Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **409,** 1057–1067 (2010) doi:10.1111/j.1365-2966.2010.17363.x

High-redshift formation and evolution of central massive objects – I. Model description

B. Devecchi,[1⋆] M. Volonteri,[2] M. Colpi[3] and F. Haardt[1]

[1]*Dipartimento di Fisica & Matematica, Università dell'Insubria, Via Valleggio 11, 22100 Como, Italy*
[2]*Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA*
[3]*Dipartimento di Fisica G. Occhialini, Università degli Studi di Milano Bicocca, Piazza della Scienza 3, 20126 Milano, Italy*

Accepted 2010 July 13. Received 2010 June 30; in original form 2010 January 7

ABSTRACT

Galactic nuclei host central massive objects either in the form of supermassive black holes or in the form of nuclear stellar clusters. Recent investigations have shown that both components co-exist in at least a few galaxies. In this paper, we explore the possibility of a connection between nuclear star clusters and black holes that establishes at the moment of their formation. We here model the evolution of high-redshift discs, hosted in dark matter haloes with virial temperatures $>10^4$ K, whose gas has been polluted with metals just above the critical metallicity for fragmentation. A nuclear cluster forms as a result of a central starburst from gas inflowing from the unstable disc. The nuclear stellar cluster provides a suitable environment for the formation of a black hole seed, ensuing from runaway collisions among the most massive stars. Typical masses for the nuclear stellar clusters at the time of black hole formation ($z \sim 10$) are in the range 10^4–10^6 $M_\odot$ and have half-mass radii $\lesssim$0.5 pc. The black holes forming in these dense, high-redshift clusters can have masses in the range $\sim$300–2000 $M_\odot$.

Key words: black hole physics – instabilities – galaxies: formation – galaxies: nuclei – galaxies: star clusters: general.

1 INTRODUCTION

Central massive objects are known to reside in the centres of a large fraction of galaxies either in the form of supermassive black holes (BHs) or nuclear stellar clusters (NCs). BHs (and to a lesser extent NCs) have masses that correlate with those of their hosts (Marconi & Hunt 2003; Häring & Rix 2004; Ferrarese et al. 2006; Wehner & Harris 2006; Seth et al. 2008), suggesting a symbiotic evolution between these nuclear components and the larger structure of their galaxy hosts.

NCs, with masses, $M_{\rm NC}$, ranging between 10^5 and 10^7 $M_\odot$ (Böker et al. 2004; Côté et al. (2006); Walcher et al. 2006; Balcells, Graham & Peletier 2007) appear to be more common in galaxies with stellar masses below $\sim 10^{11}$ $M_\odot$, while BHs ($M_{\rm BH} > 10^6$ $M_\odot$) have been suggested to preferentially inhabit galaxies more massive than 10^{10} $M_\odot$ (Ferrarese et al. 2006; Wenher et al. 2006). Despite this apparent dichotomy, NCs and BHs can co-exist in the same system. NCs with embedded BHs are indeed observed along the all-galaxy mass range (Seth et al. 2008; Graham & Spitler 2009; Kormendy et al. 2009).

Two main scenarios have been proposed to explain the origin of NCs. In the first scenario, NCs are the end-product of multiple mergers of globular clusters following their sinking by dynamical friction into the central part of the host galaxy (Tremaine, Ostriker & Spitzer 1975; Lotz et al. 2001; Capuzzo-Dolcetta & Miocchi 2008). In the second scenario, NCs form *in situ* after inflow of gas from the outer part of the host (Milosavljevic 2004). The latter possibility seems more consistent with stellar population studies in NCs (Côté et al. 2006).

Whereas models of NCs in a cosmological context are still lacking, models of BH seed formation have been developed in the Lambda cold dark matter (ΛCDM) cosmogony. How BH seeds form and how they evolve into the supermassive variety has been the subject of many studies (Begelman & Rees 1978). A natural path to BH formation relies on the first generation of stars (Population III) born in metal-free environments (Madau & Rees 2001; Volonteri, Haardt & Madau 2003). These stars are expected to be more massive than their present counterparts (Abel, Bryan & Norman 2000; Omukai & Palla 2003; Freese et al. 2008; Iocco et al. 2008). Population III stars that form above a threshold mass of $\simeq$260 $M_\odot$ are believed to directly implode, leaving a seed BH of similar mass (Heger et al. 2003). Alternative routes to BH seed formation typically exploit gas-dynamical processes in metal-free galaxies (direct formation models) (Loeb & Rasio 1994; Eisenstein & Loeb 1995; Sellwood & Moore 1999; Bromm & Loeb 2003; Koushiappas, Bullock & Dekel 2004; Begelman, Volonteri & Rees 2006; Lodato & Natarajan 2006, also see the discussion in Schleicher, Spaans

⋆E-mail: bernadetta.devecchi@mib.infn.it

& Glover 2010). Gravitational instabilities can indeed lead to a vigorous gas inflow into the very central region of protogalactic discs, supplying the necessary matter for the formation of BH seeds.

Most of the above scenarios have a common ingredient, that is, the gas hosted in the progenitor halo has not been polluted by metals. Metal pollution marks the end of the formation of Population III stars and of their seed BHs. Fragmentation and formation of low-mass stars start as soon as the gas is polluted above a critical metallicity threshold, $Z_{\rm crit}$ (Bromm, Coppi & Larson 1999; Bromm et al. 2002; Santoro & Shull 2006; Schneider et al. 2006; Clark, Glover & Klessen 2008; Omukai, Schneider & Haiman 2008; Smith et al. 2009). Population III stars with masses between 140 and 260 $\rm M_\odot$ are believed to be responsible for starting metal pollution, as they do not leave a relic BH, but explode releasing all metals in the surrounding medium. Star formation is also a hindrance for strong gas inflows in the centre of massive haloes as it limits the available mass supply. This does not necessary means the end of the seed formation epoch as a new channel for BH formation opens (Omukai et al. 2008; Devecchi & Volonteri 2009, hereinafter DV09).

DV09 showed that the first episode of low-mass, Population II star formation can provide the conditions for the formation of NCs and BH seeds arising in the most compact of these clusters. Gravitationally unstable protogalactic discs, whose gas is only mildly polluted by metals can still produce strong gas inflows into their central regions. Stars start to form in the central few pc, where gas has accumulated and density is high enough for star formation to set in. Clusters formed in this way are crowded places. Star–star collisions in their core can proceed in a runaway fashion and a very massive star (VMS) of $\sim$1000 $\rm M_\odot$ can build up before the first supernova (SN) explodes (Portegies Zwart & McMillan 2002; Gürkan, Freitag & Rasio 2004; Freitag et al. 2006a,b). At low metallicity, the final fate of this VMS is to collapse into a BH with mass similar to its progenitor – as at low metallicity, mass-loss is negligible.

One of the factors that regulate the competition, at a given redshift, between different formation scenarios is the metal-enrichment history. Both observational and theoretical work has been performed in order to investigate how metals are distributed as cosmic structures form and evolve (Prochaska et al. 2003; Scannapieco, Schneider & Ferrara 2003; Kulkarni et al. 2005; Savaglio et al. 2005; Savaglio 2006; Tornatore, Ferrara & Schneider 2007; Li 2008; Prochaska et al. 2007). These studies have shown that metal enrichment proceeds in a very inhomogeneous fashion. This opens the possibility that different modes of BH seed formation co-exist at different regions of the Universe at the same redshift, instead of being mutually exclusive. As the mean metallicity increases with cosmic time, growing bubbles of polluted gas, where the first NCs and their BH seeds form, co-exist with relatively pristine regions where Population III star formation can still be feasible.

This is the first of this series of papers in which we develop a model aimed at tracing self-consistently metal enrichment, following the hierarchical build-up of haloes. We consider metal enrichment and radiative feedback from the first stars together with feedback of successive generations of stars. Our study is applied in the context of BH and NC formation in order to determine: (i) if one of the mechanisms outlined above is dominant in the overall seed formation scenario; (ii) if a clear transition in redshift exists between different formation channels or if they co-exist in the same cosmic epoch; (iii) the characteristic redshift at which each process reaches its maximum efficiency; and (iv) when the first NCs are expected to form.

In order to address the above questions, we develop in this paper a semi-analytical scheme, aimed at describing the evolution of a single halo in which a central massive object forms. We adopt a ΛCDM cosmology with $\Omega_{\rm b} = 0.041$, $\Omega_{\rm m} = 0.258$, $\Omega_\Lambda = 0.742$, $h = 0.742$ and $n_{\rm s} = 0.963$ as given by five-year *WMAP* data (Dunkley et al. 2009).

The outline of this paper is as follows. In Section 2, we describe our recipes for the formation and evolution of a gas disc, hosted in a high-redshift dark matter halo. In Section 3, we set our model for the formation of NCs resulting from dynamically unstable discs and of BH seeds ensuing from runaway collapse of stars in NCs. In Section 4, we illustrate the results and compare the results of DV09 with those found with our new technique. Finally, in Section 5, we briefly summarize our model.

In the follow-up papers, we will apply our semi-analytical code (together with recipes for Population III star formation) to merger tree histories extracted from simulations run with the code PINOCCHIO (Monaco et al. 2002a). This will allow us to trace halo evolution, keeping track of their spatial positions also, a condition necessary in order to understand the relevance of radiative and SN feedback.

2 DISC FORMATION

We here consider a dark matter halo of mass $M_{\rm h}$, virializing at redshift $z_{\rm vir}$, with the density profile of an isothermal sphere. The virial radius, $R_{\rm vir}$, circular velocity, $V_{\rm h}$, and virial temperature, $T_{\rm vir}$, are inferred from the top-hat collapse model (see e.g. Barkana & Loeb 2001).

Gas falling into the potential well created by the dark matter is heated by adiabatic compression and shocks, attaining a mean temperature equal to $T_{\rm vir}$. We assign an initial gas mass $(\Omega_{\rm b}/\Omega_{\rm m})M_{\rm h}$ and assume that the gas density, $\rho_{\rm gas}(r)$, follows an isothermal profile as well.

