Cosmological Simulations of the Intergalactic Medium Evolution. III. SPH Simulations

Hugo Martel1,2, Benoit Côté3,4, and Laurent Drissen1,2

1 Département de physique, de génie physique et d’optique, Université Laval, Québec, QC, G1V 0A6, Canada; hmartel@phy.ulaval.ca
2 Centre de Recherche en Astrophysique du Québec, C. P. 6128, Succ. Centre-Ville, Montréal, QC H3C 3J7, Canada
3 Department of Physics and Astronomy, University of Victoria, Victoria, BC, V8P 5C2, Canada
4 Konkoly Observatory, Research Centre for Astronomy and Earth Sciences, Eötvös Loránd Research Network (ELKH), Konkoly Thege Miklós út 15-17, H-1121 Budapest, Hungary

Received 2022 February 23; revised 2022 July 18; accepted 2022 July 19; published 2022 August 23

Abstract

We have developed a new numerical algorithm to study the joint evolution of galaxies and the intergalactic medium (IGM) in a cosmological context, with the specific goal of studying the deposition and dispersion of metals in the IGM. This algorithm combines a standard gasdynamical algorithm to simulate the evolution of the IGM, a semi-analytical model to describe the evolution of galaxies, and prescriptions for galaxy formation, accretion, mergers, and tidal disruption. The main goal in designing this algorithm was performance. In its current version, the algorithm can simulate the evolution of cosmological volumes containing thousands of galaxies in a few days, using between 12 and 32 processors. This algorithm is particularly suited for parameter surveys (both numerical parameters and physical parameters) since a large number of simulations can be completed in a fairly short amount of time. Furthermore, the algorithm provides a platform for the development and testing of new treatments of subgrid physics, which could then be implemented into other algorithms. In this paper, we describe the algorithm and present, for illustration, two simulations of the evolution of a (20 Mpc)3 cosmological volume containing ∼1200 galaxies at z = 0.

Unified Astronomy Thesaurus concepts: Intergalactic medium (813); Galaxy evolution (594); Intergalactic abundances (2003)

1. Introduction

Observations of metal absorption lines, such as CIV and O IV, in the spectra of distant sources reveal that the intergalactic medium (IGM) we see today is filled with heavy elements (Meyer & York 1987; Cowie et al. 1995; Tytler et al. 1995; Songaila & Cowie 1996; Cowie & Songaila 1998; Songaila 1998; Ellison et al. 2000; Songaila 2001; Aguirre et al. 2002; Schaye et al. 2003; Aguirre et al. 2004; Aracil et al. 2004; Pieri & Haehnelt 2004; Simcoe et al. 2004; Danforth & Shull 2005; Pieri et al. 2006; Aguirre & Schaye 2007; Aguirre et al. 2008; Danforth & Shull 2008; Pieri et al. 2010, 2010; Tumlinson et al. 2011; Muzahid et al. 2012; Turner et al. 2014, 2015). Since these metals have been synthesized by stars that formed inside high-density regions, a transport mechanism must be invoked to explain their presence in the low-density IGM. Three different mechanisms have been proposed: enrichment by Population III stars (Gnedin & Ostriker 1997; Wasserburg & Qian 2000; Qian et al. 2002; Bromm et al. 2003; Mackey et al. 2003; Venkatesan & Truran 2003; Daigne et al. 2004; Yoshida et al. 2004), early ejection by low-mass galaxies (Cen & Ostriker 1999; Madau et al. 2001; Fujita et al. 2004), and galactic winds powered by supernovae explosions or active galactic nuclei (AGNs; Adelberger et al. 2003, 2005; Veilleux et al. 2005; Songaila 2006; Aguirre & Schaye 2007; Barai et al. 2011). Galaxy mergers and tidal disruption could also result in the dispersion of metals (Domainko et al. 2006), and could explain the high metallicities observed in the intracluster medium of massive clusters, but could not explain the presence of metals in the more distant, low-density IGM.

Galactic winds, on the other hand, could enrich the IGM. Spectacular outflows have been observed in local dwarf starburst galaxies, including the extremely metal-poor I Zw 18 (Péquignot 2008; Jamet et al. 2010) and NGC1569 (Westmoquette et al. 2009). A more distant and compelling example is provided by the low-metallicity starburst galaxy Mrk 1486 (z = 0.034): by mapping the faint, but critical [OIII] λ3763 line (as well as brighter lines) using the Keck KCWI imaging spectrograph, Cameron et al. (2021) clearly showed evidence of chemically enriched outflows along the minor axis of the galaxy. Turner et al. (2015) found evidence of metal-rich outflows in the spectra of 15 hyper-luminous background QSOs that have been observed with Keck/HIRES. More massive spirals, such as NGC7213 (Hameed et al. 2001), also show evidence of global outflows.

Galactic winds can deposit a substantial amount of energy into the IGM. This modifies the physical conditions of the IGM, with important consequences for its subsequent evolution. By increasing the metallicity of the IGM, galaxies increase the cooling efficiency of the gas, which favors the formation of other galaxies. But in some cases, the energy deposited in the IGM can prevent the formation of galaxies by counterbalancing the cooling of pre-galactic clouds, or by disrupting them, which reduces the number of galaxies we see today (Scannapieco et al. 2000; Ciardi & Ferrara 2005; Sigward et al. 2005). Feedback by galactic outflows can provide an explanation for the observed high mass-to-light ratio of dwarf galaxies and the abundance of dwarf galaxies in the Local Group, and can solve various problems with galaxy formation models, such as the overcooling and angular momentum problems (see Benson 2010 and references therein).
Galactic outflows can be launched by various processes, such as supernovae (SNe; Murray et al. 2005; Nath & Silk 2009; Hopkins et al. 2012a; Stringer et al. 2012; Baumgartner & Breitschwerdt 2013; Creasey et al. 2013; Nath & Shchekinov 2013; Roy et al. 2013; Sharma & Nath 2013; von Glasow et al. 2013), radiative pressure (Murray et al. 2011; Sharma et al. 2011; Hopkins et al. 2012a; Sharma & Nath 2012), AGNs (Murray et al. 2005; King et al. 2011; Faucher-Giguère & Quataert 2012; Sharma & Nath 2013; Zubovas & Nayakshin 2014), and cosmic rays (Iapichino 1975; Dorfi 2004; Völk 2007; Samui et al. 2010; Ugliengo et al. 2012; Booth et al. 2013; Salem & Bryan 2014). In all cases, it is ultimately the stellar evolution or the AGN activity of the galaxy that will determine the extent and strength of the outflow, but the outflow, in return, can greatly affect the stellar evolution or the AGN activity of the galaxy, including its ability to form stars. Star formation in the universe has an efficiency of only 10% (Salucci & Persic 1999; Cole et al. 2001; Bell et al. 2003; Kirkman et al. 2003; Pérez-González et al. 2008). By limiting the amount of gas available to form stars, galactic outflows are probably the most important feedback mechanism for reducing the star formation rate (SFR) in galaxies (Veilleux et al. 2005; Dalla Vecchia & Schaye 2008; Bower et al. 2012; Falcke-Gonçalves 2013; Puchwein & Springel 2013), though other mechanisms might also contribute, such as gas heating (McKee & Ostriker 1977; Efstathiou 2000; Hopkins et al. 2012b), gas stripping (Bekki 2009; Benzé & Llambay et al. 2013), and gas evaporation (Pieri & Martel 2007).

While galactic winds transport metal-enriched gas into the IGM, accretion transports gas in the opposite direction, from the IGM to galaxies, including gas that has been enriched in metals by previous generations of galaxies (Kereš et al. 2005; Dekel et al. 2009). Indeed, a study by L’Huillier et al. (2012) suggests that accretion, not galaxy mergers, is the dominant process by which galaxies grow and reach their current masses. In general, the metallicity of the IGM is lower than the metallicity of star-forming regions, and the introduction of IGM gas into galaxies tends to dilute the metallicity of the galactic gas and slow down the stellar enrichment process (Finlator 2017). The evolution of galaxies and the IGM must therefore be considered as symbiotic.

Large-scale cosmological simulations have become a major tool in the study of galaxy formation and the evolution of the IGM at cosmological scales (e.g., Oppenheimer & Davé 2006; Kobayashi et al. 2007; Oppenheimer & Davé 2008; Tesari et al. 2009; Shen et al. 2010; Wiersma et al. 2010; Choi & Nagamine 2011; Davé et al. 2011; Tesari et al. 2011; Schaye et al. 2015; Pillepich et al. 2018). These simulations start at high redshift with a primordial mixture of dark and baryonic matter, and a spectrum of primordial density perturbations. The algorithm simulates the evolution of the system by solving the equations of gravity, hydrodynamics, and (sometimes) radiative transfer. Adding galactic evolution, galactic outflows, and their feedback effect of the IGM in these simulations poses a major practical problem. To simulate a large number of galaxies and interactions, in order to have a representative sample of the universe, the computational volume of the simulations must be large, typically several tens of megaparsecs. But unfortunately, by increasing the size of the simulated volume, low-mass galaxies are either neglected or poorly resolved, because of the limited computational power.

But low-mass galaxies, with their outflows, contribute significantly to the enrichment of the IGM because of their shallow gravitational potential well (e.g., Mac Low & Ferrara 1999; Madau et al. 2001; Booth et al. 2012). Moreover, low-mass galaxies are found in greater numbers than massive ones in the universe. One possible solution consists of using a so-called zoom-in technique. By using a subgrid, or a series of nested subgrids, it is possible to locally increase the resolution. It then becomes possible to resolve even subkilo-parsec scales in a cosmological simulations (see, e.g., Hopkins et al. 2014; Dubois et al. 2021). This approach provides excellent resolution, but poor statistics. A cosmological volume might contain tens of thousands of dwarf galaxies, but only one or a few of them, the ones located in the subgrid regions, will be properly resolved.

To simulate both large and small scales simultaneously, and obtain good statistics, the usual solution consists of simulating the larger scales and using a subgrid physics treatment for the smaller scales. Some studies focus primarily on the propagation of outflows and their impact on the evolution of the IGM. In this case, the algorithm needs a description of the rate of energy, momentum, and metal production, which depend on the mass and formation epoch of the galaxy. The algorithm then simulates the effect of the galactic outflows on the IGM.

Two possible approaches can be used. The first one consists of depositing momentum or thermal energy “by hand” into the system, to simulate the effect of galactic outflows on the surrounding material (Scannapieco et al. 2001; Theuns et al. 2002; Springel & Hernquist 2003; Cen et al. 2005; Kollmeier et al. 2006; Oppenheimer & Davé 2006). The algorithm determines the location of the galaxies producing the outflows and calculates the amount of energy, momentum, and metals deposited into the IGM based on the galaxy properties (mass, formation redshift, etc.). Then, in particle-based algorithms like smoothed particle hydrodynamics (SPH), energy, momentum, and metals are deposited on the nearby particles, while in grid-based algorithms they are deposited on the neighboring grid points. This will result in the formation and expansion of a cavity around each galaxy, which is properly simulated by the algorithm.

