

Alma Mater Studiorum – Università di Bologna

Dipartimento di Fisica e Astronomia
Corso di Laurea Magistrale in Astrofisica e Cosmologia

Identification and analysis of super-bubbles candidates in Milky Way-like galaxy simulations

Tesi di Laurea

Presentata da:
Saverio F. Colla

Relatore:
Dott. Federico Marinacci

Anno accademico 2019/2020

Contents

Abstract	3
Sommario	4
1 Introduction	5
1.1 Aim of the investigation	6
1.2 Super-Bubble theory	7
1.2.1 SB dynamics	7
1.2.2 Supershells evidence from observations	10
2 Simulating an isolated Milky Way-like galaxy	13
2.1 Brief description of AREPO code	13
2.1.1 Equations solved	14
2.1.2 Gravitational interactions	16
2.1.3 Magnetohydrodynamics	18
2.1.4 Computational mesh	20
2.1.5 Time integration	22
2.2 The SMUGGLE model	24
2.2.1 Physical methods	25
2.2.2 Cooling and heating	27
2.2.3 Star Formation implementation	28
2.2.4 Feedback from SNe	30
2.2.5 Radiative feedback from young massive stars	35
2.2.6 Feedback from OB and AGB stellar winds	37
2.2.7 Numerical setup of the analyzed simulation	38

3	The Astrodendro Python package	41
3.1	Dendrogram structure	41
3.2	Computing a dendrogram	42
3.3	Statistical properties	44
4	Results	47
4.1	Structures identification with Astrodendro	47
4.1.1	Stellar surface density map	48
4.1.2	Gas density map	49
4.1.3	Maps comparison	54
4.1.4	Parameters investigation	58
4.1.5	A frequentist approach	62
4.1.6	Identification using the Metropolis-Hasting MCMC	64
4.2	Statistical analysis of the identified structures	70
4.2.1	Super-Bubble radius	71
4.2.2	Age of the superbubble	72
4.2.3	Mass swept by the SB	75
4.2.4	SB histograms	77
4.2.5	Size vs. number of SuperNovae relation	79
4.3	Superbubble evolution	82
4.3.1	Bubble density profile	82
4.3.2	Radius evolution	83
4.3.3	SN-radius relation	89
5	Final discussion	94
5.1	Conclusions	95
5.1.1	SB identification	95
5.1.2	Comparison between model and theory	96
5.1.3	Future works	97
	Bibliography	99

Abstract

We present a study on the identification and analysis of the main properties of sub-dense gas structures that have an indication of being produced by collective supernovae phenomena, the so-called super-bubbles. These structures are generated in a numerical simulation of an isolated Milky Way-like galaxy performed with the SMUGGLE model for galaxy formation physics. We propose an identification method based on statistical considerations and we measure the properties of the candidate structures (such as radius, expansion rate, age and swept mass) through statistical analysis methods applied to the whole sample of identified structures as well as for a sub-sample of five individual candidates. We investigate the ability of the galactic evolutionary model to establish a relationship between the number of supernovae exploded within a super-bubble and its size, both statistically and individually for some selected structures. Finally, we compare the results obtained in this analysis with the observational data for the Milky Way and other late-type spiral galaxies and with the theory developed for super-bubble evolution.

Sommario

Questo lavoro di tesi presenta uno studio sull'identificazione e l'analisi delle principali proprietà di strutture di gas sottodenso che potrebbero essere prodotte da esplosioni multiple di supernove, le cosiddette "super-bubbles". Tali strutture sono generate dalla simulazione numerica di una galassia simile alla Via Lattea eseguita tramite il modello SMUGGLE che implementa i principali meccanismi fisici alla base dei processi di formazione galattica. Viene proposto un metodo di identificazione basato su considerazioni statistiche che permette di stimare le proprietà (come raggio, tasso di espansione, età, e massa spazzata) delle strutture candidate ad essere possibili super-bubbles e tale approccio viene applicato sia all'intero campione di strutture identificate sia ad un sotto-campione di cinque candidati. Inoltre il presente lavoro di tesi indaga la capacità del modello SMUGGLE di stabilire una relazione tra il numero di supernovae esplose all'interno di una super-bubble e le sue dimensioni, sia in maniera statistica sia considerando singolarmente alcuni particolari strutture. Infine i risultati ottenuti sono stati confrontati sia con i dati osservativi della Via Lattea ed altre galassie a spirale sia con la teoria sviluppata per l'evoluzione delle super-bubbles.

Chapter 1

Introduction

It has long been known how the domain of numerical simulations in the Astrophysical field has acquired a fundamental role in scientific research, bringing to light aspects related to fluid dynamics and gravitational interactions that would otherwise be analytically unreachable. Various simulation schemes have been implemented by research groups all over the world, establishing strong international collaborations, designed to share ideas and resources in developing more and more sophisticated models that can more accurately describe the phenomena that we observe in the Universe.

Due to limited time and resources it is usually necessary to adapt the resolution to be achieved to the dimensional scale of the physical problem being considered, which generally range from large scales of cosmological evolution to smaller scales designed to solve the fluid dynamic turbulence regime, passing through intermediate scales that focus on the physics of galactic evolutionary processes. In order to better understand the latter astrophysical scale, a galactic model (called SMUGGLE—Stars and MULTiphase Gas in GaLaxiEs model, [Marinacci et al. 2019](#)) has recently been implemented. This model is able to reproduce the physical processes that take place within a late-type galaxy such as our spiral galaxy, the Milky Way.

The work presented in this thesis focused on the analysis of a part of the results provided by the SMUGGLE simulation of the evolution of a Milky Way-like galaxy, in order to investigate the phenomenon of interaction between gas and clusters of supernovae, which creates $\sim 100 - 1000$ pc size objects that are usually referred to as “super-bubbles”.

1.1 Aim of the investigation

The primary goal of this thesis is the identification and investigation of the properties of large sub-dense structures called super-bubbles (SB), which are formed in the interstellar medium of a late-type galaxy simulated with the SMUGGLE model. To achieve this goal, a first part of the work focused on developing an appropriate method for analyzing the density maps of the gaseous and stellar components, obtained by elaborating the simulation output data, with the aim of correctly identifying the sub-dense structures to be examined (hereinafter referred to as "*SB candidates*"), which is illustrated in section 4.1. A method based on statistical considerations has been designed and implemented to perform this identification (section 4.1.5). The second part of the work was focused on the analysis of the identified structures, retrieving information from the statistical distributions owned by the SB candidates (section 4.2) and operating an individual evolutionary analysis for some of these candidates (section 4.3). Different properties have been measured for the identified SB candidates, such as the radii of these structures, their expansion velocities, the associated ages and the mass of gas swept during their expansion through the ISM. All previous quantities were then compared with theory and observational data in order to establish whether the galactic model implemented is able to reproduce the formation of these objects in accordance with the theory that has been previously developed and with the observations made in the Milky Way and other spiral galaxies. With the aim of better explaining this comparison between theory and observations with the simulation output data, we briefly describe in this chapter the super-bubble theory (section 1.2) focusing on the SB evolutionary phases (1.2.1) and on the measurements made in observations of the Milky Way and other galaxies (1.2.2).

The thesis is structured as follows. We first review in chapter 2 the numerical methods used by the AREPO code which is adopted to evolve the galaxy with the SMUGGLE model. In the same chapter, we also present in detail the physical processes implemented in SMUGGLE. A description of the analysis package, called Astrodendro, used for the identification and the study of SB candidates is provided in the chapter 3. In chapter 4 we discuss the results obtained from the study of the structures identified as SB candidates. Finally, we conclude the work in chapter 5 with an evaluation of the results obtained and a possible future improvement regarding the techniques used for the data analysis.

1.2 Super-Bubble theory

Repeated supernovae events and stellar winds from OB associations can impact the interstellar medium (ISM) of a spiral galaxy in a very substantially way, creating large hot cavities ($T \gtrsim 10^6$ K) of low-density gas ($n \lesssim 10^{-2} \text{ cm}^{-3}$) with radius greater than 100 pc (McCray and Kafatos, 1986), the so-called Super-Bubbles (SB), which are surrounded by a dense, expanding shell of cool interstellar gas (namely the "supershell").

Type II supernovae (SNII) usually occur in OB associations (rather than being randomly distributed through the galaxy) with a typical event rate of roughly one supernova per million years (McCray and Snow 1979). This rate will remains approximately constant over 50 Myr (the lifetime of the lowest mass B stars likely to become a supernova), if the association has a typical initial mass function (e.g. IMF proposed by Scalo, J. M. 1986 for OB association); at this rate, SN events inject energy into the supershell until it loses its interior pressure and enters the "snow-plow" phase, that happen when radiative cooling becomes important or when the shell bursts through the gas disk of the galaxy (Castor et al., 1975). At approximately the same time, the supershell becomes gravitationally unstable, forming giant molecular clouds, which are sites for new star formation (McCray and Kafatos, 1986). There is abundant evidence for supershells in the Milky Way and other spiral and irregular galaxies in the Local Group from 21 cm emission-line surveys, optical emission-line surveys, and studies of SN remnants (Heiles, 1979; Brinks and Bajaja, 1986; Madsen et al., 2006; Rosie Chen et al., 1999).

1.2.1 SB dynamics

Because supernova blast waves usually become subsonic before reaching the walls of the supershell or cooling radiatively (MacLow & McCray, 1987), it is reasonable to approximate the energy input from SNe as a continuous luminosity, and thus treating SBs as very large stellar wind bubbles with the contribution to the internal energy made up by the entire OB association located at about the center of the bubble itself.

Stellar wind phase: Initially, the power is dominated by the ionizing radiation and stellar wind of the most massive star (usually $\sim 35 M_{\odot}$ type O7 V star), which generates ionizing photons at rate $S_i \sim 7 \cdot 10^{48} \text{ s}^{-1}$ (Panagia, 1973) and stellar wind power of the order of

$$L_w = \frac{1}{2} \dot{M}_w V_w^2 \sim 10^{35} \text{ erg s}^{-1} \quad (1.1)$$

(assuming a reference system of a steady, spherically symmetric stellar wind with constant terminal velocity V_w and mass-loss rate \dot{M}_w), thus giving a total wind energy $E_w \sim 10^{50}$ ergs during its main-sequence lifetime ($t_{MS} \sim 5$ Myr).

The star may release another few times 10^{50} ergs in a strong stellar wind during a subsequent Wolf-Rayet phase before it terminates as a supernova or black hole.

Considering a model in which the ambient ISM consists of gas of uniform atomic density, n_0 , the combined action of all the winds will create a supershell with radius given by (Weaver et al., 1977)

$$R_s = 269 \left(\frac{L_{38}}{n_0} \right)^{1/5} t_7^{3/5} \text{ pc} \quad (1.2)$$

where $L_{38} = L_w^{tot}/(10^{38} \text{ erg}\cdot\text{s}^{-1})$ with L_w^{tot} the combined mechanical luminosity of all the stellar winds in the OB association, and $t_7 = t/(10^7 \text{ yr}) \sim 0.5$ the lifetime of a typical O star. In this way stellar winds alone can create a bubble of large radius (100 - 200 pc) even before the first SN has occurred.

The ionizing radiations and stellar winds power from the OB association then decrease rapidly in a time about the lifetime of the last O star ($t \sim 5$ Myr), leaving the next evolutionary phase to the main energy contribution of SNII events.

Adiabatic expansion phase: The SB radius evolution in the second phase is dominated by the power of all SN explosions that are quickly produced in the span of time between 5 - 50 Myr. It is possible to write the equivalent mechanical luminosity produced by all the SN events occurring in a given small fraction Δt of the time range, as (Weaver et al., 1977)

$$L_{SN} = 3.17 \times 10^{36} N_{SN} \left(\frac{E_{SN}}{10^{51} \text{ erg}} \right) \left(\frac{\Delta t}{10^7 \text{ yr}} \right)^{-1} \text{ erg s}^{-1} \quad (1.3)$$

where N_{SN} is the total number of SNII events and with each supernova explosion producing an energy of about $E_{SN} \sim 10^{51}$ ergs. In that evolutionary phase, the main growth of the supershell will be caused by supernova explosions, which continue to hammer at the shell until $t_7 \approx 5$, long after the last O stars have vanished. Then, if the energy of the hot interior E is conserved, the radius and velocity of the outer shell follow from equation (1.2) with the replacement of $L_w^{tot} = L_{SN}$, thus getting

$$R_{SB} = 1.67 \times 10^2 \left(\frac{L_{SN}}{10^{37} \text{ s}^{-1}} \right)^{1/5} \left(\frac{n_H}{1 \text{ cm}^{-3}} \right)^{-1/5} \left(\frac{\Delta t}{10^7 \text{ yr}} \right)^{3/5} \text{ pc} \quad (1.4)$$

with corresponding velocity

$$V_{SB} = 9.7 \left(\frac{L_{SN}}{10^{37} \text{ erg s}^{-1}} \right)^{1/5} \left(\frac{n_H}{1 \text{ cm}^{-3}} \right)^{-1/5} \left(\frac{\Delta t}{10^7 \text{ yr}} \right)^{-2/5} \text{ km s}^{-1} \quad (1.5)$$

where the ambient density n_0 is converted in hydrogen numerical density by the relation $n_0 = n_H + n_{He}$ (assuming the helium-hydrogen density ratio of $n_{He}/n_H = 0.1$) to highlight the dependence on the density variable n_H that is actually observed in radio maps.

Radiative cooling phase: The adiabatic phase of the supershell persists until radiative cooling becomes important in the hot interior, at a time t_c that depends on the gas metallicity ζ (McCray and Kafatos, 1986)

$$t_c \sim 4 \cdot 10^6 \zeta^{-3/2} \left(N_{SN} \frac{E_{SN}}{10^{51} \text{ erg}} \right)^{0.3} n_0^{-0.7} \text{ yr} \quad (1.6)$$

which is the order of $t_c \sim 30$ Myr for a solar metallicity $\zeta = 1$, ambient density $n_0 = 1 \text{ cm}^{-3}$ and average number of SN events $N_{SN} = 10^3$. At this time the corresponding SB radius (with the same assumptions) is of the order of

$$R_c \sim 50 \cdot \zeta^{-0.9} \left(N_{SN} \frac{E_{SN}}{10^{51} \text{ erg}} \right)^{0.4} n_0^{-0.6} \sim 400 \text{ pc} \quad (1.7)$$

After the adiabatic phase, which ends at about t_c and when the SB radius becomes comparable to the density scale height of the galactic HI layer, Rayleigh-Taylor instabilities begin to act on the polar caps of the shell, distorting them and causing the supershell to “burst” through the HI layer, discharging its internal pressure into the galactic corona. At this point equation (1.4) is no longer valid and the radius in the plane should increase according to the zero-pressure snow-plow law

$$R(t) \sim R_c \left(\frac{t}{t_c} \right)^{1/4} \text{ pc} \quad (1.8)$$

which brings the supershell radius at $R(60 \text{ Myr}) \sim 500 \text{ pc}$ in the next 30 Myr of this evolutionary phase.

Finally, following the results of MacLow & McCray (1987), it is possible to establish whether radiative cooling will be important to the dynamics of superbubbles in disk galaxies. Defining the characteristic dynamical time scale, t_D , by equating the radius of a spherical SB to approximately one stratification scale height H of the galactic disk gas:

$$t_D \approx H^{5/3} \left(\frac{\rho_0}{L_{SN}} \right)^{1/3} \quad (1.9)$$

leads to the comparison with the cooling time scale t_c , thus if $t_D \ll t_c$ the radiative cooling becomes not so important and can be neglected.

The ratio of cooling to dynamical time scales is given by (MacLow & McCray, 1987)

$$\frac{t_c}{t_D} = 13 \zeta^{-1.6} L_{38}^{0.61} \left(\frac{H}{100 \text{ pc}} \right)^{-1.7} n_0^{-1.1}, \quad (1.10)$$

which becomes less than one for $L_{38} < 1.4 \cdot 10^{-2}$. This luminosity value corresponds to a number of SN events of about $N_{SN} \sim 2$ happening in a range of time $\Delta t = 50$ Myr (assuming solar metallicity $\zeta = 1$, ambient density $n_0 = 1 \text{ cm}^{-3}$ and scale height $H = 100$ pc), which is a very small and unlikely number of events. Therefore, SB interior cooling is unimportant for typical Milky Way parameters, but it could become important for denser and cooler ISM, smaller OB associations, or enhanced cooling rates by higher metallicity values.

The main aspects of the super-bubble theory presented in this section will be used as a theoretical reference during the evolutionary analysis of the SB candidates that will be identified by processing the gas density maps from the output data of the SMUGGLE simulation. The theoretical formulas of evolution of the radius (1.4) and of the expansion velocity (1.5) during the adiabatic phase will be used in particular as a term of comparison with the data recovered in following the evolution of some individual objects (see section 4.3). In concluding this chapter we briefly propose in the next section some reference values for the characteristic quantities of the super-bubble obtained from different pieces of observational evidence.

1.2.2 Supershells evidence from observations

There is a widespread evidence for giant shells in the Milky Way and other spiral and irregular galaxies in the Local Group. For instance, the results from 21-cm emission maps for giant HI shells in the Milky Way (Heiles, 1979) show evidence of shells having radii in the range from ~ 100 pc to more than ~ 1 kpc and kinetic energies ranging from $\sim 10^{50}$ ergs to more than $\sim 10^{53}$ ergs. Velocities of the shells have been found in the range $10 - 20 \text{ km s}^{-1}$ (with sometimes small exceeding values), and with associated kinematic ages, $t \approx 0.6 R_{SB}/V_{SB}$, ranging from $5 \cdot 10^6$ yr to $8 \cdot 10^7$ yr.

Brinks and Bajaja (1986) have discovered similar structures in velocity-resolved 21 cm emission-line maps of M31 (Andromeda galaxy). They list 141 giant holes in the HI disk, concentrated at a galactocentric radius ~ 10 kpc, with SB radii typically of the order ~ 125 pc (but in several cases more than ~ 300 pc are observed), and their expansion velocities range from ~ 6 to 20 km s^{-1} . The radii, ages, and kinetic energies of the expanding HI observed shells are consistent with the theory if they are created by OB associations with number of stars approximately in the range $10 \lesssim N_{\star} \lesssim 10^3$.

Another observed feature in agreement with the theory of SBs evolution is the behavior of the supershell's radii in spiral galaxies, which tend to increase with galactocentric radius. This

correlation is explained by the trend of the scale height's profile that increases from the galactic center outwards, in fact the pressure-driven phase (eq. 1.4) of the supershells ends when the shell radius becomes comparable to the scale height of the galactic HI layer, thus leading to larger SBs for positions more distant from the center. For example, in the Milky Way the HI scale height increases from ~ 70 pc in the inner disk (Bruhweiler et al., 1980) to ~ 190 pc in the solar vicinity (Shull & Van Steenberg, 1985) to ~ 530 pc at 20 kpc from the galactic center (Kulkarni, Blitz and Heiles, 1982). Thus, one would expect that most of the supershells in the inner parts would have burst through the disk, leaving "holes" with radius comparable with the disk thickness, while the larger supershells should be found mostly in the outer parts of spiral galaxies, as observed (Kafatos et al., 1980).

Another effect that favors the development of larger shells in the outer parts of spiral galaxies is the dependence of the radiative cooling on metallicity, ζ and ambient density, n_0 . As indicated by equation (1.7), the radius R_c , at which radiative cooling removes the interior pressure increases with decreasing ζ and n_0 . In spiral galaxies both ζ (Pagel et al., 1979) and n_0 decrease with increasing galactocentric radius.

All of these effects conspire to favor the development of supershells in irregular galaxies. The interstellar gas in an irregular galaxy has large scale height and low density as a result of the low mass of the galaxy, and possibly also because the gas layer has been disturbed by tidal interactions with neighboring galaxies. The irregular galaxies tend to have lower metallicity than the giant spirals. Thus, for example, for the Large Magellanic Cloud (LMC), with $\zeta \sim 0.3$ (Dufour, 1984) and $n_0 \sim 0.35 \text{ cm}^{-3}$ (Hindman, 1967), equation (1.7) gives

$$R_c \sim 2.5 \left(\frac{N_* E_{51}}{200} \right)^{0.4} \text{ kpc} \quad (1.11)$$

for the radius at which radiative losses become important. These are the main reasons why the LMC contains so many spectacular supershells as the one shown in figure 1.1.



Figure 1.1: Superbubble in LHA 120-N 44 nebula, surrounding the star cluster NGC 1929. A large amount of material is expanding outwards due to the influence of the cluster of young stars at its heart.

Credit:ESO/Manu Mejias (<https://www.eso.org/public/images/eso1125a/>).

Chapter 2

Simulating an isolated Milky Way-like galaxy

2.1 Brief description of AREPO code

The evolution of the Milky Way-like galaxy analyzed in this work, and modelled with the SMUGGLE model (described in section 2.2), has been simulated by using the multi-purpose gravitational and finite-volume magnetohydrodynamical (MHD) AREPO code (Springel, 2010). The latter allows to solve the Poisson equation (either on a Newtonian or cosmologically expanding space-time) together with the complete set of MHD equations on an unstructured, dynamic Voronoi tessellation coupled to a tree-particle-mesh algorithm. Due to its more accurate and flexible numerical schemes, together with high performing efficiency, AREPO code has become very versatile over time and has been adopted to develop a wide range of astrophysical problems. These include cosmological simulations of galaxy formation in large volumes such as Illustris (Vogelsberger et al., 2014; Genel et al., 2014) and IllustrisTNG (Naiman et al., 2018; Nelson et al., 2018a, 2019; Marinacci et al., 2018b; Springel et al., 2018; Pillepich et al., 2018, 2019), cosmological zoom simulations of galaxy formation such as Auriga (Grand et al., 2017), isolated galaxies (Jacob et al., 2018), spiral structure in galaxies (Smith et al., 2014), stratified box simulations modeling a part of a disk galaxy (Simpson et al., 2016), turbulent boxes (Bauer & Springel, 2012; Mocz et al., 2017), Type Ia supernovae (Pakmor et al., 2013), binary stars in a common envelope phase (Ohlmann et al., 2016), tidal disruption events (Goicovic et al., 2019), protoplanetary and accretion disks (Munoz et al., 2014; Fiacconi et al., 2018) and astrophysical jets (Bourne & Sijacki, 2017; Weinberger et al., 2017).

Time-integration is performed adopting local time-step constraints for each cell individually, solving the fluxes only across active interfaces, and calculating gravitational forces only between active particles, using an operator-splitting approach. This allows simulations with high dynamic range to be performed efficiently, focusing the numerical resolution on the most expensive regions from the computational point of view. Another important efficiency feature of AREPO code is its development as a massively distributed-memory parallel code, using the Message Passing Interface (MPI) communication standard and employing a dynamical workload and memory balancing scheme to allow optimal use of multi-node parallel computers.

In order to give the reader an overview about the AREPO code, we are now going to present in more detail the different components of the code, focusing on the various type of equations that it is able to solve (sections 2.1.1, 2.1.2 and 2.1.3), the computational mesh on which the code operates (section 2.1.4) and the method adopted for the time integration (section 2.1.5).

2.1.1 Equations solved

AREPO solves the equations of a collision-less particle component and of (magneto)hydrodynamics on a uniformly expanding, flat Friedmann-Lemaitre-Robertson-Walker space-time, which is described by the scale-factor $a(t)$ defined by the line-element (see e.g. Mo et al., 2010)

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

The evolution of the scale-factor a is given by the Friedmann equation, assuming a Λ cold dark matter cosmology:

$$H = H_0[\Omega_0 a^{-3} + (1 - \Omega_0 - \Omega_\Lambda)a^{-2} + \Omega_\Lambda]^{1/2} \quad (2.2)$$

with $H = \dot{a}a^{-1}$ and where Ω_0 is the matter density in the Universe relative to the critical density, and Ω_Λ represents the corresponding density parameter for the cosmological constant. Note that the contribution of radiation to the cosmic expansion history can be neglected for the redshifts of interest in cosmological simulation of galaxy formation.

The gravitational potential Φ in an expanding space-time can be obtained from the Poisson equation in proper space, replacing the variables with the comoving ones, and using the Friedmann equation to connect second derivatives of the scale factor to the mean density:

$$\nabla^2\Phi = 4\pi G(\rho_{total} - \rho_{mean}) \quad (2.3)$$

with G being the gravitational constant, and ρ_{total} and ρ_{mean} the total and mean density, respectively.

The collision-less component (i.e. dark matter and stars in galaxies) can be described by the Vlasov equation:

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

Applying the method of characteristics yields simple equations of motion,

$$\dot{\mathbf{x}} = \mathbf{v}/a \quad (2.5)$$

$$\dot{\mathbf{v}} = -\frac{\nabla \Phi}{a^2} - \frac{\dot{a}}{a} \mathbf{v} \quad (2.6)$$

which can be used to integrate the trajectories of discrete particles that sample the initial phase-space distribution function. Here $\dot{\mathbf{x}}$ is the velocity \mathbf{v} rescaled by the cosmic scale factor a and $\dot{\mathbf{v}}$ is the acceleration of an object subject to a gravitational potential Φ in an expanding space-time.

The equations of ideal MHD can be written as follow (Pakmor & Springel, 2013):

$$\frac{\partial \rho}{\partial t} + \frac{1}{a} \nabla \cdot (\rho \mathbf{v}) = 0 \quad (2.7)$$

$$\frac{\partial \rho \mathbf{w}}{\partial t} + \nabla \cdot \left(\mathbf{v} \mathbf{v}^T + \mathbf{I} p_{tot} - \frac{\mathbf{B} \mathbf{B}^T}{a} \right) = -\frac{\rho}{a} \nabla \Phi \quad (2.8)$$

$$\frac{\partial \mathcal{E}}{\partial t} + a \nabla \cdot \left[\mathbf{v} (E + p_{tot}) - \frac{1}{a} \mathbf{B} (\mathbf{v} \cdot \mathbf{B}) \right] = \frac{\dot{a}}{2} \mathbf{B}^2 - \rho (\mathbf{v} \cdot \nabla \Phi) + a^2 (\mathcal{H} - \Lambda) \quad (2.9)$$

$$\frac{\partial \mathbf{B}}{\partial t} + \frac{1}{a} \nabla \cdot (\mathbf{B} \mathbf{v}^T - \mathbf{v} \mathbf{B}^T) = 0 \quad (2.10)$$

where \mathbf{B} is the magnetic field vector and

$$\mathbf{w} = a \mathbf{v} \quad (2.11)$$

$$\mathcal{E} = a^2 E \quad (2.12)$$

reduces the number of MHD source terms to a single magnetic field term in the energy equation. Note that possible external heating and cooling, described through $a^2(\mathcal{H} - \Lambda)$, appear as source terms in the energy equation (2.9). Finally the total energy and pressure are calculated as follows:

$$E = \rho u_{th} + \frac{1}{2} \rho \mathbf{v}^2 + \frac{\mathbf{B}^2}{2a} \quad (2.13)$$

$$p_{tot} = (\gamma - 1) \rho u_{th} + \frac{\mathbf{B}^2}{2a} \quad (2.14)$$

In this section we have presented the most general version of the code, nevertheless no cosmological background nor magnetic field are presents in the galactic simulation that we have analyzed in this work. Therefore, the equation solved in the particular case analyzed in this work are obtained by fixing the expansion factor to $a = 1$ and the magnetic field to $\mathbf{B} = 0$.

2.1.2 Gravitational interactions

The Vlasov equation (2.4) describes the time evolution of the phase space distribution function of a collisionless fluid. When we are dealing with a finite and discrete number N of particles, interacting with each other via gravitational forces, the main effort is spent to compute the acceleration $-\nabla\Phi$ to solve the equation of motion (equations 2.5 and 2.6) for each individual particle. This leads to, in its pure form, a computational effort scaling with N^2 . To avoid this unfavorable scaling behavior with particle number, AREPO uses two different techniques: the first one is an oct-tree algorithm which groups distant particles and calculates their collective contribution to the overall force, while the second method is a grid-based approach, where the gravitational force is calculated on a Cartesian grid via Fourier methods and then interpolated to each particle position. For the purpose of this work we are going to focusing here on the first method, but we refer the interested reader to [Springel \(2010\)](#) for more details on the second method.

The Barnes-Hut oct-tree algorithm ([Barnes & Hut, 1986](#)) starts with an all-enclosing root node which is then split up recursively into eight sub-nodes of equal volume, provided there are still particles contained in these volumes. The tree-construction starts with an enclosing cubical root node in which the particles are inserted sequentially one by one. Each time a particle ends up in an already occupied leaf node, the corresponding sub-nodes are created and the particles are distributed in the sub-nodes recursively until all leaf-nodes contain at most one particle. For each tree node, multipole moments are computed recursively and then used to approximate the gravitational forces. If the multipole expansion of a visited node is considered sufficiently accurate, it is evaluated and added to the accumulated force, otherwise the node is opened and the daughter nodes are considered in turn. Following [Springel \(2005\)](#), a relative opening criterion is normally used where a node is opened if

$$\frac{Gm}{d^2} \frac{L_{node}^2}{d^2} > \alpha |\mathbf{a}| \quad (2.15)$$

where m is the mass of the node, d the distance of the particle to the node's center-of-mass,

L_{node} is the node side length, α is a tolerance parameter, and $|\mathbf{a}|$ is the absolute value of the acceleration, estimated from the previous timestep. This criterion tries to limit the absolute force error introduced in each particle-node interaction by comparing a rough estimate of the truncation error with the size of the total expected force. For the very first force calculation (where an estimated of the acceleration from the previous time-step is not available) a classic geometrical criterion

$$L_{node} > d\theta_{opening} \quad (2.16)$$

may be used instead of the relative criterion, where the opening angle $\theta_{opening}$ is a parameter effectively controlling the resulting force accuracy.

Normally, the oct-tree algorithm takes into account the gravitational interactions of all particles in the simulation volume using a Newtonian force law, implying non-periodic boundary conditions. For isolated self-gravitating systems at rest with respect to the simulation coordinate system, the opening of nodes will be similar between different timesteps. This means that the force errors made due to the approximation of the force via the tree are correlated at different times. Because all force errors in the tree algorithm do not necessarily add up to zero, this can lead to a net force on the system, and thus a non-conservation of momentum of the entire system that builds up with time. To reduce this effect, AREPO randomizes the placement of the tree domain center for each tree construction for simulations with non-periodic boundary conditions for gravity, which de-correlates the force errors in time and greatly improves the global momentum conservation (Weinberger et al., 2019).

To avoid artificial two-body interactions a gravitational force softening is introduced, which replaces the Newtonian force below a softening scale ϵ_{soft} with a reduced force that smoothly declines to zero for vanishing distance. The force becomes fully Newtonian at $2.8\epsilon_{soft}$, but strong reductions of the force due to softening only occur for $d \sim \epsilon_{soft}$ and smaller distances. This is a standard approach used to limit the maximum acceleration one particle can cause, simplify the orbit integration, and protect against the formation of bound particle pairs.

