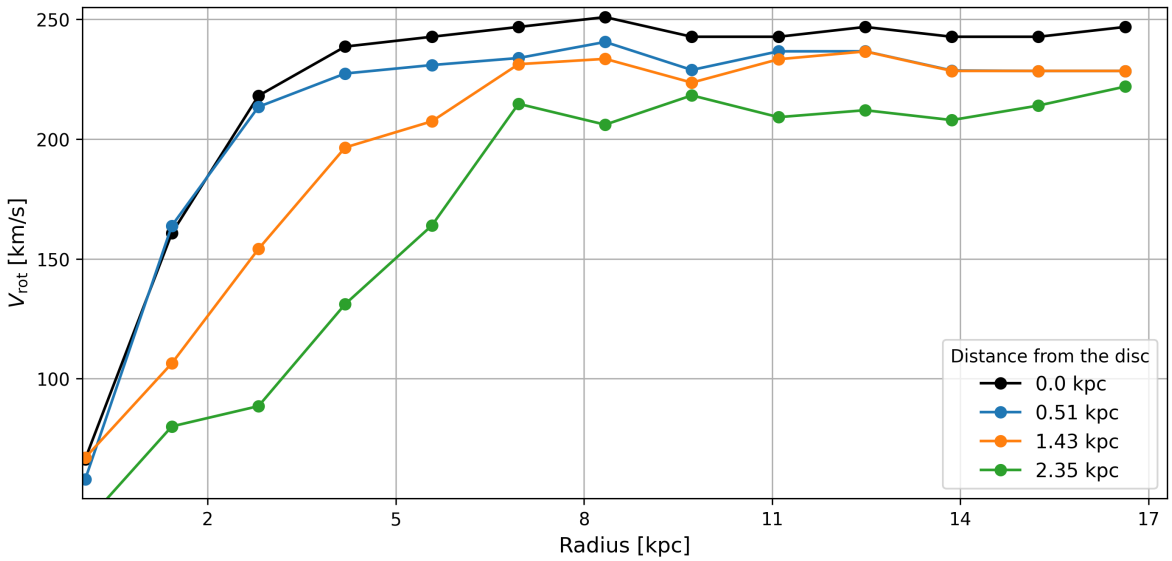


Figure 4.26: Observed rotation curves of extra-planar HI gas at different heights from the mid-plane: black: 0.0 kpc, blue: 0.51 kpc, orange: 1.43 kpc, green: 2.35 kpc.

and kinematics of the galaxy–halo interface. In fact, a gradual decline in rotational velocity with height may suggest a nearly co-rotating EPG, whereas a steeper or more irregular gradient could indicate a dynamically distinct or decoupled halo component. By analyzing the vertical velocity gradient, we gain valuable insight into the spatial extent, coherence, and origin of the EPG.

Using the rotation curves extracted at different heights (see Fig. 4.26), we estimated the value of the vertical velocity gradient $\partial v/\partial z$ by applying a second-order accurate central difference scheme. The gradients were calculated point-by-point across the range of heights and are shown in Fig. 4.27. We then computed the mean vertical gradient by averaging the values over the radial range of 7–17 kpc, where the rotation curves are well resolved and free from significant uncertainties.

We find a mean vertical velocity gradient of approximately ~ -10 km s^{-1} kpc^{-1}. This

value is consistent with observational determinations in nearby edge-on galaxies, including NGC891 (-12 km s^{-1} kpc^{-1} Oosterloo et al., 2007), NGC4152 (-11.18 km s^{-1} kpc^{-1} Li et al., 2021, results from the HALOGAS survey (-10 km s^{-1} kpc^{-1}, Marasco et al., 2019), and even MW estimates (-6 to -9 km s^{-1} kpc^{-1}, Bish et al., 2019). This agreement reinforces the presence of a lagging halo component, likely shaped by angular momentum redistribution through galactic fountain flows or interactions with the hot CGM.

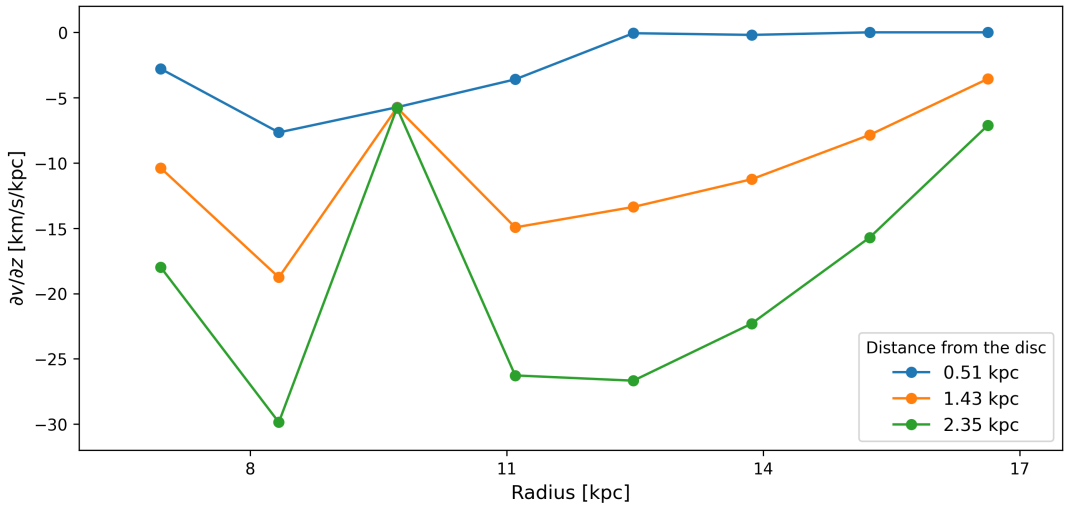


Figure 4.27: Inferred vertical gradient of rotational velocity for $R > 7$ kpc at different height: [0.51 kpc–2.35 kpc].

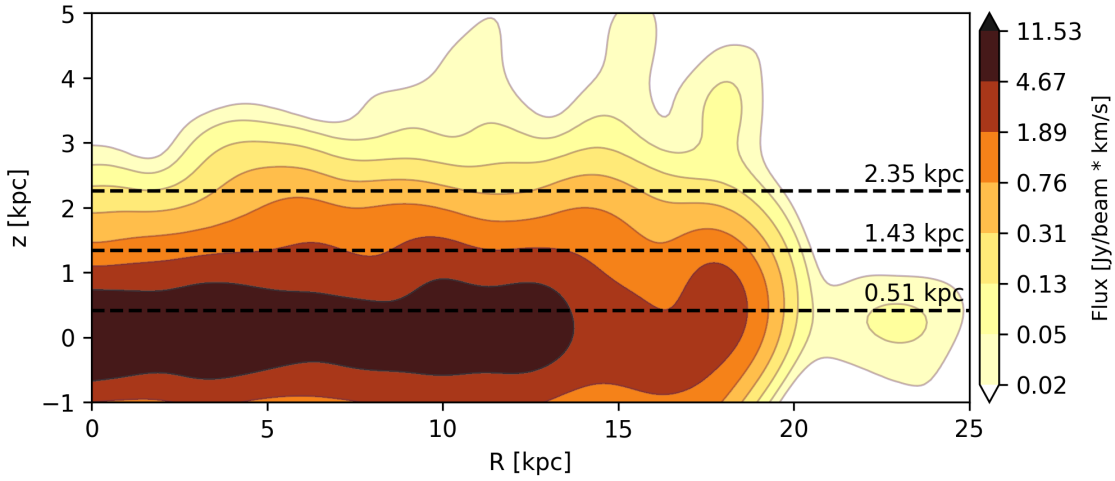


Figure 4.28: Zoom-in of moment-0 map of the edge-on galaxy, showing the vertical structure of the HI emission. The dotted lines indicate the heights above the disc plane along which rotation curves were extracted to estimate the vertical velocity gradient.