The second approach consists of combining dark-matter-only simulations with an analytical prescription to model the impact of outflows originating from high-density regions. Tegmark et al. (1993) developed an analytical model to describe the propagation of galactic outflows in an expanding universe. In this model, a certain amount of energy is released into the interstellar medium (ISM) by SNe during an initial starburst. This energy drives the expansion of a spherical shell that propagates into the surrounding IGM, until it reaches pressure equilibrium. This model, or variations of it, has been used extensively to study the effect of galactic outflows on the IGM (Furlanetto & Loeb 2001; Madau et al. 2001; Scannapieco & Broadhurst 2001; Scannapieco et al. 2002; Scannapieco & Oh 2004; Levine & Gnedin 2005; Pieri et al. 2007; Samui et al. 2008; Germain et al. 2009; Pinsonneault et al. 2010). In this approach, the evolution of the IGM and the propagation of the outflows are calculated separately, but not independently, as they can influence one another. The presence of density inhomogeneities in the IGM can affect the propagation of the outflows (Levine & Gnedin 2005; Pieri et al. 2007), while energy and metals carried by the outflows...
can modify the evolution of the IGM and the location where outflows can take place.

These approaches are appropriate to capture and study the global evolution of the IGM. However, to study the evolution of the galaxies, and how this evolution is influenced by processes such as mergers, tidal disruption, harassment, ram-pressure stripping, and accretion of matter from the IGM, we need a more sophisticated galactic evolution model. Semi-analytical models (SAMs) have been used in the past to reproduce the global properties of galaxies (e.g., White & Frenk 1991; Cole et al. 2000; Springel et al. 2001; Hatton et al. 2003; Baugh 2006; Croton et al. 2006; Monaco et al. 2007; Somerville et al. 2008; Croton et al. 2016; Henriques et al. 2020). These models are significantly more sophisticated than the basic subgrid models that have been used in earlier cosmological simulations. Using the merger trees extracted from the dark matter halos of an N-body cosmological simulation, SAMs can predict the evolution of the galaxies hosted by these halos (e.g., Bertone et al. 2005, 2007). However, although such models can study the propagation of outflows into the IGM, as well as their impact on the stellar content of galaxies, this approach does not include a hydrodynamic treatment of the baryonic physics.

1.1. Long-term Objectives and Present Goals

These various algorithms all have their advantages, and limitations. Our objective is to develop a new cosmological hydrodynamical algorithm that can simulate the joint evolution of the galaxy population and the IGM, while requiring a small fraction of the computer resources used by the multi-billion-particles algorithms. The main advantage of this algorithm will be the ability to easily implement additional physical processes, test the implementation, tune the parameters of the algorithm, and conduct numerous simulations to study the impact of these physical processes. More specifically, the algorithm will be able to:

1. Simulate a comoving cosmological volume containing several thousands of galaxies, taking into account galaxy formation, mergers, tidal disruption, accretion, feedback, and the gasdynamical and chemical evolution of the IGM.

2. Use an SAM for the treatment of subgrid physics, in order to simulate the evolution of the stellar populations and interstellar matter on the fly (not as post-processing), for these thousands of galaxies.

3. Reproduce key observational constraints.

4. Require a modest amount of computer resources, of the order of a few days of wall time per simulation.

We have developed a new SAM that can describe efficiently the concurrent evolution of thousands or even tens of thousands of galaxies, and their interaction with the IGM. This SAM was then implemented into an existing P3M/SPH algorithm in order to provide a treatment of the subgrid physics. In this paper, we present the details of this implementation. Apart from some early attempts (e.g., Metzler & Evrard 1994, who introduced the concept of “galaxy particles”; see, also, Arieli et al. 2010), using an SAM as a subgrid model in a cosmological hydrodynamical simulation to evolve the galaxy population and the IGM concurrently is a novel approach. For this reason, our goal in this paper is not to present a final, definitive version of the algorithm, but rather to demonstrate the feasibility of this approach. Therefore, this work should be seen as a feasibility study. Our goal is to demonstrate that by implementing an SAM in a cosmological algorithm, we can reach all of the aforementioned targets, in terms of number of simulated galaxies, physical processes included, observational constraints satisfied, and computer time.

The remainder of this paper is organized as follows. In Section 2, we describe in detail the implementation of our SAM with its different components into a cosmological gasdynamical algorithm. In Section 3, we present the minimum set of observational constraints that the algorithm must be able to reproduce in order to be considered satisfactory. In Section 4, we describe the setup of two particular simulations that are presented in this paper. In Section 5, we present the results of these simulations, and compare them with observations. In Section 6, we discuss issues that will be addressed in future work. A summary and conclusions are presented in Section 7.

2. The Numerical Algorithm

We use a cosmological P3M/SPH algorithm to describe the evolution of the large-scale structures in a cubic comoving volume with periodic boundary conditions. We implement into this algorithm a subgrid treatment of galaxies, based on the SAM described in Côté et al. (2013) and Côté et al. (2015; hereafter CMD15). Each galaxy is represented by one single, massive particle. These “galaxy particles” have a position, a velocity, a mass, and several internal variables (halo gas mass, mass of cold and hot gas components, stellar mass, metal content, etc.) whose evolution will be handled by the SAM. Representing galaxies with individual particles is a technique that has been used in the past (Metzler & Evrard 1994; Tutukov et al. 2007; Barai et al. 2009; Martel et al. 2012, 2014). Also, a similar approach, called GALCONS, has been developed for Eulerian codes (Arieli et al. 2008, 2010). What most distinguishes our algorithm from these earlier ones is the level of detail and sophistication of the SAM describing the evolution of the galaxies that the particles represent.

The code is divided into three separate modules: a cosmological module, a galaxy module, and a galactic evolution module, each one composed of several routines. The cosmological module handles the evolution of the IGM. The galaxy module handles galaxy formation, and interactions between galaxies. The galactic evolution module handles the internal processes in galaxies using the SAM.

2.1. Basic P3M/SPH Algorithm

The cosmological module is a standard P3M/SPH code (Evrard 1988; Hockney & Eastwood 1988; Monaghan 1992) that simulates the evolution of the large-scale structures and the IGM. The code includes gravity, hydrodynamics, and radiative cooling. Dark matter is represented with $N_{\text{DM}}$ particles of mass $m_{DM}$, and gas is represented with $N_{\text{gas}}$ particles of mass $m_{\text{gas}}$. The simulations are performed inside a comoving cubic volume of size $L_{\text{box}}$ with periodic boundary conditions, using the supercomoving coordinates of Martel & Shapiro (1998). For radiative cooling, we use the metallicity-dependent cooling tables computed by Mappings III (Sutherland & Dopita 1993). Individual time steps are used, and radiative cooling is handled using the implicit integration scheme of Monaghan & Lattanzio (1991). Our chemical model includes 31 elements, ranging from atomic number $Z = 1$ (hydrogen) to $Z = 31$ (gallium). We do not distinguish between various isotopes of the same
At the beginning of the simulation, we initialize these values using primordial abundances. These values are modified over time as metals get deposited into the IGM by halo outflows. While only $\text{[Fe/H]}$ is used to calculate the cooling rate of the IGM gas, we keep track of other elements in order to connect our simulations with observables (e.g., C IV and O VI absorbers).

The SPH algorithm automatically adjusts the smoothing length of gas particles in order to roughly maintain a constant number of neighbors $N_{\text{neigh}}$, typically between 60 and 80. Galaxy particles are also treated as SPH particles, as described in Section 2.7 below.

### 2.2. Galaxy Structure

We use a multizone model to describe the internal properties of the galaxies, as described in CMD15. Each galaxy is made of six components: a dark matter halo; a gaseous halo; a galactic disk containing a hot gaseous component, a cold gaseous component, and a galactic stellar component; and a central supermassive black hole. Figure 1 illustrates the internal variables of each galaxy particle. The global variables are the position, velocity, and total mass of the galaxy particle. $r_{\text{disk}}$ and $z_{\text{disk}}$ are the scale length and scale height of the disk, and concern both the hot and cold component.

Figure 1. Internal variables of each galaxy particle. The global variables are the position, velocity, and total mass of the galaxy particle. $r_{\text{disk}}$ and $z_{\text{disk}}$ are the scale length and scale height of the disk, and concern both the hot and cold component.
distributed according to the initial mass function of Chabrier (2003), in the range 0.1–100 $M_\odot$. For each population, the code keeps track of 16 variables, most of them pertaining to the interstellar bubble that the population generates (for details, we refer the reader to CMD15). During the simulation, new populations are formed while the existing ones are evolved. The total mass of the galaxy particle is $M_{\text{vir}} = M_{\text{DM}} + M_{\text{halo}} + M_{\text{hot}} + M_{\text{cold}} + M_{\text{star}} + M_{\text{BHF}}$. Overall, each galaxy particle carries $23 + 3N_{\text{chem}} + 16N_{\text{pop}}$ internal variables, where $N_{\text{chem}}$ is the number of elements included in the chemical models (currently 31), and $N_{\text{pop}}$ is the number of stellar populations in the galaxy.

### 2.3. Time-stepping

The algorithm evolves the system forward in time using a series of iterations. Each iteration advances the system by a basic time step ($\Delta t_{\text{basic}}$), which is recomputed at the beginning of every iteration. The basic time step is chosen to be the time step that would be appropriate if gravity was the only interaction present in the system (no hydrodynamics, no galaxy formation/evolution).

During each iteration, the algorithm performs the following tasks:

1. Advance the position and velocity of all particles, advance the internal energy and smoothing length of all gas particles, and the smoothing length of all galaxy particles.
2. Handle accretion of IGM material (dark matter and gas) onto galaxies.
3. Calculate the internal evolution of galaxies, using the SAM.
4. Handle IGM outflows by depositing energy and metals onto surrounding gas particles.
5. Handle collisions and tidal destructions of galaxies.
6. Form new galaxies.

The basic P$^3$M/SPH algorithm handles step 1, using individual time steps. Each particle $i$ has its own time step ($\Delta t_i = \Delta t_{\text{basic}}/2^n$, $n = 0, 1, 2, \ldots, n_{\text{max}}$). Figure 2 illustrates the various time steps used during a simulation. Each row represents one basic time step. Particles in Bin 0 are advanced using the basic time step. In other bins, the particles are advanced using a series of shorter time steps, down to the last bin, where the time step $\Delta t_{\text{fund}} = \Delta t_{\text{basic}}/2^{n_{\text{max}}}$ is the fundamental time step. Dark matter particles are put in Bin 0 since they are sensitive to gravity only. Gas and galaxy particles might require shorter time steps, and will be put in the appropriate bins (as explained below, the SPH part of the algorithm treats galaxy particles as gas particles). The internal evolution of galaxies is handled by the galactic evolution module, and each galaxy has its own time step $\Delta t_{\text{evol}}$. The two bottom intervals in Figure 2 illustrate two particular cases. In one case, $\Delta t_{\text{evol}} > \Delta t_{\text{basic}}$, and the SAM will use $\Delta t_{\text{basic}}$ to evolve the galaxy. In the other case, $\Delta t_{\text{evol}} < \Delta t_{\text{basic}}$, and the SAM will use several iterations to evolve this galaxy for a full basic time step. Finally, interactions between galaxies, IGM accretion, and galaxy formation are handled at the end of the basic time step, by the galaxy module.

