

G.A.S.

I. A prescription for turbulence-regulated star formation and its impact on galaxy properties

M. Cousin¹, P. Guillard^{2,3}, and M. D. Lehnert²

¹ Aix Marseille Univ., CNRS, CNES, LAM, Marseille, France
e-mail: morgane.cousin86@gmail.com

² Sorbonne Université, CNRS UMR 7095, Institut d'Astrophysique de Paris, 98 bis bd Arago, 75014 Paris, France

³ Institut Universitaire de France, Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche, 1 rue Descartes, 75231 Paris Cedex 05, France

Received 16 November 2018 / Accepted 29 January 2019

ABSTRACT

Context. Star formation in galaxies is inefficient, and understanding how star formation is regulated in galaxies is one of the most fundamental challenges of contemporary astrophysics. Radiative cooling, feedback from supernovae and active galactic nuclei (AGN), and large-scale dynamics and dissipation of turbulent energy act over various time and spatial scales and all regulate star formation in a complex gas cycle.

Aims. This paper presents the physics implemented in a new semi-analytical model of galaxy formation and evolution called the Galaxy Assembler from dark-matter Simulation (G.A.S.).

Methods. The fundamental underpinning of our new model is the development of a multiphase interstellar medium (ISM) in which energy produced by supernovae and AGN maintains an equilibrium between a diffuse, hot, and stable gas and a cooler, clumpy, and low-volume filling factor gas. The hot gas is susceptible to thermal and dynamical instabilities. We include a description of how turbulence leads to the formation of giant molecular clouds through an inertial turbulent energy cascade, assuming a constant kinetic energy transfer per unit volume. We explicitly modelled the evolution of the velocity dispersion at different scales of the cascade and accounted for thermal instabilities in the hot halo gas. Thermal instabilities effectively reduce the impact of radiative cooling and moderates accretion rates onto galaxies, and in particular, for those residing in massive haloes.

Results. We show that rapid and multiple exchanges between diffuse and unstable gas phases strongly regulates star formation rates in galaxies because only a small fraction of the unstable gas is forming stars. We checked that the characteristic timescales describing the gas cycle, gas depletion timescale, and star-forming laws at different scales are in good agreement with observations. For high-mass haloes and galaxies, cooling is naturally regulated by the growth of thermal instabilities, so we do not need to implement strong AGN feedback in this model. Our results are also in good agreement with the observed stellar mass function from $z \approx 6.0$ to $z \approx 0.5$.

Conclusion. Our model offers the flexibility to test the impact of various physical processes on the regulation of star formation on a representative population of galaxies across cosmic times. Thermal instabilities and the cascade of turbulent energy in the dense gas phase introduce a delay between gas accretion and star formation, which keeps galaxy growth inefficient in the early Universe. The main results presented in this paper, such as stellar mass functions, are available in the GALAKSIENN library.

Key words. galaxies: formation – galaxies: evolution – galaxies: star formation – galaxies: ISM – methods: numerical

1. Introduction

Galaxies are defined by their stellar populations, that is “when and where” their stars formed. Therefore, if models are to capture accurately the process of galaxy formation and evolution, researchers must determine how star formation is regulated locally and globally in galaxies. However, star formation is one of the most challenging processes to characterise in galaxy evolution models, essentially because the formation of stars involves many non-linear processes that occur over a wide range of temporal and spatial scales in for example density, velocity, and magnetic field strength and regularity (Kritsuk et al. 2013; Krumholz & McKee 2005). Observations show that star formation in galaxies is a very inefficient process. Only 0.1%–10% of the available gas is converted into stars per local free-fall time (e.g. Lada et al. 2010; Hennebelle & Falgarone 2012; Agertz & Kravtsov 2015; Lee et al. 2016). On the other hand, numerical simulations of molecular clouds indicate that star formation efficiency is

highly dependent on how ionisation and kinetic feedback is injected into the interstellar (ISM; e.g. Krumholz & Thompson 2012; Gatto et al. 2015; Hennebelle & Iffrig 2014; Dale 2015; Geen et al. 2017; Gavagnin et al. 2017).

The global regulation of star formation is one of the greatest challenges in modelling galaxy formation and evolution due to the difficulty in simulating feedback-regulated star formation, and the absence of a detailed physical description of processes responsible for large-scale feedback [whether driven by active galactic nuclei (AGN) or intense star formation or both] and gas accretion onto galaxies. Recently, there is growing observational and theoretical evidence that turbulent pressure injected by young stars is comparable to gravitational pressure in distant disks, enabling self-regulated star formation with low efficiency (Lehnert et al. 2013). Warm diffuse and cold dense gas evolve under the influence of compression by passing spiral arms, thermal and gravitational instabilities, and supernovae-driven shocks. High resolution hydrodynamic simulations, using sub-grid turbulence

models (Schmidt et al. 2013; Semenov et al. 2016), show complex, multiphase, turbulent structures within the ISM with a realistic global Kennicutt–Schmidt relation on kiloparsec scales and gas depletion times in star-forming regions over scales of 10–50 pc consistent with observations. In these models, the global gas depletion time is long, i.e. $\tau_{\text{depl}} = M_{\text{g}}/\dot{M}_{\star} \approx 1\text{--}10$ Gyr, where M_{g} and \dot{M}_{\star} are the gas mass and star formation rate, respectively. This lengthy depletion time is because the gas spends most of the time in a state that does not rapidly lead to star formation. The gas is recycled many times since the lifetime of star-forming clouds or the local gas depletion time in star-forming regions is typically 1–500 Myr. This wide range of local depletion times, which lead to significant gas cycling and recycling is due to dynamical disruption, dispersal by feedback (Raskutti et al. 2016; Semenov et al. 2017), and supersonic turbulence (Guillard et al. 2009). This complexity must be captured in some way in galaxy evolution models to generate realistic galaxies.

In high resolution hydrodynamical simulations (e.g. Hopkins et al. 2014; Schaye et al. 2015; Kimm et al. 2017; Mitchell et al. 2018) or in semi-analytical models (SAMs; e.g. Cole 1991; Cole et al. 2000; Hatton et al. 2003; Baugh 2006; Croton et al. 2006; Cattaneo et al. 2006; Somerville et al. 2008; Guo et al. 2011; Henriques et al. 2013), depending on the model and mass of the galaxy, different prescriptions for various feedback mechanisms have been invoked to regulate star formation. For low stellar masses, $M_{\star} < 10^9 M_{\odot}$, feedback is used to either regulate gas accretion, mostly by heating the gas through photoionisation (e.g. Doroshkevich et al. 1967; Couchman & Rees 1986; Ikeuchi 1986; Rees 1986), or ejecting the gas using the mechanical energy generated by supernovae (SNe; e.g. White & Rees 1978). For high stellar masses, $M_{\star} > 10^{10} M_{\odot}$, models rely on the action of supermassive black holes (SMBH) to inhibit gas accretion onto galaxies. A significant part of the power generated by AGN is used to limit the cooling of the hot gas phase surrounding the galaxy.

Despite the inclusion of some of these processes to regulate the gas content of galaxies, galaxy evolution models fail to reproduce the star formation histories and physical properties of galaxies, mainly because a robust theory of star formation including a dynamical coupling between gas phases is still lacking. As explained in Cousin et al. (2015a), current semi-analytical models overestimate the number of low-mass galaxies, especially at high redshift ($z > 2.0$), where the gap between models and observations is roughly an order of magnitude or more. For high-mass galaxies, AGN feedback, initially used to limit the growth of massive galaxies at low redshift, has also an impact on the star formation rate and history at higher redshift. Consequently, even the massive galaxies, those with $M_{\star} \approx 10^{11} M_{\odot}$ observed at redshifts greater than 3, are not robustly reproduced in current models.

We present a new semi-analytical model G.A.S. – the Galaxy Assembler from dark-matter Simulation – which is based, in part, on a previous version described in Cousin et al. (2015b, 2016). This paper (Paper I) provides an overview of the physical processes considered in G.A.S. and how they are implemented as phenomenological rate equations. In addition, we have two complementary companion papers: Cousin et al. (2019) G.A.S. II, in which we describe and model the mechanisms of dust attenuation and Cousin (2019) G.A.S. III, in which we explore the panchromatic emission of galaxies from the far-ultraviolet to the submillimetre bands.

In this first paper we focus mainly on the regulation of star formation. In previous models, we have adopted an ad hoc recipe to generate a delay between accretion and star formation

(Cousin et al. 2015a). We implement a physical prescription based on the inertial cascade of turbulent energy from large to small scales. Accreted gas onto galaxies is initially considered as mainly diffuse. We compute the mass fraction of the gas subject to phase separation and fragmentation following Sharma et al. (2012) and Cornuault et al. (2018) and references therein, which also depends on the disk properties. Star formation occurs in the fragmented gas at a scale of 0.1 pc. In a large set of semi-analytical models (e.g. Croton et al. 2006; Cattaneo et al. 2006; Somerville et al. 2008), star formation in massive galaxies is regulated by a strong reduction, or even complete suppression, of gas accretion by AGN feedback. In our new model, we do not need efficient AGN feedback, but instead our regulation process is a natural outcome of both the growth of thermal instabilities in the hot halo phase and the dissipation of turbulent energy within the denser, fragmented gas reservoir. Those two processes delay gas accretion onto galactic disks and star formation.

The paper is organised as follows. In Sect. 2, we provide a brief description of the dark-matter simulation we use, as well as the prescription we adopt to implement baryonic accretion rates. In Sect. 3, we focus on the turbulent inertial cascade. We describe how we compute the energy and mass transfer rates between physical scales; we define and compute the gas fragmentation timescale and show that it is a key parameter for the regulation of star formation in low-mass galaxies. In Sect. 4, we describe our model for the gas cycle as it accretes onto galactic disks, from diffuse accreted gas to potentially star-forming gas. In Sect. 5, we describe our implementation of SN and AGN feedback. Section 6 focusses on the thermal instabilities arising in the hot gas phase in the halo. We assume that gas accretion onto galaxies is limited by turbulent mixing in the range of radii, where thermal instabilities develop because gas acquires high velocity dispersions. This allows us to define an effective cooling rate. In Sect. 7 we present and discuss our results, mainly focussing on the evolution of the galaxy stellar mass function with redshift, and the evolution of relevant timescales of physical processes (e.g. gas cooling, fragmentation, and orbital timescales). We also discuss the impact of our implementation of thermal instabilities on the quenching of massive galaxies in massive haloes.

2. From dark matter to baryons

2.1. Dark matter

The G.A.S. model is built upon a set of dark-matter merger trees extracted from a pure N-body simulation. The current simulation uses WMAP-5 yr cosmology ($\Omega_{\text{m}} = 0.28$, $\Omega_{\Lambda} = 0.72$, $f_{\text{b}} = 0.16$, $h = 0.70$, Komatsu et al. 2009) with a volume of $[100 \text{ h}]^3 \text{ Mpc}$ in which 1024^3 particles evolve. Each particle has a mass of $m_{\text{p}} = 1.025 \times 10^8 M_{\odot}$. Haloes and substructures (satellites) are identified via the HaLoMaker code (Tweed et al. 2009). In our merger trees, we only consider haloes with at least 20 dark-matter particles leading to a minimal dark-matter halo mass of $2.050 \times 10^9 M_{\odot}$. Dark-matter haloes grow from smooth accretion. The dark-matter accretion rate, \dot{M}_{dm} , only includes particles that are newly detected in the halo and have never been previously identified in another halo.

2.2. Baryonic accretion

Based on the dark-matter accretion rate \dot{M}_{dm} , the accretion rate of baryons is defined by

$$\dot{M}_{\text{b}} = f_{\text{b}}^{\text{ph-ion}}(M_{\text{h}}, z)\dot{M}_{\text{dm}}, \quad (1)$$

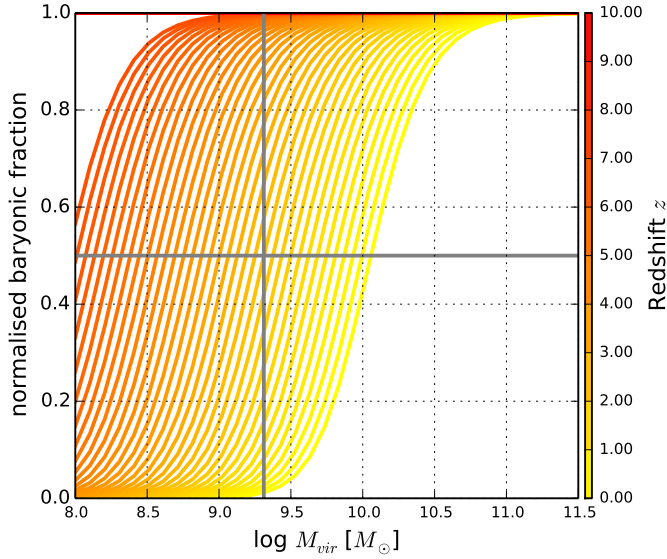


Fig. 1. Relative mass fraction of baryons which are accreted smoothly. The fraction is plotted as a function of both the dark-matter virial mass and redshift. The redshift of each curve is colour coded by the colour bar on the right. Each curve, starting at $z = 0.0$, is separated by $\Delta z = 0.1$. For $z > z_{\text{reion}} = 7.0$, we set $f_b^{\text{ph-ion}}/f_b = 1.0$. The grey vertical bar indicates the minimum halo mass in our dark-matter model. The grey horizontal line indicates $f = 0.5$.

where $f_b^{\text{ph-ion}}(M_h, z)$ is the effective baryonic fraction. The baryon accretion rate depends on the gas ionisation state and we adopt in this work the photoionisation model based on the Gnedin (2000) and Kravtsov et al. (2004) prescription, but we use the effective filtering mass given by Okamoto et al. (2008). We assume that re-ionisation occurs at $z = 7.0$, we therefore set $f_b^{\text{ph-ion}}(M_h, z > 7) = f_b = 0.16$.

Figure 1 shows the normalised mass baryonic fraction that is associated with the dark-matter smooth accretion; we assume a universal baryonic fraction $f_b = 0.16$. Following the minimal dark-matter halo mass used in our model, the main impact ($f_b^{\text{ph-ion}}/f_b < 0.5$) of the photoionisation prescription occurs at low redshift ($z < 0.9$).

As initially proposed by Khochfar & Silk (2009), in G.A.S. we define two different modes of accretion: a cold¹ mode, and a hot mode. Depending on the dark-matter halo mass, the fraction of accreted, hot gas is computed as

$$f_h(M_{\text{vir}}) = \frac{1}{2} [1 + \text{erf}(\log M_{\text{vir}} - \log M_{\text{mix}})] \quad (2)$$

where M_{mix} is the transition mass when the cold and hot gas mass accretion rates are equal and M_{vir} is the halo virial mass. The evolution of the hot gas fraction is inspired by the study of Lu et al. (2011; see their Eqs. (24) and (25)), but we do not account for evolution with redshift since it is very weak and does not strongly impact how the gas is accreted in our model.

The baryonic accretion is divided into two parts that feed two different reservoirs: M_{cold} for the cold mode and M_{hot} for the hot mode. Both the cold and hot reservoirs are fed by metal-free gas. During the evolution of any galaxy, metal-rich ejecta coming from the galaxy are added to the hot reservoir. The metal content of the hot gas therefore depends directly on the rates and timescales over which galaxies create metals. The metallicity of the hot reservoir evolves with time. The chemo-dynamical

¹ The temperature of the accreted gas in this mode is close to 10^4 K.

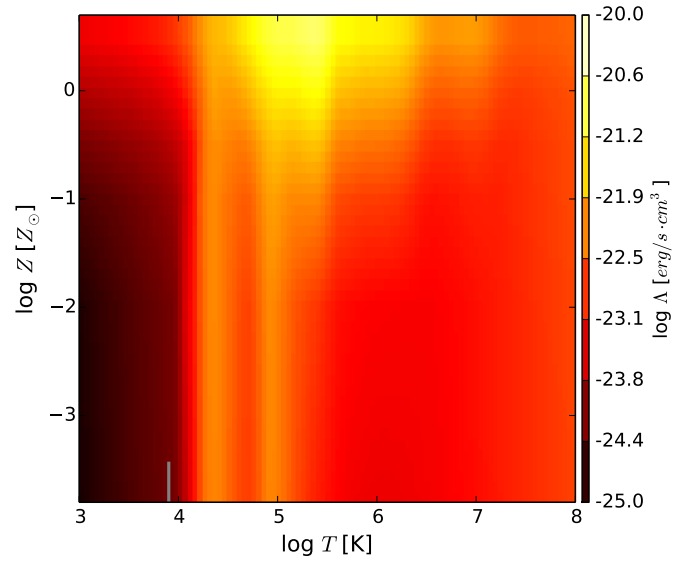


Fig. 2. Cooling efficiency, $\Lambda(T_{\text{hot}}, Z_{\text{hot}})$, as a function of both gas temperature and gas metallicity (see De Rijcke et al. 2013). The values of $\Lambda(T_{\text{hot}}, Z_{\text{hot}})$ are colour coded using the colour bar on the right (in units of $\text{erg s}^{-1} \text{cm}^{-3}$).

model included in G.A.S. tracks the abundance of the main elements in the gas phase. The production and the re-injection of these metals are taken into account for stars with initial masses between $0.1 M_{\odot}$ and $100 M_{\odot}$ over metallicities from zero to super-solar.

2.2.1. Cold accretion mode

For the cold accretion mode, we assume that gas falls directly onto the galaxy (Dekel & Birnboim 2006), and we compute the cold accretion rate via

$$\dot{M}_{\text{streams}} = \frac{M_{\text{cold}}}{2t_{\text{dyn}}}, \quad (3)$$

where M_{cold} is the mass stored in the cold reservoir and t_{dyn} is the dynamical time of the dark-matter halo: $t_{\text{dyn}} = r_{\text{vir}}/v_{\text{vir}}$.

2.2.2. Hot accretion: Radiative cooling

As in the previous versions of this model (Cousin et al. 2015b, 2016), we assume that the hot gas surrounding the galaxy is confined in the potential well of the dark-matter halo and in hydrostatic equilibrium. The hot gas density profile $\rho_{\text{hot}}(r)$ is computed following the prescriptions in Suto et al. (1998), Makino et al. (1998), Komatsu & Seljak (2001) and Capelo et al. (2012). We assume for such hot atmospheres a constant temperature T_{hot} , and a average gas metallicity Z_{hot} . Details are provided in Cousin et al. (2015b, Sect. 4). Radiative cooling and the associated gas condensation is computed using the prescription in White & Frenk (1991). The cooling time of the hot gas phase is defined – as a function of the radius – as

$$t_{\text{cool}}(r) = 10.75 \frac{\mu m_p T_{\text{hot}}}{\rho_{\text{hot}}(r) \Lambda[T_{\text{hot}}, Z_{\text{hot}}]}. \quad (4)$$

In previous versions of this model, we adopted the Sutherland & Dopita (1993) cooling efficiencies $\Lambda(T, Z)$. In the present version, we use those computed by De Rijcke et al. (2013), tabulated as a function of gas temperature and metallicity,

which we interpolate between $T = 10^3$ K and 10^8 K, and for gas metallicities between the range $10^{-4} Z_{\odot}$ and $2Z_{\odot}$. We assume the solar metal mass fraction of $Z_{\odot} = 0.02$. Figure 2 shows the cooling efficiency as a function of both gas temperature and average gas metallicity.

At a given time, the mass of warm gas that can condensate and feed the galaxy is enclosed within the cooling radius, r_{cool} . This radius is calculated using the cooling time equation

$$t(r_{\text{cool}}) = \mathcal{T}_{\text{cool}}^{\text{hot}}. \quad (5)$$

In this equation, $\mathcal{T}_{\text{cool}}^{\text{hot}}$ is a ‘‘cooling clock’’, which is a measure of the effective cooling time of the hot gas. After each time step, Δt , the mass-weighted cooling time is updated via

$$\mathcal{T}_{\text{cool}}^{\text{hot},n} = \underbrace{\left(\mathcal{T}_{\text{cool}}^{\text{hot},n-1} + \Delta t \right)}_{\text{hot halo gas}} \left(1 - \frac{\Delta M}{M} \right) + \underbrace{\frac{\Delta t}{2} \frac{\Delta M}{M}}_{\text{newly incoming hot halo gas}}, \quad (6)$$

where M is the total mass in the hot phase after the latest time step, Δt . The quantity ΔM is the net mass variation of hot gas during this time step (accretion–ejection). We assume that cooling takes place during the overall last time step for the gas already in the hot atmosphere. However, accretion is continuous², and incoming gas also starts to cool. Therefore, taking into account the time for the incoming gas to enter the hot phase, this new gas cools during only half a time step on average. Therefore, the effective cooling time of the hot gas halo can increase or decrease between two time steps depending on the relative fraction of halo gas to incoming gas. During mergers, the cooling clock of the remnant hot phase is set to the value of the most massive progenitor at the time of the merger (see Table 1).