AREPO allows collisionless particles to have different softening lengths, either assigned through a concept of “particle types”, where the softening of each type can be set separately, or through properties of individual particles, such as the particle mass. For gas cells, the gravitational softening length ϵ_{cell} is always chosen in a variable way, depending on the volume V of the Voronoi cell:

$$\epsilon_{cell} = f_h \left(\frac{3V}{4\pi} \right)^{1/3} \quad (2.17)$$

where f_h is an input parameter that controls the size of the softening in relation to the cell size (usually $f_h = 2.8$). Interactions between particles with different softening lengths are symmetrized by adopting the larger of the two. If a node contains particles with different softening lengths, and the target particle's softening is smaller than the node's maximum softening length, and the distance to the node's center-of-mass is smaller than the node's maximum softening, a node is always opened, because otherwise the multipole expansion may effectively account for some interactions with the wrong symmetrized softening.

The tree-algorithm in its standard form assumes non-periodic boundary conditions (which is the case of the simulation that we are going to analyze) and extending this to periodic boundaries is computationally costly, on the other hand it can naturally and efficiently handle large dynamic ranges in spatial scales, allowing very high spatial force resolutions. In addition, it can be easily combined with local and adaptive time-stepping.

2.1.3 Magnetohydrodynamics

To solve the equations of (magneto-)hydrodynamics, AREPO uses a second-order accurate finite-volume discretization. To this end, volume-averaged primitive variables ρ , \mathbf{v} and \mathbf{B} are stored as properties of the cell at its center. Gradients are estimated with the corresponding values for neighbouring cells (Pakmor et al., 2016b), allowing a piecewise-linear reconstruction of the solution. Using the gradients, the primitive variables are extrapolated to all mesh interfaces, for which fluxes are calculated by solving a Riemann problem locally at each interface. The flux calculation can be done either with an exact, iterative Riemann solver or the approximate HLLC solver in the case of pure hydrodynamics and an HLLD solver in the case of magnetohydrodynamics (more details in Toro, 1997).

To avoid the gradient estimation become inaccurate for highly distorted cells, the gradient ∇W_i of a primitive variable W_i can be determined such that the linearly extrapolated values from cell i to the positions of the neighbouring cells j ,

$$W_j = W_i + \mathbf{d}_{ij} \nabla W_i \quad (2.18)$$

agree with the actual values W_j found there as well as possible. Here \mathbf{d}_{ij} is the position vector of cell j relative to cell i .

Therefore AREPO employs an improved weighted least-square gradient estimation by mini-

mizing the residuals of the over-determined set of equations:

$$S_{tot} = \sum_j g_j (W_j - W_i - \mathbf{d}_{ij} \cdot \nabla W_i)^2 \quad (2.19)$$

The adopted weights are given by $g_j = A_{ij}/|\mathbf{d}_{ij}|$, where A_{ij} is the area of the interface between i and j .

The property of magnetic field divergence conservation ($\nabla \cdot \mathbf{B} = 0$) is generally not true for discretizations of MHD equations (2.7 - 2.10), thus AREPO adopts the divergence-cleaning method introduced by Powell et al. (1999). This method advects numerically induced divergences away, and has the advantage of using cell-centered magnetic fields. Thus, the momentum, energy and induction equations (2.8 - 2.10) become:

$$\frac{\partial \rho \mathbf{w}}{\partial t} + \nabla \cdot \left(\rho \mathbf{u} \mathbf{u}^T + p_{tot} - \frac{\mathbf{B} \mathbf{B}^T}{a} \right) = -\frac{1}{a} (\nabla \cdot \mathbf{B}) \mathbf{B} \quad (2.20)$$

$$\frac{\partial \mathcal{E}}{\partial t} + a \nabla \cdot \left[\mathbf{u} (E + p_{tot}) - \frac{1}{a} \mathbf{B} (\mathbf{u} \cdot \mathbf{B}) \right] = \frac{\dot{a}}{2} \mathbf{B}^2 - (\nabla \cdot \mathbf{B}) (\mathbf{u} \cdot \mathbf{B}) \quad (2.21)$$

$$\frac{\partial \mathbf{B}}{\partial t} + \frac{1}{a} \nabla \cdot (\mathbf{B} \mathbf{u}^T - \mathbf{u} \mathbf{B}^T) = -\frac{1}{a} (\nabla \cdot \mathbf{B}) \mathbf{u} \quad (2.22)$$

The divergence of the magnetic field in a cell i is calculated as

$$\nabla \cdot \mathbf{B}_i = V_i^{-1} \sum_{faces} \mathbf{B}_{face} \cdot \hat{\mathbf{n}} A_{face} \quad (2.23)$$

with $\hat{\mathbf{n}}$ being the normal vector to the cell, \mathbf{B}_{face} the value of the magnetic field on the interface with area A_{face} , and V_i the volume of the i -th cell.

The advantage of this method is that it is very flexible in terms of mesh geometry, and the possibility to discretize it on local timesteps makes it computationally inexpensive. The latter is essential to reach the dynamic range required for modern (cosmological) simulations of galaxy formation.

To compute the gravitational forces of the fluid, the gas cells are included in the gravitational force calculation as described above, using a cell-centred softened forces with softening lengths tied to the cell radii. The gravitational forces are coupled as source terms to the MHD equations, in an operator split approach, which means that the source terms are applied for half a timestep at the beginning of a step, then the hyperbolic set of fluid equations are evolved for one timestep in their conservative form, followed by another gravity half-step. Errors in the total energy quickly decrease for better resolution, and the source term approach can avoid unphysical drifts in thermal energy in poorly resolved flows for conservative treatments (Weinberger et al., 2019).

2.1.4 Computational mesh

A characteristic feature of the AREPO code is its discretization on a fully adaptive, dynamic Voronoi mesh. The starting point for creating this Voronoi mesh are mesh-generating points, each one representing the position of a Voronoi cell. Each Voronoi cell is defined as the region in space that is closer to a given mesh-generating point than to any other mesh-generating point. This implies that mesh interfaces are normal to the connection of two neighboring cells, and lie at equal distance to the mesh-generating points on either side.

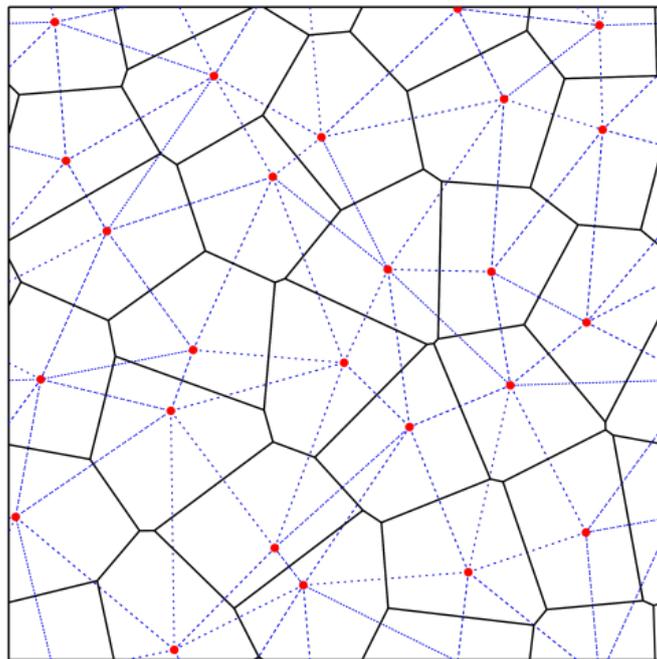


Figure 2.1: 2D representation of a Voronoi mesh (black lines) and the corresponding topological dual, the Delaunay triangulation (blue dashed lines) with mesh-generating points indicated by red dots (figure taken from [Springel 2011](#)).

The ensemble of connections to neighboring cells forms a so-called Delaunay triangulation, consisting of tetrahedra in three dimensions. The circumsphere of each tetrahedron does not contain any other mesh-generating point, and it is this mathematical property that makes the Delaunay tessellation special and unique among all other possible triangulations of space, as shown in figure 2.1.

Each mesh generating point normally moves with the bulk velocity of the fluid in its cell, taking also into account accelerations due to local pressure gradients, the Lorentz force and gravity. This naturally leads to an approximate equal-mass discretization and to quasi-Lagrangian

behavior. This, however, does not ensure by itself that the mesh remains reasonably regular at all times: due to local fluid motions, high aspect ratios of mesh cells can occur, which increase discretization errors and mesh noise, and may thus negatively impact the accuracy of the solution. To avoid this, velocity corrections can be applied for highly distorted cells to steer the mesh motion, usually these corrective velocity \mathbf{v}_{corr} are parametrized as follows:

$$\mathbf{v}_{corr} = \begin{cases} 0 & \text{for } \alpha_{max} \leq 0.75\beta \\ f_{shaping} \frac{\alpha_{max} - 0.75\beta}{0.25\beta} v_{char} \hat{\mathbf{n}} & \text{for } 0.75\beta < \alpha_{max} \leq \beta \\ f_{shaping} v_{char} \hat{\mathbf{n}} & \text{for } \alpha_{max} > \beta \end{cases} \quad (2.24)$$

where v_{char} indicates the characteristic speed in the cell, which is defined by

$$v_{char} = \frac{d}{\Delta t} \quad (2.25)$$

where d being the distance between mesh-generating point and center of mass of a cell and Δt the timestep, while $f_{shaping}$ is an input parameter that represents the cell deformability and $\hat{\mathbf{n}}$ is the normal vector pointing towards the cell center of mass. Here α_{max} denotes the maximum angle under which any of the cell faces is seen from the mesh-generating point, determinable by computing:

$$\alpha_{face} = \frac{(A_{face}/\pi)^{1/2}}{h_{face}} \rightarrow \alpha_{max} = \max(\alpha_{face}) \quad (2.26)$$

in which h_{face} is the distance from the mesh-generating point to the interface with area A_{face} . Finally β is an input parameter that represent the cell roundness criterion.

Even when using quasi-Lagrangian mesh-motion, over the course of many simulation timesteps, cells can diverge from their desired mass content or intended size. For this reason Arepo offers the possibility to refine and derefine the mesh locally, thus offering full spatial adaptivity. To refine a cell, the mesh-generating point is split into a pair of two very close points, offsetting their location in a random direction by $0.025 r_{cell}$, where

$$r_{cell} = \left(\frac{3V}{4\pi} \right)^{1/3} \quad (2.27)$$

This effectively splits the original Voronoi cell into two cells, without affecting the geometry of the neighboring cells. The conserved quantities of the split cell are then subdivided conservatively among the two new cells according to their volume ratio. For derefinement, a mesh-generating point is taken out, thereby removing the corresponding Voronoi cell from the mesh, with the neighboring cells claiming its volume. The conserved quantities associated with

the eliminated cell are conservatively distributed among these neighbors, in proportion to the volume overlap they realize with the removed cell. After a refinement event, the mesh steering motions take a couple of timesteps to re-establish a locally regular mesh geometry, at which point the next refinement may then proceed (Weinberger et al., 2019).

2.1.5 Time integration

For gravity and ideal MHD, AREPO uses explicit time-integration, which puts certain constraints on the size of individual local timesteps both with respect to the accuracy and stability of the scheme. For each cell, there are separate timestep constraints for hydrodynamics, gravitation, and potentially for source and sink terms. Generally, the most restrictive constraint is applied for all the calculations of a single cell.

For gas cells, we define a local Courant-Friedrichs-Levy (CFL) timestep criterion:

$$\Delta t \leq C_{CFL} \frac{r_{cell}}{v_{signal}} \quad (2.28)$$

with the Courant factor C_{CFL} as a free parameter, and where V is the volume of the Voronoi cell. The signal speed, in the case of a moving mesh, is calculated taking into account sound speed and Alfvén speed as follows:

$$v_{signal} = \left(\gamma \frac{p}{\rho} + \frac{B^2}{\rho} \right)^{1/2} \quad (2.29)$$

In particular cases the gas bulk velocity is added to the signal speed, which leads to a more restrictive timestep criterion.

For gravitational part, the following criterion is adopted:

$$\Delta t \geq \left(\frac{2C_{grav}\epsilon_{soft}}{|\mathbf{a}|} \right)^{1/2} \quad (2.30)$$

with C_{grav} being a free parameter. We stress again that ϵ_{soft} is the gravitational softening length introduced in section 2.1.2 and $|\mathbf{a}|$ is the absolute value of the acceleration.

The gravitational time integration is done with a second-order accurate leapfrog scheme, expressed through alternating "drift" operation, which modify the positions \mathbf{x} , and "kick" operation, which modify the velocities \mathbf{v} (the method is usually denoted with the acronym DKD):

$$\mathbf{drift} : \quad \mathbf{x}(t_{j+\frac{1}{2}}) = \mathbf{x}(t_j) + \frac{1}{2}\mathbf{v}(t_j)\Delta t \quad (2.31)$$

$$\mathbf{kick} : \quad \mathbf{v}(t_{j+1}) = \mathbf{v}(t_j) + \mathbf{g}[\mathbf{x}(t_{j+\frac{1}{2}})]\Delta t \quad (2.32)$$

$$\mathbf{drift} : \quad \mathbf{x}(t_{j+1}) = \mathbf{x}(t_{j+\frac{1}{2}}) + \frac{1}{2}\mathbf{v}(t_j)\Delta t \quad (2.33)$$

where $\mathbf{g}[\mathbf{x}(t_{j+\frac{1}{2}})]$ is the gravitational acceleration given to the particle in position $\mathbf{x}(t_{j+\frac{1}{2}})$ half timestep forward. An equivalent technique consists to invert in time the positions and velocities computations, which results in a "kick" followed by a "drift" equations (KDK), as follow

$$\mathbf{kick} : \quad \mathbf{v}(t_{j+\frac{1}{2}}) = \mathbf{v}(t_j) + \frac{1}{2}\mathbf{g}[\mathbf{x}(t_j)]\Delta t \quad (2.34)$$

$$\mathbf{drift} : \quad \mathbf{x}(t_{j+1}) = \mathbf{x}(t_j) + \mathbf{v}(t_{j+\frac{1}{2}})\Delta t \quad (2.35)$$

$$\mathbf{kick} : \quad \mathbf{v}(t_{j+1}) = \mathbf{v}(t_{j+\frac{1}{2}}) + \frac{1}{2}\mathbf{g}[\mathbf{x}(t_{j+1})]\Delta t \quad (2.36)$$

The DKD and KDK methods are equivalent, but the first one is generally used because the gravity acceleration, the most expensive part of the calculation, needs to be computed only once per timestep. Two main strengths of this method are time-reversibility and a symplectic nature which implies that it conserves the energy of dynamical systems. These features are especially useful when computing orbital dynamics.

The time integration of the hydrodynamic quantities uses Heun's method, ensuring that only a single mesh-construction is required for each timestep while obtaining formal second-order convergence in time (see for details [Pakmor et al., 2016b](#)). The update from time n to time $n+1$ of the conserved variables Q_i of the i -th cell is done conservatively through

$$Q_i^{n+1} = Q_i^n - \frac{\Delta t}{2} \left(\sum_j A_{ij}^n F(W_{ij}^n, W_{ji}^n) + A_{ij}^{n+1} F(W'_{ij}, W'_{ji}) \right) \quad (2.37)$$

where A_{ij} is the area of the interface between cells i and j , while $F(W_{ij}^n, W_{ji}^n)$ and $F(W'_{ij}, W'_{ji})$ are respectively the flux computed at the beginning of the current time step n and the flux computed on an updated mesh at the end of the timestep using time-extrapolated quantities indicated by the superscript $'$. Finally W represents the set of primitive variables (ρ, \mathbf{v}, P) , computed between adjacent cells i and j in the two cases as follow

$$W_{ij}^n = W_i^n + \mathbf{d}_{ij} \cdot \frac{\partial W_i}{\partial \mathbf{x}} \quad (2.38)$$

$$W'_{ij} = W_i^n + \mathbf{d}'_{ij} \cdot \frac{\partial W_i}{\partial \mathbf{x}} + \Delta t \frac{\partial W_i}{\partial t} \quad (2.39)$$

where \mathbf{d}_{ij} is the position vector of the geometric center of the interface between cells i and j relative to the center of cell i . The time derivatives are estimated via the spatial gradients using the continuity equations for ρ, \mathbf{v}, P ([Pakmor et al., 2016b](#)).

In section 2.1 we presented the main numerical aspects of the AREPO code, regarding the methods and the abilities in accurately solving the equations of gravity and MHD interactions,

the numerical mesh on which these equations are solved and the method adopted for compute the time integration of the characteristic physical quantities. We are going to present now in section 2.2 how to implement a physical model, the SMUGGLE model, which is able to realistically simulate the evolution of a spiral galaxy, by including all the physical processes that intervene simultaneously in determining the mode of evolution of the galactic ecosystem.

2.2 The SMUGGLE model

In addition to gravity and MHD, to correctly simulate the evolution of a galaxy it is necessary to make models of other physical processes such as radiative cooling, star formation, feedback from SNe, etc. In this section we describe a model that has been used to develop the galactic simulation analyzed in this thesis, the SMUGGLE model, which includes all the previous mentioned physical processes and much more. Adopting the methods of AREPO code described in the previous section, and alongside a realistic physical model, a Milky Way-like galactic simulation has been implemented to explore in high-resolution the evolution of an isolated late-type spiral galaxy (Marinacci et al. 2019).

The **SMUGGLE** model (**S**tars and **M**Ultiphase **G**as in **G**aLaxi**E**s) resolves the multiphase gas structure of the interstellar medium (ISM) implementing crucial aspects of stellar feedback, which include photoionization, radiation pressure, energy and momentum injection from stellar winds and from supernovae (SN). An important feature of this model is the ability to naturally reproduce the Kennicutt-Schmidt relation (Kennicutt, 1998) and self-consistently generates gaseous outflows that can eventually rain down onto the disc, forming the so-called “galactic fountain”, which sustains late-time star formation in the star-forming galaxies. A challenging task in implementing this kind of simulation models is to simultaneously modelling large volumes, while still resolving small scales in order to capture the relevant galaxy formation physics. In fact for galaxy formation, the most important physical processes beyond gravity and hydrodynamics are the formation of stars and their interactions with the surrounding ISM, which needs more numerical resolution to be simulated in detail and therefore properly modelled. For instance, the physics of the interaction between a SN explosion and its surrounding ISM occurs on sub-parsec scales, while on the other hand to capture the gravitational collapse and assembly of a halo such as that of the Milky Way, modelling scales up to ~ 1 Mpc is required. Therefore,

it is important to separately evaluate simulation limitations imposed by numerical resolution versus limitations imposed by inadequate modelling of physical processes. Resolving the complex multiphase structure of the ISM and detailed modelling of stellar feedback are therefore key targets in modern studies of galaxy formation. The model we are going to describe try to accomplish this task by implementing several physical galactic processes with an appropriate resolution.

2.2.1 Physical methods

From the observational data we know that most galaxies must have an active mechanism to prevent gas from cooling, modulating the star formation during the evolution of the galaxy. Beyond the usually adopted Galactic Nuclei (AGN) feedback, which is a fundamental feedback channel in cluster of galaxies and in massive galaxies ($M > 10^{12} M_{\odot}$), there is another leading process to explain the regulation of star formation in galaxies with masses up to those of the Milky Way ($\sim 10^{12} M_{\odot}$): the stellar feedback (SF). To correctly model this type of feedback several energy and momentum injection channels must be considered. These include contributions from SN explosions at the end of the lifetime of massive ($M > 8 M_{\odot}$) stars, but also the impact of their early ionizing radiation on the surrounding medium. Detailed modelling suggests that the feedback energy budget is dominated by radiation while the direct momentum injection budget is comparable between SN and radiation (Agertz et al. 2013). Radiation pressure and photoionization are responsible in a very deep way in the change of density and temperature of the gas. Moreover, dust in the ISM can boost the impact of radiation pressure via photon trapping where a single photon will undergo multiple scatterings (). All these processes should be included in models that aim to correctly resolve the multiphase ISM.

The ISM structure in galaxies is very complex, with hot, warm and cold phases of gas coexisting and interacting with each other. In particular, the modelling of the cold dense regions is challenging since the computational effort increases due to the shorter (hydro-dynamical) time-step. A common solution to this problem is to impose an effective equation of state for the dense gas, which means that, instead of directly resolving the individual gas phases, the ISM is treated as two-fluid gas composed of cold clouds embedded into a hot and diffuse medium and, to prevent unresolved numerical fragmentation, a polytropic relation between density and temperature $T \propto \rho^{\gamma}$ is imposed as well as a minimum gas temperature comparable to warm ionized gas, $T \sim 10^4 K$. Although this approach is numerically stable and with desirable

convergence properties, it simply lacks in resolving the ISM complexity. Because of this, efforts in building models with more resolution in the multiphase gas structure must proceed together with the development of a stellar feedback model that can regulate and prevent the runaway collapse of the cold ISM gas.

What remains poorly understood in these explicit ISM models is the efficiency with which the available energy and momentum couples to the surrounding gas. This coupling efficiency determines how important each feedback channel is in regulating star formation and launching galactic outflows. Idealized radiative transfer experiments in [Sales et al. \(2014\)](#) indicate that although radiation pressure has the ability to eventually push the gas to galactic-outflow speeds, the timescales required to accelerate the gas to these speeds are much longer than those of photoionization. As a result, stars will heat their surrounding gas via photoionization, driving gas expansion and lowering the surrounding gas density.

Current state-of-the-art zoom-in simulations have reached a resolution where it becomes desirable to remove the effective ISM approach and instead model the multiphase gas structure and related stellar feedback in more detail, even full-volume simulations have reached mass resolutions of the order of $10^5 M_{\odot}$ and therefore are able to resolve L_{\star} galaxies with millions of gas and stellar resolution elements (i.e. Illustris TNG50 [Pillepich, A., 2019](#)). The aim of the SMUGGLE model is to construct a new ISM model that captures the multiphase gas structure alongside a more explicit local stellar feedback model. Recent successful attempts in this direction have been presented (e.g. [Hopkins et al., 2011](#); [Agertz et al., 2011, 2013](#); [Hopkins et al., 2014a](#)), demonstrating the possibility to achieve a more detailed treatment of the gas and stars in galaxies, within the cosmological context and with high numerical resolution, on scales of MW galaxies and below.

The goal of the SMUGGLE model is to move the scales of numerical closure to smaller scales within the ISM. In contrast, effective models do not resolve structure within the ISM even if the numerical resolution is increased. The numerical closure scale is therefore essentially set by the ISM and the SMUGGLE model can yield converged results at the resolution of current state-of-the-art cosmological simulation.

To correctly implement the SMUGGLE model, and other similar models, there is a need to complement fundamental modules for gravity and hydrodynamics in the AREPO code with some methods to describe the main physical processes that influence the phase structure of the ISM and thus the star formation. These methods include several mechanisms such as: gas heat-

ing and cooling, a stochastic implementation for the formation of stellar particles, stellar feedback processes from three main channels (SNe, radiation and stellar winds). In the following subsections the physical processes considered in this model and their numerical implementation will be described in more detail.

2.2.2 Cooling and heating

The pristine mix of hydrogen and helium that compose ISM gas is modified by a cooling and heating network that changes the thermal state of the gas through two-body processes such as: collisional excitation, collisional ionization, recombination, dielectric recombination and free-free emission, Compton cooling off CMB photons, and photoionization from a spatial-uniform UV background. Furthermore if the temperatures are high ($T \gtrsim 10^4$ K) the cooling from metal line is accomplished with a self-consistently update of the gas metallicity field. The net metal cooling rates (which are tabulated as a function of temperature, gas density and redshift based on CLOUDY calculation) are added to this network, scaled by the gas total metallicity relative to solar (more details in [Vogelsberger et al., 2013](#)).

An important feature of the SMUGGLE model is the creation of gas at low temperature ($T \lesssim 10^4$ K). In particular the model is able to cool the gas up to ~ 10 K via low-temperature metal lines, fine-structure and molecular cooling processes, under the assumption of chemical and ionization equilibrium. This is particularly important in order to generate molecular clouds of gas with high densities and low temperature which make possible the star formation.

Moreover, to ensure an effective accumulation of dense gas and reduce the dissociating effect of ionizing radiation, it is essential to include a process of self-shielding of the gas at high densities ($n \gtrsim 10^{-3} \text{ cm}^{-3}$). In order to model the self-shielding mechanism, the molecular cooling function Λ_{mol} (taken from the cooling tables presented in [Hopkins et al., 2018c](#)) is multiplied by a factor $1 - f_{ssh}(z)$ (which depends from the redshift z) computed by the parametrization presented in [Rahmati et al. \(2013\)](#). In this way the self-shielding factor f_{ssh} is used to suppress the photoionization and photoheating rates in the range of high density of molecular hydrogen gas, modulating the strength of the dissociation mechanisms with the redshift. The limit set in considering the self-shielding correction is $z = 6$, after which no correction is applied. More details about self-shielding can be found in [Vogelsberger et al. \(2013\)](#).

Cosmic ray and photo-electric heating are also taken into account. These processes are

important for both cold ($T \sim 50K$) and warm ($T \sim 8000K$) phase of the ISM, entering in the thermal balance and stability mechanisms (see e.g. [Field et al., 1969](#); [Wolfire et al., 1995](#)). Cosmic ray heating is implemented using the form given by [Guo & Oh \(2008\)](#) where the heating function has the form

$$\Lambda_{CR} = -10^{-16}(0.98 + 1.65)n_e e_{CR} n_H^{-1} \text{ erg s}^{-1} \text{ cm}^{-3} \quad (2.40)$$

where n_e is the electron density, n_H is the hydrogen number density and the e_{CR} is the cosmic ray energy density. The cosmic ray energy density is parametrized as

$$e_{CR} = \begin{cases} 9 \times 10^{-12} & \text{erg cm}^{-3} \text{ for } n_H > 0.01 \text{ cm}^{-3} \\ 9 \times 10^{-12} \left(\frac{n_H}{0.01 \text{ cm}^{-3}} \right) & \text{erg cm}^{-3} \text{ for } n_H \leq 0.01 \text{ cm}^{-3} \end{cases} \quad (2.41)$$

in such a way that it becomes progressively less important for low density gas (see also [Hopkins et al., 2018c](#)).

The emission of electrons from dust grains, in particular polycyclic aromatic hydrocarbons (PAHs), caused by the photoelectric effect due to the interstellar radiation field, becomes important for the photoelectric heating of the gas. The implementation of the photoelectric heating function Λ_{phot} has been made in the form given by [Wolfire et al. \(2003\)](#) (equation 19 in their paper) where

$$\Lambda_{phot} = -1.3 \times 10^{-24} e_{\nu}^{pe} n_H^{-1} n_e \left(\frac{Z}{Z_{\odot}} \right) \times \left(\frac{0.049}{1 + (x_{pe}/1925)^{0.73}} + \frac{0.037(T/10^4 \text{ K})^{0.7}}{1 + (x_{pe}/5000)} \right) \text{ erg s}^{-1} \text{ cm}^3$$

$$x_{pe} \equiv \frac{e_{\nu}^{pe} T^{0.5}}{\Phi_{PAH} n_e n_H}$$

in which e_{ν}^{pe} is the photon energy density and Φ_{PAH} is a factor which takes into account the uncertainties in the interaction rates between molecular atoms and dust grains, and a value of $\Phi_{PAH} = 0.5$ comes from observations of carbon density ratios in diffuse clouds ([Jenkins & Tripp, 2001](#)).

2.2.3 Star Formation implementation

To implement the star formation a probabilistic approach is adopted. Star particles are created stochastically according to the probability derived from the star formation rate \dot{M}_{\star} of each gas cell computed as follows

$$\dot{M}_\star = \begin{cases} 0 & \text{for } \rho < \rho_{th} \\ \epsilon \frac{M_{gas}}{t_{dyn}} & \text{for } \rho \geq \rho_{th} \end{cases} \quad (2.42)$$

where ρ_{th} is a density threshold parameter, ϵ is an efficiency factor that has been set equal to 0.01 (a value obtained from observations, e.g. [Krumholz & Tan, 2007](#)) and t_{dyn} is the gravitational dynamical time of any given gas cell, defined as

$$t_{dyn} = \left(\frac{3\pi}{32G\rho_{gas}} \right)^{1/2} \quad (2.43)$$

with ρ_{gas} and M_{gas} the density and mass of the gas within the cell, respectively. To be eligible for star formation a gas cell has to fulfill several criteria. A first criterion for a gas cell to be eligible for star formation is to have a density above a specific density threshold ρ_{th} , the latter is set to the value of 100 cm^{-3} according to the range of average density in the giant molecular clouds. ([Ferriere, 2001](#)). An additional criterion is added to restrict star formation only to gravitationally bound regions. Defining the virial parameter α_i for every i -th cell, which is computed as follow

$$\alpha_i = \frac{\|\nabla \times \mathbf{v}_i\|^2 + (c_{s,i}/\Delta x_i)^2}{8\pi G\rho_i} \quad (2.44)$$

where v_i , $c_{s,i}$, ρ_i are the gas velocity, sound speed, gas density respectively within the i -th cell and Δx_i is taken as the radius of the sphere having the same volume as the i -th cell. The numerator represents the sum of the turbulent support, which is defined by

$$\|\nabla \times \mathbf{v}_i\|^2 \equiv \sum_{i,j} \left(\frac{\partial v_i}{\partial x_j} \right)^2 \quad (2.45)$$

computed over all the x_j spatial dimensions, and the thermal support $(c_{s,i}/\Delta x_i)^2$ within the resolution scale Δx_i . If $\alpha_i < 1$ implies that the gas is not able to overcome gravitational collapse via gas motion and thermal support so the cell i becomes eligible for star formation. This criterion ensures the identification of gas which is collapsing under self-gravity at the resolution scale (i.e. the gas that is able to form stars; [Hopkins et al., 2018c](#)).

Once the star formation rate \dot{M}_\star is determined according to equation (2.42), the stellar mass that would be formed in the current time-step Δt by the cell i is given by

$$M_{\star,i} = M_i \left[1 - \exp \left(-\frac{\dot{M}_\star \Delta t}{M_i} \right) \right] \quad (2.46)$$

where M_i is the mass of gas in the cell. To convert a gas cell into a star particle equation (2.46) is sampled stochastically by adopting the factor

$$p \equiv 1 - \exp \left(-\frac{\dot{M}_\star \Delta t}{M_i} \right) \quad (2.47)$$

as the probability of a given mass M_i has to form stars consistently with the local star formation rate (SFR) \dot{M}_* ; in other words p is the fraction of the cell mass that will be converted to stars. Hence, to decide whether the conversion occurs, a uniformly distributed random variable p^* in the interval $[0, 1]$ is extracted and compared with p , and if $p^* < p$ the cell is converted to a star particle.

2.2.4 Feedback from SNe

SN feedback is fundamental in regulating star formation and the resulting ISM structure. The generation galactic-scale outflows (Li et al., 2017) and gas turbulence (Martizzi et al., 2016) are two main effects that come from the momentum injection by single SN and furthermore multiple consecutive SNe events.

The primary difficulty associated to the modelling SN feedback is to properly capture the early energy-conserving expansion phase of a SN remnant, the so-called Sedov-Taylor phase, in which the momentum imparted to the ISM gas is generated by an overpressurized central gas bubble that expands, sweeping up and accelerating the surrounding material. The major difficulty in properly capture this evolutionary phase come from the fact that, usually, galaxy formation simulations do not have the resolution required to model this phase. If this issue is not addressed, the momentum imparted to the surrounding ISM can be largely underestimated. Therefore, an effective SN implementation that captures in the proper way the total momentum and energy injection is important because it is this momentum injection that is largely responsible for the regulation of SF.