### 2.4. Galaxy Formation

Our subgrid treatment of galaxy formation, galaxy mergers, and destruction of galaxies by tides is described in detail in Barai et al. (2009) and Martel et al. (2012). This method successfully reproduces the halo mass function of galaxies, and provides a full description of the history of galaxy and cluster formation. The original treatment was for gravity-only simulations, but here we generalize it to simulations with hydrodynamics.

The code forms galaxies with a mass $M_{\text{vir}}^{\text{GF}}$, corresponding to $N_{\text{dm}}^{\text{GF}}$ dark matter particles and $N_{\text{g}}^{\text{GF}}$ gas particles. We typically use equal numbers of dark matter and gas particles, $N_{\text{dm}} = N_{\text{g}}$, in the whole simulation, and use equal-mass dark matter particles and equal-mass gas particles. Then, setting $N_{\text{dm}}^{\text{GF}} = N_{\text{g}}^{\text{GF}}$, ensures that each galaxy will form with a gas fraction equal to the universal value $\Omega_{\text{b}}/\Omega_0$. We define a virial radius $R_{\text{vir}}^{\text{GF}}$ using

$$R_{\text{vir}}^{\text{GF}} = \left(\frac{3M_{\text{vir}}^{\text{GF}}}{800\pi\bar{\rho}_0}\right)^{1/3} (1 + z)^{-1},$$

where $\bar{\rho}_0$ is the present mean density of the universe, and $z$ is the redshift. $R_{\text{vir}}^{\text{GF}}$ is the radius of a sphere of mass $M_{\text{vir}}^{\text{GF}}$ whose mean density is 200 times the mean density of the universe at redshift $z$. The galaxy module identifies potential sites for galaxy formation and forms galaxies of mass $M_{\text{vir}}^{\text{GF}}$ at those sites. More massive galaxies will eventually form as a result of mergers and accretion. These processes are handled by other routines. To handle galaxy formation, the code uses a list of dense particles. Dense particles are defined as gas particles with density above a density threshold $\rho_{\text{thr}} = 200\bar{\rho}_0(z)$, where $\rho_{\text{thr}} = \Omega_{\text{b}}\rho_0/\Omega_0$ is the mean baryon density. At the end of each basic time step $\Delta t_{\text{basic}}$, the code fills up the list of dense particles. The code then sorts this list in decreasing order of the density, and loops over all particles in the list, starting with the densest. For each particle:

1. The code identifies all gas and dark matter particles located within a radius $R_{\text{vir}}^{\text{GF}}$ of the dense particle. It then selects the $N_{\text{dm}}^{\text{GF}}$ dark matter particles and $N_{\text{g}}^{\text{GF}}$ gas particles nearest to the dense particle. If there are not enough particles within $R_{\text{vir}}^{\text{GF}}$ to form a galaxy (fewer than

---

5 The latter includes the dense particle itself, which is a gas particle.
2. If a galaxy forms, these \( N_{\text{DM}}^{\text{GF}} \) dark matter particles and \( N_{\text{g}}^{\text{GF}} \) gas particles are removed from the simulation and replaced by a galaxy particle. This particle is given the total mass of the particles being removed, and the position and velocity of their center of mass. Therefore, mass and momentum are conserved when the code forms a galaxy.

3. The code checks the list of dense particles to see if any particle remaining on the list has been removed when the galaxy formed. If so, these particles are crossed out from the list. This is usually the case, as dense particles are often neighbors. If two particles on the list are separated by a distance \( r < R_{\text{vir}}^{\text{GF}} \), they will end up in the same galaxy.

This process continues until it reaches the end of the list of dense particles. The internal variables of each galaxy are initialized as follows: the dark matter mass \( M_{\text{DM}} \) and gas halo mass \( M_{\text{halo}} \) are set, respectively, to the total mass of the dark matter particles and the total mass of the gas particles removed during the formation of the galaxy. Also, the mass of each chemical element in the halo is set to the sum of the masses of this element in the gas particles; hence, the mass of each element is conserved separately. The masses of the hot, cold, and stellar components \( (M_{\text{hot}}, M_{\text{cold}}, \text{and } M_{\ast}) \) are set to zero. The average velocity \( V_{\text{vir}} \) of the dark matter and the halo temperature \( T_{\text{vir}} \) are calculated using:

\[
V_{\text{vir}}^2 = \frac{GM_{\text{vir}}}{R_{\text{vir}}}, \tag{2}
\]

\[
k_B T_{\text{vir}} = \frac{1}{2} \mu m_{\text{H}} V_{\text{vir}}^2, \tag{3}
\]

with \( M_{\text{vir}} = M_{\text{vir}}^{\text{GF}} \) and \( R_{\text{vir}} = R_{\text{vir}}^{\text{GF}} \). In Equation (3), \( k_B \) is the Boltzmann constant, \( \mu \) is the mean molecular weight, and \( m_{\text{H}} \) is the mass of a hydrogen atom, respectively. The central black hole is given an initial mass \( M_{\text{seed}} \).

There is an important difference between our treatment of galaxy formation and the treatment of “star particles” in some SPH algorithms. We only impose a condition on the density \((\rho > 200 \rho_{\odot})\), while other algorithms also impose a condition on the temperature \((\text{e.g., } \text{Marri & White } 2003)\). In our algorithm, the cooling of the gas inside galactic halos is handled by the SAM, not by the SPH code. Hence, there will be a delay between the formation of a galaxy and star formation inside that galaxy, due to the time required for the gas to cool. One could argue that our galaxy formation routine is actually forming protogalaxies.

2.5. Galaxy Mergers

To handle galaxy mergers, the algorithm searches for pairs of galaxies with separation \( r_{ij} < \max(R_{\text{vir},i}, R_{\text{vir},j}) \), where \( R_{\text{vir},i}, R_{\text{vir},j} \) are the radii of galaxies \( i \) and \( j \), respectively. If this condition is satisfied, the galaxies are considered candidates for a merger. The algorithm then computes the total energy of the pair of galaxies, in the center-of-mass rest frame:

\[
E_{ij} = K_i + K_j + W_{ij} + U_i + U_j, \tag{4}
\]

where \( K_i, K_j \) are the kinetic energies of the galaxies in the center-of-mass rest frame, \( W_{ij} = -GM_{\text{vir},i}M_{\text{vir},j}/r_{ij} \) is the gravitational potential energy of the pair, and \( U_i, U_j \) are the internal energies of the galaxies, defined by \( U = -\zeta GM_{\text{vir}}^2/2R_{\text{vir}} \), with \( \zeta = 1 \) (see Barai et al. 2009). Note that we neglect the thermal energy of the gas. If the condition \( E_{ij} < 0 \) is satisfied, a merger takes place. Galaxies \( i \) and \( j \) are removed from the simulation and replaced by a new galaxy \( k \).

This galaxy is initialized by merging the various components separately, as follows:

\[
M_{\text{DM},k} = M_{\text{DM},i} + M_{\text{DM},j}, \tag{5}
\]

\[
M_{\text{halo},k} = M_{\text{halo},i} + M_{\text{halo},j}, \tag{6}
\]

\[
M_{\text{cold},k} = M_{\text{cold},i} + M_{\text{cold},j}, \tag{7}
\]

\[
M_{\text{hot},k} = M_{\text{hot},i} + M_{\text{hot},j}, \tag{8}
\]

\[
M_{\ast,k} = M_{\ast,i} + M_{\ast,j}, \tag{9}
\]

\[
M_{\text{BH},k} = M_{\text{BH},i} + M_{\text{BH},j}. \tag{10}
\]

Alternatives to this approach are discussed in Sections 6.3 and 6.5 below. Note that these equations imply \( M_{\text{vir},k} = M_{\text{vir},i} + M_{\text{vir},j} \). The virial radius is initialized using:

\[
R_{\text{vir},k} = \frac{\zeta GM_{\text{vir},k}^2}{2|E_{ij}|}. \tag{11}
\]

This equation comes from the fact that the energy of the system in the center-of-mass frame is entirely contained in the internal energy \( U_k \) of the merger remnant after the merger is completed. The new galaxy is given the position and velocity of the center of mass in order to conserve momentum. The temperature \( T_{\text{vir}} \) of the halo is then calculated using Equations (2) and (3), with the new values of \( M_{\text{vir}} \) and \( R_{\text{vir}} \). Finally, the stellar populations of galaxies \( i \) and \( j \) are all incorporated into galaxy \( k \).

Equation (10) requires some explanation. When two galaxies collide and merge, the black holes will lose energy and angular momentum by dynamical friction, and sink toward the center of the system, forming a binary black hole. This binary will keep losing energy by dynamical friction, until the separation becomes shorter than the local mean separation between stars (Quinlan 1996; Makino 1997). From this point, the binary will keep losing energy by gravitational radiation, until the two black holes merge. The timescale of this latter process, however, can well exceed a Hubble time. Hence, it is conceivable that many massive galaxies host a central binary black hole. By using Equation (10), we are assuming that accretion onto a binary black hole would result in the same amount of feedback as accretion onto a single black hole with the same total mass.

2.6. Tidal Destruction

As the simulation proceeds, galaxies can tidally disrupt one another. Tidal disruption is a complex process. The outer parts of the galaxy might get peeled off and dispersed into the IGM, while the central parts might survive the encounter. If the galaxy is not entirely destroyed, its structure can be strongly altered. Furthermore, the presence of a tidal field can affect the SFR (Renaud et al. 2009). Modeling tidal disruption of individual galaxies with an SAM would be a daunting task. However, we are dealing in our simulations with thousands of galaxies and hundreds of thousands of encounters that might lead to tidal disruption. This enables us to use a simple approach that statistically provides a correct description of the
global effect of tidal disruption (Barai et al. 2009; Martel et al. 2012). We essentially adopt an all-or-nothing approach: a galaxy experiencing an external tidal field will either be totally destroyed or else left undisturbed. In practice, we estimate the minimum tidal field required to unbind 50% of the mass of the galaxy. If the actual tidal field exceeds that value, the algorithm considers that the galaxy is totally destroyed. If not, the galaxy is left untouched.