It has been proposed that a critical metallicity, $Z_{\rm crit}$, exists above which star formation in the low-mass mode starts. We here focus only on those haloes whose gas has been enriched above $Z_{\rm crit}$. In the following, we provide the details of our treatment for the evolution of the halo gas component, star and BH seed formation.

2.1 Gas cooling

We assume the standard model for disc formation (White & Rees 1978; Mo, Mao & White 1998; Cole et al. 2000). To this purpose, for a given $T_{\rm vir}$ and metallicity, Z, of the gas, the cooling time-scale $\tau_{\rm cool}$ can be computed as

$$\tau_{\rm cool} = \frac{3}{2}\frac{\mu m_{\rm H}}{\rho_{\rm gas}(r)}\frac{k_{\rm B}T_{\rm vir}}{\Lambda(T_{\rm vir}, Z)}, \quad (1)$$

where $\Lambda(T_{\rm vir}, Z)$ is the cooling function. We adopt $\Lambda(T_{\rm vir}, Z)$ as tabulated in Sutherland & Dopita (1993). The cooling radius $r_{\rm cool}(t)$ at time t is defined by imposing $\tau_{\rm cool} = t$. Matter inside $r_{\rm cool}$ has enough time to cool down from its initial temperature.

For the disc to form, the gas has to both cool down and collapse. The time-scale for disc formation will therefore be the longest between the cooling and the free-fall timescales. In analogy with the computation of $r_{\rm cool}(t)$, we can define a free-fall radius $r_{\rm ff}(t)$ that sets the radius inside which the gas has sufficient time to free-fall into the central structure. During the initial phase of the collapse, a

mass of cold gas, $M_{\rm cold}$, develops at a rate

$$\dot{M}_{\rm cold} = 4\pi\rho_{\rm gas}(r_{\rm min}(t))r_{\rm min}^2(t)\dot{r}_{\rm min}, \quad (2)$$

where $r_{\rm min} = \min[r_{\rm cool}, r_{\rm ff}]$.

The cold gas is assumed to condense into a rotationally supported disc, whose angular momentum, $J_{\rm d}$, is a fraction $j_{\rm d}$ of the angular momentum $J_{\rm h}$ of the halo. If the specific angular momentum of baryons is conserved during collapse, $j_{\rm d} = M_{\rm cold}/M_{\rm h}$.

During the first phase of collapse, we assume the surface density profile of the disc to follow a Mestel, isothermal profile: $\Sigma(R) = \Sigma_0(R_0/R)$. Scale parameters, Σ_0 and R_0, are calculated by imposing that the disc mass is $M_{\rm cold}$ and its angular momentum is $J_{\rm d}$ (see Mo et al. 1998; Lodato & Natarajan 2006; DV09). Σ_0 and R_0 as a function of the halo parameters are given by[1]

$$\Sigma_0 = \frac{10\, m_{\rm d}(m_{\rm d}/j_{\rm d})^2\, H(z)\, V_{\rm h}}{16\pi G\lambda^2} = 70\left(\frac{m_{\rm d}}{0.05}\right)\left(\frac{V_{\rm h}}{15\,{\rm km\,s^{-1}}}\right)\left(\frac{\lambda}{0.05}\right)^{-1}{\rm M_\odot\,pc^{-2}}, \quad (3)$$

$$R_0 = 2\sqrt{2}\left(\frac{j_{\rm d}}{m_{\rm d}}\right)\lambda R_{\rm h} = 100\left(\frac{\lambda}{0.05}\right)\left(\frac{R_{\rm h}}{700\,{\rm pc}}\right)\,{\rm pc}, \quad (4)$$

where $H(z)$ is the Hubble constant as a function of redshift, calculated here at redshift 10, $m_{\rm d} = M_{\rm cold}/M_{\rm h}$, $c_{\rm s}$ is the sound speed and λ is the spin parameter. The vertical structure of the disc is determined by solving the equation for hydrostatic equilibrium. For a disc that is isothermal and self-gravitating, this implies a z-dependency for the density $\propto \cosh^{-2}(z/H)$, where H is the vertical scaleheight and is equal to $H(R) = c_{\rm s}^2/\pi G\Sigma(R)$. The scaleheight of the disc depends on its thermodynamical state, and it finally relies on the cooling channels available. In this series of papers, we focus on the behaviour of low-metallicity systems. At the early stages of collapse, the main coolant for the gas would be molecular hydrogen. In this first paper, however, we concentrate on haloes illuminated by a flux in the Lyman–Werner (LW) band (11.2–13.6 eV). LW photons can photodissociate H_2 molecules, thus suppressing cooling below 8000 K. The critical fluxes required to suppress H_2 in massive haloes ($T_{\rm vir} > 10^4$ K) are high compared to the expected background level at the redshifts of interest. Dijkstra et al (2008) and Shang, Bryan & Haiman (2010) suggested that haloes prone to H_2 suppression are a small subset of all haloes that sample the bright-end tail of the fluctuating cosmic ultraviolet (UV) background, due to the presence of a close luminous neighbour. We will consider this more general treatment in Paper II (the second paper in this series of papers), where LW flux ensuing from star formation history in a cosmological context will be calculated. Given $T_{\rm gas} = 8000$ K and for typical parameters of our systems, the scaleheight at the transition radius (see below) is $H(R)/R \sim 0.3$ in the discs.

In the following sections, we describe how scale parameters of the disc vary with cosmic time since (i) the disc mass keeps increasing due to the infall of fresh cold gas accreting from the halo (as described by equation 2); and (ii) the gas content changes due to gravitational instabilities that trigger episodes of inflow and star formation (outlined in Section 2.2).

[1] Note that the evolution of the collapsing structure and the initial assembly of the disc is described in a way similar to DV09. The main difference is that the assembly of the disc is followed here by solving explicitly for the time-evolution of the cold gas, instead of assuming that a fixed fraction of the halo gas cools down instantaneously.

2.2 Gravitational instabilities

The stability of the disc is described in terms of the Toomre parameter Q (Toomre 1964). As the surface density increases with more cold gas mass added to the disc, Q decreases and eventually drops below the critical value for disc stability, $Q_{\rm c}$. At this point the disc develops bar-like structures that can lead to a redistribution of mass and angular momentum. Note that the discs we are considering are rather thick discs, while the Toomre parameter is calculated for thin discs. A proper treatment for the 'fat' systems we are studying here should then rely on different criteria (see e.g. Ostriker & Peebles 1973; Efstathiou, Lake & Negroponte 1982; Christoudoulou, Shlosman & Tohline 1995). We assume $Q_{\rm c} = 2$ in our reference model; in Section 4.3 and in Paper II, we will consider different values of $Q_{\rm c}$ in order to take into account the uncertainties in the stability criterion.

Inflow of gas driven by gravitational instabilities can be described in terms of an effective viscosity ν. The inflow rate is given by Lin & Pringle (1987) (see also Lodato 2007):

$$\dot{M}_{\rm grav} = 2\pi\nu\Sigma(R)\left|\frac{{\rm d}\ln\Omega(R)}{{\rm d}\ln R}\right|, \quad (5)$$

where Ω is the angular velocity and

$$\nu = \eta\left(\frac{Q_{\rm c}^2}{Q^2} - 1\right)\frac{c_{\rm s}^3}{\pi G\Sigma(R)}. \quad (6)$$

In 'fat' discs, such as those considered here, the inflow might be driven by global, rather than local, gravitational instabilities (Shlosman, Frank & Begelman 1989; Shlosman, Begelman & Frank 1990; Balbus & Papaloizou 1999). Global wave transport is expected to contribute to the mass and angular momentum redistribution in this type of structures, working on a dynamical time-scale. We note that the assumption of a viscosity-driven behaviour of the disc is valid only for discs with $H/R < 0.1$ (Lodato & Rice 2004) and is a poor description for most of the systems considered here. In order to take into account the uncertainties in our treatment of $\dot{M}_{\rm grav}$, we adopt different values of η. We set in our reference model $\eta = 0.3$ (Gammie 2001), but we also compare our result with simulations run with $\eta = 1$ (high-efficiency inflow) and $\eta = 0.1$ (low-efficiency inflow).

The inflow rate that channels mass in the central region of the disc is then

$$\dot{M}_{\rm grav} = \frac{2\eta c_{\rm s}^3}{G}\left[\left(\frac{Q_{\rm c}}{Q}\right)^2 - 1\right] = \frac{\eta c_{\rm s}}{GV_{\rm h}^2}\left[\pi^2G^2Q_{\rm c}^2\Sigma_0^2R_0^2 - 2c_{\rm s}^2V_{\rm h}^2\right], \quad (7)$$

which corresponds to $7.5\times10^{-3}\,{\rm M_\odot\,yr^{-1}}$ for $Q = 1.9$, $\eta = 0.3$ and $T_{\rm gas} = 8000$ K.

We assume that the inner surface density profile changes its scaling with radius as $\propto R^{-5/3}$, according to simulations by Mineshige & Umemura (1997). The transition radius, $R_{\rm tr}$, between the two profiles is given by

$$R_{\rm tr} = \frac{M_{\rm inf}}{4\pi\Sigma_0R_0} = 1\,{\rm pc}\left(\frac{M_{\rm inf}}{10^5\,{\rm M_\odot}}\right)\left(\frac{70\,{\rm M_\odot\,pc^{-2}}}{\Sigma_0}\right)\left(\frac{100\,{\rm pc}}{R_0}\right), \quad (8)$$

where $M_{\rm inf}$ is the net mass inflowing from $R > R_{\rm tr}$. $R_{\rm tr}$ is calculated at each time t requiring that $M_{\rm inf}$ is added to the mass already in place inside $R_{\rm tr}$ (DV09).

For a given set of $M_{\rm h}$, $T_{\rm gas}$ and $Q_{\rm c}$, a maximum angular momentum exists below which a disc becomes unstable $\lambda < 0.125(\Omega_{\rm b}/\Omega_{\rm m})Q_{\rm c}\sqrt{T_{\rm vir}/T_{\rm gas}}$, where $T_{\rm gas}$ is the temperature of the cold gas disc. Discs with higher rotation are more stable against fragmentation and the amount of mass that needs to be removed from the Mestel disc in order to reach $Q = Q_{\rm c}$ is lower. The inflow mass therefore decreases for haloes with large spin parameters.