The SN feedback implementation in the SMUGGLE model is based on few assumptions. First, the total energy injected in a single SN event is given by

$$E_{SN} = f_{SN} E_{51} \quad (2.48)$$

where $E_{51} = 10^{51}$ ergs and f_{SN} is a model parameter that encodes the SN feedback efficiency. Second, the blast wave velocity at explosion is

$$v_{SN} = \left(\frac{2E_{SN}}{M_{SN}} \right) \quad (2.49)$$

such that the momentum carried by the blast wave at explosion is given by

$$p_{SN} = M_{SN} v_{SN} = \sqrt{2E_{SN} M_{SN}} \quad (2.50)$$

where M_{SN} is the ejecta mass per SN.

Due to the possibility that, in the SMUGGLE model, multiple SN events may occur in a time-step Δt within a single stellar particle (that represent a stellar population rather than a single star, because of the resolution of the simulation), the computation of the proper amount of energy and momentum injected begins from the number of SN events and the total associated ejecta masses. Therefore, for each stellar particle and for each Δt the total energy for the total number of SNe is defined as

$$E_{SN,tot} = f_{SN} E_{51} (N_{SNII} + N_{SNIa}) \quad (2.51)$$

which includes N_{SNII} and N_{SNIa} that are the number of type II and type Ia SN, respectively. On the other hand, to the ejected masses of both supernova channels ($M_{SNII,tot}$ and $M_{SNIa,tot}$) is associated a total momentum given by

$$\begin{aligned} p_{SN,tot} &= p_{SNII,tot} + p_{SNIa,tot} \\ &= \sqrt{2E_{SNII}E_{SN}M_{SNII,tot}} + \sqrt{2E_{SNIa}E_{SN}M_{SNIa,tot}}. \end{aligned} \quad (2.52)$$

The number of type II SNe (N_{SNII}) is computed by assuming that the stellar population represented by a single star particle follows a [Chabrier \(2001\)](#) initial mass function $\Phi(m)$, and by integrating it as follow

$$N_{SNII} = M_{\star} \int_{M(t+\Delta t)}^{M(t)} \Phi(m) dm. \quad (2.53)$$

In the previous equation M_{\star} is the mass of the star particle at birth, and $M(t)$ is the mass of a star that leaves the main sequence at an age t (which values are tabulated [Vogelsberger et al., 2013](#) and references therein). In the same way it is possible to define the associated ejecta mass as

$$M_{SNII,tot} = M_{\star} \int_{M(t+\Delta t)}^{M(t)} M f_{rec}(m, Z) \Phi(m) dm \quad (2.54)$$

in which $f_{rec}(m, Z)$ is the amount of mass given back to the ISM by stellar evolution, which depends on stellar mass and metallicity ([Portinari et al., 1998](#)). The minimum main sequence mass for a type II SNe explosion is set to $8M_{\odot}$ and an IMF upper limit of $100M_{\odot}$ is adopted.

The calculation of SNIa events, instead, has been implemented by the parametrization of the events temporal distribution using a delay time distribution (DTD) and deriving the number of type Ia SNe events as

$$N_{SNIa} = \int_t^{t+\Delta t} DTD(t') dt' \quad (2.55)$$

A functional form of $DTD(t)$ that agrees with theoretical models is given by

$$DTD(t) = \Theta(t - \tau_8) N_0 \left(\frac{t}{\tau_8} \right)^{-s} \frac{s-1}{\tau_8} \quad (2.56)$$

where $\tau_8 = 40$ Myr approximates the main sequence life time of an $8 M_\odot$ star, $N_0 = 2.6 \times 10^{-3}$ $\text{SN } M_\odot^{-1}$, $s = 1.12$ and Θ is the Heaviside function that parameterizes the delay between the birth of the stellar population and the first SNIa event (Maoz et al., 2012). This form for the DTD comes from the link between SNIa rates and the orbital energy and angular momentum loss rate due to gravitational wave emission (Greggio, 2005). Finally, each SNIa releases the same amount of ejecta $M_{\text{SNIa}} \sim 1.37 M_\odot \text{SN}^{-1}$, and therefore the total mass return is computed as

$$M_{\text{SNIa,tot}} = M_{\text{SNIa}} N_{\text{SNIa}} \quad (2.57)$$

To capture the discrete nature of SNe explosions, a time-step constraint is imposed in order to ensure that the expectation value for the number of SN events per time-step is of the order of unity. For each stellar particle the timestep constrain Δt_\star is computed as

$$\Delta t_\star = \min(\Delta t_{\text{grav}}, \Delta t_{\text{evol}}) \quad (2.58)$$

where Δt_{evol} represents the timestep constraint given by stellar evolution and is determined by taking the minimum

$$\Delta t_{\text{evol}} = \min\left(\Delta t_{\text{SNII}}, \frac{t_{\text{age}}}{300}\right) \text{ yr} \quad (2.59)$$

with t_{age} the age of the stellar particle and

$$\Delta t_{\text{SNII}} = \frac{\tau_8}{N_{\text{SNII}}} \quad (2.60)$$

is the average time between two SN explosion events, computed as the ratio between the main sequence life time of an $8 M_\odot$ star, previously defined as $\tau_8 = 40$ Myr (see equation 2.56) and the expected number of SN events N_{SNII} ($\sim 10^{-2} \text{SN } M_\odot^{-1}$) for the stellar particle over that time. The latter value comes from the assumed IMF function and a minimum mass for a type II SN progenitor of $8 M_\odot$. Integrating the equations (2.53) and (2.55) in the range of Δt_\star gives the expected number of SNe over the time-step. This value is taken as the expectation value (λ) of a Poisson distribution, which is sampled to obtain the actual number of discrete SN events per time step, given the probability of n events as

$$p(n; \lambda) = \frac{\lambda^n e^{-\lambda}}{n!} \quad (2.61)$$

It is immediately obvious that the two supernova types have two different distributions as a function of the stellar age. In particular, type II supernovae display a very steep rise in the first 40 Myr after which they reach a constant trend (as shown in figure 2.2) that reflects the particular choice of $8 M_\odot$ as the lower mass of type II SN progenitor and the relative main sequence

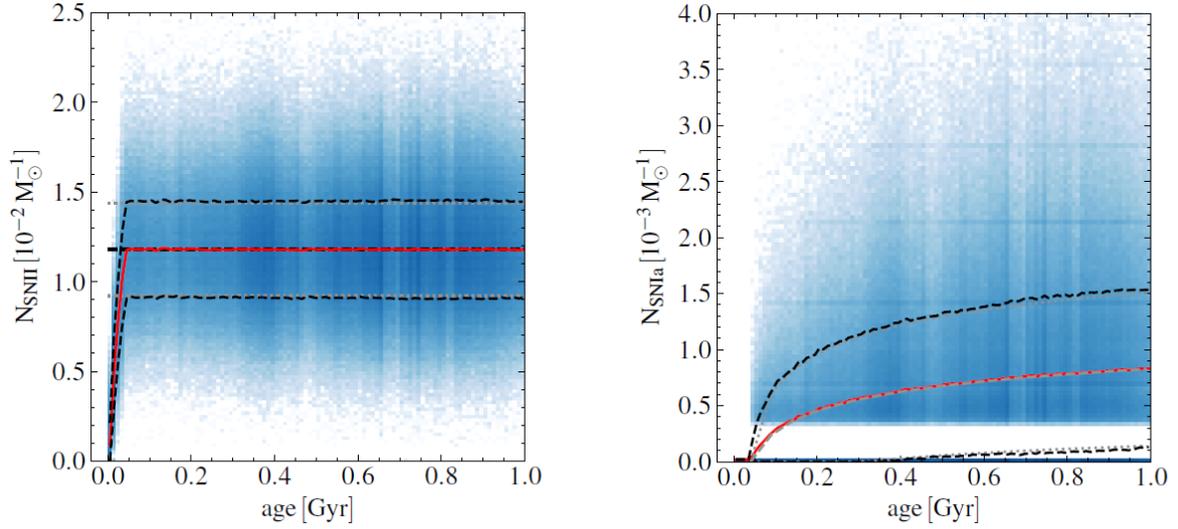


Figure 2.2: Cumulative number of supernova type II events (left) and type Ia (right) per unit formed stellar mass as a function of stellar age. The red solid line represents the average supernova events as a function of age, with the corresponding standard deviation around the mean value drawn as dashed black lines. On the background of each plot is present a two-dimensional histogram (blue shades) counting the number of stellar particles falling into each 2D-bin having a width of 10 Myr in age and 2×10^{-4} and 4×10^{-5} supernova events per unit mass for supernova type II and type Ia, respectively (figure taken from [Marinacci et al. 2019](#)).

life time of ~ 40 Myr. Basically all the SN events per stellar particle occur within this age. Type Ia supernovae instead have a different trend: events start only after 40 Myr – which is the minimum age for a star in a binary system to leave the main sequence and being of sufficiently low mass to become a white dwarf – and then follow the trend described in equation (2.55). The evaluation of these behaviours is important when we will perform, during the analysis discussed in section 4.2, a stellar particle selection based on their ages in order to identify the particles that belong to stellar cluster responsible for the recent (in term of the particular time explored of the simulation analyzed) creation of a super-bubble.

In order to properly capture the momentum injection into the gas (which is fundamental in the regulating star-formation feedback) there is the need to explicitly account for momentum that is generated by the PdV work done by the hot post-shocked gas during the adiabatic Sedov-Taylor expansion phase of the SN blast. In SMUGGLE model this is implemented by boosting

the momentum imparted to each gas cell i influenced by SN feedback as

$$\Delta p_i = \tilde{w}_i \min \left[p_{SN,tot} \left(1 + \frac{m_i}{\Delta m_i} \right)^{1/2}, p_t \right] \quad (2.62)$$

where \tilde{w}_i is a weight function partitioning the energy and momentum injection among gas cells, Δm_i is the total SN mass from type II and type Ia weighted by \tilde{w}_i , the mass of the gas cell is m_i and p_t is the so-called terminal momentum, i.e. the final value of the momentum of the SN blast at about the cooling radius, when the evolution of the blast wave transitions from the Sedov-Taylor phase to a momentum-conserving phase (see [Cioffi et al., 1988](#), equation 4.7).

The SN energy and momentum injection are spread over a number of nearest gas particles using weight functions. For each star particle a predefined effective number of neighbours

$$N_{ngb} = \frac{4\pi}{3} h^3 \sum_i W(|\mathbf{r}_i - \mathbf{r}_s|, h) \quad (2.63)$$

is set to $N_{ngb} = 64 \pm 1$ and the coupling radius h is obtained iteratively among the the position vectors \mathbf{r}_i of the i -th gas neighbour respect the position vector \mathbf{r}_s of the stellar particle. W is the standard cubic spline SPH kernel ([Monaghan & Lattanzio, 1985](#)). In addition it is also imposed a feedback limiter, namely, a maximum radius for the coupling of energy and momentum from the stars, R_{SB} . In practice the effective scale for coupling is defined as

$$h_{coupling} = \min(h, R_{SB}) \quad (2.64)$$

Once $h_{coupling}$ has been determined, weights are defined in such a way that each gas cell within $h_{coupling}$ receives SN energy and momenta proportionally to the fraction of the 4π solid angle that it covers as seen from the star position, namely

$$w_i \equiv \frac{\Delta \Omega_i}{4\pi} \quad (2.65)$$

and to ensure that the correct amounts of energy and momentum are imparted each cell receives a fraction

$$\tilde{w}_i \equiv \frac{w_i}{\sum_i w_i} \quad (2.66)$$

of the amount of $p_{SN,tot}$ and $E_{SN,tot}$ determined in equations (2.52) and (2.51). Momenta are directed radially away from the star position. The scale R_{SB} is actually the size of the super-bubble estimated from equation (1.4), which depends from the total number of SNe events occurred in a single stellar particle (that namely represents an association of massive OB stars) and from the life time of this stellar population (~ 40 Myr for a lower SN progenitor mass of

8 M_{\odot}). The ambient average gas density n_h also influences the size of the SB expansion, in fact when the pressure associated with the SN ejecta becomes comparable to the ambient ISM pressure $P_{env} \propto n_H T_{env}$, the impact of the SN blast wave will have on the surrounding ISM gas becomes negligible. In general, typical values of R_{SB} are several hundreds of parsecs, depending on the numerical resolution, for instance assuming $n_H = 1 \text{ cm}^{-3}$, the corresponding super-bubble radius is in the range $\sim 0.3 - 1 \text{ kpc}$ for the numerical resolution adopted in the simulation that we are going to analyze in this work (see [Marinacci et al., 2019](#)). In the original paper describing the SMUGGLE model implementation it is also noted that the numerical convergence of the model improves when R_{SB} is kept constant instead of adjusting its value on the fly. Therefore a value $R_{SB} = 0.86 \text{ kpc}$ is set as a default value. However, tests have shown that varying this parameter in the above range did not have a significant impact on the final results.

2.2.5 Radiative feedback from young massive stars

Radiation, especially from young and massive stars, has an important impact on both the thermal and the dynamical state of the ISM gas. There are two main processes involved: photoionization and radiation pressure momentum imparted to the ISM. The first can alter the ionization state and therefore the gas temperature of the ISM, moreover each absorbed photon imparts additional momentum to the gas via radiation pressure. The latter can be particularly relevant at high densities, such those found in giant molecular clouds, the birth clouds of stars. This process is particularly important because of its timing, in fact it can be responsible for the dispersal of such clouds before any SN goes off ([Murray et al., 2010](#); [Lopez et al., 2011](#); [Walch et al., 2012](#)), thereby rendering SN feedback more effective due to the reduced gas densities in which SNe subsequently explode.

The most relevant observational feature produced by the ISM gas photoionization from young massive stars is the formation of ionized hydrogen HII regions. The formation of such regions is captured in the SMUGGLE model by defining the ionizing photon rate as follows

$$N_{\star} = \frac{L_{\star}}{\langle h\nu \rangle} = \frac{\gamma_{\star} M_{\star}}{\langle h\nu \rangle} \quad (2.67)$$

that is the ionizing luminosity of the star divided by the mean photon energy emitted above the hydrogen ionization level of 13.6 eV. Assuming an averaged mass-to-light ratio $\gamma_{\star} = 10^3 L_{\odot}/M_{\odot}$ it is possible to estimate the luminosity L_{\star} of the stellar particle starting from its mass M_{\star} ; in addition it is reasonable to choose $\langle h\nu \rangle = 17\text{eV}$ that is the energy corresponding to the

peak emission of a black-body Planck spectrum with temperature $T \sim 40,000$ K (Rybicki & Lightman, 1986), consistent with massive OB stars.

Because the mass of the gas within the Strömngren radius¹ is usually much smaller than the gas mass contained in within $h_{coupling}$ (the scale at which the model couples the feedback energy), the photoionization process acting on each gas cell nearby a ionizing stellar particle is implemented in a probabilistic way in the SMUGGLE model. In particular at every cell is assigned a probability of being photoionized as

$$p_{phot} = \frac{n_{\star}}{\alpha_{rec} n_H^2 V} \quad (2.68)$$

where $\alpha_{rec} \sim 2.6 \times 10^{-13} \text{ cm}^3 \text{ s}^{-1}$ is the hydrogen recombination rate, and

$$n_H = \frac{X\rho}{m_p}, \quad n_{\star} = \tilde{w}_i N_{\star} \quad (2.69)$$

are the average hydrogen number density of the cell and the rate of ionizing photons injected into the cell, respectively (X is the hydrogen mass fraction and V is the volume of the cell). In this way the probability is given by comparing the number of recombinations expected in the cell to the total photon number emitted by the source scaled by the same solid angle weighting scheme adopted to model the supernova feedback. Here again a random number p' is then selected from a uniform distribution in the range $[0,1]$ and cells where $p' < p_{phot}$ are tagged for photoionization. which effect is to impose a temperature floor $T_{phot} = 1.7 \times 10^4$ K (consistent with a gaseous medium made of hydrogen and photoionized by $\langle h\nu \rangle = 17\text{eV}$ photons) and to disable their radiative cooling for a duration t_{off} equal to the star particle time-step.

The heating of eligible gas cells due to photoionization will generate some over-pressurization of the gas compared to their colder surroundings, leading to the expansion of the photoionized HII regions around young stellar systems, thus imparting momentum to the gas (which is a different physical mechanism than a direct transfer of momentum by radiation pressure described below).

Radiation pressure, which becomes important in dense and optically-thick regions, represents a source of momentum that has an additional impact on the dynamical state of the gas. The momentum injection rate from radiation pressure can be written as (see Agertz et al., 2013 equation 5)

$$\dot{p}_{rad} = (1 + \tau_{IR}) \frac{L_{\star}}{c} \quad (2.70)$$

¹The Strömngren radius define the ionized region around a very energetic ionizing star (usually of class O and B), and is analytically obtained by balancing the recombination and ionization rates of the gas surrounding the star.

where τ_{IR} is the infrared optical depth and c is the speed of light. Thus the total momentum injected into the ISM at every time-step simply becomes

$$p_{rad}^{tot} = \dot{p}_{rad} \Delta t . \quad (2.71)$$

A simple approach to estimate this quantity is to assume a constant infrared optical depth $\tau_{IR} = \kappa_{IR} \Sigma_{gas}$ by setting the opacity in infrared band $\kappa_{IR} = 10(Z/Z_{\odot}) \text{ cm}^2 \text{ g}^{-1}$ from observational data (Hopkins et al., 2018c) and computing the gas column density Σ_{gas} for each star particle j using the Sobolev approximation as

$$\Sigma_{gas,j} = \langle \rho_s \rangle_j \ell_j \quad (2.72)$$

with the average density $\langle \rho_s \rangle_j$ given by (see equation 2.63)

$$\langle \rho_s \rangle_j = \sum_i W(|\mathbf{r}_i - \mathbf{r}_j|, h) m_i \quad (2.73)$$

and the Sobolev length as

$$\ell_j = h_{coupling,j} + \frac{\rho_j}{\|\nabla \rho_j\|} \quad (2.74)$$

in which ρ_j and $\nabla \rho_j$ are respectively the gas density and the associated gradient determined by adopting a standard SPH approach (Hopkins et al., 2018c). The total radiation momentum determined in equation (2.71) is coupled to the gas contained within $h_{coupling}$ in the same way as it is done for SN feedback (see equations 2.65-2.66) and directed radially away from the star position.

2.2.6 Feedback from OB and AGB stellar winds

Stars also contribute to feedback through stellar winds, in particular the short-lived OB stars association and the asymptotic giant branch (AGB) stars are the most important class in this type of stellar feedback. In contrast to the young OB stars class, which pre-process the gas before SNe explosions take place, the AGB winds feedback continue to act at later times than radiation and is associated to older stellar populations. A detailed accounting of the mechanical power carried by winds of massive stars shows that their momentum injection rate per unit stellar mass formed is comparable to that of the SNe (Agertz et al., 2013) but it is released before the first SN explosion ($t \lesssim 5 \text{ Myr}$) and rapidly declines thereafter. This behavior has the potential of enhancing the effect of subsequent SN explosions because of the reduction in the ambient density.

The implementation of these two feedback contributions in the SMUGGLE model follows two steps:

1. The mass lost by a stellar particle is computed over a time step Δt as

$$M_{loss} = M_{\star}[m_{c,loss}(t + \Delta t) - m_{c,loss}(t)] \quad (2.75)$$

where M_{\star} is the initial mass of the stellar particle and $m_{c,loss}(t)$ is the cumulative mass loss per unit stellar mass as a function of time t (see [Hopkins et al., 2018c](#)).

2. Energy and momentum injection from stellar winds are then computed by setting

$$E_{winds} = L_{kin}\Delta t \quad (2.76)$$

as the energy of the winds, where $L_{kin} = \frac{1}{2}M_{loss}v_{winds}^2$ is the kinetic energy per unit of time associated to M_{loss} determined above and the wind velocity v_{winds} . Thus the total momentum injection rate is determined as

$$p_{winds} = \sqrt{2M_{loss}E_{winds}} \quad (2.77)$$

The same approach adopted in SNe coupling radii is applied here with the same maximum values. Finally winds feedback is only performed if the returned mass over a given time-step is larger than 10^{-4} times the mass M_{\star} of the stellar particle at birth, otherwise mass loss is accumulated until this threshold is reached.

So far we have presented the ways in which the SMUGGLE model reproduces the fundamental physical processes that take place within a galaxy and that determine its evolution by influencing the structure of the ISM and regulating star formation. In the next section we will describe the numerical setup of the isolated Milky Way-like galaxy simulation that we have analyzed in this thesis.

2.2.7 Numerical setup of the analyzed simulation

The implementation of all the previously described physical processes in the SMUGGLE model has been tested in a set of simulations of an isolated late-type galaxy, similar to our Milky Way ([Marinacci et al., 2019](#)). This was done in order to explore the ability of the model to regulate the star formation and to investigate the impact of various stellar feedback processes on the ISM structure. It is on the intermediate resolution run of that set of simulations that the analysis

conducted in this work will be performed. In what follows we just briefly recap the main properties of the simulation that will be analyzed in the next chapter of this thesis.

The galaxy model has been set up with a dark matter halo, a bulge and a stellar and gaseous discs components with different profiles to ensure the hydrostatic equilibrium at the beginning of the simulation. The dark matter halo and the bulge are modelled with a Hernquist profile (Hernquist, 1990) while the stellar and gaseous discs are exponential in the radial direction, with the first one following a sech^2 distribution in the vertical direction and the second one computed self-consistently by putting the gas, at the initial temperature of 10^4 K, in hydrostatic equilibrium with the gravitational potential (see for more details Springel et al., 2003).

The dark matter mass distribution is modelled by a density profile that follow the law

$$\rho_{halo}(r) = \frac{M_{halo}}{2\pi} \frac{a}{r(r+a)^3} \quad (2.78)$$

with the scale parameter a define as

$$a = r_s \sqrt{[2\ln(1+c) - c/(1+c)]} \quad (2.79)$$

where $c \equiv \frac{r_{200}}{r_s}$ is the concentration index defined by the scale-length r_s of the halo and its radius r_{200} corresponding to the distance from the center at which the average density is 200 times the critical density. The bulge is taken with spherical geometry and thus modelled with a Hernquist profile

$$\rho_b(r) = \frac{M_b}{2\pi} \frac{b}{r(r+b)^3} \quad (2.80)$$

with b as bulge scale parameter. The disk components of gas and stars are modelled with an exponential surface density profile of scale-lengths r_g and r_d respectively, following

$$\Sigma_{gas}(r) = \frac{M_{gas}}{2\pi h^2} \exp(-r/r_g) \quad (2.81)$$

and

$$\Sigma_{\star}(r) = \frac{M_{\star}}{2\pi h^2} \exp(-r/r_d) \quad (2.82)$$

The vertical mass distribution of the stars in the disc is specified by giving it the profile of an isothermal sheet with radially constant scale height h , thus obtaining a 3-dimensional stellar density given by

$$\rho_{\star}(R, z) = \frac{M_{\star}}{4\pi h r_d^2} \text{sech}^2\left(\frac{z}{2h}\right) \exp\left(-\frac{R}{r_d}\right) \quad (2.83)$$

As mentioned above, the vertical structure of the gas is obtained by numerically solving the hydrostatic equilibrium equation with the gravitational potential generated by the stellar and the dark matter mass components in an iterative way.

M_{halo}	c	M_b	b	M_*	r_d	h	M_{gas}	r_g
[M_\odot]		[M_\odot]	[kpc]	[M_\odot]	[kpc]	[pc]	[M_\odot]	[kpc]
1.53×10^{12}	12	1.5×10^{10}	1.0	4.73×10^{10}	3.0	300	9×10^9	6.0

Table 2.1: Structural parameters for the Milky Way-like galaxy model of the simulation. From the left to the right: mass of the dark matter halo M_{halo} , concentration index c , mass of the bulge M_b , bulge scale length b , mass of the stellar disc M_* , stellar disk radial scale length r_d , stellar disk scale height h , mass of the gaseous disk M_{gas} , gas disk radial scale length r_g .

Once the parameters of the various distributions of galactic components have been set, in order to build the various structures that make up the galaxy, different sample techniques are applied. The stellar disk and the bulge are sampled with non-collisional particles. The gas particles are sampled in a similar way, but each particle is used as a mesh-generating point for the Voronoi mesh of AREPO. The dark matter halo, instead, is considered only in terms of an external (static) potential. The same parameters values presented in [Hopkins et al. \(2012\)](#) were used to construct the Milky Way-like galaxy model. Regarding the structural parameters, the model has a total mass of $1.6 \times 10^{12} M_\odot$ subdivided between: M_{halo} the mass of the dark matter halo, M_b the bulge mass, M_* the stellar disc mass and M_{gas} the gaseous disc mass. These parameters values and the remaining structural parameters, which correspond to the various scale lengths (b, r_d, h, r_g) and the concentration c for the dark matter component, are summarized in [Table 2.1](#). With the chosen values the resulting disc gas fraction is the order of $f_{gas} \simeq 10$ within a $R = 8.5$ kpc. The gas in the disc has a metallicity equal to the solar value of 0.0127 ([Asplund et al., 2009](#)).

Finally the system is enclosed in a cubic volume with side of 857 kpc, with a background grid consisting of low-density gas cells to avoid vacuum boundary conditions. The resulting initial conditions are then evolved for ~ 1 Gyr, saving the outputs of the evolution as snapshot of the system every time interval of $\Delta t \sim 7$ Myr. From each snapshot is possible to retrieve information about all the different components of the system, such as gas density values for every cell of the hydrodynamic mesh and position of stellar particle as well as other important characteristics (stellar particle mass, SN events per particle, age, etc.).

Chapter 3

The Astrodendro Python package

In developing the structures identification regarding the SB candidates and the stellar clusters that generate these sub-dense regions (henceforth called “clusters generators”), a software package, implemented in Python language, called Astrodendro has been used. The Astrodendro package (<http://www.dendrograms.org/>) has been adopted to compute dendrograms of the astronomical data retrieved by the SMUGGLE simulation, and to perform the SB identification by analyzing the sub-dense regions associated to the stellar clusters formed through the evolution of the simulation. The algorithm method to compute the dendrograms is described in section 3.2 where a complete description of the computational parameters is illustrated, while section 3.1 presents how the dendrogram tree is structured and how it is possible to explore it in an iterative way, thus retrieving the structures geometrical information and their associated statistical properties as well (see section 3.3).

3.1 Dendrogram structure

A dendrogram is a diagram that shows the hierarchical relationship between objects. The main use of a dendrogram is to work out the best way to allocate objects to clusters, thus it is most commonly created as an output from hierarchical clustering.

A dendrogram is composed of two types of structures: *branches*, which are structures that can be split into multiple sub-structures, and *leaves*, which are structures that have no sub-structure. Branches can split up into branches and leaves, which allows hierarchical structures to be adequately represented. The term *trunk* is used to refer to a structure that has no parent structure, as illustrated in figure (3.1). In order to explore the dendrogram, an identification

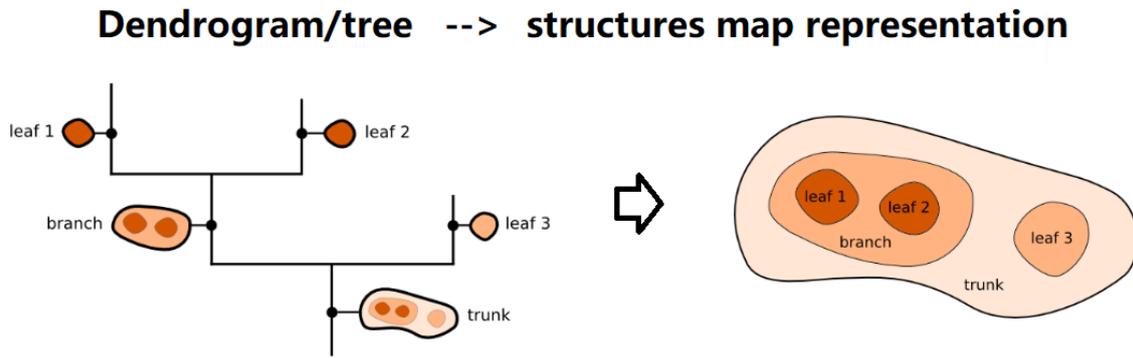


Figure 3.1: The hierarchical relationship (dendrogram/tree) between clusters on the left and the corresponding map representation on the right (figure taken from <https://dendrograms.readthedocs.io>).

number (ID) is assigned to each structure. The first place to start is the *trunk* of the tree, which is a list of all the structures at the lowest level, thus it is a collection of items each of which could be a *leaf* or a *branch*. Branches have a *children* attribute that returns a list of all sub-structures, which can include branches and leaves, while the latter can be accessed by the homonym attribute applied directly to the main dendrogram instance. To retrieve information about the leaves (which will be the main analyzed sub-structures of the tree in finding the sub-dense regions of interest, see chapter 4) it is simply possible to iterate over the list returned by `"dendrogram.leaves"`, where each element of the list is a *leaf* instance with two attributes:

- *leaf.indices* : return an array of pixels indices along the map that belong to the *leaf* structure under examination.
- *leaf.values* : return an array of pixels values (in our case gas column density) which make up the *leaf* structure under examination.

The first attribute is important to evaluate the area of the leaf and the associated border (both used to calculate the equivalent radius and the related error, see section 4.2.1), while the second attribute has been used to generate the density profile of the structure by retrieving pixels values within each areola of the domain subdivision (see section 4.3.1 for a detailed explanation).

3.2 Computing a dendrogram

In elaborating the 2D density maps to construct the hierarchical tree, the way the Astrodendro algorithm works is starting from the brightest pixels in the dataset, and progressively adding

fainter and fainter pixels, evaluating at each step how to concatenate the regions under examination in an hierarchical structure (i.e. the dendrogram/tree). The three main parameters that are used to compute the dendrogram are the following:

- Σ^{low} : the minimum value to consider in the dataset, any value lower than this will not be considered in the dendrogram.
- Δ^{min} : how significant a *leaf* has to be in order to be considered as an independent entity. The significance is measured from the difference between its peak flux and the value at which it is being merged into the tree.
- N_{px}^{min} : the minimum number of pixels/values needed for a *leaf* to be considered as an independent entity.