2.7. Accretion and the Doughnut Hole Effect

Galaxies can increase their mass significantly by accreting matter from their surroundings. A study by L’Huillier et al. (2012) suggests that three-fourths of the present mass of galaxies was assembled by accretion, and only one-fourth by mergers with smaller galaxies. We implement accretion in our algorithm as follows: a particle (gas or dark matter) is considered a candidate for accretion when it gets within a distance \( r < \eta R_{\text{vir}} \) of the center of a galaxy, where \( R_{\text{vir}} \) is the virial radius of the galaxy and \( \eta \) is an accretion parameter between 0 and 1. If we use \( \eta = 1 \), particles could be acccreted as soon as they enter the virial radius. With \( \eta < 1 \), particles will have to penetrate the galaxy to a certain depth before they can be accreted. We treat \( \eta \) as a free parameter, and experiment with different values to determine the appropriate one.

To determine if there is accretion, we use the same approach as for mergers: the energy \( E_{ij} \) of the system is calculated using Equation (4), where \( i \) and \( j \) now stand for the galaxy and the particle, respectively (obviously, \( U_j = 0 \)). If \( E_{ij} < 0 \), the particle is accreted. The mass of the particle is added to the mass \( M_{\text{vir}} \) of the galaxy, and the virial radius \( R_{\text{vir}} \) is recalculated using Equation (11). The internal variables are updated as follows: if the accreted particle is a dark matter particle, its mass is added to \( M_{\text{DM}} \). If the accreted particle is a gas particle, we distinguish between two accretion modes, hot and cold. In the hot accretion mode, the gas is accreted by the halo; in the cold accretion mode, the accreted gas reaches the disk and is added to the cold component. There is no gas accretion onto the hot component since that component is confined inside interstellar bubbles. The mass of the accreted gas particle is added to \( M_{\text{halo}} \) in the hot accretion mode, and to \( M_{\text{cold}} \) in the cold accretion mode (see CMD15 for details).

Since accretion removes gas particles located within a radius of \( \eta R_{\text{vir}} \), it essentially creates a cavity around each galaxy particle. This can lead to a numerical problem that we refer to as “the doughnut hole effect.” An SPH particle located at the surface of the cavity will feel a pressure force from neighbors located outside the cavity, but no pressure force coming from inside the cavity, since there are no neighbors there. The net force will push gas particles inside the cavity, where they will be accreted, and eventually, this could lead to runaway accretion of gas. This is a problem that can potentially plague any SPH code that uses sink particles. In their simulations of fragmentation of molecular clouds, Bate et al. (1995), Bromm et al. (2002), and Martel et al. (2006) dealt with this problem by treating sink particles as SPH particles and including them in the SPH equations. We tried this approach, by treating each galaxy particle as an SPH particle with a mass \( M_{\text{halo}} \). This prevented runaway accretion, but the effect was so strong for high-mass galaxies that accretion was totally suppressed. We found a simple fix to this problem: the SPH algorithm adjusts the smoothing length of each particle such that it roughly keeps a fixed number of neighbors \( N_{\text{neigh}} \). An SPH particle located near the surface of a cavity will be “missing” about \( N_{\text{neigh}}/2 \) neighbors, the ones that would have been on the cavity side if they had not been accreted. To compensate for the absence of these neighbors, we treat the galaxy particles as SPH particles, but with a mass \( m = (N_{\text{neigh}}/2)m_{\text{gas}} \). This enters only in the calculation of pressure forces on accreting gas particles, and nowhere else in the SPH algorithm or the SAM. Tests showed that this simple approach works remarkably well.

2.8. Galactic Evolution

We have developed a multizone SAM, in which the overall shape and dimension of galaxies are considered, but not the internal structure. The model was presented in CMD15, and earlier phases of its development were presented in Côté et al. (2012, 2013). The model successfully reproduces a wide variety of observational results:

1. The observed central black hole mass–dark matter halo mass relation of galaxies, for halo masses in the range \( 11.0 < \log(M_{\text{DM}}/M_*) < 13.8 \).
2. The observed stellar mass–halo mass relation of galaxies, for halo masses in the range \( 10.0 < \log(M_{\text{DM}}/M_*) < 14.0 \).
3. The observed mass–metallicity relation of galaxies, for stellar masses in the range \( 6 < \log(M_*/M_*) < 12 \).
4. The observed ISM hydrogen mass versus stellar mass relation of galaxies, for stellar masses in the range \( 6 < \log(M_*/M_*) < 12 \).
5. Several observed elemental ratios (e.g., \([\text{Mg/Fe}], [\text{C/O}]\)) in the Milky Way).

Once galaxies are formed, the algorithm simulates their internal evolution using the SAM of CMD15. The galactic evolution module is called at the end of each basic time step, and evolves all galaxies simultaneously, updating the values of their internal variables. In this section, we give an outline of the various steps involved, and refer the reader to CMD15 for a full description of the SAM.

For each galaxy, an evolution time step is calculated using

\[
(\Delta t)_{\text{evol}} = 0.002 t_{\text{ff}} = 0.002 \frac{t_{\text{dyn}}}{\sqrt{2}}
\]

where \( t_{\text{dyn}} = R_{\text{vir}}/V_{\text{vir}} \) is the dynamical time and \( t_{\text{ff}} \) is the freefall time. If \( (\Delta t)_{\text{evol}} > (\Delta t)_{\text{basic}} \), the galaxy will be evolved for a time step \( \Delta t = (\Delta t)_{\text{basic}}. \) Otherwise, several time steps of length \( \Delta t = (\Delta t)_{\text{evol}} \) will be used to complete the basic time step. During each time step \( \Delta t \), the code performs the following steps:

1. Initialization: the code initializes the amount of energy and masses that will be ejected into the halo by galactic winds:

\[
E_{GW} = 0, \quad M_{GW} = 0.
\]

2. Cooling: an amount of mass \( \Delta M_{\text{cool}} = M_{\text{cool}} \Delta t \) is transferred from the halo to the cold component:

\[
M_{\text{halo}} \leftarrow M_{\text{halo}} - \Delta M_{\text{cool}}, \quad M_{\text{cold}} \leftarrow M_{\text{cold}} + \Delta M_{\text{cool}}.
\]
\[ r_{\text{cool}} = \left[ \frac{M_{\text{halo}} n(T_{\text{vir}}, Z)}{12\mu m H k B T_{\text{vir}} V_{\text{vir}}} \right]^{1/2}, \]

where \( V_{\text{vir}} = (GM_{\text{vir}}/R_{\text{vir}})^{1/2} \). If \( r_{\text{cool}} < R_{\text{vir}} \), we are in the slow cooling regime, and \( r_{\text{cool}} \) is given by Equation (8) of CMD15; otherwise, we are in the rapid cooling regime, and we use Equation (11) of CMD15.

3. Galaxy mass and structure: the total mass of the galaxy might have changed during the previous time step, as a result of mergers, accretion, and outflows. The code uses that mass to recalculate the scale length and scale height of the disk, and the concentration parameter of the dark matter halo (see Section 2.3.1 in CMD15).

4. Star formation: the code calculates the star formation \( M_\star \), and an amount of mass \( \Delta M_\star = M_\star \Delta t \) is transferred from the cold component to the stellar component:

\[ M_{\text{cold}} \leftarrow M_{\text{cold}} - \Delta M_\star, \]

\[ M_\star \leftarrow M_\star + \Delta M_\star. \]

The code creates a stellar population of mass \( \Delta M_\star \), initializes its metallicity, and calculates the number of interstellar bubbles that population will generate.

5. Momentum-driven galactic outflows: the code calculates the amount of mass \( \Delta M_{\text{GW,M}} \) and energy \( \Delta E_{\text{GW,M}} \) ejected into the halo by M-driven outflows, updates the masses of the cold and halo components, and the energy and mass are added to the galactic wind:

\[ M_{\text{cold}} \leftarrow M_{\text{cold}} - \Delta M_{\text{GW,M}}. \]

\[ M_{\text{halo}} \leftarrow M_{\text{halo}} + \Delta M_{\text{GW,M}}. \]

\[ E_{\text{GW}} \leftarrow E_{\text{GW}} + \Delta E_{\text{GW,M}}. \]

\[ M_{\text{GW}} \leftarrow M_{\text{GW}} + \Delta M_{\text{GW,M}}. \]

We have modified the treatment of momentum-driven outflows. Apart from the treatment of accretion (Section 2.7 above), this is the only modification we have made to the original SAM of CMD15. The mass ejected is calculated using:

\[ M_{\text{GW, M}} = \sigma_0 n_{\text{f,d}} \sigma_0 \Delta t, \]

where \( \sigma \) is the velocity dispersion of the galaxy, and \( \sigma_0 \) sets the strength of the outflows. In CMD15, we used a constant value for \( \sigma_0 \), but pointed out that it is an average over the lifetime of galaxies. The coupling between radiation and gas depends on the presence of dust, and since the dust content of galaxies increases with time, so should \( \sigma_0 \) do. With a fixed value of \( \sigma_0 \), we get too much feedback at early times and too little at late times, making it difficult to reproduce the CSFR-\( z \) relation (see Section 3.4 below). Since dust is made of carbon and silicon, whose abundances increase with time, we replaced a constant \( \sigma_0 \) by:

\[ \sigma_0 = K_0 \frac{m_{\text{C,cold}} + m_{\text{Si,cold}}}{m_{\text{H,cold}}}, \]

where \( m_{X,\text{cold}} \) is the mass of element \( X \) in the cold component, and \( K_0 \) is a constant.

6. Stellar evolution and feedback: This is a major part of the algorithm. Each stellar population drives the expansion of an interstellar bubble in which energy and metal-enriched gas are deposited. The amount of energy and the mass and composition of the gas deposited are calculated by the SAM, taking into account the contribution of stellar winds, AGBs, and type Ia and type II SNe. The stellar population is then considered “active.” The code loops over all active stellar populations, and evolves their interstellar bubbles by a step \( \Delta t \). Each bubble will go through a series of phases (adiabatic expansion, pressurized thin shell, galactic outflow, and thin cold shell). If, during the first two phases, the radius of a bubble reaches the scale height of the galaxy, an energy-driven outflow is generated. During this process, a mass \( \Delta M_\text{cold} \) of cold gas and a mass \( \Delta M_\text{hot,1} \) of hot gas are ejected into the halo, while a mass \( \Delta M_\text{hot,2} \) of hot gas is transferred to the cold gas. The energy and mass are added to the galactic wind:

\[ M_{\text{cold}} \leftarrow M_{\text{cold}} - \Delta M_{\text{cold}} + \Delta M_{\text{hot,2}}, \]

\[ M_{\text{hot}} \leftarrow M_{\text{hot}} - \Delta M_{\text{hot,1}} - \Delta M_{\text{hot,2}}, \]

\[ M_{\text{halo}} \leftarrow M_{\text{halo}} + \Delta M_{\text{cold}} + \Delta M_{\text{hot,1}}, \]

\[ E_{\text{GW}} \leftarrow E_{\text{GW}} + \Delta E_{\text{cold}} + \Delta E_{\text{hot}}, \]

\[ M_{\text{GW}} \leftarrow M_{\text{GW}} + \Delta M_{\text{cold}} + \Delta M_{\text{hot,1}}. \]

The stellar population becomes “inactive” when either the galactic wind is generated, or the interior of the bubble entirely cools down, stopping the expansion. The code then loops over all inactive populations, calculates the mass loss by stellar winds and SNe, and adds it to the cold component.