The detailed kinematics of galaxies observed at intermediate inclinations (e.g., 30°, 60°) could be further investigated using advanced modeling techniques. For example, one approach involves generating synthetic data for the surface density distribution based on empirical models, such as the one proposed by Oosterloo et al. (2007):

$$\Sigma(R) = \Sigma_0 \left(1 + \frac{R}{R_g}\right) \exp\left(-\frac{R}{R_g}\right), \tag{4.4}$$

where Σ_0 is the central surface density and R_g is the radial scale length. To describe the vertical distribution of the HI gas, an empirical function was used:

$$\zeta(z) = \zeta_0 \frac{\sinh(|z|/h_{\text{halo}})}{\cosh^2(|z|/h_{\text{halo}})}, \tag{4.5}$$

where ζ_0 is the mid-plane surface density, z is the vertical coordinate, and h_{halo} is the halo scale height.

These models have successfully reproduced the vertical HI distribution in galaxies like NGC891 and could be employed to interpret observations or synthetic data at various inclinations. Applying such techniques would allow for a more comprehensive analysis of anomalous gas kinematics and facilitate the recovery of vertical gradients even in non-edge-on systems, where projection effects obscure direct measurements.

The vertical velocity gradient, as discussed previously, is a key dynamical signature of EPG and supports the physical plausibility of our simulation. While our gradient measurement was limited to the edge-on case, the extension of this analysis using model-based deprojections for inclined galaxies represents a promising avenue for future work. Such studies could help generalize the detectability and interpretation of lagging halos, providing a more complete picture of the multiphase circumgalactic environment.

5

Conclusions

In this thesis work, we created and analyzed mock HI observations of a simulated MW-like galaxy viewed at different inclinations. The simulations were performed with the AREPO hydrodynamical N-body code (Weinberger et al., 2020), which solves the Euler equations on a moving Voronoi mesh using an exact Riemann solver and an oct-tree algorithm for gravity. The AREPO code was used in conjunction with the SMUGGLE model (Marinacci et al., 2019; Barbani et al., 2025) so that the resulting a MW-like galaxy has a realistic multiphase ISM, and self consistently generates gaseous outflows that interact with the (hot) CGM material surrounding the disk. The setup of the simulations and a more detailed description of the AREPO code and the SMUGGLE model can be found in Chapter 2.

To generate mock HI observations of the simulated galaxy, we modified the MARTINI (Oman, 2024) software package to support output from SMUGGLE (see Section 3.1). MARTINI is a tool for the Python programming language that generates mock HI observations by convolving simulation data with user-specified spectral and angular resolutions and noise levels. To ensure realistic conditions, we used the observational parameters of the nearby star-forming galaxy NGC891 as a reference. The only variation introduced was the inclination angle of the generated mock observations.

This setup enabled us to investigate the morphological and kinematic features of the HI gas emission from the galactic disc and form the extra-planar regions using relatively simple and well tested observational methods (see Chapter 3). It also provided a robust

way to study projection effects and to identify degeneracies in observable quantities as a function of inclination. In Chapter 4, we present the results of our mock HI observations, which can be summarized as follows:

- Based on the observational features of the edge-on galaxy, we find that the HI emission of our simulated galaxy is observed to extend up to $\sim 75''$ (or ~ 3.5 kpc) above and below the plane. This is significantly larger than the expected scale height of the cold HI disc, which suggests the detection of an EPG layer in our mock observations. Most of the emission (about 90%) originates from the central region with a vertical extent of $\pm 25''$ (or ~ 1.1 kpc). From the surface density map, we observe slightly enhanced emission on the approaching side of the simulated galaxy. Several HI clouds are detected at heights of up to $180''$ (or ~ 8 kpc), which are likely associated with outflows or so-called galactic fountains. The velocity dispersion of the HI gas is notable, ranging from ~ 50 to 80 km s^{-1}, particularly in the central regions. This is most likely a result of projection effects due to the high inclination. The mean velocity field exhibits a largely symmetrical "butterfly" pattern, which is characteristic of rotating discs viewed edge-on. For edge-on galaxies, the rotation curve can be extracted using the envelope-tracing method. This technique involves measuring the velocity at a specified terminal intensity level on the PV diagram, and we inferred an average velocity of "flat" part of rotation curve of 236 km s^{-1}, thus recovering the "actual" rotation velocity value of the simulated galaxy.

- In the mildly inclined galaxies (30°, 60°), we can see the effects of projection, i.e., an increase in surface brightness, as well as increased values in the velocity fields and velocity dispersion with respect to less inclined galaxies. These effects are especially evident in the galaxy with 60° inclination. In this case, resolved structures, such as the central region patches, seen in less inclined galaxies, are no longer visible. On the PV diagrams, particularly those along the major axis, we observe emission from the EPG component and indications of gas outflows/inflows. These features are more recognizable in the more inclined galaxies due to their irregular velocity signatures.

- From the mock observations of the face-on galaxy, we can identify several morphological characteristics. The surface density map reveals a patchy structure of the cold gas where high surface density regions alternate with “holes". Most of the emission is concentrated within the velocity range of ± 30 km s^{-1}, with only a few filaments reaching higher velocities of up to $\sim$100 km s^{-1}.

- We obtained the radial profiles of rotation velocity ($V_{\rm rot}$), velocity dispersion ($\sigma_{\rm HI}$), and surface brightness ($\Sigma_{\rm HI}$) using $^{\rm 3D}$BAROLO, for galaxy inclinations of 0°, 30°, and 60° as mentioned earlier, in the edge-on case $V_{\rm rot}$ was derived using the envelope-tracing method, whereas the surface brightness profile was recovered through Abel de-projection (Hickstein et al., 2019). The results were consistent across different inclinations except for surface brightness profile at 60° inclination. In this case the surface brightness was overestimated compared to other inclinations in the inner regions. While we lack a definite explanation for this discrepancy, it likely due to projection effects and the difficulties in making an optimal choice of fit parameters in $^{\rm 3D}$BAROLO.

- By separating the cold disc from the raw data cube using the method described in Section 3.3, we were able to isolate the anomalous gas component, characterized by kinematics inconsistent with that of the disc. Observationally, this gas is associated with the extra-planar HI component found in nearby galaxies. The separation procedure was applied to mock observations with inclinations of 0°, 30°, and 60°, and yielded consistent results not only across these cases, but also in comparison with observed values from nearby spiral galaxies in terms of the mass content of this component. The mass of the anomalous gas was estimated in the range 0.49–0.55 $10^9 M_\odot$, corresponding to a fraction of $f_{halo} \approx 0.1$ of the total HI mass.