In order to provide a clear description of the core algorithm that generates the dendrogram, we consider here an example of a one-dimensional dataset. Even though this dataset is very simple, what is described applies to datasets with any number of dimensions. As it is illustrated in figure (3.2a) the starting point of the algorithm is highlighted by the dashed blue line (namely "current value") drawn on the figure. On this image the dataset values are represented by the black solid line which makes it evident the oscillating nature of the density values along the spatial dimension, showing peaks and gorges of density. If the data are noisy, in order to

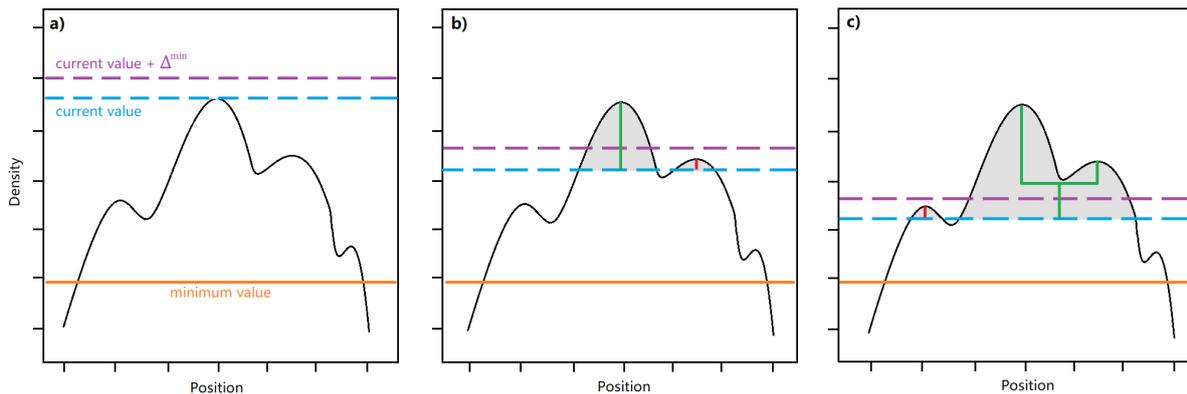


Figure 3.2: Schematic representation of the process of implemented in Asterendro to build the hierarchy tree. The solid black line represents a one-dimensional density profile as a function of the position. A minimum density value is set by the horizontal orange line which determines the limit value Σ^{low} in computing the dendrogram, while the dashed blue and purple lines represent the current value of the computation and the minimum threshold Δ^{min} to chain a leaf to a branch of the tree, respectively.

avoid including local maxima that are only identified because of noise, it is possible to define a minimum height required for a structure to be retained (i.e. the significance of the structure Δ^{min}), represented by the purple dashed line which indicates the sum of the current value and the set minimum height. Starting from the current value the algorithm moves to the pixel with the next largest value, and each time, it decide whether to join the pixel to an existing structure, or create a new structure. A new structure is only started if the value of the pixel is greater than its immediate neighbors, and therefore is a local maximum. Every times a new structure has to be created it needs also to satisfy the minimum significance criteria Δ^{min} as shown in figure (3.2b) where next to the first main identified structure (green line on the figure) a new structure is being created but still did not reach the minimum level of significance (red line on the figure). As the current value moves further down the red structure gets to reach the value of Δ^{min} that has been set, becoming part of the hierarchical tree. Then both structures keep growing until a pixel that is not a local maximum is reached, and is adjacent to both existing structures. At this point the structures under consideration are merged into a branch, which is shown by a green horizontal line on figure (3.2c). The described procedure is repeated until the minimum value of density Σ^{low} is reached, incorporating in this way all the structures that are found along the path. All the branches found up to the minimum value eventually merge in a single tree.

3.3 Statistical properties

For 2D position-position (PP) datasets, the *Astro dendro.analysis* module can be used to compute basic properties for each dendrogram structure. By using the *PPStatistic* class contained in the *Astro dendro.analysis* module is possible to analyze the pixels values that belong to each structure under examination, which can be identified by the ID number as mentioned in section 3.1. Specifically, the first three moments of the associated statistical distribution were used for the purpose of our analysis. Considering a 2-dimensional case, basically adding one spatial dimension to the example illustrated in figure 3.2, and denoting as d_k the density value of the k -th pixel of the structure under consideration, it is possible to define the moment μ_0 of order zero as simply the sum of all the n pixels values belonging to that structure, i.e.

$$\mu_0 = \sum_{k=1}^{k=n} d_k, \quad (3.1)$$

the moment μ_1 of order one as the intensity-weighted mean position

$$\mu_1^i = \sum_{k=1}^{k=n} \frac{x_k^i d_k}{\mu_0} \quad (3.2)$$

where x_k^i is the position of the k -th pixel along i -th dimension, thus μ_1^i denotes the intensity-weighted mean position along the i -th dimension, and finally the moment μ_2 of order two as the intensity-weighted covariance matrix

$$\mu_2^{i,j} = \frac{1}{\mu_0} \sum_{k=1}^{k=n} d_k (x_k^i - \mu_1^i)(x_k^j - \mu_1^j). \quad (3.3)$$

where $\mu_2^{i,j}$ denotes the element of the symmetric covariance matrix in position (i, j) . These moments of pixels distribution are necessary to determine various geometrical properties of the structures that we are going to analyze in chapter 4. In particular, the application of the *PPStatistic* class to a structure results in an object that has methods to compute various statistics, such as:

- *major_sigma* : Major axis of the projection onto the position-position (PP) plane, computed from the intensity-weighted second moment $\mu_2^{i,i}$ in direction i of greatest elongation in the PP plane.
- *minor_sigma* : Minor axis of the projection onto the position-position (PP) plane, computed from the intensity-weighted second moment $\mu_2^{j,j}$ perpendicular to the major axis (which means along j direction) in the PP plane.
- *x_cen* : The mean position of the structure in the i direction as moment of order one μ_1^i
- *y_cen* : The mean position of the structure in the j direction as moment of order one μ_1^j
- *to_mpl_ellipse* : Returns a Matplotlib ellipse representing the first and second moments of the structure.

The major and minor axes have been used to calculate the "elliptical radius" of a structure, which represents the radius of a circle with the same area of the representative ellipse of the structure (see for details section 4.2.1), while the center coordinates (x_{cen}, y_{cen}) of a structure have been adopted as reference points to compute the "equivalent radius", which is defined as the radius of the circle with the same total pixels area composing the structure (see for details section 4.2.1). Finally using the *to_mpl_ellipse* method it is possible to visualize directly on

the density map the ellipse associated to a particular structure, used to visually evaluating the roundness of each leaf.

The description of how the Astrodendro package works provided in the chapter 3 is, albeit short, fundamental to understand how the algorithm is able to identify complex structures on a density map, creating a hierarchical tree based on three fundamental input parameters. We are going now to present in the next chapter (4) how the computational dendrogram algorithm has been used in identify the gaseous structure of interest working on the column density map of the total gas of the galaxy in analysis, and how the Astrondendro statistical properties have been adopted in order to retrieve important results from the identified gas structure.

Chapter 4

Results

4.1 Structures identification with Astrodendro

Here follows a complete description regarding the methods and strategies adopted to elaborate the stellar and gas images in order to correctly identify super-bubble candidates. To accomplish this task, stellar particles and gas density data from relevant snapshots of the SMUGGLE simulation have been elaborated applying image processing methods in order to produce density maps which have been used to retrieve information about the position of stellar clusters and their influence on the surrounding gas density.

The region of interest for the elaborated maps is limited by a cube of side length of 40 kpc centered on the potential minimum of the galaxy and comprising the entire gas disk, from which projected maps are recovered by elaborating gas density data from the Voronoi mesh. The chosen side length of the cube allows us to explore the density distributions of both components (stars and gas) up to the galaxy's boundaries, where the star formation is essentially no longer active due to the reduced gas density.

On the one side two different gas projections are generated for every snapshot: a face-on and an edge-on view of the total gas column density along the galaxy. On the other side stellar particles are used to generate a surface mass density map, which has been then superimposed on the gas density map to evaluate the relation between denser stellar regions and sub-dense gas regions. We continued later using the Astrodendro algorithm alongside a probability measure on the correlation existing between stellar and gaseous structures, by designing and implementing an iterative method to explore the space of the dendrogram's computation parameters in such a way as to establish the parameters values which maximize the probability of obtaining strong

evidence of sub-dense gas structures developed by the clusters stellar activity within them (see section 4.1.5 for a description of this procedure). Finally, since the generated probability distribution does not follow any analytical trend attributable to a known statistical distribution, a Metropolis-Hasting Markov Chain Monte Carlo (MH-MCMC) algorithm has been adopted to realize the search procedure of the maximum probability (more details in section 4.1.6), finding in this way the best set of astrodendro parameters that generate the best SB candidates identification on the gas density map.

Before proceeding with the analysis performed on the SB candidates (4.2), we present in this section the methodology adopted in generating the stellar density map (section 4.1.1), the gas column density map (4.1.2) and how these two maps have been compared (4.1.3), concluding this section with the dendrogram parameters investigation (4.1.4) and the statistical approach adopted in determining the SB candidates identification method (sections 4.1.5-4.1.6).

4.1.1 Stellar surface density map

In developing the super-bubbles research method, the first step was achieved by building a representative map of the stellar distribution within the galaxy. Starting from the Cartesian coordinates on the galactic plane of all the stellar particles and after dividing the projected 2D spatial domain into a 512^2 pixels image, positional data and mass of each stellar particle were recovered for a specific snapshot, in order to build a stellar density map. In the beginning the procedure used was that of adding the masses of all the star particles that fall within a certain pixel and thus dividing the total stellar mass for the pixel area (opportunedly converted in physical units). Later on, since the supernovae activity is the main reference factor for establishing how much a stellar particle influences the surrounding gas, each mass contribution was weighted by the number of SN events of the corresponding stellar particle, thus obtaining an average total mass weighted by the total number of SN events that occurred within a certain pixel. Denoting as m_i the mass of the i -th stellar particle that falls within the j -th pixel with area A_j , it is possible to determine the surface stellar mass density Σ_j^{star} weighted by the number of SN events w_i for every stellar particle within the pixel, as follows

$$\Sigma_j^{star} = \frac{\sum_i m_i w_i}{A_j \cdot \sum_i w_i}, \quad (4.1)$$

where every weight w_i includes the contribution of both type of SN events occurred from the birth of the i -th stellar particle:

$$w_i = (N_{SNII} + N_{SNIa})_i . \quad (4.2)$$

Since the total number of pixels per image side is fixed in both dimensions the area A_j is equal for every pixel, and corresponds to

$$A_j \equiv A_{px} = \left(\frac{40 \text{ kpc}}{512 \text{ px}} \right)^2 \approx 6 \cdot 10^3 \text{ pc}^2 \quad (4.3)$$

Using equation (4.1) and iterating the procedure for all the image pixels leads to a well-defined SN-weighted stellar surface density map that includes also the information regarding the SN events, thus emphasizing the denser and more active supernovae stellar regions.

For a given snapshot an age limit has been set in order to retrieve information only from young stellar particles (i.e. $5 < t_{age} < 50$ Myr), thus limiting the temporal information in the stellar density map to a period of time in which SN explosions have had the opportunity to transfer energy to the gas, imparting momentum and thus excavating the sub-dense regions that we are looking for, and ensuring that most of the SN events per particle occurred (as discussed in section 2.2.4 on the time range of supernova events per stellar particle, also illustrated in figure 2.2). A reference snapshot at ~ 700 Myr (hereinafter referred to as "reference snapshot") after the begin of the simulation has been used to evaluate the overall SN activity during an evolutionary era of the galaxy in which the star formation rate has reached an almost stationary average value around $\sim 2 M_{\odot} \text{ yr}^{-1}$ (as shown in figure 4.1). The stellar density map created for the reference snapshot is reported in figure 4.2 where a Gaussian smooth with a radius of 3 pixels has also been applied in order to enhance the visibility of stellar structures. This smoothing operation is important to obtain a better identification of the stellar clusters through the Astrodendro algorithm, which operates on an iterative comparison of the density variation between two adjacent pixels. In fact, by smoothing the density values between adjacent pixels it is possible to better delimit the stellar clusters areas, incorporating unpaired pixels within the belonging clusters, as shown in figure 4.3.

4.1.2 Gas density map

The procedure used to generate the gas density map from the outputs data of the reference snapshot is described in this sub-section. Following the methodology adopted for the stellar

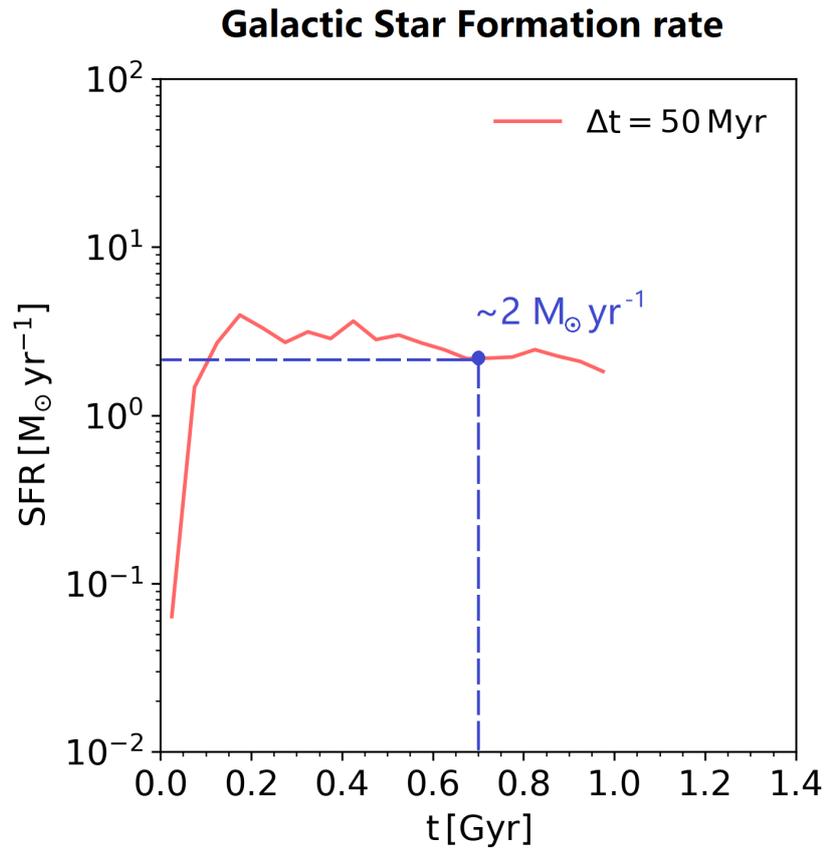


Figure 4.1: Star formation rate SFR (solid red line) as a function of the simulation time. The SFR curve has been computed using a sampling bin size of $\Delta t = 50 \text{ Myr}$. The blue dot on the curve indicates the position of the reference snapshot at $\sim 700 \text{ Myr}$ and its SFR average value of $\sim 2 M_{\odot} \text{ yr}^{-1}$ respectively indicated by the vertical and horizontal dashed lines.

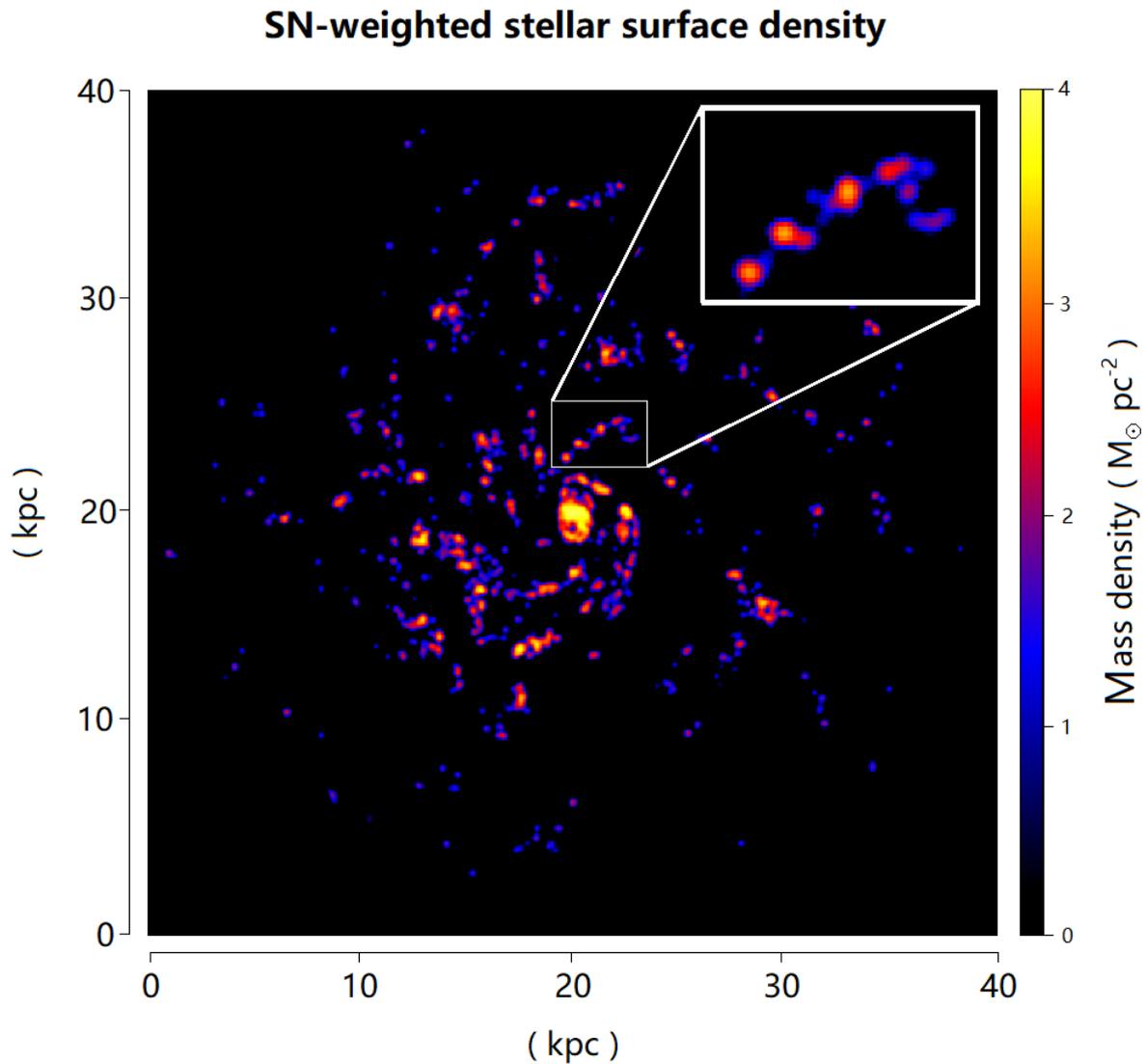


Figure 4.2: Stellar mass surface density map weighted by the number of SN events within every pixel (reference snapshot n° 100, $t = 698.9$ Myr). A Gaussian smooth with radius of 3 pixels has been applied to the original map in order to enhance stellar clusters, as it is remarked by the zoom inset on the upper right corner where four different clusters are clearly distinguishable by their peak of density and quite symmetric round shape.

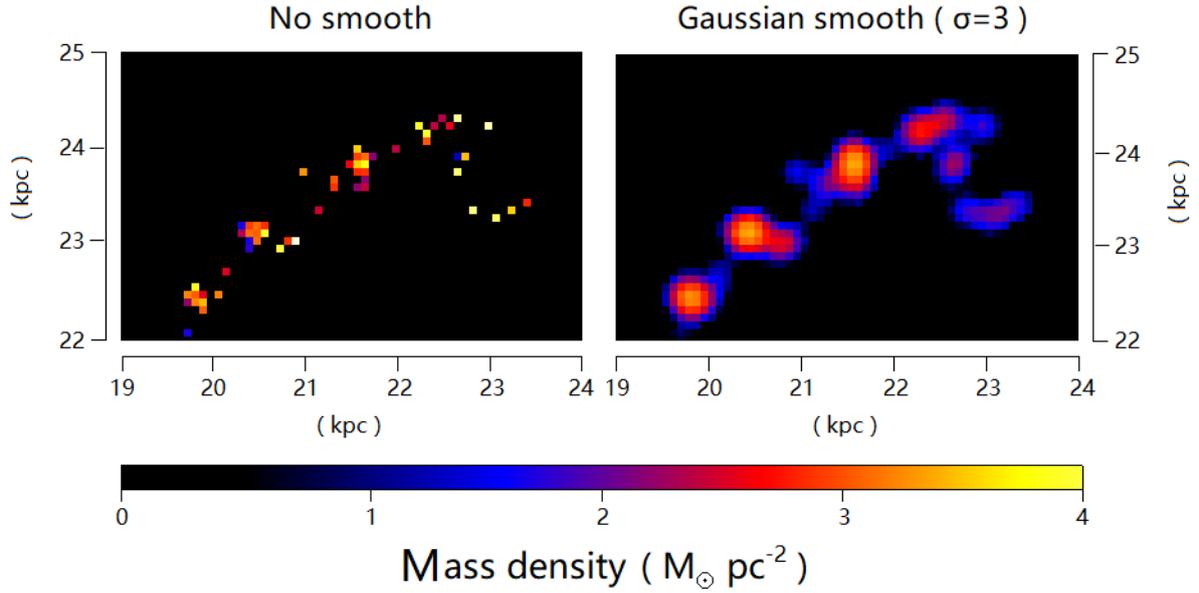


Figure 4.3: Comparison between maps of the stellar mass surface density weighted by the number of SN events, with (right) and without (left) a Gaussian smoothing (with a radius of 3 pixels) to incorporate unpaired pixels. The zone represented in the comparison corresponds to the zoomed inset on the upper right corner of figure 4.2.

density map, the gas density map has been computed by retrieving the density values of every Voronoi cell within the spatial cube domain, and projecting onto a 512^2 pixels 2D map by adding the density contribution of each cell to the nearest pixel along the line of sight. In a reference system in which the galactic plane is parallel to the Cartesian plane $x - y$, choosing a line of sight parallel to the \hat{z} coordinate of the cubic domain leads to generate a face-on gas density projection, while choosing the \hat{x} or \hat{y} line of sight leads to an edge-on galaxy projection. In particular, the first projection is used to investigate the relation between stellar clusters and sub-dense gas regions, while the second one brings information about the scale height profile of the galactic disk.

The explicit method to generate the a density map is to associate the i -th gas cell to the j -th pixel if it contains the center point of that cell, after properly projecting all the cell's center positions on the 2D plane of reference. Denoting the cell volume V_i containing a total gas mass m_i , the column density value associated to the pixel is calculated as follow

$$\Sigma_j^{gas} = dx_j \cdot \sum_i^N \frac{m_i}{V_i} \quad (4.4)$$

with N being the total number of projected gas cells that fall within the j -th pixel and dx_j is the

length of the pixel's edge. In this way, every pixel contains the total gas column density in the chosen line of sight. Once again, using a constant map resolution of 512^2 pixels, the length of each pixel results of the order of

$$dx_j \equiv \Delta x = \sqrt{A} \approx 78 \text{ pc} \quad (4.5)$$

All the performed analysis about the SB identification is focused on the face-on projection map, while the edge-on projection can be used in the analysis of the galactic gas disk scale height evolution (which is left to future investigation).

Once the gas density map has been created it is necessary to reverse all the pixel's density values in order to transform the sub-dense regions into high density regions. This procedure is fundamental to correctly adopt the Astrodendro algorithm, which iteratively separates density structures starting from the highest pixel value and creating a hierarchic structure's tree going towards the lowest density value (as described in section 3.2 and illustrated in figure 3.2). In this way, by inverting the map density values with an suitable method, allows to isolate the sub-dense regions of interest by delimiting fictitious high density regions on the reversed map, just right at the beginning of the hierarchical tree, resulting in a faster and more accurate searching method.

The method adopted to reverse the density map is based upon the maximum Σ_{max} and minimum Σ_{min} pixel density values in the original density projected map of reference (the suffix "gas" has been removed to lighten the notation). Every new pixel values Σ_j^{new} belonging to the reversed map can be computed by taking the sum of these extreme values and associate to the j -th pixel a density of

$$\Sigma_j^{new} = \Sigma_{max} + \Sigma_{min} - \Sigma_j^{old} \quad (4.6)$$

where Σ_j^{old} denotes the density value of the j -th pixel in the original map. Adopting this technique results in preserving the density scale values, which is important to attribute significant physical units to the Astrodendro parameters values.

Here again a smoothing method has been applied to the reversed density map in order to perform a better structures identification. The logic followed in applying an appropriate smoothing is based on the nominal resolution that radio telescopes have in observing the trace in neutral hydrogen from the 21-centimeter emission lines. Since a typical size of the radio-telescope beam full width half maximum (FWHM) at this wavelength is approximately 15 arcsec in interferometric 21-cm observations (Oosterloo et al., 2007), the smoothing radius is given from the

conversion of this FWHM into a projected length (in kpc) after placing the galaxy at a fictitious distance of 5 Mpc and taking the standard deviation σ of the Gaussian profile associated to the beam using the statistic relation

$$\text{FWHM} = \sigma \sqrt{8 \ln 2} \sim 2.35\sigma \quad (4.7)$$

If we denote the angular beam aperture as $\delta = 15''$, the conversion into a radio resolution scale length L_{FWHM} , corresponding to the projected FWHM on the sky plane, is given by

$$L_{\text{FWHM}} = 2D \cdot \tan\left(\frac{\delta}{2}\right) \approx 0.5 \text{ kpc} \quad (4.8)$$

where $D = 5 \text{ Mpc}$ is the distance of the galaxy from the observer. Thus dividing by 2.35 and converting into pixels number results in a smoothing radius equal to

$$\sigma = \frac{L_{\text{FWHM}}}{2.35} \cdot \frac{512 \text{ px}}{40 \text{ kpc}} \approx 3 \text{ px} \quad (4.9)$$

comparable to the smoothing radius adopted for the stellar density map. However, this smoothing procedure is based on the radio band observations of hydrogen, but we are considering in our map the column density of the total gas, thus to be consistent we should convert the values of the map into the density of only hydrogen (rather than the total gas) by applying appropriate conversion factors. This conversion procedure has not been investigated, however we expect only slight changes in the conformation of the gas structures observing only hydrogen, but the conversion should still be analyzed in future works.

The produced gas density map with inverted values and smoothed following the previous described method is shown in figure (4.4), where the same reference snapshot adopted for the stellar map has been used.

4.1.3 Maps comparison

The comparison between the stellar and the gas map provides a first check, albeit qualitatively, on the possible correlation between the stellar clusters and the sub-dense gas structures, namely the SB candidates. It is somewhat plausible that not all cavities dug in the gas are attributable to the recent activity of the young stellar clusters, but some of them are the sign of past SN activities whose generating clusters have now vanished leaving behind only some areas of rarefied gas. By superimposing the stellar map onto the gas density map, it is possible to visually identify which areas of rarefied gas are more attributable to recent SN explosions (i.e. occurred

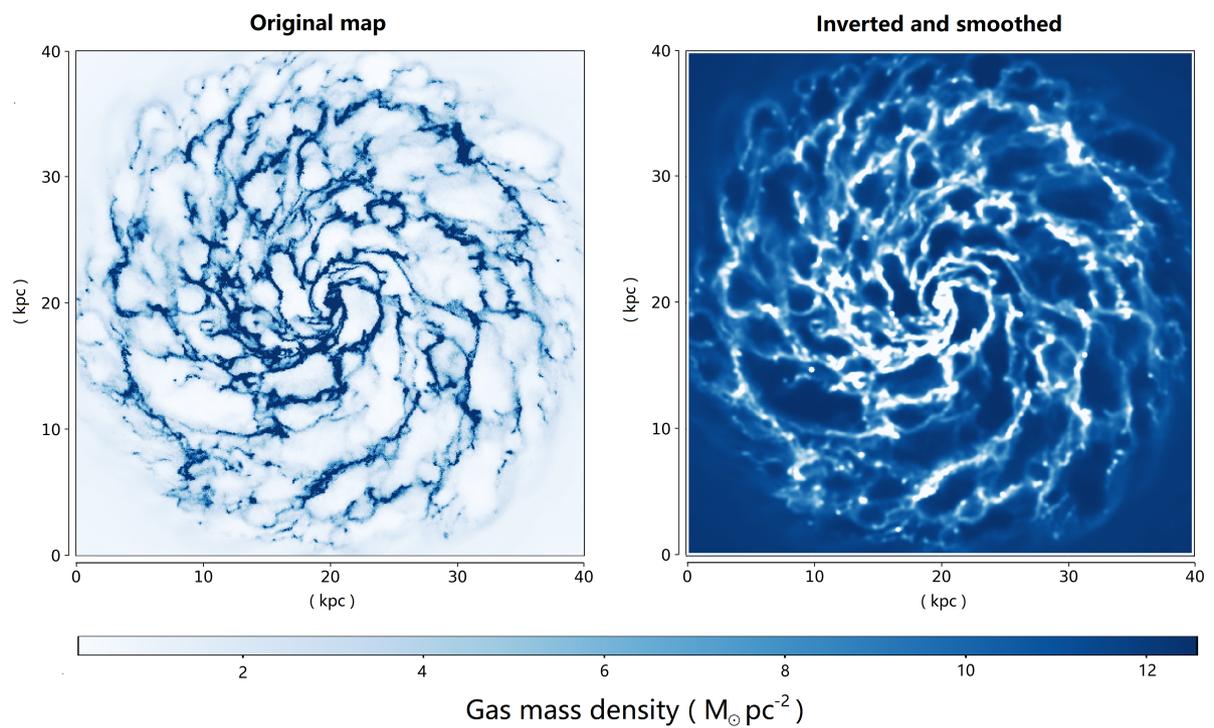


Figure 4.4: Original (left panel) and inverted values (right panel) gas surface density maps in a face-on projection. A Gaussian smoothing with radius $\sigma = 3$ pixels has also been applied to the inverted values map. The bluer colors indicate higher surface density, which correspond to star formations zones on the original map and to cavities excavated by SN activity on the inverted map.

within 50 Myr) and which are the remnants of past activity. However, a more quantitative comparison is necessary to establish an effective correlation between SB candidates and the clusters¹ generator at their centers.

The qualitative comparison between the gas map and the stellar map is proposed in figure (4.5) where it is immediately possible to note that some stellar clusters, especially those with a peaked density, are located more or less in the center of cavities in the gas presumably created by them during their initial evolution, when the event rate of SNs was high. On the other hand, many other sub-dense structures seem to have no young stellar clusters within them. Presumably these structures were created in the past and subsequently left to evolve according to the dynamics of the gas rotating around the galactic center, in fact they are more oblong and deformed, a sign of the torsion imparted by the rotating action of the moving gas.

It is therefore necessary to establish a selection criterion among the sub-dense gas regions to determine which structures are actually a product of many recent sequential SN explosions, to correctly identify the regions that can be labeled as "SB candidates". In fact by choosing the right SB candidates we can obtain a significant statistic on the properties of these cavities, and in the same way we could follow their evolution individually. All these investigations are necessary to check that the simulation model produces results in agreement with theory and observations, in particular verifying whether the super-bubbles size distribution reflects that observed in our and other similar spiral galaxies or adjustments to the model might be needed.

The first approach followed in analyzing the density maps, was to investigate how the number and the shape of the identified structures vary by changing the Astrodendro computational parameters: the lower limit in density value Σ^{low} , the minimum density variation between pixels values Δ^{min} and the minimum number of pixel N_{px}^{min} to accept a significant structure. By generating simultaneously a dendrogram for the gas and the stellar map, it is clear that when the computational parameters vary the algorithm is able to identify different structures in the both maps, regardless of the relationship that the gas structures have with the stellar structures. An example of how it is possible to identify many structures beyond their significance is presented in the figure (4.6) for the gas density map, where an overestimation of the number of the identified structures is clearly visible; this leads to consider also structures spontaneously formed by the turbulent action of the gas (in the inner regions) and from the drift of cavities formed

¹We indicate with the name of "cluster generator" a stellar cluster that generates the surrounding sub-dense gas region by sweeping the gas through its supernovae stellar activity.