7. Black hole accretion: the algorithm considers two accretion modes: quasar mode and radio mode. The code calculates the accreted masses \( \Delta M_Q \) and \( \Delta M_R \), and updates the masses accordingly:

\[ M_{\text{cold}} \leftarrow M_{\text{cold}} - \Delta M_Q, \]

\[ M_{\text{halo}} \leftarrow M_{\text{halo}} - \Delta M_R, \]

\[ M_{\text{BH}} \leftarrow M_{\text{BH}} + \Delta M_Q + \Delta M_R. \]

8. Halo feedback and metallicity update: using the values of \( E_{\text{GW}} \) and \( M_{\text{GW}} \), the code calculates the energy \( E_{\text{IGM}} \) and mass \( M_{\text{IGM}} \) that will escape the halo and get deposited onto the IGM (see Section 2.3 above), and updates the halo mass:

\[ M_{\text{halo}} \leftarrow M_{\text{halo}} - M_{\text{IGM}}. \]

The values of \( E_{\text{IGM}} \) and \( M_{\text{IGM}} \) are passed to the cosmological module, which handles the evolution of the IGM.

Note that every time the mass of a baryonic component \( (M_{\text{cold}}, M_{\text{hot}}, M_{\star}, M_{\text{halo}}) \) is updated, the mass of every element from \( Z = 1 \) to \( Z = 31 \) is updated separately.

2.9. Feedback and Metal Deposition into the IGM

Galactic outflows (transfer of mass and energy from the galaxy to the halo) are handled by the SAM. Halo outflows (transfer of mass and energy from the halo to the IGM) are handled by the cosmological module, and therefore are not part of the SAM, which only provides the values of \( E_{\text{IGM}} \) and \( M_{\text{IGM}} \), as well as the chemical composition of the ejected gas. Our approach to feedback is fairly standard, and has been used in several algorithms in the past (see, e.g., Navarro & White 1993; Mosconi et al. 2001; Kawata & Gibson 2003; Scannapieco et al. 2005; Stinson et al. 2006; Dorier & Dalla Vecchia 2012). The code identifies the gas particles located near each galaxy,
and deposits a fraction of the mass $M_{\text{IGM}}$ and energy $E_{\text{IGM}}$ onto those particles. The chemical abundances of the gas particles are updated. Since galaxy particles are treated as SPH particles (see Section 2.7 above), each galaxy particle possesses a smoothing kernel, which is used to weight the fraction of mass and energy each gas particles receives. For the simulations presented in this paper, we used thermal energy feedback: the energy deposited onto gas particles goes entirely into thermal energy. To prevent that energy from being immediately radiated away, we turn off radiative cooling for those gas particles, for a time interval $t_{\text{nocool}}$.

### 3. Observational Constraints

We have selected a set of five critical observational constraints that the algorithm must be able to reproduce to be considered satisfactory. These constraints are particularly useful for tuning the free parameters of the model.

#### 3.1. Halo Mass Function

The halo mass function $\phi(M_h, z)$ is defined as the number of halos per unit mass per unit comoving volume at redshift $z$. This is not really an “observational” result since the mass of dark matter halos cannot be determined precisely from observations. The halo mass function can be determined analytically, or can be calculated using high-resolution $N$-body simulations, and the agreement between the two approaches can be quite remarkable. We use the online calculator **HMFcalc** (Murray et al. 2013), which combines the two approaches, to estimate $\phi(M_h, z)$. The results are shown in Figure 3, at six different redshifts, from $z = 0$ (top line) to $z = 6$ (bottom line). We note that $\phi(M_h, z)$ increases with decreasing redshift at all masses. Reproducing the halo mass function is the first constraint that the algorithm must satisfy.

#### 3.2. Gas Mass–Stellar Mass Relation

In recent years, several studies have estimated the gas mass–stellar mass relation from measurement of the mass of neutral hydrogen, $M_{\text{HI}}$, first in the range $\log M_h > 9$ (Catinella et al. 2010; Parkash et al. 2018), and later at lower masses, down to $\log M_h = 6.60$ (Jaiswal & Omar 2020; Naluminsa et al. 2021). These observational results are summarized in Figure 4. Overall, all of these studies are fairly consistent, both at low and high stellar masses. The exception is the galaxy IC 10 (also called UGC 192), an outlier in the sample used by Naluminsa et al. (2021), with a low gas mass $M_{\text{HI}} = 2.5 \times 10^7 M_\odot$. However, Namumba et al. (2019) reported a value of $M_{\text{HI}} = 7.8 \times 10^7 M_\odot$ for that galaxy (open red square in Figure 4).

Huang et al. (2012) and Maddox et al. (2015) published measurements in the range $\log M_h = 6.8 – 11.2$, from the ALFALFA survey. Compared to all other studies, their estimates of $M_{\text{gas}}$ are significantly higher (gray symbols and gray line in Figure 4), at all stellar masses. Parkash et al. (2018) argue that the ALFALFA survey is HI-selected, which creates a bias toward HI-rich galaxies, and against massive elliptical galaxies with little gas content.

Reproducing the gas mass–stellar mass relation is the second constraint that the algorithm must satisfy. Because our simulations do not suffer from any selection effect, we will not consider the ALFALFA results when comparing the simulations to observations.

#### 3.3. Stellar Mass–Halo Mass Relation

The relation between the stellar mass of central galaxies and the mass of their host dark matter halos can be determined using abundance matching (see, e.g., Behroozi et al. 2010). Numerous studies based on this technique have been conducted in the past (Yang et al. 2009; Behroozi et al. 2010; Guo et al. 2010; Moster et al. 2010; Behroozi et al. 2013; Moster et al. 2013). These results are shown in Figure 5. They are remarkably consistent over four orders of magnitude in halo mass, and together reveal the inefficiency of star formation in galaxies: The most star-rich galaxies have a halo of mass $M_h = 10^{12} M_\odot$, and contain a stellar mass $M_*=3 \times 10^{10} M_\odot$, for a ratio $M_*/M_h=0.03$, while the universal baryon/dark matter ratio $\Omega_b/\Omega_m=0.20$. The ratio $M_*/M_h$ drops at lower and higher halo masses, which is presumably caused by stellar feedback (low end) and AGN feedback (high end). Reproducing the stellar mass–halo mass relation is the third constraint that the algorithm must satisfy.

#### 3.4. Cosmic Star Formation Rate

The cosmic SFR (CSFR) is the SFR per unit comoving volume. Numerous studies have shown that the CSFR, $\psi(z)$, increases with redshift and reaches a peak of the order of $\psi = 0.1–0.2 M_\odot \text{yr}^{-1} \text{Mpc}^{-3}$ at redshift $z \sim 2$, and then drops down to $\psi = 0.01–0.02 M_\odot \text{yr}^{-1} \text{Mpc}^{-3}$ at $z = 0$. These results were summarized in Madau & Dickinson (2014), who...
proposed the following fitting formula:

$$\psi(z) = \frac{a(1+z)^b}{1 + [(1+z)/c]^d}.$$  \hspace{1cm} (35)

The parameters are listed in Table 1.

As Madau & Dickinson (2014) pointed out, there is, however, an inconsistency between the CSFR and the \(M_*-M_h\) relation, which are determined separately using different methods. To quote these authors, “Our model predicts an SMD\(^{\text{a}}\) that is somewhat high \((\sim 0.2 \text{ dex on average}, \text{ or } 60\%)\) compared with many, but not all, of the data at \(0 < z \lesssim 3\).”

The discrepancy between the CSFR and the SMD can be explained by the uncertainty in converting observable tracers like H\(\alpha\), far-UV, and thermal infrared luminosities to the actual number of recently formed stars required to produce them. Revised studies that include observations of nearby massive stars and improved stellar population synthesis models (Yu & Wang 2016; Wilkins et al. 2019) predict a stronger and harder stellar ultraviolet output for a given stellar mass in galaxies, therefore resulting in lower values of the CSFR to account for the observed properties. The fits proposed by these authors are also based on Equation (35), and their parameters are listed in Table 1. Both fits predict a peak value of the CSFR, at \(z \sim 2\), that is significantly lower than the value predicted by Madau & Dickinson (2014).

Figure 6 shows the observational data used by Madau & Dickinson (2014; symbols) and their fit (blue line), as well as the fits provided by Yu & Wang (2016; red line) and Wilkins et al. (2019; cyan line). Reproducing the cosmic stellar mass function is the fourth constraint that the algorithm must satisfy, keeping in mind the actual uncertainties on the form of the function.

\(\text{a}\) Stellar mass density.

### Table 1

| Fit                  | \(a\) | \(b\) | \(c\) | \(d\) |
|----------------------|-------|-------|-------|-------|
| Madau & Dickinson (2014) | 0.015 | 2.7   | 2.9   | 5.6   |
| Yu & Wang (2016)      | 0.023 | 1.66  | 2.81  | 3.67  |
| Wilkins et al. (2019) | 0.0103| 2.48  | 3.10  | 6.26  |

Figure 5. Stellar mass vs. halo mass. Symbols show observational measurements from Yang et al. (2009; gray); Behroozi et al. (2010; magenta); Guo et al. (2010; green); Moster et al. (2010; red); Behroozi et al. (2013; orange); and Moster et al. (2013; blue). Yellow circle: Milky Way (Cautun et al. 2020). Dotted line: relation \(M_*/M_h = \Omega_m/\Omega_\Lambda\).

3.5. Central Black Hole Mass–Halo Mass Relation

While it is well established that more massive halos tend to host more massive central black holes, the exact form of this relation remains uncertain, and its physical origin is still debated. Results from various studies are shown in Figure 7. In spite of large differences, there is some consistency among those results. In particular, the results of Baes et al. (2003) and Treu & Richtler (2012) are consistent over the range \(\log M_h = 11.4–12.8\), and are consistent with about half of the results of Davis et al. (2019). These results suggest a power law \(M_{BH} \propto M_{DM}^{0.55}\) in that mass range. At larger masses, there is moderate agreement between the results of Bandara et al. (2009) and Dutton et al. (2010), suggesting a power law with a shallower slope. Reproducing the black hole mass–halo mass relation is the fifth constraint that the algorithm must satisfy. Because of the large uncertainty in the actual relation, this is the least constraining of the five constraints.