2.3 Star formation

If the disc grows strongly unstable, star formation sets in, consuming part of the gas that would otherwise flow into the central region. Further evolution of $M_{\rm inf}$ is then regulated by the competition between star formation in the outer Mestel disc and gas inflow in its centre.

An unstable disc starts fragmenting into bound clumps, once the gravitationally induced stress exceeds a critical value corresponding to $\dot{M}_{\rm crit} = 0.06\, c_{\rm s}^3/G$ (Lodato & Natarajan 2006).

Many authors have suggested that the presence of a certain amount of metals is the key ingredient in order to produce efficient fragmentation and star formation in the low-mass mode (Bromm et al. 1999, 2002; Schneider et al. 2006; Clark et al. 2008; Omukai et al. 2008 and references therein). In order to estimate the value of $Z_{\rm crit}$ at which gas starts to fragment, one can compare the cooling rate $\Lambda_{\rm cool}$ to the adiabatic heating rate $\Gamma_{\rm ad}$. Fragmentation requires that $\Lambda_{\rm cool} \geqq \Gamma_{\rm ad}$. Once this happens, the temperature starts to decrease as the density increases, until cooling becomes inefficient. The condition $\Gamma_{\rm ad} < \Lambda_{\rm cool}$, in fact, constrains the metallicity at a given temperature and density to be greater than a critical value $Z_{\rm crit}$. For an initially isothermal gas, this condition provides a relationship between metallicity and density. Equivalently, if the gas is characterized by a fixed metallicity Z, then only those regions with density greater than a given threshold, $n_{{\rm crit},Z}$, are able to cool down efficiently and fragment into stars. This is the condition that we apply to collapsing discs to determine if they can form stars. The behaviour of a gas cloud collapsing in presence of a UV flux has been studied by Omukai et al. (2008). They showed that the collapse of the clouds initially proceeds, thanks to atomic hydrogen. Only at higher enough densities, H_2 forms and cooperates in the cooling process. The presence of even a small amount of metals allows fragmentation of the gas and star formation in the low-mass mode.

As the value of $Z_{\rm crit}$ depends on the exact cooling channels, that are not followed by our simulations, we can only give a rough estimate of the characteristic density at which star formation is allowed. The $Z_{\rm crit}$–$n_{{\rm crit},Z}$ relations adopted in our study are those found by Santoro & Shull (2006). They assume that Population III stars are the main sources of pollution and consider different species of coolants as expected from Population III stellar evolution (see fig. 9 in Santoro et al. 2006). We adopt, as a reference, the curve corresponding to Population III stars in the intermediate-mass range (mass range 185–205 $\rm M_\odot$). In DV09, we studied how the uncertainties in the $Z_{\rm crit}$–$n_{{\rm crit},Z}$ relation affect the ability of our system to form NCs. We found that different relations correspond to different characteristic formation redshifts for BH seeds. We will further address this topic in Paper II where we will adopt different $Z_{\rm crit}$–$n_{{\rm crit},Z}$ relations in order to better understand how the uncertainties on $Z_{\rm crit}$ can influence the NC and BH seed formation.

Once a $Z_{\rm crit}$–$n_{{\rm crit},Z}$ relation is assumed, given the density profile and Z of a disc, we can infer a characteristic radius, $R_{\rm SF}$, inside which fragmentation sets in. For the systems we are interested in

$$R_{\rm SF} = \frac{\Sigma_0 R_0}{c_{\rm s}}\sqrt{\frac{\pi G}{2\mu m_{\rm u} n_{{\rm crit},Z}}} = 30\,{\rm pc}\left(\frac{\Sigma_0}{70\,{\rm M_\odot\,pc^{-2}}}\right)\left(\frac{R_0}{100\,{\rm pc}}\right), \quad (9)$$

and we allow star formation only inside $R_{\rm SF}$.

We denote the time-lag between cloud formation and collapse to form stars as $t_{\rm lag}$. This is basically the lifetime of molecular clouds; in our model, it corresponds to the delay between the onset of instability (that we assume to be the driver leading to the formation of the clouds) and the start of the star formation event. Observational estimates for lag times range from 1 to 7 Myr (Klessen, Krumholz & Heitsch 2009). We assume $t_{\rm lag} = 3.5$ Myr. We have checked that our results are not strongly affected by the assumed value of $t_{\rm lag}$: final BH seed masses change by a factor of 2–3 for $t_{\rm lag}$ ranging between 1 and 7 Myr.[2]

We assume a star formation rate surface density, $\Sigma_{\rm SFR}$, that follows the empirical Schmidt–Kennicutt law (Kennicutt 1998):

$$\Sigma_{\rm SFR} = 2.5\times10^{-4}\left[\frac{\Sigma(R)}{1\,{\rm M_\odot\,pc^{-2}}}\right]^{1.4}\,{\rm M_\odot\,yr^{-1}\,kpc^{-2}}, \quad (10)$$

and calculate the star formation rate, $\dot{M}_{*,{\rm d}}$, by integrating over the region of the disc, where stars form. We thus integrate $\Sigma_{\rm SFR}$ between $R_{\rm tr}$ and $R_{\rm SF}$ to infer a star formation rate, $\dot{M}_{*,{\rm d}}$, in the outer disc:

$$\dot{M}_{*,{\rm d}} = 0.002\left[\frac{\Sigma_0}{78\,{\rm M_\odot\,pc^{-2}}}\frac{R_0}{170\,{\rm pc}}\right]^{7/5}\left[\left(\frac{R_{\rm SF}}{20\,{\rm pc}}\right)^{3/5} - \left(\frac{R_{\rm tr}}{0.3\,{\rm pc}}\right)^{3/5}\right]{\rm M_\odot\,yr^{-1}}. \quad (11)$$

Star formation decreases the amount of gas that can be funnelled in the inner disc. For this reason and as long as $\dot{M}_{\rm grav} > \dot{M}_{\rm crit}$, we assume a net inflow rate, $\dot{M}_{\rm inf}$, equal to

$$\dot{M}_{\rm inf} = \dot{M}_{\rm grav} - \dot{M}_{*,{\rm d}}. \quad (12)$$

2.4 Supernova feedback

Approximately 3 Myr after an episode of star formation, the most massive stars explode as SNe. Metals processed in their centres are released into the surrounding gas, increasing the metal content. Metal yields and energy production depend on the initial mass function (IMF) of the stars. We here assume a Salpeter IMF with minimum ($M_{\rm l}$) and maximum ($M_{\rm u}$) masses of 0.1 and 100 $\rm M_\odot$, respectively, normalized to 1 $\rm M_\odot$.

We correlate the rate of metal production with the star formation rate in the disc, assuming that the total mass in metals scales as

$$\dot{M}_{\rm met} = Y_{\rm met}\,\nu_{\rm SN}\,\dot{M}_*, \quad (13)$$

where $Y_{\rm met}$ is the total IMF-averaged metal yield and $\nu_{\rm SN}$ is the number of SNe exploding after the formation of a mass M_* in stars divided by M_*. We assume that SN progenitors are stars with masses between 10 and 50 $\rm M_\odot$ (Scannapieco et al. 2003).[3] With our prescriptions, $Y_{\rm met} = 1.6\,\rm M_\odot$ and $\nu_{\rm SN} = 0.004\,84$.

[2] Note that the initial absence of H_2 could further delay the collapse of the clouds.

[3] We here follow Tsujimoto et al. (1995) and assume that stars with masses greater than 50 $\rm M_\odot$ collapse into a BH without ejecting heavy metals.

SN explosions in high-redshift haloes can be extremely destructive. SN-driven bubbles can push away part of the gas in the halo, eventually fully depriving the host of its gas. The efficiency of gas depletion depends on the depth of the potential well of the host and on the energy released during the explosion.

Using energy conservation, we compute the amount of mass that is removed from the disc plus halo as a result of SN explosions, assuming that each single explosion releases an energy $E_{\rm SN} = 10^{51}$ erg (Woosley & Weaver 1995). The resulting outflow rate, $\dot{M}_{\rm sh}$, is

$$\dot{M}_{\rm sh} = \frac{f_{\rm w} \nu_{\rm SN} E_{\rm SN} \dot{M}_*}{2\,(1+\Omega_{\rm b}/\Omega_{\rm m})\,G M_{\rm h}/R_{\rm vir} + (1/2) v_{\rm sh}^2}, \quad (14)$$

where $f_{\rm w}$ is the fraction of energy channelled into the outflow and $v_{\rm sh}$ is the velocity of the outflow at the virial radius. We refer the reader to Appendix A for details of our recipes for gas depletion. Finally, we define the retention fraction coefficient $f_{\rm ret} \equiv (M_{\rm gas} - M_{\rm sh})/M_{\rm gas}$. After the explosion, we assume that a fraction $f_{\rm ret}$ of the metals yields produced remains in the host.

2.5 Disc evolution

The evolution of an unstable disc is followed by tracing the time dependence of $M_{\rm cold}$, $M_{\rm inf}$, $M_{*,\rm d}$ and $M_{\rm sh}$, integrating equations (2), (11), (12) and (14), respectively, assuming that initially all the gas mass is hot ($T = T_{\rm vir}$).