Star-gas comparison map

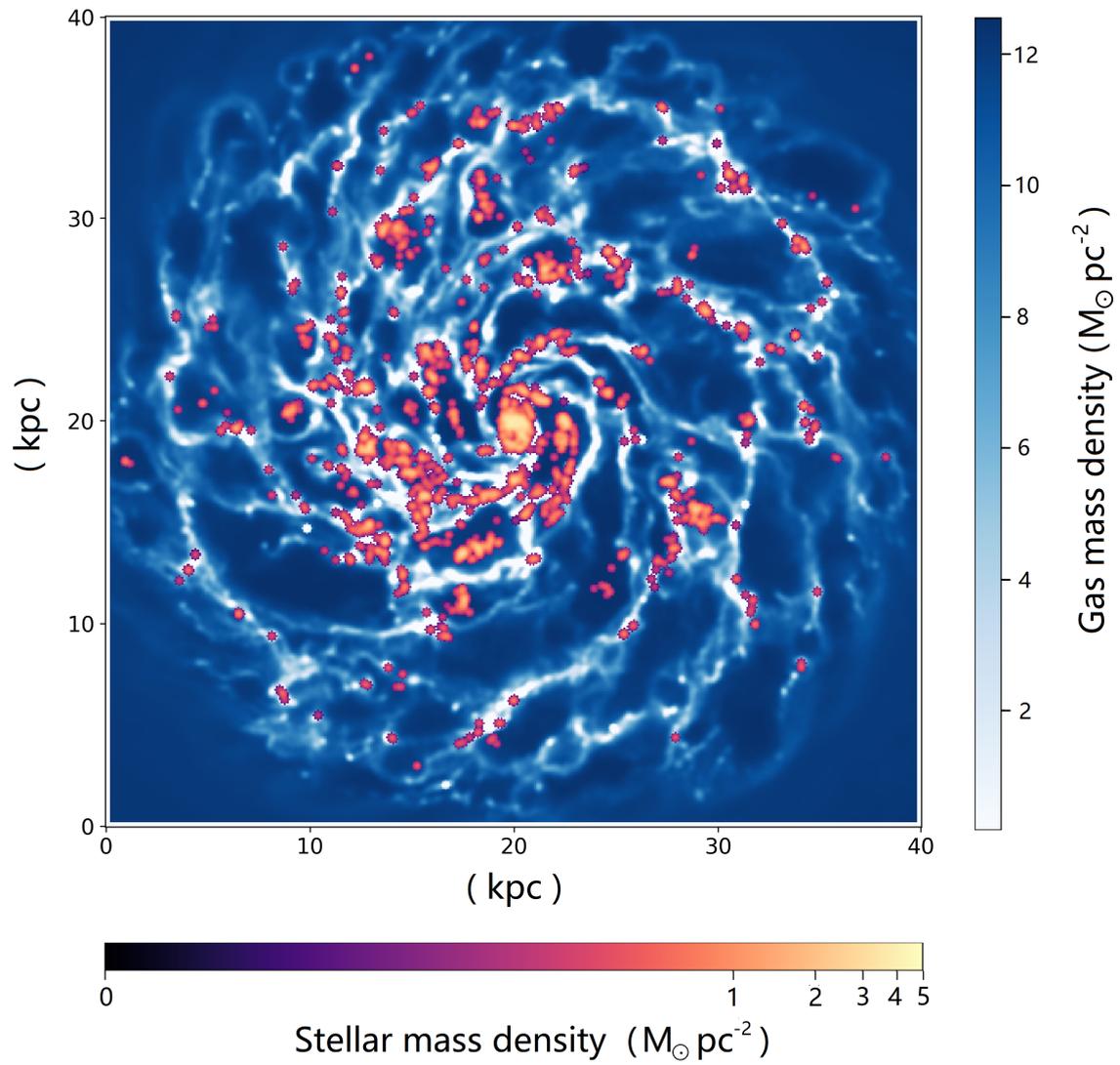


Figure 4.5: Inverted gas column density map (white-blu scale) with superimposed SN-weighted stellar density map (purple-pink scale).

in the past and transported to the edges of the galaxy, even in areas where stellar clusters are not formed. The dendrogram computation result shown in figure (4.6) has been generated from the reference snapshot by setting a reasonable set of values for the parameters, estimated by qualitative considerations. For instance setting the lower limits in density values and number of pixels to zero ($\Sigma_{gas}^{low} = 0 \text{ M}_{\odot} \text{ pc}^{-2}$, $N_{px}^{min} = 0$) leads to consider the whole range of identifications based on these two parameters (which means computing the dendrogram on the entire density scale and keeping the structures with any number of pixels), while limiting the significance of the structures only to the typical minimum column density fluctuation observable in HI using radio-interferometric observations of 21-cm ISM emission (i.e. mean square fluctuation of $\Delta_{gas} \approx 0.5 \text{ M}_{\odot} \text{ pc}^{-2}$, Dutta & Bharadwaj, 2013).

The result presented in figure (4.6) shows how it is necessary to devise a practical method, also adaptable to other snapshots, which is able to establish the structures to discard and maintain based on their relationship with the young formed stellar structures.

4.1.4 Parameters investigation

The Astrondendro algorithm can be applied as well to the stellar density map, where once again the variation of the choice of computational parameters influences the identification of stellar structures, this time. An example of two different dendrograms computed on the stellar map associated to the reference snapshot is presented in figure (4.7) where the choice of the lower limiting parameters follows the consideration adopted with the gas map (i.e. $\Sigma_{\star}^{low} = 0 \text{ M}_{\odot} \text{ pc}^{-2}$, $N_{px}^{min} = 0$), and the minimum stellar density variation parameter has been arbitrarily set to the middle range value of the whole map density scale $\Delta_{\star} \approx 2 \text{ M}_{\odot} \text{ pc}^{-2}$ and compared to a result obtained setting the minimum value as possible $\Delta_{\star} = 0 \text{ M}_{\odot} \text{ pc}^{-2}$. The comparison is made to emphasize how delicate is the choice of the parameter related to the density variation among different pixels in identifying stellar clusters, in fact as can be seen from the figure (4.7), the criterion adopted discriminating between two neighboring stellar clusters is very sensitive to parameter Δ_{\star} which thus results the fundamental parameter to be considered to correctly identify the stellar agglomerations. This is also true for the corresponding parameter Δ_{gas} in the gas density maps. During the course of the identification method, they were therefore used as reference parameters, while the remaining two parameters were set to zero.

A first quantitative comparison method to establish how much the identification of the structures is influenced by the choice of the parameters values is to explore the parameters space

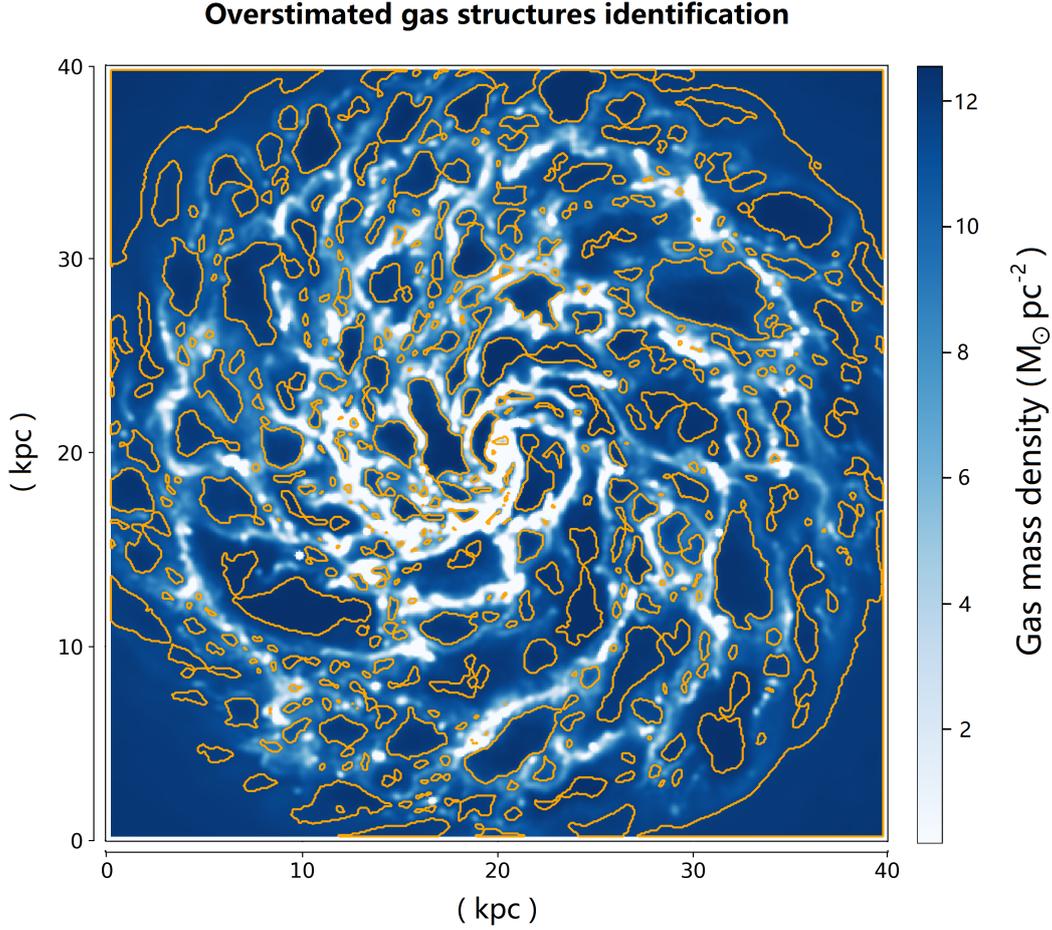


Figure 4.6: Gas column density map (inverted values) with identified sub-dense structures (orange lines). Clearly an overestimation of the gaseous structures found, which also includes not significant peripheral areas, is visible. The set of Astrodendro parameters values used are $(\Sigma_{gas}^{low}, N_{px}^{min}, \Delta_{gas}) = (0, 0, 0.5)$.

$(\Delta_{gas}, \Delta_{\star})$ by comparing the number of gas and stellar structures found for the corresponding set of reference parameter values. The procedure has been performed in an iterative way by counting the number of gas structures N_{gas} that contain at least one stellar structure for every set of reference parameter values, then dividing N_{gas} by the total number of stellar structures N_{\star} , found for the same parameters value. In this way it is possible to compare the values of the ratio

$$R(\Delta_{gas}, \Delta_{\star}) \equiv \frac{N_{gas}}{N_{\star}} \quad (4.10)$$

by varying the choice of the reference parameters values, thus creating a 2D surface curve that is presented in figure (4.8). This curve contains information regarding the correspondence in the number of SB candidates compared to the number of active stellar clusters found by the Astrodendro algorithm with different choice of the reference parameters values. We can

Stellar structures identification with Astrodendro

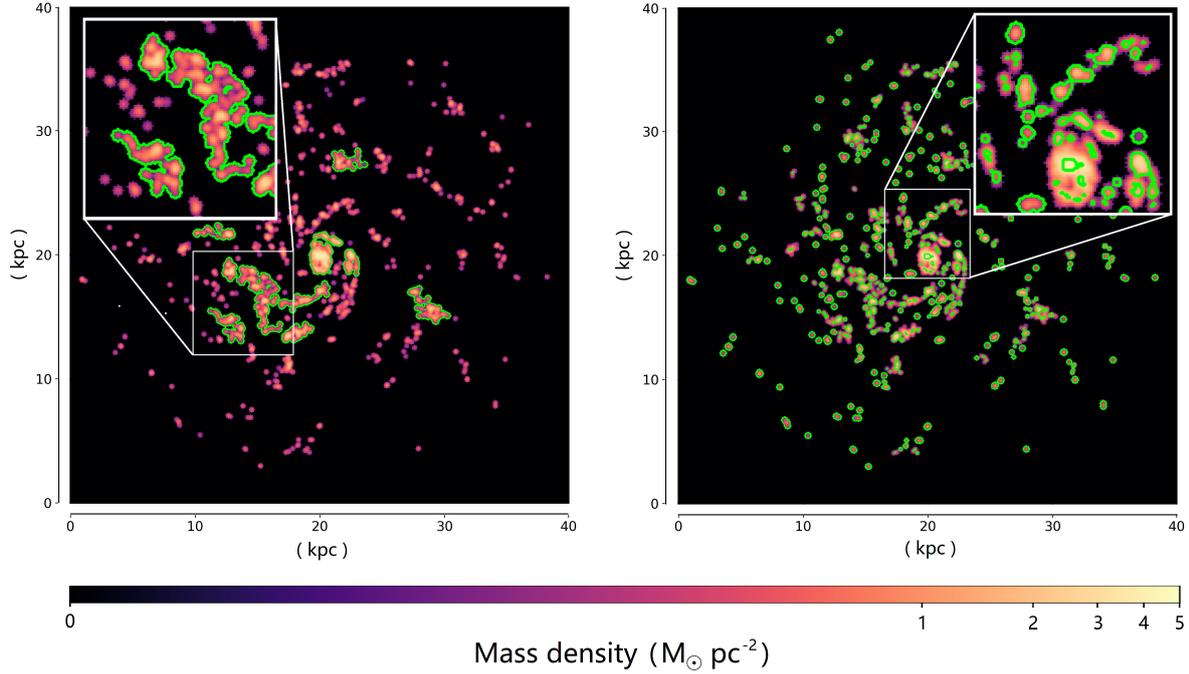


Figure 4.7: Stellar surface density maps with two examples of structures identification produced by the Astrodendro algorithm. **Left:** minimum density variation parameter set to $\Delta_* = 2 M_\odot \text{pc}^{-2}$ resulting in a too much coarse criterion for which few big unlikely structures are identified. **Right:** minimum density variation parameter set to $\Delta_* = 0 M_\odot \text{pc}^{-2}$ leads to a situation with too many identified stellar clusters, with an underestimation of the dimensions related to some of them, as clearly visible in the zoomed inset on the upper right corner.

imagine that in the best possible scenario a 1 : 1 match should be found between sub-dense gas structures and stellar agglomerates, in such a way as to be able to attribute the formation of each individual SB to the SN activity produced by a single stellar cluster. Following this consideration the choice of the reference parameters should be evaluated in the range of values for which $R = 1$, however this approach is not entirely exhaustive as it does not take into account the fluctuations in the counts of these structures caused by variations in the shape and number of stellar clusters that are located within the sub-dense gas regions identified. To clarify this last point, for example, we could imagine that an SB with a large area was generated by successive SN events belonging to different clusters (identified by the Astrodendro algorithm applied to the stellar map with a particular value of Δ_*) that contributed in a coordinated way to expand it on different distant fronts, in this way the identified sub-dense structure (applying Astrodendro to the gas density map with a particular value of Δ_{gas}) could contain more than one

Stellar/gas structures ratio

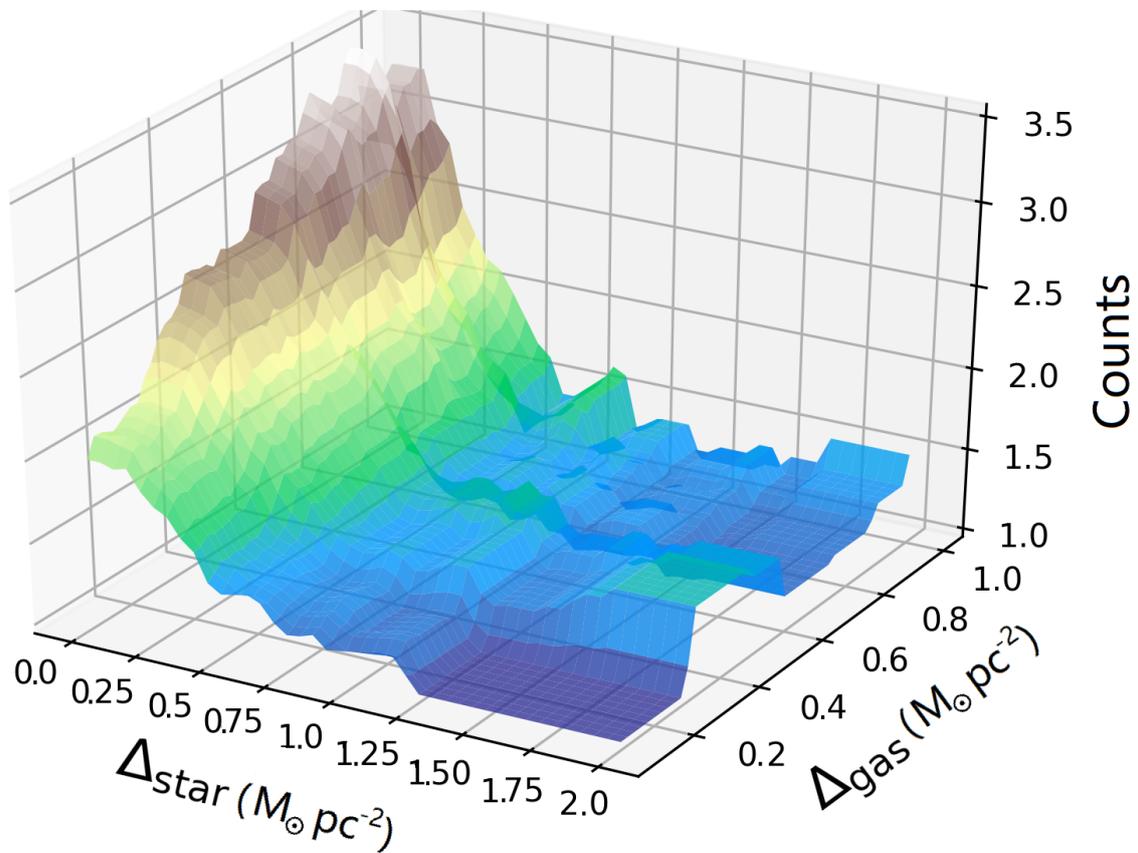


Figure 4.8: Ratio of number of gas structure that contains at least one stellar structure and the total number of stellar structures found by Astrodendro's algorithm by varying the main parameters Δ_{gas} e Δ_{*}

stellar structure, leading to $R > 1$ because of the wrong choice of the parameter Δ_* , but being a good SB candidate. On the other hand, for an unfortunate choice of both parameters, it might happen that two gas structures are so close to each other to be identified as one large sub-dense region which could contain a large stellar structure, thus maintaining a ratio value of $R = 1$ but clearly being a wrong result due to the missed separation of two effective SB candidates. Furthermore, since we have performed a cut in age on the stellar particles we must also consider the possibility that sub-dense regions were generated in the past by stellar activities of no longer existing clusters, thus lacking a correspondence with a stellar structure in the snapshot that we are considering.

All these considerations lead to the need to include in identification method for the SB candidates a series of statistical information in order to better handle all the possible situations that can happen, like those ones described in the previous examples. This is done by including in the method information from the distribution in shape of the structures and from the distribution in counts of stellar structures that fall within a gas structure, as described in section 4.1.5, in such a way as to attribute a frequentist probability measure to every star-gas match identification, with the aim of obtaining the best estimate of the reference parameters (Δ_{gas} , Δ_*) by maximizing the probability associated with the identification.

4.1.5 A frequentist approach

Given the difficulties in finding a set of computational parameters appropriate for a correct search of SB candidates carried out through the Astrodendro package, it is necessary to establish an iterative method of image analysis that allows to generate a mapping of sub-dense gas regions and that assigns a certain probability to the found sample of structures, to be considered as a valid set of SB candidates. We need a method which would avoid going to investigate individually the sub-dense gas regions and to evaluate their possible correspondence with the stellar structures, resulting thus in an extremely laborious work.

The method that we present below assigns a probability to a certain star-gas match result generated by a particular set of reference parameters values. One possible approach is to calculate a frequentist probability starting from the distributions in shape (i.e. axis ratio) of the gas structures and the stellar structures found for a given set of (Δ_{gas} , Δ_*), together with the number of stellar clusters counts which fall within a sub-dense gas region. We are therefore interested in finding the probability that the shape of a sub-dense gas structure is influenced by

the stellar structure(s) that reside(s) within it, evaluating in this way how plausible it is that a sub-dense region was actually created by the SN activity of the stellar cluster(s) within it.

The shape of a structure can be evaluated by the ratio between the minor and the major axis of the best elliptical fit associated with it, which can be retrieved from the Astrodendro statistical computation (as described in section 3.3). In this way, by denoting with r_{gas} and r_* the axis ratios respectively for gas and stellar structures, it is possible to evaluate the roundness of a structure simply comparing the proximity to $r_{gas} = 1$ and $r_* = 1$ of the measured values for all the structures found. In principle, due to the large number of structures found in the dendrogram, there will be a statistical distribution for these values, with a corresponding mean value (\bar{r}_{gas} , \bar{r}_*) and variance (σ_{gas} , σ_*) for gas and stellar structures, respectively. Moreover, a statistical distribution can be attributed to the count of the number N_* of stellar structures falling within a sub-dense gas region, with a corresponding mean value \bar{N}_* and a variance σ_N . Is it thus possible to assigne a conditional probability to every pair of identification parameters values (Δ_{gas} , Δ_*) following the statistical relation

$$P(\bar{r}_{gas}|\bar{N}_*, \bar{r}_*) = \frac{P(\bar{r}_{gas}, \bar{N}_*, \bar{r}_*)}{P(\bar{N}_*, \bar{r}_*)}, \quad (4.11)$$

where $P(\bar{r}_{gas}, \bar{N}_*, \bar{r}_*)$ is the probability of getting a gas structure with axis ratio in the range $\bar{r}_{gas} \pm \sigma_{gas}$ and containing a number of stellar structures between $\bar{N}_* \pm \sigma_N$ with axis ratios in the range $\bar{r}_* \pm \sigma_*$, which is defined with a frequentist approach as

$$P(\bar{r}_{gas}, \bar{N}_*, \bar{r}_*) = \frac{n^\circ SB : r_{gas} \in [\bar{r}_{gas} \pm \sigma_{gas}] \wedge N_* \in [\bar{N}_* \pm \sigma_N] \wedge r_* \in [\bar{r}_* \pm \sigma_*]}{n^\circ SB : N_* > 0}, \quad (4.12)$$

whereas $P(\bar{N}_*, \bar{r}_*)$ is the probability of getting a gas structure containing a number of stellar structures between $\bar{N}_* \pm \sigma_N$ with axis ratio in the range $\bar{r}_* \pm \sigma_*$ and is defined as

$$P(\bar{N}_*, \bar{r}_*) = \frac{n^\circ SB : N_* \in [\bar{N}_* \pm \sigma_N] \wedge r_* \in [\bar{r}_* \pm \sigma_*]}{n^\circ SB : N_* > 0}. \quad (4.13)$$

In practice, the conditional probability comes out by counting all the sub-dense gas structures with a shape reflecting the overall SB candidates distribution found by the Astrodendro algorithm and that contain a number of stellar structures with a given shape in agreement with the stellar statistic distributions. The number of all the sub-dense gas structures that contain at least one stellar structure (i.e $N_* > 0$) cancels out, leaving a conditional probability of the form

$$P(\bar{r}_{gas}|\bar{N}_*, \bar{r}_*) = \frac{n^\circ SB : r_{gas} \in [\bar{r}_{gas} \pm \sigma_{gas}] \wedge N_* \in [\bar{N}_* \pm \sigma_N] \wedge r_* \in [\bar{r}_* \pm \sigma_*]}{n^\circ SB : N_* \in [\bar{N}_* \pm \sigma_N] \wedge r_* \in [\bar{r}_* \pm \sigma_*]}. \quad (4.14)$$

This probability reflects the goodness of a match between gas and stellar structures, quantitatively evaluating how plausible it is that the shape of the sub-dense gas regions has been influenced by the stellar clusters located within them.

Although this probability can be used as a reference value, it is in fact not attributable in any way to any known probability distribution, as it is derived from considerations purely related to the problem in question. It is therefore impossible to use an analytical method to determine the maximum probability of the distribution of values in the parameter space $(\Delta_{gas}, \Delta_{\star})$ and we need to use a random sampling approach to obtain a numerical result that is as close as possible to the true maximum probability value. Therefore, it is necessary to use a Markov Chain type Monte Carlo (MCMC) method to efficiently explore the space of the reference parameters and the associated probability. The method used requires the implementation of a Metropolis-Hasting (Robert, 2015) algorithm, which is largely adopted in problems that require to find the maximum of a probability distribution (section 4.1.6).

4.1.6 Identification using the Metropolis-Hasting MCMC

A Markov chain is a chain of states where the conditional probability of any element \mathbf{x}_n can be expressed as a function of the previous element \mathbf{x}_{n-1} only. The probability $p(\mathbf{x}_{n+1}|\mathbf{x}_n)$ is known as the Markov chain's transition kernel and represents the conditional probability that bounds two adjacent states along the chain. If we consider an ergodic chain², it is possible to demonstrate that the chain has a unique stationary distribution $f(\mathbf{x})$ such that

$$\int_{-\infty}^{+\infty} dx_n f(\mathbf{x}_n) p(\mathbf{x}_{n+1}|\mathbf{x}_n) = f(\mathbf{x}_{n+1}), \quad (4.15)$$

which means that we can produce chains whose states are distributed according to $f(\mathbf{x})$ if we can find a transition kernel that satisfies this requirement.

The Metropolis-Hastings (MH) algorithm is based on finding a transition kernel that will have any desired stationary distribution. The kernel has to satisfy the detailed balance

$$p(\mathbf{x}_{n+1}|\mathbf{x}_n) f(\mathbf{x}_n) = p(\mathbf{x}_n|\mathbf{x}_{n+1}) f(\mathbf{x}_{n+1}) \quad (4.16)$$

for all n . This balance can be imagined as a flow of points out of state \mathbf{x}_n into \mathbf{x}_{n+1} and a counter flow out of \mathbf{x}_{n+1} into \mathbf{x}_n . The flow is proportional to the probability of being in the

²An ergodic chain is a chain in which it exists a number N such that any state of the chain can be reached from any other state in a number of steps equal to or less than N .

first state times the probability of transitioning; in this way to maintain the balance the flow and counter flow between every pair of states must be equal. The stationary state will then be the steady state of the flow and the time the chain spends in a given state will be proportional to $f(\mathbf{x})$, that represents the target distribution to be investigated. The MH algorithm can be implemented following 3 main steps, starting at given state \mathbf{x}_n :

1. Choose a new trial point \mathbf{x}_t from a proposal distribution $q(\mathbf{x}_t|\mathbf{x}_n)$
2. Calculate the rejection rate:

$$\alpha(\mathbf{x}_t, \mathbf{x}_n) = \min \left\{ 1, \frac{q(\mathbf{x}_n|\mathbf{x}_t) f(\mathbf{x}_t)}{q(\mathbf{x}_t|\mathbf{x}_n) f(\mathbf{x}_n)} \right\} \quad (4.17)$$

3. Check the value of α :

- If $\alpha > 1$ accept the trial state and set $\mathbf{x}_{n+1} = \mathbf{x}_t$
- If $\alpha < 1$ draw a uniform deviate D between 0 and 1.

Accept the trial state only if $\alpha > D$ and set $\mathbf{x}_{n+1} = \mathbf{x}_t$, otherwise set $\mathbf{x}_{n+1} = \mathbf{x}_n$

4. Repeat steps 1-3 until the chosen convergence criterion is reached.

The proposal distribution $q(\mathbf{x}_t|\mathbf{x}_n)$, which selects a new trial point \mathbf{x}_t according on the current state \mathbf{x}_n , must be chosen appropriately to allow the convergence of the method in a reasonable time. The chain moves around the parameter space in a random walk and if it does not reach every region of significant probability many times it will not be a good approximation of an independent sampling from the stationary distribution $f(\mathbf{x})$. To achieve good mixing the rejection rate of proposed moves must not be too high or too low. If it is too high, the chain will have many duplicated points that will not fill parameter space in an even way. If the rejection rate is too low, the chain will move but not fast enough to get around the parameter space. A rule of thumb is to reach a rejection rate around $\sim 80\%$, which can be changed adjusting the proposal distribution $q(\mathbf{x}_t|\mathbf{x}_n)$ by varying its definition parameters (Metcalf, R. 2019). Often the proposal distribution is symmetric, $q(\mathbf{x}_t|\mathbf{x}_n) = q(\mathbf{x}_n|\mathbf{x}_t)$, so that it does not come into the definition of rejection rate α at all. Since we need to sample from $q(\mathbf{x}_t|\mathbf{x}_n)$ it makes sense to use a standard distribution with a well implemented random deviate generator, like the multivariate Gaussian centered on the current point, so $\mathbf{x}_{n+1} = \mathbf{x}_n + \mathbf{y}$ where \mathbf{y} is sampled from this latter distribution.

In our case the stationary distribution to be investigated is the conditional probability associated to the structures identification for the set of parameters $\mathbf{x}_n = (\Delta_{gas}, \Delta_*)_n$, that is

$$f(\mathbf{x}_n) = P(\bar{r}_{gas} | \bar{N}_*, \bar{r}_*) \Big|_{(\Delta_{gas}, \Delta_*)_n} \quad (4.18)$$

and the proposal distribution has been chosen as a multivariate Gaussian in two dimensions, which can select a new trial point $\mathbf{x}_t = (\Delta_{gas}, \Delta_*)_t$ depending on the chosen variances values $(\sigma_{\Delta_{gas}}, \sigma_{\Delta_*})$ in both directions. This is formally written as

$$q(\mathbf{x}_t | \mathbf{x}_n) = \mathcal{G}(\mathbf{x}_n, (\sigma_{\Delta_{gas}}, \sigma_{\Delta_*})). \quad (4.19)$$

The variances need to be adjusted until an acceptable rejection rate is found. Reducing σ 's tends to decrease the rejection rate, while when the σ 's are large steps tend to put the proposed new point into regions that are far away from a peak in the probability and thus are rejected. To initialize the chain a guess point in parameter space has to be chosen, then the chain will be attracted by the high probability regions assuming there is some gradient in the local domain around the starting point.

Due to the large sensitivity of the final number of structures identified on the stellar density map on the parameter Δ_* , the variances of the proposal distribution have been chosen equal to $\sigma_{\Delta_{gas}} = 0.1$ and $\sigma_{\Delta_*} = 0.01$. This results in an anisotropic exploration of the parameter space, with finer steps in the stellar density variations. The MH-MCMC algorithm thus implemented has been left to evolve for a thousand steps. For every step the probability data and the acceptance rate have been recorded in order to create a 2D histogram recording the number of steps that occurred within a given range of values in both dimensions of parameters space.