Knowing the cooling radius r_{cool} (deduced from Eq. (5)), we can write the condensation rate of the gas,

$$\dot{M}_{\text{cool}} = \frac{v(r_{\text{cool}})}{2r_{\text{cool}}} \int_0^{r_{\text{cool}}} \rho_{\text{hot}}(r) r^2 dr. \quad (7)$$

The mass within r_{cool} decreases with a timescale $r_{\text{cool}}/v(r_{\text{cool}})$, where $v(r_{\text{cool}})$ is the circular velocity of the dark-matter halo measured at $r = r_{\text{cool}}$. We assume that the hot atmosphere extends up to the virial radius r_{vir} of the dark-matter halo, thus the cooling radius cannot be larger than r_{vir} .

Figure 3 shows the average galaxy gas accretion rates produced by the two different modes as a function of both the dark-matter virial mass and the redshift. At low dark-matter halo masses, $M_{\text{vir}} < 10^{10.5} M_{\odot}$, accretion on the galaxy is dominated by the cold mode. The contribution to the accreted mass of the hot mode increases progressively with halo mass. Around $M_{\text{vir}} \sim 10^{10.5} M_{\odot}$, the contribution of the two modes are equal. This transition occurs at approximately the same halo mass for all redshifts we considered. For both the cold mode and the hot mode accretion, the average accretion rate decreases with the redshift. For dark-matter haloes with $M_{\text{h}} = 10^{10.5} M_{\odot}$, the average accretion due to the cold mode decreases from $20 M_{\odot} \text{ yr}^{-1}$ at $z \simeq 9.0$ to $0.2 M_{\odot} \text{ yr}^{-1}$ at $z = 0.3$; for the hot mode, the accretion rate decreases from $30 M_{\odot} \text{ yr}^{-1}$ to less than $0.3 M_{\odot} \text{ yr}^{-1}$ between $z \simeq 9.0$ and $z = 0.3$.

² Using the adaptive time-step method described in Cousin et al. (2015b), we ensure that all exchange rates between reservoirs are constant during each time step.

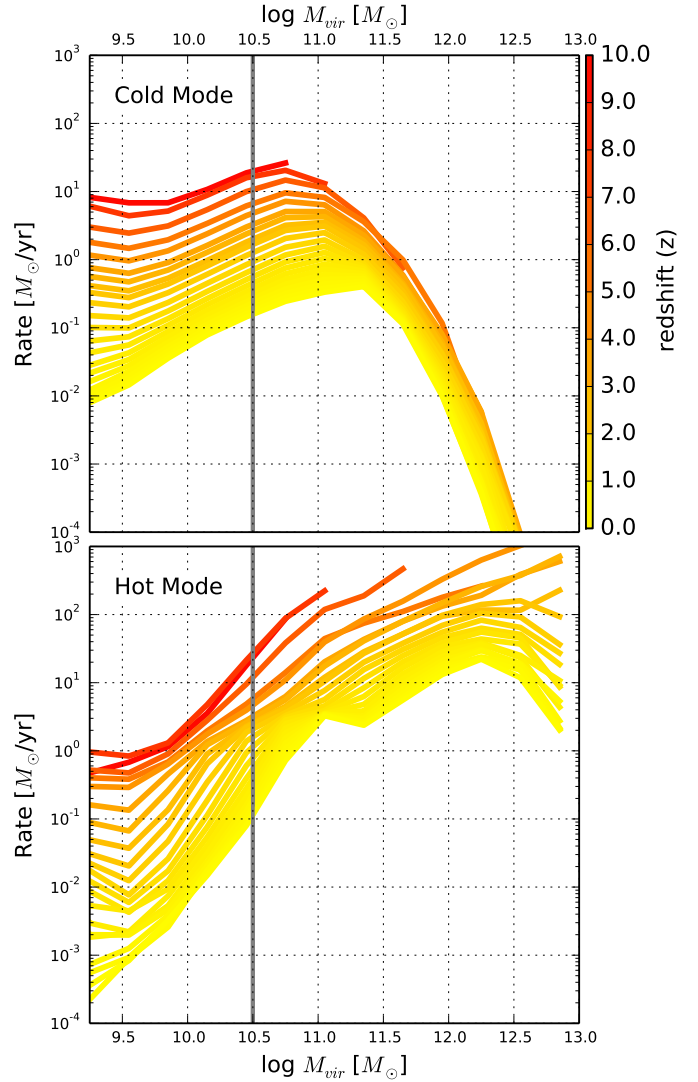


Fig. 3. Mean accretion rates into the galaxy halo galaxy from the cold (*upper panel*) and hot mode accretion (*lower panel*) as a function of virial mass. The mean values are computed by selecting only haloes which are actively accreting gas. Colour code indicates redshift. The grey solid vertical line indicates the dark-matter halo mass for which contributions of cold and hot modes are equivalent.

3. Turbulent inertial cascade

Both observations and numerical simulations of the ISM at high spatial resolutions show that turbulence plays a fundamental role in star formation (e.g. Bergin & Tafalla 2007; Miville-Deschênes et al. 2010; Renaud et al. 2014). Turbulence controls the rate at which kinetic energy is dissipated (e.g. Krumholz & McKee 2005; Padoan & Nordlund 2011; Federrath & Klessen 2012) and leads to the development of multiphase morphological structures in the gas (e.g. André 2013; Levrier et al. 2018). Recent numerical models of galaxy formation have adopted a gravo-turbulent sub-grid model for star formation (e.g. Hopkins 2012; Kimm et al. 2017), but those time-consuming simulations are limited to low cosmological volumes. In this section, we develop our modelling of the mass and energy transport from the large scale of injection to the small dissipative scale, through a hierarchy of structures, using a phenomenological prescription for the cascade of turbulent energy. This allows us to compute the mass of gas that may form stars

Table 1. Definitions, values, and associated references of parameters used in the current model.

Symbol	Definition	Eq./Sect.	Values
	– Reference masses –		
$M_{h,\min}$	Minimal dark-matter halo mass (20 particles)	Sect. 2.1	$2.05 \times 10^9 M_\odot$
M_{mix}	Transition mass from cold to hot accretion mode	Sect. 2.2, Eq. (2)	$10^{11} M_\odot$
M_\bullet^{init}	Initial SMBH mass (seed)	Sect.	$300 M_\odot$
	– Thermal instability –		
ε_{TI}	Thermal instability propagation efficiency	Sect. 6.1.2, Eq. (35)	0.63
	– Feedback: Repartition –		
f_{ME}	Fraction of SMBH infall rate converted in power	Sect. 5.3, Eq. (26)	0.1
$f_{k,\text{SN}}$	Kinetic fraction of SN energy	Sect. 24, Eq. (24)	2/3
f_w	Fraction affected to large-scale wind	Sect. 24, Eq. (24)	0.2
$f_{\text{Th,SN}}$	Thermal fraction of non-kinetic SN energy	Sect. 5.4.4, Eq. (27)	1/2
$f_{k,\text{AGN}}$	Kinetic fraction of AGN energy	Sect. 5.4.1, Eq. (25)	10^{-3}
$f_{\text{Th,AGN}}$	Thermal fraction of non-kinetic AGN energy	Sect. 5.4.4	0.6
	– Accretion –		
$f_{\text{frag}}^{\text{in}}$	Fraction (in mass) of gas which is already fragmented when accreted	Sect. 4	1/3
	– Turbulent kinetic energy budget –		
f_{incr}	Fraction of the rotational energy associated with the newly diffuse accreted gas converted in turbulent kinetic energy	Sect. 5.5.1, Eq. (30)	1/3
f_{disp}	Fraction of the turbulent kinetic energy budget dissipated per dynamical time step	Sect. 5.5.1, Eq. (30)	1/2
	– Feedback: Additional parameters –		
v_w	Large-scale wind velocity	Sect. 5.4.1, Eq. (24)	[100,200] km s ⁻¹
E_{SN}	Total SN energy	Sect. 5.2	10^{44} J
η_\bullet	Inflow-outflow ratio for SMBH activity	Sect. 5.4.1	1.0
η_m	Minimal mass ratio for major merger events		1/3

and to test the impact of these phenomena on the properties of galaxies. In our prescription, the mass of gas that is able to form stars is the mass of gas reaching the dissipation scale l_\star (see Sect. 3.1), calculated from the mass flow rate between scales, under the assumption of a constant energy transfer (Sect. 3.2).

3.1. Self-similar scaling relations and energy transfer rate

We assume that the two Larson (1981) self-similar scaling relations are satisfied to compute the mass and energy transfer rates during the inertial turbulent cascade. The two scaling relations link the mass surface density μ and the 1D-velocity dispersion σ , respectively, to the wave number $k = 1/l$ as follows:

$$\mu_k = \mu_\star \left(\frac{k}{k_\star} \right)^{-a} \quad \text{and} \quad \sigma_k = \sigma_\star \left(\frac{k}{k_\star} \right)^{-b}, \quad (8)$$

where k_\star is the wave number associated with the star formation scale, $l_\star = 1/k_\star$. μ_\star and σ_\star are two normalisation parameters and a and b are the two slope indices; both depend on the gas phase considered (Fleck 1996; Hoffmann & Romeo 2012). The second self-similar scaling relation gives the energy transfer rate per unit volume, as in Kritsuk et al. (2013),

$$\dot{e} \propto \rho(k)\sigma(k)^3 k \propto \mu(k)\sigma(k)^3 k^2 [M \cdot L^{-1} \cdot T^{-3}]. \quad (9)$$

We set the energy transfer rate \dot{e} to be constant (e.g. Hennebelle & Falgarone 2012), which gives a relation between

the two slopes of Larson's relations as follows: $b = \frac{1}{3}(2 - a)$. Following Romeo et al. (2010), we assume $(a, b) = (1/5, 3/5)$. These values reproduce observations of the dynamics of clouds at the atomic-molecular transition. The energy transfer rate during the inertial cascade is now proportional to

$$\dot{e} \propto \mu_\star \sigma_\star^3 k_\star^2. \quad (10)$$

The normalisation parameters used in Eq. (8) are listed in Table 2 and correspond to standard values for dense, supersonic, and compressible gases (Jog & Solomon 1984a,b; Romeo et al. 2010; Hoffmann & Romeo 2012). We assume $l_\star = 0.1$ pc for the star formation scale, which corresponds to the characteristic width of interstellar filaments hosting prestellar cores (e.g. André 2013; Palmeirim et al. 2013). Although the detailed physical interpretation of this width is still debated, this scale is of the order of the scale below which the turbulence becomes subsonic in star-forming filaments (e.g. Padoan et al. 2001). At this scale, we adopt a typical velocity dispersion $\sigma_\star = 0.3$ km s⁻¹ (e.g. Orkisz et al. 2017), a value slightly higher than the speed of sound for a molecular gas at 10 K ($c_s(10 \text{ K}) = 0.2$ km s⁻¹), which corresponds to the observed transition between bound and unbound filaments (Arzoumanian et al. 2013). The critical mass surface density above which the gas is gravitationally unstable and converted into stars is set to $\mu_\star \approx 150 M_\odot \text{ pc}^{-2}$ (Gao & Solomon 2004; Heiderman et al. 2010; Lada et al. 2012).

Table 2. Definitions, values, and associated references of the parameters used in the turbulent cascade scaling relations (Sect. 3 and Eq. (8).)

Symbol	Definition	Values	Refs
k_*	Star formation wave number	$1/l_* = 10 \text{ pc}^{-1}$	André (2013)
μ_*	Mass surface density threshold for star formation	$150 M_\odot \text{ pc}^{-2}$	Lada et al. (2012)
σ_*	Velocity dispersion at the star formation scale	0.3 km s^{-1}	Arzoumanian et al. (2013)
a	Slope index of the Larson surface density law	$1/5$	Romeo et al. (2010)
b	Slope index of the Larson 1D velocity dispersion law	$3/5$	Romeo et al. (2010)

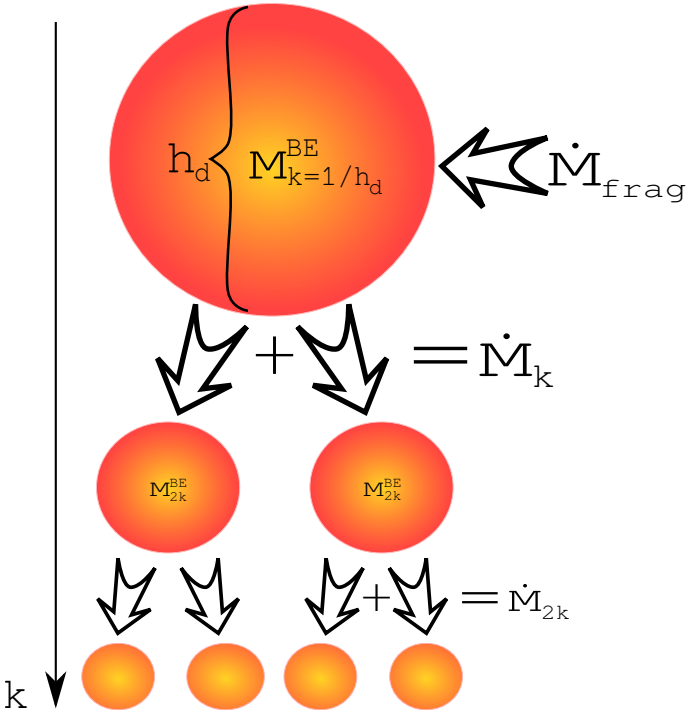


Fig. 4. Diagram illustrating how the gas fragments during the turbulent inertial cascade. When a structure formed at the scale k reaches the Bonnor–Ebert mass, it breaks into smaller structures at the next level, $2k$. The mass transfer rate, \dot{M}_k , between scales is given by Eq. (12). The largest structures are fed at a rate, \dot{M}_{frag} , as discussed in Sect. 4.1 and given by Eq. (14).

3.2. Mass transfer rate during the inertial cascade

In star-forming galactic disks, the structure of the gas is observed to be self-similar over a wide range of scales, from a few kiloparsec to the length scale l_* over which star formation occurs (Dickman et al. 1990; Miville-Deschênes et al. 2010). In this section, we compute the mass transfer rate from the energy transfer rate (Eq. (10)). Starting from the scale of the disk scale height, h_d , which is the largest possible injection scale, the gas mass is progressively distributed over smaller scales. Following the inertial cascade, when going from a scale $1/k$ to $1/2k$, large structures break into smaller ones. A diagram of this progressive fragmentation of the gas in the ISM of our modelled galaxies is shown in Fig. 4.

The Bonnor–Ebert mass at a given scale $1/k$, M_k^{BE} , sets the critical mass above which the structure becomes unstable and collapses into smaller structures (Bonnor 1956), until we reach the dissipation scale l_* , i.e.

$$M_k^{\text{BE}} = 1.5 \frac{\sigma_k^4}{G^2 \mu_k} = 1.5 \frac{\sigma_*^4}{G^2 \mu_*} \left(\frac{k}{k_*} \right)^{-11/5}. \quad (11)$$

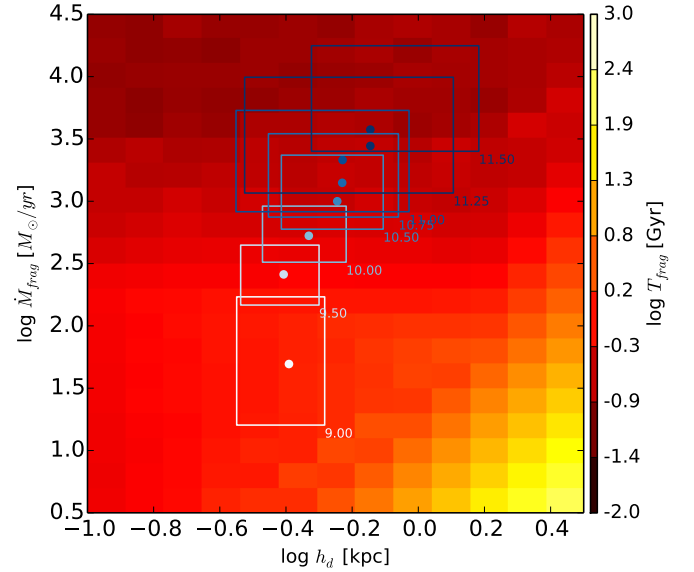


Fig. 5. Gas fragmentation rate as a function of both the maximum energy injection scale, l_{max} , which we assume is equal to the disk scale height, and the growth rate of GMCs (Eq. (14)). Each coloured box shown in the diagram represents the gas fragmentation rate for a specific stellar mass bin as indicated by the labels at the lower right corner of each box. The mass bins range from $M_* = 10^{9.0} M_\odot$ to $10^{11.5} M_\odot$. The limits of each box encompass the 15% and 85% percentiles of the distribution for each mass bin. Coloured points within each box indicate the value of the median of the distribution. The regions shown are for modelled galaxies at $z = 2.1$.

We derive the mass transfer rate \dot{M}_k between wave numbers k and $2k$, using the conservation of energy

$$\frac{3}{2} \dot{M}_k \sigma_k^2 = \dot{e} V_k E \left(\frac{M_k}{M_k^{\text{BE}}} \right), \quad (12)$$

where $V_k = \frac{\pi}{6} k^{-3}$ is the volume of the cloud at scale k , M^k is the total mass stored at scale k , and \dot{e} is the constant kinetic turbulent energy transfer rate per unit volume (Eq. (10)).

3.3. Gas fragmentation timescale

We define the gas fragmentation timescale, T_{frag} , as the time needed to transfer all the gas from the disk scale height, h_d , to the star formation length scale, l_* , assuming that the fragmented gas reservoir is fed at a constant rate \dot{M}_{frag} . During the process of fragmentation, we assume that the disk scale height is constant. In practice, we track the mass of the initial Bonnor–Ebert sphere along the cascade given by Eq. (12), until a steady state is reached.

Figure 5 shows our estimated gas fragmentation timescale $T_{\text{frag}}[h_d, \dot{M}_{\text{frag}}]$ as a function of the instantaneous disk scale height h_d and mass flow rate at which the largest structure is

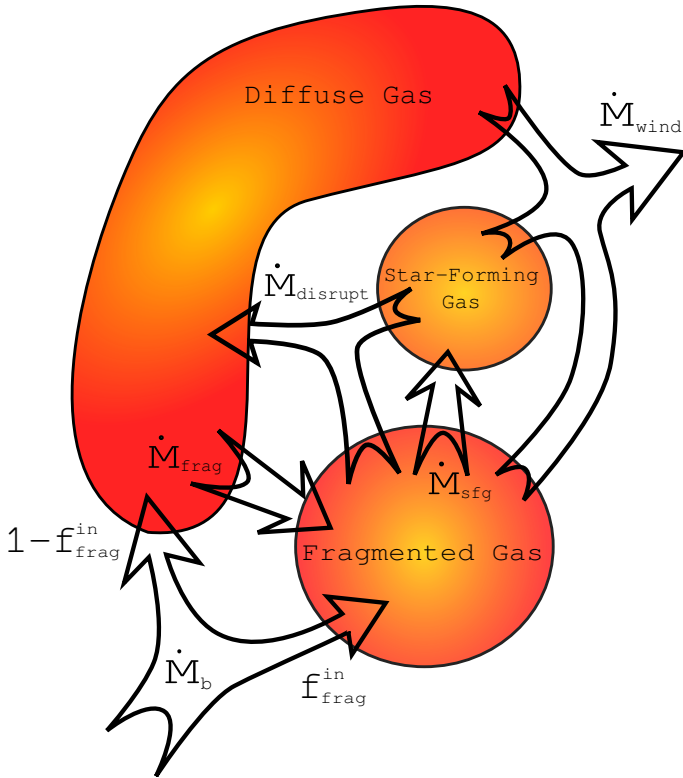


Fig. 6. Diagram illustrating the flow of mass between the three gas reservoirs considered in our multiphase G.A.S. model. The star-forming reservoir contains the mass that is immediately available for star formation. It is fed by a turbulent cascade and gas fragmentation at a mass flow rate, \dot{M}_{sfg} (Eq. (18)). The gas reservoir fragments at a rate, \dot{M}_{frag} (Eq. (14)). The diffuse gas reservoir is fed through the rates of two mechanisms: (1) gas accretion rate $(1 - f_{\text{frag}}^{\text{in}})\dot{M}_{\text{b}}$ (Eq. (1)); and (2) the rate at which the fragmented and star-forming gas is disrupted by the energy injected by SNe and AGN, \dot{M}_{disrupt} (Eqs. (27) and (28)). Each of these three gas reservoirs contributes to the outflow rate, \dot{M}_{wind} (Eqs. (24) and (26)).

fed, \dot{M}_{frag} (see Sect. 4.1 and Eq. (14)). Both the average disk scale height and the giant molecular clouds (GMCs) growth rate increase with the stellar mass, and the average fragmentation timescale decreases. Obviously, gas accretion history, star formation, gas ejection modify those two parameters, h_{d} and \dot{M}_{frag} . To take this into account, we define an effective disk fragmentation timescale $\mathcal{T}_{\text{frag}}$, which depends on the history of the disk (see Sect. 4.2 and Eq. (19)).