Changes in these quantities over cosmic time correspond to changes in the scale parameters of the Mestel disc (Σ_0, R_0) and $R_{\rm tr}$. Angular momentum conservation and total mass conservation in the system are imposed to infer the following equations linking the present values of $\Sigma_0(t)$ and $R_0(t)$ to their value at the time, $\tau_{\rm inst}$, in which the disc reaches the instability regime [$\Sigma_0(\tau_{\rm inst})$ and $R_0(\tau_{\rm inst})$]:

$$\Sigma_0(t) = \Sigma_0(0) \left[1 - \frac{m_{\rm a}(t)}{m_{\rm d}(t)}\right]^3, \quad (15)$$

$$R_0(t) = R_0(0) \left[1 - \frac{m_{\rm a}(t)}{m_{\rm d}(t)}\right]^{-1}, \quad (16)$$

where $m_{\rm d}(t) = M_{\rm cool}(t)/M_{\rm h}$ and $m_{\rm a}(t)$ is the fraction of gas mass removed from the outer disc, via star formation, inflow and outflows from SNe. The disc alternates episodes of instability to states of quiescence: as Q increases above $Q_{\rm c}$, the instability is quenched. Disc parameters evolve regulated by the amount of new mass that is added by accretion from the halo.

We halt the simulations as soon as the hot halo gas is consumed and the disc is in a stable state. In the next section, we describe the evolution of the gas that has been funnelled at the centre of the disc, where we expect that an NC and/or a BH seed forms.

3 CENTRAL MASSIVE OBJECT FORMATION

In the previous section, we focused on the evolution of Toomre unstable discs. Here, we study the formation of an NC from a mass reservoir, $M_{\rm inf}$, funnelled into $R_{\rm tr}$ after $t_{\rm lag}$.

Star formation in the central part of the disc can proceed with a higher efficiency with respect to the outer disc as higher pressure and density are expected there (Elmegreen & Efremov 1997; Li, Haiman & Mac Low 2007). We treat the gas mass accumulated within $R_{\rm tr}$ as a single collapsing cloud, provided that the metallicity–density threshold for star formation is reached within $R_{\rm tr}$. Here, the efficiency factor, $\epsilon_{\rm SF}$, with which gas is converted into stars is kept equal to 1 (Li et al. 2007). In Paper II, we will also show a model assuming $\epsilon_{\rm SF} = 0.25$ in order to evaluate how our predictions change, if we adopt lower efficiencies. The mass and radius of the cluster are $M_{\rm cl} = \epsilon_{\rm SF} M_{\rm inf}$ and $R_{\rm cl} = \min(R_{\rm SF}, R_{\rm tr})$. The latter equation ensures that NC formation is allowed only if the proper thermodynamical conditions are reached.

NCs formed in this way are very compact and unstable objects. Numerical studies of stellar clusters have shown that gravitational encounters between stars can drive the clusters towards core collapse, leading to a contraction of the core, where physical collisions between stars are frequent.[4] Core collapse is driven by dynamical friction of the more massive stars. Star–star collisions between massive stars in the cluster core can happen in a runaway fashion, leading to the build-up of a VMS, whose mass can be as high as $1000\,{\rm M}_\odot$ (Portegies Zwart et al. 1999, 2004; Gürkan et al. 2004; Freitag et al. 2006a,b).

After $\sim$3 Myr, the first SN explodes. For the systems we are interested in, this happens before the inflow process ends. We assume that an NC is able to produce a BH seed only if it undergoes core collapse before the first SN explosion. We compute the mass of the nuclear star cluster, $M_{\rm cl,0}$, 3 Myr after the onset of star formation, and we infer the core collapse time-scale, $t_{\rm cc}$. For those systems with $t_{\rm cc} \leq 3$ Myr, we calculate the mass for the VMS, $M_{\rm VMS}$, as

$$M_{\rm VMS} = m_* + 4 \times 10^{-3} M_{\rm cl} f_{\rm c} \ln\lambda_{\rm C} \ln\left(\frac{3\,{\rm Myr}}{t_{\rm cc}}\right) = 765 \left(\frac{M_{\rm cl}}{10^5\,{\rm M}_\odot}\right) \ln\left(\frac{3\,{\rm Myr}}{t_{\rm cc}}\right) {\rm M}_\odot, \quad (17)$$

where $f_{\rm c} = 0.2$ and $\ln\lambda_{\rm C} = \log(0.1 M_{\rm cl,0}/\langle m\rangle)$ (Portegies Zwart & McMillan 2002, see also DV09).

At low metallicity, stars more massive than $260\,{\rm M}_\odot$ are expected to leave a remnant BH that retains most of its progenitor mass. Each time $M_{\rm VMS} > 260\,{\rm M}_\odot$, we assume that a BH seed is formed with $M_{\rm BH} = M_{\rm VMS}$. For a BH seed to form, the stellar cluster needs a total mass $\geq 260\,{\rm M}_\odot$ in massive stars that can segregate. Here, we assume that stars more massive than $10\,{\rm M}_\odot$ efficiently segregate in the cluster core. Once all these more massive stars are consumed, collisions are assumed to stop. For our chosen IMF, a minimum VMS mass of $260\,{\rm M}_\odot$ corresponds to a minimum cluster mass of $1.5 \times 10^4\,{\rm M}_\odot$. We here include this caveat and assume that BH seeds can form only in those nuclear clusters that are more massive than this threshold.

The ability of runaway collisions to form an intermediate-mass BH also depends on the metallicity of the gas from which the stars originated. Recent simulations of star–star collisions have shown that at metallicity higher than $\sim 10^{-1}\,{\rm Z}_\odot$, physical collisions are too disruptive and the build-up of a star with mass greater than $260\,{\rm M}_\odot$ is inhibited (Glebbeek et al. 2009). Mass-loss depends on the metallicity of the star. Wind models predict that mass-loss depends on metallicity as a power law, Z^α, with α ranging between 0.5 and 0.8. We assume that a BH seed forms only at metallicity $<10^{-3}\,{\rm Z}_\odot$. For this value of Z, mass-loss during the stellar lifetime

[4] Note that these studies concentrate on halo stellar clusters. NC dynamics can be altered with respect to their halo counterparts by (i) the presence of the external stellar environment; and (ii) continuous gas accretion. These two effects work in opposite directions: cluster contraction can be reverted as stars of the more rarefied outer system are trapped into the central clusters. Inflow of gas can instead drive an acceleration of core collapse as it deepens the potential well. The final outcome resulting from the competition between them is not easy to infer and it is beyond the scope of this paper. We will address this issue in a follow-up paper.

should be negligible and the final mass of the remnant is expected to differ only slightly from that of its progenitor star, unless rotation is taken into account.

Summarizing, BHs form, provided that the following conditions are met: (i) $Z > Z_{\rm crit}$, so that an NC can form but $Z < 10^{-3}\,{\rm Z}_{\odot}$; (ii) λ is sufficiently small for the disc to reach gravitational instability; (iii) clusters form with $t_{\rm cc} < 3$ Myr; and (iv) massive enough to host an adequate amount of stars with $m_* > 10\,{\rm M}_{\odot}$.

4 RESULTS

In this section, we describe how disc properties are shaped by gravitational instabilities and star formation. Our reference halo has mass $M_{\rm h} = 6.5 \times 10^7\,{\rm M}_{\odot}$, $R_{\rm vir} = 1.2$ kpc, $T_{\rm vir} = 1.46 \times 10^4$ K, $V_{\rm h} = 15\,{\rm km\,s^{-1}}$ and forms at $z_{\rm vir} = 10$; we further adopt $\lambda = 0.05$, $Z/{\rm Z}_{\odot} = 10^{-4}$, $\eta = 0.3$, $Q_{\rm c} = 2$, $\epsilon_{\rm SF} = 1$ and $t_{\rm lag} = 3.5$ Myr.

4.1 Disc and cluster properties

In this subsection, we illustrate the evolution of the reference disc model. After a time $\tau_{\rm inst} = 1.3 \times 10^8$ yr, the disc has grown enough to hit the instability threshold, as it has assembled enough mass so that the Toomre parameter drops below $Q_{\rm c}$. At this time $\Sigma_0 = 78\,{\rm M}_{\odot}\,{\rm pc}^{-2}$ and $R_0 = 170$ pc. The upper (lower) panel in Fig. 1 shows $M_{\rm inf}(\dot{M}_{\rm inf})$ and $M_{*,{\rm d}}(\dot{M}_{*,{\rm d}})$ as a function of time t, from the time the disc becomes unstable till the time of formation of the central BH. The evolution proceeds along different phases:

(1) $\tau_{\rm inst} < t < \tau_{\rm inst} + t_{\rm lag}$: at $t = \tau_{\rm inst}$, the disc starts collecting gas into the centre and this leads to an enhancement of $M_{\rm inf}$. The disc parameters are such that $\dot{M}_{\rm grav} > \dot{M}_{\rm crit}$ (horizontal dotted line in the right-hand panel of the figure). Clump formation in the disc sets in but these newly formed clouds have not had enough time to collapse and form stars until a time $t_{\rm lag}$ has elapsed.

(2) $\tau_{\rm inst} + t_{\rm lag} < t < \tau_{\rm inst} + t_{\rm lag} + 3$ Myr: star formation begins. The critical density for star formation remains constant as metals have not been released yet and the star formation radius still remains close to its initial value. Both $M_{\rm inf}$ and $M_{*,{\rm d}}$ increase until the first SN explodes. We label the nuclear mass at this time as $M_{\rm cl,0}$.

(3) $t > \tau_{\rm inst} + t_{\rm lag} + 3$ Myr: the first SN explode affecting the subsequent evolution of the system. Part of the metals produced in stars are released in the surrounding gas, increasing the metallicity above $Z_{\rm crit}$. At the same time, SN explosions evacuate gas from the system, lowering the surface density. As a result, the Toomre parameters increase above $Q_{\rm c}$.

The disc then lingers on a marginally stable configuration, thanks to the high $\dot{M}_{\rm sh}$ that is able to counterbalance $\dot{M}_{\rm cool}$. Fragmentation in the disc is temporary halted as $\dot{M}_{\rm grav}$ drops below $\dot{M}_{\rm crit}$. At later times, the surface density starts to increase again, as hot gas in the halo continues to cool. The disc then enters a new cycle of instability and evolution terminates when the hot gas reservoir has been consumed.

4.2 Dependence on input parameters

We here explore disc evolution as a function of the halo mass, $M_{\rm h}$, for three different values of inflow efficiency, η, and as a function of the initial metallicity Z.