- The kinematical analysis of the EPG was performed for the mock edge-on galaxy, where we obtained rotation curves at different heights above the disc and estimated the corresponding vertical velocity gradient. The rotation curves exhibited a declining trend with increasing height. The velocity gradient was calculated in the external parts of the rotation curves (beyond the rising portion, i.e., > 7 kpc), yielding a mean gradient of -10 km s^{-1} kpc^{-1}. These results support the model of a nearly co-rotating, extended gaseous halo and are consistent with observational determinations.

While the analysis performed in this thesis offers valuable insights into the kinematics and structure of HI gas using mock observations, several limitations should be acknowledged. First, the simulations on which the mock observations were based are idealized (in the sense that the galaxy evolves in isolation) and lacked magnetic fields, which are known to influence the structure and dynamics of gaseous halos (e.g., Beck, 2011). Moreover, our mock HI observations are restricted to a single snapshot in time, and this narrowed the scope of our investigation at a specific time in the evolution of

the simulated galaxy, preventing the analysis of how the HI properties are linked to the evolution of physical process within the galaxy such as varying levels of star formation activity. Furthermore, our analysis was limited to four discrete inclination angles, which does not capture the full spectrum of projection effects encountered in real observations. It should be mentioned that more advanced methods exist for separating the EPG component from disc emission in data cubes (Marasco et al., 2019), which could yield more refined results, including the potential to model more accurately the properties of anomalous gas. These limitations point to clear directions for future work, such as applying more advanced techniques to analyze mock data, time-resolved analysis, and a more comprehensive exploration of projection effects. Another promising direction is the extension of mock observations to include other gas tracers, which would provide a multi-phase view of outflows and inflows, offering a more complete picture of the baryon cycle. Finally, it is also possible to extend our analysis toward mock observations of the MW, by generating data cubes similar to the ones of the Leiden-Argentina-Bonn (LAB) survey (Kalberla et al., 2005). This would allow us to interpret some observational features, such as high- and intermediate-velocity clouds (HVCs and IVCs), of our own galaxy with greater clarity.

In summary, this work demonstrates the potential of mock HI observations as a powerful tool for probing the kinematics and morphology of neutral hydrogen in star-forming galaxies across a range of inclinations. We analyzed a MW-like galaxy using synthetic data to explore the distribution and dynamics of EPG. By employing well-defined separation techniques, we isolated the anomalous gas component from the cold HI disc, allowing us to estimate its mass and characterize its distinct kinematic features. The cold disc was studied independently, revealing how the EPG alters the overall velocity field and morphological appearance of the galaxy. In the edge-on configuration, we quantified the vertical velocity gradient of the halo gas, consistent with the presence of a lagging halo, as observed in real galaxies. This comprehensive decomposition emphasizes the need to treat EPG as a distinct and dynamically relevant component in both simulations and observational analyses of galactic gas kinematics.

Despite the idealized nature of the simulations, the results are consistent with observational trends in nearby MW-like galaxies. These findings provide a solid foundation for more advanced studies, bridging the gap between theoretical models and observational techniques. With further development, such mock observations can help refine our understanding of the multi-phase interstellar medium, the galactic fountain cycle, and the complex interplay between disc, EPG and the CGM. This work represents a step toward that broader goal.

Bibliography

Armillotta, L. et al. (2016). "Efficiency of gas cooling and accretion at the disc-corona interface". In: Monthly Notices of the Royal Astronomical Society 462.4, pp. 4157–4170.

Asplund, M. et al. (2009). "The Chemical Composition of the Sun". In: Annual Review of Astronomy and Astrophysics 47.1, pp. 481–522.

Barbani, F. et al. (2023). "Galactic coronae in Milky Way-like galaxies: the role of stellar feedback in gas accretion". In: Monthly Notices of the Royal Astronomical Society 524.3, pp. 4091–4108.

Barbani, F. et al. (2025). "Understanding the baryon cycle: Fueling star formation via inflows in Milky Way-like galaxies". In: Astronomy & Astrophysics 697, A121, A121.

Barnes, J. and P. Hut (1986). "A hierarchical O(N log N) force-calculation algorithm". In: Nature 324.6096, pp. 446–449.

Beck, R. (2011). "Magnetic Fields in Galaxies". In: Large-Scale Magnetic Fields in the Universe. Springer New York, pp. 215–230.

Begeman, K.G. (1989). "HI rotation curves of spiral galaxies. I. NGC 3198." In: Astronomy & Astrophysics 223, pp. 47–60.

Behroozi, P. et al. (2019). "UNIVERSEMACHINE: The correlation between galaxy growth and dark matter halo assembly from z = 0-10". In: Monthly Notices of the Royal Astronomical Society 488.3, pp. 3143–3194.

Benson, A.J. et al. (2002). "The effects of photoionization on galaxy formation - I. Model and results at z=0". In: Monthly Notices of the Royal Astronomical Society 333.1, pp. 156–176.

Bilimogga, P.V. et al. (2022). "Using EAGLE simulations to study the effect of observational constraints on the determination of H I asymmetries in galaxies". In: Monthly Notices of the Royal Astronomical Society 513.4, pp. 5310–5327.

Bish, H.V. et al. (2019). "Galactic Gas Flows from Halo to Disk: Tomography and Kinematics at the Milky Way's Disk–Halo Interface". In: The Astrophysical Journal 882.2, p. 76.

Blitz, L. and E. Rosolowsky (2006). "The role of pressure in GMC formation II: the H2-pressure relation". In: The Astrophysical Journal 650.2, pp. 933–944.

Bogdán, Á. et al. (2017). "Probing the Hot X-Ray Corona around the Massive Spiral Galaxy, NGC 6753, Using Deep XMM-Newton Observations". In: The Astrophysical Journal 850.1, p. 98.

Boomsma, R. (2007). "The disk-halo connection in NGC 6946 and NGC 253". PhD thesis. University of Groningen.

Boomsma, R. et al. (2004). High velocity HI in NGC 6946 and extra-planar gas in NGC 253.

Bregman, J.N. (1980). "The galactic fountain of high-velocity clouds". In: The Astrophysical Journal 236, pp. 577–591.

– (2007). "The Search for the Missing Baryons at Low Redshift". In: The Annual Review of Astronomy & Astrophysics 45.1, pp. 221–259.

Cárcamo, M. et al. (2018). "Multi-GPU maximum entropy image synthesis for radio astronomy". In: Astronomy and computing 22, pp. 16–27.

Chabrier, G. (2001). "The Galactic Disk Mass Function: Reconciliation of the Hubble Space Telescope and Nearby Determinations". In: The Astrophysical Journal 554, pp. 1274–1281.

Chengalur, J.N. et al. (2013). "Accurate measurement of the H I column density from H I 21 cm absorption-emission spectroscopy". In: Monthly Notices of the Royal Astronomical Society 432.4, pp. 3074–3079.

Cimatti, A. et al. (2019). Introduction to Galaxy Formation and Evolution. From Primordial Gas to Present-Day Galaxies.

Cioffi, D. et al. (1988). "Dynamics of radiative supernova remnants". In: The Astrophysical Journal 334, pp. 252–265.

Collins, J.A. et al. (2002). "Kinematics of Diffuse Ionized Gas Halos: A Ballistic Model of Halo Rotation". In: The Astrophysical Journal 578.1, pp. 98–108.