4. The Simulations

We consider a concordance \(\Lambda\) cold dark matter (CDM) model, with density parameter \(\Omega_m = 0.275\), baryon density parameter \(\Omega_b = 0.0458\), cosmological constant \(\Lambda_0 = 0.725\), Hubble constant \(h = 0.702\), primordial tilt \(n = 0.968\), and rms density fluctuations \(\sigma_8 = 0.816\), in accordance with the results of WMAP7. We have performed a series of simulations, in order to test the algorithm and tune the numerical parameters. Two of these simulations (Runs A and B) are presented in Section 5 below. In all cases, we simulate the evolution of a comoving volume of size 20 Mpc, using 128\(^3\) equal-mass dark matter particles and 128\(^3\) equal-mass gas particles, and a 256\(^3\) PM grid. The length resolution is 23.4 kpc. The total mass in the computational volume is \(3.009 \times 10^{14} M_\odot\). The masses of dark matter and gas particles are \(m_{\text{dm}} = 1.196 \times 10^8 M_\odot\) and \(m_g = 2.390 \times 10^4 M_\odot\), respectively. The corresponding ratio...
$m_d/m_{dm}$ is equal to the universal ratio $\Omega_{b0}/(\Omega_{b0} - \Omega_{c0})$, which enables us to use equal numbers for $N_{dm}^{GF}$ and $N_g^{GF}$. We set these numbers equal to 50. Therefore, galaxies form with an initial mass given by:

$$M_{\ast}^{GF} = N_{dm}^{GF}m_{dm} + N_g^{GF}m_g = 7.173 \times 10^9 M_\odot. \quad (36)$$

### 4.1. Initial Conditions

We generate initial conditions using the code COSMICS (Bertschinger 1995). This code generates the initial positions and velocities of the particles in order to reproduce a Gaussian random density field with a prescribed density fluctuation power spectrum. The particles are laid down on a cubic grid, one dark matter particle and one gas particle per grid point, and COSMICS calculates their displacements and initial velocities. We then calculate the gas density at the location of each gas particle, and use it to initialize the internal energy and SPH smoothing length. The simulations start at redshift $z = 45$, long before any significant IGM enrichment has taken place. Hence, we give to each gas particle a primordial composition with hydrogen, helium, and lithium abundances of 0.7552, 0.2448, and 4.648 $\times 10^{-10}$, respectively (Cyburt et al. 2016; Valerdi et al. 2021).

### 4.2. Performance

The current version of the code is written in FORTRAN, and parallelized with OpenMP. It therefore runs on systems with shared memory. All simulations were performed on a local computer named Stargate. This computer has 16 processors 6 Intel(R) Xeon(R) CPU E5-2630 v3 @ 2.40 GHz, and 132 GB of RAM. For the actual simulations, we actually used 14 processors. For comparison, one simulation was rerun on two separate computers that are part of the Compute Canada network. The first one, named Beluga, has 974 nodes with 40 processors 2 x Intel Gold 6148 Skylake @ 2.40 GHz and a minimum of 192 GB of RAM per node. The second one, named Narval, has 1142 nodes with 64 processors 2 x AMD Rome 7532 @ 2.40 GHz and a minimum of 256 GB of RAM per node. Table 2 shows the performance of the algorithm on these various computing platforms. It is interesting to note that the performance, measured by the product of the computing time by the number of processors, is about the same on all platforms, of the order of 70 days-processors, for three of the runs. Doubling the number of processors on Narval reduces the CPU time by a factor of 1.36.

### 5. Testing and Tuning the Algorithm

The parameters of the algorithm were tuned in order to produce the best fit to the observational constraints described in the Section 3. The optimal values we have determined are listed in Table 3, with the last column indicating where these various computing platforms. It it interesting to note that the parameters $\sigma_0$ used in the original SAM is now being calculated using Equation (25), with the value of $K_0$ given in Table 3.
When a merger takes place, the lower-mass progenitor is the one considered "destroyed." The evolution of the galaxy population is summarized in Table 4. The first galaxy forms at redshift \( z = 11.23 \) in both runs. During the course of each simulation, \( N_{\text{form}} \) galaxies of mass \( M = M_{\text{vir}}^{\text{GF}} = 7.173 \times 10^9 M_\odot \) formed by monolithic collapse. \( N_{\text{merg}} \) galaxies were eventually destroyed by mergers\(^9\), and \( N_{\text{tide}} \) galaxies were tidally destroyed by a more massive galaxy. In \( N_{\text{acc}} \) cases, the tidal fragment accreted onto the larger galaxy, while in \( N_{\text{inc}} \) cases, they were dispersed into the IGM. At the end of the simulation, the computational volume contains \( N_{\text{gal},f} \) galaxies. The most massive galaxy has a mass \( M_{\text{vir}} = M_{\text{max}} \approx 1.6 \times 10^{13} M_\odot \), while the least massive galaxy has a mass \( M_{\text{vir}} = M_{\text{min}} \approx 6 \times 10^9 M_\odot \). Some galaxies end up with a mass lower than the mass \( M_{\text{vir}} \) they started up with, because of mass loss through halo outflows. Overall, the simulations tracked the formation and evolution of more than 7000 galaxies from redshift \( z \approx 11 \) to the present, resulting in a final volume containing over 1000 galaxies, with a mass range covering 3.4 orders of magnitude.

### 5.1. Large-scale Structures

Figure 8 shows the computational volume inside a slice of thickness 3.0 Mpc at \( z = 0 \). The black and red dots represent dark matter and galaxies, respectively. The gas particles are not shown, but follow closely the dark matter. Note that all red dots have the same size, and do not reflect the actual virial radius \( R_{\text{vir}} \) of the galaxies.

We see the usual filamentary structure typical of CDM universes. The galaxies have formed at the densest locations, inside knots and the filaments connecting them. Low-density filaments do not contain any galaxies, because the gas density in these filaments never reached the threshold for galaxy formation.

We used a standard friends-of-friends algorithm to identify clusters in the simulation. Figure 9 shows a massive system found at \( z = 0 \). With 25 galaxy members and a total mass of \( 6.29 \times 10^{12} M_\odot \), it is more a group than a cluster. In any case, we did not expect to find any massive cluster in a \((20 \text{ Mpc})^3\) volume. This group contains a massive galaxy, with virial mass \( M_{\text{vir}} = 8.90 \times 10^{11} M_\odot \) (comparable to the Milky Way). The next most massive galaxies have virial masses of \( 1.24 \times 10^{12} M_\odot, 0.90 \times 10^{11} M_\odot, 0.89 \times 10^{11} M_\odot, \) and \( 0.80 \times 10^{11} M_\odot \), respectively. Fourteen galaxies have a virial mass below \( 1.2 \times 10^{10} M_\odot \), which would classify them as dwarfs.

This group is composed of three separate subgroups in the process of merging. The bottom subgroup, centered at \( x = -0.7 \text{ Mpc}, y = -1.2 \text{ Mpc} \), hosts the two most massive galaxies. We note that the virial radius of the merger remnants depends on both the virial mass and the relative velocity of the progenitors (see Equation (11)). The green dots show the galaxies that have been destroyed by tides. The algorithm no longer evolves those galaxies, but follows their trajectories in order to determine the location of

### Table 4

| Run | \( z_f \) | \( N_{\text{form}} \) | \( N_{\text{merg}} \) | \( N_{\text{tide}} \) | \( N_{\text{acc}} \) | \( N_{\text{inc}} \) | \( N_{\text{gal},f} \) | \( M_{\text{max}}[10^9 M_\odot] \) | \( M_{\text{min}}[10^9 M_\odot] \) |
|-----|---|---|---|---|---|---|---|---|---|
| A   | 11.23 | 8113 | 6125 | 771 | 720 | 51 | 1217 | 6.041 | 6.043 |
| B   | 11.23 | 7944 | 5975 | 743 | 695 | 48 | 1127 | 6.043 | 6.040 |

---

\(^9\) When a merger takes place, the lower-mass progenitor is the one considered "destroyed."
the intracluster light. Interestingly, all of these galaxies (seven of them) are located in the upper subgroup. The massive galaxies present in the other subgroups have reaccreted the fragments of galaxies destroyed by tides. We note that intracluster stars are usually found in massive clusters, and not in groups.

### 5.2. Halo Mass Function

Figure 10 shows the mass function of halos at various redshifts (histograms). For comparison, we show the theoretical mass function, calculated using HMFcalc. At $z = 0$ the simulated halo mass function reproduces the theoretical one very well over 3.5 orders of magnitude in mass. The slight deficit in the lowest-mass bin, just above $M_{\text{vir, min}}$, is an effect of resolution. Normally, this bin would contain halos that form with an initial mass $M < M_{\text{vir, min}}$, but end up reaching a mass $M > M_{\text{vir, min}}$ by accretion, but halos do not form with a mass $M < M_{\text{vir, min}}$ in the simulations. The high-mass cutoffs are a consequence of the limited computational volume; in each histogram, the rightmost bin contains only one or two galaxies.

The halo mass functions is also well reproduced at higher redshifts, up to $z = 3.99$. At $z = 5.91$, the simulation slightly underestimates the halo mass function, which again can be ascribed to limited resolution. At that redshift, the system contains only 221 galaxies.

### 5.3. Stellar and Gas Content of Galaxies

Figure 11 shows the gas mass–stellar mass relation for our simulated galaxies (black circles), together with the observational measurements shown in Figure 4. The gas mass of the simulated galaxies is obtained by adding the masses of the cold and hot components. It therefore includes all elements, though the total gas mass is dominated by neutral hydrogen.

Except for a few outliers with a very low gas content, our results are fully consistent with observations over a mass range $\log M_{\text{b}} = 7.4$–11.5. The most extreme outlier is the irregular galaxy Mrk 475, from the sample of Jaiswal & Omar (2020), which has a very low gas content. That galaxy has an elongated HI feature, resulting in an offset between the HI center and the optical center. Another major outlier is the blue compact dwarf galaxy UGC 192 (a.k.a., IC 10), from the sample of Namumba et al. (2019), which also has a low gas content. These authors point out that UGC 192 has a significant amount of neutral gas outside of the main disk, which was not included in their calculation of $M_{\text{HI}}$. Namumba et al. (2019) measured $M_{\text{HI}}$ for the same galaxy, including the external region, and came up with a value 0.44 dex higher (open red square). While that value is still low compared to our simulated galaxies, it is much less of an outlier.

The plateau seen at $\log M_{\text{gas}} \approx 7.8$ is a result of the star formation criterion used by the algorithm. The SFR is calculated using:

$$\frac{dM_{\text{s}}}{dt} = f_\text{s} \left( \frac{M_{\text{cold}} - M_{\text{crit}}}{r_\text{h}} \right)$$

(37)

$$M_{\text{crit}} = M_{\text{crit,0}} \left( \frac{V_{\text{vir}}}{200 \text{ km s}^{-1}} \right) \left( \frac{R_{\text{disk}}}{10 \text{ kpc}} \right) M_\odot$$

(38)

(see Section 2.4 of CMD15). If $M_{\text{cold}}$ is less than the critical mass $M_{\text{crit}}$, we set $dM_{\text{s}}/dt = 0$. This limits the amount of star formation taking place in low-mass galaxies.