4. G.A.S. cycle: Evolution of the gas reservoirs

To model the gas cycle in galactic disks, we follow the mass content of three gas reservoirs: (i) a diffuse gas reservoir; (ii) a fragmented gas reservoir; and (iii) a star-forming gas reservoir. In the following, we describe each of these gas reservoirs and their mass flow rates, of which we give a schematic view in Fig. 6. We consider that the accreted gas onto the galaxy is multiphase and we set the mass fraction of fragmented gas to $f_{\text{frag}}^{\text{in}} = 1/3$ (van de Voort & Schaye 2012). We further discuss the physical origin and consequences of that assumption in Sect. 8.

4.1. Diffuse gas reservoir

The first main reservoir stores a diffuse ($\approx 1 \text{ cm}^3$) and warm (10^4 K) gas which is traced by emission lines of the warm,

ionised medium such as [O III] and [N II] (e.g. Béthermin et al. 2016; Laporte et al. 2017; Suzuki et al. 2017). As illustrated in Fig. 6, the diffuse gas reservoir is fed by two different sources: (i) the accretion of warm gas coming from both the cold and the hot mode, and (ii) the disruption of fragmented gas by the injection of energy from a SN or an actively accreting SMBH (see Sect. 5.4.3).

The initial temperature of the warm, diffuse gas is assumed to be 10^4 K , and we compute its isobaric cooling using the same approach as with the hot gas phase (Sect. 2.2.2). The effective cooling time of this phase $\mathcal{T}_{\text{cool}}^{\text{diff},n}$ is computed after each time step Δt , as the sum of the mass-weighted cooling times of the halo and the newly incoming gas, i.e.

$$\mathcal{T}_{\text{cool}}^{\text{diff},n} = \underbrace{\left(\mathcal{T}_{\text{cool}}^{\text{diff},n-1} + \Delta t \right) \left(1 - \frac{\Delta M}{M} \right)}_{\text{warm gas}} + \underbrace{\frac{\Delta t}{2} \frac{\Delta M}{M}}_{\text{newly incoming gas}}, \quad (13)$$

where M is the total mass of diffuse gas after the time step Δt . The quantity ΔM is the mass increase of diffuse gas coming from accretion and disrupted gas coming from the fragmented phase. We assume that radiative cooling acts on the warm diffuse gas during all the previous time step (Eq. (13)). However, for the newly incoming gas, we assume that the radiative cooling occurs only during half of the previous time step (as for the hot gas phase). Therefore, the fraction of halo and freshly acquired gas can increase or decrease after each step.

The mass transfer rate between the diffuse gas phase to the fragmented gas phase \dot{M}_{frag} is computed using the cooling timescale

$$\dot{M}_{\text{frag}} = (1 - f_{\text{frag}}) f_{\text{Q}} \phi_{\text{m}} \frac{M_{\text{diff}}}{t_{\text{cool}}^{\text{diff}}}, \quad (14)$$

where M_{diff} is the mass of diffuse gas. The cooling timescale is computed as a function of the average metallicity Z_{diff} (Eq. (4)). We assume a temperature of 10^4 K and an average volume of the warm, diffuse gas component³ $\mathcal{V}_{\text{diff}} = 22\pi r_{\text{d}}^2 h_{\text{d}}$. The other parameters in Eq. (14) are computed as follows:

- ϕ_{m} , the mass fraction of diffuse gas that condenses in an effective cooling time, $\mathcal{T}_{\text{cool}}^{\text{diff}}$, depends on the characteristic cooling timescale of this reservoir $t_{\text{cool}}^{\text{diff}}$. The parameter ϕ_{m} was calculated in Cornuault et al. (2018) and for computational purposes we fitted their computation by an error function given by,

$$\phi_{\text{m}}(x) = \frac{1}{2} \left[1 + \text{ERF} \left(\frac{\log_{10} x - \log_{10} x_t}{\sqrt{s}} \right) \right], \quad (15)$$

where $x = \mathcal{T}_{\text{cool}}^{\text{diff}}/t_{\text{cool}}^{\text{diff}}$. The best fit gives $x_t = 0.55$ and $s = 0.13$.

- f_{frag} , the fraction of the gas that is fragmented, is defined as

$$f_{\text{frag}} = \frac{M_{\text{frag}} + M_{\text{sfg}}}{M_{\text{diff}} + M_{\text{frag}} + M_{\text{sfg}}}, \quad (16)$$

where M_{diff} , M_{frag} and M_{sfg} are the gas masses stored in the three different reservoirs (Fig. 6). This factor accounts for the fact that the more the gas is fragmented, the lower the mass flow to bound structures.

³ The diffuse gas is assumed to evolve in a thick disk with scale height h_{d} and a total radius of the stellar disk equal to $11r_{\text{d}}$. We assume that the warm diffuse gas can extend to a radius that is up to two times larger than total radius of the stellar disk.

– f_Q , the Toomre disk instability criterion, is calculated as,

$$f_Q = \text{MAX} \left[1.0; \frac{Q_{\text{crit}}}{Q} \right]. \quad (17)$$

This factor accounts for the fact that the mass flow rate to bound structures increases as the diffuse gas becomes gravitationally unstable. We adopt a standard value of $Q_{\text{crit}} = 1.0$.

4.2. Fragmented gas reservoir

The clumpy gas phase has low filling factor (typically less than $\approx 10\%$) with densities $n_{\text{H}} = 1-10^3 \text{ cm}^{-3}$ and temperatures 10^3-10 K . It is traced by atomic and molecular gas lines such as CO, [C I], and [C II] (e.g. Aravena et al. 2016; Bothwell et al. 2016, 2017; Decarli et al. 2016; Popping et al. 2017). We assume that the fragmented gas is contained within spherical and bound structures with initial radii $r = h_d/2$, and must contain at least a mass $M_{\text{BE}}^{1/h}$ to be unstable. At each time step, the maximum number of bound structures formed in the disk is $N_{\text{GMC}} \approx (M_{\text{frag}} + M_{\text{sfg}})/M_{1/h_d}^{\text{BE}}$.

The gas fragmentation process feeds the star-forming gas reservoir at the rate

$$\dot{M}_{\text{sfg}} = \frac{M_{\text{frag}}}{\mathcal{T}_{\text{frag}}}. \quad (18)$$

The effective disk fragmentation timescale $\mathcal{T}_{\text{frag}}$ is updated at each time step and allows us to follow the full history of the gas cycle in relation with the current disk properties (h_d and \dot{M}_{frag}). If M is the total mass of fragmented gas after the last time step Δt , and $\Delta M = \Delta t \dot{M}_{\text{frag}}$ is the gas mass incorporated in the fragmented gas reservoir, thus the disk fragmentation timescale is updated as

$$\mathcal{T}_{\text{frag}}^n = \mathcal{T}_{\text{frag}}^{n-1} \left(1 - \frac{\Delta M}{M} \right) + \frac{\Delta M}{M} T_{\text{frag}} [h_d, \dot{M}_{\text{frag}}]. \quad (19)$$

This timescale is a function of $T_{\text{frag}}[h_d, \dot{M}_{\text{frag}}]$, which depends on values of h_d and \dot{M}_{frag} , the disk scale height and the growth rate of GMCs, respectively (see Sect. 3.3).

4.3. Star-forming gas reservoir

Progressively, the fragmented gas is converted into star-forming gas. The gas contained in the star-forming reservoir is very dense and cold, typically traced by molecular lines as HCN (e.g. Oteo et al. 2017). The star-forming gas reservoir is characterised by its very short timescale before it forms stars. We assume a constant star formation timescale which is linearly dependent on the length scale over which star formation occurs, $t_{\text{sf}} = l_{\star}/\sigma_{\star} = 0.1 \text{ Myr}$. The star formation rate is then simply given by

$$\text{SFR} = \dot{M}_{\star} = \frac{M_{\text{sfg}}}{t_{\text{sf}}}. \quad (20)$$

In our prescription, the rate of star formation is mainly limited by the rate at which gas becomes clumpy.

4.4. Schmidt–Kennicutt laws

The G.A.S. model follows the evolution of three interacting gas reservoirs. Star formation is assumed to occur in the reservoir containing the densest gas after a progressive structuring starting at disk scale height in some GMCs. We can estimate the

Schmidt–Kennicutt laws at two different scales within this context. At the galaxy scale, the gas surface density Σ_{gas} is computed assuming that half of the mass of the fragmented gas is enclosed in the half mass radius ($1.68r_d$) of the disk. The star formation rate surface density Σ_{SFR} is computed in the same way. Using these two definitions, our predictions of the galaxy-scale Schmidt–Kennicutt law is shown in Fig. 7.

The median trend of our star-forming galaxy sample at $z = 2.1$ is in good agreement with the Kennicutt (1998) relation over four orders of magnitude (Fig. 7). Below $\log_{10}\Sigma_{\text{gas}} = 1.0$, our star formation rates are slightly higher than those deduced from Kennicutt (1998). We measure an average scatter of ≈ 0.46 dex. Our star-forming galaxy distribution is fully consistent with individual measurements galaxies measurements (Kennicutt 1998; Bigiel et al. 2008). Using the whole star-forming sample of galaxies in the redshift range $z = 1.5-2.1$, we estimate a power-law index, $N = 1.232 \pm 0.002$. This slope is slightly shallower than the standard Kennicutt (1998) slope. In addition, we find a clear trend of an increasing gas surface density as the fragmented gas fraction increases (Fig. 7).

In our model, we assume that star formation is initiated in GMCs by the progressive fragmentation of the gas. GMCs are scaled to the disk scale height h_d . By assuming that the fragmented gas mass and the star formation rate is homogeneously distributed in all GMCs formed into the disk, we define the gas surface density and the star formation rate surface density as follows:

$$\Sigma_{\text{gas}} = \frac{1}{N_{\text{GMC}}} \frac{M_{\text{frag}} + M_{\text{sfg}}}{\pi(h_d/2)^2} \quad \text{and} \quad \Sigma_{\text{SFR}} = \frac{1}{N_{\text{GMC}}} \frac{\text{SFR}}{\pi(h_d/2)^2}. \quad (21)$$

The clear correlation we found at the galaxy scale disappears at the GMC scale (Fig. 7). Compared to galaxy scale, the fragmented gas fraction shows the opposite trend: the higher the gas surface density in a GMC, the lower its fragmented gas fraction. This trend can be translated as follows: galaxies with a relatively higher (lower) fraction of fragmented gas host relatively more (fewer) GMCs. The mass of gas and star formation is homogeneously distributed in all GMCs (Eq. (21)). The average gas surface density is therefore lower in galaxies with highly fragmented gas, which also happen to host more GMCs than galaxies with relatively low levels of fragmentation.

5. Model feedback

Massive stars and AGN in galaxies inject significant amounts of energy into the ISM and the circumgalactic medium (CGM) of galaxies. One of main challenges in galaxy evolution models is to distribute this power into the various gas phases in these media (efficiency parameters are listed in Table 1). We now describe how we distribute the SN and AGN power within the galaxy and its surroundings in the G.A.S. model, and how this power is used to regulate star formation.

5.1. Morphology and the efficiency of outflows

Galaxies in the early universe frequently have a clumpy morphology, which suggests there are interacting regions of dense gas and stars (e.g. Elmegreen et al. 2009; Elmegreen 2009). The clumps in these distant galaxies are not typically observed in the same range of mass and size as star-forming regions in local disk galaxies. This morphological evolution provides clues as to how star formation in distant galaxies may have proceeded and was regulated, but overall observations are consistent with mass and

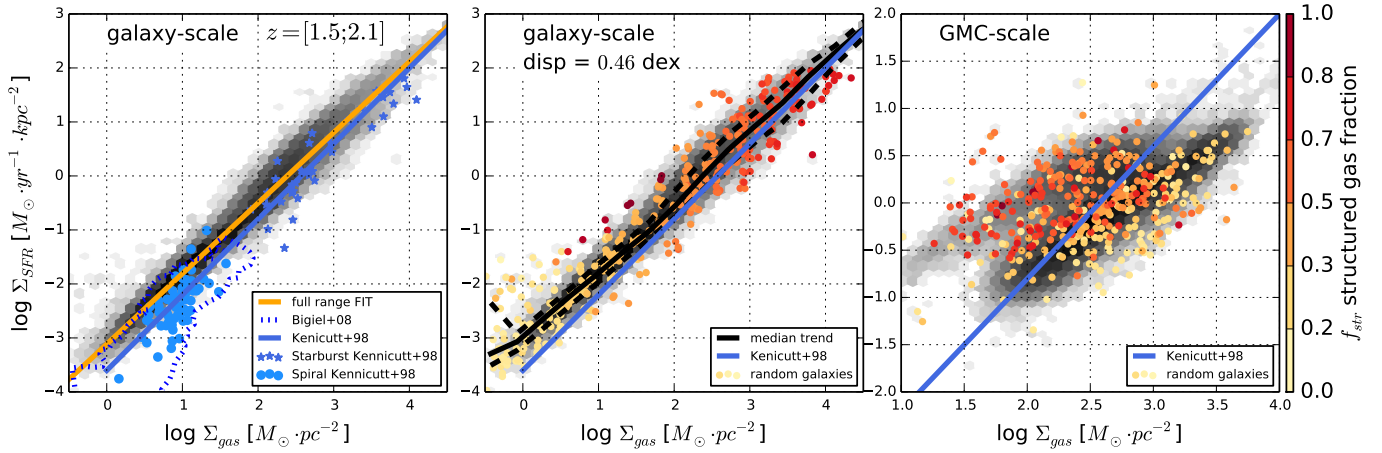


Fig. 7. Schmidt–Kennicutt relations estimated from our model galaxies at both the galaxy scale and GMC scale. In the three panels the solid blue line indicates the **Kennicutt (1998)** relation. The grey shaded area is the distribution of a full sample of star-forming galaxies in the redshift range $z = 1.5\text{--}2.1$. Each coloured point indicates the position of an individual star-forming galaxy, uniformly selected within the galaxy-scale $\Sigma_{\text{gas}} - \Sigma_{\text{SFR}}$ plan. The colour bar indicates the fragmented gas fraction in each representative galaxy. *Left panel:* the solid orange line is the best fit computed for our entire sample of star-forming galaxies within this redshift range. The dashed contour indicates measurements of nearby galaxies (**Bigiel et al. 2008**). The blue stars and points are from the starburst and normal disk galaxy samples from **Kennicutt (1998)**, respectively. *Central panel:* the solid black line indicates the median trend of our star-forming galaxy sample. The dashed black lines indicate the 15% and 85% percentiles ranges of our model galaxies. We estimate an average scatter in the distribution of 0.46 dex. *Right panel:* distribution of galaxies in the $\Sigma_{\text{gas}} - \Sigma_{\text{SFR}}$ estimated over the GMC-scale of star formation (see text for details). We show this estimate for the same galaxy sample used in the galaxy-scale Schmidt–Kennicutt relation in the two leftmost plots in this figure.

energy injected into the ISM playing an important role in regulating star formation in galaxies (self-regulation; e.g. **Lehnert et al. 2015**, and references therein).

To account for this morphological evolution, we assign to each galaxy a specific morphological type. Galaxies that formed recently are assumed to be clumpy. The morphological type can then change, from clumpy to a smooth(er) disk, only during a merger. If two clumpy galaxies merge, the remnant galaxy can be converted to a smooth disk with a probability of 25%. If two different morphologies merge, the remnant galaxy is assigned the morphology of the most massive progenitor (clumpy or smooth). Following these rules, clumpy galaxies become progressively regular, smoother disk galaxies. Furthermore, galaxies with these two different morphologies also behave differently as they evolve. For low-mass, clumpy galaxies, we assume that gas is more compact and therefore that ejection of gas through outflows is more efficient. The terminal velocity of a wind is proportional to the square root of the energy injection rate divided by the total mass flow rate of the wind \dot{M}_{wind} . Thus, we can translate this efficiency in terms of the average terminal wind velocity. For clumpy galaxies, we assume the terminal velocity of the wind $V_w = 100 \text{ km s}^{-1}$, and $V_w = 200 \text{ km s}^{-1}$ for disk galaxies. The mass loading factor $\dot{M}_{\text{wind}}/\text{SFR}$ is higher in clumpy galaxies than in rotating smooth disks. At $z = 2.1$, distributions of the mass loading factor are characterised by the probabilities that they lie above the 25th, 50th, and 75th percentiles for all star-forming galaxies, which values are $\dot{M}_{\text{wind}}/\text{SFR} = 6.5$, 11.5, and 26.4, respectively.

5.2. Supernovae feedback

In **Cousin et al. (2016)** the instantaneous SN event rate ($\eta_{\text{SN}} \text{ Gyr}^{-1}$) is proportional to the star formation rate, which is a function of time (related to the star formation history; see Sect. 3.5.2). Based on the SN rate, the instantaneous power generated by SNe is simply given by $Q_{\text{SN}} = \eta_{\text{SN}} E_{\text{SN}}$. We assume the standard value $E_{\text{SN}} = 10^{44} \text{ erg s}^{-1}$ (e.g. **Aguirre et al. 2001**). In

G. A. S., outside of the morphological division in the efficiency of outflows, we use the same prescriptions as in **Cousin et al. (2016)**. Please see that paper for details.

5.3. Active nuclei feedback

In addition to the energy input from SNe, the different gas phases are also impacted by the energy produced by AGN. We assume that SMBHs are created during a major merger if the remnant galaxy have a bulge of at least $10^6 M_{\odot}$. The seed of the SMBH is given an initial mass, $M_{\bullet}^{\text{init}} = 300 M_{\odot}$. We associate a gas torus with each SMBH formed. The torus is the gas reservoir that feeds the growth and energy output of the SMBH. This torus is fed by diffuse gas during merger events using the following prescription:

$$\Delta M_{\text{torus}} = 0.1 \mu_g \mu_m M_{\text{diff}}(r < 3r_{\text{torus}}), \quad (22)$$

where μ_g is the mass fraction of the gas that is diffuse in the remnant galaxy disk; $\mu_m = \frac{\text{MIN}(M_1, M_2)}{\text{MAX}(M_1, M_2)}$ is the merger mass ratio and M_i is the total mass (dark-matter + galaxy) enclosed in the half mass radius of the halo; $M_{\text{diff}}(r < 3r_{\text{torus}})$ is the mass of diffuse gas enclosed in $r < 3r_{\text{torus}}$. The torus radius, $r_{\text{torus}} = 10 \text{ pc}$.

From Eq. (22) it is clear that major mergers with gas-rich progenitors accrete the largest amount of gas onto the torus of the remnant, while minor mergers of gas-poor progenitors lead to very little accretion. The infall rate is chosen to be the maximum value of the Bondi infall rate (**Bondi 1952**) and the free-fall rate given by

$$\dot{M}_{\text{infall}} = \text{MAX} \left(\frac{3\pi G \mu k_B T_{\text{torus}}}{4\Lambda(T_{\text{torus}}, Z_{\text{torus}})} M_{\bullet}; \frac{M_{\text{torus}}}{2t_{\bullet}} \right), \quad (23)$$

where the gas temperature of the torus is fixed to $T_{\text{torus}} = 10^{6.5} \text{ K}$. Z_{torus} is the gas-phase metallicity of the torus. The metallicity and temperature determine the cooling efficiency $\Lambda(T, Z)$ (see Sect. 2.2.2). The values M_{\bullet} and M_{torus} are the SMBH mass and

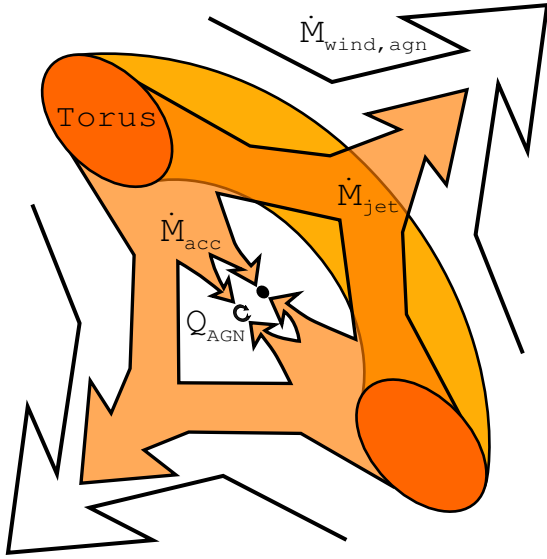


Fig. 8. Illustration showing the distribution of accretion and energy output of a model AGN. Gas infalling from the torus to the AGN is divided between accretion onto the SMBH \dot{M}_{acc} and ejection \dot{M}_{jet} . We assume the diffuse gas in the disk can be driven outwards by the AGN to give the overall ejection rate of gas $\dot{M}_{\text{wind,agn}}$. The accretion onto the SMBH \dot{M}_{acc} is itself divided in two parts: a fraction $(1-f_{\text{ME}})\dot{M}_{\text{acc}}$ of the infalling gas goes to increasing the mass of the SMBH, while the remaining gas $Q_{\text{agn}} = f_{\text{ME}}\dot{M}_{\text{acc}}$ is converted into energy (see Sect. 5.3 for details). The radiative and mechanical energy of the AGN is then distributed in various ways as illustrated in Fig. 9.