It is critical to know when the chain has converged, thus establishing a criterion by which we can be sure that the desired result has been achieved. One often used approach is to evaluate the autocorrelation for each of the parameters as a function of the lag, m , which represents the separation along the chain. Denoting the parameters values as β and γ , the correlation between them along a chain separation of m can be calculated as

$$C_{\beta, \gamma}(m) = \frac{\sum_{i=1}^{N-m} (\beta_i - \bar{\beta})(\gamma_{i+m} - \bar{\gamma})}{\sqrt{\left(\sum_{i=1}^{N-m} (\beta_i - \bar{\beta})^2\right) \left(\sum_{i=m}^N (\gamma_i - \bar{\gamma})^2\right)}} \quad (4.20)$$

where $N = 1000$ is the chosen total number of chain's steps. In the case of the autocorrelation $\beta = \gamma$. Distant points along the chain should not be correlated so this function should oscillate about zero for large values of the lag m . The first time this function drops to zero or near

zero is an estimate of the correlation length N_{corr} . This reference number tells us what is the minimum separation along the chain that makes two points independent from each other, making it possible to define an effective number of independent samples in the chain as

$$N_{eff} = \frac{N}{N_{corr}}. \quad (4.21)$$

In other words the larger is N_{eff} the more we can be sure that the chain has converged.

In our case the MH-MCMC algorithm has been executed with a number of steps $N = 1000$, obtaining a rejection rate of $\sim 76\%$. The parameters correlation curve computed from the output data of the MH-MCMC algorithm is presented in figure (4.9) from which a correlation length of $N_{corr} = 66$ has been estimated. From this value we can retrieve an effective number of independent samples of $N_{eff} \sim 15$, which can be increased by choosing a larger number of steps in order to improve the convergence of the chain. However, from the rapid decrease of the trend of the cross-correlation between the parameters Δ_{gas} and Δ_* (blue solid line in figure 4.9) we can be quite certain that the chain has reached convergence.

Once the convergence of the chain is established a 2D histogram of the counts of steps that fall within a given bin size has been created. Then the histogram has been smoothed with a Gaussian kernel density estimator in order to retrieve the probability contour levels at constant values. Figure (4.10) shows the result of this procedure. A number of bins equal to 30 has been used in each dimension to produce the figure.

From the recovered probability distribution and the associated contour levels it is possible to get the set of parameters values that correspond to the probability peak of obtaining a strong correlation between gaseous and stellar structures. The maximum evidence of correlation is $P(\bar{r}_{gas}|\bar{N}_*, \bar{r}_*) = 86\%$, which corresponds to the set of parameters $(\Delta_{gas}, \Delta_*) = (3.13, 0.14)$. These are thus the best choice of reference search parameters for gaseous and stellar structures that generate the highest star-gas correlation according to the associated conditional probability $P(\bar{r}_{gas}|\bar{N}_*, \bar{r}_*)$. By inserting these values into the Astrodendro algorithm it was possible to generate two different dendrograms, one from the gas map and one from the stellar map, for which the corresponding structures identification are presented in figure (4.11a). With this set of parameters the number of gas sub-dense structures has been found to be $N_{gas} = 56$ and the corresponding number of correlated stellar structures is $N_* = 90$.

As it is possible to see from the stellar density map (4.11a), in which the corresponding sub-dense gas correlated regions are drawn using red lines, quite a good match has been achieved. For instance, on the upper and bottom left corners of the stellar density map, the zoomed insets

Correlation coefficient in the MCMC chain

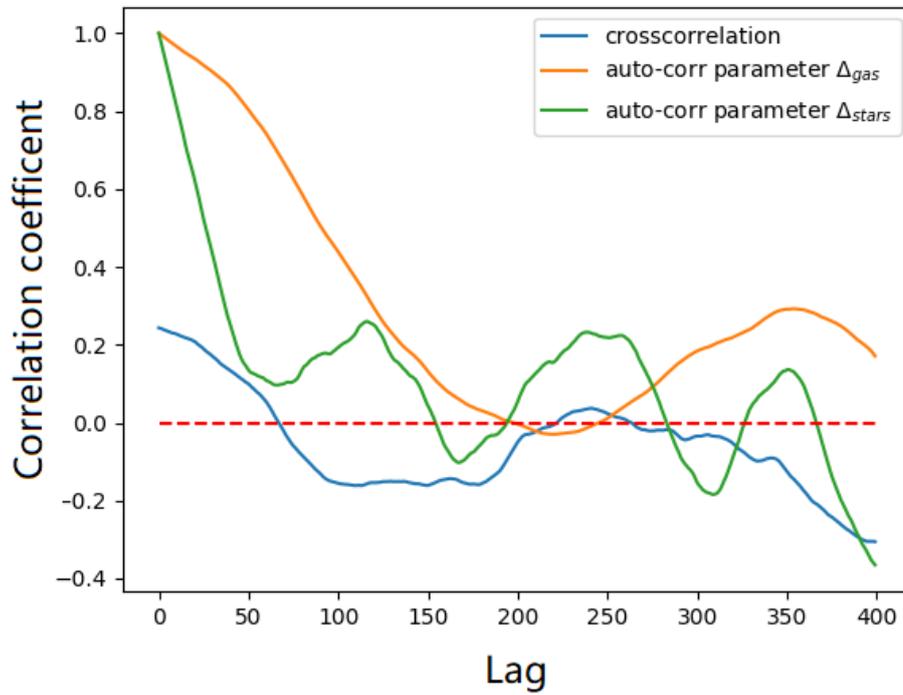


Figure 4.9: Correlation coefficient as a function of the lag in the MCMC chain produced using the Metropolis-Hasting algorithm with a number of steps $N = 1000$. The orange and green solid lines represent the auto-correlations trend for the parameters Δ_{gas} and Δ_* , respectively, while the blue solid line indicates the cross-correlation between the latter two parameters. An effective number of independent samples of $N_{eff} = 66$ is calculated by the intersection of the cross-correlation trend with zero (horizontal dashed red line).

Conditional probability distribution

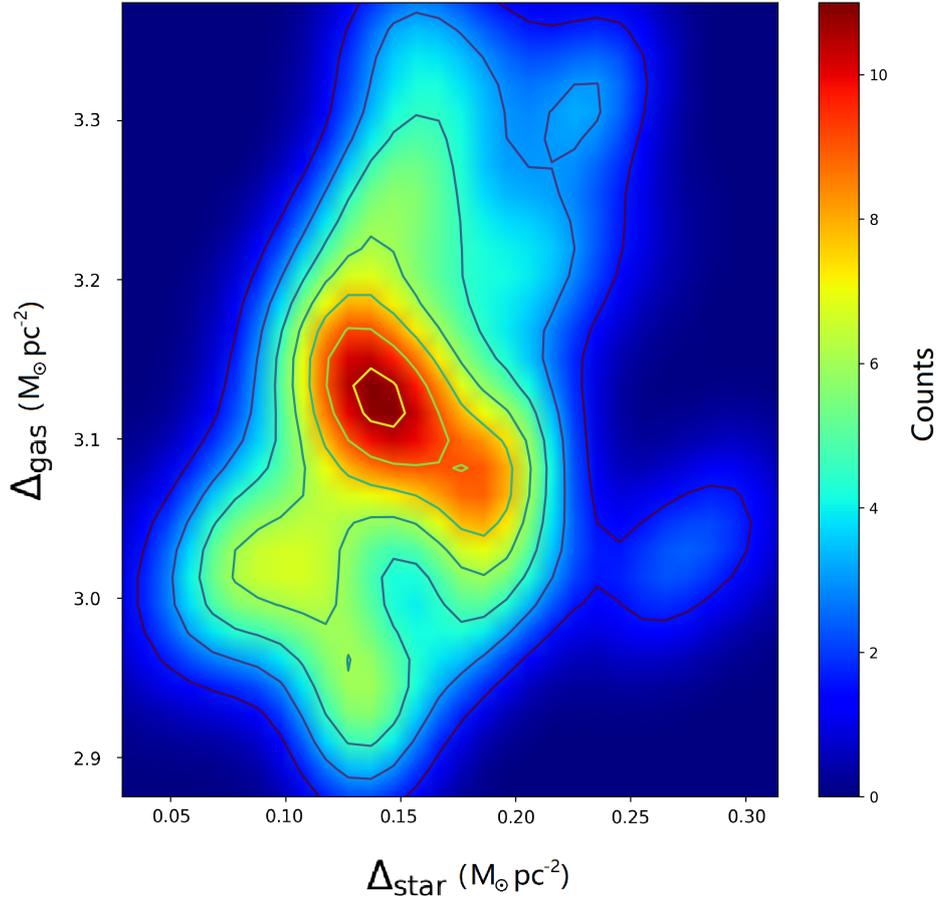


Figure 4.10: Metropolis-Hasting MCMC algorithm result. The explored probability distribution regards the conditional probability $P(\bar{r}_{gas}|\bar{N}_*, \bar{r}_*)$ of obtaining a correlation between gaseous and stellar structures for a given set of reference parameters. The color scale corresponds to the number of counts occurred in a certain 2D bin, and contour levels of the Gaussian kernel density estimator are drawn on the distribution. The maximum probability value is $P(\bar{r}_{gas}|\bar{N}_*, \bar{r}_*) = 86\%$ which corresponds to the pair of parameters $(\Delta_{gas}, \Delta_*) = (3.13, 0.14)$.

show an example of star-gas match where it is possible to notice how well the stellar identified cluster fills the sub-dense gas region, leading to consider these gaseous structures as good SB candidates. In this case it appears very plausible that the gaseous region is being excavated by the stellar activity of the cluster under consideration, creating in this way a super-bubble. A large number of similar situations among all the found star-gas matches are present. In particular, 56 SB candidates have been found to correlate well with the stellar clusters within them. On the other hand, a number of 10 star-gas matches are found to be very unlikely due to their structures shapes in both components (star and gas) or the lack of centering between the structures, as it is shown for example in the zoomed inset on the bottom right corner of the stellar map.

Finally, the result of all the identified sub-dense gas structures is presented in figure (4.11b). Although it is necessary to individually analyze the more irregular structures in order to perform a possible refinement that could lead to consider some of them as good SB candidates, the result achieved by the probabilistic method implemented has proved to be a good starting point to generate in a semi-automated way a search for gaseous structures that can be considered as candidates for super-bubbles. In sections 4.2 a statistical analysis on these identified SB candidates in order to retrieve information from the obtained SB sample, such as dimensions, roundness and positions with respect to the galactic center will be discussed. Furthermore in section 4.2.5 the correlation between the SB radius and the number of SN that generate the corresponding sub-dense region will be investigated, also comparing the results of the simulation with SB theory.

4.2 Statistical analysis of the identified structures

In order to retrieve statistical information from the identified SB sample, the analysis of these structures has focused on four parameters: the radius R_{SB} of each SB candidate, the axis ratio R_{ax} of the best elliptical fit associated to each gas structure, the age t_{SB} of individual sub-dense region and the mass of gas M_{SB} swept from the birth of the structure to t_{SB} . All these parameters have been chosen as quantities of comparison with the observational data, with the aim to establish in a statistical way if the SB generated from the star-gas evolution of the SMUGGLE simulation are in agreement with galaxies observations or model refinement might be needed.

SB candidates identification

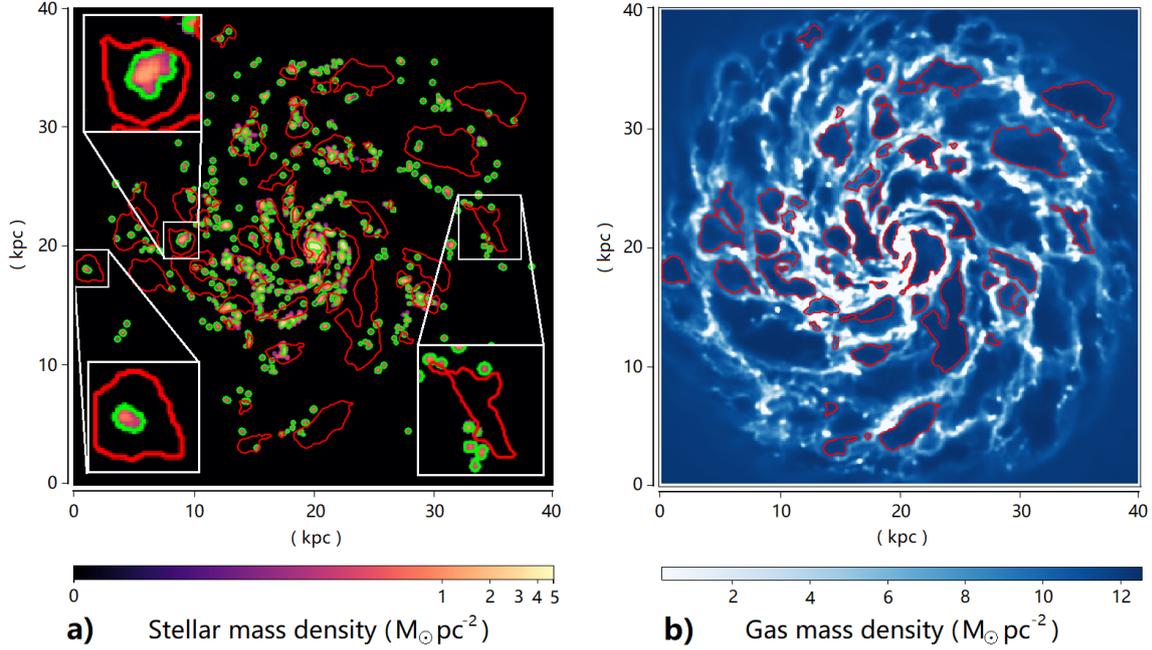


Figure 4.11: Best SB candidates identification generated by applying the Astrorendro computation using parameters $\Delta_{gas} = 3.13 M_{\odot} \text{pc}^{-2}$ and $\Delta_{\star} = 0.14 M_{\odot} \text{pc}^{-2}$, which were obtained from the maximum conditional probability $P(\bar{r}_{gas} | \bar{N}_{\star}, \bar{r}_{\star}) = 86\%$. **Left:** Stellar SN-weighted surface density map with identified stellar clusters (green lines) to which the correspondent match with the sub-dense gas structures (namely SB candidates) was superimposed (red lines). Three zoomed insets are shown to better visualize the identification performed. **Right:** Gas column density map with the corresponding Astrorendro identification of the most evident (in terms of maximum probability) sub-dense regions (red lines, same structures on the left plot).

4.2.1 Super-Bubble radius

To estimate the SB radius a first approach has focused on generating a circle with the same area corresponding to the best elliptical fit area produced by the statistical computation of Astrondendro. Thus taking as SB radius the corresponding "elliptical radius" which has been calculated using the major a and minor b axis values of the ellipse associated to the gas structure, i.e.

$$R_{SB} = \sqrt{ab}. \quad (4.22)$$

However, from a visual comparison, it was found that this method underestimate the true radius of the SB, as shown in figure (4.12a) where is clearly visible that the SB boundary (delimited by a rapid increment of the density) is larger then the radius length estimated in this way.

A more pragmatic approach has then adopted to better identify the SB dimension, which consists in calculating an equivalent radius taking the area delimited by the structure borders identified by the Astrodendro algorithm. Following this method all the pixels belonging to the gas structure were considered in producing a circle with the same area and from which an equivalent radius has been recovered. In this way, if we denote as A_{tot} the area (in physical units) covered by the total number of pixels that make up the gas structure, the equivalent radius can be calculated simply as

$$R_{SB} = \sqrt{\frac{A_{tot}}{\pi}} \quad (4.23)$$

where a conversion to physical units is first applied, following equation (4.3)

$$A_{tot} = N_{px} \cdot A_{px} = N_{px} \cdot \left(\frac{40 \text{ kpc}}{512 \text{ px}}\right)^2 \quad (4.24)$$

with N_{px} being the total number of pixels belonging to the gas structure. This results in a more realistic estimation of the SB radius, as it is shown in figure (4.12b), where a net improvement is clearly visible. To associate a measurement error σ_{SB} to the radius thus determined, a mean square deviation was taken by evaluating the positive and negative deviations from the radius length and the distance of every border pixel with respect to the center of the equivalent circle associated. If we indicate the i -th pixel distance from the center with r_i , it is possible to calculate the equivalent radius error as

$$\sigma_{SB} = \sqrt{\frac{1}{n_{px}} \sum_{i=1}^{n_{px}} (r_i - R_{SB})^2} \quad (4.25)$$

in this way we can evaluate the average deviation between the equivalent radius and the actual outline of the structure. In figure (4.12b) is presented an example to illustrate the procedure adopted to identify the pixels at the border of the structure (black dots on the figure). In the figure every corresponding position with respect to the center of the equivalent circle has been converted in physical units in order to calculate an average displacement from the equivalent radius length.

4.2.2 Age of the superbubble

Another fundamental parameter that characterizes the evolution of these large structures is represented by their age. The SB age can be estimated by considering its mean expansion velocity from the birth to the end of its life when the last SN has exploded. A first approach in determining the SB expansion velocity is to consider the velocity at which the expansion becomes

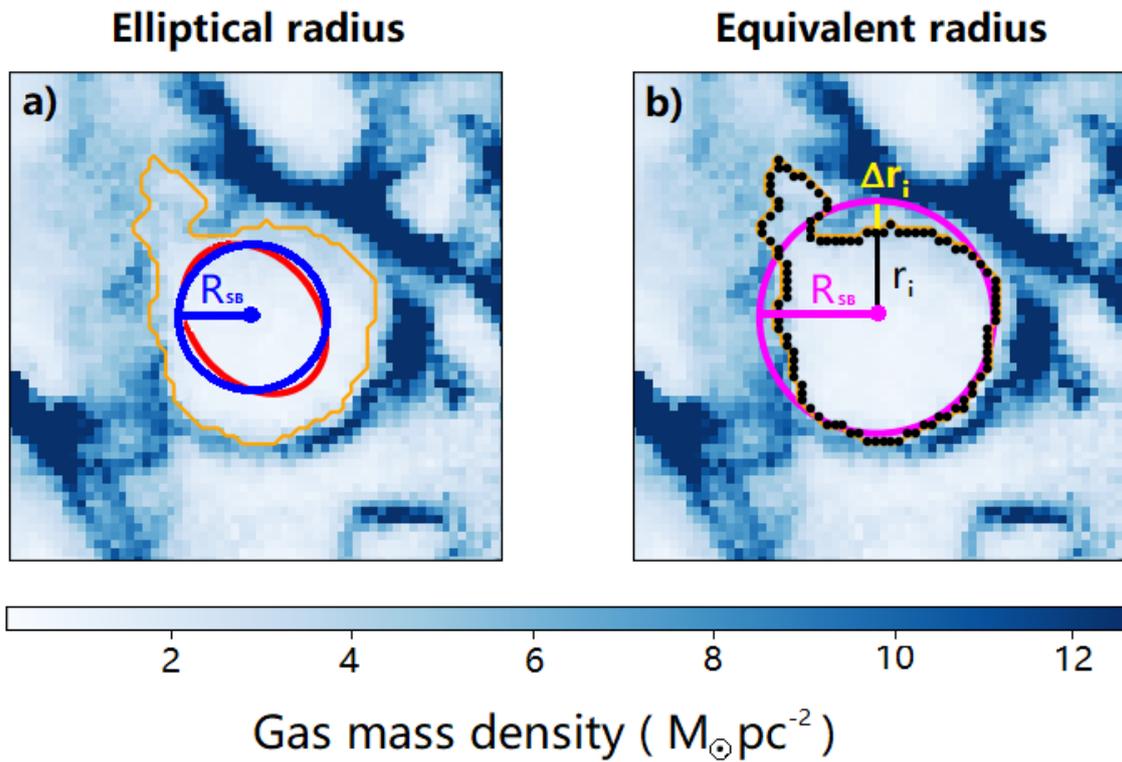


Figure 4.12: Comparison between different methods to compute the gas sub-dense region size. **a)** Elliptical radius (blue line) calculated as radius of the circle with equal area to the elliptical fit (red line) of the structure. **b)** Equivalent radius (magenta line) calculated as radius of the circle with equal area to the structure region. The error associated to the radius is taken as standard deviation of the differences Δr_i of every border pixel (black dots) radius r_i compare to the equivalent radius.

sub-sonic (i.e. the shell velocity becomes less than the sound speed in the gas surrounding the bubble). The ISM sound speed depends on its density and temperature. If we take into account the warm phase ($T \sim 10^4$ K, $n \sim 0.2 \text{ cm}^{-3}$) of the ISM in our galaxy, a typical value of the adiabatic sound speed is of the order of $c_s \sim 15 \text{ km s}^{-1}$. Thus it is possible to calculate the age of a superbubble from the value of its radius R_{SB} , as follows

$$t_{SB} \approx 50 \cdot \left(\frac{R_{SB}}{\text{kpc}} \right) \text{ Myr} \quad (4.26)$$

Although this turns out to be a good approximation, it does not take into account the evolutionary phase at which the SB is situated and the corresponding real expansion velocity value. As it is explained in section 1.2.1, different phases correspond to different expansion velocities and it is important to understand for each individual element in the SB sample retrieved from Astrodendro computation, at which evolutionary phase every SB belongs to. A more accurate approach in determining the SB velocity is thus to follow the radius expansion for different subsequent times and to calculate in this way an average expansion velocity during the main evolutionary phase, when the set of SN explosions is pushing forward the shell of the bubble at almost constant velocity. However, due to the relatively low time resolution between simulation outputs of about ~ 7 Myr, it has been possible to follow the SB evolution for few snapshots only during the main phase. For every pair of consecutive snapshots a mean velocity v_n is calculated by differentiating the radius R_{SB}^{n+1} at time t_{n+1} respect the previous snapshot values, as follows

$$v_n = \frac{R_{SB}^{n+1} - R_{SB}^n}{t_{n+1} - t_n} \approx 140 \cdot \left(\frac{\Delta R_{SB}}{\text{kpc}} \right) \text{ km s}^{-1} \quad (4.27)$$

where the last approximation is evaluated with the constant output time resolution of the SMUGGLE simulation. Thus taking as reference expansion velocity the average

$$\bar{v}_{SB} = \frac{1}{m} \sum_{n=1}^{n=m} v_n \quad (4.28)$$

where m is the number of snapshots considered as reference for the almost flat gradient velocity phase. Finally the age of the SB can be calculated simply by dividing the radius R_{SB} at the last considered snapshot to the mean expansion velocity, namely

$$t_{SB} = \frac{R_{SB}}{\bar{v}_{SB}} \quad (4.29)$$

All the SB radii adopted in the age calculation correspond to the equivalent radii associated to the gaseous structures found in the reference snapshot at $t \sim 700$ Myr, after properly checking

for each SB that was in the correct evolutionary phase and discarding all the SBs that already reached the maximum expansion (i.e. which have perforated the gaseous galactic disk) and for which applying the previous formulae is no longer a valid age estimation.

From the error σ_{SB} associated to the equivalent radius (see equation 4.25) is possible to propagate the error to each measurement v_n of the expansion velocity adopting a zero uncertainty $\delta t = 0$ for the time variable (because is a given value of the simulation, so it is formally exact), obtaining as velocity error δv_n for the n -th measurement

$$\delta v_n = v_n \cdot \sqrt{\left(\frac{\delta \Delta R_{SB}}{\Delta R_{SB}}\right)^2 + \left(\frac{\delta t}{t}\right)^2} = v_n \cdot \frac{\delta \Delta R_{SB}}{\Delta R_{SB}} \quad (4.30)$$

where the error of the difference of SB radii at consecutive times is given by adding in quadrature

$$\delta \Delta R_{SB} = \sqrt{(\sigma_{SB}^{n+1})^2 + (\sigma_{SB}^n)^2} \quad (4.31)$$

with the suffix n and $n + 1$ indicating the two consecutive times under consideration. In this way it is possible to retrieve a mean error for the average expansion velocity \bar{v}_{SB} of the bubble, adding in quadrature all the measurement errors obtained from equation (4.30), as follows

$$\delta \bar{v}_{SB} = \frac{1}{m} \sqrt{\sum_{n=1}^{n=m} \delta v_n^2} \quad (4.32)$$

Finally to the SB age t_{SB} calculated from equation (4.29) is possible to associate an error δt_{SB} of measurement using the propagation

$$\delta t_{SB} = t_{SB} \cdot \sqrt{\left(\frac{\delta R_{SB}}{R_{SB}}\right)^2 + \left(\frac{\delta \bar{v}_{SB}}{\bar{v}_{SB}}\right)^2} \quad (4.33)$$

4.2.3 Mass swept by the SB

The last parameter taken into account to retrieve an appropriate statistics is the mass of gas swept by the SB shell expansion into the ISM, from its birth until the time of the reference snapshot. The mass of each SB has been evaluated by adding up all the mass contributions of each pixel that fall within the equivalent radius R_{SB} at the snapshot referred to the birth time of the SB (determined by the method outlined in section 4.3.2). Using this approach is possible to accurately calculate the mass of gas surrounding the stellar cluster from the center of the equivalent circle to the radius R_{SB} before the SN activity start to excavate the sub-dense SB structure. All the mass of gas thus calculated will then be swept out by the bubble's expansion

during its future evolution until the time of the reference snapshot. From the density values of the pixels belonging to the equivalent circle drawn on the gas density map of the snapshot taken as the birth time of the sub-dense structure, it is possible to calculate the total gas mass of the ISM included by the circle as

$$M_{SB} = \sum_{i=1}^k m_i = \sum_{i=1}^k \sigma_{gas,i} \cdot A_i \quad (4.34)$$

where $\sigma_{gas,i}$ is the gas (column) density value of the pixel i -th with area A_i , which is the same for all pixels (see section 4.1.1 equation 4.3). Although in principle this approach is very simple and accurate, in practice it is somewhat laborious as the birth snapshot must be assessed individually for each structure. Therefore, a less laborious, but still quite accurate method in calculating the mass of gas swept by the SB during its expansion was identified. This method is based on the assumption that all the gas swept by the SB lies on the surface of the shell within a certain thickness. The mass of swept gas can thus be calculated by taking the average gas surface density value within this thickness and multiply this value by the area of the equivalent circle associated to the SB under consideration. Formally it is possible to calculate the mean gas (column) density within a given areola from the formula

$$\bar{\sigma}_{\Delta r} = \frac{1}{k} \sum_{i=1}^k \sigma_{gas,i} \quad (4.35)$$

where k is the number of pixels that belongs to the areola of thickness $\Delta r = R_{SB}/3$. The total mass swept by the SB expansion is thus

$$\tilde{M}_{SB} = \bar{\sigma}_{\Delta r} \cdot \pi R_{SB}^2. \quad (4.36)$$

calculated value of the mass in equation (4.36) has been compared with the mass value for the same SB estimated using the first method, finding a difference of about $\sim 5\%$ on average among all the SBs of the analyzed sample.

Due to the formal accuracy of the density values $\sigma_{gas,i}$ (given as output of the simulation) the error $\delta\tilde{M}_{SB}$ associated to the mass estimated from the equation (4.36) can be retrieved by the radius error σ_{SB} (see equation 4.25) using the error propagation on the power of the radius, i.e.

$$\delta\tilde{M}_{SB} = M_{SB} \bar{\sigma}_{\Delta r} \cdot 2\pi \frac{\delta\sigma_{SB}}{R_{SB}} \quad (4.37)$$

For most of the gas structures taken under examination on the reference snapshot, the two methods in determining the mass swept by the bubble expansion were found to be consistent within the measurement error.

4.2.4 SB histograms

From the previous properties estimation and together with the calculated axis ratios of the elliptical fit associated to every SB candidate, four histograms have been built in order to collect statistical information on the sample under examination. The result is presented in figure (4.13) where data have been grouped using 10 bins for each histogram. For each physical quantity represented in the histograms, the mean and median values are calculated from the sample under consideration, which contains a total number of $N_{SB} = 56$ elements. The median of the SB radius distribution has a value of $\bar{R}_{SB} = 0.49$ kpc but actually this value belongs to the bin range in which only 5 structures appear, thus the mean value of $\langle R_{SB} \rangle = 0.57$ kpc is probably more reliable in representing a typical radius value for the SB candidates because a group of 10 sub-dense structures fall in the same bin range where $\langle R_{SB} \rangle$ belongs to. However particular caution must be applied in the analysis of this distribution of values, because there seem to be two distribution peaks. The second bin of the distribution, corresponding to a middle range of $\tilde{R}_{SB} \simeq 0.35 \pm 0.07$ kpc, appear to be the most probable value for a typical SB radius, as this range has the largest number of structures (i.e. 13). This double-peaked behavior could be due to several factors, not investigated in this work, which could have a physical meaning or which could even prove to be artifacts due to incorrect identification for some structures. In any case a further investigation must be carried out to determine the cause of this type of behavior.