Conroy, C. and R.H. Wechsler (2009). "Connecting Galaxies, Halos, and Star Formation Rates Across Cosmic Time". In: The Astrophysical Journal 696.1, pp. 620–635.

Davé, R. et al. (2019). "simba: Cosmological simulations with black hole growth and feedback". In: Monthly Notices of the Royal Astronomical Society 486.2, pp. 2827–2849.

Di Teodoro, E.M. and F. Fraternali (2015). "3D BAROLO: a new 3D algorithm to derive rotation curves of galaxies". In: Monthly Notices of the Royal Astronomical Society 451.3, pp. 3021–3033.

Ferland, G.J. et al. (1998). "CLOUDY 90: Numerical Simulation of Plasmas and Their Spectra". In: The Publications of the Astronomical Society of the Pacific 110.749, pp. 761–778.

Field, G.B. et al. (1969). "Cosmic-Ray Heating of the Interstellar Gas". In: Astrophysical Journal 155, p. L149.

Fraternali, F. and J.J. Binney (2006). "A dynamical model for the extraplanar gas in spiral galaxies". In: Monthly Notices of the Royal Astronomical Society 366.2, pp. 449–466.

– (2008). "Accretion of gas on to nearby spiral galaxies". In: Monthly Notices of the Royal Astronomical Society 386.2, pp. 935–944.

Fraternali, F. et al. (2002). "Deep H I Survey of the Spiral Galaxy NGC 2403". In: The Astronomical Journal 123.6, p. 3124.

Freeman, K.C. (1970). "On the Disks of Spiral and S0 Galaxies". In: The Astrophysical Journal 160, p. 811.

Heald, G.H. et al. (2006). "Imaging Fabry-Perot Spectroscopy of NGC 5775: Kinematics of the Diffuse Ionized Gas Halo". In: The Astrophysical Journal 636.1, pp. 181–199.

Hernquist, L. (1990). "An Analytical Model for Spherical Galaxies and Bulges". In: The Astrophysical Journal 356, p. 359.

Hickstein, D.D. et al. (2019). "A direct comparison of high-speed methods for the numerical Abel transform". In: Review of Scientific Instruments 90.6.

Hoopes, C.G. et al. (1999). "Diffuse Ionized Gas in Edge-on Spiral Galaxies: Extraplanar and Outer Disk Hα Emission". In: The Astrophysical Journal 522.2, pp. 669–685.

Hopkins, P.F. et al. (2018). "FIRE-2 simulations: physics versus numerics in galaxy formation". In: Monthly Notices of the Royal Astronomical Society 480.1, pp. 800–863.

Jenkins, E.B. and T.M. Tripp (2001). "The Distribution of Thermal Pressures in the Interstellar Medium from a Survey of C I Fine-Structure Excitation". In: The Astrophysical Journal Supplement Series 137.2, pp. 297–340.

Jones, C. and W.R. Forman (1984). "The Structure of Clusters of Galaxies Observed with Einstein". In: The Astrophysical Journal 276, pp. 38–55.

Kalberla, P.M.W. et al. (2005). "The Leiden/Argentine/Bonn (LAB) Survey of Galactic HI: Final data release of the combined LDS and IAR surveys with improved stray-radiation corrections". In: Astronomy & Astrophysics 440.2, pp. 775–782.

Karakas, A.I. (2010). "Updated stellar yields from asymptotic giant branch models". In: Monthly Notices of the Royal Astronomical Society 403.3, pp. 1413–1425.

Kim, C.-G. and E.C. Ostriker (2018). "Numerical Simulations of Multiphase Winds and Fountains from Star-forming Galactic Disks. I. Solar Neighborhood TIGRESS Model". In: The Astrophysical Journal 853.2, 173, p. 173.

Lee, S.-W. et al. (2001). "NGC 5775: Anatomy of a disk-halo interface". In: Astronomy & Astrophysics 377, pp. 759–777.

Lelli, F. et al. (2012). "Dynamics of starbursting dwarf galaxies. II. UGC 4483". In: Astronomy & Astrophysics 544, A145, A145.

Lemasle, B. et al. (2018). "Milky Way metallicity gradient from Gaia DR2 F/1O double-mode Cepheids". In: Astronomy & Astrophysics 618, A160.

Leroy, A.K. et al. (2013). "Molecular gas and star formation in nearby disk galaxies". In: The Astronomical Journal 146.2, 19, p. 19.

Li, A. et al. (2021). "A kinematic analysis of ionized extraplanar gas in the spiral galaxies NGC 3982 and NGC 4152". In: Monthly Notices of the Royal Astronomical Society 504.2, pp. 3013–3028.

Marasco, A. and F. Fraternali (2010). "Modelling the HI halo of the Milky Way". In: Astronomy & Astrophysics 525, A134.

Marasco, A. et al. (2019). "HALOGAS: the properties of extraplanar HI in disc galaxies". In: Astronomy & Astrophysics 631, A50.

Marasco, A. et al. (2025). "HI within and around observed and simulated galaxy discs: Comparing MeerKAT observations with mock data from TNG50 and FIRE-2". In: Astronomy & Astrophysics 697, A86, A86.

Marinacci, F. et al. (2011). "Galactic fountains and the rotation of disc-galaxy coronae". In: Monthly Notices of the Royal Astronomical Society 415.2, pp. 1534–1542.

Marinacci, F. et al. (2019). "Simulating the interstellar medium and stellar feedback on a moving mesh: implementation and isolated galaxies". In: Monthly Notices of the Royal Astronomical Society 489.3, pp. 4233–4260.

Matthews, L.D. and K. Wood (2003). "High-Latitude H I in the Low Surface Brightness Galaxy UGC 7321". In: The Astrophysical Journal 593.2, pp. 721–732.

McKee, C.F. and M.R. Krumholz (2009). "The Atomic-to-Molecular Transition in Galaxies. III. A New Method for Determining the Molecular Content of Primordial and Dusty Clouds". In: The Astrophysical Journal 709.1, pp. 308–320.

McKee, C.F. and E.C. Ostriker (2007). "Theory of Star Formation". In: Annual Review of Astronomy and Astrophysics 45.1, pp. 565–687.

Moster, B.P. et al. (2011). "The effects of a hot gaseous halo in galaxy major mergers". In: Monthly Notices of the Royal Astronomical Society 415.4, pp. 3750–3770.

Nelder, J.A. and R. Mead (1965). "A simplex method for function minimization". In: The computer journal 7.4, pp. 308–313.

Nelson, D. et al. (2021). The IllustrisTNG Simulations: Public Data Release.

Oman, K.A. (2024). "MARTINI: Mock Array Radio Telescope Interferometry of the Neutral ISM". In: The Journal of Open Source Software 9.98, 6860, p. 6860.

Oosterloo, T. et al. (2007). "The Cold Gaseous Halo of NGC 891". In: The Astronomical Journal 134.3, p. 1019.

Pakmor, R. et al. (2016). "Improving the convergence properties of the moving-mesh code AREPO". In: Monthly Notices of the Royal Astronomical Society 455.1, pp. 1134–1143.

Pelupessy, F.I. et al. (2012). "N-body integrators with individual time steps from Hierarchical splitting". In: New Astronomy 17.8, pp. 711–719.