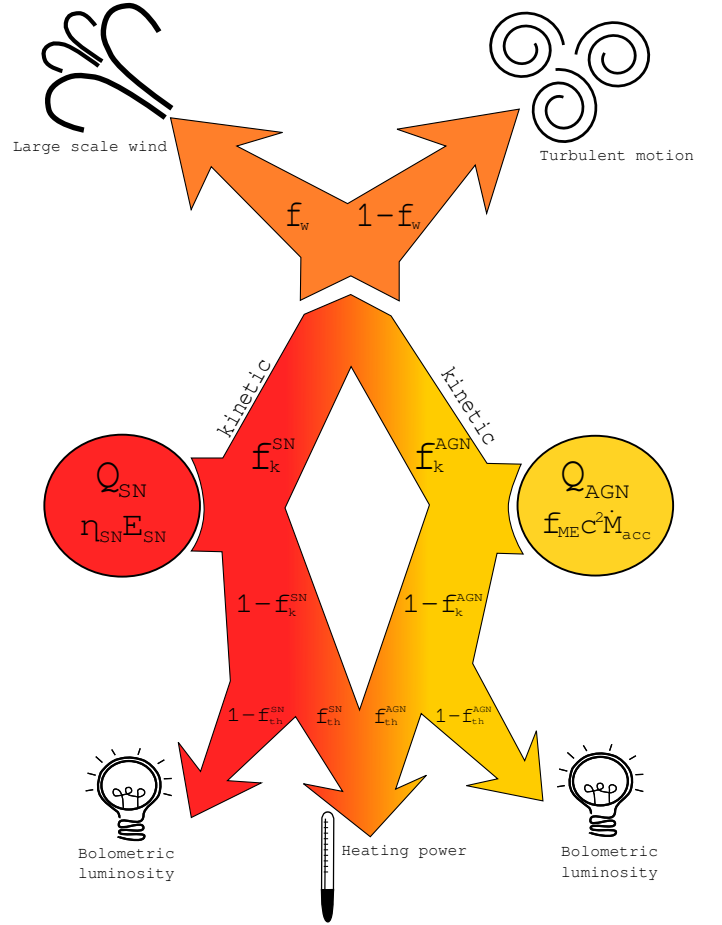


Fig. 9. Illustration of how the energy output from AGN and SNe are distributed. The SNe and AGN power Q_{SN} and Q_{AGN} are shown in red and yellow, respectively. Specifically, the power is distributed through four different channels: (1) as a large-scale wind, (2) in generating turbulence, (3) in heating the gas, and (4) output bolometric luminosity.

the torus mass, respectively; t_{\bullet} is the orbital time of gas at the radius of the torus r_{torus} .

Our prescription for the infall of gas onto the SMBH follows closely that presented in Ostriker et al. (2010). The total mass infall flux into the region of influence of the SMBH is the sum of the mass flux that is accreted onto the SMBH \dot{M}_{acc} , and the fraction that is driven out of the region of influence of the SMBH \dot{M}_{ej} . This yields the total infall rate $\dot{M}_{\text{infall}} = \dot{M}_{\text{ej}} + \dot{M}_{\text{acc}}$. In our model, we assume that the relative fraction of ejected to accreted mass flow rates is $\eta_{\bullet} = \frac{\dot{M}_{\text{ej}}}{\dot{M}_{\text{acc}}} = 1.0$. This division results in equal shares of the infalling gas to (i) be accreted onto the SMBH, driving the increase of the SMBH mass; and (ii) generate power by converting a fraction of the accreted mass $f_{\text{ME}} = 0.1$ (Fig. 8). The power produced by the AGN is then $Q_{\text{AGN}} = f_{\text{ME}}\dot{M}_{\text{acc}}c^2$.

5.4. Distribution of feedback power

We use the power output from AGN and SNe to regulate the gas cycle in galaxies in three different ways in our model. The two different sources of power – the AGN power Q_{AGN} and the mechanical energy of SN Q_{SN} – are each divided in two parts: kinetic power $f_{k,\text{AGN}}$ and $f_{k,\text{SN}}$, and the bolometric power (Fig. 9). A fraction of the kinetic power f_w is used to drive a large-scale wind (Sect. 5.4.1). The residual fraction $1-f_w$ is used to disrupt the fragmented gas of the disk (Sect. 5.4.3) and power the turbulence of the diffuse gas. A fraction of both the AGN and starburst bolometric luminosity, $f_{\text{th,AGN}}$ and $f_{\text{th,SN}}$, respectively, are used to heat the ejected gas (Sect. 5.4.4).

5.4.1. Large-scale ejecta

We derive the instantaneous ejection rate using the conservation of the kinetic energy released by SNe (e.g. Dekel & Silk 1986;

Kauffmann et al. 1993; Efsthathiou 2000), i.e.

$$\dot{M}_{\text{wind,SN}} = 2f_w f_{k,\text{SN}} f_{\text{esc}} \frac{Q_{\text{SN}}}{v_w^2} \quad (24)$$

where v_w is the average velocity of large-scale wind (Sect. 5.1) and f_{esc} is the fraction of mass that escapes the disk.

The ejection rate due to the energy output of the AGN is determined by the fraction of the mass infalling towards the SMBH driven out (\dot{M}_{ej} ; Sect. 5.3). We derive the velocity of the jet using conservation of energy,

$$v_{\text{jet}}^2 = \frac{2f_w f_{k,\text{AGN}} Q_{\text{AGN}}}{\dot{M}_{\text{ej}}} = \frac{2f_w f_{k,\text{AGN}} f_{\text{ME}}}{\eta_{\bullet}} c^2. \quad (25)$$

The outflowing jet is coupled to the diffuse gas in the disk which reduces its velocity. To obtain the coupling efficiency between the jet and the diffuse gas, we assume that jet velocity is equal to the escape velocity of the galaxy V_{esc} . Thus, the kinetic energy of the jet is simply used to drive the diffuse gas out of the galaxy. This leads to an mass outflow rate given by

$$\dot{M}_{\text{wind,AGN}} = \dot{M}_{\text{ej}} \text{MAX} \left[1.0, \left(\frac{v_{\text{jet}}}{V_{\text{esc}}} \right)^2 \right]. \quad (26)$$

The total instantaneous ejection rate of the combined action of SNe and AGN is $\dot{M}_{\text{wind}} = \dot{M}_{\text{wind,SN}} + \dot{M}_{\text{wind,AGN}}$. Each gas

reservoir, M_{diff} , M_{frag} , and M_{sf} , contributes to the instantaneous ejection rate in proportion of its mass fraction (Fig. 6).

5.4.2. Overestimating the gas escape fraction: Re-accretion timescale

When the gas is ejected from the disk, a fraction of the gas remains in the hot CGM. The remaining fraction is ejected from the dark-matter potential. As in Cousin et al. (2015b, see their Sect. 4.3 and Fig. 6) to compute the escape fraction of the gas, we adopt a ‘‘ballistic’’ approach based on the comparison of the dark-matter escape fraction to a shifted Maxwell–Boltzmann velocity distributions. However, as shown by for example Oppenheimer & Davé (2008), the ejected gas is not only affected by gravitational forces (V_{esc}) but also by ram pressure of gas in the CGM. Because of this affect, our ballistic approach overestimates the escape fraction. To correct this bias, we store all the hot gas ejected from a halo in an extended circumgalactic reservoir $M_{\text{ex-cir}}$. The mass stored in this reservoir is then progressively re-accreted and added to the hot gas trapped into the dark-matter potential well.

The gas that is re-accreted out of this reservoir has been implemented in various semi-analytical models (e.g. De Lucia et al. 2004; Somerville et al. 2008; Guo et al. 2011; Henriques et al. 2013) and potentially plays a major role in contributing to the overall gas supply of galaxies. The formulation in the various semi-analytical models can vary. For example, Guo et al. (2011) adopted a prescription depending on both of the dark-matter halo mass and redshift. In Henriques et al. (2013) the re-accretion of gas is inversely proportional to the dark-matter virial mass without any dependence on redshift. In G.A.S., we adopt a prescription similar to De Lucia et al. (2004) and Somerville et al. (2008). We assume that the gas is re-accreted on a timescale which is twice the halo crossing time $t_{\text{reacc}} = 2r_{\text{vir}}/v_{\text{vir}}$.

5.4.3. Disruption rate

The power generated by SNe that does not contribute to driving a wind is assumed to be injected directly into the ISM. This remaining power is distributed between the different ISM gas phases in proportion to their mass. Massive stars are assumed to remain mostly embedded in their dense birth clouds during their short lifetimes (3×10^6 yr; Cousin et al. 2016). The fraction of the SN energy which is injected into the fragmented and star-forming gas f_{frag} disrupts this phase, feeding the diffuse gas reservoir (Eq. (16) and Fig. 6). The disruption rate is defined as

$$\dot{M}_{\text{disrupt,SN}} = 2f_{\text{frag}}f_k^{\text{SN}} \frac{Q_{\text{SN}}}{\sigma_v^2} \frac{[(1-f_w) + f_w(1-f_{\text{esc}})]}{1-f_w f_{\text{esc}}}, \quad (27)$$

where σ_v is the average velocity dispersion of the diffuse gas (Sect. 5.5.1). We assume that the velocity dispersion of the disrupted gas is equal to that of the diffuse gas. Assuming this implies that if the diffuse gas is highly turbulent (high σ_v) then it takes more energy to disrupt the fragmented gas. The disruption rate depends on two terms specifically related to the gas phase. The first term, $1 - f_w$, corresponds to the minimum fraction of SN power which disrupts the fragmented gas. The second term, $f_w(1 - f_{\text{esc}})$, corresponds to the fraction of power that remains in the gas because of the limit imposed by the galaxy escape fraction. The fraction of power which does not contribute to the wind is therefore re-injected to disrupt the fragmented gas. This fraction increases with the galaxy mass. The fraction of the power of SNe injected into the diffuse gas phase maintains or increases

the level of turbulence of the diffuse gas (i.e. σ_v). During a time step Δt , the possible increase of the turbulent energy of the diffuse gas is given by $\Delta E_{\sigma}^{\text{SN}} = (1 - f_{\text{frag}})f_k^{\text{SN}}(1 - f_w f_{\text{esc}})Q_{\text{SN}}\Delta t$ (see Sect. 5.5.1 for all other contributions).

Simultaneously, we also include the energy output from AGN in disrupting the gas. The contribution from any AGN to the disruption rate is given by

$$\dot{M}_{\text{disrupt,AGN}} = 2f_{\text{frag}}f_{k,\text{AGN}}(1 - f_w) \frac{Q_{\text{AGN}}}{\sigma_v^2}. \quad (28)$$

As for SNe, the residual power $1 - f_{\text{frag}}$ is injected into the diffuse gas as turbulent energy, $\Delta E_{\sigma}^{\text{AGN}} = (1 - f_{\text{frag}})f_{k,\text{AGN}}(1 - f_w)Q_{\text{AGN}}\Delta t$. The total mass disruption rate is the sum of the SN and AGN contributions, i.e. $\dot{M}_{\text{disrupt}} = \dot{M}_{\text{disrupt,SN}} + \dot{M}_{\text{disrupt,AGN}}$.

5.4.4. Radiative heating and bolometric luminosity

In the previous two sections we discussed our prescriptions for the kinetic power of SNe and AGN. The non-kinetic fraction of the total power is also divided in two different parts. A fraction $f_{\text{th,SN}}$ ($f_{\text{th,AGN}}$) of the non-kinetic SN (AGN) power is used to heat the ejected gas. These fractions are adjusted so that the average temperature of the ejected gas is between 10^6 and 10^7 K. The residual power $(1 - f_{k,\text{SN}})(1 - f_{\text{th,SN}})$ for SNe is then assumed to be emitted as bolometric luminosity in each galaxy (i.e. the part of the total radiative power that does not go directly into heating the gas).

5.5. Average velocity dispersion and disk scale height

5.5.1. Average velocity dispersion

The gas disruption rates, as given in Eqs. (27) and (28), depend on the average velocity dispersion of the diffuse gas, σ_v . We compute and continuously update the velocity dispersion by taking into account simultaneously the kinetic energy injected by gas accretion, SNe, and AGN.

Our prescription is based on the evolution of the ‘‘turbulent’’ kinetic energy budget E_{σ_v} , as follows:

$$2E_{\sigma_v} = M_{\text{diff}}\sigma_v^2, \quad (29)$$

where σ_v is the 3D gas velocity dispersion. Between two time steps, we assume that

- A fraction, $f_{\text{disp}} = 1/2$, of the turbulent energy is dissipated per orbital time.
- The turbulent energy is increased by $2\Delta E_{\sigma_v}^{\text{acc}} = f_{\text{incr}}\Delta M v_{\text{acc}}^2$, corresponding to a fractional increase, $f_{\text{incr}} = 1/3$, per dynamical time, of the kinetic rotational energy of the newly accreted diffuse gas, ΔM . We assume that the freshly accreted diffuse gas (ΔM) forms a thin rotating disk (e.g. Mo et al. 1998). This thin gas disk merges with the pre-existing disk hosting a mass, M_{diff} , of diffuse gas. The quantity v_{acc} is the orbital velocity of the freshly accreted diffuse gas computed at the half mass radius.
- The total budget of the turbulent energy is finally increased by the energy injected in the diffuse gas by SNe and AGN, $\Delta E_{\sigma_v}^{\text{SN}}$ and $\Delta E_{\sigma_v}^{\text{AGN}}$ (see Sect. 5.4.3).

Considering all of these energy terms, after a time step Δt , we have

$$E_{\sigma_v}^{n+1} = \left[1 - f_{\text{disp}} \left(\frac{\Delta t}{t_{\text{dyn}}}\right)\right] E_{\sigma_v}^n + \Delta E_{\sigma_v}^{\text{acc}} \left(\frac{\Delta t}{t_{\text{dyn}}}\right) + \Delta E_{\sigma_v}^{\text{SN}} + \Delta E_{\sigma_v}^{\text{AGN}}. \quad (30)$$

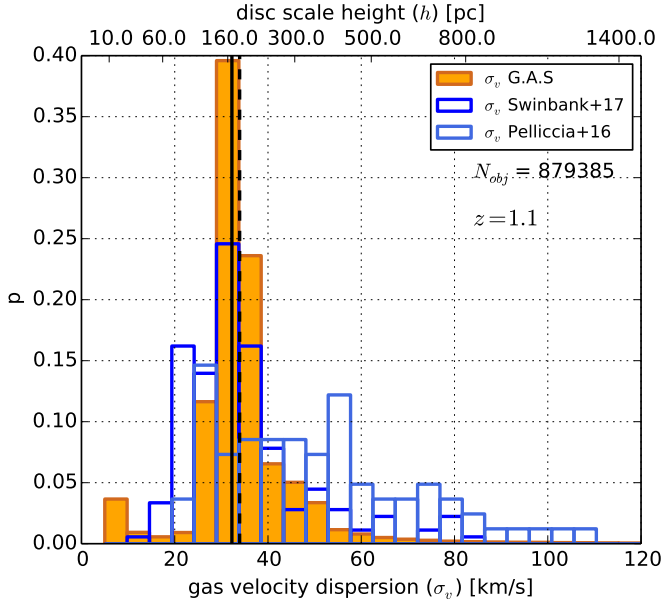


Fig. 10. Probability distribution of the velocity dispersion of the diffuse gas. This distribution represents model galaxies with $M_\star > 10^7 M_\odot$ and at $z = 1.1$. We compare this distribution with observational measurements from Swinbank et al. (2017, dark blue histogram) and Pelliccia et al. (2017, light blue histogram). The solid and dashed vertical lines represent the median and mean value of the distribution of our modelled galaxies, respectively.

Based on this updated total turbulent kinetic energy, we calculate the average velocity dispersion of the diffuse gas σ_v (Eq. (29)). The two free parameters, f_{incr} and f_{disp} , have been adjusted to reproduced observed values of velocity dispersion (see Fig. 10).

During a merger, the total turbulent kinetic energy of the two progenitors are added. We also add a fraction of the gravitational energy due to the interaction between the two galaxies,

$$E_{\text{int}} = G \frac{M_1 M_2}{1.68(r_d^1 + r_d^2)}, \quad (31)$$

where M_i is the mass (baryon + dark matter) included in the galaxy half mass radius and r_d^i is the disk exponential radius (Hatton et al. 2003).

We compare our predictions of the gas velocity dispersion with recent observational measurements of the velocity dispersion of the warm ionised media from Swinbank et al. (2017) and Pelliccia et al. (2017). Objects targeted by these two programmes are distributed in redshift between $z \simeq 0.3$ and $z \simeq 1.7$. Our predictions are in good agreement with these observational results but our results reproduce some of the most extreme dispersions observed (both low and high values; Fig. 10). At $z = 1.1$, the median value of the velocity distribution of the diffuse gas is close to 35 km s^{-1} .

5.5.2. Disk scale height

The disk fragmenting timescale (Eq. (19)) depends on the average disk scale height. We use the disk scale height to define the initial energy injection scale of the inertial turbulent cascade. Our gas fragmenting scenario is mainly based on scaling relations and the disk scale height is crucial in calculating this parameter of the turbulent cascade. The disk scale height h_d is deduced from $\sigma_v = \sigma(h_d^{-1})$. At $z = 1.1$, the median value of

the velocity dispersion leads to a median value of the disk scale height is $\approx 160 \text{ pc}$ (Fig. 10).

6. Thermal instabilities

For galaxies with high masses, previous semi-analytical models invoked powerful AGN feedback to reduce greatly or even completely quench accretion of cooling gas in their haloes (e.g. Cattaneo et al. 2006; Somerville et al. 2008; Guo et al. 2011; Benson 2012; Henriques et al. 2013). Such prescriptions assume a strong coupling between AGN power, the ISM, and the hot circumgalactic gas, which results in powerful outflows expelling the gas and heating the halo gas. However, observations are not entirely clear concerning the precise impact of outflows on distant galaxies (see e.g. Mullaney et al. 2015; Netzer et al. 2014, 2016; Scholtz et al. 2018; Falkendal et al. 2019). In Cousin et al. (2015a,b), we also assume that AGN can have an impact on the hot halo phase but we have limited the effect to simply heating the halo gas.

In the following we present a new prescription to reduce efficiently the gas accretion onto massive galaxies, $M_\star > 10^{11} M_\odot$. This new mechanism is based on the growth of thermal instabilities in the hot halo phase surrounding the galaxy.

6.1. Thermal instabilities: Description

Our new mechanism assumes that the condensation of gas and therefore the gas accretion onto the galaxy is progressively and strongly limited by the growth of thermal instabilities in the hot gaseous CGM. In the standard model of cooling presented previously and also used in, for example Croton et al. (2006), Baugh (2006), and Somerville et al. (2008), there is a strong dichotomy between the gas stored within and beyond the cooling radius, r_{cool} . At radii greater than the cooling radius, the gas is assumed to remain hot, while below r_{cool} , the gas is assumed to be warm, 10^4 K , and dense enough to condense and feed the galaxy. Of course, the reality is more complex, and obviously hot and warm gas co-exist around the cooling radius. In this transition region, the gas can be thermally unstable. These thermal instabilities can generate warm clouds that are orbiting in the hotter gas. Cornuault et al. (2018), in a phenomenological model of accreting gas, have found that the cloud–cloud velocity dispersion can be comparable to the characteristic virial velocity of the dark-matter halo in which they formed. The mixing and dynamics of warm clouds into the hot atmosphere can moderate the effective accretion rate of gas onto a galaxy.