Fig. 2 shows the NC mass, $M_{\rm cl,0}$ (solid lines), and BH mass, $M_{\rm BH}$ (long-dashed lines), at the time of BH formation and the stellar disc mass, $M_{*,{\rm d}}$ (dash–dotted lines), and net inflow mass, $M_{\rm inf}$ (dotted lines), at the end of the simulation, as a function of $M_{\rm h}$, for different values of η.

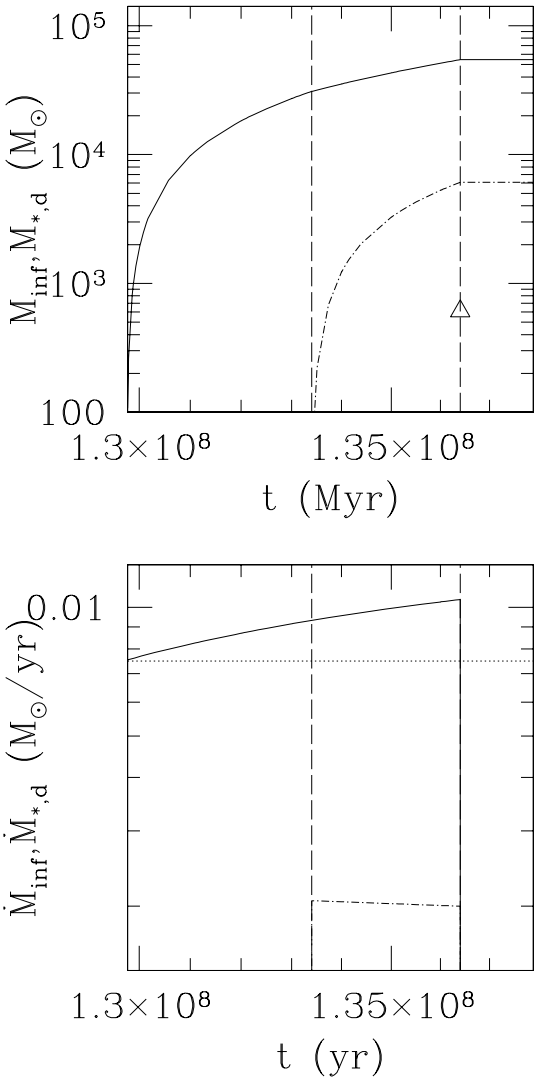

Figure 1. Upper panel: $M_{\rm inf}$ (solid line) and $M_{*,{\rm d}}$ (dash–dotted line) versus time t for our reference models. The triangle highlights the BH mass at the moment of its formation. Lower panel: $\dot{M}_{\rm inf}$ (solid line) and $\dot{M}_{*,{\rm d}}$ (dash–dotted line) for our reference model as a function of time. The horizontal dotted line marks the critical threshold for fragmentation $\dot{M}_{\rm crit}$. The vertical lines define the different evolutionary phases of the disc properties as described in the text.

The NC mass, $M_{\rm cl,0}$, depends weakly on the host mass, $M_{\rm h}$, and somewhat more strongly on the inflow efficiency, η. By contrast, the final stellar disc mass depends significantly on $M_{\rm h}$. This is because (i) more massive haloes have a larger gas mass and retention efficiency ($f_{\rm ret}$); and (ii) gas cools more rapidly in larger haloes. This allows for a larger departure of the Toomre parameter Q from $Q_{\rm c}$ and larger inflow ($\dot{M}_{\rm grav}$, see equation 7). $\dot{M}_{\rm grav}$ is larger than the critical inflow triggering fragmentation, $\dot{M}_{\rm crit}$, and more stars form in the outer disc. For $\eta = 0.1$, a threshold in $M_{\rm h}$ exists, below which $\dot{M}_{\rm grav} < \dot{M}_{\rm crit}$ always, and $M_{*,{\rm d}}$ drops to zero.

The dotted lines in Fig. 2 show the net inflow mass, $M_{\rm inf}$, at the end of the simulations. $M_{\rm inf}$ increases with increasing $M_{\rm h}$ as a result of both initially larger gas masses and deeper potential wells.

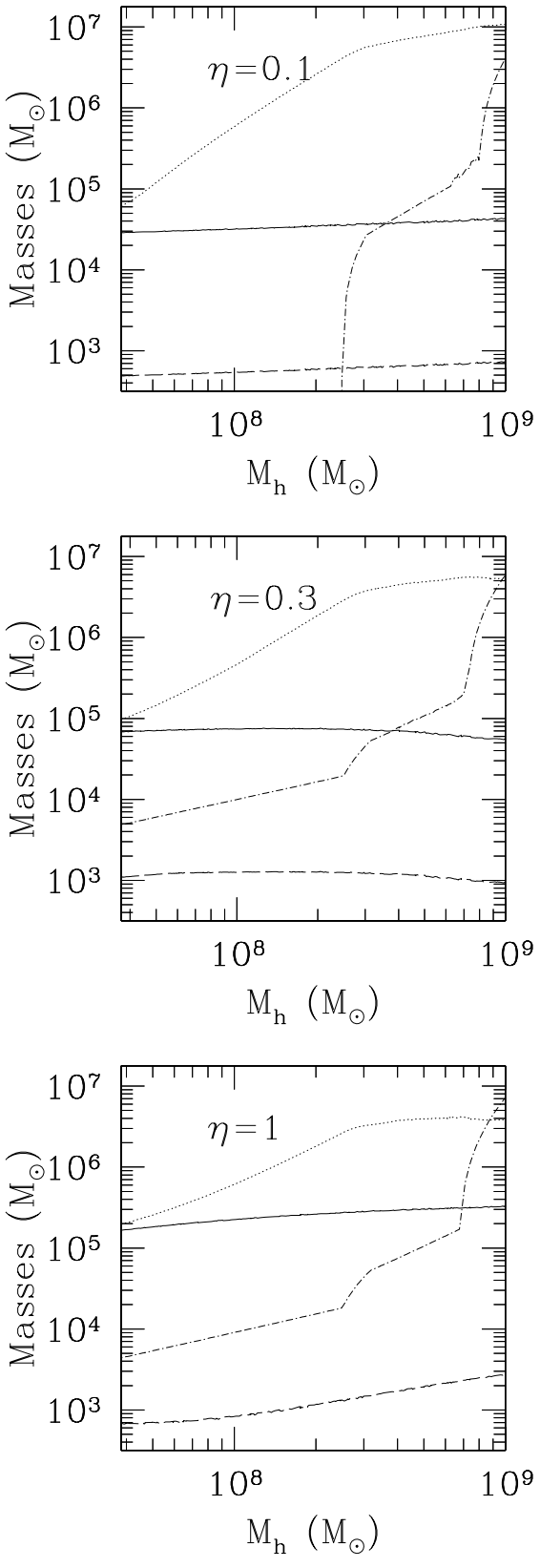

Figure 2. Mass $M_{\rm cl,0}$ (solid lines) of the NC and $M_{\rm BH}$ (dashed lines) at the time the BH forms, and $M_{\rm inf}$ (dotted lines) and $M_{*,\rm d}$ (dash–dotted lines) at the end of the simulations, as a function of $M_{\rm h}$ for different inflow efficiencies. The upper, middle and lower panels refer to $\eta = 0.1$, 0.3 and 1, respectively. Reference properties for the halo are $z_{\rm vir} = 10$ and $\lambda = 0.05$, and for the baryonic component $Z/{\rm Z}_\odot = 10^{-4}$, they are $Q_{\rm c} = 2$ and $\epsilon_{\rm SF} = 1$.

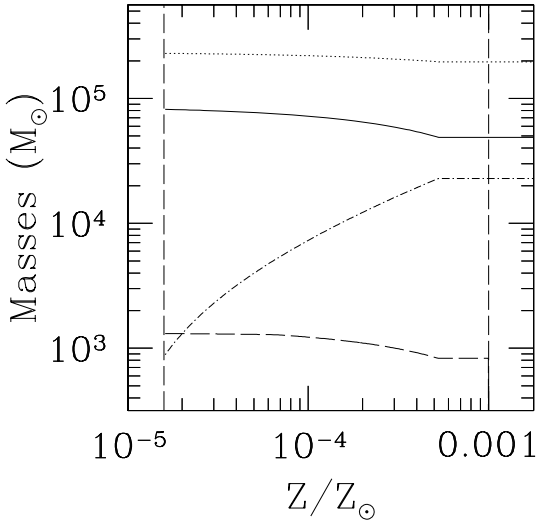

Figure 3. $M_{\rm cl,0}$ (solid lines), $M_{\rm BH}$ (dashed lines), $M_{\rm inf}$ (dotted lines) and $M_{*,\rm d}$ (dash–dotted lines) as a function of the initial metallicity of the gas for our reference halo. The halo mass is kept constant at its reference value $M_{\rm h} = 6.7 \times 10^7\,{\rm M}_\odot$. The vertical lines mark the minimum metallicity for low-mass star formation, $Z_{\rm crit} \simeq 10^{-4.87}\,{\rm Z}_\odot$, and the maximum metallicity that allows BH formation, $Z_{\rm crit} \simeq 10^{-3}\,{\rm Z}_\odot$.

$M_{\rm inf}$ depends weakly on η. For higher η, extended star formation is allowed for longer timescales. As a consequence, more gas is converted into stars, before it can reach the nucleus and final $M_{\rm inf}$ are lower.

Disc evolution also depends on the initial metallicity Z. The metallicity determines when stars can form and thus it is a key parameter: in our model, there is a minimum metallicity, where 'normal' stars start forming, $Z_{\rm crit} \simeq 10^{-4.87}\,{\rm Z}_\odot$ and a maximum metallicity that allows BH formation, $Z_{\rm crit} \simeq 10^{-3}\,{\rm Z}_\odot$.