Regarding the SB swept mass, the median value is probably more reliable in representing the sample property due to the long tail of the distribution. A number of 34 SB candidates seem to have swept, in the course of their evolution, a mass of about $M_{SB} \simeq 1.1 \times 10^7 M_{\odot}$. This value is in agreement (as an order of magnitude) with observational data, for example with what observed for the spiral galaxy NGC 6946 (Boomsma et al. 2008). The number of structures decreases rapidly for higher mass values, outlining a small range of possible values between $10^7 \div 10^8 M_{\odot}$. This behavior of the distribution is probably due to the impossibility of sweeping a further significant quantity of mass, by the expansion of a SB, after the latter has perforated the galactic gas disk, thus resulting in a loss of internal pressure of the bubble. The swept mass of each element in the sample under examination has been determined using the method regarding the mean density in the areola of thickness $\Delta r = R_{SB}/3$ and surrounding the equivalent circle (as discussed in section 4.2.3), thus adopting equation (4.36). As illustrated in figure (4.14) the areola thickness result to be in a qualitatively way a good representation of the outer region, evaluating an area in proximity of the bubble surface. Of course different

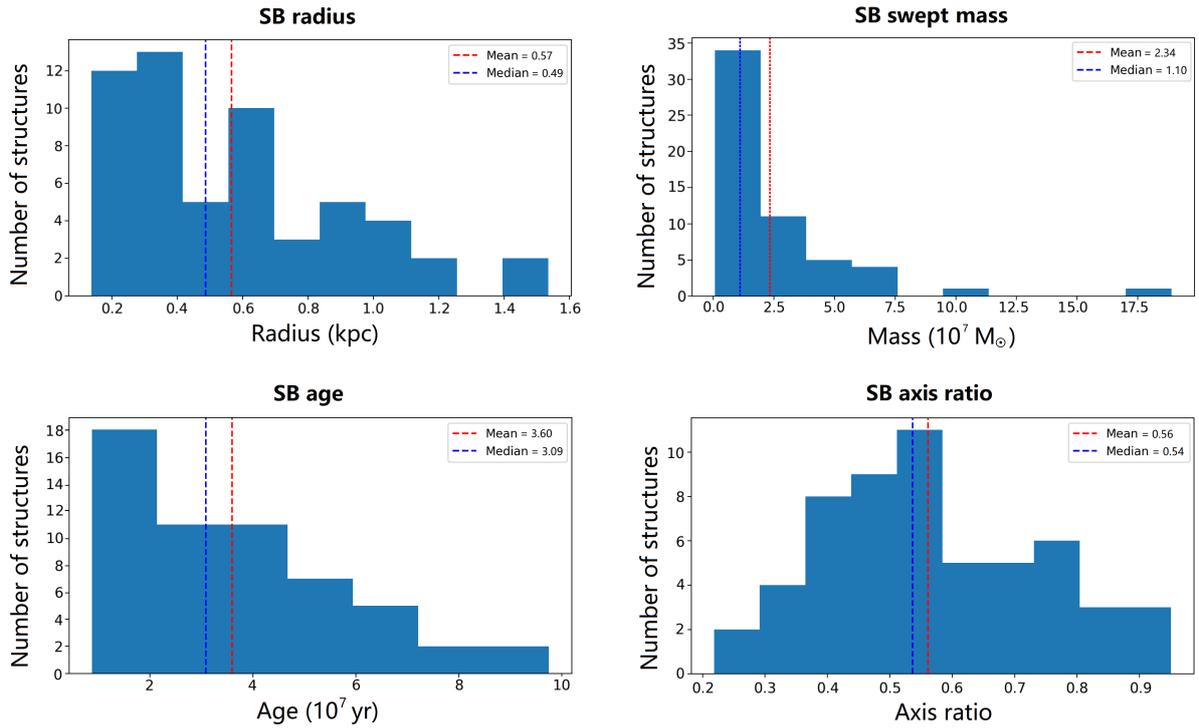


Figure 4.13: Statistical results associated to the SBs sample recovered from the Astrodendro computation with the best star-gas match obtained, as discussed in section 4.1.5. From the first upper left histogram and in clock-wise direction: SB equivalent radius, swept mass, axis ratios and SB age. Mean (red) and median (blue) of the distributions are drawn as reference vertical lines, and the corresponding values are displayed on the legend. A total elements number of $N_{SB} = 56$ from the original sample have been used.

choice of Δr lead to different values of mean gas density inside the areola, thus resulting in different values of swept mass. However some tests have been carried out varying the thickness of the areola between 1/10 and 1/3 of the equivalent radius, finding a relatively small percentage variation in the mass value of approximately $\sim 8\%$ respect to the highest value found for the adopted parameter $\Delta r = R_{SB}/3$

The histogram of the ages has been built by calculating the age of each SB candidate following the methodology discussed in section 4.2.2 regarding the application of a constant nominal velocity attributable to the speed of sound in the ISM. A value of velocity $c_s = 20 \text{ km s}^{-1}$ has been adopted for every SB candidate, deriving an age estimate from the equivalent radius R_{SB} using equation (4.23). The histogram is peaked on the first bin, with middle value of $t_{SB} = 17 \pm 2 \text{ Myr}$, while the median of the distribution is equal to $\bar{t}_{SB} \simeq 31 \pm 5 \text{ Myr}$, where the latter error is calculated from equation (4.33) placing $\delta\bar{v}_{SB} = 0$. Both values are in agreement with observational data ranges (Heiles, 1979), however the median value is more reliable in representing the sample due to the long tail of the distribution.

Finally the axes ratios histogram has been obtained by dividing the minor with major axis of the ellipse associated to each of the gas structure by the Astrodendro computational statistic (section 3.3). The histogram is peaked around a ratio value of ~ 0.5 , with a quite symmetric distribution shape. Both the mean and the median values are representative of the sample under examination given their close proximity to the peak of the distribution. However this reference value must be taken with caution because it probably incorporate the axes ratios of structure that need to be individually analyzed in order to establish the correctness of their identification as SB candidates. From the figure (4.14) the reader can see that some of the structures in the sample have a very irregular shape, this distortion is probably caused by the torsion imparted by the rotating gas. Therefore, even though each of these structures has a stellar cluster within its sub-dense region, particular caution must be taken in its border identification. A further refinement on each individual irregular structure will result in a different sample distribution of the axis ratios, which probably will become peaked to higher value (i.e. more circular structures).

4.2.5 Size vs. number of SuperNovae relation

As discussed in section 1.2.1, during the intermediate evolutionary phase of a superbubble, the bubble shell is in expansion pushed forward by the approximately constant rate of SN explosions. The bubble radius R_{SB} is thus increasing according to the total SN energy produced until

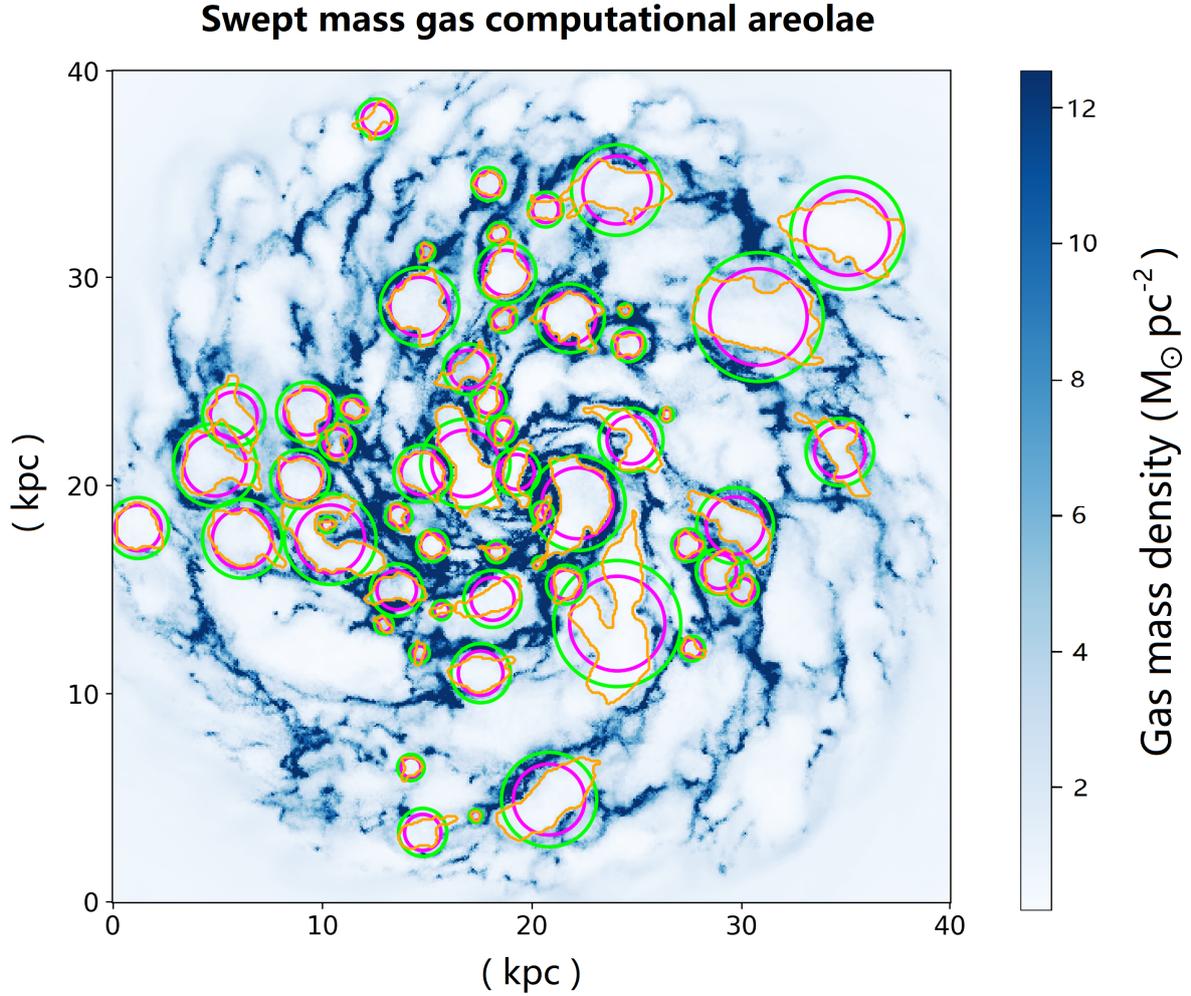


Figure 4.14: Illustration of the method applied to the SBs sample recovered from the Astrodendro computation with the best star-gas match (obtained as discussed in section 4.1.5), to compute the swept mass of gas by each individual SB candidate. The magenta color represents the equivalent circles with radius R_{SB} of each individual SB candidate, while the green areola outlines the region in which the mean gas column density value has been computed. The thickness Δr of the areolae varying with the radius of the SB candidate associated, according to the relation $\Delta r = R_{SB}/3$.

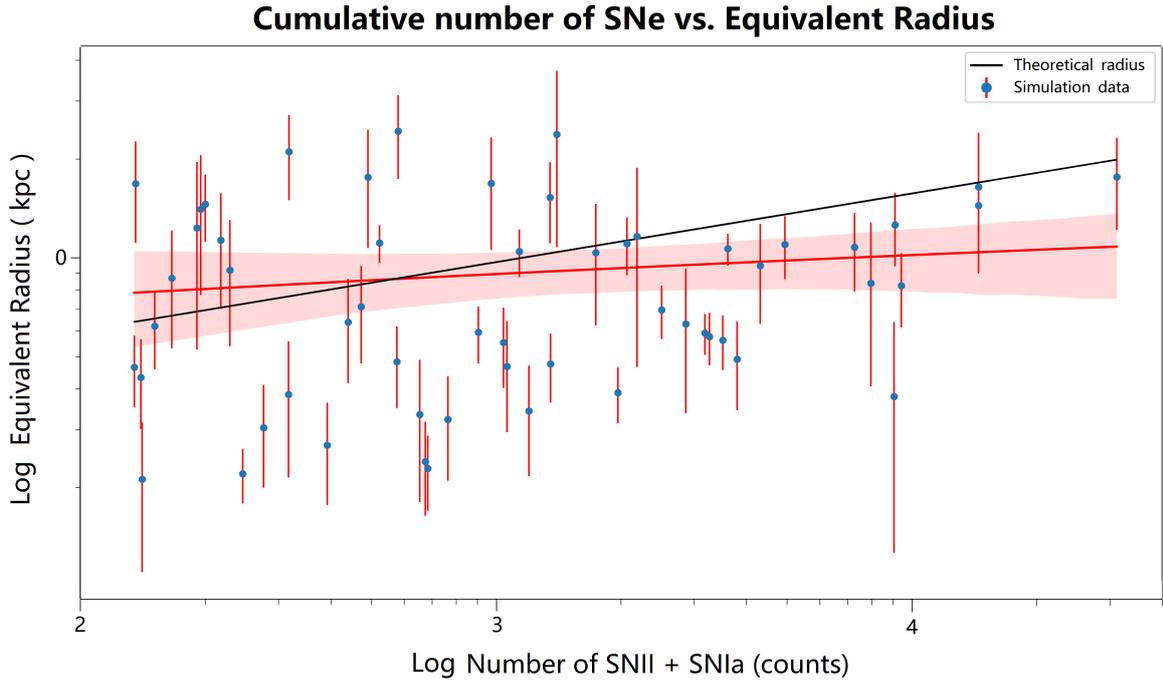


Figure 4.15: Statistical regression of the equivalent radius R_{SB} versus the cumulative number of SNe occurred within 50 Myr for every SB candidate of the best match star-gas sample under examination. The radii values (blue dots) retrieved from the simulation data are plotted with the corresponding error bar (red vertical lines). The fit regression is display as a red continuous line that pass trough the simulation data, with a corresponding red shade area of 95% confidence interval. For comparison, the theoretical trend (black line) is drawn on the plot, obtained from the relation (1.4).

a certain time, resulting in a relation between the number of supernovae exploded within the bubble and its radius (so called "SN-radius" correlation) as proposed by Weaver et al., (1977), in which equation (1.4) was outlined. In order to evaluate the correctness of the SBs evolution developed by the SMUGGLE simulation, the SN-radius relation has been checked in a statistical way, retrieving information on the number of all the SN events occurred in every stellar particle belonging to the stellar structures that formed the sub-dense regions, for all the SB candidates of the sample under consideration. In this way at every bubble radius R_{SB} has been associated a corresponding number of SN events of both type (SNIi and SNIa) by adding up all the contributions from the stellar particles within the sub-dense region of the bubble. The data thus retrieved have been plotted on bilogarithmic scale as presented in figure (4.15) and compared with the theoretical formula (1.4) also plotted on the same figure. Errors in the equivalent radii values are associated to each simulation data using equation (4.25), drawn on the figure as vertical red lines in correspondence of each blue dot.

A comparison between the theoretical trend and the simulation output data can be evaluated by fitting the points on the plot with a linear regression, shown on figure (4.15) as a red solid line that pass trough the data. As it is visible from the figure, the slope of the line fit is lower respect to the theoretical one, being consistent only for cumulative number of SN values below 10^3 , according to the coincidence of the theoretical line within a region (red shade area on the plot) delimited by the 95% confidence interval. For higher SN values the lines tend to diverge from each other. A possible interpretation of the different slope is attributable to the SB candidates that are no longer in the intermediate evolutionary phase, and for which the theoretical relation (1.4) is no longer valid. These structures have already perforated the galactic gas disk and thus their radius slower increases while further SN events occur within them. On the other hand a more accurate estimation of the ambient density value (in which the SB is expanding) needs to be taken in account. In fact in determining the theoretical trend, following the equation (1.4), a mean ambient density nominal value of $n_{gas} = 1 \text{ cm}^{-3}$ has been used. In both cases a further analysis of each individual structure need to be applied, in order to discard the wrong SB candidates or to keep in consideration the true value of the ambient gas density for each individual structure.

4.3 Superbubble evolution

To conclude the analysis related to the identified SB structures, we present in this section a complete description of how a particular SB candidate has been followed in its evolution since the birth of the structure. The methodology applied to generate the gas density profile of the bubble is illustrated in section 4.3.1, while the procedure adopted to follow the evolution of the bubble radius (and velocity) as a function of time is described in section 4.3.2 together with the comparison between the simulation results and theory. Finally, in section 4.3.3 the SN-radius relation for some SB examples is presented, describing the achieved results as well as possible further improvements left for future analysis.

4.3.1 Bubble density profile

For every analyzed SB candidate a gas column density profile has been generated in order to evaluate the positioning of the radii determined by the two approaches discussed in section 4.2.1. In particular we want to investigate if the choice of taking the equivalent radius as repre-

sentative radius for the SB candidate is consistent with the size of the sub-dense region delimited by the peak of density in its profile. From the retrieved properties of the statistical computation performed by Astrodendro (see section 3.3), the elliptical radius has been calculated from equation 4.22 as a reference point to generate the array of areolae in which the SB domain has been subdivided. Denoting as $r_{ellipse}$ the equivalent elliptical radius, the areola array is created starting from zero to a value of $1.5r_{ellipse}$ and with steps width of $r_{ellipse}/10$, thus resulting in a sampling that is fine enough to allow a density profile to be correctly delineated determined for each bubble examined. The density in each areola is then calculated by averaging the density values of all the pixels that fall within the region of that particular areola, following equation 4.35 adapted to this case. Performing this procedure for every areola in the previously defined array leads to a well defined density profile along the radius of the structure, starting from the center of the elliptical fit. The density profile generated in this way for an example case of SB candidate is shown in figure 4.16. The vertical lines corresponding to the position of the elliptical radius (blue line) and the equivalent radius (magenta line) are drawn on the density profile and the corresponding circles are highlighted on the 2D map zoomed inset with respective colors.

From the density profile plot shown in figure (4.16 right) is possible to compare the positioning of the radii obtained with both methods. It is clearly visible that the elliptical radius is an underestimation of the true SB size, while the equivalent radius tends to better approach the peak of density on the profile. This further reinforces the basic soundness in our choice of the equivalent radius as representative dimension of the SB candidate size. An equivalent radius value of $R_{SB} = 1.1 \pm 0.1$ kpc has been obtained for this example of SB candidate. This is probably a case in which the superbubble has already perforated the galactic gas disk due to high value of the radius, however a further investigation is needed in order to evaluate the thickness of the galactic gaseous disk at this evolutionary stage. Turbulent motions caused by repeating large SN events that can generate galactic fountains tend to increase the thickness of the gas disk of the galaxy during its evolution, especially in the outer part of the disk where the gravitational potential generated by the stellar components and by the gas itself is less strong.

4.3.2 Radius evolution

To check to evolution of the equivalent radius length as a function of time and to retrieve the mean expansion velocity value, some of the SB candidates have been monitored under their

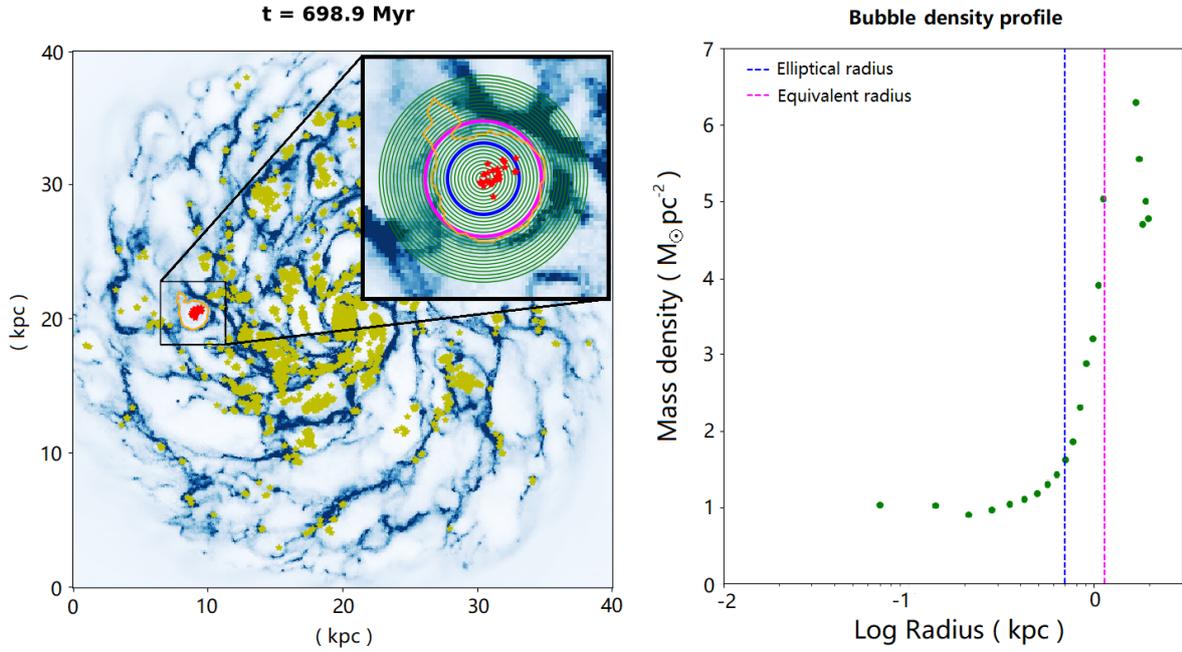


Figure 4.16: Right: mass column gas density profile for a typical SB candidate, with comparison between different methods to determine the radius (drawn as vertical dashed lines). The blue line correspond to the elliptical radius and the magenta line to the equivalent radius, as shown in the legend on the plot. For this example a value of $\simeq 0.7$ kpc has been found for the elliptical radius, while a value of $R_{SB} \simeq 1.1 \pm 0.1$ kpc is associated to the equivalent radius. **Left:** The 2D total gas density map used to extrapolate the bubble profile, referred to the snapshot at $t \sim 700$ Myr. Superimposed on the map are the stellar particles (drawn as greenish points) with ages between 5 and 50 Myr. Highlighted in red are the stellar particles which belong to the sub-dense gas structure under examination. The zoomed inset on the upper right corner contains the analyzed region with areolae subdivision superimposed as green circles and stellar particles (that generated the gas structure) in red. The blue and magenta circles corresponds to the different radii calculated with thw two methods, in order to visually compare them respect to the structure border drawn with an orange line.

evolution through different snapshots. Once a sub-dense gas region has been identified on the reference snapshot at $t \sim 700$ Myr, it has been followed back in time to its birth in order to measure its radius from the beginning of the evolution. Since the structure ID associated by the Astrodendro algorithm (as explained in section 3.1) differs for different dendrograms, which have been generated on different snapshots, the stellar particles belonging to the cluster generator have been used as a tracker to correctly identify the same SB candidate through each snapshot taken under examination. In fact, the stellar particle ID in the simulation does not change from the birth of the particle, thus resulting in an appropriate parameter to follow the SB evolution developed by the SN activity. The gas structure has been followed back in time keeping as reference point the stellar particles that make up the cluster generator, producing a dendrogram structure for every snapshot. The birth of the gas structure has been set to the last snapshot back in time where the Astrodendro algorithm could recognize a sub-dense region surrounding the particle trackers. The equivalent radius has been then calculated (see section 4.2.1) for every snapshot from the birth and forward in time as long as the SB was still recognizable (i.e. when the torsion imparted by the gas motion to the structure mixed it up with other adjacent sub-dense regions). The radius values for a typical SB identified are plotted as a function of time in figure 4.17 where also the differential expansion velocity values have been calculated and presented on the same plot. Each expansion velocity point is calculated following the equation (4.27) and represented on the plot as middle point in time between two adjacent snapshots, in order to properly visualize a differential quantity. A dashed horizontal line drawn on the plot represents the average expansion velocity value, calculated from equation (4.28), which takes into account all the velocity values presented in the plot. As it is possible to see from the velocity curve, an evolutionary phase in which the velocity gradient is almost flat is present, after which a rapid decrease is evident on the last point value. At this final time the SB has probably perforated the galactic disk of gas resulting in a termination of the expansion supported by the internal pressure of the bubble.

Regarding the comparison between the theoretical evolution of the radius and the equivalent radius obtained from the analysis of the same SB candidate tracked through different snapshots, we can notice a quite good agreement at the beginning of the evolution and a small deviation from the theory for the last three points at the end of the SB evolution. If we take as reference point for the bubble to perforate the galactic gaseous disk the point at the highest value of expansion velocity, around $t \sim 30$ Myr, after which is evident a rapid decrement of this quantity,

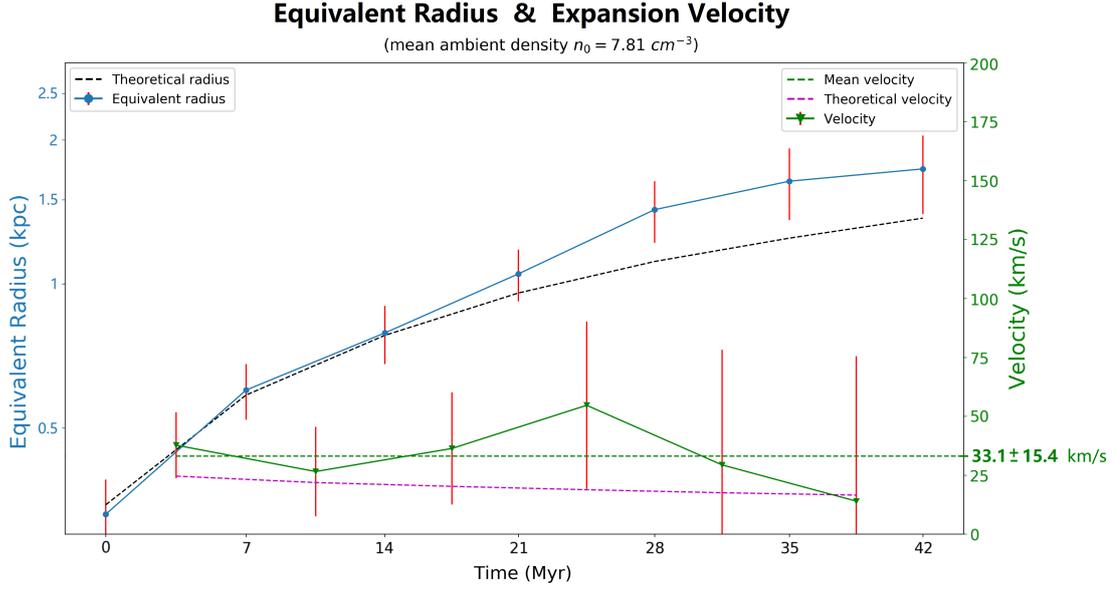


Figure 4.17: Time evolution of the equivalent radius R_{SB} and the corresponding differential expansion velocity v_{SB} , for a typical SB candidate. Data retrieved from the simulation outputs are represented with blue and green dots for the radius and the velocity respectively. The theoretical trends are plotted as dashed black and magenta lines for the radius and velocity respectively, using equation (1.4) and (1.5) with the ambient gas density values obtained from equation (4.35) applying an areola thickness of $\Delta r = R_{SB}/3$. The mean ambient density value of $n_0 = 7.81 \text{ cm}^{-3}$ is calculated from all the snapshots in which the SB has been tracked. Error bars are displayed as red vertical lines associated to each data point, and are calculated using the error propagation formulae (4.25) and (4.30). The average expansion velocity has been calculated to a value of $\bar{v}_{SB} = 33.1 \pm 15.4 \text{ km s}^{-1}$ using equation (4.28) and (4.32) respectively for the mean value and the associated error. A green horizontal dashed line represent this mean expansion velocity value obtained from all the data in which the line pass trough.

it is possible to attribute the deviation from the theory to a no longer valid theoretical law for the evolution of the radius associated to the SB candidate. In fact, due to the loss of internal pressure which is released after the perforation of the gaseous galactic disk, equation (1.4) is no longer applicable to the next evolutionary phase. However this consideration is only a suggestion and a further analysis has to be performed, particularly in determining the thickness profile of the galactic gas disk which is decisive for establishing the actual perforation of the disc by the super bubble. This type of investigation is not part of this work and is left for future analysis.

Although there is a slight difference between the average theoretical expansion velocity value $\bar{v}_{th} = 21.2 \text{ km s}^{-1}$ and the observed one $\bar{v}_{SB} = 33.1 \pm 15.4 \text{ km s}^{-1}$, it can be concluded that the two values are consistent within the error estimate. Using \bar{v}_{SB} as reference expansion velocity it is possible to determine the age of the bubble up to the final radius value $R_{SB} = 1.74 \pm 0.34 \text{ kpc}$, obtaining an age estimation of $t_{SB} = 53 \pm 27 \text{ Myr}$, where equations (4.29) and (4.33) are used respectively for the age value and its associated error. Due to the large uncertainty in the expansion velocity value, the age of the SB candidate has a large error too, however is still consistent as an order of magnitude with observational data (Heiles, 1980). Considering this value as a good estimation of the SB candidate age, we can conclude by comparing it with the final time $t_{fin} = 42 \text{ Myr}$ of the observed structure evolution, that the birth of the structure took place approximately $\sim 10 \text{ Myr}$ before the first recognition as such SB candidate.

In generating the plot shown in figure (4.17) the error bars associated the equivalent radius and the expansion velocity values are calculated following the methodology outlined in section 4.2.1 and 4.2.2 using the error propagation starting from the uncertainty in the equivalent radius measurement. While the ambient density value computation has been performed for every snapshot of the evolution using equation (4.35) and adopting an areola thickness of $\Delta r = R_{SB}/3$. In order to transform the mean gas column density within the areola into a volumetric density, the density value obtained from equation (4.35) has been divided by the scale height $h = 6 \text{ kpc}$ of the galactic gas disk at the initial conditions, thus obtaining a value of $n_0 = 7.81 \text{ cm}^{-3}$. Although the chosen value for h is an approximation of the real thickness value of the gas disk, it leads to a good estimation of the ISM gas density n_0 in which the bubble is expanding, according to the results obtained for the SN-radius correlation fit shown in figure (4.19). More details are provided in section 4.3.3 with a comparison between correlation fit with and without adopting the estimated ambient density values.

Finally a summary of the equivalent radius evolution and the expansion velocity investigated for a number of four SB candidates is presented in figure (4.18). Although some SB candidates show a good agreement with the SB theory, for example the structures called *SB3* and *SB4* show an excellent agreement with the theoretical values for times greater than 14 Myr of evolution, other structures need a more detailed investigation as they show an overestimation of the theoretical radius, as for the structure called *SB1*, or even a completely discordant course with the theory, as for the structure called *SB2*. In relation to the deviation from the theoretical trend as regards the initial evolution of the *SB3* and *SB4* radius, it is probably associated with the failure to reach the intermediate evolutionary phase for which the theoretical equation (1.4) is applicable. In fact it is possible to notice from the two subplots under consideration that a steep gradient in the expansion velocity appears at the beginning of the evolution, which indicates an expansion not yet at constant speed. This initial phase however needs further investigation in order to understand if it is coherent with what observed in real galaxies. For what concerns instead the results discordant with the theory obtained for the structures *SB1* and *SB2* an individual analysis must be performed to understand the reason for these behaviors. An underestimation of the number of SN events due to an incorrect identification of the stellar particles belonging to the cluster generator, an incorrect ambient density value computation or even a wrong structure boundary identification with Astrodendro would lead to a result inconsistent with the theory; we leave further investigations on this aspect for future works.

A summary of the computed values regarding the average expansion velocity \bar{v}_{SB} , the final radius of evolution R_{SB}^{fin} and the age estimation t_{SB} for the structures denoted as *SB1*, *SB2*, *SB3*, *SB4*, is presented in table 4.1. From the expansion velocity values we note, except for the structure *SB3*, a good agreement (within the measurement errors) with the theoretical values \bar{v}_{th} calculated from equation (1.5) and taken as an average between the values from all the snapshots under consideration. While the SB ages and the velocities are consistent as an order of magnitude with observational data from the Milky Way galaxy (Heiles, 1979), the final radius of evolution looks a little bit larger than what observed in the same reference paper. However the agreement is still good within a factor of 2/3. The slight inconsistency between radii values from simulation data and observations could come from the fact of comparing the SB radius in a late evolutionary phase, therefore a more sensible comparison could be made for the SB that have not yet perforated the galactic gas disk. In any case, in order to establish more accurately the agreement between data and theory, measurement errors should be reduced. The relatively

Equivalent Radius & Expansion Velocity

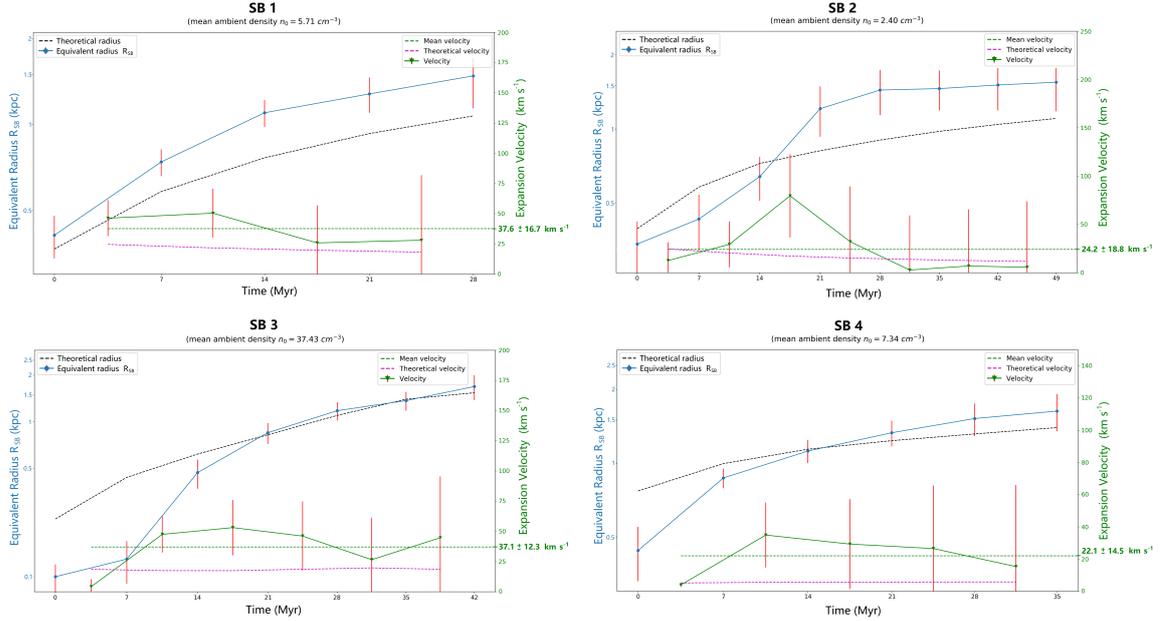


Figure 4.18: Summary of the results regarding the equivalent radius evolution and the expansion velocity obtained for a number of 4 samples taken from the original best star-gas match sample of $N = 56$ SB candidates. In each subplot a comparison with the theory is performed for both the radius and the velocity (dashed black and magenta lines), in order to evaluate the goodness of the simulation output data. Error bars are associated to every data point for both the quantity under consideration, and the average expansion velocity \bar{v}_{SB} values are written on the plots beside the green dashed line which represents the position of the corresponding value on the right scale. The logarithmic scale set for the equivalent radius emphasizes the rapid increment at the beginning of the evolution. All the four values for the mean expansion velocity are consistent with the theoretical mean value, within the measurement error, and are presented in table 4.1 as well as the ages of the structures and the final radius.

large error on the expansion velocity could come from the limitation in time sequence between one snapshot and the next, which for reasons of optimization of memory space they were generated every ~ 7 Myr. A narrower temporal resolution would lead to a decrease in the error on speed and by reflection also on the error of the age of the structure.