6.1.1. Thermal instability clock

As shown by Sharma et al. (2012), the timescale over which gas becomes thermally unstable is related to the cooling time. Following their approach, and similar to gas cooling generally, we define a thermal instability timescale, \mathcal{T}_{TI} . This timescale evolves concomitantly with the cooling timescale. After each time step, Δt , the thermal instability clock is updated following a mass-weighted prescription, similar to that applied to the cooling clock, namely,

$$\mathcal{T}_{\text{TI}}^n = \underbrace{\left(\mathcal{T}_{\text{TI}}^{n-1} + \Delta t_{\text{TI}} \right)}_{\text{hot gas in the halo}} \left(1 - \frac{\Delta M}{M} \right) + \underbrace{\frac{\Delta t_{\text{TI}}}{2} \frac{\Delta M}{M}}_{\text{freshly accreted hot gas}}, \quad (32)$$

where M is the total mass stored in the hot phase after the last time step, Δt . The quantity ΔM is the mass accreted by the hot

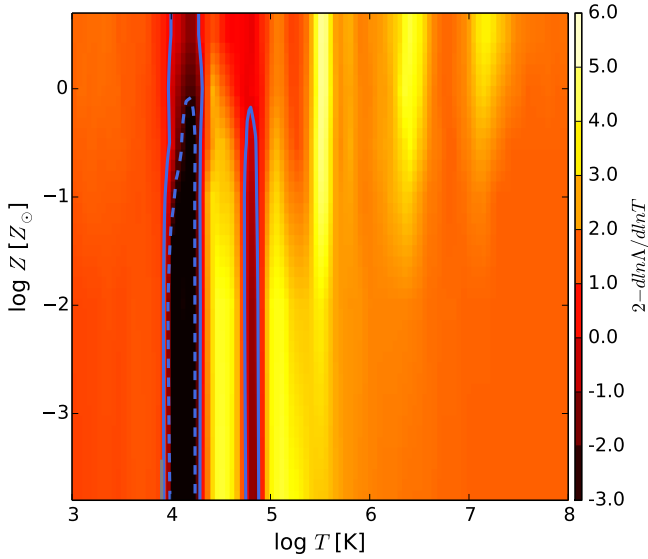


Fig. 11. Thermal instability efficiency, $F(\Lambda)$, as a function of both gas temperature and gas metallicity. The amplitudes of thermal instabilities are computed as given in Sharma et al. (2012) and the values are indicated in the colour bar on the right side of the plot. When $F(\Lambda) < 0$, the gas is thermally unstable. The solid and dashed blue contours indicate where $2 - d \ln \Lambda / d \ln T = 0$ and -3.0 , respectively.

gas phase or ejected from the galaxy. For each time step, Δt , the time increment Δt_{TI} is calculated as

$$\Delta t_{\text{TI}} = 0.1 F(\Lambda) \Delta t \quad (33)$$

where $F(\Lambda)$ is the thermal instability function (Sharma et al. 2012), depending on the logarithmic derivative of the cooling efficiency function Λ ,

$$F(\Lambda) = 2 - \frac{d \ln \Lambda}{d \ln T}. \quad (34)$$

We assume that the hot gas becomes thermally unstable during all the previous time steps (Eq. (32)). However, as for the effective cooling clock, we consider that thermal instabilities can evolve with the addition of freshly accreted gas only during half a time step⁴. The thermal instability function, $F(\Lambda)$, determines if the hot gas is stable or unstable. Gas is assumed to be unstable if $F(\Lambda) \geq 0$ and stable if $F(\Lambda) < 0$. When hot gas is stable, the time increment Δt_{TI} is set to 0. According to the pre-existing/incoming mass ratio, the thermal instability clock can increase or decrease during a given time step.

As for the cooling efficiency, the thermal instability function has been interpolated and tabulated. Figure 11 shows the thermal instability function dependence on gas temperature between 10^3 K and 10^8 K and gas metallicity between $10^{-4} Z_{\odot}$ and $2 Z_{\odot}$. As for the effective cooling clock, the thermal instability clock of the hot phase is assumed to have the value that the most massive progenitor had prior to merging.

6.1.2. Mixing zone

The thermal instability clock provides at any time the effective timescale of the growth of thermal instabilities. Starting at the

⁴ Following a scheme where gas is continuously added into the hot gas reservoir.

cooling radius, we assume that thermal instabilities are propagating through the gas and ultimately reach the centre of the halo. We define the instantaneous size of the mixing zone as

$$\Delta r_{\text{TI}}^n = \varepsilon_{\text{TI}} t_{\text{TI}}^n \sqrt{\frac{\gamma R T_{\text{hot}}}{\mu}}. \quad (35)$$

We assume an adiabatic index $\gamma = 1.4$ and a mean molecular mass $\mu = 0.62$. R is the specific ideal gas constant and T_{hot} is the average temperature of the hot gas. We assume that the propagation speed of these instabilities is close to the sound speed of the gas. Following this hypothesis, the value of the free parameter ε_{TI} is set to 0.63.

In the mixing zone, we assume that thermal instabilities lead to the formation of warm gas clouds orbiting around the galaxy. As already stated, the cloud–cloud velocity dispersion is assumed to be close to the virial velocity (Cornuault et al. 2018). In addition, even if we do not take into account explicitly such a mechanism, the warm gas clouds probably interact strongly with the large-scale wind escaping the galaxy. Through this interaction, we can reasonably assume that kinetic energy of the outflowing gas is injected into a dynamic and cloudy CGM. Thus we assumed that the condensation of warm clouds is suppressed in the mixing zone. The condensation of clouds is only efficient in the inner region of the hot atmosphere and the effective cooling rate is then computed using Eq. (7). For this effective cooling rate, the initial cooling radius calculated is substituted with the expression, $r_{\text{TI}} = r_{\text{cool}} - \Delta r_{\text{TI}}$.

6.2. Impact of thermal instabilities on the gas cooling

To understand and quantify the progression of thermal instabilities in the hot gas phase, we define the thermal instability volume filling factor as

$$\phi_v = 1.0 - \left(\frac{r_{\text{TI}}}{r_{\text{cool}}} \right)^3. \quad (36)$$

The parameter ϕ_v measures the volume fraction in which thermal instabilities are developed; $\phi_v = 0$ indicates that there is no impact on the gas cooling. As in the original model, all the gas enclosed in the cooling radius can condense and therefore feed the galaxy. The case $\phi_v = 1$ indicates that the entire region with $r < r_{\text{cool}}$ is thermally unstable. When $\phi_v = 1$, since the clouds are orbiting at the virial velocity, is the point at which the cooling flow would otherwise occur is effectively stopped.

Figure 12 shows the mean trend of ϕ_v (Eq. (36)) as a function of both the dark-matter halo mass and redshift. Depending on the redshift, the first resolved haloes, $M_{\text{vir}} > 10^9 M_{\odot}$, have ϕ_v distributed between 0.3 and 0.8. At $z \geq 9.0$, the average ϕ_v of the first haloes formed is close to 0.6. This value progressively decreases with the redshift. At $z \approx 4.0$, the average ϕ_v reaches a minimum around 0.3. Then the average ϕ_v of first haloes detected increases as the redshift decreases. The maximum value, 0.8, is reached at low redshift. In this low-mass regime, hot atmospheres are only formed via large-scale winds coming from the galaxy. The average ϕ_v is closely linked to galaxies driving winds. Indeed, as mentioned in Sect. 6.1.1, the effective thermal instability (TI) clock runs in proportion to the freshly added to pre-existing gas mass ratio in the hot halo phase (Eq. (32)). The higher the proportion of freshly added ejecta, the slower the TI clock advances and thus the slower ϕ_v increases. Between $z \approx 9.0$ and $z = 4.0$, the intensity of ejecta affecting the first resolved haloes, $M_{\text{vir}} \approx 10^9 M_{\odot}$, increases continuously

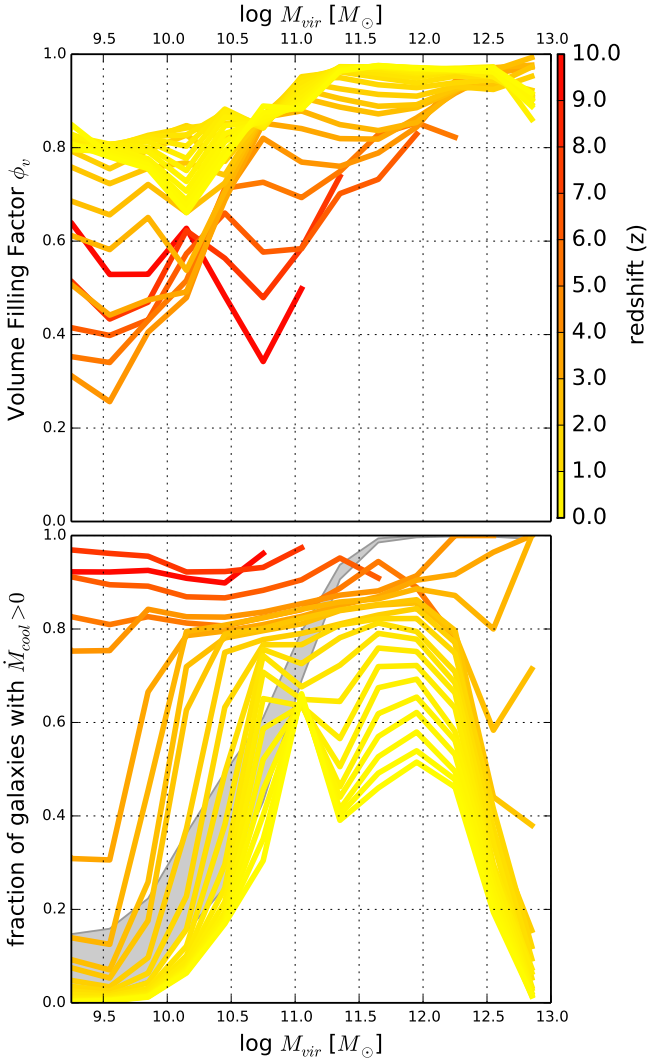


Fig. 12. *Upper panel:* mean value of the volume filling factor, ϕ , of gas that is thermally unstable (Eq. (36)). The case $\phi_v = 1$ implies that all the region within $r = r_{\text{cool}}$ is thermally unstable in our model. *Lower panel:* fraction of galaxies that are accreting radiatively cooling halo gas as a function of halo virial mass, M_{vir} . Different colours indicate this fraction as a function of redshift. The value of the redshift for each line is indicated in the colour bar on the right side of the upper panel. The grey area spans the range of the mean fraction of galaxies that are accreting hot gas for each halo mass summed over redshifts 0.1–10 (Eq. (2)). For example, for a halo mass of $10^{10.5} M_{\odot}$, the fraction of galaxies that have hot gas in the halo which is radiatively cooling can range between 30% and ~55% depending on redshift.

(related to both the gas accretion and the SFR). Then the intensity progressively decreases until the lowest redshifts. We caution that the number of haloes in the low-mass bins at very high redshift, $z \approx 9$, is low and thus these bins are greatly affected by statistical noise.

Between the redshift range, $z = 4.0$ to $z = 1.5$, the volume filling factor of the most massive haloes reaches very high ϕ_v , ≈ 0.95 . At $z < 1.0$, the average volume filling factor of thermally unstable gas in massive haloes starts to decrease. In these massive haloes, some accretion still occurs. This accretion produces a new star formation and therefore some large-scale ejection events. In these massive haloes we therefore observe an increase of freshly acquired hot halo gas. This increase in freshly acquired

hot halo gas results in a progressive decrease of the volume filling factor of thermally unstable gas.

Figure 12 also shows the change in the dark-matter halo mass and the evolution with redshift of the fraction of galaxies in which condensation of the hot halo gas is occurring, i.e. $\dot{M}_{\text{cool}} > 0$. (Eq. (7)). At the halo detection mass threshold of our model, radiatively cooling gas is accreting over the whole mass range. The higher the redshift, the higher the fraction of galaxies that accrete radiatively cooled hot halo gas. For $z > 5.0$, the fraction remains high, > 0.8 . Then this fraction sharply decreases and reaches values lower than 0.1 at $z < 3.0$. At low virial halo masses, the hot gas phase is only fed by large-scale ejecta from galactic outflows ($f_{\text{hot}} \leq 0.18$). In low virial-mass haloes, the mass of hot halo gas is small, $M_{\text{hot}} \leq 10^7 M_{\odot}$. The cold mode is the dominant accretion mode at low masses. Even if at $z > 5.0$ radiative cooling produces an effective accretion in a large fraction of the galaxies, the average radiative cooling rate is less than $1 M_{\odot} \text{ yr}^{-1}$ (Fig. 3). The fraction of dark-matter haloes hosting actively and radiatively cooling hot gas increases (or remains constant) as a function of the dark-matter virial mass. The largest population of actively cooling hot haloes is reached between $\approx 10^{10.25} M_{\odot}$ and $\approx 10^{12} M_{\odot}$. Between $z = 4.0$ and $z = 3.0$, in the most massive haloes $> 10^{12} M_{\odot}$ the fraction reaches 1.0. However by $z < 3.0$ for the most massive haloes, we find a strong decrease in the fraction of galaxies with actively cooling haloes. This decrease is directly linked to the high ϕ_v reached in these haloes and this decrease is responsible for the reduction in the effective cooling rate estimated for the most massive haloes at $z < 2$ (Fig. 3).

7. Results and analysis

7.1. Stellar mass assembly and co-moving density of galaxies

We compare the predictions of the G.A.S. model for the co-moving stellar mass density as a function of redshift with a compilation of observations (Fig. 13). Overall, the stellar mass growth of the Universe predicted in our model agree with observations. These predictions are in reasonable agreement with estimates even at the highest redshifts, $z > 3$, although they are generally consistent with the highest total co-moving stellar mass density estimates. Furthermore, this general agreement between G.A.S. and observations is also reflected in a comparison of the co-moving number density of galaxies binned by redshift (Fig. 14). We compare our results with a variety of observed distributions covering the redshift range $z \approx 0.1$ – $z \approx 6$ (Baldry et al. 2008; Yang et al. 2009; Caputi et al. 2011; Ilbert et al. 2013; Grazian et al. 2015; Duncan & Conselice 2015; Song et al. 2016; Davidzon et al. 2016). In addition, we compare our best G.A.S. predictions for the co-moving number density with those of Guo et al. (2011) and also those obtained with the previous ad hoc regulation prescriptions of Cousin et al. (2015a). These comparisons allow us to gauge the impact of including the star formation and gas-dissipation prescriptions that are in G.A.S. with models that do not include such prescriptions (Fig. 15).

Between $z \approx 1.0$ and $z \approx 6.0$, our model agrees very well with observational measurements of both the cosmic co-moving mass densities and shape of the co-moving number density of galaxies for a wide range of redshifts (cf. Figs. 13 and 14). At $z < 1.0$, our model systematically under-predicts the co-moving density of intermediate-mass galaxies and over-predicts the co-moving density of low-mass galaxies. We find that in the low- and intermediate-mass range, $10^{8.5} \leq M_{\star} \leq 10^{11.5} M_{\odot}$, the

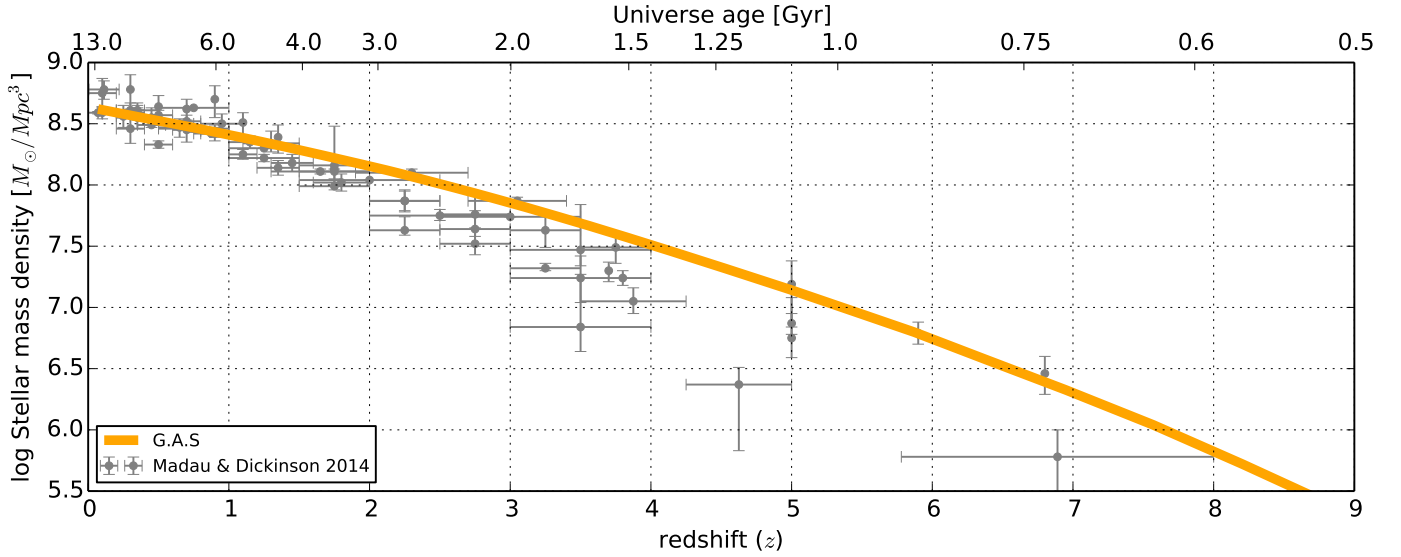


Fig. 13. Redshift evolution of the co-moving stellar mass volume density. The solid orange line is the prediction of the G.A.S. model. We compare our predictions with the compilation of observational results (see Madau & Dickinson 2014, and references therein).

co-moving density functions predicted by our model show a double power-law shape, with a steeper slope at the low-mass end of the model distribution. At the highest redshifts, $z \approx 6.0$, the density function keeps this shape even for the most massive model galaxies, $M_{\star} \geq 10^{11.5} M_{\odot}$. This power-law shape is the signature of the continuous competition between the disruption and progressive fragmentation of the gas through the turbulent cascade. Some galaxies, those with stellar masses above $10^{10.5} M_{\odot}$, have already substantially formed at $z \approx 6.0$. The turbulent cascade regulates the star formation in intermediate- and low-mass objects and keeps a sufficient amount of diffuse and fragmented non-star-forming gas that actively feeds star formation until the formation of the most massive observed galaxies.

As the redshift decreases, the power-law shape is broken at the high-mass end of the distribution and an exponentially declining function is progressively formed. In agreement with observational estimates, the knee of our predicted co-moving number density functions develops around $10^{10.75} M_{\odot}$ at $z \approx 5.0$ and progressively shifts to $10^{11.0} M_{\odot}$ by $z \approx 0.1$. The break appears when the star formation is strongly reduced or quenched, i.e. when all the fresh gas contained in the disk is consumed. At the high-mass end of the galaxy distribution, gas is accreted via the hot mode (Sect. 2.2 and Fig. 3). In, for example, Guo et al. (2011), Croton et al. (2006), or Somerville et al. (2008), the reduction of the gas accretion is because of the power of the AGN heating the halo gas and halting accretion. To obtain the necessary heating, such prescriptions need an increasing and constant production of power from AGN. In G.A.S., much of the reduction or halting of gas accretion is a result of the increase in the thermally unstable hot gas phase surrounding galaxies. As shown in Fig. 16, the impact of thermal instabilities appears to be progressive and its impact starts to become significant at $z \sim 4.0$ for the most massive galaxies. The effect is clearly visible for the galaxy mass function at $z = 2.1$. Without such regulation, the continuous accretion of gas leads to the formation of very massive, $10^{12.0} M_{\odot}$, galaxies at $z \approx 1.5$. At very high masses, the reduction of the gas accretion onto galaxies is not sufficient to quench completely the star formation. Indeed, some galaxies have stellar masses more than $10^{11.5} M_{\odot}$ are predicted in our model but are not observed (e.g. Davidson et al.

2016; Ilbert et al. 2013; Yang et al. 2009; Baldry et al. 2008). This is due to a progressive decrease of the volume filling factor of thermal instabilities in the most massive galaxies (Fig. 12). A slight increase of the thermal instabilities propagation efficiency, ϵ_{TI} , could limit the growth of these very massive model galaxies.

Compared to the ad hoc recipe ($\propto M_{\text{h}}^3$) presented previously in Cousin et al. (2015a), the new regulation prescription based on the turbulent cascade is in better agreement with the co-moving densities of low-mass galaxies. The ad hoc recipe was clearly not appropriate for capturing the necessary ingredients to produce sufficient numbers of low-mass galaxies and especially those at high redshifts (Fig. 15). On the contrary, the stellar mass functions predicted by Guo et al. (2011) clearly over-produce the co-moving density of low- or intermediate-mass galaxies, $M_{\star} < 10^{10.5} M_{\odot}$. Their prescriptions for regulating star formation do not have the correct dependencies as a function of the stellar mass. Gas is converted to stars too efficiently in low- or intermediate-mass galaxies, hence producing an excess. Furthermore, a lack of gas in more massive galaxies means that their models underestimate the number of massive galaxies.