In Fig. 3, the NC mass, $M_{\rm cl,0}$ (solid line), BH mass, $M_{\rm BH}$ (dashed line), net inflow mass, $M_{\rm inf}$ (dotted line), and stellar disc mass, $M_{*,\rm d}$ (dash–dotted line), are plotted as a function of the initial gas metallicity. Star formation sets in only for $Z > Z_{\rm crit}$ (left-hand vertical dashed line in Fig. 3). For the majority of the systems we are considering here, $R_{\rm SF} > R_{\rm tr}$. The first three quantities are almost independent of metallicity, as long as Z is in the range between the minimum and maximum value. The disc stellar mass is instead very sensitive to the exact value of the metallicity. $M_{*,\rm d}$ increases with Z (see equation 11). Mass is consumed instead of flowing into the nucleus and $M_{\rm cl,0}$ weakly decreases.[5] When the metallicity is large enough that $R_{\rm SF}$ equals R_0, then the whole disc forms stars and the dependence on Z disappears.

In our reference model, we set $\lambda = 0.05$. The distribution of spin parameters can be described as a lognormal distribution with peak at $\lambda \sim 0.03$–0.05 and dispersion $\sim$0.5. For our reference halo, the maximum λ allowing for instability is 0.07. Approximately, 96 per cent of the haloes in the interesting mass range have $\lambda < \lambda_{\rm max}$ and are therefore prone to disc instabilities.

[5] Note that for systems with $R_{\rm SF} < R_{\rm tr}$, the initial cluster mass, $M_{\rm cl,0}$, increases with Z as: (i) extended star formation is not allowed in the outer disc and the amount of mass funnelled into the nucleus does not depend on Z; and (ii) higher metallicity corresponds to higher $R_{\rm SF}$, so that a larger amount of nuclear gas can be converted into stars. This case was considered in DV09; we refer to that work for a discussion of this regime.

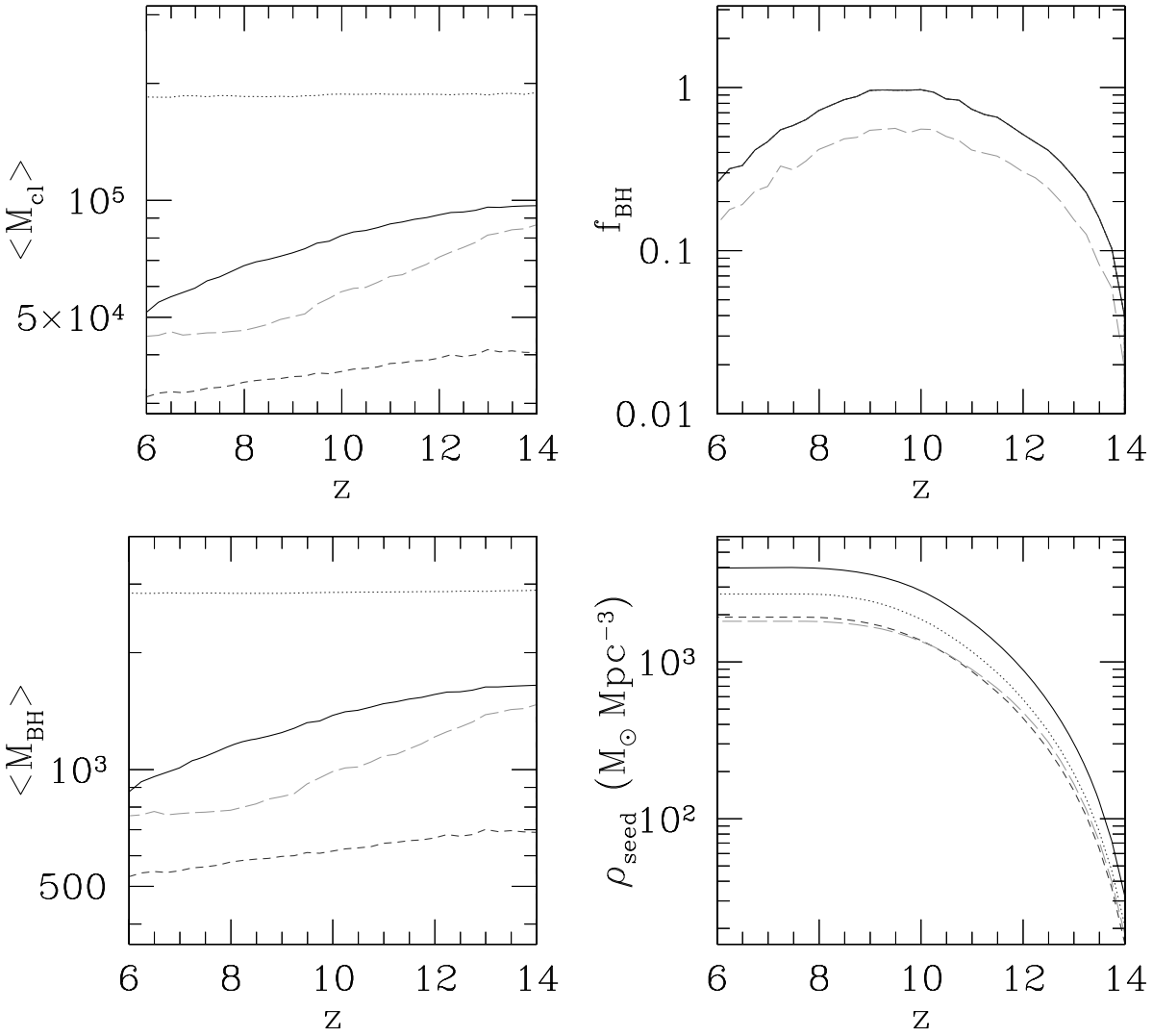

Figure 4. Panels from top to bottom (left-hand to right-hand side): $\langle M_{\rm cl,0}\rangle$ (for those clusters able to form a seed BH), $f_{\rm BH}$, $\langle M_{\rm BH}\rangle$ and $\rho_{\rm seed}$ as a function of redshift for different values of η and $Q_{\rm c}$. Line styles: solid lines refer to a model with $\eta = 0.3$ and $Q_{\rm c} = 2$, dotted lines refer to a model with $\eta = 1$ and $Q_{\rm c} = 2$, short-dashed lines refer to a model with $\eta = 0.1$ and $Q_{\rm c} = 2$ and long-dashed lines refer to a model with $\eta = 0.3$ and $Q_{\rm c} = 1$.

If $Q_{\rm c}$ is lowered to unity, instabilities are more easily quenched and the process becomes less efficient (see the upper right-hand panel in Fig. 4).

4.3 Comparison with previous models

In this section, we apply our prescriptions for gas cooling, star formation and inflow to determine the properties of the NC and BH populations in low-metallicity environments and compare them with those found in DV09.

To this purpose, we construct a sample of haloes with the same properties as those analysed in DV09. We focus on haloes with $z_{\rm vir}$ ranging between 5 and 30 corresponding to $T_{\rm vir}$ between 10^4 and 1.8×10^4 K (Barkana & Loeb 2001) and we determine their frequency using a modified version of the Press & Schechter formalism (Sheth & Tormen 1999) in a WMAP5 cosmology (Dunkley et al. 2009). To each halo, we assign a value of the spin parameter, λ, extracted from the probability distribution found in the Millennium simulation. We here adopt the same prescriptions for the metallicity evolution as in model A in DV09. Specifically, we assume a redshift-dependent metallicity that follows the relation $Z/{\rm Z}_\odot = 0.35 \times 10^{-0.36z}$ (see DV09 and Li 2008 for details). We allow a logarithmically uniform scatter in Z of $\Delta\log(Z)$, based on the observed scatter. We adopt the model described in this paper for the evolution of each disc halo. The model is an improvement of that discussed in DV09 as (i) it traces the time-evolution of the baryonic component; and (ii) it accounts for the more generic case in which star formation sets in the outer Mestel disc. More realistic models that follow metal enrichment self-consistently will be described in Paper II. We consider different inflow efficiencies $\eta = 1$ (dotted lines), 0.3 (solid lines) and 0.1 (short-dashed lines) for $Q_{\rm c} = 2$ and $\eta = 0.3$ for $Q_{\rm c} = 1$ (long-dashed lines).

In the upper left-hand panel of Fig. 4, we show the evolution of the NC mass, $\langle M_{\rm cl,0}\rangle$ (averaged over the sample), as a function of redshift. NCs start to form at $z = 14$, as haloes are polluted above the critical metallicity for fragmentation. This result is dictated by the particular metal-enrichment history and $Z_{\rm crit}$–$n_{\rm crit}$ relation chosen. As shown in DV09, a lower minimum $Z_{\rm crit}$ and/or a more gradual

metal-enrichment history lead to higher initial formation redshift of NCs and BH seeds.[6] BH seeds start to form after the formation of the first NCs. In the upper right-hand panel of Fig. 4, we compute the fraction, $f_{\rm BH}$, of haloes with $T_{\rm vir} > 10^4$ K hosting a BH seed, as a function of redshift. $f_{\rm BH}$ is regulated by the fraction of haloes with $Z_{\rm crit} < Z < 10^{-3}\,{\rm Z}_\odot$: it first increases as more and more haloes are polluted and then decreases when most haloes reach $Z = 10^{-3}\,{\rm Z}_\odot$. The fraction of haloes prone to BH formation, we find in this paper, are larger than those found in DV09 (see their fig. 8, upper panel) by a factor of 10. With the prescriptions we adopt here, $M_{\rm cool}$ evolves with time, regulated by cooling of the halo gas, whereas in DV09 a fixed $M_{\rm cool}/M_{\rm h}$ was used. This implies that for a given $Q_{\rm c}$, the fraction of unstable discs that form increases.

Mean BH masses $\langle M_{\rm BH}\rangle$ versus redshift z is plotted in Fig. 4 (lower left-hand panel). $\langle M_{\rm BH}\rangle$ spans values between a few hundred $\rm M_\odot$ and 2000 $\rm M_\odot$. For most of the NCs considered here, the growth of the VMSs is limited by the amount of stars with $m_* > 10\,{\rm M}_\odot$. Lower mean cluster masses thus lead to lower mass seed BHs. An exception occurs for the highest $M_{\rm cl,0}$ (i.e. $\eta = 1$): for these systems, $t_{\rm cc}$ is long and not all massive stars contribute in building up the VMS. $\langle M_{\rm BH}\rangle$ stays at a near constant value $\sim$900 $\rm M_\odot$ at any redshift.