Puche, D. and D. Westpfahl (1994). "Neutral Hydrogen Observations of Nearby Dwarf Galaxies". In: European Southern Observatory Conference and Workshop Proceedings. Ed. by G. Meylan and P. Prugniel. Vol. 49, p. 273.

Putman, M.E. et al. (2012). "Gaseous Galaxy Halos". In: The Annual Review of Astronomy & Astrophysics 50, pp. 491–529.

Rahmati, A. et al. (2013). "On the evolution of the Hi column density distribution in cosmological simulations". In: Monthly Notices of the Royal Astronomical Society 430.3, pp. 2427–2445.

Rogstad, D.H. et al. (1974). "Aperture-synthesis observations of HI in the galaxy M83." In: Astrophysical Journal, Vol. 193, p. 309-319 193, pp. 309–319.

Roper, F.A. et al. (2023). "The diversity of rotation curves of simulated galaxies with cusps and cores". In: Monthly Notices of the Royal Astronomical Society 521.1, pp. 1316–1336.

Rossa, J. and R.-J. Dettmar (2003). "An Hα survey aiming at the detection of extraplanar diffuse ionized gas in halos of edge–on spiral galaxies-II. The Hα survey atlas and catalog". In: Astronomy & Astrophysics 406.2, pp. 505–525.

Rybicki, G.B. and A.P. Lightman (1986). Radiative Processes in Astrophysics.

Sancisi, R. and R.J. Allen (1979). "Neutral hydrogen observations of the edge-on disk galaxy NGC 891." In: Astronomy & Astrophysics 74, pp. 73–84.

Schaap, W. et al. (2000). The vertical extent and kinematics of the HI in NGC 2403.

Schaye, J. et al. (2015). "The EAGLE project: simulating the evolution and assembly of galaxies and their environments". In: Monthly Notices of the Royal Astronomical Society 446.1, pp. 521–554.

Shapiro, P.R. and G.B. Field (1976). "Consequences of a New Hot Component of the Interstellar Medium". In: The Astrophysical Journal 205, pp. 762–765.

Sofue, Y. (2016). "Rotation and mass in the Milky Way and spiral galaxies". In: Publications of the Astronomical Society of Japan 69.1.

Springel, V. (2010). "E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh". In: Monthly Notices of the Royal Astronomical Society 401.2, pp. 791–851.

Springel, V. and L. Hernquist (2003). "Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation". In: Monthly Notices of the Royal Astronomical Society 339.2, pp. 289–311.

Springel, V. et al. (2005). "Modelling feedback from stars and black holes in galaxy mergers". In: Monthly Notices of the Royal Astronomical Society 361.3, pp. 776–794.

Springob, C.M. et al. (2005). "A Digital Archive of H I 21 Centimeter Line Spectra of Optically Targeted Galaxies". In: The Astrophysical Journal 160.1, pp. 149–162.

Swaters, R.A. et al. (1997). "The HI Halo of NGC 891". In: The Astrophysical Journal 491.1, pp. 140–145.

Thompson, A.R. et al. (2017). Interferometry and Synthesis in Radio Astronomy, 3rd Edition, pp. 528–530.

Tüllmann, R. et al. (2006). "The multi-phase gaseous halos of star-forming late-type galaxies: II. Statistical analysis of key parameters". In: Astronomy & Astrophysics 457.3, pp. 779–785.

Vogelsberger, M. et al. (2012). "Moving mesh cosmology: numerical techniques and global statistics". In: Monthly Notices of the Royal Astronomical Society 425.4, pp. 3024–3057.

Vogelsberger, M. et al. (2013). "A model for cosmological simulations of galaxy formation physics". In: Monthly Notices of the Royal Astronomical Society 436.4, pp. 3031–3067.

Weinberger, R. et al. (2020). "The AREPO Public Code Release". In: The Astrophysical Journal Supplement Series 248.2, p. 32.

Weltman, A. et al. (2020). "Fundamental physics with the Square Kilometre Array". In: Publications of the Astronomical Society of Australia 37.

Whiting, M.T. (2012). "duchamp: a 3D source finder for spectral-line data: The duchamp source finder". In: Monthly Notices of the Royal Astronomical Society 421.4, pp. 3242–3256.

Wilson, T.L. et al. (2013). Tools of Radio Astronomy. Springer Berlin, Heidelberg, pp. 177–185.

Wolfire, M.G. et al. (2003). "Neutral Atomic Phases of the Interstellar Medium in the Galaxy". In: Astrophysical Journal 587.1, pp. 278–311.

Appendix

A Main MARTINI functions

Here we present the output of the `MARTINI` functions used during data processing and the generation of mock HI observations. The key functions include:

MARTINI.source() – This function initializes the source data from the simulation snapshot. It reads the gas properties, such as positions, velocities, densities, temperatures, and metallicities, and prepares them for the computation of the HI emission. It also filters out non-relevant particles (e.g., hot gas above a given temperature threshold) and can apply additional selection criteria such as radial or vertical cutoffs. An example of the output provided by this function is provided below:

```
Smuggle v2: Using local cutout file snapshot_200.hdf5
Source module contained 10468698 particles with total HI mass of
                                                        4.96e+09 solMass.
Pruned particles that will not contribute to data cube, 1359973
             particles remaining with total HI mass of 4.95e+09 solMass.
```

MARTINI.insert_source_in_cube() – This function inserts the processed source emission into the data cube. It computes the contribution of each gas cell to the flux in the spatial and spectral bins of the cube, based on its velocity, position, and emission strength. The flux is integrated over velocity channels using a line profile (typically a Gaussian), taking into account thermal broadening and turbulent motions. The output of this functions includes information on the integrated and maximum HI flux, the selected distance and the total HI mass, as listed in the following example:

```
Source inserted.
Flux density in cube: 2.83e+01 Jy
Mass in cube (assuming distance 9.50 Mpc and a spatially resolved
                                                              source):
```

```
4.95e+09 solMass [100% of initial source mass]
Maximum pixel: 4.83e-04 Jy / arcsec2
```

MARTINI.add_noise() – Adds noise to the data cube to simulate realistic observational conditions and outputs the relevant properties of such noise as in the example provided below. The noise level can be set according to observational parameters (e.g., system temperature, integration time) or specified manually. This function ensures that the resulting synthetic data cube includes the effects of random fluctuations, as found in real radio observations.

```
Noise added.
Noise cube RMS: 2.33e-06 Jy / arcsec2 (before beam convolution).
Data cube RMS after noise addition (before beam convolution):
                                                    3.04e-06 Jy / arcsec2
```

MARTINI.convolve_beam() – Convolves the HI data cube with a 2D Gaussian kernel representing the telescope's beam. This simulates the resolution of a radio telescope and blurs the emission to match the spatial response of the instrument. The beam size (FWHM in arcseconds) is a critical parameter, typically derived from real observations (e.g., of NGC891), to make the synthetic data comparable to real data. The output of this function reports importnat properties of the data cube after the convolution has been performed, as can be seen below.

```
Beam convolved.
Data cube RMS after beam convolution: 1.18e-03 Jy / beam
Maximum pixel: 5.66e-02 Jy / beam
Median non-zero pixel: 7.42e-05 Jy / beam
```

B Observables

B.1 Alternative methods to derived hydrogen molecular fraction

To assess the reliability of the hydrogen molecular fraction values provided by the SMUGGLE model, we compared our results to those derived from an empirical method based on observations of nearby galaxies (Leroy et al., 2013). In this empirical approach, the molecular fraction is defined as:

$$f_{\mathrm{mol}} = \frac{R_{\mathrm{mol}}}{R_{\mathrm{mol}} + 1} \tag{B.1}$$

where $R_{\rm mol} = \left(\frac{P}{P_0}\right)^{\alpha}$, with P as the mid-plane pressure of the gas, and $P_0 = 1.7 \times 10^4\,{\rm K\,cm^{-3}}$ and $\alpha = 0.8$ free parameters of the model that we chose according to Blitz and Rosolowsky (2006). The impact of this prescription on the derived neutral hydrogen distribution is presented in Fig. B.1, where it is compared to our default approach.