### 5.4. Stellar Mass–Halo Mass Relation and Star Formation History

The $M_{\text{s}}$–$M_{\text{b}}$ relation derived from the simulations is shown in Figure 12, with the left and right panels corresponding to Run A and Run B, respectively. The top panels show the $M_{\text{s}}$ versus $M_{\text{b}}$ relation at $z = 0$ determined from the simulations. Each black dot represents one galaxy. The red dots and error bars show the average and standard deviation on $M_{\text{s}}$ in bins of constant width in $\log M_{\text{b}}$. In the bottom panels, we compare...
Figure 12. Stellar mass vs. halo mass for Run A (left panels) and Run B (right panels). Top panels: results of the simulations (black dots), and average and standard deviation in mass bins (red dots and error bars). Bottom panels: comparison with observations. Color symbols and error bars have the same meaning as in Figure 5. Solid black lines show the mean and the standard deviation of the simulated results (these lines go through the red symbols and tips of the error bars in the top panels). Dotted line in all panels: relation $M_*/M_h = 3\Omega_b/\Omega_c$. Insets show the transition region, around log $M_h = 11.4$.

Figure 13 shows the CSFR versus redshift and lookback time for Run A (solid black line) and Run B (dashed black line). The various symbols and color lines represent the various measurements and empirical fits also shown in Figure 6. In both runs, the CSFR peaks at redshift $z \sim 2$, in accordance with most observations. At higher redshift, the agreement with observations is reasonable, considering the small number of star-forming galaxies in the simulations at that epoch. At redshifts $z < 2$, Run A predicts a CSFR that is too low at $0.5 < z < 2$, and too high at $z < 0.5$, but remains within the error bars of most observations. Run B, which has a stronger feedback, produces a CSFR that is lower at all redshifts compared to Run A. For this run, the fit to observations is very good at $z < 0.5$, but worse than Run A at redshifts $0.5 < z < 2$. Interestingly, both runs almost perfectly reproduce the fit of Yu & Wang (2016) at $z < 1$ (red line), while at $1 < z < 2$, both runs are in good agreement with the fit of Wilkins et al. (2019; cyan line).

The results shown in Figures 12 and 13 illustrate the discrepancy that was discussed in Section 3.4. Run A provides a fairly good fit to the CSFR determined by Madau & Dickinson (2014; blue line), but overestimates the stellar mass, while Run B underestimates the CSFR, but correctly predicts the stellar mass. Except for the overall agreement that the CSFR peaks at $z \sim 2$, there remains a great deal of uncertainty in the actual CSFR history, as the blue, cyan, and red lines in Figures 6 and 13 show. By contrast, the stellar mass–halo mass relation is well constrained by observations. For this reason, we consider Run B as our “best run.”

5.5. Black Hole Growth

Figure 14 shows the central black hole mass $M_{BH}$ versus dark matter halo mass $M_{DM}$ at $z = 0$ (black dots), and for comparison, the observational measurements also shown in Figure 7. In spite of the large dispersion, both in the simulation results and the observations, the match is excellent in the interval log $M_{DM} = 11.4-13.0$, if we ignore the high $M_{BH}$...
values of Davis et al. 2019 (cyan), which are clear outliers. We have no observations at $M_{\text{log}} < 11.4$, and no simulated galaxies at $M_{\text{log}} > 13.0$. Still, a simple extrapolation would suggest a good match, though this remains to be shown. Note that the absence of simulated galaxies at large masses results from our choice of computational volume. This issue is discussed in Section 6.2 below.

We found that the results are not very sensitive to the choice of seed mass $M_{\text{seed}}$, as long as that value is sufficiently small. In Runs A and B, the final black hole masses are three to 100,000 times the seed mass.

5.6. Metal Enrichment of the IGM

Figure 15 shows the distribution of oxygen in the IGM at four different redshifts, for Run A. We used the same slice of thickness 2.5 Mpc as in Figure 8. The colors show the ratio of column densities $\Sigma_O / \Sigma_H$. At $z = 4.200$, the metals are concentrated near the few galaxies that have already formed, and trace the densest regions. At $z = 1.803$, more galaxies have formed, and the metals trace the filamentary structure, but enrichment does not affect the low-density regions at that time. At $z = 0.996$, the metals are starting to spread into low-density regions, and by $z = 0$, a significant fraction of the volume has been enriched in metals.

Figure 16 shows the distribution of four elements (carbon, oxygen, magnesium, and iron) in the IGM at $z = 0$, for Run A. The colors show the ratio of column densities $\Sigma_X / \Sigma_H$. At all redshifts, magnesium and iron have comparable abundances, while carbon and oxygen abundances are higher by factors of $\sim 3.8$ and $\sim 14.2$, respectively. In the simulations, the deposition of metals onto IGM gas particles does not distinguish between various elements (see Section 2.9). As a result, the regions enriched by IGM outflows are the same for all elements. However, the distributions of abundance ratio are not uniform, since they depend on the timing between the various sources of metals (stellar winds from low-mass and massive stars, type Ia SNe, and type II SNe). To illustrate this, we plot in Figure 17 the evolution of the oxygen-to-iron ratio, calculated using:

$$[O/Fe] = \log \frac{\Sigma_O}{\Sigma_{Fe}} - \log \left( \frac{n_O}{n_{Fe}} \right)_\odot, \quad (39)$$

where $n_X$ is the abundance by mass of element $X$. This illustrates the relative contribution of type II SNe, which produce alpha-elements like oxygen, and type Ia SNe, which produce most of the iron. Because of the difference between the lifetimes of their progenitors, type II SNe are the first ones to contribute, and at $z = 4.2$, enrichment in alpha-elements dominates. Eventually, iron enrichment by type Ia SNe catches up. At $z = 0.996$, we find small, isolated regions that are entirely oxygen-rich. These regions were enriched by low-mass galaxies that formed late, and are therefore in the early stages of their evolution. In other metal-enriched regions, there is a trend of iron-rich gas to be located in the densest central regions, and oxygen-rich gas to be located near the edges. This results from oxygen being produced earlier than iron, and therefore traveling larger distances. However, mixing and re-accretion of enriched gas tends to limit this effect.

Figure 18 shows the evolution of the mass fractions of each element in the IGM, for Run A. The simulation started with a primordial composition (hydrogen, helium, and lithium, and no heavier metals). Galaxy formation started at $z = 11.23$. Accounting for the time needed for the halo gas to cool and form stars, for the stars to generate metal-rich bubbles, and for
Figure 15. Evolution of the distribution of oxygen in the IGM, for Run A. The panels show, at various times, the ratio of column densities $\Sigma_{O}/\Sigma_{H}$ inside a slice of thickness 2.5 Mpc, corresponding to the slice shown in Figure 8. Redshifts are indicated.

Figure 16. Distribution of elements in the IGM at $z = 0$, for Run A. The panels show, for various elements, the ratio of column densities $\Sigma_{X}/\Sigma_{H}$ inside a slice of thickness 2.5 Mpc, corresponding to the slice shown in Figure 8. Elements are indicated.

Figure 17. Evolution of the oxygen-to-iron ratio in the IGM, for Run A. The panels show, at various times, the abundance ratio $[O/Fe]$ inside a slice of thickness 2.5 Mpc, corresponding to the slice shown in Figure 8. Redshifts are indicated.

Figure 18. Evolution of the abundances of elements in the IGM, vs. redshift, for Run A. Elements are indicated. Primordial elements are indicated in blue.
these bubbles to produce galactic and IGM outflows, IGM enrichment starts at $z = 9.7$. The abundances of all elements increase rapidly until $z \sim 6.5$, when two competing effects become significant: re-accretion of enriched gas by existing galaxies, and removal of enriched gas that gets incorporated into newly formed galaxies. This results in a plateau that lasts until $z \sim 4.5$, when abundances start increasing again. We note a change of slope $a \sim 2$, which corresponds to the epoch when the CSFR reaches a maximum and starts to decrease. The final abundances of nonprimordial elements are below the solar abundances by about 1 dex, simply because most metals produced by stars do not escape into the IGM.

6. Open Issues and Further Developments

The current version of the algorithm has successfully passed all critical tests that we have imposed. Still, some aspects of the algorithm could be improved, and additional physical processes could be included. We defer this to future work. In this section, we describe the various issues we intend to address in the future, in decreasing order of priority.

6.1. Galaxy Formation at the Resolution Limit

Figure 12 shows that the algorithm reproduces the observed stellar mass–halo mass relation very well, except at the lowest and highest halo masses, where the stellar mass is, on average, a little high. In the simulations presented in this paper, the lowest-mass galaxies have a total mass $M_{\text{vir}}^{\text{GF}} = 7.173 \times 10^9 M_\odot$ (Equation (36)), and form by monolithic collapse. More massive galaxies form by mergers and accretion. While galaxies of mass $M_{\text{vir}}^{\text{GF}}$ are resolved, the actual processes that forms these galaxies is not. In the real universe, these galaxies form by mergers of even less-massive galaxies, which are not resolved by the algorithm. We note that this issue is not specific to our algorithm: any numerical algorithm with a subgrid treatment of galaxy formation will have the same issue, unless that algorithm has the power to resolve galaxies down to much smaller virial masses, of the order of $10^7 M_\odot$, or even less.

Monolithic collapse results in low-mass galaxies with a large initial gas fraction, comparable to the universal gas fraction $\Omega_G / \Omega_0$. Even though these galaxies will eject most of that gas in the form of IGM outflows, the large initial gas fraction can result in a strong starburst, possibly explaining the large stellar masses (and large dispersion) seen in the lowest mass bins in Figure 12. One easy solution is to claim that the mass resolution of the algorithm is not $M_{\text{vir}}^{\text{GF}}$, but rather $4 M_{\text{vir}}^{\text{GF}}$, since after all, these galaxies form by mergers and accretion in the simulation, and satisfy the observational constraints. Another approach could be to modify the subgrid treatment of galaxy formation such that, when galaxies of mass $M_{\text{vir}}^{\text{GF}}$ form, the algorithm would adjust their gas and stellar content. This would require a significant amount of code development and testing, and we make it our top priority.

6.2. Satellite Galaxies

Each galaxy particle in the algorithm represents one single galaxy. If a massive galaxy is surrounded by several low-mass satellite galaxies, these galaxies must be represented by different galaxy particles. The algorithm does not allow for a single galaxy particle to represent a massive galaxy and its satellite together.