Comoving BH seed densities are plotted in the lower right-hand panel. $\rho_{\rm seed}$ depends on both BH formation efficiencies and masses attained. They are found to depend weakly on the parameters: final $\rho_{\rm seed}$ differs only by a factor of 2, ranging between 2000 and 4000 $\rm M_\odot\,Mpc^{-3}$.

5 SUMMARY

In this paper, the first of this series, we presented a model for the formation of high-redshift, metal-poor NCs and their embedded BH seeds. The channel that we propose can be summarized as follows:

(i) After the virialization of a dark matter halo, hot gas cools down and the presence of angular momentum leads to the formation of a cold gas disc. The disc accretes gas from the halo and grows in mass, increasing its surface density to the extent that gravitational instabilities develop. This leads to the growth of non-axisymmetric perturbations that redistribute angular momentum, driving mass into the central part of the halo.

(ii) For those systems where the metallicity Z exceeds $Z_{\rm crit}$ and the inflow rate surpasses a critical threshold for fragmentation ($\dot{M}_{\rm grav} > \dot{M}_{\rm crit}$), star formation in the extended disc can set in, consuming part of the mass that would otherwise accrete to the centre. The inner density profile steepens inside a transition radius, where a compact nuclear stellar cluster forms.

(iii) In those NCs where the core-collapse time-scale is shorter than the main-sequence lifetime of massive stars (i.e. $t_{\rm cc} < 3$ Myr), core collapse anticipates stellar evolution. Runaway collisions between the most massive stars build up a VMS that finally implodes, leaving a remnant BH seed.

(iv) The evolution continues and it is driven by the competition between star formation, triggered by prolonged cooling of hot halo gas, and SN explosions that evacuate part of it.

(v) Evolution halts as soon as the hot halo gas is consumed/blown away and the disc settles into a state of stationary equilibrium.

We apply our recipes to a set of haloes with $T_{\rm vir} > 10^4$ and properties typical of $z \sim 10$–20. We determine how changing halo properties and model parameters influence the final NC and BH seed population. The NCs have masses, at the time of BH formation, of 10^4–$10^6\,{\rm M}_\odot$ and host BHs with masses between 300 and a few $10^3\,{\rm M}_\odot$.

We compared our new prescriptions with the results of DV09. The mean NC masses are found to decrease with decreasing redshift: higher metallicity usually corresponds to larger star formation rates in the extended disc. These reduce the amount of gas channelled into the centre, leading to less-massive clusters. Except for the more massive NCs, $M_{\rm BH}$ is dictated by the availability of massive stars. The decrease in $M_{\rm cl,0}$ consequently leads to less massive BHs. The total mass densities in BH seeds, $\rho_{\rm seed}$, are higher than found in DV09 as a result of the larger number of gravitationally unstable discs. In our model, we do not take into account any BH growth by accretion, nor their assembly during galaxy clustering. Comoving densities found here are thus lower limits.

In the second paper of this series, we will apply this semi-analytical scheme to merger tree histories in order to trace NC and BH seed formation into a cosmological context. In the third paper of this series, we will follow the evolution of the emerging BH population through accretion and clustering and compare our results with the present observation of the BH occupation frequency in low-redshift galaxies, quasar luminosity functions, etc.

[6] In our reference model ($\eta = 0.3$, $Q_{\rm c} = 2$), the mean cluster mass, $\langle M_{\rm cl,0}\rangle$, decreases with redshift as the metallicity increases. This behaviour appears only because the majority of the systems we are considering here have $R_{\rm SF} > R_{\rm tr}$ (i.e. stars form in most part of the disc). If this were not the case, the behaviour of $\langle M_{\rm cl,0}\rangle$ versus redshift would have been the opposite (see DV09).

ACKNOWLEDGMENTS

We would like to thank the referee for his/her useful suggestions.

REFERENCES

Abel T., Bryan G. L., Norman M. L., 2000, ApJ, 540, 39
Balbus S. A., Papaloizou J. C. B., 1999, ApJ, 521, 650
Balcells M., Graham A. W., Peletier R. F., 2007, ApJ, 665, 1104
Barkana R., Loeb A., 2001, Phys. Rev., 349, 125
Begelman M. C., Rees M. J., 1978, MNRAS, 185, 847
Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289
Böker T., Sarzi M., McLaughlin D. E., van der Marel R. P., Rix H.-W., Ho L. C., Shields J. C., 2004, AJ, 127, 105
Bromm V., Loeb A., 2003, Nat, 425, 812
Bromm V., Coppi P. S., Larson R. B., 1999, ApJ, 527, L5
Bromm V., Ferrara A., Coppi P. S., Larson R. B., 2002, MNRAS, 328, 969
Capuzzo-Dolcetta R., Miocchi P., 2008, MNRAS, 388, L69
Christoudoulou D. M., Shlosman I., Tohline J. E., 1995, ApJ, 443, 551
Clark P. C., Glover S. C. O., Klessen R. S., 2008, ApJ, 672, 757
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Côté P. et al., 2006, ApJS, 165, 57
Devecchi B., Volonteri M., 2009, ApJ, 694, 302 (DV09)
Dunkley J. et al., 2009, ApJ, 701, 1804
Efstathiou G., Lake G., Negroponte J., 1982, MNRAS, 199, 1069
Eisenstein D. J., Loeb A., 1995, ApJ, 443, 11
Elmegreen B. G., Efremov Y. N., 1997, ApJ, 480, 235
Ferrara A., Pettini M., Shchekinov Y., 2000, MNRAS, 319, 539
Ferrarese L. et al., 2006, ApJ, 644, L21
Freese K., Bodenheimer P., Spolyar D., Gondolo P., 2008, ApJ, 685, L101
Freitag M., Gürkan M. A., Rasio F. A., 2006a, MNRAS, 368, 141
Freitag M., Rasio F. A., Baumgardt H., 2006b, MNRAS 368, 121,
Gammie C. F., 2001, ApJ, 553, 174
Glebbeek E., Gaburov E., de Mink S. E., Pols O. R., Portegies Zwart S. F., 2009, A&A, 497, 255
Graham A. W., Spitler L. R., 2009, MNRAS 397, 2148,
Greif T. H., Johnson J. L., Bromm V., Klessen R. S., 2007, ApJ, 670, 1

Gürkan M. A., Freitag M., Rasio F. A., 2004, ApJ, 604, 632
Haehnelt M. G., Rees M. J., MNRAS, 263, 168
Häring N., Rix H.-W., 2004, ApJ, 604, L89
Heger A., Fryer C. L., Woosley S. E., Langer N., Hartmann D. H., 2003, ApJ, 591, 288
Iocco F., Bressan A., Ripamonti E., Schneider R., Ferrara A., Marigo P., 2008, MNRAS, 390, 1655
Kennicutt R. C. Jr, 1998, ApJ, 498, 541
Klessen R. S., Krumholz M. R., Heitsch F., 2009, preprint (arXiv:0906.4452)
Kormendy J., Fisher D. B., Cornell M. E., Bender R., 2009, ApJS, 182, 216
Koushiappas S. M., Bullock J. S., Dekel A., 2004, MNRAS, 354, 292
Kulkarni V. P., Fall S. M., Lauroesch J. T., York D. G., Welty D. E., Khare P., Truran J. W., 2005, ApJ, 618, 68
Li Y., Haiman Z., Mac Low M.-M., 2007, ApJ, 663, 61
Li L.-X., 2008, MNRAS, 388, 1487
Lin D. N. C., Pringle J. E., 1987, MNRAS, 225, 607
Lodato G., 2007, Nuovo Cimento Rivista Serie, 30, 293
Lodato G., Natarajan P., 2006, MNRAS, 371, 1813
Lodato G., Rice W. K. M., 2004, MNRAS, 351, 630
Loeb A., Rasio F. A., 1994, ApJ, 432, 52
Lotz J. M., Telford R., Ferguson H. C., Miller B. W., Stiavelli M., Mack J., 2001, ApJ, 552, 572
Madau P., Rees M. J., 2001, ApJ, 551, L27
Marconi A., Hunt L. K., 2003, ApJ, 589, L21
Milosavljević M., 2004, ApJ, 605, L13
Mineshige S., Umemura M., 1997, ApJ, 480, 167
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Monaco P., Theuns T., Taffoni G., 2002a, MNRAS, 331, 587
Monaco P., Theuns T., Taffoni G., Governato F., Quinn T., Stadel J., 2002b, ApJ, 564, 8
Mori M., Ferrara A., Madau P., 2002, ApJ, 571, 40
Omukai K., Palla F., 2003, ApJ, 589, 677
Omukai K., Schneider R., Haiman Z., 2008, ApJ, 686, 801
Ostriker J. P., Peebles P. J. E., 1973, ApJ, 186, 467
Portegies Zwart S. F., McMillan S. L. W., 2002, ApJ, 576, 899
Portegies Zwart S. F., Makino J., McMillan S. L. W., Hut P., 1999, A&A, 348, 117
Portegies Zwart S. F., Baumgardt H., Hut P., Makino J., McMillan S. L. W., 2004, Nat, 428, 724
Prochaska J. X., Gawiser E., Wolfe A. M., Castro S., Djorgovski S. G., 2003, ApJ, 595, L
Prochaska J. X., Chen H.-W., Dessauges-Zavadsky M., Bloom J. S., 2007, ApJ, 666, 267
Santoro F., Shull J. M., 2006, ApJ, 643, 26
Savaglio S., 2006, New Journal of Physics, 8, 195
Savaglio S. et al., 2005, ApJ, 635, 260
Scannapieco E., Schneider R., Ferrara A., 2003, ApJ, 589, 35
Schleicher D. R. G., Spaans M., Glover S. C. O., 2010, ApJ, 712, L69
Schneider R., Omukai K., Inoue A. K., Ferrara A., 2006, MNRAS, 369, 1437
Sellwood J. A., Moore E. M., 1999, ApJ, 510, 125
Seth A., Agüeros M., Lee D., Basu-Zych A., 2008, ApJ, 678, 116
Shang C., Bryan G. L., Haiman Z., 2010, MNRAS, 402, 1249
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Shlosman I., Frank J., Begelman M. C., 1989, Nat, 338, 45
Shlosman I., Begelman M. C., Frank J., 1990, Nat, 345, 679
Smith B. D., Turk M. J., Sigurdsson S., O'Shea B. W., Norman M. L., 2009, ApJ, 691, 441
Sutherland R. S., Dopita M. A., 1993, ApJS, 88, 253
Toomre A., 1964, ApJ, 139, 1217
Tornatore L., Ferrara A., Schneider R., 2007, MNRAS, 382, 945
Tremaine S. D., Ostriker J. P., Spitzer L. Jr, 1975, ApJ, 196, 407
Tsujimoto T., Nomoto C., Yoshii Y., Hashimoto M., Yamgida S., Thielemann F. R., 1995, MNRAS, 277, 945
Volonteri M., Haardt F., Madau P., 2003, ApJ, 582, 559
Walcher C. J., Böker T., Charlot S., Ho L. C., Rix H.-W., Rossa J., Shields J. C., van der Marel R. P., 2006, ApJ, 649, 692
Wehner E. H., Harris W. E., 2006, ApJ, 644, L17
White S. D. M., Rees M. J., 1978, MNRAS, 183, 341
Woosley S. E., Weaver T. A., 1995, ApJS, 101, 181