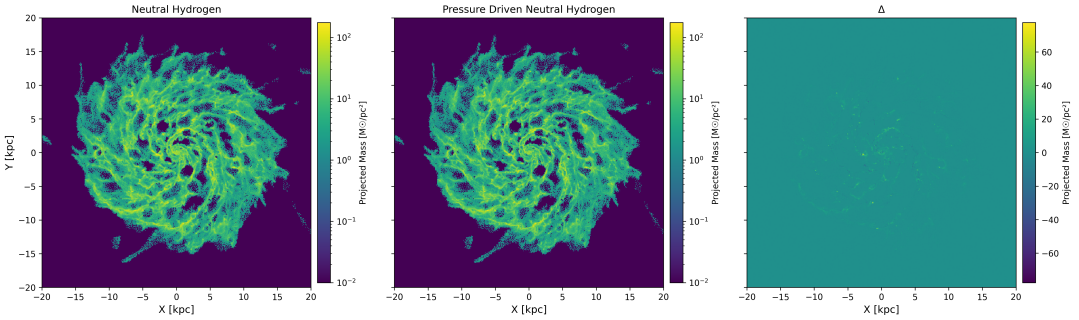


Figure B.1: Projected HI mass along the line of sight for a face-on galaxy, calculated using two different prescriptions for the molecular hydrogen fraction. Left: HI map using the molecular hydrogen fraction from the SMUGGLE output. Middle: HI map using the empirical model by Leroy et al. (2013). Right: Difference between the two maps. Resolution: 410×410 bins with a size of 97.56 pc.

Qualitatively, the comparison does not reveal significant differences between the two models. The large-scale HI distribution patterns remain similar, and the normalization levels are comparable. The difference map shows no substantial residuals. However, there is a minor quantitative discrepancy: the empirical model underestimates the total HI mass by approximately 9% compared to the SMUGGLE-based calculation. Nonetheless, we conclude that both approaches yield consistent results, and we adopt the molecular fraction provided by SMUGGLE for simplicity.

B.2 Vertical distribution of HI

In Section 4.1, we presented the normalized vertical distribution of HI column density across four regions of the galaxy (Fig. 4.3). Figure B.2 illustrates how the galaxy was divided into these four quadrants. We measured mean values of column density as a function of z and divided by the maximum emission on the disc in each quadrant, to find the normalized vertical distribution of HI gas.

B.3 Simulations without a hot corona

In Fig. B.3, we show the HI column density map of the edge-on galaxy simulated without a hot corona, using the parameters listed in Table B.1. Following the procedure outlined

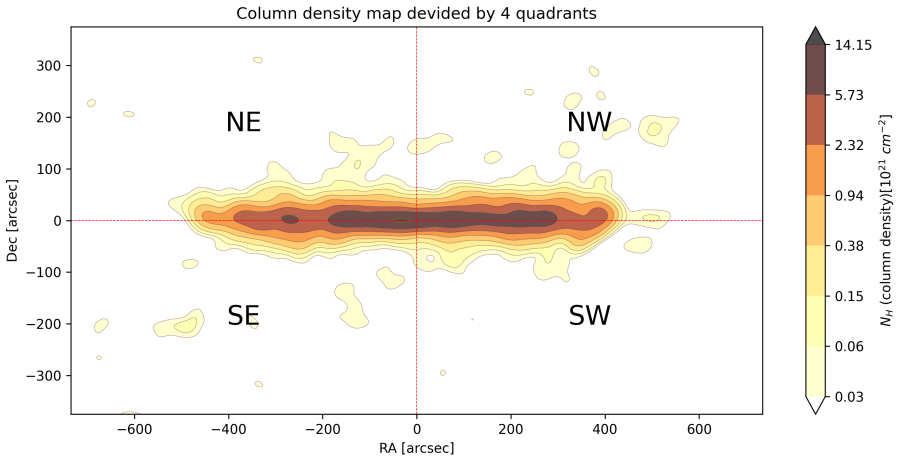


Figure B.2: HI column density map of the edge-on galaxy. Red dashed lines represents division of the map for four quadrants – NE, NW, SE, SW – as indicated in the figure.

in Appendix B.2, we derived the corresponding vertical distribution of HI in this case as well. A comparison with the configuration including the hot corona reveals notable differences in the spatial structure: the galaxy without a halo exhibits a broader radial extent, while its vertical distribution is significantly more compact. This emphasizes the role of the interaction between the gas ejected from the disc and the hot CGM in vertically thickening the neutral hydrogen layer and influencing the overall morphology of the galaxy.

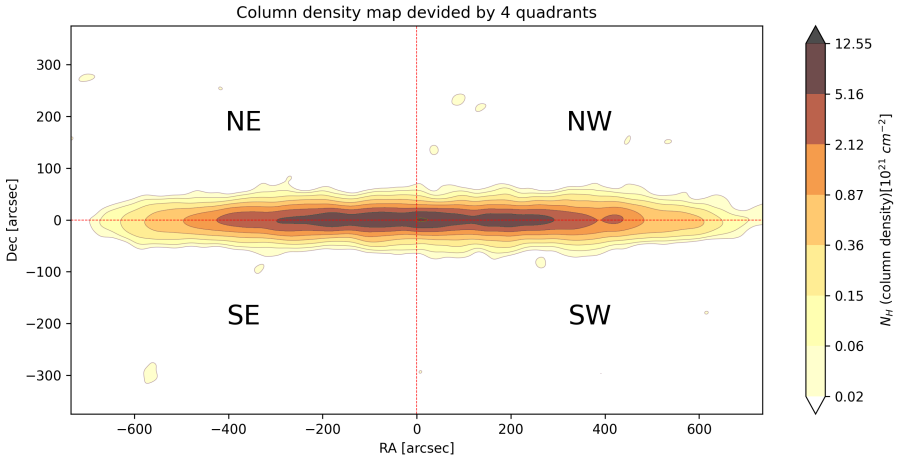


Figure B.3: The same as Fig. B.2, but for a simulation that does not include the hot corona.