This is all fine as long as massive galaxies do not swallow their satellites. The treatment of galaxy mergers described in Section 2.5 is very simple and reproduces remarkably well the halo mass function $\phi(M_h, z)$ at all masses up to at least $\log M_h = 13.6$. However, simulations in larger computational volumes (40–100 Mpc in size) reveal that the algorithm might suffer from an overmerging problem. Massive galaxies tend to split with their satellites. This might explain the small excess in stellar mass seen in the last bin in Figure 12, though that mass bin contains only two galaxies. The simple merger criterion described in Section 2.5 might just be too simple, and could benefit from some refinements.

6.3. Interstellar Bubbles

Our SAM assumes that each stellar population generates an expanding interstellar bubble, and when the bubble radius reaches the scale height of the gaseous disk, a galactic wind is generated. Each bubble is centered on the plane of symmetry of the disk (galactic height $h = 0$). This would become a problem in massive galaxies with high gas content. In that case, assuming that the bubbles are centered on the plane of symmetry might be not only incorrect, but troublesome. We might have a situation in which bubbles cannot burst out of the disk because it is too thick (note that more massive galaxies produce more bubbles, not larger bubbles). This could be solved by randomly assigning a galactic height to each population. In any case, the spatial distribution of bubbles within the disk is an issue that will need to be addressed before we simulate larger volumes containing more massive galaxies.

6.4. Starbursts and Halo Stars

Gas-dynamical simulations have revealed that gas-rich galaxy mergers, which are common at high redshift, often lead to a strong starburst (Robertson et al. 2006; Brook et al. 2007; Richard et al. 2010). Eventually, the stars formed during that starburst either end up in the halo, or settle into a thick disk whose scale height exceeds the scale height of the thin disk where the gaseous components and the remainder of the stellar component reside. To account for this, we would need to divide the stellar component into two separate components: the thin disk component already in the model, and a thick disk/halo component of mass $M_{\text{thick}}$. Equation (6) would be replaced by:

$$M_{\text{halo},k} = M_{\text{halo}, i} + M_{\text{halo}, j} - M_{\text{burst}},$$

$$M_{\text{thick},k} = M_{\text{thick}, i} + M_{\text{thick}, j} + M_{\text{burst}},$$

where $M_{\text{burst}}$ is the total mass of gas consumed during the starburst. Because of their spatial distribution, these stars would have the ability to deposit energy feedback and metals directly into the gaseous halo, while the thin disk stars deposit energy feedback and metals into the hot and cold galactic components.

6.5. Ram-pressure Stripping and Pressure Confinement

Galaxies moving into a hot IGM, or into the gaseous halo of a more massive galaxy, are subject to an external pressure, which can greatly affect their evolution. Calculating the external pressure acting on galaxies during the simulation is fairly trivial, since the algorithm “knows,” at every time step, the position, velocity, mass, and pressure of every SPH gas particle, and the position, velocity, and halo gas mass and pressure of every galaxy. Determining the effect of that
external pressure on the evolution of the galaxy is a complicated issue. Recent simulations (Williamson & Martel 2018; Du et al. 2019; Hausammann et al. 2019) reveal the existence of competing effects: on one hand, ram-pressure stripping can remove a substantial amount of gas, possibly quenching star formation or at least reducing the SFR. In the other hand, pressure confinement can shut down IGM outflows, enabling low-mass galaxies to retain their metal-enriched gas. Furthermore, the external pressure can compress the central gas of the galaxy, resulting in episodes of starbursts (Du et al. 2019). Before we can include the treatment of external pressure into our SAM, we would first need to acquire a better understanding of the various processes involved, by performing a series of high-resolution simulations of individual galaxies in a high-pressure environment, possibly using the wind-tunnel approach of Williamson & Martel (2018). We defer this to further work.

### 6.6. Mergers and Intracluster Light

When two galaxy particles merge, the approach described in Section 2.5 results in a complete fusion, with 100% of the stellar, gaseous, and dark matter content of the progenitors ending up inside the merger remnant. However, when real galaxies merge, a fraction of the content might be dispersed into the IGM. Numerical simulations show that during such encounters, 10%–30% of the mass becomes unbound, depending on the geometry of the encounter (Stanghellini et al. 2006; Wu & Jiang 2009). This process, together with tidal disruption, could explain the origin of the intracluster light.

To implement this into our algorithm, we could use an approach similar to the one described in Section 2.6, by treating the unbound matter as tidal fragments, the difference being that the fragment would not represent the entire low-mass progenitor, but rather a fraction of it. The real issue would be to determine the content of the fragment, and its velocity. As for the case of ram-pressure stripping, this would require a series of preliminary simulations similar to the ones presented in Stanghellini et al. (2006) and Wu & Jiang (2009). We note that these simulations did not include gas dynamics, and therefore do not provide information on the amount of gas that becomes unbound during the merger.

### 6.7. Dust Content

Our treatment of momentum-driven outflows (Equations (24) and (25)) is the only modification we have made to the original SAM of CMD15. The original SAM uses a constant, average value of $\sigma_0$, and here this resulted in too much feedback at high redshift and too little at low redshift. Using a varying $\sigma_0$ given by Equation (25) enabled us to reproduce the observed cosmic SFR at all redshifts (see Figure 13). Still, this treatment of the dust content is quite elementary, and could be refined. Ultimately, we might want to treat the dust as an additional component of mass $M_{\text{dust}}$. We would then need an equation for the time-evolution of that quantity, taking into account the various processes of dust formation and destruction.

### 6.8. Chemical Model

The current algorithm evolves the chemical composition of stars, galactic gas, and IGM, keeping track of 31 elements, from hydrogen to gallium. Adding more elements is possible but would increase the memory requirements and the size of the output dumps. A future improvement would be to update (or even reduce) the list of elements in order to only focus on elements that are relevant to our studies. Since our SAM can follow the chemical evolution of galaxies, we could use our simulations to simultaneously address the enrichment of the IGM and the abundances of stellar populations in local galaxies. In addition to C, N, O, Mg, Ne, Si, and Fe, which are particularly relevant for studies of the IGM (Tumlinson et al. 2017), our algorithm could focus on elements that can be detected via stellar spectroscopy (Abrahallima & Frebel 2018). For example, instead of following beryllium, boron, and gallium, we could follow heavier elements such as strontium, barium, and europium to include the contribution of the slow and rapid neutron-capture processes in our chemical enrichment prescription, two processes that are currently not included in our framework.

### 6.9. Optimization

The current version of the code is written in Fortran, and parallelized with OpenMP, and it is not yet optimized for speed. At the moment, this is not a problem because each simulation only takes a few days. In the simulations presented in this paper, the virial mass of galaxies covers 3.4 orders of magnitude, and as we discuss in Section 6.1, the actual dynamical range might be a little lower. If we wish to increase the dynamical range, we will have to use more particles, and optimization might become necessary. Currently, the P$^3$M/SPH part of the algorithm is optimized, but the SAM is not. There is potential for a $\sim 30\%$–$40\%$ speed-up. This would enable us to extend the dynamical range in mass of the simulations, and potentially include less-massive galaxies.

### 7. Summary and Conclusion

We have designed a new algorithm to simulate the joint evolution of the galaxy population and the IGM, for the specific purpose of studying the chemical enrichment of the IGM by galactic outflows. This new algorithm combines a standard P$^3$M/SPH gasdynamical algorithm, a semi-analytical treatment of galaxy formation, mergers, tidal disruption, and accretion, and an SAM to describe the evolution of the galaxies, and the deposition of energy and metal-enriched outflows into the IGM. In Section 1.1, we stated four specific objectives that this algorithm had to achieve.

1. Simulate a comoving cosmological volume: We present in this paper two simulations of the joint evolution of galaxies and the IGM, in a comoving volume of size (20 Mpc)$^3$, from redshift $z = 45$ to the present. During the course of the simulation, $\sim 8000$ galaxies formed, and as a result of mergers and tidal disruption, about 1200 galaxies are present in the volume at $z = 0$. The simulations track the deposition and propagation of 31 chemical species in the IGM.

2. Simulate the evolution of galaxies on the fly: This is precisely what the SAM enables us to achieve. At every step during the simulation, the entire galaxy population is evolved, taking into account accretion of gas and dark matter from the IGM, star formation, generation and propagation of interstellar bubbles, chemical enrichment, AGN accretion and feedback, and galactic outflows.
3. Reproduce observational constraints: We have selected five key observational constraints: (1) the mass function of dark matter halos, (2) the gas mass–stellar mass relation, (3) the stellar mass–halo mass relation, (4) the CSFR versus redshift relation, and (5) the central black hole mass–stellar mass relation. The two simulations presented in this paper satisfy all of these constraints, over 3.4 orders of magnitude in halo mass. The largest discrepancies occur at the low-mass end, where the algorithm tends to overestimate the stellar content of dwarf galaxies.

4. Require a modest amount of computer time: The simulations presented in this paper only took a few days to complete, using 32 processors or fewer, using a version of the code that is not fully optimized for speed yet.

The main strength of this new algorithm is its performance. With each simulation taking a few days to complete, and potentially a few hours once we start using more processors, we can perform large numbers of simulations. As we stated in the introduction, this algorithm gives us the ability to implement additional physical processes into the subgrid treatment of galaxies, perform large numerical parameter studies to test the implementation of these processes, and tune the numerical parameters. The algorithm can be used as a test bed for subgrid treatments that can be later implemented into other existing algorithms. But note that our original goal was to develop a tool that we intend to use to study the chemical enrichment of the IGM.

The work presented in this paper should be seen as a feasibility study. Our goal was to validate the approach of using an SAM for the treatment of subgrid physics in a cosmological gasdynamical algorithm. We have reached all of the objectives we had initially set, but this algorithm is still a work in progress. In Section 6, we presented, in decreasing order of priority, the next steps in code development, which will be presented in a forthcoming paper.

This research is supported by the Canada Research Chair program and NSERC. B.C. also acknowledges the support from the National Science Foundation (NSF, USA) under grant No. PHY-1430152 (JINA Center for the Evolution of the Elements). We thank the referee for a thorough review of this paper and valuable suggestions.

ORCID iDs

Benoit Côté © https://orcid.org/0000-0002-9986-8816
Laurent Drissen © https://orcid.org/0000-0003-1278-2591
Wilkins, S. M., Lovell, C. C., & Stanway, E. R. 2019, MNRAS, 490, 5359
Williamson, D., & Martel, H. 2018, ApJ, 867, 72
Wu, Y. T., & Jiang, I. G. 2009, MNRAS, 399, 628
Wyder, T. K., Treyer, M. A., Milliard, B., et al. 2005, ApJL, 619, L15
Yang, X., Mo, H. J., & van den Bosch, F. C. 2009, ApJ, 695, 900
Yoshida, N., Bromm, V., & Hernquist, L. 2004, ApJ, 605, 579
Yu, H., & Wang, F. Y. 2016, ApJ, 820, 114
Zubovas, K., & Nayakshin, S. 2014, MNRAS, 440, 2625