4.3.3 SN-radius relation

Finally, the SN-radius relation has been checked by comparing the theory with the simulation data. For each SB candidate examined, data regarding the total number of SN events occurred within the cluster generator have been retrieved for every snapshot in the set of the SB evolution.

	\bar{v}_{SB}	\bar{v}_{th}	R_{SB}^{fin}	t_{SB}
	[km s ⁻¹]	[km s ⁻¹]	[kpc]	[Myr]
SB1	37.6 ± 16.7	27.8	1.48 ± 0.34	38.5 ± 19.3
SB2	24.2 ± 18.8	18.6	1.55 ± 0.37	62.7 ± 51.2
SB3	37.1 ± 12.3	18.7	1.69 ± 0.31	44.6 ± 16.9
SB4	22.1 ± 14.5	8.4	1.63 ± 0.28	72.4 ± 49.4

Table 4.1: Computed values for mean expansion velocity \bar{v}_{SB} , equivalent SB radius at the end of the evolution R_{SB}^{fin} and ages t_{SB} of four SB candidates (*SB1, SB2, SB3, SB4*). The expansion velocity measured from the simulation data is compared with the theoretical mean velocity \bar{v}_{th} calculated by averaging the theoretical velocities obtained from equation (1.5) for every pair of snapshots of the SB evolution. The errors are estimated from the considerations discussed in section 4.2.1 and 4.2.2.

Both SN event types are added up from each stellar particle and related to the equivalent radius calculated for the corresponding snapshot, in order to plot the radii values as a function of total SN events obtained from the birth until the death of the stellar cluster generator. The curve thus created is presented in figure (4.19) for the same SB candidate analyzed in section 4.3.2, where at each point is associated an error bar in the radius measurement (as discussed in section 4.2.1). The dashed black line represents the theoretical curve generated using the equation (1.4), in which the supernovae luminosity L_{SN} has been calculated from equation (1.3) adopting a constant energy per SN equal to $E_{SN} = 10^{51}$ ergs, and the gas ambient density values are calculated taking an areola of thickness $\Delta r = R_{SB}/3$ in which a mean column density value can be obtained from equation (4.35) and then divided by the scale height $h = 0.3$ kpc of the galactic gas disk in order to obtain an estimation of the volumetric density surrounding the SB candidate during its expansion in the ISM. For comparison with the theoretical correlation without taking into account the ambient density, a dashed red line is drawn upon the plot. As we can see from figure (4.19) the ambient density values for each snapshot (which are displayed as green dots) must be taken under consideration to improve the goodness of the theoretical fit applied to the simulation data of equivalent radius as a function of the cumulative number of SNe events. Especially for the beginning of the evolutionary phase, considering the value of the ambient gas density brings the data to be consistent with the theoretical trend, until the fourth data point, after which the simulation data overestimate the theory. Once again a further investigation on the scale height of the galactic gas disk needs to be performed in order to

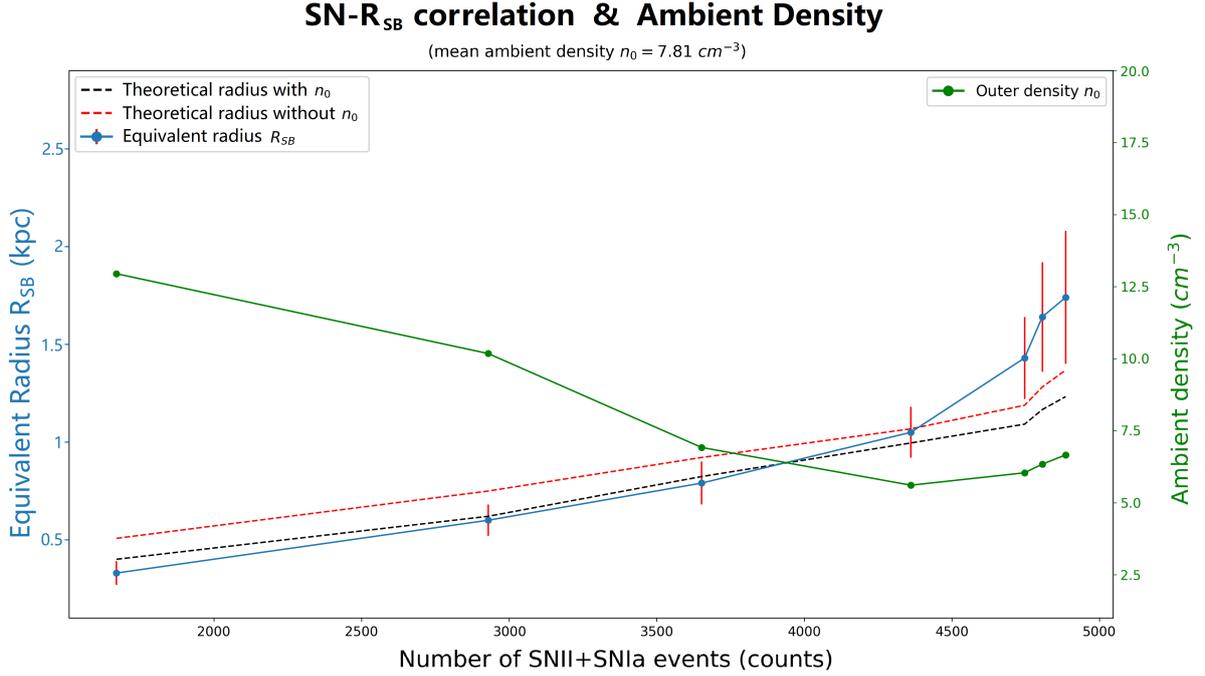


Figure 4.19: Equivalent radius vs. cumulative number of SNe events correlation regarding the SB candidate already analyzed in section 4.3.2. Simulation data for the equivalent radius R_{SB} (blue dots) are represented with their corresponding error bar. A maximum value of $R_{SB} = 1.74 \pm 0.34$ kpc is reached for a corresponding number of supernovae $N_{SN} \simeq 4.9 \times 10^3$. The theoretical radius trends, including (black dashed line) and excluding (red dashed line) the ambient gas density in the calculation, are drawn on the plot. A much better data fit with the theory is obtained taking into account the ISM gas density surrounding the bubble, which is calculated within an areola of thickness $\Delta r = R_{SB}/3$ and displayed as green dots joined by a solid green line

establish until which data point the comparison with the theoretical law (equation 1.4) is still valid.

The solid green line represents the union of the points regarding the ambient gas density estimated in a region surrounding the SB candidate for every evolutionary snapshot. As expected the trend is decreasing as the number of SN events increase and the bubble gets bigger, until it reaches a minimum density value of $n_0^{min} \sim 5.61 \text{ cm}^{-3}$ and then starts to slowly increase. A tentative explanation of this behavior could be due to a first phase in which the decrease in density is generated by the expansion of the surrounding gas due to the pressure imparted by the super-bubble enlargement front, after which the compression of the gas due to the snow-plow effect could lead to a slight increase in the average density near the supershell.

Finally a summary of the SN-radius relation investigated for a number of four SB candidates is presented in figure (4.20). Once again the theoretical radius (black dashed line) is compared with the simulation data of the equivalent radius R_{SB} (blue dots) in each of the subplots. Beside a quite good agreement with the theory for the structure *SB3* and *SB4* (already outlined in section 4.3.2), once again it is possible to notice from the subplots of the figure regarding the structure *SB1* and *SB2* a particular disagreement between simulation data and theoretical values associated to these structures. Further investigations need to be performed in order to discover the origin of this tension.

The trend of the ambient gas density is displayed by a green solid line that pass trough the simulation data, which are computed as mean column density in an areola of thickness $\Delta r = R_{SB}/3$ and divided by the scale height $h = 0.3 \text{ kpc}$ of the galactic gas disk in order to retrieve a volumetric ambient density. A decreasing trend is still visible for three out of four SB candidate, however an opposite trend has been obtained for the structure *SB1* which tends to slightly increase as the SB radius increases. An individual analysis has to be done for this structure, which is left to future investigations, to better understand what really happen to the density in the region nearby the bubble, since phenomena of interaction with other bubbles or collisions with very dense gas filamentary regions could occur. In any case a more precise exploration will have to be carried out, which is not investigated in this work.

An important feature observable for the structure *SB3* is the high initial value of the ambient density that exceed $\sim 120 \text{ cm}^{-3}$ at the first data point. This particular structure has been followed since its very first evolution, where the SB radius is very small and some stellar particles are still being born from the local peak of gas density. This ensures that the model is working

SN- R_{SB} correlation & Ambient Density

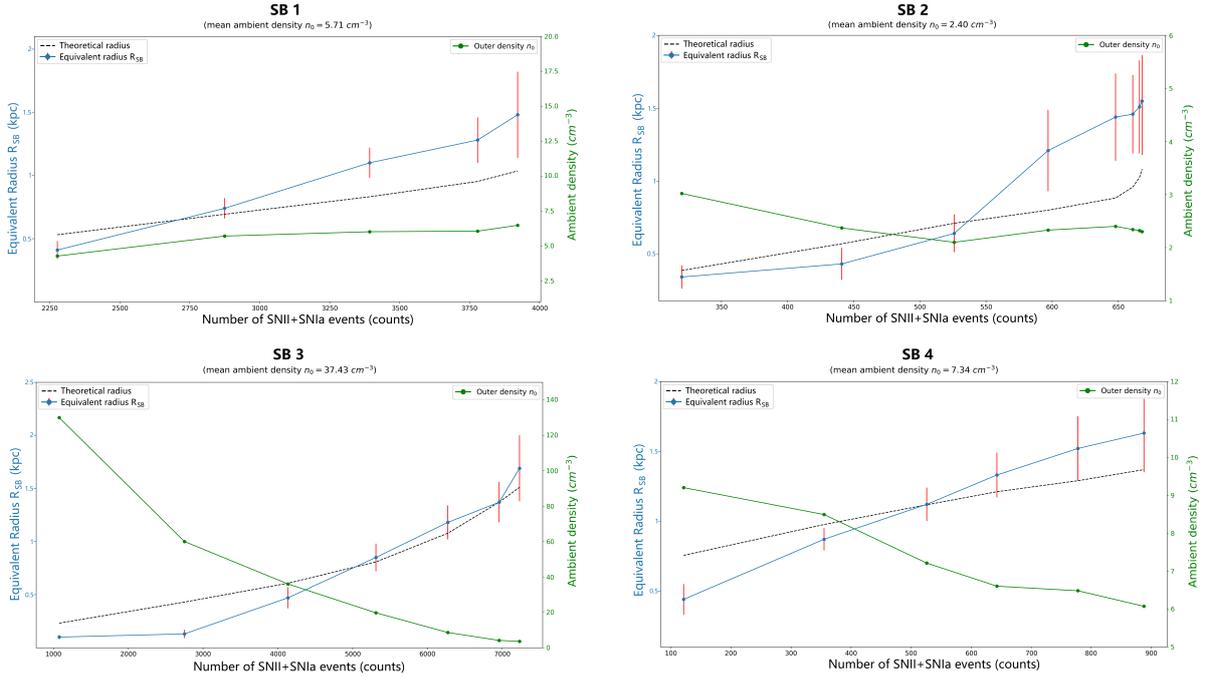


Figure 4.20: Relation between the SB equivalent radius R_{SB} and the cumulative number of SN events occurred in four different SB candidates (the already presented *SB1*, *SB2*, *SB3*, *SB4* in section 4.3.2). The equivalent radius (blue dots) is plotted as a function of total SN events, with the corresponding error bars obtained from equation (4.25). A theoretical trend (black dashed line) for the SN-radius correlation is drawn on each subplots in order to compare it with simulation data. The green scale on the right of each subplots represents the ambient density value, which trend is drawn by a solid green line that join the computed simulation data (green dots). The mean ambient density value has been computed for each SB candidate and it is displayed upon each corresponding subplots.

properly, generating stellar agglomerations in high density regions that exceed the minimum threshold value ($\sim 100 \text{ cm}^{-3}$) for creating new stars from the gas.

Chapter 5

Final discussion

A first preliminary analysis has been completed, regarding the identification and the properties of large sub-dense structures which qualify as superbubble candidates. These sub-dense gas structures are produced during the evolution of the Milky Way-like galaxy simulated using the SMUGGLE model. Although this is only a first attempt to approach the identification problem of SB candidates by analyzing the simulation output data, satisfactory results were obtained regarding the properties extrapolated from these structures in comparison with the observational data for the real Milky Way galaxy and other nearby galaxies ([Boomsma et al., 2008](#))

The performed analysis was entirely based on the use of the Astrodendro package, employed to outline and iteratively trace the sub-dense structures under examination. These structures have been then analyzed either as a statistical sample as well as individual entities, in order to retrieve information relating to the statistical distribution of various properties of the aforementioned SB candidates (such as radii, ages and roundness), or directly measuring these properties from a small sample of the structures kept under considerations. Although some individual structures exhibit slightly discordant behaviors with those observed in real spiral galaxies, most of the SB elements of the statistical sample examined have properties that are consistent with the theory and observational data.

Beyond relatively large measurement errors, associated with the radius estimation and the consequent resulting quantities (such as expansion velocity, age, etc.), we can conclude that the results obtained so far show a remarkable ability of the SMUGGLE model to produce galactic features that are consistent to what is observed in our and other similar spiral galaxies. However to attribute a more strong evidence to these features the above errors will have to be reduced by designing more precise methodologies to better measure the properties of these large structures.

5.1 Conclusions

A statistical method of investigating the identification of the most significant sub-dense structures (from the point of view of stellar activity within them), has been designed, implemented and tested on gas and stellar maps produced through the simulation output data. Conclusions and remarks regarding the applied methodology are presented in section 5.1.1 where the points of strength and the potential shortcomings of the method are outlined. Good results have been obtained by evaluating the measured properties of SB candidates in comparison with observational data and theory, demonstrating the goodness of the galactic model implemented in providing large-scale features comparable to what happens in the reality of the spiral galaxies we observe (see section 5.1.2). Although this investigation was only a preliminary work, we can affirm that it has produced good results, however a more precise evaluation of the properties of the large-scale sub-dense structures will necessarily have to be carried out in future works, to better confirm the correctness of the implemented SMUGGLE model. Possible improvements to the method presented by us and suggestions on new methodologies to be investigated are presented in the section 5.1.3.

5.1.1 SB identification

The method to chose the appropriate Astrodendro parameters which would perform the best possible SB candidates identification, described in section 4.1.5, has been developed starting from statistical considerations on the distribution of the values of some properties of the structures identified with the research performed by Astrondendro for a certain set value of the computational parameters. The conditional probability (equation 4.14) associated to the generated sample of sub-dense structures, has been built using a frequentist approach, that is by counting the structures with a given property falling in a representative range of values for the entire sample under examination, thus resulting in a well defined methodology to be implemented. The conditional probability computed in this way represents in a certain way the correlation between the sub-dense gas regions and the stellar clusters within them for the entire sample of identified structures, thus proposing a quantitative evaluation method in determining the structures that can be identified as SB candidates. However, the proposed method has some weaknesses that need improvement: on one side the method loses its meaning when a dendrogram with few leaves is generated, since the number of elements in the sample becomes statistically no

longer significant; on the other side the method by its nature does not take into consideration the individual entity of the structures, therefore being inaccurate in correctly selecting all the SB candidates, as it often happens that irregular structures fall into the sample of SB candidates due to the fact that they possess within them and perhaps in the marginal area a small stellar cluster. Therefore a final refinement needs to be applied to the SB candidates sample under examination.

For the latter considerations the implemented identification method is a good starting point but still needs improvements to correctly handle some particular situations.

5.1.2 Comparison between model and theory

From the results presented in chapter 4 it is evident that a satisfactory agreement exists between simulation output data and theory, as well as observational data. From the point of view of the statistical sample examined for SB candidates a good agreement can be established between the average (or median) values of the distributions of the SB candidates properties, such as radius, age, swept mass and axial ratio (see histograms of figure 4.13), and the data observed for instance in the spiral galaxy NGC6949 (Boomsma et al., 2008). The only parameter that deviates slightly from the observations is the axial ratio, as it turns out to be too low for our data sample (evaluated at around ~ 0.5 compared with ~ 0.8 found by the same above reference team). However a particular caution must be applied in comparing this value as the sample under examination includes very elongated structures that probably should be discarded as not significant (obviously after having been carefully evaluated individually), thus leading to an increase in the average axial ratio value of the sample. An overall valuation regarding the SN-radius correlation can be obtained by comparing the theoretical trend (from equation 1.4) with the linear regression fit applied to the statistical sample data under examination, as presented in figure (4.15). A quite good agreement is achieved, even if there is an evident wide dispersion of the data and a slight deviation from the theoretical trend is visible for large values of supernova numbers.

Concerning the individual analysis of some exemplary structures of the SB candidates sample, good results have been obtained in comparing the theoretical evolution of the radius of these structures over time and the theoretical correlation between the cumulative number of SN occurred within the structure and its increasing size due to the expansion of the supershell. Although the final radii of evolution are slightly larger than what is observed on average in

real spiral galaxies (Heiles, 1979), the mean velocity expansion and the ages of the SBs under examination are in agreement with observational data (within the measurement errors). Nevertheless two elements of the sample were found to slightly disagree with the theory, showing an overestimation of the relationship of the radius size as a function of time (*SB1* subplot in figure 4.20), or even a complete disagreement on the whole of the evolution followed for the structure (*SB2* subplot in figure 4.20). A more detailed investigation must be carried out to understand how to deal with examples of this type, if on the one hand there was a measurement error or if the simulation produced excessive and / or not physically plausible features.

5.1.3 Future works

To conclude we describe in this section some suggestions on what can be improved in the methodology used in this work and on possible new methods that are worth investigating in future works.

Starting from the limitation (discussed in section 5.1.1) of the method implemented in iteratively identify the SB candidates from the output data of the simulation, a possible way forward could focus on introducing an appropriate selection method in order to discard structures found by Astrodendro that cannot really be considered superbubble, for example by evaluating the axial ratio or the actual position of the stellar cluster within them. A second approach could be centered on generating a further dendrogram for the rectangular region in the gas map in which the structure under examination is inserted, refining the value of the Astrodendro computational parameters until the structure acquires greater significance, in this way for example two adjacent sub-dense regions could be separated into two valid SB candidates.

Regarding the analysis of the simulation output data, a new methodology could be applied to evaluate the radius of the SB candidate, for example by evaluating the point in the density profile of the single structure that reaches a value equal to the local density assessable by generating the galactic gas density profile, and taking the corresponding value of distance from the center of the structure as the representative radius of the latter. Instead the distributions of the quantities of age and mass swept could be improved by calculating these two properties for each individual structure in a precise way. For example, using in the formula for age (4.29) a velocity value taken as the average expansion velocity during the evolution of the structure under consideration (as performed in section 4.3.2) and repeat the calculation for every single structure in the SB candidates sample. For the calculation of the swept mass, on the other hand, it could be more

accurate to evaluate this quantity by taking as reference the output data of the birth snapshot of the SB candidate under examination, and adding the contribution in mass of each pixel that falls within the circle with radius equivalent to the final dimensions of the related sub-dense structure. In this way is taken under consideration all the mass of gas that was actually present in the region swept by the SB expansion.

All these considerations could lead to an effective improvement in the identification and analysis of the properties relating to SB candidates, thus leading to a more accurate assessment of the actual large-scale superbubble phenomena produced by the SMUGGLE model.

Bibliography

- Agertz O., Teyssier R., Moore B., 2011, MNRAS, 410, 1391
- Agertz O., Kravtsov A. V., Leitner S. N., Gnedin N. Y., 2013, ApJ, 770, 25
- Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, ARA&A, 47, 481
- Barnes, J., & Hut, P. 1986, Nature, 324, 446, doi: 10.1038/324446a0
- Bauer, A., & Springel, V. 2012, MNRAS, 423, 2558, doi: 10.1111/j.1365-2966.2012.21058.x
- Boomsma, R. and Oosterloo, T. A. and Fraternali, F. and van der Hulst, J. M. and Sancisi, R., 2008, Astronomy & Astrophysics, 1432-0746
- Bourne, M. A., & Sijacki, D. 2017, MNRAS, 472, 4707, doi: 10.1093/mnras/stx2269
- Brinks, E., and Bajaja, E. 1986, Astr. Ap., 169,14.
- Bruhweiler, F. C, Gull, T. R., Kafatos, M., and Sofia, S. 1980, Ap. J. (Letters), 238 L27
- Castor, J., McCray, R., Weaver, R. 1975, ApJ 200:L107-L110
- Chabrier G., 2001, ApJ, 554, 1274
- Cioffi D. F., McKee C. F., Bertschinger E., 1988, ApJ, 334, 252
- Dufour, R. J. 1984, in IAU Symposium 108, *Structure and Evolution of the Magellanic Clouds*, ed. S. van den Bergh and K. S. de Boer (Dordrecht:Reidel), p. 353.
- Dutta, Prasun, and Somnath Bharadwaj. “*The HI Column Density Power Spectrum of Six Nearby Spiral Galaxies.*” Monthly Notices of the Royal Astronomical Society: Letters 436.1 (2013): L49–L53.
- Ferriere K. M., 2001, Reviews of Modern Physics, 73, 1031
- Field G. B., Goldsmith D. W., Habing H. J., 1969, ApJ, 155, L149
- Fiacconi, D., Sijacki, D., & Pringle, J. E. 2018, MNRAS, 477, 3807, doi: 10.1093/mnras/sty893

Genel, S., Vogelsberger, M., Springel, V., et al. 2014, MNRAS, 445, 175, doi: 10.1093/mnras/stu1654

Goicovic, F. G., Springel, V., Ohlmann, S. T., & Pakmor, R. 2019, MNRAS, 487, 981, doi: 10.1093/mnras/stz1368

Grand, R. J. J., Gomez, F. A., Marinacci, F., et al. 2017, MNRAS, 467, 179, doi: 10.1093/mnras/stx071

Greggio L., 2005, A&A, 441, 1055

Guo F., Oh S. P., 2008, MNRAS, 384, 251

Heiles, C. 1979, Ap. J., 229,533.

Hernquist L., 1990, ApJ, 356, 359

Hindman, J. V. 1967, Australian J. Phys., 20,147.

Hopkins P. F., Quataert E., Murray N., 2011, MNRAS, 417, 950

Hopkins P. F., Quataert E., Murray N., 2012, MNRAS, 421, 3488.

Hopkins P. F., Keres D., Onorbe J., Faucher-Giguere C.-A., Quataert E., Murray N., Bullock J. S., 2014a, MNRAS, 445, 581

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

Jacob, S., Pakmor, R., Simpson, C. M., Springel, V., & Pfrommer, C. 2018, MNRAS, 475, 570, doi: 10.1093/mnras/stx3221

Jenkins E. B., Tripp T. M., 2001, ApJS, 137, 297

Kafatos, M., Sofia, S., Bruhweiler, F., and Gull, T. 1980, Ap. J., 242,294.

Kennicutt Jr. R. C., 1998, ApJ, 498, 541

Krumholz M. R., Tan J. C., 2007, ApJ, 654, 304

Kulkarni, S., Blitz, L., and Heiles, C. 1982, Ap. J. (Letters), 259, L63.

Lopez L. A., Krumholz M. R., Bolatto A. D., Prochaska J. X., Ramirez-Ruiz E., 2011, ApJ, 731, 91

Li M., Bryan G. L., Ostriker J. P., 2017, ApJ, 841, 101

Madsen, G., Reynolds, R., Haffner, L. 2006 ApJ. 652. 10.1086/508441.

- Maoz D., Mannucci F., Brandt T. D., 2012, MNRAS, 426, 3282
- Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018b, MNRAS, 480, 5113, doi: 10.1093/mnras/sty2206
- Marinacci, F., Sales, L. V., Vogelsberger, M., Torrey, P., Springel, V., 2019, arXiv:1905.08806v2 [astro-ph.GA]
- Martizzi D., Fielding D., Faucher-Giguere C.-A., Quataert E., 2016, MNRAS, 459, 2311
- McCray, R. & Kafatos, M. 1986, ApJ, 317:190-196
- McCray, R. & Snow, T. P. Jr. 1979, Ann. Rev. A. Ap., 17, 213
- MacLow & McCray, ApJ 324:776-785, 1987
- Metcalf, R., *Practical statistics for physics and astrophysics*, 2019(Lecture Notes)
- Mo, H., van den Bosch, F. C., & White, S. 2010, Galaxy Formation and Evolution
- Mocz, P., Burkhardt, B., Hernquist, L., McKee, C. F., & Springel, V. 2017, ApJ, 838, 40, doi: 10.3847/1538-4357/aa6475
- Monaghan J. J., Lattanzio J. C., 1985, A&A, 149, 135
- Munoz, D. J., Kratter, K., Springel, V., & Hernquist, L. 2014, MNRAS, 445, 3475, doi: 10.1093/mnras/stu1918
- Murray N., Quataert E., Thompson T. A., 2010, ApJ, 709, 191
- Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, MNRAS, 477, 1206, doi: 10.1093/mnras/sty618
- Nelson, D., Pillepich, A., Springel, V., et al. 2018a, MNRAS, 475, 624, doi: 10.1093/mnras/stx3040
- Ohlmann, S. T., Ropke, F. K., Pakmor, R., & Springel, V. 2016, ApJL, 816, L9, A&A, 599, A5
- Oosterloo, T.; Fraternali, F.; Sancisi, R.; 2007AJ....134.1019O
- Pagel, B. E. J., Edmunds, M. G., Blackwell, D. E., Chun, M. S., and Smith, G. 1979, M.N.R.A.S., 189,95.
- Panagia, N. 1973, A.J., 78,929.
- Pakmor, R., Kromer, M., Taubenberger, S., & Springel, V. 2013, ApJ, 770, L8.

- Pakmor, R., & Springel, V. 2013, MNRAS, 432, 176.
- Pakmor, R., Springel, V., Bauer, A., et al. 2016b, MNRAS, 455, 1134
- Pillepich, A., Nelson, D., Springel, V., Pakmor, R., Torrey, P., Weinberger, R., and Vogelsberger, M., and Marinacci, F., Genel, S. and van der Wel, Arjen and et al.
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, MNRAS, 475, 648
- Portinari L., Chiosi C., Bressan A., 1998, A&A, 334, 505
- Powell, K. G., Roe, P. L., Linde, T. J., Gombosi, T. I., & De Zeeuw, D. L. 1999, Journal of Computational Physics, 154, 284.
- Robert, C. P. 2015, *The Metropolis–Hastings Algorithm*, doi.org/10.1002/9781118445112.stat07834
- Rosie Chen, C. H., You-Hua, C., Gruendl, R. A., Points, S. D. 1999, ApJ 119 : 1317–1324
- Rahmati A., Pawlik A. H., Raicevic M., Schaye J., 2013, MNRAS, 430,2427
- Rybicki G. B., Lightman A. P., 1986, Radiative Processes in Astrophysics. Wiley-VCH, Weinheim, Germany
- Sales L. V., Marinacci F., Springel V., Petkova M., 2014, MNRAS, 439,2990
- Scalo, J. M. 1986, 1986IAUS..116..451S
- Shull, J. M, and Van Steenberg, M. 1985, Ap. J., 294,599.
- Simpson, C. M., Pakmor, R., Marinacci, F., et al. 2016, ApJ, 827, L29, doi: 10.3847/2041-8205/827/2/L29
- Smith, R. J., Glover, S. C. O., Clark, P. C., Klessen, R. S., & Springel, V. 2014, MNRAS, 441, 1628, doi: 10.1093/mnras/stu616
- Sparre, M., Pfrommer, C., & Vogelsberger, M. 2019, MNRAS, 482, 5401, doi: 10.1093/mnras/sty3063
- Springel V., Hernquist L., 2003, MNRAS, 339, 289
- Springel, V. 2005, MNRAS, 364, 1105,
 —. 2010a, ARA&A, 48, 391,
 —. 2010b, MNRAS, 401, 791
- Springel, V. 2010, MNRAS.401..791S

- Springel, V. 2011 *Hydrodynamic simulations on a moving Voronoi mesh*, arXiv:1109.2218 [physics.flu-dyn]
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, MNRAS, 475, 676
- Toro, E. 1997, *Riemann solvers and numerical methods for fluid dynamics* (Springer)
- Vogelsberger M., Genel S., Sijacki D., Torrey P., Springel V., Hernquist L., 2013, MNRAS, 436, 3031
- Vogelsberger, M., Genel, S., Springel, V., et al. 2014, MNRAS, 444, 1518, doi: 10.1093/mnras/stu1536
- Walch S. K., Whitworth A. P., Bisbas T., Wunsch R., Hubber D., 2012, MNRAS, 427, 625
- Weaver, R., McCray, R., Castor, J. 1977, ApJ 218:377-395)
- Weinberger, R., Springel, V., Pakmor, R. 2019, ApJ Supplement Series 248, 1538-4365
- Weinberger, R., Ehlert, K., Pfrommer, C., Pakmor, R., & Springel, V. 2017, MNRAS, 470, 4530
- Wolfire M. G., Hollenbach D., McKee C. F., Tielens A. G. G. M., Bakes E. L. O., 1995, ApJ, 443, 152
- Wolfire M. G., McKee C. F., Hollenbach D., Tielens A. G. G. M., 2003, ApJ, 587, 278