Parameters of galaxy without halo

R_{200} (kpc)	$M_{\rm dm}$ ($M_\odot$)	r_s (kpc)	$M_{\rm b}$ ($M_\odot$)	a (kpc)	$M_\star$ ($M_\odot$)	$r_\star$ (kpc)	h (pc)	$M_{\rm g}$ ($M_\odot$)	$r_{\rm g}$ (kpc)	$f_{\rm gas}$ (R<$R_\odot$)
241	1.53×10^{12}	36.46	1.5×10^{10}	1.0	4.73×10^{10}	3.0	300	9×10^{9}	6.0	0.10

Table B.1: Structural parameters of the galaxy simulation that does not include a hot corona component. $\boldsymbol{R_{200}}-$ radius where density equal to 200 times of critical density. $\boldsymbol{M}_{\rm dm}-$ dark matter halo mass. $\boldsymbol{r_s}-$ dark matter halo scale length. $\boldsymbol{M}_{\rm b}-$ bulge mass. $\mathbf{a}-$ bulge scale length. $\boldsymbol{M_\star}-$ stellar disc mass. $\boldsymbol{r_\star}-$ stellar disc scale length. $\mathbf{h}-$ stellar disk height. $\boldsymbol{M}_{\rm g}-$ gaseous disc mass. $\boldsymbol{r}_{\rm g}-$ gaseous disc scale length. $\boldsymbol{f}_{\rm gas}-$ gas fraction within R=8.5 kpc.

HI vertical distribution for the galaxy non incuding a hot corona

In Fig. B.4, we present the normalized vertical distribution of the HI gas for the simulated galaxy without a hot corona component. Compared to the distribution shown in Fig. 4.3, which includes the coronal gas, the vertical extent of the HI gas is noticeably reduced. Specifically, in the galaxy with a hot halo, the emission above the $\sigma_{\rm rms}$ threshold extends up to $1.5'$ in angular size, whereas in the halo-free case it reaches only about $1.2'$. This corresponds to a physical difference of approximately 800 pc. This comparison highlights the contribution of the halo gas to the overall vertical thickening of the neutral hydrogen distribution.

B.4 Theoretical rotation curve of the simulated galaxy

In Fig. B.5, we present the theoretical rotation curve of the modeled galaxy, calculated using the input parameters listed in Table 3.2. The individual contributions of the galaxy's components—dark matter, bulge, gas, and stars—are shown, each influencing the overall velocity profile.

The total circular velocity v_c is computed by summing the squared contributions from each mass component:

$$v_c(R) = \sqrt{v_{\rm dm}^2 + v_b^2 + v_g^2 + v_\star^2} \tag{B.2}$$

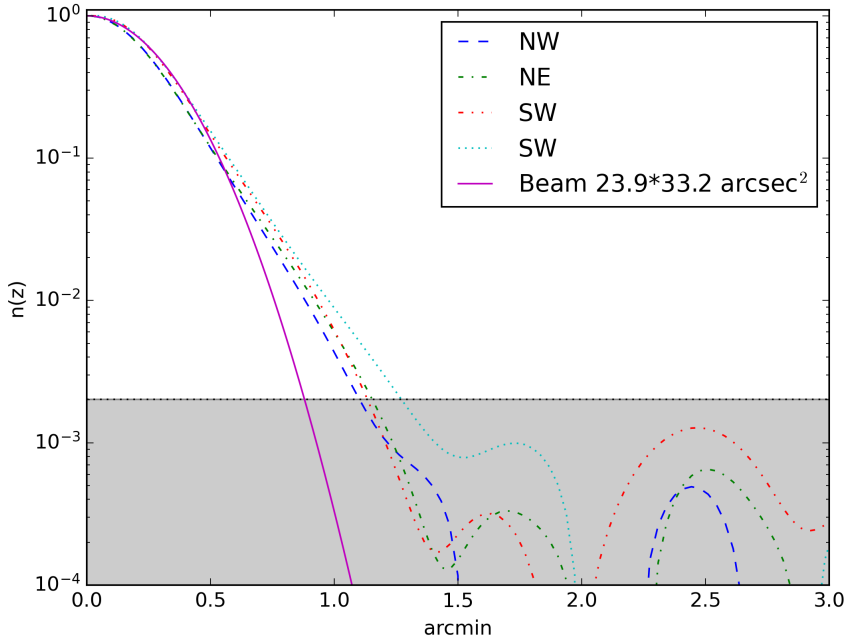


Figure B.4: Normalized HI vertical distribution of the edge-on galaxy without a hot corona. The grey box is the region below the $\sigma_{rms,\mathrm{HI}}$.

The individual terms are defined as:

$$v^2_{\mathrm{dm}}(R) = \frac{GM_{\mathrm{dm}}(< R)}{R} \tag{B.3a}$$

$$v^2_b(R) = \frac{GM_b(< R)}{R} \tag{B.3b}$$

$$v^2_g(R) = \frac{2GM_g}{h_g} y^2 \left[I_0(y)K_0(y) - I_1(y)K_1(y)\right] \tag{B.3c}$$

$$v^2_\star(R) = \frac{2GM_\star}{h_\star} y^2 \left[I_0(y)K_0(y) - I_1(y)K_1(y)\right] \tag{B.3d}$$

Here, $y = R/(2h)$, where h represents the radial scale length for the gas (h_g) or stellar ($h_\star$) disc. The functions I_n and K_n are the modified Bessel functions of the first and second kind, respectively. This formalism allows the decomposition of the rotation curve into physically motivated components, highlighting the relative dynamical importance of each.

B.5 Abel transformation

The Abel transformation is an integral method used to de-project cylindrically symmetric quantities from their line-of-sight projections. In essence, it relates the observed surface brightness $I(R)$ at projected radius R to the intrinsic emissivity profile $j(r)$ at radius r,

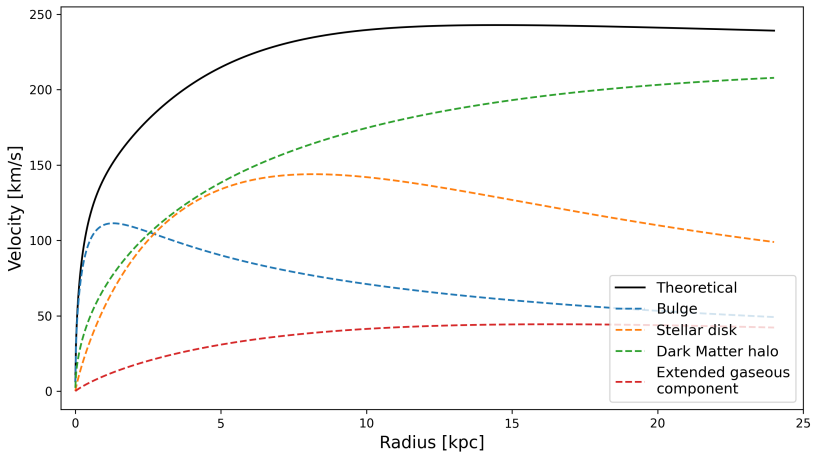


Figure B.5: Modeled rotation curve of the simulated galaxy analysed in this work, showing contributions from dark matter, bulge, gas, and stars.

assuming axisymmetry. The inverse Abel transform recovers $j(r)$ via:

$$j(r) = -\frac{1}{\pi} \int_r^\infty \frac{dI}{dR} \frac{dR}{\sqrt{R^2 - r^2}}.$$

This allows one to reconstruct the true three-dimensional distribution of emitting material from the observed two-dimensional brightness profile in the edge-on systems.