In fact, to elucidate the role played by various physical prescriptions we now include in G.A.S., we compare four different model configurations: our best complete model and three alternative versions in which: (i) the gas re-accretion prescription is disabled; (ii) thermal instabilities are not considered; and, (iii) the photoionisation prescription is switched off. These three alternative versions allow us to determine the effective impact of those prescriptions in regulating the mass growth of the ensemble of model galaxies (Fig. 16). At the highest redshift for which we made this comparison, $z = 4$, we see that turning off all three of these prescriptions has little impact on the modelled stellar mass distribution. As we decrease in redshift, we see progressively greater impact for these three processes on the stellar mass distribution. The lack of thermal instabilities in the hot gas in the halo results in overly massive galaxies by $z \approx 2-3$, but leaves essentially unchanged the number of lower mass galaxies, i.e. those with $M_{\star} \lesssim 10^{11} M_{\odot}$. Photoionisation appears to only impact the co-moving number density of galaxies at low redshift and for galaxies with low stellar masses. At

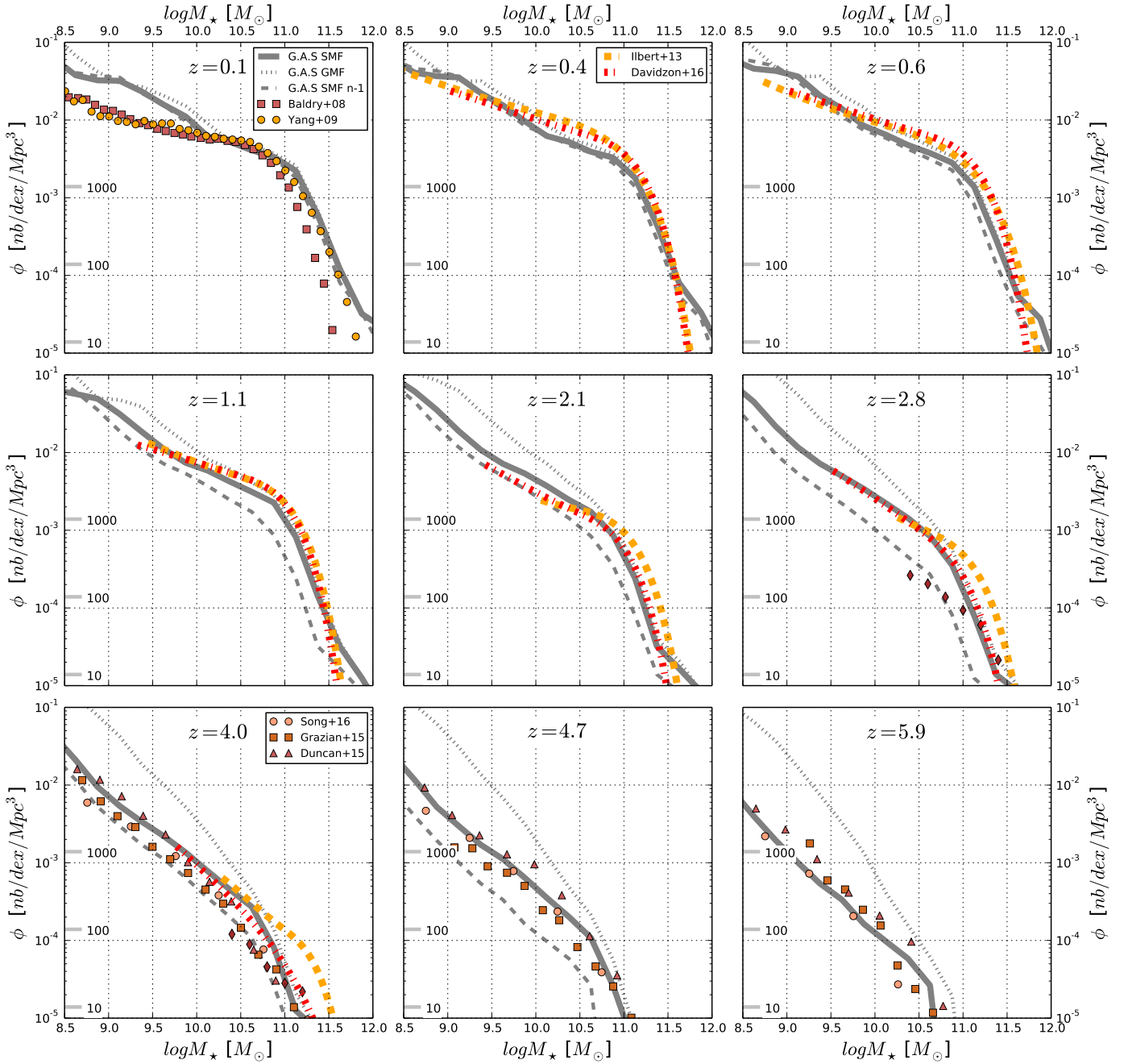


Fig. 14. Comparison between the predictions of G.A.S. with observational estimates of the co-moving number density of galaxies for 9 redshift bins spanning the range, $z \approx 0.1$ – $z \approx 6$ (from upper left to lower right and as indicated at the top of each panel). We compare the predictions of the G.A.S. model (solid grey lines) with various studies: Baldry et al. (2008, red squares), Yang et al. (2009, orange circles), Caputi et al. (2011, brown diamonds), Grazian et al. (2015, chocolate squares), Duncan & Conselice (2015, chestnut-red triangles), Song et al. (2016, salmon circles), Ilbert et al. (2013, yellow dashed-line), and Davidzon et al. (2016, red dot-dashed line). For this comparison with observations, we indicate binned measurements using points and best-fit stellar mass functions as dashed lines (Schechter and double-Schechter). All stellar mass functions are for a Chabrier IMF. The best fits of Ilbert et al. (2013) and Davidzon et al. (2016) are corrected for the Eddington bias which affects the high-mass galaxy bins in particular. The grey dotted-line indicates the total gas plus stars, baryonic mass function predicted by G.A.S. The grey dashed-line indicates the stellar mass function of the previous panel for the next highest redshift bin. The tick marks placed on the left side of each panel indicate the number of galaxies used in the stellar mass function predicted by the G.A.S. model.

$z < 1$, the power-law shape in the co-moving density of low-mass galaxies, i.e. those with $M_* < 10^{8.5} M_\odot$, is broken and becomes flatter as the redshift decreases. This trend is a result of a progressive reduction in the effective gas accretion with decreasing redshift and is a direct consequence of the photoionisation prescription (Sect. 2.2, Fig. 1). The process with the largest impact

on the co-moving number density of galaxies appears to be the re-accretion of gas (Sect. 5.4.2). Its impact on the galaxy co-moving density distribution starts to become evident at $z \approx 3$ (Fig. 16). Without re-accretion, the model strongly underpredicts the co-moving density of almost all stellar masses but especially so for intermediate-mass galaxies.

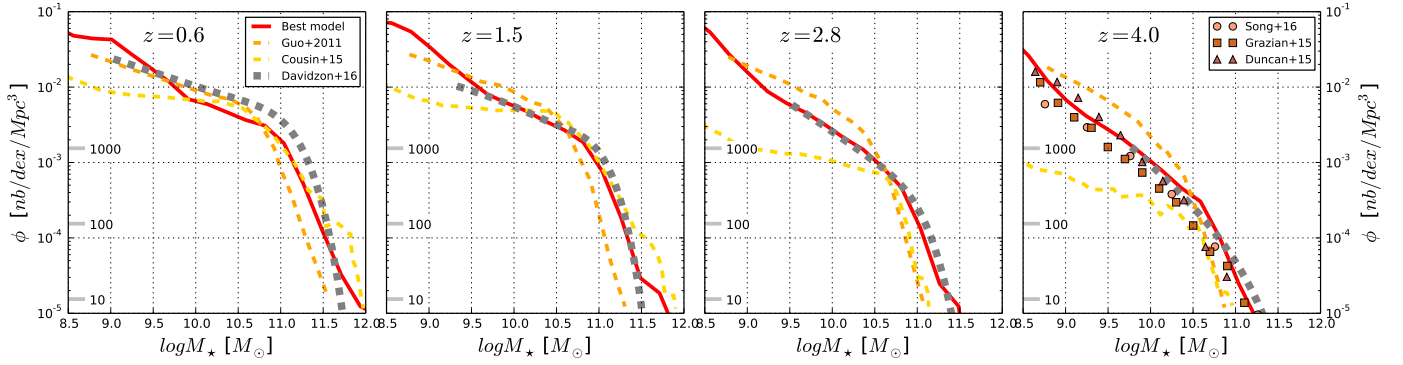


Fig. 15. Comparison of the stellar mass functions predicted by our best fiducial model with those predicted by Guo et al. (2011). Each panel shows a different redshift as indicated at the upper left in each panel; solid red, dashed orange, and dashed gold lines indicate the predictions of G. A. S., Guo et al. (2011), and Cousin et al. (2015a), respectively (as indicated in the legend of the left-most panel). We compare these models with observations (dashed grey lines; Davidzon et al. 2016). At $z = 4.0$, we also plot additional observational measurements (salmon circles, red upward pointing triangles, and orange squares; Song et al. 2016; Duncan & Conselice 2015; Grazian et al. 2015, respectively, and as indicated in the legend of the right-most panel)

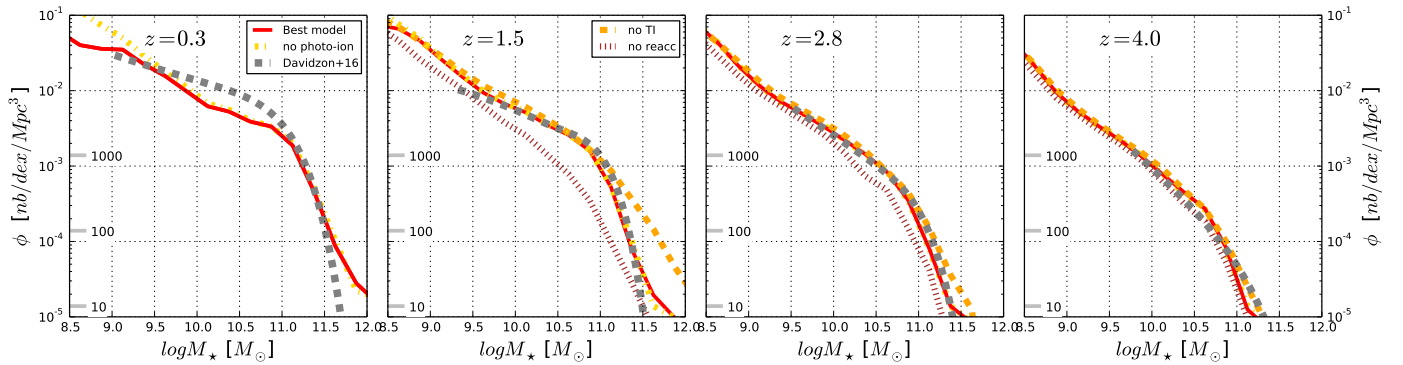


Fig. 16. Impact of our main prescriptions for regulating star formation and gas accretion, photoionisation, re-accretion, and thermal instabilities on the stellar mass assembly of model galaxies (i.e. the co-moving density of model galaxies as a function of redshift). The solid red line shows our fiducial model using the full suite of prescriptions. The dot-dashed gold lines indicate our model without photoionisation. In the three right-most panels, the dotted brown lines and dashed orange lines indicate predictions of the stellar mass functions without including the prescriptions for re-accretion of gas over two halo crossing times and thermal instabilities, respectively (as indicated in the legends of the two left-most panels). These are compared to the observational estimates of the co-moving density of galaxies from Davidzon et al. (2016). The four different panels represent four different redshift bins, $z = 0.3, 1.5, 2.8,$ and 4 (from the left-most to the right-most panel, respectively).

7.2. Evolution of disk properties

We discuss with Fig. 17 the evolution with both the stellar mass and redshift of five main properties of star-forming galaxies. We present the evolution of the following disk properties:

- The gas mass fraction, $f_{\text{gas}} = \frac{M_{\text{gas}}}{M_{\text{gas}} + M_{\star}}$. The value M_{gas} is the total mass of the three gas reservoirs combined for a given disk galaxy.
- The fragmented gas fraction, $f_{\text{frag}} = \frac{M_{\text{frag}} + M_{\text{sfg}}}{M_{\text{frag}} + M_{\text{sfg}} + M_{\text{diff}}}$.
- The galaxy half mass radius, r_{50} .
- The diffuse gas density, ρ_{diff} . This density is computed based on the mass of diffuse gas, a disk with an external radius of $22 \times r_d$ and a scale height, h_d . We also assume a mean molecular mass, $\mu = 0.62$.
- The ratio of the gas velocity dispersion to orbital velocity, $\frac{\sigma_v}{v}$. The dispersion velocity is that of the diffuse gas. The orbital velocity is computed at $2.2r_d$ (Pelliccia et al. 2017).

The following analysis is based on model star-forming galaxies that have been extracted at a variety of redshifts. We define star forming as those galaxies which lie above 25% of the mean of the relation between the star formation rate and stellar mass (the “main sequence” of star-forming galaxies Schreiber et al. 2015,

their Eq. (9)). Within this sample of model galaxies, we calculate median of the ensemble binned in redshift and stellar mass for each of the quantities listed above. We first focus on the gas mass fraction, f_{gas} . At a given stellar mass, the highest redshift galaxies have the highest gas fractions while galaxies with larger stellar masses have lower gas fractions. If we consider both the increase in mass with decreasing redshift, these tracks imply that the gas mass fraction globally decreases as galaxies grow and evolve. At low redshift, $z < 0.5$, the gas mass fractions of star-forming galaxies lie between 15%–50% depending on the stellar mass. In two samples of star-forming galaxies, evolving around $z = 1.5$, Daddi et al. (2010) measured a gas mass fraction higher than our predictions. At $z = 1.5$, for an average stellar mass of $M_{\star} = 10^{10.5} M_{\odot}$, the gas mass fraction is distributed as [p15, p50, p85] = [0.18, 0.25, 0.36] in comparison to $f_{\text{gas}} \approx 0.6$ measured by Daddi et al. (2010). In the same redshift slice, for an average stellar mass of $M_{\star} = 10^{11} M_{\odot}$ we measured [p15, p50, p85] = [0.11, 0.16, 0.26] in comparison to $f_{\text{gas}} \approx 0.5$ measured by Daddi et al. (2010).

In parallel to the progressive decrease of the gas-mass fraction, the distribution of this gas between the diffuse and the fragmented/dense gas also evolves. We first note that at a given stellar mass, the fragmented gas mass fraction is an increasing

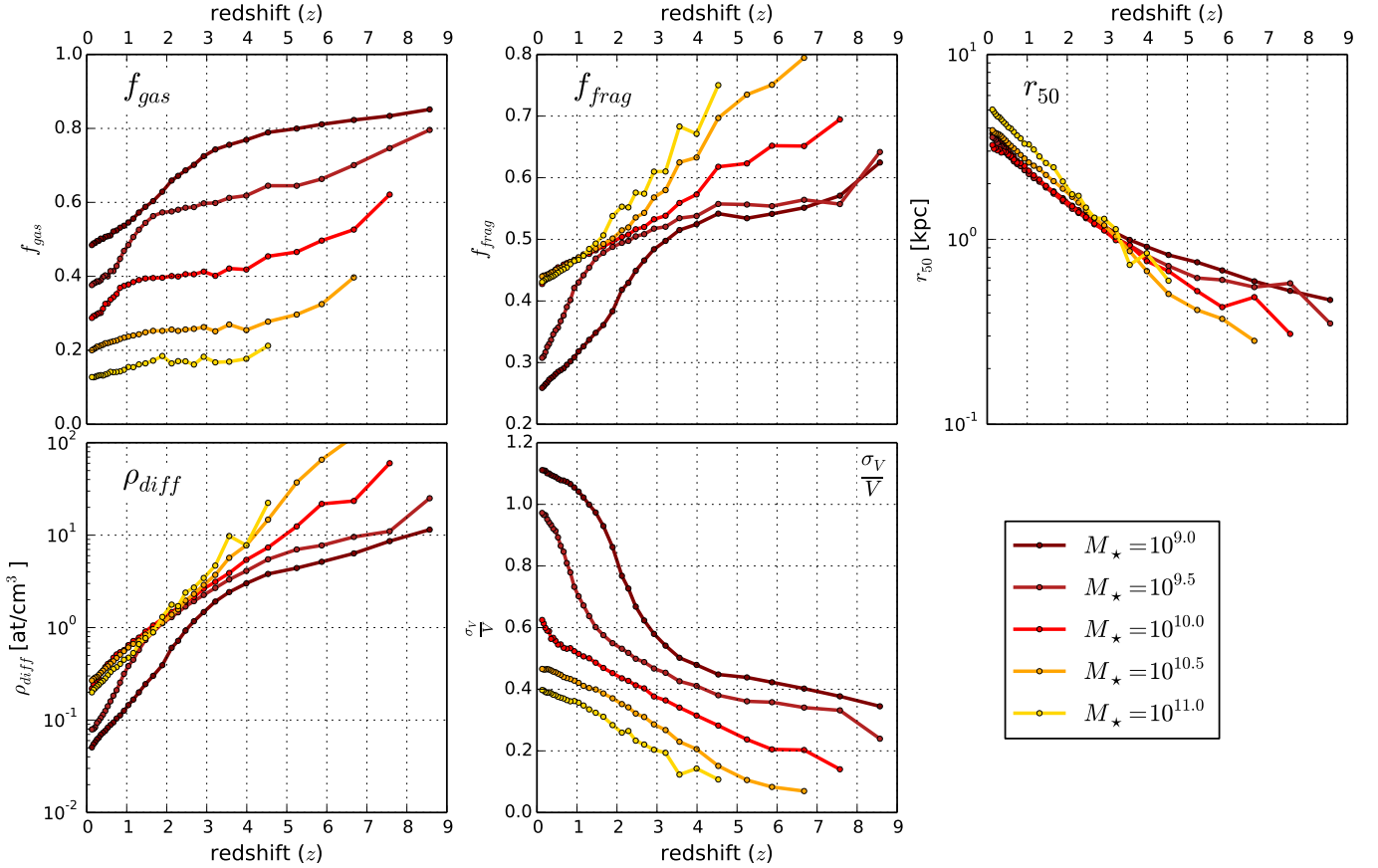


Fig. 17. Evolution with redshift of 5 main disk properties of disk galaxies which lie along the star-forming galaxy main sequence shown for a range of mass bins. Legend in the *lower right* gives the stellar mass values for each curve in the panels. *Upper left*: gas mass fraction, f_{gas} , as a function of redshift for 5 stellar mass bins. *Upper centre*: fraction of the gas which is fragmented, f_{frag} . *Upper right*: average half mass radius, r_{50} , of the ensemble of star-forming model galaxies. *Lower left*: diffuse gas density, ρ_{diff} . *Lower right*: ratio of the gas velocity dispersion to orbital velocity, $\frac{\sigma_v}{V}$.

function of the redshift. Galaxies formed at higher redshift contain a higher percentage of fragmented gas than galaxies evolving at lower redshift with a similar mass. At high redshift $z > 2.5$, the fragmented gas mass fraction is a clear decreasing function of the stellar mass. The hierarchy in mass is less clear at lower redshifts. We note that galaxies hosting a stellar mass higher than $10^{10} M_{\odot}$ stabilise their fragmented gas mass fraction at around 42%. For less massive galaxies, $<10^{10} M_{\odot}$, the fragmented gas fraction strongly decreases and reaches low values distributed between 25% and 33% at $z = 0.1$.

The gas content clearly evolves with time: gas-rich and strongly fragmented galaxies evolve through structures dominated by stars and in which the gas is mainly diffuse. This evolution in gaseous and stellar contents is following the evolution of galaxy size: At a given stellar mass, the disk half mass radius is a decreasing function of the redshift. At redshift $z > 3.0$, the disk half mass radius is a decreasing function of the stellar mass. Massive galaxies appear more compact than low-mass galaxies. For all stellar mass bin analysed, the average half mass radius is strictly lower than 1 kpc. At $z < 3.0$, we note the opposite trend: the average half mass radius is strictly larger than 1 kpc and is an increasing function of the stellar mass.

The anti-correlated evolution of the gas content and disk size directly impacts the density of the diffuse gas, which is the main driver of the GMC feeding rate (Sect. 4.1, Eq. (14)). At a given stellar mass, the diffuse gas density appears to be an increasing

function of the redshift. In galaxies formed at higher redshift, the gas is denser than in galaxies of a same stellar mass formed at lower redshift. At a $z > 3$, the diffuse gas density is an increasing function of the stellar mass. At $z < 3$, the most massive galaxies see their diffuse gas density converge to $0.2\text{--}0.3 \text{ atoms cm}^{-3}$. The less massive galaxies ($\leq 10^{10} M_{\odot}$) see their diffuse gas density strongly decrease and reach a value of $<0.1 \text{ atoms cm}^{-3}$. This evolution is mainly linked to the decrease of the fragmented gas-mass fraction discussed previously.

This decrease also impacts the dispersion to orbital velocities ratio, $\frac{\sigma_v}{V}$. This ratio appears to be a decreasing function of both the redshift and stellar mass. At $z \leq 3.0$, we measure a strong increase of $\frac{\sigma_v}{V}$ in the two lowest stellar mass bins. This strong evolution observed at low redshift is linked to the injection of kinetic energy by the latest-formed SN populations in a gas that becomes rare and diffuse. We recall that the velocity dispersion is computed at the disk scale height and takes into account only the diffuse gas mass. Less gas means a turbulent energy budget (Sect. 5.4.3) divided into a smaller number of gas atoms and therefore a higher average velocity per mass unit. In parallel, the decrease with the redshift of the ratio $\frac{\sigma_v}{V}$ is also due to the increase with the redshift of the orbital velocity traced by the decrease of the average disk orbital timescale (see upper left panel of Fig. 18). By tracing average evolutions from low mass at high redshift to high mass at low redshift, the average galaxy trends indicate that the $\frac{\sigma_v}{V}$ ratio slightly decreases through the evolution.