APPENDIX A: GAS DEPLETION FROM SUPERNOVA FEEDBACK

In this section, we describe the prescription used to infer the amount of mass lost from the host, following an SN episode. Under the assumption that the gas follows an isothermal profile

$$\rho_{\rm gas} = \frac{\Omega_{\rm b}}{\Omega_{\rm m}} \frac{V_{\rm h}^2}{4\pi G r^2}, \tag{A1}$$

we compute the initial binding energy of the system, $E_{\rm b,in}$, inclusive of the dark matter and gas component:[7]

$$E_{\rm b,in} = -\frac{G M_{\rm h}^2}{R_{\rm vir}} \left(1 + \frac{\Omega_{\rm b}}{\Omega_{\rm m}}\right)^2. \tag{A2}$$

After removing a gas mass, $M_{\rm sh}$, the new binding $E_{\rm b,f}$ is

$$E_{\rm b,f} = -\frac{G M_{\rm h}^2}{R_{\rm vir}} \left(1 + \frac{\Omega_{\rm b}}{\Omega_{\rm m}} - \frac{M_{\rm sh}}{M_{\rm h}}\right)^2. \tag{A3}$$

Energy conservation imposes the following equality: $E_{\rm b,in} + \mathcal{E}_{\rm SN} = E_{\rm b,f} + K_{\rm SN}$, where $\mathcal{E}_{\rm SN}$ is the energy injected by the SN explosions and $K_{\rm SN}$ is the kinetic energy of the outflow that is leaving the halo.

The energy $\mathcal{E}_{\rm SN}$ injected by SNe, over a time t, is given by

$$\mathcal{E}_{\rm SN}(t) = \nu_{\rm SN} E_{\rm SN} \int_0^t \dot{M}_* {\rm d}s, \tag{A4}$$

where $E_{\rm SN}$, the energy of a single SN explosion, is set equal to 10^{51} erg, $\dot{M}_* = \epsilon_{\rm SF}\dot{M}_{\rm inf} + \dot{M}_{*,\rm d}$ and $\nu_{\rm SN}$ is the ratio between the total number of SNe exploding after the formation of a mass, M_*, of stars and M_*. SNe start to explode after 3 Myr from the onset of the star formation episode. We assume that at that time all stars with masses in the SN regime explode together, giving rise to a single bubble, and that a fraction, $f_{\rm w}$, of the total energy released by SN explosions is channelled into the outflow. Following Scannapieco et al. (2003), we relate $f_{\rm w}$ with the halo mass as follows. We define $\delta_{\rm B}(M_{\rm h})$ as

$$\delta_{\rm B}(M) = \begin{cases} 1.0, & \tilde{N}_{\rm t} \leq 1 \\ 1.0 - 0.165\ln(\tilde{N}_{\rm t}^{-1}), & 1 \leq \tilde{N}_{\rm t} \leq 100 \\ [1.0 - 0.165\ln(100)]\,100\tilde{N}_{\rm t}^{-1} & 100 \leq \tilde{N}_{\rm t} \end{cases} \tag{A5}$$

where $\tilde{N}_{\rm t} \equiv 1.7 \times 10^{-7}(\Omega_{\rm b}/\Omega_{\rm m})M_{\rm h}/{\rm M}_\odot$. $f_{\rm w}$ is then given by $f_{\rm w} = \delta_{\rm B}(M_{\rm h})/\delta_{\rm B}(M_{\rm h} = 2 \times 10^8\,{\rm M}_\odot)$. This prescription follows from the scaling relation found in Ferrara, Pettini & Shchekinov (2000), normalized with the result found by Mori, Ferrara & Madau (2002).

Let us now compute $K_{\rm SN} \equiv (1/2)M_{\rm sh}v_{\rm sh}^2$. Here, the shell velocity is calculated at the moment it reaches $R_{\rm vir}$. The topology of early metal enrichment can be extremely complex. Enriched gas first propagates into the cavities created by the shock (Greif et al. 2007) and after the bubble shell leaves the host, it spreads mainly into voids, leading to a very anisotropic spatial distribution of metals. We here neglect these degrees of complexity and simply assume that SN-driven bubbles evolve in spherical symmetry. The shell

[7] Note that the actual gas density profile at the moment of SN explosions would be quite different from the initial one. Differences would arise both from the fact that the gas needs to cool down in order to form stars and from the radiative feedback of stars during their lifetimes.

radius is assumed to follow the Sedov–Taylor solution as long as the pressure of the bubble drops below the value of the surrounding medium. The shell radius $R_{\rm sh}$ evolves with time t as

$$R_{\rm sh} = 1.15 \left(\frac{f_{\rm w} \mathcal{E}_{\rm SN} t^2}{\rho_{\rm bk}} \right)^{1/5}, \tag{A6}$$

where t is the time after the explosion and $\rho_{\rm bk}$ is the ambient gas density, scaling as $(\Omega_{\rm b}/\Omega_{\rm m}) M_{\rm h}/R_{\rm vir}^3$. The velocity of the shell, $v_{\rm sh}$, in the Sedov–Taylor approximation is given by $(2/5)R_{\rm sh}/t$. Under these conditions, $v_{\rm sh}(R_{\rm vir})$ can be easily inferred. At the time $R_{\rm sh} = R_{\rm vir}$, the kinetic energy is

$$K_{\rm SN} = 0.67 \frac{\Omega_{\rm m}}{\Omega_{\rm b}} \frac{M_{\rm sh}}{M_{\rm h}} f_{\rm w} \mathcal{E}_{\rm SN}. \tag{A7}$$

To compute $M_{\rm sh}$ for a given halo, we resort to energy conservation leading to

$$\left(1 - 0.67 \frac{\Omega_{\rm m}}{\Omega_{\rm b}} \frac{M_{\rm sh}}{M_{\rm h}} \right) f_{\rm w} \mathcal{E}_{\rm SN} = \frac{G M_{\rm h} M_{\rm sh}}{R_{\rm vir}} \left[2 \left(1 + \frac{\Omega_{\rm b}}{\Omega_{\rm m}} \right) - \frac{M_{\rm sh}}{M_{\rm h}} \right]. \tag{A8}$$

We note that $M_{\rm sh}/M_{\rm h}$ is always much less than $2(1 + \Omega_{\rm b}/\Omega_{\rm m})$. Neglecting the term $M_{\rm sh}/M_{\rm h}$ in parentheses on the right-hand side of the equation, the outflow mass of the first single bubble is given by

$$M_{\rm sh} = \frac{f_{\rm w} \mathcal{E}_{\rm SN}}{2 (1 + \Omega_{\rm b}/\Omega_{\rm m}) G M_{\rm h}/R_{\rm vir} + \frac{1}{2} v_{\rm sh}^2}. \tag{A9}$$

We assume that an outflow develops only when $v_{\rm sh} > v_{\rm esc}$; otherwise $M_{\rm sh} = 0$.

Massive haloes are not completely evacuated after this first SN episode. Further star formation is often allowed and to account for subsequent SN explosion episodes, we compute the total number of exploding SNe, $N_{\rm SN}$, between the time t_1 and t_2 as

$$N_{\rm SN} = \int_{t_1}^{t_2} \nu_{\rm SN} \left(\epsilon_{\rm SF} \dot{M}_{\rm inf} + \dot{M}_{*,\rm d} \right) {\rm d}s. \tag{A10}$$

We define $M_{\rm sh,s}$ as the contribution of a single explosion (i.e. equation A9 with $\mathcal{E}_{\rm SN} = E_{\rm SN}$). An Outflow develops only if it injects sufficient energy so that $v_{\rm sh} > v_{\rm esc}$. This requires a minimum energy in a single bubble, that is, a minimum number of SNe. For a given $\dot{M}_*$, we calculate the expected number of SNe on a time-scale of the order of their progenitor lifetimes (i.e. 3 Myr). If the resulting $v_{\rm sh}$ is greater than the escape velocity from the halo, outflows are allowed and

$$\dot{M}_{\rm sh} = M_{\rm sh,s} \frac{{\rm d}N_{\rm SN}}{{\rm d}t} = \frac{f_{\rm w} \nu E_{\rm SN} (\epsilon_{\rm SF} \dot{M}_{\rm inf} + \dot{M}_{*,\rm d})}{2 \left[1 + (\Omega_{\rm b}/\Omega_{\rm m}) \right] (G M_{\rm h}/R_{\rm vir}) + (1/2) v_{\rm sh}^2}; \tag{A11}$$

otherwise $\dot{M}_{\rm sh} = 0$.

This paper has been typeset from a TEX/LATEX file prepared by the author.