C Comparing fitting procedures of ^{3D}BAROLO

We compare the results obtained with ^{3D}BAROLO for a galaxy inclined at 30°, using both available normalization methods for fitting: azimuthal averaging (AZIM) and pixel-by-pixel normalization (LOCAL). The resulting moment-0 maps are displayed in Fig. C.6, while the fitted kinematic parameters are shown in Fig. C.7.

As illustrated in Fig. C.6, the residuals are noticeably reduced when using pixel-by-pixel normalization, suggesting a better fit to the observed flux distribution. However, the derived kinematic parameters remain largely consistent between the two methods, with the exception of the velocity dispersion. Although some deviations in the velocity dispersion profiles are observed, these are minor and follow a similar overall trend. Therefore, while pixel-by-pixel normalization improves the residual distribution, the choice of normalization does not significantly alter the kinematic interpretation.

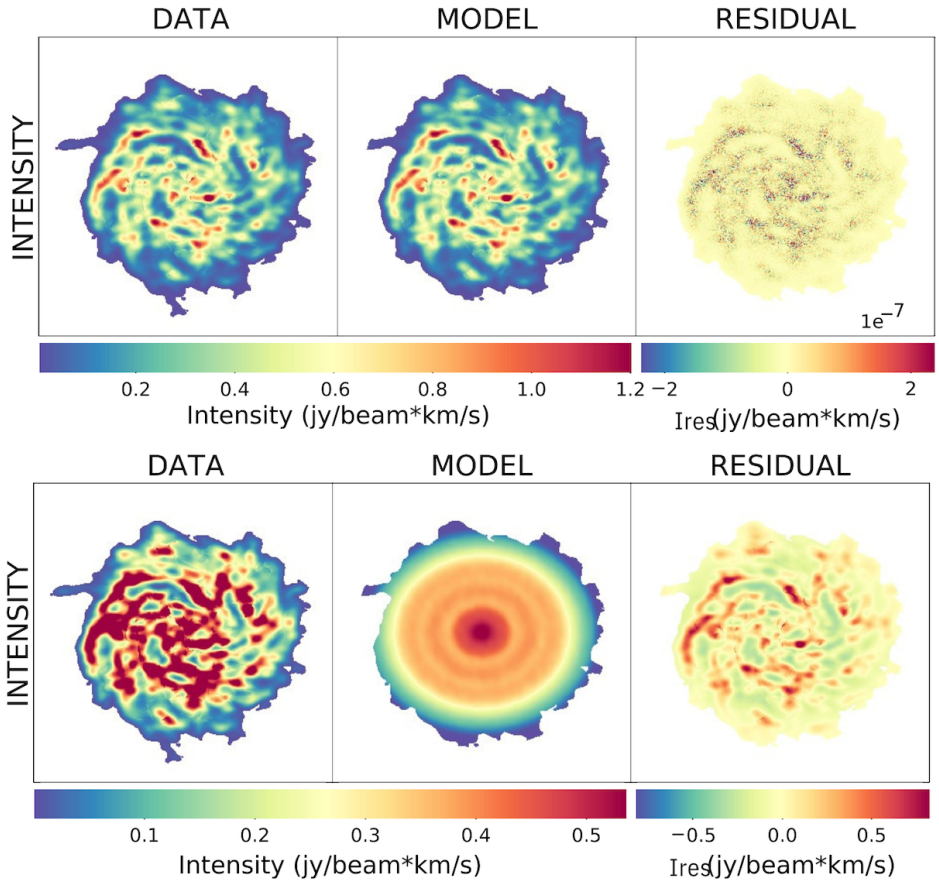


Figure C.6: Moment-0 maps obtained with ^{3D}BAROLO using pixel-by-pixel (top) and azimuthal (bottom) normalization. Left column: observational data; middle column: fitted model; right column: residuals.

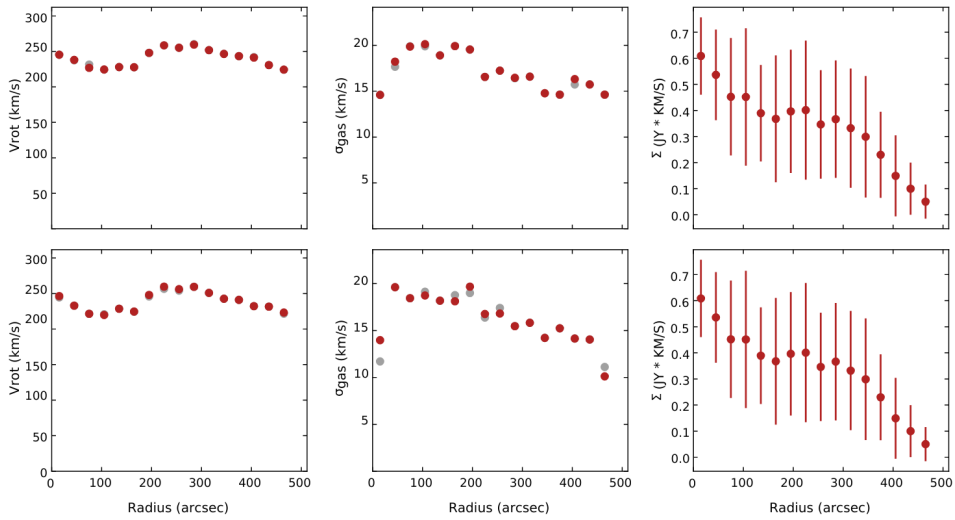


Figure C.7: Kinematic properties and HI surface brightness profile inferred with BAROLO, with pixel-by-pixel (top) and azimuthal (bottom) normalization. Left column: rotational velocity. Middle column: velocity dispersion. Right column: surface brightness.

Acknowledgements

First and foremost, I would like to express my sincere gratitude to Professor Federico Marinacci for his invaluable guidance, support, and above all, his patience throughout my thesis work. Over this time, I have witnessed his high level of professionalism, and there is no doubt about the quality of his work – he is truly a master of his field.

I am deeply grateful to my parents for their unconditional love and encouragement, which have been my foundation. They supported me from the beginning of my journey and were patient with all of my decisions – some of which were not entirely logical.

Special thanks go to my friends – both those here in Bologna ("It. banda") and those back in Kazakhstan – for their constant support and companionship during this journey. I will not name everyone — and there are too many, surprisingly, simply because I do not want to accidentally forget someone and cause offense. But of course, I must mention those with whom I've lived for the past two years, and I have absolutely no regrets. Adilet and Azamat, I wish you a swift completion of your studies and all the best for the future.

I would also like to acknowledge my colleagues who accompanied me along this journey – especially Krishna and Thimea, with whom I shared the demanding workload of the Multiwavelength Astrophysics Laboratory.

My heartfelt gratitude also goes to the "P.gnezdo" who have always supported me and remembered me at every opportunity. Let us not forget their wives, who have shown me no less kindness. There are also individuals who deserve special mention. My gratitude goes to Dias, a true "Zhylan" – not easy to find someone like him – who has picked me up and dropped me off more times than I can count (>10). Artem, even though we haven't seen each other in years, we've maintained a close connection. Unfortunately, your trip to Italy didn't happen, but I hope that was only so that one day you can come here with your family.

Finally, I would like to thank Karakat, whose presence brought joy and warmth to my daily life and helped brighten my routine.