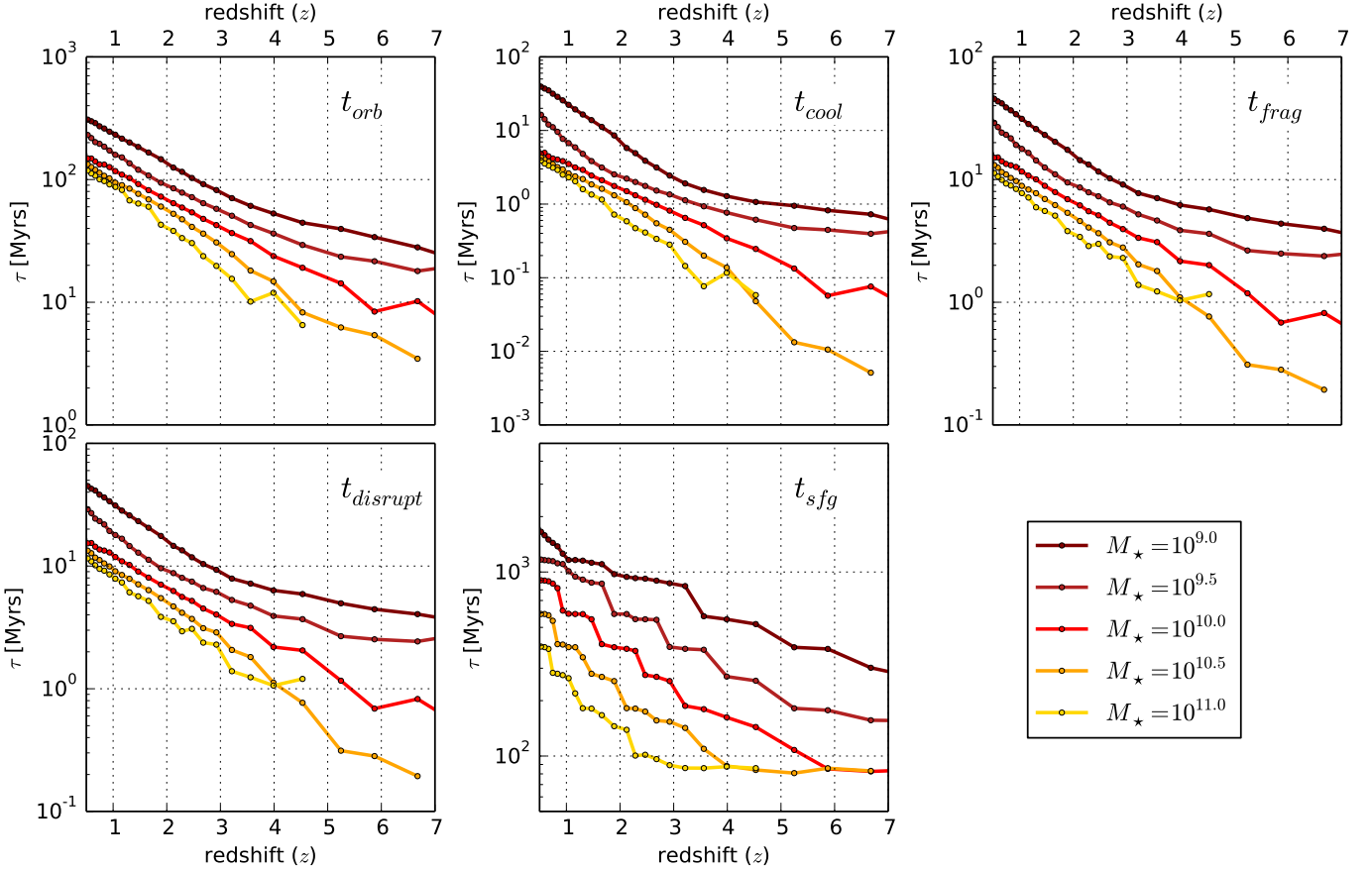


Fig. 18. Evolution with redshift of 5 main timescales for regulating the growth of disk galaxies which lie along the star-forming galaxy main sequence shown for a range of redshift bins. The legend in the *lower right panel* gives the stellar mass value for each curve in the panels. *Upper left*: orbital timescale, t_{orb} . *Upper centre*: characteristic cooling timescale, t_{cool} . *Upper right*: fragmentation timescale of the gas, t_{frag} . *Lower left*: gas disruption timescale, t_{disrupt} . *Lower right*: gas depletion timescale, t_{sfg} .

7.3. Timescales

The gas cycle presented in Sect. 4 is based on three gas reservoirs and different exchange rates between these reservoirs. From these transfer rates and knowing the mass in each of these gas reservoirs, we can deduce the various timescales. Evolution with both the redshift and the stellar mass of five main timescales are shown in Fig. 18. We focus on

- the orbital time t_{orb} estimated at a radius, $2.2r_d$ (e.g. Pelliccia et al. 2017);
- the cooling timescale, t_{cool} , given by Eq. (4);
- the fragmentation timescale, $t_{\text{frag}} = \frac{M_{\text{frag}} + M_{\text{sfg}}}{M_{\text{frag}}}$, defined as the time required to double the fragmented gas mass;
- the disruption timescale, $t_{\text{disrupt}} = \frac{M_{\text{frag}} + M_{\text{sfg}}}{M_{\text{disrupt}}}$, defined as the time required to deplete the fragmented gas via the kinetic energy provided by SNe and/or AGN; and
- the gas depletion time of the fragmented gas, $t_{\text{sfg}} = \frac{M_{\text{frag}}}{M_{\text{sfg}}}$, defined as the time required to convert all the fragmented gas and star-forming gas into stars. As the star formation timescale is shorter than all other timescales, t_{sfg} , is essentially the time required to convert all of the fragmented gas into stars.

We present details of the model output by providing the medians and the 15% and 85% percentiles of these characteristic timescales (Table 3). We compile the statistics for both the full and star-forming samples of model galaxies at $z = 2.1$. These

estimates are given in the four different stellar mass bins 10^9 , 10^{10} , $10^{10.5}$, and $10^{11.25} M_{\odot}$. Even if the median trends presented in Fig. 18 suggest some regularity in galaxy properties as a function of mass, all timescales at constant mass have significant scatter (see Table 3). This large scatter indicates that the dynamics for any one galaxy or ensemble of galaxies is not regular or simple, but is a complex interaction of the various processes included in the model.

First, we first focus on the disk orbital timescale which plays a key role in controlling processes that occur over long timescales. Depending on redshift and stellar mass, the orbital timescale ranges between ≈ 4 Myr and ≈ 300 Myr. At fixed stellar mass, the orbital timescale increases with the redshift. At fixed redshift, it is a decreasing function of the stellar mass. The range of orbital timescales calculated using the G.A.S. model are in good agreement with orbital timescales estimated for local galaxies (e.g. Kennicutt 1998; Leroy et al. 2008; Colombo et al. 2018).

The fragmentation of the ISM is driven by energy injection on scales of and larger than the disk height via the condensation of the diffuse gas. The rate of fragmentation depends on the radiative cooling rate of the diffuse warm gas. The characteristic cooling time is dependent on redshift and galaxy mass and ranges over five orders of magnitude (Eq. (4)). At a given stellar mass, the characteristic cooling timescale is a decreasing function of the redshift (Fig. 18). This trend is driven by the increase with the redshift of the diffuse gas density. At a given

Table 3. Statistics of the main timescales in the gas cycle.

	$10^{9.0} \pm 0.2 M_{\odot}$	$10^{10.0} \pm 0.2 M_{\odot}$	$10^{10.25} \pm 0.2 M_{\odot}$	$10^{10.75} \pm 0.2 M_{\odot}$	$10^{11.25} \pm 0.2 M_{\odot}$
Full sample					
259090	11.1%	2.0%	1.4%	0.6%	0.0%
Timescales	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]
t_{orb} (Myr)	[71.48, 132.42, 210.06]	[32.38, 67.23, 118.72]	[29.12, 61.81, 108.64]	[17.38, 45.52, 100.14]	[16.63, 36.94, 106.27]
t_{cool} (Myr)	[2.281, 7.913, 28.065]	[0.616, 1.791, 5.333]	[0.527, 1.527, 4.041]	[0.263, 1.140, 4.030]	[0.326, 1.325, 5.042]
t_{frag} (Myr)	[7.929, 16.453, 31.759]	[3.152, 6.811, 13.585]	[2.684, 5.865, 11.290]	[1.587, 4.311, 9.970]	[1.790, 3.883, 11.050]
t_{disrupt} (Myr)	[7.732, 15.892, 28.138]	[3.111, 6.672, 12.628]	[2.642, 5.874, 10.966]	[1.481, 4.328, 9.919]	[1.661, 3.733, 10.815]
t_{sfg} (Myr)	[864.5, 974.5, 1577.5]	[255.8, 390.2, 903.2]	[181.0, 280.2, 615.8]	[139.5, 182.3, 878.4]	[91.7, 181.1, 897.2]
Ratios	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]
$t_{\text{cool}}/t_{\text{orb}}$	[0.029, 0.055, 0.143]	[0.018, 0.025, 0.042]	[0.017, 0.023, 0.035]	[0.012, 0.022, 0.042]	[0.014, 0.029, 0.058]
$\mathcal{T}_{\text{cool}}/t_{\text{cool}}$	[1.781, 2.694, 3.556]	[2.982, 3.488, 3.746]	[3.171, 3.472, 3.757]	[3.058, 3.445, 4.199]	[2.963, 3.312, 4.314]
$t_{\text{frag}}/t_{\text{orb}}$	[0.105, 0.118, 0.151]	[0.088, 0.096, 0.115]	[0.083, 0.092, 0.107]	[0.075, 0.088, 0.113]	[0.075, 0.093, 0.124]
$t_{\text{sfg}}/t_{\text{orb}}$	[5.48, 8.30, 18.55]	[4.13, 6.55, 13.47]	[3.36, 5.36, 11.69]	[2.57, 5.36, 18.16]	[2.12, 6.16, 32.75]
$t_{\text{sfg}}/t_{\text{frag}}$	[43.63, 69.07, 137.39]	[41.00, 65.86, 123.26]	[35.63, 57.21, 115.10]	[29.59, 57.33, 169.20]	[24.12, 60.52, 235.34]
Star-forming sample					
76762	2.2%	1.5%	1.1%	0.3%	0.0%
Timescales	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]
t_{orb} (Myr)	[65.87, 129.84, 209.79]	[30.95, 65.21, 115.34]	[28.80, 60.72, 105.58]	[18.18, 45.44, 98.73]	[11.16, 47.99, 109.13]
t_{cool} (Myr)	[1.738, 6.264, 16.940]	[0.547, 1.504, 3.506]	[0.471, 1.355, 2.933]	[0.191, 0.804, 2.395]	[0.070, 0.801, 3.186]
$\mathcal{T}_{\text{cool}}$ (Myr)	[6.024, 17.445, 36.808]	[1.967, 5.358, 11.648]	[1.673, 4.771, 9.871]	[0.794, 2.929, 8.081]	[0.355, 3.290, 9.476]
t_{frag} (Myr)	[7.059, 14.864, 26.348]	[2.914, 6.132, 11.352]	[2.542, 5.504, 9.915]	[1.407, 3.806, 8.482]	[0.713, 3.976, 9.133]
t_{disrupt} (Myr)	[7.166, 15.119, 26.448]	[2.937, 6.252, 11.630]	[2.584, 5.608, 10.130]	[1.429, 3.906, 8.714]	[0.714, 4.146, 9.505]
t_{sfg} (Myr)	[834.8, 943.1, 1162.4]	[255.4, 382.5, 590.6]	[174.4, 269.8, 400.7]	[105.2, 166.7, 262.0]	[69.4, 99.3, 175.9]
Ratios	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]	[p15, p50, p85]
$t_{\text{cool}}/t_{\text{orb}}$	[0.029, 0.055, 0.143]	[0.018, 0.025, 0.042]	[0.017, 0.023, 0.035]	[0.012, 0.022, 0.042]	[0.014, 0.029, 0.058]
$\mathcal{T}_{\text{cool}}/t_{\text{cool}}$	[1.781, 2.694, 3.556]	[2.982, 3.488, 3.746]	[3.171, 3.472, 3.757]	[3.058, 3.445, 4.199]	[2.963, 3.312, 4.314]
$t_{\text{frag}}/t_{\text{orb}}$	[0.104, 0.115, 0.132]	[0.087, 0.094, 0.104]	[0.082, 0.091, 0.099]	[0.072, 0.083, 0.092]	[0.067, 0.081, 0.095]
$t_{\text{sfg}}/t_{\text{orb}}$	[5.33, 7.61, 12.45]	[3.88, 5.81, 9.11]	[3.17, 4.76, 7.60]	[2.07, 3.70, 7.31]	[1.32, 2.30, 7.58]
$t_{\text{sfg}}/t_{\text{frag}}$	[43.41, 64.93, 114.44]	[39.95, 60.73, 97.30]	[34.20, 51.90, 85.50]	[24.63, 42.88, 89.36]	[14.79, 30.06, 124.96]

Notes. The model galaxies used for this analysis all lie at $z = 2.1$. Median values and 15% and 85% percentiles are given for five different stellar mass bins. The upper table shows statistics of the fully resolved galaxy sample. The lower table shows the star-forming sample. In each sub-table, the upper part shows absolute values of the timescales and the lower part shows the ratios with the orbital time or gas fragmentation timescales.

redshift, the characteristic cooling time is a decreasing function of the mass. These two trends together means that the characteristic cooling timescale of a median galaxy track decreases with both redshift and stellar mass (Fig. 18). At $z = 2.1$, the characteristic cooling timescale of our star-forming galaxies sample is distributed between ≈ 0.8 Myr and ≈ 6.2 Myr. These values are slightly smaller than in the fully resolved galaxy sample of ≈ 1.3 Myr and ≈ 8.0 Myr (Table 3). At all redshifts and for all stellar masses, the characteristic cooling timescale is significantly shorter than the orbital timescale ($< 6\%$). If no additional mass were added to the reservoir of diffuse gas, its condensation, which is mainly governed by the characteristic cooling timescale, would deplete completely in less than an orbital time. The short cooling and condensation timescales help to drive the overall complex dynamics of the gas cycle in model galaxies as alluded to earlier.

To provide an indication of how much of the diffuse gas is condensing at any given time, we tabulate the mass fraction of the diffuse gas (Table 3). The mass fraction of the diffuse gas that is condensing can be estimated by comparing the effective cooling time (Eq. (6)) to the characteristic cooling timescale (Eq. (4)). At $z = 2.1$, for the five stellar mass bins tabulated, the median values of the ratio, $\mathcal{T}_{\text{cool}}/t_{\text{cool}}$, range between ≈ 2.7 and ≈ 3.5 for both model galaxy samples. Low-mass galaxies have

the lowest ratios. The ratio is approximately constant, ≈ 3.5 , for galaxies with $M_{\star} \geq 10^{10} M_{\odot}$.

The fragmentation and disruption timescales have very similar dependencies on redshift and stellar mass (Fig. 18). These two timescales show a decreasing trend with both redshift and stellar mass. At $z = 2.1$, for the star-forming galaxy sample, the fragmenting timescale and the disrupting timescale are distributed between 4 and 15 Myr. Those values correspond to $\approx 10\%$ of the disk orbital timescale. As for the cooling timescale, the fragmentation and disruption timescales of our star-forming galaxy sample are marginally smaller than those of the full galaxy sample. The disruption timescale is always (slightly) larger than the fragmentation timescale (Table 3) and thus gas fragmentation is always faster than gas disruption in G.A.S. The short fragmentation timescales indicate that the growth of GMCs is a continuous and efficient process. Timescales measured in our model are fully compatible with GMC gas accretion rates found via numerical simulations (e.g. Vázquez-Semadeni et al. 2010; Colin et al. 2013).

Working against the gas cooling and fragmentation is the injection of energy by SNe and AGN. This energy injection is used in G.A.S. to disrupt the gas and transfer some of the fragmented gas back into the reservoir of diffuse gas. Hydrodynamical simulations find that disruption timescales of

about 10–20 Myr (Colin et al. 2013; Dobbs 2015). This range of values is fully consistent with our predictions.

7.3.1. Impact of fragmentation and disruption on the life cycle of GMCs

The small differences between the condensation and disruption timescales implies that the total mass of fragmented gas in GMCs stays approximately constant and that significant fraction of the fragmented mass is continuously regenerated. From this equilibrium, there are two evolutionary paths for GMCs that depend on both the scale and mass of GMCs.

- Low-mass GMCs, i.e. those formed in galaxies with short or intermediate disk-scale heights, $h_d \leq 50$ pc; the mass disrupted by SN kinetic energy injection is close to the total of low-mass GMCs. In such circumstances, after only few SN cycles, GMCs are fully disrupted. These disrupted GMCs are constantly replenished through the formation of new GMCs through the condensation of diffuse gas. Thus the disruption/fragmentation timescales are approximately the lifetime of the GMCs in our model. Our results are in reasonable agreement with the sizes and short GMC lifetimes estimated in local galaxies, 25–70 pc and 17 ± 4 Myr (Murray 2011; Miura et al. 2012; Meidt et al. 2015).
- More massive GMCs resist being disrupted by SNe. The competition between accretion and disruption regulates the mass of GMCs (Vázquez-Semadeni et al. 2010). Because of this competition, GMCs have to be therefore considered as dynamical structures where the gas is just passing through from the diffuse-gas state to the star-forming gas state.

The SN/AGN-driven gas disruption slows significantly the rate at which gas fragments. In small GMCs, the cloud structure and therefore the sites of star formation can be completely destroyed after only a few SN cycles. In more massive GMCs, the fragmentation is also significantly slowed. Indeed, gas recently added to the reservoir of diffuse gas starts to fragment on larger scale than the gas within the GMCs. In our prescription, this constant renewal and exchange of the gas in these reservoirs and the impact of the added gas to the overall cascade of fragmenting gas is fully accounted for via the mass-weighted fragmentation clock (Eq. (19)).

7.3.2. Gas depletion timescale

Depending on the redshift and stellar mass of a model galaxy, the gas depletion timescale, t_{sfg} , ranges from 80 Myr to 2 Gyr (Fig. 18). The upper end of this range is consistent with the gas depletion timescale measured in local spiral galaxies (e.g. Leroy et al. 2013; Colombo et al. 2018) and the values predicted at high redshift, $z \approx 2-3$, are also generally consistent with the values estimated for distant star-forming galaxies (e.g. Daddi et al. 2010; Genzel et al. 2010). Gas depletion timescales are a decreasing function of both the redshift and stellar mass. Above $M_\star \geq 10^{10.5} M_\odot$ and at $z > 3.0$, the average gas depletion timescale reaches a lower limit of close to 80 Myr.

The $t_{\text{sfg}}/t_{\text{frag}}$ ratio provides an estimate of the efficiency of star formation in GMCs. The number of condensation/disruption cycles and the depletion timescale of fragmented gas imply that faster cycles lead to shorter depletion timescales. At $z = 2.1$, median values extracted from our star-forming sample indicate that between ≈ 30 and ≈ 65 cycles are required to convert fragmented gas into star-forming gas (and therefore into stars). This number of gas cycles is consistent with gas dynamics and number of cycles estimated in hydrodynamic simula-

tions (e.g. Semenov et al. 2017). In the star-forming sample, the average number of condensation/disruption-cycles is strictly a decreasing function of the stellar mass. However, in the full galaxy sample, a minimum is reached for models galaxies with mass of $10^{10.25} M_\odot$. The fragmenting gas in galaxies less and more massive than $10^{10.25} M_\odot$ need more cycles to convert the fragmented gas into stars. During each cycle only a small fraction, 1%–5%, of the fragmented gas stops fragmenting and is available for the star formation. In fact, the majority of the gas that forms GMCs is recycled in of order one to a few dynamical times or simply remains in a state that cannot form stars. We obtain star formation efficiencies between $\approx 1\%$ and $\approx 5\%$. Such values are consistent with observational estimates (e.g. Wong & Blitz 2002; Leroy et al. 2008; Murray et al. 2010; André 2013; Heiderman et al. 2010) and efficiencies measured in hydrodynamic simulations (e.g. Semenov et al. 2017; Kimm et al. 2017).

In summary, the gas cycle implemented in the new G.A.S. model attempts to capture the complex dynamics of the gas in galaxies. In our prescriptions, gas is continuously exchanged between the diffuse and the fragmented non-star-forming phases. A large number of condensation and disruption cycles, 30–70, are needed to progressively convert diffuse gas into very dense star-forming gas and then stars. The characteristic timescales of ≈ 15 Myr of these exchange rates is only a small fraction of the orbital timescale. Under such conditions, the fragmented gas is converted into star-forming gas and stars continuously and the gas spends the majority of its time in a non-star-forming gas phase.

8. Discussion and conclusions

We present a new semi-analytical model, G.A.S., in which we implemented a more realistic gas cycle than has been previously implemented in a semi-analytical model. We introduced a prescription for delaying star formation which is underpinned by progressively fragmenting the gas. The formation of giant molecular clouds, filaments, and cold cores takes time and therefore at a given instant only a small fraction of the total gas is in the form of prestellar cores. The majority of the gas is not in a form that is immediately available to form stars. Within this framework, we account for the continuous dissipation of the turbulent kinetic energy through the different ISM scales and phases. We implemented an approximate multiphase ISM in which the gas cycles between a warm diffuse phase, a cold fragmented non-star-forming phase, and a very dense star-forming gas phase. Diffuse warm gas is progressively fragmented following the effective cooling time. Even if the overall depletion timescale of the fragmented gas is, on average, proportional to the disk orbital timescale, we measure a large scatter in both the depletion timescale and the disk orbital timescale. This proves that prescriptions that rely solely on the disk orbital timescale do not properly capture the complexity of gas cycles in galaxies. Smaller characteristic timescales, such as the cooling and/or energy injection timescales due to SNe and AGN are also important. This energy input on shorter timescales, efficiently disrupts the star-forming gas. The large-scale velocity dispersion of the diffuse gas is also maintained by SNe/AGN kinetic energy injection. A fraction of the gas can also be ejected from disks by SN explosions but the characteristic timescale of ejection is on average ten times larger than the fragmenting/disruption cycle timescale. While the SN/AGN kinetic energy injection in the ISM regulates the star formation rate in low-mass galaxies, our model galaxies retain a high percentage of the gas. This “local” gas is then used with a progressively increasing efficiency

necessary to build massive, $M_{\star} > 10^{11} M_{\odot}$, galaxies in the early Universe, $z \in [4; 6]$.

Star formation occurs only in the very dense gas, which in our model is a product of continuous fragmentation over a wide range of scales. This new complete gas cycle strongly regulates the star formation in our modelled galaxies. Only a few percent of the available fragmented gas is converted into stars during a GMC life cycle. By taking into account the fragmented gas on a galaxy scale, our model is able to reproduce the standard Schmidt–Kennicutt law. Our estimated gas depletion timescales, which are directly related to the time it takes for gas to fragment, are in good agreement with observational estimates.

This new gas regulation cycle leads to very good agreement with the observed stellar mass function over a wide range of redshifts, $z \in [0.8; 6]$. The ability of our model to catch the stellar mass assembly in high redshift galaxies as been already used with success in Lagache et al. (2018). Galaxy properties (gas/stars contents, metallicities, and far-ultraviolet fluxes ...) predicted by our model have been post-processed to successfully predict the [CII] luminosity functions at high redshifts and have allowed us to explore the main characteristics of this emission.

At $z < 0.8$, we find some discrepancies between the predicted and the observed stellar mass functions. Specifically, an over-density of low-mass galaxies and the under-density of intermediate-mass galaxies modelled at $z < 0.8$ could probably be solved by slightly increasing the efficiency of photoionisation in our prescription.

At $z < 4.0$, to reduce the efficiency of or to stop the growth of stellar mass in massive galaxies, the accretion of gas and its transformation into stars has to be reduced or stopped. Previous semi-analytical models implemented strong AGN feedback to quench gas accretion onto massive galaxies. In those models, a significant fraction of the power produced by the AGN is directly used to reduce the cooling of the hot halo gas. This implies a constant AGN power production, which may not be consistent with our understanding of AGN variability (see e.g. Hickox et al. 2014; Stanley et al. 2015; Volonteri et al. 2015, and references therein). For massive galaxies, we propose an alternative model which regulates the radiative cooling and gas accretion onto galaxies. In parallel to radiative cooling, we implemented the development of thermal instabilities creating warm gas surrounding the galaxy. The growth of these thermal instabilities in the central region of the halo progressively reduces or halts the cooling and dissipation of kinetic energy in the halo and therefore reduces or stops the accretion onto galaxies. This process of regulation is a natural outcome of the growth of thermal instabilities in the hot halo gas and does not depend on the power produced by AGN.

However, our actual efficiency parameters do not fully stop gas accretion onto very massive galaxies. In some massive dark-matter haloes, radiative cooling restarts ($VFF < 1.0$) even if its hot gaseous halo has been fully quenched previously ($VFF = 1.0$). The recovery of the gas accretion leads to the formation of some (< 10) unobserved overly massive galaxies at $z \simeq 0.3$. Our prescription of thermal instability growth in the hot gas phase is governed by a set of two parameters which can be further refined to overcome this problem.

In our model of the gas cycle, we assume that the gas initially fragments at the disk scale height with the progressive and continuous formation of over-densities which are akin to observed GMCs. The fragmentation of the gas will progress down to the scale at which stars form or ~ 0.1 pc. However, this fragmentation can start at larger scales in the cold streams or in the hot(warm) gas phase surrounding the galaxy. Some clumps of

warm gas could already be formed around the galaxy. We assume in our model that a fraction of the newly accreted gas is already fragmented but we do not include any interaction between this already fragmented gas in the halo and the large-scale wind. This kind of interaction could reduce, perhaps substantially, the gas-accretion efficiency (Cornuault et al. 2018). Indeed such coupling could increase the cloud-cloud velocity dispersion and maintain the turbulence in the hot and warm gas contained in the CGM of the most massive galaxy. Such a hypothesised mechanism will be complementary to the development of thermal instabilities, could in fact act as a catalyst to the formation of more clouds in the CGM, and could contribute in quenching the radiative cooling and dissipation in the most massive haloes.

Acknowledgements. MC acknowledges Thomas Fenouillet for his much appreciated help in the use of the LAM computation clusters and wishes to thank Mathieu Géniois for his help in the proofreading of this article and for all physical and technical discussions. MC thanks Olivier Ilbert for the numerous and very helpful discussions about stellar mass function (estimations, errors, limits) and Benoit Epinat for a useful discussion about galaxy disk dynamics. MC also wishes to express his appreciation to Alexandre Beelen and Yanick Roehly for very useful physical and technical discussions. The authors thank the Centre National d’Etudes Spatiales (CNES), Aix Marseille Université, Sorbonne Université, the Programme National de Cosmologie and Galaxies (PNCG) and Programme de Physico-chimie du Milieu Interstellaire (PCMI) of CNRS/INSU for their financial support. PG gratefully acknowledges the support of the Institut Universitaire de France (IUF).

References

- Agertz, O., & Kravtsov, A. V. 2015, *ApJ*, 804, 18
 Aguirre, A., Hernquist, L., Schaye, J., et al. 2001, *ApJ*, 560, 599
 André, P. 2013, *Proc. IAU*, 10, 31
 Aravena, M., Decarli, R., Walter, F., et al. 2016, *ApJ*, 833, 71
 Arzoumanian, D., André, P., Peretto, N., & Könyves, V. 2013, *A&A*, 553, A119
 Baldry, I. K., Glazebrook, K., & Driver, S. P. 2008, *MNRAS*, 388, 945
 Baugh, C. M. 2006, *Rep. Progr. Phys.*, 69, 3101
 Benson, A. J. 2012, *New Astron.*, 17, 175
 Bergin, E. A., & Tafalla, M. 2007, *ARA&A*, 45, 339
 Béthermin, M., De Breuck, C., Gullberg, B., et al. 2016, *A&A*, 586, L7
 Bigiel, F., Leroy, A., Walter, F., et al. 2008, *AJ*, 136, 2846
 Bondi, H. 1952, *MNRAS*, 112, 195
 Bonnor, W. B. 1956, *MNRAS*, 116, 351
 Bothwell, M. S., Maiolino, R., Peng, Y., et al. 2016, *MNRAS*, 455, 1156
 Bothwell, M. S., Aguirre, J. E., Aravena, M., et al. 2017, *MNRAS*, 466, 2825
 Capelo, P. R., Coppi, P. S., & Natarajan, P. 2012, *MNRAS*, 422, 686
 Caputi, K. I., Cirasuolo, M., Dunlop, J. S., et al. 2011, *MNRAS*, 413, 162
 Cattaneo, A., Dekel, A., Devriendt, J., Guiderdoni, B., & Blaizot, J. 2006, *MNRAS*, 370, 1651
 Cole, S. 1991, *ApJ*, 367, 45
 Cole, S., Lacey, C. G., Baugh, C. M., & Frenk, C. S. 2000, *MNRAS*, 319, 168
 Colin, P., Vázquez-Semadeni, E., & Gómez, G. C. 2013, *MNRAS*, 435, 1701
 Colombo, D., Kalinova, V., Utomo, D., et al. 2018, *MNRAS*, 475, 1791
 Cornuault, N., Lehnert, M. D., Boulanger, F., & Guillard, P. 2018, *A&A*, 610, A75
 Couchman, H. M. P., & Rees, M. J. 1986, *MNRAS*, 221, 53
 Cousin, M. 2019, *A&A*, submitted (Paper III)
 Cousin, M., Lagache, G., Béthermin, M., Blaizot, J., & Guiderdoni, B. 2015a, *A&A*, 575, A32
 Cousin, M., Lagache, G., Béthermin, M., & Guiderdoni, B. 2015b, *A&A*, 575, A33
 Cousin, M., Buat, V., Boissier, S., et al. 2016, *A&A*, 589, A109
 Cousin, M., Buat, B., Lagache, G., & Béthermin, M. 2019, *A&A*, 627, A132
 Croton, D. J., Springel, V., White, S. D. M., et al. 2006, *MNRAS*, 365, 11
 Daddi, E., Bournaud, F., Walter, F., et al. 2010, *ApJ*, 713, 686
 Dale, J. E. 2015, *New Astron. Rev.*, 68, 1
 Davidzon, I., Cucciati, O., Bolzonella, M., et al. 2016, *A&A*, 586, A23
 Decarli, R., Walter, F., Aravena, M., et al. 2016, *ApJ*, 833, 70
 Dekel, A., & Birnboim, Y. 2006, *MNRAS*, 368, 2
 Dekel, A., & Silk, J. 1986, *ApJ*, 303, 39
 De Lucia, G., Kauffmann, G., & White, S. D. M. 2004, *MNRAS*, 349, 1101
 De Rijcke, S., Schroyen, J., Vandenbroucke, B., et al. 2013, *MNRAS*, 433, 3005
 Dickman, R. L., Horvath, M. A., & Margulis, M. 1990, *ApJ*, 365, 586

- Dobbs, C. L. 2015, *MNRAS*, **447**, 3390
- Doroshkevich, A. G., Zel'dovich, Y. B., & Novikov, I. D. 1967, *Sov. Astron.*, **11**, 233
- Duncan, K., & Conzelmann, C. J. 2015, *MNRAS*, **451**, 2030
- Efstathiou, G. 2000, *MNRAS*, **317**, 697
- Elmegreen, B. G. 2009, in *Galaxy Evolution: Emerging Insights and Future Challenges*, eds. S. Jogee, I. Marinova, L. Hao, & G. A. Blanc, *ASP Conf. Ser.*, **419**, 23
- Elmegreen, D. M., Elmegreen, B. G., Marcus, M. T., et al. 2009, *ApJ*, **701**, 306
- Falkendal, T., De Breuck, C., Lehnert, M. D., et al. 2019, *A&A*, **621**, A27
- Federrath, C., & Klessen, R. S. 2012, *ApJ*, **761**, 156
- Fleck, Jr., R. C. 1996, *ApJ*, **458**, 739
- Gao, Y., & Solomon, P. M. 2004, *ApJ*, **606**, 271
- Gatto, A., Walch, S., Low, M.-M. M., et al. 2015, *MNRAS*, **449**, 1057
- Gavagnin, E., Bleuler, A., Rosdahl, J., & Teyssier, R. 2017, *MNRAS*, **472**, 4155
- Geen, S., Soler, J. D., & Hennebelle, P. 2017, *MNRAS*, **471**, 4844
- Genzel, R., Tacconi, L. J., Gracia-Carpio, J., et al. 2010, *MNRAS*, **407**, 2091
- Gnedin, N. Y. 2000, *ApJ*, **542**, 535
- Grazian, A., Fontana, A., Santini, P., et al. 2015, *A&A*, **575**, A96
- Guillard, P., Boulanger, F., Pineau Des Forêts, G., & Appleton, P. N. 2009, *A&A*, **502**, 515
- Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, *MNRAS*, **413**, 101
- Hatton, S., Devriendt, J. E. G., Ninin, S., et al. 2003, *MNRAS*, **343**, 75
- Heiderman, A., Evans, II, N. J., Allen, L. E., Huard, T., & Heyer, M. 2010, *ApJ*, **723**, 1019
- Hennebelle, P., & Falgarone, E. 2012, *A&ARv*, **20**, 55
- Hennebelle, P., & Iffrig, O. 2014, *A&A*, **570**, A81
- Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2013, *MNRAS*, **431**, 3373
- Hickox, R. C., Mullaney, J. R., Alexander, D. M., et al. 2014, *ApJ*, **782**, 9
- Hoffmann, V., & Romeo, A. B. 2012, *MNRAS*, **425**, 1511
- Hopkins, P. F. 2012, *MNRAS*, **423**, 2016
- Hopkins, P. F., Kereš, D., Oñorbe, J., et al. 2014, *MNRAS*, **445**, 581
- Ikeuchi, S. 1986, *Ap&SS*, **118**, 509
- Ilbert, O., McCracken, H. J., Le Fevre, O., et al. 2013, *A&A*, **556**, A55
- Jog, C. J., & Solomon, P. M. 1984a, *ApJ*, **276**, 127
- Jog, C. J., & Solomon, P. M. 1984b, *ApJ*, **276**, 114
- Kauffmann, G., White, S. D. M., & Guiderdoni, B. 1993, *MNRAS*, **264**, 201
- Kennicutt, Jr., R. C. 1998, *ApJ*, **498**, 541
- Khochfar, S., & Silk, J. 2009, *ApJ*, **700**, L21
- Kimm, T., Katz, H., Haehnelt, M., et al. 2017, *MNRAS*, **466**, 4826
- Komatsu, E., & Seljak, U. 2001, *MNRAS*, **327**, 1353
- Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2009, *ApJS*, **180**, 330
- Kravtsov, A. V., Gnedin, O. Y., & Klypin, A. A. 2004, *ApJ*, **609**, 482
- Kritsuk, A. G., Lee, C. T., & Norman, M. L. 2013, *MNRAS*, **436**, 3247
- Krumholz, M. R., & McKee, C. F. 2005, *ApJ*, **630**, 250
- Krumholz, M. R., & Thompson, T. A. 2012, *ApJ*, **760**, 155
- Lada, C. J., Lombardi, M., & Alves, J. F. 2010, *ApJ*, **724**, 687
- Lada, C. J., Forbrich, J., Lombardi, M., & Alves, J. F. 2012, *ApJ*, **745**, 190
- Lagache, G., Cousin, M., & Chatzikos, M. 2018, *A&A*, **609**, A130
- Laporte, N., Ellis, R. S., Boone, F., et al. 2017, *ApJ*, **837**, L21
- Larson, R. B. 1981, *MNRAS*, **194**, 809
- Lee, E. J., Miville-Deschênes, M.-A., & Murray, N. W. 2016, *ApJ*, **833**, 229
- Lehnert, M. D., Le Tiran, L., Nesvadba, N. P. H., et al. 2013, *A&A*, **555**, A72
- Lehnert, M. D., van Driel, W., Le Tiran, L., Di Matteo, P., & Haywood, M. 2015, *A&A*, **577**, A112
- Leroy, A. K., Walter, F., Brinks, E., et al. 2008, *AJ*, **136**, 2782
- Leroy, A. K., Walter, F., Sandstrom, K., et al. 2013, *AJ*, **146**, 19
- Levrier, F., Neveu, J., Falgarone, E., et al. 2018, *A&A*, **614**, A124
- Lu, Y., Kereš, D., Katz, N., et al. 2011, *MNRAS*, **416**, 660
- Madau, P., & Dickinson, M. 2014, *ARA&A*, **52**, 415
- Makino, N., Sasaki, S., & Suto, Y. 1998, *ApJ*, **497**, 555
- Meidt, S. E., Hughes, A., Dobbs, C. L., et al. 2015, *ApJ*, **806**, 72
- Mitchell, P. D., Blaizot, J., Devriendt, J., et al. 2018, *MNRAS*, **474**, 4279
- Miura, R. E., Kohno, K., Tosaki, T., et al. 2012, *ApJ*, **761**, 37
- Miville-Deschênes, M.-A., Martin, P. G., Abergel, A., et al. 2010, *A&A*, **518**, L104
- Mo, H. J., Mao, S., & White, S. D. M. 1998, *MNRAS*, **295**, 319
- Mullaney, J. R., Alexander, D. M., Aird, J., et al. 2015, *MNRAS*, **453**, L83
- Murray, N. 2011, *ApJ*, **729**, 133
- Murray, N., Quataert, E., & Thompson, T. A. 2010, *ApJ*, **709**, 191
- Netzer, H., Mor, R., Trakhtenbrot, B., Shemmer, O., & Lira, P. 2014, *ApJ*, **791**, 34
- Netzer, H., Lani, C., Nordon, R., et al. 2016, *ApJ*, **819**, 123
- Okamoto, T., Gao, L., & Theuns, T. 2008, *MNRAS*, **390**, 920
- Oppenheimer, B. D., & Davé, R. 2008, *MNRAS*, **387**, 577
- Orkisz, J. H., Pety, J., Gerin, M., et al. 2017, *A&A*, **599**, A99
- Ostriker, J. P., Choi, E., Ciotti, L., Novak, G. S., & Proga, D. 2010, *ApJ*, **722**, 642
- Oteo, I., Zhang, Z.-Y., Yang, C., et al. 2017, *ApJ*, **850**, 170
- Padoan, P., & Nordlund, Å. 2011, *ApJ*, **730**, 40
- Padoan, P., Juvela, M., Goodman, A. A., & Nordlund, Å. 2001, *ApJ*, **553**, 227
- Palmeirim, P., André, P., Kirk, J., et al. 2013, *A&A*, **550**, A38
- Pelliccia, D., Tresse, L., Epinat, B., et al. 2017, *A&A*, **599**, A25
- Popping, G., Decarli, R., Man, A. W. S., et al. 2017, *A&A*, **602**, A11
- Raskutti, S., Ostriker, E. C., & Skinner, M. A. 2016, *ApJ*, **829**, 130
- Rees, M. J. 1986, *MNRAS*, **218**, 25
- Renaud, F., Bournaud, F., Kraljic, K., & Duc, P.-A. 2014, *MNRAS*, **442**, L33
- Romeo, A. B., Burkert, A., & Agertz, O. 2010, *MNRAS*, **407**, 1223
- Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, **446**, 521
- Schmidt, W., Collins, D. C., & Kritsuk, A. G. 2013, *MNRAS*, **431**, 3196
- Scholtz, J., Alexander, D. M., Harrison, C. M., et al. 2018, *MNRAS*, **475**, 1288
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, *A&A*, **575**, A74
- Semenov, V. A., Kravtsov, A. V., & Gnedin, N. Y. 2016, *ApJ*, **826**, 200
- Semenov, V. A., Kravtsov, A. V., & Gnedin, N. Y. 2017, *ApJ*, **845**, 133
- Sharma, P., McCourt, M., Quataert, E., & Parrish, I. J. 2012, *MNRAS*, **420**, 3174
- Somerville, R. S., Hopkins, P. F., Cox, T. J., Robertson, B. E., & Hernquist, L. 2008, *MNRAS*, **391**, 481
- Song, M., Finkelstein, S. L., Ashby, M. L. N., et al. 2016, *ApJ*, **825**, 5
- Stanley, F., Harrison, C. M., Alexander, D. M., et al. 2015, *MNRAS*, **453**, 591
- Sutherland, R. S., & Dopita, M. A. 1993, *ApJS*, **88**, 253
- Suto, Y., Sasaki, S., & Makino, N. 1998, *ApJ*, **509**, 544
- Suzuki, T. L., Kodama, T., Onodera, M., et al. 2017, *ApJ*, **849**, 39
- Swinbank, A. M., Harrison, C. M., Trayford, J., et al. 2017, *MNRAS*, **467**, 3140
- Tweed, D., Devriendt, J., Blaizot, J., Colombi, S., & Slyz, A. 2009, *A&A*, **506**, 647
- van de Voort, F., & Schaye, J. 2012, *MNRAS*, **423**, 2991
- Vázquez-Semadeni, E., Colín, P., Gómez, G. C., Ballesteros-Paredes, J., & Watson, A. W. 2010, *ApJ*, **715**, 1302
- Volonteri, M., Capelo, P. R., Netzer, H., et al. 2015, *MNRAS*, **449**, 1470
- White, S. D. M., & Frenk, C. S. 1991, *ApJ*, **379**, 52
- White, S. D. M., & Rees, M. J. 1978, *MNRAS*, **183**, 341
- Wong, T., & Blitz, L. 2002, *ApJ*, **569**, 157
- Yang, X., Mo, H. J., & van den Bosch, F. C. 2009, *ApJ*, **695**, 900

Appendix A: The GALAKSIENN library

The GALAKSIENN library stores the main results produced by our new G.A.S. semi-analytical model, especially MOCK galaxy catalogues and sky maps. It is available online through the ZENODO platform: <https://zenodo.org/>, DOI: [10.5281/zenodo.1451229](https://doi.org/10.5281/zenodo.1451229). A complete description of the GALAKSIENN library is given in Paper III. In association with this Paper I, we distribute the ASCII tables of the stellar mass functions (Fig. 14).